



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

March 2020

Generalized Sparse Grid Interpolation Based on the Fast Discrete Fourier Transform

Michael Griebel and Jan Hamaekers

Abstract In [9], an algorithm for trigonometric interpolation involving only so-called *standard information* of multivariate functions on generalized sparse grids has been suggested and a study on its application for the interpolation of functions in periodic Sobolev spaces of dominating mixed smoothness has been presented. In this complementary paper, we now give a slight modification of the proofs, which yields an extension from the pairing $(\mathcal{H}^s, \mathcal{H}_{mix}^t)$ to the more general pairing $(\mathcal{H}^s, \mathcal{H}_{mix}^{t,r})$ and which in addition results in an improved estimate for the interpolation error. The improved (constructive) upper bound is in particular consistent with the lower bound for sampling on regular sparse grids with $r = 0$ and $s = 0$ given in [4, 5].

1 Introduction

This is an addendum to our previous paper [9]. Throughout this article we will use the definitions and notation given therein. As noted in [9], so-called *sparse grid* based approaches [2, 12] have emerged as useful techniques to tackle higher-dimensional problems, since they allow to break the curse of dimensionality under certain conditions. For example, for the periodic Sobolev spaces

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

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

of bounded mixed smoothness on the n -dimensional torus \mathbb{T}^n which, e.g. indeed appear with $t = \frac{3}{4} - \varepsilon$, $r = 1$ for the solution of electronic Schrödinger equation [20], a specific generalization of the regular sparse grid spaces based on Fourier polynomials $e^{i\mathbf{k}^T \mathbf{x}}$ and associated Fourier coefficients $\hat{f}_{\mathbf{k}}$ with frequencies \mathbf{k} from the generalized hyperbolic cross

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

were introduced in [18] and further discussed in [9, 11, 12, 14, 19]. 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 $\mathcal{O}(2^L L^{n-1})$ frequencies are involved. Furthermore, the case $T = -\infty$ corresponds to the full tensor grid. The case $T \rightarrow 1$ 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 $\mathcal{O}(2^L)$ only.

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 function values, which is of interest in many practical applications. In this case, it is in general not clear if the approximation error of an associated *interpolant* exhibits the same order of the convergence rate as that of the best linear approximation [15, 25].

In [9], we have shown that, if $L \in \mathbb{N}_0$, $T < 1$, $s < t$, $t > \frac{1}{2}$ and $f \in \mathcal{H}_{\text{mix}}^{t,0} := \mathcal{H}_{\text{mix}}^{t,0}$ with a pointwise convergent Fourier series, it holds

$$\|f - I_{\mathcal{J}_L^T} f\|_{\mathcal{H}^s} \lesssim \begin{cases} 2^{-((t-s)+(Tt-s)\frac{n-1}{n-T})L} L^{n-1} \|f\|_{\mathcal{H}_{\text{mix}}^t} & \text{for } T \geq \frac{s}{t}, \\ 2^{-(t-s)L} \|f\|_{\mathcal{H}_{\text{mix}}^t} & \text{for } T < \frac{s}{t}, \end{cases} \quad (1)$$

where $I_{\mathcal{J}_L^T}$ denotes the general sparse grid interpolation operator, as defined via (13) with

$$\mathcal{J}_L^T := \{\mathbf{1} : \|\mathbf{1}\|_1 - T\|\mathbf{1}\|_\infty \leq (1-T)L\}, \quad T < 1. \quad (2)$$

Analogous estimates for the specific case of regular sparse grids, i.e. $T = 0$, can be found e.g. in [23, 24, 27]. Moreover, for $T = 0$, estimates with an improved logarithmic term, i.e. $L^{\frac{n-1}{2}}$, are given in [3, 6, 22, 26, 28].

In this addendum to our previous article [9] we now show the following: 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 $f \in \mathcal{H}_{\text{mix}}^{t,r}$ with a pointwise convergent Fourier series. Then it holds

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

Thus, we extend (1) from the pairing $(\mathcal{H}^s, \mathcal{H}_{mix}^t)$ to the more general case of $(\mathcal{H}^s, \mathcal{H}_{mix}^{t,r})$. This allows us to also treat the case $r > 0$, in contrast to the previous estimate (1). At the same time we prove for the first clause a logarithmic term of type $L^{(n-1)/2}$ only, which improves on the L^{n-1} term contained in the first clause of (1). Moreover, for the specific case $T = 0, s = 0, r = 0$, this upper bound is in particular consistent with the lower bound given in case of sampling on the Smolyak grid (regular sparse grid) in [4, 5, 17].

