Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Endenicher Allee 19b • 53115 Bonn • Germany
phone +49 228 73-69828 • fax +49 228 73-69847
www.ins.uni-bonn.de

M. Griebel, J. Hamaekers

Generalized sparse grid interpolation based on the fast discrete Fourier transform

INS Preprint No. 1902

February 2019

Generalized Sparse Grid Interpolation Based on the Fast Discrete Fourier Transform

Michael Griebel and Jan Hamaekers

Abstract In this paper, we present an algorithm for the trigonometric interpolation of multivariate functions on generalized sparse grids and study its application for the approximation of functions in periodic Sobolev spaces of dominating mixed smoothness. Moreover, we derive estimates for the resulting error and the involved cost. We construct interpolants with a computational cost complexity which is substantially lower than for the standard full tensor grid case. The associated generalized sparse grid interpolants have the same approximation order as the standard full tensor grid interpolants, provided that certain additional regularity assumptions on the considered functions are fulfilled. Numerical results validate our theoretical findings.

1 Introduction

In many application areas of numerical simulation, like e.g. physics, chemistry, finance and statistics, high-dimensional approximation problems arise. Here, a conventional numerical approach encounters the so-called curse of dimensionality [5], i.e. the rate of convergence with respect to the number of degrees of freedom deteriorates exponentially with the dimension n. For example, a conventional discretization on uniform grids with $\mathscr{O}(2^L)$ points in each coordinate direction would involve $M = \mathscr{O}(2^{nL})$ degrees of freedom. Correspondingly, only a convergence rate of the type

---

Michael Griebel
Institute for Numerical Simulation, University of Bonn, Wegelerstr. 6, 53115 Bonn, Germany, and Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Schloss Birlinghoven, 53754 Sankt Augustin, Germany

Jan Hamaekers
Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Schloss Birlinghoven, 53754 Sankt Augustin, Germany, e-mail: `jan.hamaekers@scai.fraunhofer.de`

$$\|f - f_L^{\mathrm{FG}}\|_{\mathscr{H}^r} \leq c \cdot M^{-\frac{s-r}{n}} \|f\|_{\mathscr{H}^s} \leq c \cdot 2^{-L(s-r)} \|f\|_{\mathscr{H}^s}$$

can be achieved, where $\|\cdot\|_{\mathscr{H}^r}$ and $\|\cdot\|_{\mathscr{H}^s}$ are the usual Sobolev norms in $\mathscr{H}^r$ and $\mathscr{H}^s$, respectively, s denotes the isotropic smoothness of f and c is a constant which may depend on n and the underlying domain Ω but not on the discretization parameter L.

So-called sparse grid based approaches [7, 28] have emerged as useful techniques to tackle higher dimensional problems, since they open the possibility to break the curse of dimensionality under certain conditions. They date back to the work of Korobov, Bakhvalov, Babenko and Smolyak starting in the late 1950s [56]. For example, if f is in a Sobolev space of bounded mixed smoothness $\mathscr{H}^t_{\mathrm{mix}}(\Omega)$, i.e. if the t-th mixed derivatives of f are bounded, and Ω is a product domain, an error estimate of the type

$$\|f - f_L^{\mathrm{SG}}\|_{\mathscr{H}^r} \leq c \cdot 2^{-(t-r)L} L^{n-1} \|f\|_{\mathscr{H}^t_{\mathrm{mix}}}$$

can be achieved using so-called regular sparse grids where only $M := \mathscr{O}(2^L L^{n-1})$ degrees of freedom are involved. Here, the rate of convergence with respect to the number of degrees of freedom does no longer exponentially deteriorate with the number n of dimensions, except for the logarithmic terms[1] L^{n-1}. Also, in specific cases, the use of so-called energy-norm based sparse grids [7] may even result in an error estimate of type

$$\|f - f_L^{\mathrm{ESG}}\|_{\mathscr{H}^r} \leq c \cdot 2^{-(t-r)L} \|f\|_{\mathscr{H}^t_{\mathrm{mix}}},$$

where only $M := \mathscr{O}(2^L)$ degrees of freedom are involved. Hence, compared to the regular sparse grid case, even the logarithmic term L^{n-1}, in the degrees of freedom is eliminated.[2]

For the discretization with sparse grids, Galerkin type methods, finite difference approaches and the so-called combination technique have been developed over the last two decades [7]. Furthermore, these approaches were used in the context of moderate higher-dimensional elliptic, parabolic and hyperbolic differential equations. In addition, sparse grid techniques were successfully applied for the solution of integral equations [29], for quadrature [17], for regression [15] and for time series prediction [6]. Moreover, the sparse grid method was supplemented with adaptive refinement schemes [6, 17, 26], was used for the construction of anisotropic sparse tensor product spaces [25, 24] and was applied in the context of weighted mixed spaces [22, 26]. Sparse grid based collocation schemes were for example discussed in [3, 30, 34, 35, 38, 45]. They recently found widespread use in the important field

---

[1] Here, in case of the best linear approximation, estimates of type $\|f - f_L^{\mathrm{SG}}\|_{\mathscr{H}^r} \lesssim 2^{-(t-r)L} L^{(n-1)/2} \|f\|_{\mathscr{H}^t_{\mathrm{mix}}}$ and even of type $\|f - f_L^{\mathrm{SG}}\|_{\mathscr{H}^r} \lesssim 2^{-(t-r)L} \|f\|_{\mathscr{H}^t_{\mathrm{mix}}}$ could be achieved for certain types of basis sets [29, 39, 61]. This holds, e.g. for wavelets and the Fourier basis, respectively. The involved degrees of freedom are still of order $M := \mathscr{O}(2^L L^{n-1})$.

[2] The constants in the bounds of the degrees of freedom and in the accuracy estimates still depend on n, though.

of uncertainty quantification [44]. On a theoretical level, sparse grids are closely related to ANOVA-like decompositions [14, 19, 31] which are well-known from statistics. A detailed survey on sparse grids is for example given in [7].

Note that the adaption of the sparse grid techniques to Fourier based methods is obtained by means of Fourier polynomials from the hyperbolic cross and hence sparse grid methods are also known under the name hyperbolic cross approximation [51, 58]. For an extensive survey see [19]. The properties of such approximations of periodic functions in Sobolev spaces on the n-dimensional torus $\mathbb{T}^n$ have been studied by several authors in, e.g. [8, 12, 18, 36, 37, 39, 42, 49, 52, 53, 58].

The sparse grid approach can be further generalized. To this end, periodic spaces of generalized mixed Sobolev smoothness

$$\mathscr{H}^{t,r}_{\text{mix}}(\mathbb{T}^n) := \left\{ f : \sqrt{\sum_{\mathbf{k}\in\mathbb{Z}^n} \prod_{d=1}^{n} (1+|k_d|)^{2t}(1+|\mathbf{k}|_\infty)^{2r}|\hat{f}_{\mathbf{k}}|^2} < \infty \right\}$$

and a specific generalization of the regular sparse grid spaces based on Fourier polynomials $e^{i\mathbf{k}^T\mathbf{x}}$ with frequencies $\mathbf{k}$ from the generalized hyperbolic cross

$$\Gamma^T_{2^L} := \left\{ \mathbf{k} \in \mathbb{Z}^n : \prod_{d=1}^{n} (1+k_d)\cdot(1+|\mathbf{k}|_\infty) \le 2^{L(1-T)} \right\}$$

were introduced in [39] and further discussed in [27, 28, 31, 40]. Moreover, a generalization to Banach spaces was given in [10]. Here, $T \in [-\infty, 1)$ is an additional parameter that controls the mixture of isotropic and mixed smoothness: The case $T = 0$ corresponds to the conventional hyperbolic cross (or regular sparse grid) discretization space, where $\mathscr{O}(2^L L^{n-1})$ frequencies are involved. Furthermore, the case $T = -\infty$ corresponds to the full tensor grid. The case $T \to 1$ corresponds to a latin hypercube and the case $0 < T < 1$ resembles certain energy-norm based sparse grids where the order of the amount of included frequencies does not depend on the number of dimensions n, i.e. it is $\mathscr{O}(2^L)$.

Usually, these generalized sparse grid spaces are based on *linear information*, i.e. on the Fourier coefficients of f which involve an explicit integration of f against the respective Fourier basis function. In contrast to that, so called *standard information* involves only the use of function values. In this case, it is in general not clear if the approximation error of an associated *interpolant* exhibits the same convergence rate as that of the best linear approximation.

In this paper, we now mainly deal with trigonometric interpolation on generalized sparse grids and its application for the approximation of multivariate functions in certain periodic Sobolev spaces of bounded mixed smoothness. For functions on the torus, regular sparse grid interpolation methods based on the fast Fourier transform were for example introduced by Hallatschek in [30] and are also discussed in [4, 18, 35, 36, 42, 52]. Here, for a function in a periodic space of dominated mixed smoothness $\mathscr{H}^t_{\text{mix}}(\mathbb{T}^n)$ with $t > 1$, i.e. a Korobov space, it was proved in [30] that the approximation error in the maximum norm of its regular sparse grid interpolant is of the order $\mathscr{O}(2^{-L(t-1)}L^{n-1})$ and in [18] a (suboptimal) upper bound estimate of the same order is shown for the approximation error in the $\mathscr{L}_2$ norm. Here, the involved degrees of freedom are of the order $\mathscr{O}(2^L L^{n-1})$ and the corresponding computational cost complexity (including the work count involved in the FFT) is of the order[3] $\mathscr{O}(2^L L^n)$. Based on the results of [58] it was furthermore shown in [42, 43] that the approximation error in the $\mathscr{H}^r$-norm of the interpolant associated with a regular sparse grid is of the order $\mathscr{O}(2^{-(t-r)L} L^{n-1})$, if the function is in a periodic Sobolev space $\mathscr{H}^t_{\text{mix}}$ with $t > \frac{1}{2}$.

In this work, we present an extension of the algorithm of Hallatschek given in [30] to the case of interpolation (standard information) on the *generalized* sparse grids as introduced in [27] for $f \in \mathscr{H}^{t,r}_{\text{mix}}$. We further study the error for the corresponding best linear information algorithm and give estimates for the error and involved degrees of freedom for functions in different variants of periodic Sobolev spaces of dominating mixed smoothness. Moreover, for functions of mixed Sobolev smoothness $\mathscr{H}^t_{\text{mix}}$, we show estimates for the approximation error of the interpolant in the $\mathscr{H}^r$-norm. To our knowledge this has been done so far only for the regular sparse grid case $T = 0$, but not yet for the case of generalized sparse grids with $0 < T < 1$, which resemble certain energy-norm based sparse grids. Altogether, it turns out that under specific conditions, both, the rate of the error and the rate of the corresponding degrees of freedom are independent of the dimension n of the function. For example, let $f \in \mathscr{H}^2_{\text{mix}}(\mathbb{T}^n)$ and let us measure the approximation error in the $\mathscr{H}^1$-norm, where $\mathbb{T}^n$ denotes the n-dimensional torus. Then, an error of the order $\mathscr{O}(2^{-L})$ and an amount of degrees of freedom of the order $\mathscr{O}(2^L)$ can be achieved[4] for interpolants corresponding to a generalized sparse grid with $0 < T < \frac{1}{2}$ for any dimension n.

Note that, compared to our previous article [23], we now present a slight modification in the proofs, which yields to improved estimates for the approximation error of the trigonometric regular sparse grid interpolants and its overall complexity. For the special case $T = 0$, the new result corresponds to the lower bounds shown in case of sampling on the Smolyak grids (regular sparse grids) in [11].

The remainder of this paper is organized as follows: In section 2 we introduce the fast Fourier transform on general sparse grids with hierarchical Fourier bases. In particular, we recall the conventional Fourier basis representation of periodic functions in subsection 2.1 and the so-called hierarchical Fourier basis representation in subsection 2.2. Furthermore, in subsection 2.3, we present generalized sparse grids and discuss the construction and application of associated trigonometric interpolation operators and computational cost complexities. In section 3 we introduce different variants of periodic Sobolev spaces. We discuss their associated best linear approximation error in subsection 3.2, the approximation error of the trigonometric general sparse grid interpolants in subsection 3.3 and its overall complexities in subsection 3.4. In section 4 we apply our approach to some test cases to show that our theoretical results are also matched in practical numerical experiments. Finally we give some concluding remarks in section 5.

---

[3] We here do not have $\mathscr{O}(2^L L^{n-1})$ but we have $\mathscr{O}(2^L L^n)$ as one L stems from the computational cost complexity of the *one-dimensional* FFT involved.

[4] Moreover, the computational cost is of order $\mathscr{O}(2^L L)$, where the additional factor L stems from the involved FFT algorithm.

2 Fourier transform on general sparse grids with hierarchical bases

To construct a trigonometric interpolation operator for generalized sparse grids, we follow the approach of Hallatschek [30]. To this end, we first recall the conventional Fourier basis, introduce the so-called hierarchical Fourier basis and then discuss its use in the construction of a generalized sparse grid interpolant.

2.1 Fourier basis representation

First, let us shortly recall the conventional Fourier basis representation of periodic functions. To this end, let $\mathbb{T}^n$ be the n-torus, which is the n-dimensional cube $\mathbb{T}^n \subset \mathbb{R}^n$, $\mathbb{T} = [0, 2\pi]$, where opposite sides are identified. We then have n-dimensional coordinates $\mathbf{x} := (x_1, \ldots, x_n)$, where $x_d \in \mathbb{T}$. We define the basis function associated to a multi-index $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{Z}^n$ by

$$\omega_{\mathbf{k}}(\mathbf{x}) := \left(\bigotimes_{d=1}^{n} \omega_{k_d} \right)(\mathbf{x}) = \prod_{d=1}^{n} \omega_{k_d}(x_d), \quad \omega_k(x) := e^{ikx}. \tag{1}$$

The set $\{\omega_{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^n}$ is a complete orthogonal system of the space $\mathscr{L}_2(\mathbb{T}^n)$ and hence every $f \in \mathscr{L}_2(\mathbb{T}^n)$ has the unique expansion

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_{\mathbf{k}} \omega_{\mathbf{k}}(\mathbf{x}), \tag{2}$$

where the Fourier coefficients are given by

$$\hat{f}_{\mathbf{k}} := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} \omega_{\mathbf{k}}^*(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x}. \tag{3}$$

Note that it is common to characterize the smoothness class of a function f by the decay properties of its Fourier coefficients, see [32]. In this way, we introduce for $r \in \mathbb{R}$ the periodic Sobolev space of isotropic smoothness as

$$\mathscr{H}^r(\mathbb{T}^n) := \left\{ f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_{\mathbf{k}} \omega_{\mathbf{k}}(\mathbf{x}) \,:\, \|f\|_{\mathscr{H}^r} := \sqrt{\sum_{\mathbf{k}\in\mathbb{Z}^n} (1+|\mathbf{k}|_\infty)^{2r} |\hat{f}_{\mathbf{k}}|^2} < \infty \right\}.$$

For discretization purposes, let us now define finite-dimensional subspaces of the space $\mathscr{L}_2(\mathbb{T}^n) = \mathscr{H}^0(\mathbb{T}^n)$. To this end, we set

$$\sigma : \mathbb{N}_0 \to \mathbb{Z} : j \mapsto \begin{cases} -j/2 & \text{if } j \text{ is even,} \\ (j+1)/2 & \text{if } j \text{ is odd.} \end{cases} \tag{4}$$

For $l \in \mathbb{N}_0$ we introduce the one-dimensional nodal basis

$$\mathscr{B}_l := \{\phi_j\}_{0\le j\le 2^l-1} \quad \text{with} \quad \phi_j := \omega_{\sigma(j)} \tag{5}$$

and the corresponding spaces $V_l := \mathrm{span}\{\mathscr{B}_l\}$. For a multi-index $\mathbf{l} \in \mathbb{N}_0^n$ we define finite-dimensional spaces by a tensor product construction, i.e. $V_{\mathbf{l}} := \bigotimes_{d=1}^n V_{l_d}$. Finally, we introduce the space[5]

$$V := \sum_{\mathbf{l}\in\mathbb{N}^n} V_{\mathbf{l}}.$$

In the following we shortly recall the common one-dimensional trigonometric interpolation. Let the Fourier series $\sum_{k\in\mathbb{Z}} \hat{f}_k \omega_k$ be pointwise convergent to $f(x)$. Then, for interpolation points $\mathscr{S}_l := \{m\frac{2\pi}{2^l} : m = 0,\dots,2^l-1\}$ of level $l \in \mathbb{N}_0$, the interpolation operator can be defined by

$$I_l : V \to V_l : f \mapsto I_l f := \sum_{j\in\mathscr{G}_l} \hat{f}_j^{(l)} \phi_j$$

with indices $\mathscr{G}_l := \{0,\dots,2^l-1\}$ and discrete nodal Fourier coefficients

$$\hat{f}_j^{(l)} := 2^{-l} \sum_{x\in\mathscr{S}_l} f(x)\phi_j^*(x), \tag{6}$$

which only involve point evaluations $f(x)$ at points $x \in \mathscr{S}_l$. This way, the 2^l interpolation conditions

$$f(x) = I_l f(x) \quad \text{for all } x \in \mathscr{S}_l$$

are fulfilled. In particular, from (6) and (2) one can deduce the well-known aliasing formula

$$\hat{f}_j^{(l)} = \sum_{k\in\mathbb{Z}} \hat{f}_k 2^{-l} \sum_{x\in\mathscr{S}_l} \omega_{\sigma(j)}^*(x)\omega_k(x) = \sum_{m\in\mathbb{Z}} \hat{f}_{\sigma(j)+m2^l}. \tag{7}$$

Next, let us consider the case of multivariate functions. To this end, let the Fourier series $\sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_{\mathbf{k}} \omega_{\mathbf{k}}$ be pointwise convergent to f. Then, according to the tensor product structure of the n-dimensional space, we introduce the n-dimensional interpolation operator on full tensor grids as

$$I_{\mathbf{l}} := I_{l_1} \otimes \cdots \otimes I_{l_n} : V \to V_{\mathbf{l}} : f \mapsto I_{\mathbf{l}} f = \sum_{\mathbf{j}\in\mathscr{G}_{\mathbf{l}}} \hat{f}_{\mathbf{j}}^{(\mathbf{l})} \phi_{\mathbf{j}},$$

with

$$\mathscr{G}_{\mathbf{l}} := \mathscr{G}_{l_1} \times \cdots \times \mathscr{G}_{l_n} \subset \mathbb{N}_0^n$$

and multi-dimensional discrete nodal Fourier coefficients

[5] Except for the completion with respect to a chosen Sobolev norm, V is just the associated Sobolev space.

$$\hat{f}_{\mathbf{j}}^{(\mathbf{l})} := 2^{-|\mathbf{l}|_1} \sum_{\mathbf{x} \in \mathscr{S}_{\mathbf{l}}} f(\mathbf{x}) \phi_{\mathbf{j}}^*(\mathbf{x}), \tag{8}$$

where

$$\mathscr{S}_{\mathbf{l}} := \mathscr{S}_{l_1} \times \cdots \times \mathscr{S}_{l_n} \subset \mathbb{T}^n.$$

Similar to (7) there holds the multi-dimensional aliasing formula

$$\hat{f}_{\mathbf{j}}^{(\mathbf{l})} = \sum_{\mathbf{m} \in \mathbb{Z}^n} \hat{f}_{\sigma(\mathbf{j}) + \mathbf{m} 2^{\mathbf{l}}}, \tag{9}$$

where $\sigma(\mathbf{j}) := (\sigma(j_1), \ldots, \sigma(j_n))$ and $\mathbf{m} 2^{\mathbf{l}} := \left(m_1 2^{l_1}, \ldots, m_n 2^{l_n}\right)$.

2.2 One-dimensional hierarchical Fourier basis representation

Now we discuss a hierarchical variant of the Fourier basis representation. Let us first consider the one-dimensional case. To this end, we introduce a univariate Fourier hierarchical basis function for $j \in \mathbb{N}_0$ by

$$\psi_j := \begin{cases} \phi_0 & \text{for } j = 0, \\ \phi_j - \phi_{2^l - 1 - j} & \text{for } 2^{l-1} \le j \le 2^l - 1,\, l \ge 1, \end{cases} \tag{10}$$

and we define the one-dimensional hierarchical Fourier basis including basis functions up to level $l \in \mathbb{N}_0$ by

$$\mathscr{B}_l^{\mathrm{h}} := \{\psi_j\}_{0 \le j \le 2^l - 1}.$$

Let us further introduce the difference spaces

$$W_l := \begin{cases} \operatorname{span}\{\mathscr{B}_0^{\mathrm{h}}\} & \text{for } l = 0, \\ \operatorname{span}\{\mathscr{B}_l^{\mathrm{h}} \setminus \mathscr{B}_{l-1}^{\mathrm{h}}\} & \text{for } l > 0. \end{cases}$$

Note that there holds the relation $V_l = \operatorname{span}\{\mathscr{B}_l\} = \operatorname{span}\{\mathscr{B}_l^{\mathrm{h}}\}$ for all $l \in \mathbb{N}_0$. Thus, we have the decomposition of V_l into the direct sum $V_l = \bigoplus_{\nu=0}^{l} W_l$. Now, let $l \in \mathbb{N}_0$ and $u \in V_l$. Then, for any $u \in V_l$, one can easily switch from its hierarchical representation

$$u = \sum_{0 \le j \le 2^l - 1} u_j^h \psi_j, \quad \text{where } u_j^h \in \mathbb{R}, \tag{11}$$

to the nodal representation

$$u = \sum_{0 \le j \le 2^l - 1} u_j \phi_j \tag{12}$$

by a linear transform. This transform just maps the hierarchical coefficient vector $(u_0^h, u_1^h, \ldots, u_{2^l-1}^h)^T$ to the nodal coefficient vector $(u_0, u_1, \ldots, u_{2^l-1})^T$. For example for $l = 0,1,2,3$, the corresponding *de-hierarchization* matrices read as

$$\begin{pmatrix}1\end{pmatrix}, \quad \begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}, \quad \begin{pmatrix}1 & -1 & 0 & -1\\ 0 & 1 & -1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}, \quad \begin{pmatrix}1 & -1 & 0 & -1 & 0 & 0 & 0 & -1\\ 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0\\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\end{pmatrix},$$

respectively. For all $l \in \mathbb{N}_0$ the de-hierarchization matrix can be easily inverted and its determinant is equal to one. Analogously one can switch from the nodal representation (12) to the hierarchical representation (11) by the associated inverse linear transform. Here, the corresponding *hierarchization* matrices read as

$$\begin{pmatrix}1\end{pmatrix}, \quad \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \quad \begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}, \quad \begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\end{pmatrix},$$

for $l = 0,1,2,3$ respectively. Simple algorithms with computational cost complexity $\mathscr{O}(2^l)$ for hierarchization and de-hierarchization are given in [30].

Let us now define the operator

$$\Delta_l := (I_l - I_{l-1}) : V \to W_l, \text{ for } l \geq 0,$$

where we set $I_{-1} = 0$. Note that the image of Δ_l is a subspace of W_l. Hence, we define the corresponding hierarchical Fourier coefficients $\check{f}_j$ by the unique representation

$$\Delta_l f = \sum_{0\leq v<l} \sum_{j\in\mathscr{J}_v} (\hat{f}_j^{(l)} - \hat{f}_j^{(l-1)})\phi_j + \sum_{j\in\mathscr{J}_l} \hat{f}_j^{(l)}\phi_j =: \sum_{j\in\mathscr{J}_l} \check{f}_j \psi_j \tag{13}$$

with

$$\mathscr{J}_v := \begin{cases}\{0\} & \text{for } v = 0,\\ \{2^{v-1}, \ldots, 2^v - 1\} & \text{for } v \geq 1.\end{cases} \tag{14}$$

Moreover, we can write the interpolation operator associated with level l in the form

$$\begin{aligned} I_l f &= (I_l - I_{l-1} + I_{l-1} - \cdots - I_0 + I_0 - I_{-1}) f\\ &= (\Delta_l + \cdots + \Delta_0) f\\ &= \sum_{0\leq v\leq l} \sum_{j\in\mathscr{J}_v} \check{f}_j \psi_j = \sum_{0\leq j\leq 2^l-1} \check{f}_j \psi_j. \end{aligned}$$

In particular, let us note that the interpolation relation

$$\Delta_l f(x) = f(x) \quad \text{for all } x \in \mathscr{S}_l^{\mathrm{h}} \quad \text{and} \quad \Delta_l f(x) = 0 \quad \text{for all } x \in \mathscr{S}_{l-1}$$

holds, where

$$\mathscr{S}_l^{\mathrm{h}} := \mathscr{S}_l \setminus \mathscr{S}_{l-1}, \quad \text{with } \mathscr{S}_{-1} := \emptyset.$$

For $l \in \mathbb{N}_0$ the equation

$$\sum_{0 \le v < l} \sum_{j \in \mathscr{J}_v} (\hat{f}_j^{(l)} - \hat{f}_j^{(l-1)}) \phi_j + \sum_{j \in \mathscr{J}_l} \hat{f}_j^{(l)} \phi_j = - \sum_{0 \le v < l} \sum_{j \in \mathscr{J}_v} \check{f}_{2^l - 1 - j} \phi_j + \sum_{j \in \mathscr{J}_l} \check{f}_j \phi_j$$

follows by the definitions (10) and (13). Therefore, the hierarchical Fourier coefficient $\check{f}_j$ is equal to the discrete nodal Fourier coefficient $\hat{f}_j^{(l)}$ associated with level l for $j \in \mathscr{J}_l$. Hence in the case $l \in \mathbb{N}_0$, $j \in \mathscr{J}_l$, we obtain the relation

$$\check{f}_j = \hat{f}_j^{(l)} = \sum_{m \in \mathbb{Z}} \hat{f}_{\sigma(j) + m2^l} \tag{15}$$

with the help of the aliasing formula (7).

Let us remark that the one-dimensional standard Fourier basis representation is sufficient to define multi-dimensional full tensor grids, whereas the hierarchical Fourier basis representation is indeed necessary for the definition of sparse grids.

2.3 Generalized sparse grids

Now we consider the case of multivariate functions. Here, we use a tensor product ansatz to construct n-dimensional basis functions as well as spaces for the n-dimensional case. To this end, we set $\psi_{\mathbf{j}} := \bigotimes_{d=1}^n \psi_{j_d}$ and $W_{\mathbf{l}} := \bigotimes_{d=1}^n W_{l_d}$ for $\mathbf{l} \in \mathbb{N}_0^n$. In particular, we have the direct sum decomposition

$$V = \bigoplus_{\mathbf{l} \in \mathbb{N}^n} W_{\mathbf{l}}.$$

Moreover, we define $W_{\mathscr{I}} := \bigoplus_{\mathbf{l} \in \mathscr{I}} W_{\mathbf{l}}$ for an index set $\mathscr{I} \subset \mathbb{N}_0^n$.