The remainder of this paper is organized as follows: In section 2, we summarize necessary definitions. In section 3, we present our improved error estimate. In section 4, we discuss our new result and give some concluding remarks.

2 Fourier-based approximation for general sparse grids

In this section we will shortly recall necessary definitions given in our previous article [9]. 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 $\mathbf{x} := (x_1, \dots, x_n)$, where $x_d \in \mathbb{T}$. We define the basis function associated to a multi-index $\mathbf{k} = (k_1, \dots, k_n) \in \mathbb{Z}^n$ by

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

Every $f \in \mathcal{L}_2(\mathbb{T}^n)$ has the unique expansion $f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} \hat{f}_{\mathbf{k}} \omega_{\mathbf{k}}(\mathbf{x})$, where the Fourier coefficients are given by

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

In the following, let us now define finite-dimensional subspaces of the space $\mathcal{L}_2(\mathbb{T}^n) \cong \mathcal{H}^0(\mathbb{T}^n)$ and associated appropriate interpolation operators. To this end, we first consider the one-dimensional case, i.e. $n = 1$, and we set

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

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

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

and the corresponding spaces $V_l := \text{span}\{\mathcal{B}_l\}$.

Now, let the Fourier series $\sum_{\mathbf{k} \in \mathbb{Z}^n} \hat{f}_{\mathbf{k}} \omega_{\mathbf{k}}$ be pointwise convergent to $f(x)$. Then, for interpolation points $\mathcal{S}_l := \{m \frac{2\pi}{2^l} : m = 0, \dots, 2^l - 1\}$ of level $l \in \mathbb{N}_0$, the interpolation operator can be defined by $I_l : V \rightarrow V_l : f \mapsto I_l f := \sum_{j \in \mathcal{G}_l} \hat{f}_j^{(l)} \phi_j$ with indices $\mathcal{G}_l :=$

$\{0, \dots, 2^l - 1\}$ and discrete nodal Fourier coefficients

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

which only involve point evaluations $f(x)$ at points $x \in \mathcal{S}_l$. This way, the 2^l interpolation conditions $f(x) = I_l f(x)$ for all $x \in \mathcal{S}_l$ are fulfilled. In particular, from (6) and (3), one can deduce the well-known aliasing formula

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

Next, we introduce an univariate Fourier hierarchical basis function for $j \in \mathbb{N}_0$ by

$$\psi_j := \begin{cases} \phi_0 & \text{for } j = 0, \\ \phi_j - \phi_{2^{l-1}-j} & \text{for } 2^{l-1} \leq j \leq 2^l - 1, l \geq 1, \end{cases} \quad (8)$$

and we define the one-dimensional hierarchical Fourier basis including basis functions up to level $l \in \mathbb{N}_0$ by $\mathcal{B}_l^h := \{\psi_j\}_{0 \leq j \leq 2^l - 1}$.

Let us further introduce the difference spaces

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

Note that there holds the relation $V_l = \text{span}\{\mathcal{B}_l\} = \text{span}\{\mathcal{B}_l^h\}$ for all $l \in \mathbb{N}_0$. Thus, we have the decomposition of V_l into the direct sum $V_l = \bigoplus_{v=0}^l W_v$. 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 $u = \sum_{0 \leq j \leq 2^l - 1} u_j^h \psi_j$, where $u_j^h \in \mathbb{R}$, to the nodal representation $u = \sum_{0 \leq j \leq 2^l - 1} u_j \phi_j$ by a linear transform.

Next, we define the difference operator

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

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

$$\Delta_l f = \sum_{0 \leq v < l} \sum_{j \in \mathcal{J}_v} (\hat{f}_j^{(l)} - \hat{f}_j^{(l-1)}) \phi_j + \sum_{j \in \mathcal{J}_l} \hat{f}_j^{(l)} \phi_j =: \sum_{j \in \mathcal{J}_l} \check{f}_j \psi_j \quad (9)$$

with

$$\mathcal{J}_l := \begin{cases} \{0\} & \text{for } l = 0, \\ \{2^{l-1}, \dots, 2^l - 1\} & \text{for } l \geq 1. \end{cases}$$

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

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

In particular, let us note that the interpolation relation

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

holds, where $\mathcal{S}_l^h := \mathcal{S}_l \setminus \mathcal{S}_{l-1}$, with $\mathcal{S}_{-1} := \emptyset$.