For the general sparse grid construction, we restrict ourselves to index sets, which obey the following condition [17, 30]: An index set $\mathscr{I} \subset \mathbb{N}_0^n$ is called admissible if there holds the relation

$$\{\mathbf{v} \in \mathbb{N}_0^n : \mathbf{v} \le \mathbf{l}\} \subset \mathscr{I} \tag{16}$$

for all $\mathbf{l} \in \mathscr{I}$. Here, the inequality $\mathbf{v} \le \mathbf{w}$ is to be understood componentwise, i.e. $\mathbf{v} \le \mathbf{w} :\Leftrightarrow v_d \le w_d$ for all $1 \le d \le n$. Now, for any admissible index set $\mathscr{I}$, we define generalized sparse grid space by

$$V_{\mathscr{I}} := \sum_{\mathbf{l} \in \mathscr{I}} V_{\mathbf{l}} = \bigoplus_{\mathbf{l} \in \mathscr{I}} W_{\mathbf{l}} = W_{\mathscr{I}}. \tag{17}$$

Due to property (16) of $\mathscr{I}$ we are able to introduce the corresponding general sparse grid trigonometric interpolation operator by

$$I_{\mathscr{I}} := \sum_{\mathbf{l} \in \mathscr{I}} \Delta_{\mathbf{l}} : V \to V_{\mathscr{I}}, \quad \text{where } \Delta_{\mathbf{l}} := \Delta_{l_1} \otimes \cdots \otimes \Delta_{l_n} : V \to W_{\mathbf{l}}.$$

The associated set of interpolation points is given by

$$\mathscr{S}_{\mathscr{I}} := \bigcup_{\mathbf{l}\in\mathscr{I}} \mathscr{S}^{\mathrm{h}}_{\mathbf{l}}, \quad \text{where } \mathscr{S}^{\mathrm{h}}_{\mathbf{l}} := \mathscr{S}^{\mathrm{h}}_{l_1} \times \cdots \times \mathscr{S}^{\mathrm{h}}_{l_n}.$$

Let us note that this general sparse grid construction includes generalized sparse grids as introduced in [27], i.e. we may employ for $\mathscr{I}$ the index set

$$\mathscr{I}_L^T := \{\mathbf{l} : |\mathbf{l}|_1 - T|\mathbf{l}|_\infty \leq (1-T)L\}, \quad T < 1, \tag{18}$$

where the choice of the parameter T allows the indication of specific degrees of sparsification. Here full tensor grids, i.e. $\mathscr{I}_L^{-\infty} = \{\mathbf{l} : |\mathbf{l}|_\infty \leq L\}$ and conventional regular sparse grids, i.e. $\mathscr{I}_L := \mathscr{I}_L^0 = \{\mathbf{l} : |\mathbf{l}|_1 \leq L\}$ of level $L \in \mathbb{N}_0$ are covered as special cases. But also, energy-norm type [7] sparse grid patterns are included by $0 < T < 1$ or a Latin hypercube by $T \to 1$.

For a function f with a pointwise convergent Fourier series, the multi-dimensional hierarchical coefficients $\breve{f}_{\mathbf{j}}$ are given by the unique representation

$$\Delta_{\mathbf{l}} f = \sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \breve{f}_{\mathbf{j}} \psi_{\mathbf{j}}, \quad \text{where } \mathscr{J}_{\mathbf{l}} := \mathscr{J}_{l_1} \times \cdots \times \mathscr{J}_{l_n}.$$

In particular, the hierarchical Fourier series

$$\sum_{\mathbf{l}\in\mathbb{N}_0^n} \sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \breve{f}_{\mathbf{j}} \psi_{\mathbf{j}} \tag{19}$$

converges pointwise to f on all grids $S_{\mathbf{l}}$, $\mathbf{l} \in \mathbb{N}_0^n$. Furthermore, with the help of the multi-dimensional aliasing formula (9), a relation similar to (15) can easily be deduced, that is, for $\mathbf{l} \in \mathbb{N}_0^n$ and $\mathbf{j} \in \mathscr{J}_{\mathbf{l}}$, it holds

$$\breve{f}_{\mathbf{j}}^{(\mathbf{l})} = \hat{f}_{\mathbf{j}}^{(\mathbf{l})} = \sum_{\mathbf{m}\in\mathbb{Z}^n} \hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}}}. \tag{20}$$

With the definition (17) of the general sparse grid space $V_{\mathscr{I}}$, we can estimate its number of degrees of freedom by

$$|V_{\mathscr{I}}| = \sum_{\mathbf{l}\in\mathscr{I}} |W_{\mathbf{l}}| = \sum_{\mathbf{l}\in\mathscr{I}} 2^{\sum_{d=1}^n \max(0, l_d - 1)} \lesssim \sum_{\mathbf{l}\in\mathscr{I}} 2^{|\mathbf{l}|_1}. \tag{21}$$

Starting from relation (21) the following estimate is shown in the case of the general index sets $\mathscr{I}_L^T$ of (18) in [27, 28]:

Lemma 1. *Let $L \in \mathbb{N}_0$ and $T < 1$. The number of degrees of freedom of the general sparse grid spaces $V_{\mathscr{I}_L^T}$ with respect to the discretization parameter L is*

$$\left|V_{\mathscr{I}_L^T}\right| \lesssim \sum_{\mathbf{l}\in\mathscr{I}_L^T} 2^{|\mathbf{l}|_1} \lesssim \begin{cases} 2^L & \textit{for } 0<T<1,\\ 2^L L^{n-1} & \textit{for } T=0,\\ 2^{L\frac{T-1}{T/n-1}} & \textit{for } T<0,\\ 2^{Ln} & \textit{for } T=-\infty. \end{cases} \tag{22}$$

Furthermore, analogously to the well-known case of a multi-dimensional discrete Fourier transform, we can utilize the tensor product structure of the underlying spaces and operators to efficiently compute the general sparse grid interpolant $I_{\mathscr{I}} f$ for a given $f \in V$. Here, the multi-dimensional transformation is expressed in terms of one-dimensional discrete Fourier transforms, hierarchizations and de-hierarchizations of different size, cf. Algorithm 1 and see also [30].

Algorithm 1: A procedure to apply the general sparse grid interpolation operator $I_{\mathscr{I}}$ for a given admissible index set $\mathscr{I}$ and given interpolation values $\{u_{\mathbf{j}} := f(x_{\mathbf{j}}) \in \mathbb{C}\}_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}},\mathbf{l}\in\mathscr{I}}$ associated to the general sparse grid interpolation points $\mathscr{S}_{\mathscr{I}}$. The algorithm works in-place on the given input coefficients, where we use an additional temporary array to perform the involved one-dimensional FFTs.

input : $\{u_{\mathbf{j}} := f(x_{\mathbf{j}})\}_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}},\mathbf{l}\in\mathscr{I}}$
for $d \leftarrow 1$ **to** n **do**
  forall $\mathbf{l} \in \mathscr{M}_d(\mathscr{I})$ **do**
    forall $\mathbf{j} \in \mathscr{J}_{l_1} \times \cdots \times \mathscr{J}_{l_{d-1}} \times \mathscr{J}_0 \times \mathscr{J}_{l_{d+1}} \times \cdots \times \mathscr{J}_{l_n}$ **do**
      One-dimensional FFT for $(u_{j_1,\ldots,j_{d-1},0,j_{d+1},\ldots,j_n}, \ldots, u_{j_1,\ldots,j_{d-1},2^{l_d}-1,j_{d+1},\ldots,j_n})$
      Hierarchization for $(u_{j_1,\ldots,j_{d-1},0,j_{d+1},\ldots,j_n}, \ldots, u_{j_1,\ldots,j_{d-1},2^{l_d}-1,j_{d+1},\ldots,j_n})$
for $d \leftarrow 1$ **to** n **do**
  forall $\mathbf{l} \in \mathscr{M}_d(\mathscr{I})$ **do**
    forall $\mathbf{j} \in \mathscr{J}_{l_1} \times \cdots \times \mathscr{J}_{l_{d-1}} \times \mathscr{J}_0 \times \mathscr{J}_{l_{d+1}} \times \cdots \times \mathscr{J}_{l_n}$ **do**
      De-hierarchization for $(u_{j_1,\ldots,j_{d-1},0,j_{d+1},\ldots,j_n}, \ldots, u_{j_1,\ldots,j_{d-1},2^{l_d}-1,j_{d+1},\ldots,j_n})$
Finally, the non-hierarchical sparse grid Fourier coefficients are given in $\{u_{\mathbf{j}}\}_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}},\mathbf{l}\in\mathscr{I}}$.

In Algorithm 1 we give a procedure to apply the general sparse grid interpolation operator $I_{\mathscr{I}}$ associated to an admissible index set $\mathscr{I}$, where we define for $d \in \{1,\ldots,n\}$ the set

$$\mathscr{M}_d(\mathscr{I}) := \{\mathbf{l} \in \mathscr{I} : \mathbf{l} + \mathbf{e}_d \notin \mathscr{I}\},$$

with the d-th unit vector $\mathbf{e}_d$.

Note that the application of a fast Fourier transform algorithm for the computation of a one-dimensional discrete Fourier transform of length 2^l results in a computational cost complexity of order $\mathscr{O}(l2^l)$. Note furthermore that the complexity for a one-dimensional hierarchization or de-hierarchization of length 2^l is of linear order $\mathscr{O}(2^l)$. Now, an upper estimate for the resulting overall computational cost complexity $\mathscr{R}[I_{\mathscr{I}}]$ of Algorithm 1 can be easily deduced in the form

$$\begin{aligned}\mathscr{R}[I_{\mathscr{I}}] &\lesssim \sum_{d=1}^{n} \sum_{\mathbf{l}\in\mathscr{M}_d(\mathscr{I})} 2^{l_d} l_d 2^{|\mathbf{l}|_1 - l_d} = \sum_{d=1}^{n} \sum_{\mathbf{l}\in\mathscr{M}_d(\mathscr{I})} 2^{|\mathbf{l}|_1} l_d \le \sum_{d=1}^{n} l_{\max} \sum_{\mathbf{l}\in\mathscr{M}_d(\mathscr{I})} 2^{|\mathbf{l}|_1} \\ &\le n l_{\max} \sum_{\mathbf{l}\in\mathscr{I}} 2^{|\mathbf{l}|_1}, \end{aligned} \tag{23}$$

where $l_{\max} := \max_{\mathbf{l}\in\mathscr{I}} |\mathbf{l}|_\infty$. Note that the inverse operator $I_{\mathscr{I}}^{-1}$ can easily computed by performing the algorithm in a reverse way [30]. In the case of the general sparse grid index sets $\mathscr{I}_L^T$, relation (23) and Lemma 1 lead directly to the following computational cost complexity estimate:

Lemma 2. *Let $L \in \mathbb{N}_0$ and $T < 1$. An upper estimate for the computational cost complexity of the general sparse grid interpolation operator $I_{\mathscr{I}_L^T}$ with respect to the discretization parameter L is given by*

$$\mathscr{R}[I_{\mathscr{I}_L^T}] \lesssim L \sum_{\mathbf{l}\in\mathscr{I}_L^T} 2^{|\mathbf{l}|_1} \lesssim \begin{cases} L2^L & \textit{for } 0 < T < 1, \\ L2^L L^{n-1} & \textit{for } T = 0, \\ L2^{L\frac{T-1}{T/n-1}} & \textit{for } T < 0, \\ L2^{Ln} & \textit{for } T = -\infty. \end{cases}$$

Let us remark that the case $T = 0$ was already presented in [30], i.e. $\mathscr{R}[I_{\mathscr{I}_L}] = \mathscr{O}(L^n 2^L)$. Note in particular that both, the asymptotic number of degrees of freedom of V_L^T in Lemma 1 and the asymptotic computational cost complexity of $I_{\mathscr{I}_L^T}$ in Lemma 2, are not exponentially dependent on the dimension n in the case $0 < T < 1$.[2]

Let us finally note that, alternatively, the interpolation operator $I_{\mathscr{I}}$ can be applied using the so-called combination technique or the blending scheme [4, 16, 39]. For an admissible index set $\mathscr{I}$ it holds

$$I_{\mathscr{I}} f = \sum_{\mathbf{l}\in\mathscr{I}} r_{\mathscr{I}}(\mathbf{l}) I_{\mathbf{l}} f, \quad \text{where } r_{\mathscr{I}}(\mathbf{l}) := \sum_{\mathbf{v}\in\{0,1\}^n} (-1)^{|\mathbf{v}|} \chi_{\mathscr{I}}(\mathbf{l}+\mathbf{v}), \tag{24}$$

with the characteristic function

$$\chi_{\mathscr{I}}(\mathbf{l}) := \begin{cases} 1 & \text{for } \mathbf{l} \in \mathscr{I}, \\ 0 & \text{otherwise.} \end{cases}$$

This way, just a linear combination of interpolation operators $I_{\mathbf{l}}$ is to be formed to obtain $I_{\mathscr{I}}$, which is based on the well-known inclusion-exclusion principle [57]. It is easy to see that the degrees of freedom involved in (24) can again be estimated from above as in Lemma 1. The computational cost of the combination technique (24) can be estimated by

$$\mathscr{R}[I_{\mathscr{I}}] \lesssim \sum_{\mathbf{l}\in\mathscr{I}, r_{\mathscr{I}}(\mathbf{l})\neq 0} 2^{|\mathbf{l}|_1} |\mathbf{l}|_1,$$

which again gives the bounds of Lemma 2, albeit with slightly different constants in the $\mathscr{O}$-notation.

3 Approximation estimates

In this section, we first define different variants of (periodic) Sobolev spaces on the torus via Fourier series, i.e. we classify functions via the decay of their Fourier coefficients and hence by their smoothness. Then, we give approximation estimates for these spaces. Here, we discuss the best linear approximation error and the approximation error of the interpolant. Based on the derived estimates we further study the resulting error and the involved degrees of freedom.

3.1 Periodic Sobolev spaces

As already noted in section 2.1 we characterize the smoothness class of a function f by the decay properties of its Fourier coefficients [32]. To this end, let $\beta : \mathbb{Z}^n \to \mathbb{R}_+$ be a continuous and positive function, which implicitly expresses some smoothness class. Then we define

$$\mathscr{H}_\beta(\mathbb{T}^n) := \left\{ f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_\mathbf{k}\omega_\mathbf{k}(\mathbf{x}) \,:\, \|f\|_\beta := \sqrt{\sum_{\mathbf{k}\in\mathbb{Z}^n} \beta(\mathbf{k})^2|\hat{f}_\mathbf{k}|^2} < \infty \right\}. \quad (25)$$

Here, e.g. for $r,t \in \mathbb{R}$, the smoothness functions

$$\beta(\mathbf{k}) = \lambda_{\text{iso}}(\mathbf{k})^r, \quad \text{where} \quad \lambda_{\text{iso}}(\mathbf{k}) := 1 + |\mathbf{k}|_\infty,$$

and

$$\beta(\mathbf{k}) = \lambda_{\text{mix}}(\mathbf{k})^t, \quad \text{where} \quad \lambda_{\text{mix}}(\mathbf{k}) := \prod_{d=1}^{n}(1+|k_d|),$$

result in the conventional isotropic Sobolev spaces $\mathscr{H}^r$ [2] and in the standard Sobolev spaces with dominating mixed smoothness $\mathscr{H}^t_{\text{mix}}$ [50], respectively. A further example is the multiplicative combination of these functions, i.e.

$$\beta(\mathbf{k}) = \lambda_{\text{iso}}(\mathbf{k})^r\lambda_{\text{mix}}(\mathbf{k})^t,$$

which leads to the generalized Sobolev spaces of dominating mixed smoothness [27]

$$\mathscr{H}_\beta(\mathbb{T}^n) = \mathscr{H}^{t,r}_{\text{mix}}(\mathbb{T}^n) := \left\{ f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_\mathbf{k}\omega_\mathbf{k}(\mathbf{x}) \,: \right.$$
$$\left. \|f\|_{\mathscr{H}_\beta} = \|f\|_{\mathscr{H}^{t,r}_{\text{mix}}} := \sqrt{\sum_{\mathbf{k}\in\mathbb{Z}^n} \left(\lambda_{\text{mix}}(\mathbf{k})^t\lambda_{\text{iso}}(\mathbf{k})^r\right)^2|\hat{f}_\mathbf{k}|^2} < \infty \right\}. \quad (26)$$

In particular, these spaces include the conventional spaces as special cases, i.e.

$$\mathscr{H}^r(\mathbb{T}^n) = \mathscr{H}^{0,r}_{\text{mix}}(\mathbb{T}^n) \quad \text{and} \quad \mathscr{H}^t_{\text{mix}}(\mathbb{T}^n) = \mathscr{H}^{t,0}_{\text{mix}}(\mathbb{T}^n),$$

respectively. Hence, the parameter r from equation (26) governs the isotropic smoothness, whereas t governs the mixed smoothness.

Moreover, the spaces $\mathscr{H}^{t,r}_{\text{mix}}(\mathbb{T}^n)$ can be generalized to the case of n-dimensional smoothness parameters $\mathbf{t},\mathbf{r} \in \mathbb{R}^n$ with $\mathbf{r} \geq \mathbf{0}$ [28]. To this end, for $\mathbf{t},\mathbf{r} \in \mathbb{R}^n$ with $\mathbf{r} \geq \mathbf{0}$, we set $\beta(\mathbf{k}) = \lambda^{(\mathbf{t})}_{\text{mix}}(\mathbf{k})\lambda^{(\mathbf{r})}_{\text{iso}}(\mathbf{k})$, where

$$\lambda^{(\mathbf{t})}_{\text{mix}}(\mathbf{k}) := \prod_{d=1}^{n}(1+|k_d|)^{t_d} \quad \text{and} \quad \lambda^{(\mathbf{r})}_{\text{iso}}(\mathbf{k}) := \sum_{d=1}^{n}(1+|k_d|)^{r_d},$$

and introduce the spaces

$$\mathscr{H}_\beta(\mathbb{T}^n) = \mathscr{H}^{\mathbf{t},\mathbf{r}}_{\text{mix}}(\mathbb{T}^n) := \Bigg\{ f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_{\mathbf{k}}\omega_{\mathbf{k}}(\mathbf{x}) : \\ \|f\|_{\mathscr{H}_\beta} = \|f\|_{\mathscr{H}^{\mathbf{t},\mathbf{r}}_{\text{mix}}} := \sqrt{\sum_{\mathbf{k}\in\mathbb{Z}^n}\left(\lambda^{(\mathbf{t})}_{\text{mix}}(\mathbf{k})\lambda^{(\mathbf{r})}_{\text{iso}}(\mathbf{k})\right)^2 |\hat{f}_{\mathbf{k}}|^2} < \infty \Bigg\}. \quad (27)$$

In this way, for $r \geq 0$, the spaces $\mathscr{H}^{t,r}_{\text{mix}}$ are up to norm equivalency[6] special cases of the spaces $\mathscr{H}^{\mathbf{t},\mathbf{r}}_{\text{mix}}$, i.e. $\mathscr{H}^{t,r}_{\text{mix}} = \mathscr{H}^{(t,\dots,t),(r,\dots,r)}_{\text{mix}}$. We use the short form $\mathscr{H}^{\mathbf{r}} := \mathscr{H}^{\mathbf{0},\mathbf{r}}_{\text{mix}}$ and $\mathscr{H}^{\mathbf{t}}_{\text{mix}} := \mathscr{H}^{\mathbf{t},\mathbf{0}}_{\text{mix}}$.

Furthermore, following [55, 62], for a set of weights $\Gamma := \{\gamma_u\}_{u\subset\{1,\dots,n\}}$ with $\gamma_u \geq 0$ and a smoothness function $\tilde{\beta}$ we introduce a weighted periodic Sobolev space with the help of the definition

$$\beta(\mathbf{k}) = \tfrac{1}{\sqrt{\gamma_{v(\mathbf{k})}}}\beta'(\mathbf{k}), \quad \text{with } v : \mathbb{Z}^n \to 2^{\{1,\dots,n\}} : \mathbf{k} \mapsto v(\mathbf{k}),$$

where $v(\mathbf{k}) \subset \{1,\dots,n\}$ is uniquely defined by the following relation for all $1 \leq d \leq n$: if $k_d \neq 0$ it follows $d \in v(\mathbf{k})$. This way we can define the following space

$$\mathscr{H}_\beta(\mathbb{T}^n) = \mathscr{H}^{\Gamma}_{\beta'}(\mathbb{T}^n) := \Bigg\{ f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^n} \hat{f}_{\mathbf{k}}\omega_{\mathbf{k}}(\mathbf{x}) : \\ \|f\|_{\mathscr{H}^{\Gamma}_{\beta'}} := \sqrt{\sum_{\mathbf{k}\in\mathbb{Z}^n}\frac{1}{\gamma_{v(\mathbf{k})}}\beta'(\mathbf{k})^2|\hat{f}_{\mathbf{k}}|^2} < \infty \Bigg\}. \quad (28)$$

Let us remark that the orthogonal decomposition

$$f = \sum_{v\subset\{1,\dots,n\}} f_v, \quad \text{with } f_v := \sum_{\mathbf{k}\in\Omega_v} \hat{f}_{\sigma(\mathbf{k})}\omega_{\mathbf{k}} \quad (29)$$

[6] For $r \geq 0$ we could also use the function $\prod_{d=1}^{n}(1+|k_d|)^t\,(\sum_{d=1}^{n}(1+|k_d|)^r)$ instead of $\prod_{d=1}^{n}(1+|k_d|)^t\,(1+|\mathbf{k}|_\infty)^r$ to define the space $\mathscr{H}^{t,r}_{\text{mix}}$. The function $\prod_{d=1}^{n}(1+|k_d|)^t\,(\sum_{d=1}^{n}(1+|k_d|)^r)$ is equal to the definition of the special case $\lambda^{(t,\dots,t)}_{\text{mix}}(\mathbf{k})\lambda^{(r,\dots,r)}_{\text{iso}}(\mathbf{k})$ and hence also the associated spaces $\mathscr{H}^{t,r}_{\text{mix}}$ and $\mathscr{H}^{(t,\dots,t),(r,\dots,r)}_{\text{mix}}$ would be equivalent. However, many of the given proofs would get more technical and thus, for reasons of simplicity, we restrict ourselves to the definition (26).

is well-known in statistics under the name ANOVA (analysis of variance) [13], where

$$\Omega_v := \{\mathbf{k} \in \mathbb{Z}^n : k_d = 0 \text{ for all } d \in \{1,\dots,n\} \setminus u\}.$$

and f_u in particular depends on the coordinates $\{x_d\}_{d\in u}$ only. Note that for $f \in \mathscr{H}_\beta^\Gamma$ the weight γ_u prescribes the importance of the term f_u and hence the importance of different dimensions and of correlations between dimensions. In particular for a weight $\gamma_v \to 0$ the norm $\|f_v\|_{\mathscr{H}_\beta}$ is forced to be zero. If the size of the terms $\|f_v\|_{\mathscr{H}_\beta}$ decays fast with e.g. $|v|$, then a proper restriction onto certain lower dimensional functions results in a substantial reduction in computational cost. For example a set of weights $\Gamma_q := \{\gamma_v\}_{v\subset\{1,\dots,n\}}$ with $\gamma_v = 0$ for all $v \subset \{1,\dots,n\}$, $|v| > q$ results in a periodic Sobolev space of finite-order q, cf. [55, 62, 63]. Thus, all terms f_v with $|v| > q$ are either not present at all or can be neglected due to the decay with $|v|$. Then, the problem of approximating a n-dimensional function reduces to the problem of approximating q-dimensional functions.

3.2 Best linear approximation error

In the following, we consider the error of the best linear approximation in finite-dimensional general sparse grid discretization spaces. Here, we restrict ourselves to some specific Sobolev spaces of dominating mixed smoothness.

For $\mathbf{l} \in \mathbb{N}_0^n$ we define an approximation operator $Q_\mathbf{l}$ with respect to the $\mathscr{L}_2$-norm by

$$Q_\mathbf{l} := Q_{l_1} \otimes \ldots \otimes Q_{l_n} : \mathscr{L}_2(\mathbb{T}^n) \to V_\mathbf{l},$$

where

$$Q_l : \mathscr{L}_2(\mathbb{T}) \to V_l : f \mapsto \sum_{0\le j\le 2^l-1} \hat{f}_{\sigma(j)} \phi_j.$$

For an admissible index set $\mathscr{I}$, as introduced in section 2.3, we define a general sparse grid approximation operator $Q_\mathscr{I} : \mathscr{L}_2(\mathbb{T}^n) \to V_\mathscr{I}$ by

$$Q_\mathscr{I} f := \sum_{\mathbf{l}\in\mathscr{I}} \sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \hat{f}_{\sigma(\mathbf{j})} \omega_{\sigma(\mathbf{j})}.$$

Now let us consider two smoothness functions β and $\tilde{\beta}$ with associated Sobolev spaces $\mathscr{H}_\beta$ and $\mathscr{H}_{\tilde{\beta}}$ and norms $\|f\|_{\mathscr{H}_\beta}$ and $\|f\|_{\mathscr{H}_{\tilde{\beta}}}$, respectively. It should hold $\mathscr{H}_\beta \subset \mathscr{H}_{\tilde{\beta}} \subset \mathscr{L}_2$ and thus $\beta(\mathbf{k}) \lesssim \tilde{\beta}(\mathbf{k})$. Then, let us consider $f \in \mathscr{H}_\beta(\mathbb{T}^n) \subset \mathscr{L}_2(\mathbb{T}^n)$ with the unique representation $f = \hat{f}_\mathbf{k} \omega_\mathbf{k}$. Now, if

$$\max_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I}, \mathbf{j}\in\mathscr{J}_\mathbf{l}} \frac{\tilde{\beta}(\sigma(\mathbf{j}))^2}{\beta(\sigma(\mathbf{j}))^2} < \infty,$$

we obtain for the best linear approximation in $V_{\mathscr{I}}$ the estimate

$$\begin{aligned}
\inf_{\tilde{f}\in V_{\mathscr{I}}} \|f-\tilde{f}\|^2_{\mathscr{H}_{\tilde{\beta}}} \leq \|f-Q_{\mathscr{I}}f\|^2_{\mathscr{H}_{\tilde{\beta}}} &= \|\sum_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I}}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \hat{f}_{\sigma(\mathbf{j})}\omega_{\sigma(\mathbf{j})}\|^2_{\mathscr{H}_{\tilde{\beta}}} \\
&= \sum_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I}}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \tilde{\beta}(\sigma(\mathbf{j}))^2|\hat{f}_{\sigma(\mathbf{j})}|^2 \\
&= \sum_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I}}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \frac{\tilde{\beta}(\sigma(\mathbf{j}))^2}{\beta(\sigma(\mathbf{j}))^2}|\hat{f}_{\sigma(\mathbf{j})}|^2\beta(\sigma(\mathbf{j}))^2 \\
&\leq \left(\max_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I},\mathbf{j}\in\mathscr{J}_\mathbf{l}} \frac{\tilde{\beta}(\sigma(\mathbf{j}))^2}{\beta(\sigma(\mathbf{j}))^2}\right)\sum_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I}}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} |\hat{f}_{\sigma(\mathbf{j})}|^2\beta(\sigma(\mathbf{j}))^2 \\
&\leq \left(\max_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I},\mathbf{j}\in\mathscr{J}_\mathbf{l}} \frac{\tilde{\beta}(\sigma(\mathbf{j}))^2}{\beta(\sigma(\mathbf{j}))^2}\right)\sum_{\mathbf{l}\in\mathbb{Z}^n}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} |\hat{f}_{\sigma(\mathbf{j})}|^2\beta(\sigma(\mathbf{j}))^2 \\
&= \left(\max_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I},\mathbf{j}\in\mathscr{J}_\mathbf{l}} \frac{\tilde{\beta}(\sigma(\mathbf{j}))^2}{\beta(\sigma(\mathbf{j}))^2}\right)\|f\|^2_{\mathscr{H}_\beta}. \qquad (30)
\end{aligned}$$

This general result allows us to derive error estimates for a wide range of situations. We shortly consider two specific cases, namely the pairings $(\mathscr{H}^{t',r'}_{\text{mix}}, \mathscr{H}^{t,r}_{\text{mix}})$ and $(\mathscr{H}^r, \mathscr{H}^{\mathbf{t}}_{\text{mix}})$.