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

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

follows by the definitions (8) and (9). 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 $j \in \mathcal{J}_l$. Hence in the case $l \in \mathbb{N}_0$, $j \in \mathcal{J}_l$, we obtain the relation

$$\check{f}_j = \hat{f}_j^{(l)} = \sum_{m \in \mathbb{Z}} \hat{f}_{\sigma(j) + m 2^l} \quad (10)$$

with the help of the aliasing formula (7).

Now, let us consider the multivariate case. 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, we introduce the n -dimensional interpolation operator on full tensor grids as

$$I_1 := I_1 \otimes \cdots \otimes I_n : V \rightarrow V_1 : f \mapsto I_1 f = \sum_{\mathbf{j} \in \mathcal{G}_1} \hat{f}_{\mathbf{j}}^{(1)} \phi_{\mathbf{j}}, \quad \text{with } \mathcal{G}_1 := \mathcal{G}_1 \times \cdots \times \mathcal{G}_n \subset \mathbb{N}_0^n$$

and multi-dimensional discrete nodal Fourier coefficients

$$\hat{f}_{\mathbf{j}}^{(1)} := 2^{-|\mathbf{l}_1|} \sum_{\mathbf{x} \in \mathcal{S}_1} f(\mathbf{x}) \phi_{\mathbf{j}}^*(\mathbf{x}), \quad \text{where } \mathcal{S}_1 := \mathcal{S}_1 \times \cdots \times \mathcal{S}_n \subset \mathbb{T}^n.$$

Moreover, for the n -dimensional case, we use a tensor product ansatz to construct n -dimensional basis functions as well as spaces. To this end, for a multi-index $\mathbf{l} \in \mathbb{N}_0^n$, we define finite-dimensional spaces by a tensor product construction, i.e. $V_1 := \bigotimes_{d=1}^n V_{l_d}$. Furthermore, we introduce the space $V := \sum_{\mathbf{l} \in \mathbb{N}_0^n} V_1$ and we set $\psi_{\mathbf{j}} := \bigotimes_{d=1}^n \psi_{j_d}$ and $W_1 := \bigotimes_{d=1}^n W_{l_d}$ for $\mathbf{l} \in \mathbb{N}_0^n$. In addition, we define $W_{\mathcal{J}} := \bigoplus_{\mathbf{l} \in \mathcal{J}} W_1$ for an index set $\mathcal{J} \subset \mathbb{N}_0^n$. Now, similar to (7), there holds the multi-dimensional aliasing formula

$$\hat{f}_{\mathbf{j}}^{(1)} = \sum_{\mathbf{m} \in \mathbb{Z}^n} \hat{f}_{\sigma(\mathbf{j}) + \mathbf{m} 2^1}, \quad (11)$$

where $\sigma(\mathbf{j}) := (\sigma(j_1), \dots, \sigma(j_n))$ and $\mathbf{m} 2^1 := (m_1 2^{l_1}, \dots, m_n 2^{l_n})$.

For the general sparse grid construction, we restrict ourselves to index sets, which obey the admissibility condition given in [7, 13]. For any admissible index set \mathcal{J} , we define generalized sparse grid space by

$$V_{\mathcal{J}} := \sum_{\mathbf{l} \in \mathcal{J}} V_{\mathbf{l}} = \bigoplus_{\mathbf{l} \in \mathcal{J}} W_{\mathbf{l}} = W_{\mathcal{J}}. \quad (12)$$

Moreover, we introduce the corresponding general sparse grid trigonometric interpolation operator by

$$I_{\mathcal{J}} := \sum_{\mathbf{l} \in \mathcal{J}} \Delta_{\mathbf{l}} : V \rightarrow V_{\mathcal{J}}, \quad \text{where } \Delta_{\mathbf{l}} := \Delta_{l_1} \otimes \cdots \otimes \Delta_{l_n} : V \rightarrow W_{\mathbf{l}}. \quad (13)$$