First, for the linear approximation in general sparse grid spaces $V_{\mathscr{I}^T_L}$ with the index $\mathscr{I}^T_L$ set given in (18), the following error estimate for functions in optimized Sobolev spaces of dominating mixed smoothness $\mathscr{H}^{t,r}_{\text{mix}}$ can be derived:

Lemma 3. *For $L\in\mathbb{N}_0$, $T<1$, $t'+r'<t+r$, $t-t'\geq 0$ and $f\in\mathscr{H}^{t,r}_{\text{mix}}(\mathbb{T}^n)$ it holds*

$$\inf_{\tilde{f}\in V_{\mathscr{I}^T_L}} \|f-\tilde{f}\|_{\mathscr{H}^{t',r'}_{\text{mix}}} \leq \|f-Q_{\mathscr{I}^T_L}f\|_{\mathscr{H}^{t',r'}_{\text{mix}}}$$

$$\lesssim \begin{cases} 2^{L\left((r'-r)-(t-t')+\left(T(t-t')-(r'-r)\right)\frac{n-1}{n-T}\right)}\|f\|_{\mathscr{H}^{t,r}_{\text{mix}}} & \text{for } T\geq\frac{r'-r}{t-t'}, \\ 2^{L\left((r'-r)-(t-t')\right)}\|f\|_{\mathscr{H}^{t,r}_{\text{mix}}} & \text{for } T\leq\frac{r'-r}{t-t'}. \end{cases}$$

Proof. According to (30), the estimation of

$$\max_{\mathbf{l}\in\mathbb{Z}^n\setminus\mathscr{I},\mathbf{j}\in\mathscr{J}_\mathbf{l}} \frac{\lambda_{\text{mix}}(\sigma(\mathbf{j}))^{t'}\lambda_{\text{iso}}(\sigma(\mathbf{j}))^{r'}}{\lambda_{\text{mix}}(\sigma(\mathbf{j}))^{t}\lambda_{\text{iso}}(\sigma(\mathbf{j}))^{r}}$$

leads to the desired result, see also [28, 31]. □

Second, for $\mathbf{t}\geq\mathbf{1}$ we can define an anisotropic admissible index set by

$$\mathscr{I}^{\mathbf{t}}_L := \{\mathbf{l}\in\mathbb{N}^n_0 : \sum_{d=1}^n t_d l_d \leq L\},$$

see also [25, 24]. Then, the following error estimate can be derived analogously to Lemma 3:

Lemma 4. *For $L \in \mathbb{N}_0$, $\mathbf{t} > \mathbf{0}$, $f \in \mathscr{H}^{\mathbf{t}}_{\mathrm{mix}}$, $|\mathbf{t}|_{min} - r > 0$ and with $\mathbf{t}' := \frac{\mathbf{t}}{|t|_{min}}$, where $|\mathbf{t}|_{min} := \min_{d=1}^{n} t_d$, it holds*

$$\inf_{\tilde{f} \in V_{\mathscr{I}_L^{\mathbf{t}'}}} \|f - \tilde{f}\|_{\mathscr{H}^r} \leq \|f - Q_{\mathscr{I}_L^{\mathbf{t}'}} f\|_{\mathscr{H}^{\mathbf{t}}_{\mathrm{mix}}} \lesssim 2^{-(|\mathbf{t}|_{min} - r)} \|f\|_{\mathscr{H}^{\mathbf{t}}_{\mathrm{mix}}}.$$

Note that for e.g. $t_1 = |\mathbf{t}|_{min} < t_2 \leq \ldots \leq t_n$, the number of degrees of freedom $|V_{\mathscr{I}_L^{\mathbf{t}'}}|$ is of order $\mathscr{O}(2^L)$, which is independent of the number of dimensions n.

So far we have considered the best linear approximation of a function. However, its coefficients are given by Fourier integrals (3), which can be evaluated by analytic formulae only in special cases. Alternatively, the interpolant of the function could be employed. Then, however, it is in general not clear if the associated approximation error exhibits the same convergence rate as that of the best linear approximation. This issue is discussed in the next section.

3.3 Approximation error of interpolant

In the following we consider the error of the approximation by trigonometric *interpolation*. To this end, let us first recall the following two lemma:

Lemma 5. *For $L \in \mathbb{N}_0$, $f \in \mathscr{H}^r$, $s > \frac{n}{2}$ and $0 \leq s < r$ it holds*

$$\|f - I_{\mathscr{I}_L^{-\infty}} f\|_{\mathscr{H}^s} \lesssim 2^{-(r-s)L} \|f\|_{\mathscr{H}^r}. \tag{31}$$

Note that analogous lemmata are given in [42, 43] based on the works [48, 58]. Let us further recall the lemma:

Lemma 6. *For $L \in \mathbb{N}_0$, $f \in \mathscr{H}^t_{\mathrm{mix}}$, $t > \frac{1}{2}$ and $0 \leq s < t$ it holds*

$$\|f - I_{\mathscr{I}_L^0} f\|_{\mathscr{H}^s} \lesssim 2^{-(t-s)L} L^{\frac{n-1}{2}} \|f\|_{\mathscr{H}^t_{\mathrm{mix}}}. \tag{32}$$

Analogous lemmata are given in [9, 12, 41, 54, 59] and ealier weaker versions, i.e.

$$\|f - I_{\mathscr{I}_L^0} f\|_{\mathscr{H}^r} \lesssim 2^{-(t-s)L} L^{n-1} \|f\|_{\mathscr{H}^t_{mix}},$$

can be found in [42, 43, 58].

Next, we extend Lemma 6 from the pairing $(\mathscr{H}^s, \mathscr{H}^t_{mix})$ to the case of more general smoothness indices functions and to general sparse grids. Let $f \in \mathscr{H}_\beta$ obey a pointwise convergent (hierarchical) Fourier series. Then, the relation

$$\begin{aligned}\|f - I_{\mathscr{I}} f\|_{\mathscr{H}_{\tilde\beta}} &= \|\sum_{\mathbf{l}\in\mathbb{N}_0^n}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \breve f_\mathbf{j}\psi_\mathbf{j} - \sum_{\mathbf{l}\in\mathscr{I}}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \breve f_\mathbf{j}\psi_\mathbf{j}\|_{\mathscr{H}_{\tilde\beta}} \\ &= \|\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}}\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \breve f_\mathbf{j}\psi_\mathbf{j}\|_{\mathscr{H}_{\tilde\beta}} \\ &\le \sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} \|\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \breve f_\mathbf{j}\psi_\mathbf{j}\|_{\mathscr{H}_{\tilde\beta}} \end{aligned} \tag{33}$$

holds. By definition of the hierarchical basis (10) we obtain

$$\begin{aligned}\|\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \breve f_\mathbf{j}\psi_\mathbf{j}\|^2_{\mathscr{H}_{\tilde\beta}} &= \|\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}}\sum_{\mathbf{v}\in\{0,1\}^n} \breve f_\mathbf{j}\bigotimes_{d=1}^n \phi_{\mu_{v_d}^{l_d}(j_d)}\|^2_{\mathscr{H}_{\tilde\beta}} \\ &= \sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}}\sum_{\mathbf{v}\in\{0,1\}^n,\mathbf{l}-\mathbf{v}\ge\mathbf{0}} |\breve f_\mathbf{j}|^2\tilde\beta(\sigma(\mu_\mathbf{v}^\mathbf{l}(\mathbf{j})))^2 \\ &= \sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}}\sum_{\mathbf{v}\in\{0,1\}^n,\mathbf{l}-\mathbf{v}\ge\mathbf{0}} \left|\sum_{\mathbf{m}\in\mathbb{Z}^n}\hat f_{\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l}}\right|^2\tilde\beta(\sigma(\mu_\mathbf{v}^\mathbf{l}(\mathbf{j})))^2 \\ &= \sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}}\sum_{\mathbf{v}\in\{0,1\}^n,\mathbf{l}-\mathbf{v}\ge\mathbf{0}} \left|\sum_{\mathbf{m}\in\mathbb{Z}^n}\hat f_{\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l}}\frac{\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})}{\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})}\right|^2\tilde\beta(\sigma(\mu_\mathbf{v}^\mathbf{l}(\mathbf{j})))^2,\end{aligned}$$

where $\mu_0^l(j) = j$,

$$\mu_1^l(j) = \begin{cases}-1 & \text{if } l\le 0,\\ 2^l-1-j & \text{if } l\ge 1,\end{cases}$$

$\mu_\mathbf{v}^\mathbf{l} = (\mu_{v_1}^{l_1},\dots,\mu_{v_n}^{l_n})$ and $\phi_{-1}=0$. With the Cauchy-Schwarz inequality it follows that

$$\begin{aligned}&\left|\sum_{\mathbf{m}\in\mathbb{Z}^n}\hat f_{\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l}}\frac{\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})}{\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})}\right|^2 \\ &\qquad\le\left(\sum_{\mathbf{m}\in\mathbb{Z}^n}\left|\hat f_{\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l}}\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})\right|^2\right)\left(\sum_{\mathbf{m}\in\mathbb{Z}^n}\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})^{-2}\right)\end{aligned}\tag{34}$$

and hence it holds

$$\begin{aligned}\|\sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}} \breve f_\mathbf{j}\psi_\mathbf{j}\|^2_{\mathscr{H}_{\tilde\beta}} \le \sum_{\mathbf{j}\in\mathscr{J}_\mathbf{l}}\sum_{\mathbf{v}\in\{0,1\}^n,\mathbf{l}-\mathbf{v}\ge\mathbf{0}}&\left(\sum_{\mathbf{m}\in\mathbb{Z}^n}\left|\hat f_{\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l}}\right|^2\left|\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})\right|^2\right)\times \\ &\times\left(\sum_{\mathbf{m}\in\mathbb{Z}^n}\beta(\sigma(\mathbf{j})+\mathbf{m}2^\mathbf{l})^{-2}\right)\tilde\beta(\sigma(\mu_\mathbf{v}^\mathbf{l}(\mathbf{j})))^2.\end{aligned}$$

Now, let us assume that there is a function $g:\mathbb{N}_0^n\to\mathbb{R}$ such that it holds

$$\tilde{\beta}(\sigma(\mu_{\mathbf{v}}^{\mathbf{l}}(\mathbf{j})))^2 \sum_{\mathbf{m}\in\mathbb{Z}^n} \left|\beta(\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}})\right|^{-2} \leq C^2 g(\mathbf{l})^2 \tag{35}$$

for all $\mathbf{j} \in \mathscr{J}_{\mathbf{l}}$ and $\mathbf{v} \in \{0,1\}^n, \mathbf{l}-\mathbf{v} \geq \mathbf{0}$ with a constant C independent of $\mathbf{j}$ and $\mathbf{v}$. Then, with $|\{0,1\}^n| = 2^n$, we have

$$\|\sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \breve{f}_{\mathbf{j}}\psi_{\mathbf{j}}\|_{\mathscr{H}_{\tilde{\beta}}} \leq 2^n C g(\mathbf{l}) \left(\sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \sum_{\mathbf{m}\in\mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}}}|^2 \beta(\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}})^2\right)^{\frac{1}{2}}$$

which results with relation (33) in

$$\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} \|\sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \breve{f}_{\mathbf{j}}\psi_{\mathbf{j}}\|_{\mathscr{H}_{\tilde{\beta}}} \leq 2^n C \sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} g(\mathbf{l}) \left(\sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \sum_{\mathbf{m}\in\mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}}}|^2 \beta(\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}})^2\right)^{\frac{1}{2}}$$

and with Hölder's inequality in

$$\begin{aligned}\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} \|\sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \breve{f}_{\mathbf{j}}\psi_{\mathbf{j}}\|_{\mathscr{H}_{\tilde{\beta}}} \leq 2^n C &\left(\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} g(\mathbf{l})^2\right)^{\frac{1}{2}} \times \\ &\times \left(\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} \sum_{\mathbf{j}\in\mathscr{J}_{\mathbf{l}}} \sum_{\mathbf{m}\in\mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}}}|^2 \beta(\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}})^2\right)^{\frac{1}{2}}.\end{aligned}$$

This leads finally to the estimate

$$\|f - I_{\mathscr{I}} f\|_{\mathscr{H}_{\tilde{\beta}}} \lesssim \left(\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}} g(\mathbf{l})^2\right)^{\frac{1}{2}} \|f\|_{\mathscr{H}_{\beta}}. \tag{36}$$

Let us now consider the approximation error in the $\mathscr{H}^s$-norm for approximating $f \in \mathscr{H}_{\text{mix}}^{t,r}$ in the sparse grid space $V_{\mathscr{I}_L^T}$ by interpolation. To this end, let us first recall the following upper bound:

Lemma 7. *For $L \in \mathbb{N}_0$, $T < 1$, $s < t$ and $t \geq 0$ it holds*

$$\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}_L^T} 2^{-t|\mathbf{l}|_1 + s|\mathbf{l}|_\infty} \lesssim \begin{cases} 2^{-\left((t-s)+(Tt-s)\frac{n-1}{n-T}\right)L} L^{n-1} & \text{for } T \geq \frac{s}{t}, \\ 2^{-(t-s)L} & \text{for } T < \frac{s}{t}. \end{cases}$$

Proof. A proof is given in Theorem 4 in [39]. □

Now, we can give the following lemma:

Lemma 8. *Let $L \in \mathbb{N}_0$, $T < 1$, $s - r < t$, $t + \frac{r}{n} > \frac{1}{2}$, $t \geq 0$,$r \geq 0$ and $f \in \mathscr{H}_{\text{mix}}^t$ with a pointwise convergent Fourier series. Then it holds*

$$\|f - I_{\mathscr{I}_L^T} f\|_{\mathscr{H}^s} \lesssim \begin{cases} 2^{-\left((t-(s-r))+(Tt-(s-r))\frac{n-1}{n-T}\right)L} L^{\frac{n-1}{2}} \|f\|_{\mathscr{H}^{t,r}_{\text{mix}}} & \text{for } T \geq \frac{s-r}{t}, \\ 2^{-(t-(s-r))L} \|f\|_{\mathscr{H}^{t,r}_{\text{mix}}} & \text{for } T < \frac{s-r}{t}. \end{cases} \tag{37}$$

Proof. For $\mathbf{j} \in \mathscr{J}_{\mathbf{l}}$ and $\mathbf{v} \in \{0,1\}^n$ with $\mathbf{l} - \mathbf{v} \geq \mathbf{0}$ it follows the relation

$$\begin{aligned} &\sum_{\mathbf{m}\in\mathbb{Z}^n} \prod_{d=1}^n \left(1 + |\sigma(j_d) + m_d 2^{l_d}|\right)^{-2t} \left(1 + |\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}}|_\infty\right)^{-2r} \\ &\quad \lesssim \sum_{\mathbf{m}\in\mathbb{Z}^n} \prod_{d=1}^n \left(2^{l_d}(1+|m_d|)\right)^{-2t} \left(2^{|\mathbf{l}|_\infty}(1+|\mathbf{m}|_\infty)\right)^{-2r} \\ &\quad \lesssim 2^{-2t|\mathbf{l}|_1} 2^{-2r|\mathbf{l}|_\infty} \sum_{\mathbf{m}\in\mathbb{Z}^n} \prod_{d=1}^n (1+|m_d|)^{-2t} (1+|\mathbf{m}|_\infty)^{-2r}. \end{aligned} \tag{38}$$

For $t \geq 0$ and $r \geq 0$ it follows with the inequality of arithmetic and geometric means

$$\begin{aligned} \sum_{\mathbf{m}\in\mathbb{Z}^n} \prod_{d=1}^n (1+|m_d|)^{-2t} (1+|\mathbf{m}|_\infty)^{-2r} &\lesssim \sum_{\mathbf{m}\in\mathbb{Z}^n} \left(\frac{1}{n}\sum_{d=1}^n (1+|m_d|)^{-2t}\right)^n (1+|\mathbf{m}|_\infty)^{-2r} \\ &\lesssim \sum_{\mathbf{m}\in\mathbb{Z}^n\setminus\mathbf{0}} |\mathbf{m}|_1^{-2tn} |\mathbf{m}|_\infty^{-2r} \\ &\lesssim \sum_{m\in\mathbb{N}} \sum_{|\mathbf{m}|=m} |\mathbf{m}|_\infty^{-2tn} |\mathbf{m}|_\infty^{-2r} \\ &\lesssim \sum_{m\in\mathbb{N}} m^{n-1} m^{-2tn} m^{-2r} \\ &= \sum_{m\in\mathbb{N}} m^{-2tn-2r+n-1} \end{aligned}$$

and hence with $t + \frac{r}{n} > \frac{1}{2}$ and (38) we obtain

$$\begin{aligned} (1+|\sigma(\mu_{\mathbf{v}}^{\mathbf{l}}(\mathbf{j}))|_\infty)^{2s} \sum_{\mathbf{m}\in\mathbb{Z}^n} \prod_{d=1}^n \left(1+|\sigma(j_d)+m_d 2^{l_d}|\right)^{-2t} \left(1+|\sigma(\mathbf{j})+\mathbf{m}2^{\mathbf{l}}|_\infty\right)^{-2r} \\ \lesssim 2^{-t|\mathbf{l}|_1 + (s-r)|\mathbf{l}|_\infty}. \end{aligned}$$

According to (35) and (36) this yields

$$\|f - I_{\mathscr{I}_L^T} f\|_{\mathscr{H}^s} \lesssim \left(\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}_L^T} 2^{-2t|\mathbf{l}|_1 + 2(s-r)|\mathbf{l}|_\infty}\right)^{\frac{1}{2}} \|f\|_{\mathscr{H}^{t,r}_{\text{mix}}}$$

and with Lemma 7 we obtain the upper estimate

$$\sqrt{\sum_{\mathbf{l}\in\mathbb{N}_0^n\setminus\mathscr{I}_L^T} 2^{-2t|\mathbf{l}|_1+2(s-r)|\mathbf{l}|_\infty}} \lesssim \begin{cases} 2^{-\left((t-(s-r))+(Tt-(s-r))\frac{n-1}{n-T}\right)L} L^{\frac{n-1}{2}} & \text{for } T \geq \frac{s-r}{t}, \\ 2^{-(t-(s-r))L} & \text{for } T < \frac{s-r}{t}. \end{cases}$$

and thus the desired result. □

Note that our previous article [23] included a weaker version of Lemma 8, which just involved the logarithmic term L^{n-1} instead of $L^{(n-1)/2}$. Note here furthermore that Lemma 8 does not cover all pairings $(\mathscr{H}_{\text{mix}}^{t',r'}, \mathscr{H}_{\text{mix}}^{t,r})$, which are included in Lemma 3 for the general sparse grid linear approximation operator $Q_{\mathscr{I}_L^T}$. In particular, besides restrictions to allow for point evaluation (i.e. $t > \frac{1}{2}$), pairings $(\mathscr{H}_{\text{mix}}^{t',r'}, \mathscr{H}_{\text{mix}}^{t,r})$ with $t' \neq 0$ are not considered.

Nevertheless, let us now compare the error bounds for standard information and linear information. We see that for regular sparse grids, i.e. $T = 0$, there is a difference in the error behavior between the best approximation and the approximation by interpolation: In the $\mathscr{L}_2$-norm error estimate for the interpolant resulting from Lemma 8 with $t > \frac{1}{2}$, $r = 0$ and $T = 0$, there is a logarithmic factor present, i.e. $L^{(n-1)/2}$. In contrast, for the best linear approximation error in the $\mathscr{L}_2$-norm, there is no such logarithmic term involved, compare Lemma 3 with $t > 0$, $t' = r' = r = 0$, $T = 0$ and c.f. Table 1.

The question, if this gap between the sampling number (standard information) and the approximation number (linear information) exists or if it can be closed, is for general sampling point distributions an open problem, see [33, 47]. However, in the case of sampling on Smolyak grids (i.e. for regular sparse grids with $T = 0$) it has been shown that estimate (32) is sharp. Hence the corresponding sampling number and approximation number are not equal in this case, see [11] and also Table 1.

However, for general sparse tensor spaces with $T \neq 0$ and for full tensor spaces (i.e. $T = -\infty$), the error of the best linear approximation and the error by interpolation behave asymptotically equal, see also Table 2 and Table 3. There is no gap in these situations.

Table 1: Convergence behavior in case of regular sparse grid tensor spaces for best linear approximation and for interpolation, i.e. $Q_{\mathscr{I}_L^0}$ and $I_{\mathscr{I}_L^0}$, respectively.

	$f \in$	$\mathscr{H}^0$-error	dof M	convergence
$Q_{\mathscr{I}_L^0}$	$\mathscr{H}_{mix}^t, t > 0$	$\mathscr{O}(2^{-tL})$	$\mathscr{O}(2^L L^{n-1})$	$\mathscr{O}\left(M^{-t} L^{t(n-1)}\right)$
$I_{\mathscr{I}_L^0}$	$\mathscr{H}_{mix}^t, t > \frac{1}{2}$	$\mathscr{O}(2^{-tL} L^{(n-1)/2})$	$\mathscr{O}(2^L L^{n-1})$	$\mathscr{O}\left(M^{-t} L^{(t+1/2)(n-1)}\right)$

Table 2: Convergence behavior in case of full tensor spaces for best linear approximation and for interpolation, i.e. $Q_{\mathscr{I}_L^{-\infty}}$ and $I_{\mathscr{I}_L^{-\infty}}$, respectively.

	$f \in$	$\mathscr{H}^s$-error	dof M	convergence
$Q_{\mathscr{I}_L^{-\infty}}$	$\mathscr{H}^{s+r}, r > 0$	$\mathscr{O}(2^{-rL})$	$\mathscr{O}(2^{Ln})$	$\mathscr{O}(M^{-\frac{r}{n}})$
$I_{\mathscr{I}_L^{-\infty}}$	$\mathscr{H}^{s+r}, s+r > \frac{n}{2}, s \geq 0$	$\mathscr{O}(2^{-rL})$	$\mathscr{O}(2^{Ln})$	$\mathscr{O}(M^{-\frac{r}{n}})$

Table 3: Convergence behavior in case of general sparse grid tensor spaces for best linear approximation and for interpolation, i.e. $Q_{\mathscr{I}_L^T}$ and $I_{\mathscr{I}_L^T}$ with $T > 0$, respectively.

	$f \in$	$\mathscr{H}^s$-error	dof M	convergence
$Q_{\mathscr{I}_L^T}, 0 < T < \frac{s}{t}$	$\mathscr{H}^t_{mix}, t > 0$	$\mathscr{O}(2^{-(t-s)L})$	$\mathscr{O}(2^L)$	$\mathscr{O}(M^{-(t-s)})$
$I_{\mathscr{I}_L^T}, 0 < T < \frac{s}{t}$	$\mathscr{H}^t_{mix}, t > \frac{1}{2}$	$\mathscr{O}(2^{-(t-s)L})$	$\mathscr{O}(2^L)$	$\mathscr{O}(M^{-(t-s)})$

3.4 Convergence rates with respect to the degrees of freedom

Next, we cast the estimates on the degrees of freedom and the associated error of approximation by interpolation into a form which measures the error with respect to the involved degrees of freedom. We restrict ourselves to special cases, where the rates are independent of the dimension. We have the following results:

Lemma 9. *Let $L \in \mathbb{N}_0$, $0 < s < t$, $t > \frac{1}{2}$, $0 < T < \frac{r}{t}$, and $f \in \mathscr{H}^t_{\text{mix}}(\mathbb{T}^n)$ with a pointwise convergent Fourier series. Then it holds*

$$\|f - I_{\mathscr{I}_L^T} f\|_{\mathscr{H}^s} \lesssim M^{-(t-s)} \|f\|_{\mathscr{H}^t_{\text{mix}}},$$

with respect to the involved number of degrees of freedom $M := |V_{\mathscr{I}_L^T}|$.

Proof. This is a simple consequence of the Lemmata 1 and 8. First, we use the relation (22), that is

$$M = |V_{\mathscr{I}_L^T}| \leq c_1(n) \cdot 2^L$$

for $0 < T < \frac{s}{t}$, which results in $2^{-L} \leq c_1(n) M^{-1}$. We now plug this into (37), i.e. into the relation

$$\|f - I_{\mathscr{I}_L^T} f\|_{\mathscr{H}^r} \leq c_2(n) \cdot 2^{-L(t-s)} \cdot \|f\|_{\mathscr{H}^t_{\text{mix}}}$$

and arrive at the desired result with the order constant $C(n) = c_1(n)^{t-s} \cdot c_2(n)$. $\square$

Note that in [21] a result analogous to Lemma 9 is shown in the case of measuring the best linear approximation error with respect to the involved degrees of freedom.

There we have for $0 < T \leq \frac{s}{t}$, $t \geq 0$ the estimate

$$\|f - Q_{\mathscr{I}_L^T} f\|_{\mathscr{H}^s} \lesssim M^{-(t-s)} \|f\|_{\mathscr{H}^t_{mix}}.$$

Finally, let us discuss shortly a case of regular sparse grids with logarithmic terms involved. Here, again a simple consequence of the Lemmata 2 and 8 is, that for $L \in \mathbb{N}_0$, $t > \frac{1}{2}$ and $f \in \mathscr{H}^t_{\text{mix}}$ with a pointwise convergent Fourier series, there holds the relation

$$\|f - I_{\mathscr{I}_L^0} f\|_{\mathscr{L}_2} \lesssim M^{-t} L^{(t+1/2)(n-1)} \|f\|_{\mathscr{H}^t_{\text{mix}}} \lesssim M^{-t} \log(M)^{(t+1/2)(n-1)} \|f\|_{\mathscr{H}^t_{\text{mix}}}. \tag{39}$$

Altogether our theory shows that, in quite many cases, general sparse grids are a suitable approach to avoid the curse of dimensionality at least to a certain level of extent.

4 Numerical experiments and results

We implemented the generalized sparse grid trigonometric interpolation operator $I_{\mathscr{I}}$ for general admissible index sets $\mathscr{I} \subset \mathbb{N}_0^n$ according to Algorithm 1 in a software library called *HCFFT* [1]. This library also includes the functionality for the application of dimension-adaptive approaches, see [17, 46]. In addition to the fast discrete Fourier transform, which we deal with in this paper, it includes actually the following variants: fast discrete sine transform, fast discrete cosine transform, fast discrete Chebyshev transform, Legendre transform, generalized Hermite transform, Jacobi transform and Laguerre transform. The latter ones allow to deal with non-priodic situations, compare [60].

In the following, we present the results of some numerical calculations performed by the *HCFFT* library. We restrict ourselves to the case of the FFT based application of the interpolation operator $\mathscr{I}_L^T$. Here, we in particular study the dependence of the convergence rates on the number of dimensions for the regular sparse grid case $T = 0$ and the energy-norm like sparse grid case $T > 0$. To this end, we consider the approximation by interpolation of functions in the periodic Sobolev spaces of dominating mixed smoothness $\mathscr{H}^t_{\text{mix}}(\mathbb{T}^n)$. As test cases we use the functions

$$G_p : \mathbb{T}^n \to \mathbb{R} : \mathbf{x} \mapsto \bigotimes_{d=1}^{n} g_p(x_d),$$

with

$$g_p : \mathbb{T} \to \mathbb{R} : x \mapsto N_p \cdot (2 + \operatorname{sgn}(x - \pi) \cdot \sin(x)^p)$$

for $p = 1, 2, 3, 4$. Here, sgn denotes the sign function, i.e.

$$\operatorname{sgn}(x) := \begin{cases} -1 & x < 0, \\ 0 & x = 0, \\ 1 & x > 0 \end{cases}$$

and N_p denotes a normalization constant such that $\|g_p\|_{\mathscr{L}_2} = 1$. Note that for $\varepsilon > 0$ we have $g_p \in \mathscr{H}^{\frac{1}{2}+p-\varepsilon}(\mathbb{T})$ and thus $G_p \in \mathscr{H}_{\mathrm{mix}}^{\frac{1}{2}+p-\varepsilon}(\mathbb{T}^n)$. In particular, the $\mathscr{L}_2$- and $\mathscr{H}^1$-error can be computed by analytic formulae and the relative $\mathscr{L}_2$-error is equal to the absolute $\mathscr{L}_2$-error, i.e. $\|G_p - I_{\mathscr{I}_L^T} G_p\|_{\mathscr{L}_2} / \|G_p\|_{\mathscr{L}_2} = \|G_p - I_{\mathscr{I}_L^T} G_p\|_{\mathscr{L}_2}$. Note that these test functions are of simple product form, but the decay behavior of their Fourier coefficients reflects that of the considered Sobolev spaces of dominating mixed smoothness. The numerical results for more complicated functions of non-product structure from these Sobolev spaces were basically the same.

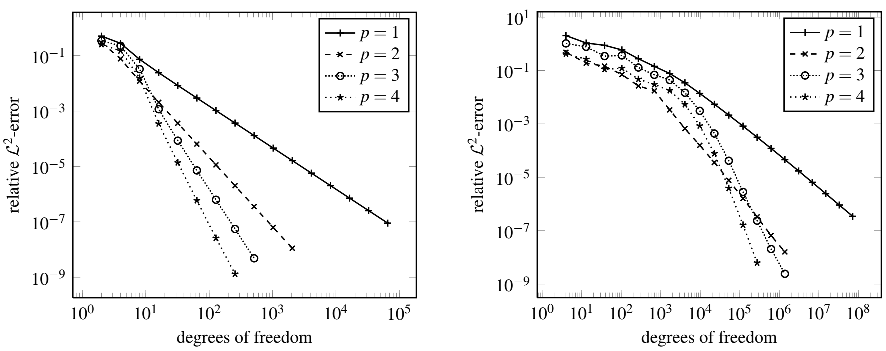

Fig. 1: Convergence behavior for approximating the functions $G_p \in \mathscr{H}_{\mathrm{mix}}^{\frac{1+2p}{2}-\varepsilon}$ by trigonometric interpolation on regular spare grids, i.e. $\|G_p - I_{\mathscr{I}_L^0} G_p\|_{\mathscr{L}_2}$ versus $|V_{\mathscr{I}_L^0}|$. *Left:* Case of $n = 1$. *Right:* Case of $n = 3$.

For validation we first performed numerical calculations in the one dimensional case for G_p with $p = 1,2,3,4$. We show the error measured in the $\mathscr{L}_2$-norm versus the number of degrees of freedom in Figure 1. To estimate the respective convergence rates, we computed a linear least square fit to the results of the three largest levels. This way, we obtained rates of values about 1.50, 2.50, 3.51 and 4.40, respectively, which coincide with the theoretically expected rates (i.e. $\frac{1}{2} + p, p = 1,2,3,4$) in the one-dimensional case, cf. Lemmata 1 and 8. Then, we performed calculations in the three-dimensional case for the regular spase grid case, i.e. $T = 0$. The values $p = 1,2,3,4$ result in numerically measured convergence rates of about 1.25, 1.87, 2.83 and 3.90, respectively.

Moreover, for the approximation of the test functions G_2 for up to six dimensions by trigonometric interpolation on regular sparse grids, we observe the rates which are given in Figure 2 and Table 4. For example in case of the $\mathscr{L}_2$-error the rates deteriorate from a value of 2.50 for $n = 1$ to a value of 1.56 for $n = 6$. Note that this

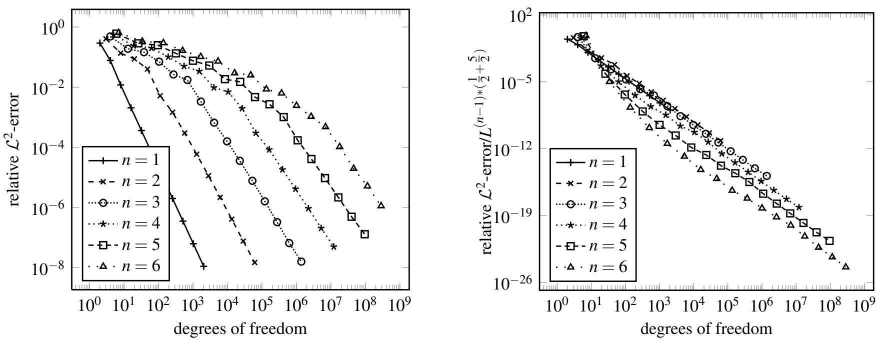

Fig. 2: Convergence behavior for approximating the function $G_2 \in \mathscr{H}_{\text{mix}}^{\frac{5}{2}-\varepsilon}$ by trigonometric interpolation on regular spare grids. *Left:* Case of relative/absolute $\mathscr{L}_2$-error, i.e. $\|G_2 - I_{\mathscr{I}_L^0} G_2\|_{\mathscr{L}_2}$. *Right:* Case of relative/absolute $\mathscr{L}_2$-error divided by $L^{(\frac{5}{2}+1/2)(n-1)}$.

decrease in the rates with respect to the number of dimensions is to be expected from our theory, since the estimates in Lemmata 1 and 8 involve dimension-dependent logarithmic terms for the regular sparse grid case, i.e. for $T=0$.

Table 4: Numerically measured convergence rates with respect to the number of degrees of freedom according to the relative $\mathscr{L}_2$-norm error and the relative $\mathscr{H}^1$-norm for the approximation of the function $G_2 \in \mathscr{H}_{\text{mix}}^{5/2-\varepsilon}$ by trigonometric interpolation on regular spare grids, i.e. $T=0$. Furthermore rates for the relative error divided by the respective logarithmic term versus the number of degrees of freedom, see also estimate (39).

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$\mathscr{L}_2$-norm	2.50	2.17	1.87	1.73	1.60	1.56
$\mathscr{L}_2$-norm / $L^{(\frac{5}{2}+1/2)(n-1)}$	2.50	2.50	2.40	2.44	2.46	2.59

We additionally give in Table 4 the computed rates associated to the relative errors divided by the respective logarithmic term versus the number of degrees of freedom. Here, the derived values fit quite well to the rates which could be expected from theory, that is, 2.5 for the error measured in the $\mathscr{L}_2$-norm.

Note that, according to Lemma 9, we can get rid of the logarithmic term in quite many cases.

As an example, if we measure the error of the approximation of G_2 by the general sparse grid interpolant $I_{\mathscr{I}_L^T} G_2$ in the $\mathscr{H}^1$-norm, then Lemma 9 leads with $r=1$,

$t = \frac{5}{2} - \varepsilon$ to a convergence rate of $\frac{3}{2} - \varepsilon$ for $0 < T < \frac{2}{5+2\varepsilon}$. To study such situations, we performed numerical calculations for the generalized sparse grids with $T = \frac{1}{8}$ and $T = \frac{1}{4}$. The obtained errors are plotted in Figure 3 and the corresponding rates are given in Table 5. The results show that the rates are substantially improved compared to the regular sparse grid case. Note that we still observe a slight decrease of the rates with the number of dimensions. This is surely a due to the fact that we are still in the pre-asymptotic regime for the higher dimensional cases. Note

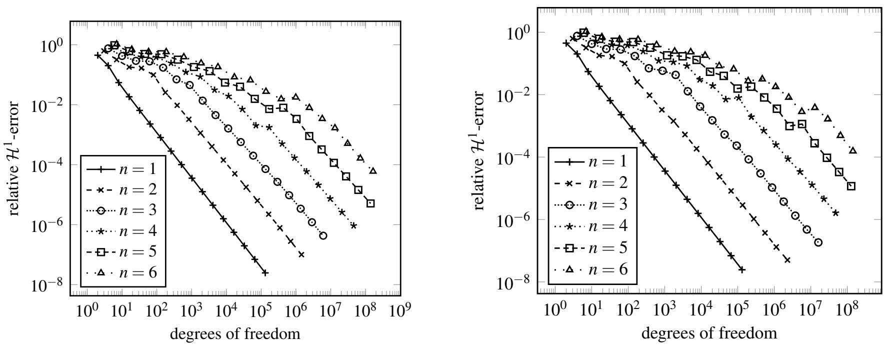

Fig. 3: Convergence behavior for approximating the function $G_2 \in \mathscr{H}^{5/2-\varepsilon}_{\mathrm{mix}}$ by the general sparse grid interpolation operator $I_{\mathscr{I}_L^T}$ with respect to the relative $\mathscr{H}^1$-error. *Left:* Case of $T = \frac{1}{8}$. *Right:* Case of $T = \frac{1}{4}$.

Table 5: Numerically measured convergence rates with respect to the number of degrees of freedom for the approximation of the function $G_2 \in \mathscr{H}^{5/2-\varepsilon}_{\mathrm{mix}}$ by trigonometric interpolation on generalized spare grids with $0 < T < \frac{2}{5}$.

error	T	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$\mathscr{H}^1$	$\frac{1}{8}$	1.50	1.44	1.39	1.29	1.28	1.27
$\mathscr{H}^1$	$\frac{1}{4}$	1.50	1.41	1.36	1.39	1.37	1.49

furthermore that the constant involved in the complexity estimate in Lemma 9 probably depends exponentially on the number n of dimensions. This explains the offset of the convergence with rising n in Figure 3.

In [39] it is noted that the involved order constant in the convergence rate estimate for the case $0 < T < 1$ is typically increasing with n and T and it is in particular larger than in the case of regular sparse grids with $T = 0$. In contrast, the convergence rate is superior in the case $0 < T < \frac{r}{t}$ to that of the regular sparse grid with $T = 0$, compare Table 3 and Table 1. Hence, in the pre-asymptotic regime, the effects of constants and order rates counterbalance each other a bit in practice. As an example, let us consider the $\mathscr{H}^1$-error of the interpolant $I_{\mathscr{I}_L^T} G_1$ for $T = 0, \frac{1}{8}, \frac{1}{4}$ and $n = 3, 4$. The associated computed rates are given in Table 6. Here, a break-even point can be seen from our numerical results depicted in Figure 4, i.e. for $n = 4$ the computed $\mathscr{H}^1$-error is slightly smaller in the case $T = \frac{1}{4}$ than in the cases $T = 0$ and $T = \frac{1}{8}$ for a number of involved degrees of freedom greater than about $|V_{\mathscr{I}_L^T}| \approx 10^6$. A similar effect is also present, albeit barely visible, for $n = 3$ and $|V_{\mathscr{I}_L^T}| \approx 10^5$. Nevertheless, in any case, the various rates are nearly the same anyway and these differences are quite small.

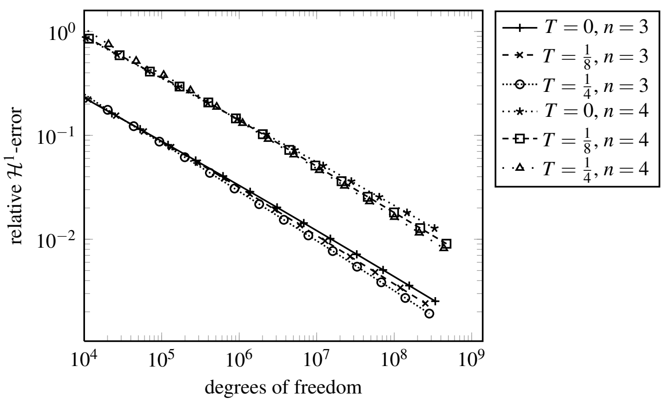

Fig. 4: Convergence behavior for the approximation of the function $G_1 \in \mathscr{H}^{3/2-\varepsilon}_{\text{mix}}$ by trigonometric interpolation on generalized spare grids, i.e. $\|G_1 - I_{\mathscr{I}_L^T} G_1\|_{\mathscr{L}_2}$ versus $|V_{\mathscr{I}_L^T}|$, where the error is measured in the relative $\mathscr{H}^1$-error, i.e. $\|G_1 - I_{\mathscr{I}_L^T} G_1\|_{\mathscr{H}^1} / \|G_1\|_{\mathscr{H}^1}$.

Table 6: Numerically measured convergence rates with respect to the number of degrees of freedom of the approximation of the function $G_1 \in \mathscr{H}^{3/2-\varepsilon}_{\text{mix}}$ by trigonometric interpolation on regular and generalized spare grids.

error	T	$n = 3$	$n = 4$
$\mathscr{H}^1$	0.0	0.45	0.42
$\mathscr{H}^1$	$\frac{1}{8}$	0.47	0.44
$\mathscr{H}^1$	$\frac{1}{4}$	0.49	0.47

5 Concluding Remarks

In this article, we discussed several variants of periodic Sobolev spaces of dominating mixed smoothness and we constructed the general sparse grid discretization spaces $V_{\mathscr{I}_L^T}$. We gave estimates for their number of degrees of freedom and the best linear approximation error for multivariate functions in $\mathscr{H}^{t,r}_{\mathrm{mix}}(\mathbb{T}^n)$ and $\mathscr{H}^{\mathbf{t}}_{\mathrm{mix}}(\mathbb{T}^n)$. Note here that these periodic Sobolev spaces, can be also generalized to the case of many-particle spaces [20, 21, 22, 31] in a straightforward way.

Moreover, we presented an algorithm for the general sparse grid interpolation based on the fast discrete Fourier transform, gave estimates for the involved number of degrees of freedom, estimates for the computation costs of the corresponding algorithm and the resulting error estimates for the general sparse grid interpolant $I_{\mathscr{I}_L^T}$ of functions in $\mathscr{H}^{t}_{\mathrm{mix}}(\mathbb{T}^n)$. Specifically, we identified smoothness assumptions that make it possible to choose $I_{\mathscr{I}_L^T}$ in such a way that the number of degrees of freedom is $\mathscr{O}(2^L)$ compared to $\mathscr{O}(2^L L^{n-1})$ and $\mathscr{O}(2^{nL})$ for the regular sparse grid (i.e. $T=0$) and full tensor grid spaces (i.e. $T=-\infty$), respectively, while keeping the optimal order of approximation. For this case, we also showed that the asymptotic computational cost complexities rates are independent of the number of dimensions. The constants involved in the $\mathscr{O}$-notation may still depend exponentially on n however.

Let us furthermore note that we mainly discussed the sparse grid interpolation operator $I_{\mathscr{I}_L^T}$ for the periodic case in the present paper. However, our implemented software library HCFFT allows us to deal with discretization spaces associated with arbitrary admissible index sets and also features spaces with finite-order weights and dimension-adaptive methods. Furthermore, discrete cosine, sine, Chebyshev, Legendre, generalized Hermite, Jacobi transform and Laguerre transforms, which allow in particular to deal also with non-periodic situations, compare [60]. We will discuss these approaches and their possible applications in the area of uncertainty quantification [44] in a forthcoming paper.

Acknowledgements This work was supported in part by the Collaborative Research Centre 1060 of the Deutsche Forschungsgemeinschaft.

References

1. HCFFT Manual v2.0. `http://www.hcfft.org`.
2. R. Adams. *Sobolev Spaces*. Academic Press, 1975.
3. V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12:273–288, 2000.
4. G. Baszenski and F. Delvos. A discrete Fourier transform scheme for Boolean sums of trigonometric operators. *International Series of Numerical Mathematics*, 90:15–24, 1989.
5. R. Bellmann. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961.
6. B. Bohn and M. Griebel. An adaptive sparse grid approach for time series predictions. In J. Garcke and M. Griebel, editors, *Sparse grids and applications*, volume 88 of *Lecture Notes in Computational Science and Engineering*, pages 1–30. Springer, 2012.

7. H. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:1–123, 2004.
8. D. Dũng. Best multivariate approximations by trigonometric polynomials with frequencies from hyperbolic crosses. *Journal of Approximation Theory*, 91(2):205–225, 1997.
9. D. Dũng. B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness. *Journal of Complexity*, 27(6):541–567, 2011.
10. D. Dũng. Sampling and cubature on sparse grids based on a B-spline quasi-interpolation. *Foundations of Computational Mathematics*, 16(5):1193–1240, 2016.
11. D. Dũng, V. Temlyakov, and T. Ullrich. *Hyperbolic Cross Approximation*. Advanced Courses in Mathematics - CRM Barcelona. Birkhäuser/Springer, 2018.
12. D. Dũng and T. Ullrich. N-widths and ε-dimensions for high-dimensional approximations. *Foundations of Computational Mathematics*, 13:965–1003, 2013.
13. B. Efron and C. Stein. The Jackknife estimate of variance. *The Annals of Statistics*, 9(3):586–596, 1981.
14. C. Feuersänger. *Sparse Grid Methods for Higher Dimensional Approximation*. Dissertation, Institute for Numerical Simulation, University of Bonn, 2010.
15. C. Feuersänger and M. Griebel. Principal manifold learning by sparse grids. *Computing*, 85(4):267–299, 2009.
16. J. Garcke. *Maschinelles Lernen durch Funktionsrekonstruktion mit verallgemeinerten dünnen Gittern*. Dissertation, Institute for Numerical Simulation, University of Bonn, 2004.
17. T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71(1):65–87, 2003.
18. V. Gradinaru. Fourier transform on sparse grids: Code design and the time dependent Schrödinger equation. *Computing*, 80:1–22, 2007.
19. M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. In L. Pardo, A. Pinkus, E. Suli, and M. Todd, editors, *Foundations of Computational Mathematics (FoCM05), Santander*, pages 106–161. Cambridge University Press, 2006.
20. M. Griebel and J. Hamaekers. A wavelet based sparse grid method for the electronic Schrödinger equation. In M. Sanz-Solé, J. Soria, J. Varona, and J. Verdera, editors, *Proceedings of the International Congress of Mathematicians*, volume III, pages 1473–1506, Madrid, Spain, August 22–30 2006. European Mathematical Society.
21. M. Griebel and J. Hamaekers. Sparse grids for the Schrödinger equation. *Mathematical Modelling and Numerical Analysis*, 41(2):215–247, 2007.
22. M. Griebel and J. Hamaekers. Tensor product multiscale many-particle spaces with finite-order weights for the electronic Schrödinger equation. *Zeitschrift für Physikalische Chemie*, 224:527–543, 2010.
23. M. Griebel and J. Hamaekers. Fast discrete Fourier transform on generalized sparse grids. In *Sparse grids and Applications*, volume 97 of *Lecture Notes in Computational Science and Engineering*, pages 75–108. Springer, 2014.
24. M. Griebel and H. Harbrecht. A note on the construction of L-fold sparse tensor product spaces. *Constructive Approximation*, 38(2):1–17, 2013.
25. M. Griebel and H. Harbrecht. On the construction of sparse tensor product spaces. *Mathematics of Computations*, 82(268):975–994, 2013.
26. M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *Journal of Complexity*, 26:455–489, 2010.
27. M. Griebel and S. Knapek. Optimized tensor-product approximation spaces. *Constructive Approximation*, 16(4):525–540, 2000.
28. M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Mathematics of Computation*, 78:2223–2257, 2009.
29. M. Griebel, P. Oswald, and T. Schiekofer. Sparse grids for boundary integral equations. *Numerische Mathematik*, 83(2):279–312, 1999.
30. K. Hallatschek. Fourier-transform on sparse grids with hierarchical bases. *Numerische Mathematik*, 63(1):83–97, 1992.
31. J. Hamaekers. *Sparse Grids for the Electronic Schrödinger Equation: Construction and Application of Sparse Tensor Product Multiscale Many-Particle Spaces with Finite-Order Weights for Schrödinger's Equation*. Südwestdeutscher Verlag für Hochschulschriften, Saarbrücken, 2010.

32. D. Haroske and H. Triebel. *Distributions, Sobolev Spaces, Elliptic Equations*. European Mathematical Society, 2007.
33. A. Hinrichs, E. Novak, and J. Vybíral. Linear information versus function evaluations for L_2-approximation. *Journal of Approximation Theory*, 153(1):97–107, 2008.
34. J. D. Jakeman and S. G. Roberts. Local and dimension adaptive stochastic collocation for uncertainty quantification. In *Sparse grids and applications*, pages 181–203. Springer, 2012.
35. Y. Jiang and Y. Xu. Fast discrete algorithms for sparse Fourier expansions of high dimensional functions. *Journal of Complexity*, 26(1):51–81, 2010.
36. L. Kämmerer and S. Kunis. On the stability of the hyperbolic cross discrete Fourier transform. *Numerische Mathematik*, 117:581–600, 2011.
37. L. Kämmerer, S. Kunis, and D. Potts. Interpolation lattices for hyperbolic cross trigonometric polynomials. *Journal of Complexity*, 28(1):76–92, 2012.
38. A. Klimke and B. Wohlmuth. Algorithm 847: spinterp: Piecewise multilinear hierarchical sparse grid interpolation in MATLAB. *ACM Transactions on Mathematical Software*, 31(4):561–579, 2005.
39. S. Knapek. *Approximation und Kompression mit Tensorprodukt-Multiskalenräumen*. Dissertation, University of Bonn, 2000.
40. S. Knapek. Hyperbolic cross approximation of integral operators with smooth kernel. 2000. Technical Report 665, SFB 256, University of Bonn.
41. T. Kühn, W. Sickel, and T. Ullrich. Approximation of mixed order Sobolev functions on the d-torus – Asymptotics, preasymptotics and d-dependence. *ArXiv e-prints*, 2014. arXiv:1312.6386 [math.NA].
42. F. Kupka. *Sparse Grid Spectral Methods for the Numerical Solution of Partial Differential Equations with Periodic Boundary Conditions*. PhD thesis, University of Wien, 1997.
43. F. Kupka. Sparse grid spectral methods and some results from approximation theory. In C. Lai, P. Bjørstad, M. Cross, and O. Widlund, editors, *Proceedings of the 11th International Conference on Domain Decomposition Methods in Greenwich*, pages 57–64, England, 1999.
44. O. Le Maître and O. Knio. *Spectral Methods for Uncertainty Quantification*, volume XVI of *Scientific Computation*. Springer, 2010.
45. X. Ma and N. Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *Journal of Computational Physics*, 228(8):3084–3113, 2009.
46. K. Matuschke. Trigonometrische Interpolation auf verallgemeinerten dünnen Gittern mit beliebiger Levelstruktur. Diplomarbeit, Institut für Numerische Simulation, Universität Bonn, 2014.
47. E. Novak and H. Wozniakowski. On the power of function values for the approximation problem in various settings. *Surveys in Approximation Theory*, 6:1–23, 2011.
48. J. Pasciak. Spectral and pseudospectral methods for advection equations. *Math. Comput.*, 35:1081–1092, 1980.
49. H. Schmeisser and W. Sickel. Spaces of functions of mixed smoothness and approximation from hyperbolic crosses. *Journal of Approximation Theory*, 128(2):115–150, 2004.
50. H. Schmeisser and H. Triebel. *Fourier Analysis and Functions Spaces*. John Wiley, 1987.
51. J. Shen and L. Wang. Sparse spectral approximations of high-dimensional problems based on hyperbolic cross. *SIAM Journal on Numerical Analysis*, 48(3):1087–1109, 2010.
52. W. Sickel and F. Sprengel. Interpolation on sparse grids and tensor products of Nikol'skij–Besov spaces. *Journal of Computational Analysis and Applications*, 1:263–288, 1999.
53. W. Sickel and T. Ullrich. Tensor products of Sobolev–Besov spaces and applications to approximation from the hyperbolic cross. *Journal of Approximation Theory*, 161(2):748–786, 2009.
54. W. Sickel and T. Ullrich. Spline interpolation on sparse grids. *Applicable Analysis*, 90(3-4):337–383, 2011.
55. I. Sloan, X. Wang, and H. Woźniakowski. Finite-order weights imply tractability of multivariate integration. *Journal of Complexity*, 20(1):46–74, 2004.

56. S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Math. Dokl.*, 4:240–243, 1963. Russian original in Dokl. Akad. Nauk SSSR, 148 (1963), pp. 1042–1045.
57. R. Stanley. What is enumerative combinatorics? In *Enumerative combinatorics*, pages 1–63. Springer, 1986.
58. V. Temlyakov. *Approximation of Periodic Functions*. Nova Science, New York, 1993.
59. H. Triebel. *Bases in function spaces, sampling, discrepancy, numerical integration*, volume 11. European Mathematical Society, 2010.
60. V. Velikov. Fast Sparse Pseudo-spectral Methods for High-dimensional Problems. Master thesis, Institute for Numerical Simulation, Universität Bonn, 2016.
61. T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *Mathematical Modelling and Numerical Analysis*, 38:93–127, 2004.
62. G. Wasilkowski and H. Woźniakowski. Finite-order weights imply tractability of linear multivariate problems. *Journal of Approximation Theory*, 130(1):57–77, 2004.
63. A. Werschulz and H. Woźniakowski. Tractability of quasilinear problems I: General results. *Journal of Approximation Theory*, 145(2):266–285, 2007.