The associated set of interpolation points is given by

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

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

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

In particular, the hierarchical Fourier series $\sum_{\mathbf{l} \in \mathbb{N}_0^n} \sum_{\mathbf{j} \in \mathcal{I}_{\mathbf{l}}} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}}$ converges pointwise to f on all grids $\mathcal{I}_{\mathbf{l}}$, $\mathbf{l} \in \mathbb{N}_0^n$. Furthermore, with the help of the multi-dimensional aliasing formula (11), a relation similar to (10) can easily be deduced, that is, for $\mathbf{l} \in \mathbb{N}_0^n$ and $\mathbf{j} \in \mathcal{I}_{\mathbf{l}}$, it holds

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

3 Approximation error of the interpolant

In the following, we consider the error of the approximation by trigonometric *interpolation*. We proceed in two steps. First, we introduce functions $\tilde{\beta}, \beta$, which determine Sobolev spaces of general smoothness $\mathcal{H}_{\tilde{\beta}}, \mathcal{H}_{\beta}$ and we derive an upper bound of the interpolation operator $I_{\mathcal{J}}$ in terms of the pairing $(\mathcal{H}_{\tilde{\beta}}, \mathcal{H}_{\beta})$. Then, we invoke for $\tilde{\beta}$ and β the specific functions which characterize the norm of \mathcal{H}^s in which the interpolation error will be measured and the regularity assumption $f \in \mathcal{H}_{mix}^{l,r}$, respectively. Finally, we derive our new bound in Lemma 1.

To this end, let $\beta : \mathbb{Z}^n \rightarrow \mathbb{R}_+$ be a continuous and positive function, which implicitly expresses some smoothness class and defines the Sobolev space given by

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

Now, let us consider two smoothness functions $\tilde{\beta}$ and β with associated Sobolev spaces $\mathcal{H}_{\tilde{\beta}}$ and \mathcal{H}_β and norms $\|f\|_{\mathcal{H}_{\tilde{\beta}}}$ and $\|f\|_{\mathcal{H}_\beta}$, respectively. It should hold $\mathcal{H}_\beta \subset \mathcal{H}_{\tilde{\beta}} \subset \mathcal{L}_2$ and thus $\beta(\mathbf{k}) \lesssim \tilde{\beta}(\mathbf{k})$.

Let furthermore $f \in \mathcal{H}_\beta$ obey a pointwise convergent (hierarchical) Fourier series. Then, the relation

$$\begin{aligned} \|f - I_{\mathcal{J}} f\|_{\mathcal{H}_{\tilde{\beta}}} &= \left\| \sum_{\mathbf{l} \in \mathbb{N}_0^n} \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} - \sum_{\mathbf{l} \in \mathcal{J}} \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}} = \left\| \sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}} \\ &\leq \sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} \left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}} \end{aligned} \quad (14)$$

holds. By definition of the hierarchical basis (8) we obtain

$$\begin{aligned} \left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}}^2 &= \left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{v} \in \{0,1\}^n} \check{f}_{\mathbf{j}} \bigotimes_{d=1}^n \phi_{\mu_{\mathbf{v}}^d(j_d)} \right\|_{\mathcal{H}_{\tilde{\beta}}}^2 \\ &= \sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{v} \in \{0,1\}^n, 1-\mathbf{v} \geq \mathbf{0}} |\check{f}_{\mathbf{j}}|^2 \tilde{\beta}(\sigma(\mu_{\mathbf{v}}^1(\mathbf{j})))^2 \\ &= \sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{v} \in \{0,1\}^n, 1-\mathbf{v} \geq \mathbf{0}} \left| \sum_{\mathbf{m} \in \mathbb{Z}^n} \hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^1} \right|^2 \tilde{\beta}(\sigma(\mu_{\mathbf{v}}^1(\mathbf{j})))^2 \\ &= \sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{v} \in \{0,1\}^n, 1-\mathbf{v} \geq \mathbf{0}} \left| \sum_{\mathbf{m} \in \mathbb{Z}^n} \hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^1} \frac{\beta(\sigma(\mathbf{j})+\mathbf{m}2^1)}{\beta(\sigma(\mathbf{j})+\mathbf{m}2^1)} \right|^2 \tilde{\beta}(\sigma(\mu_{\mathbf{v}}^1(\mathbf{j})))^2, \end{aligned}$$

where

$$\mu_0^l(j) = j, \quad \mu_1^l(j) = \begin{cases} -1 & \text{if } l \leq 0, \\ 2^l - 1 - j & \text{if } l \geq 1, \end{cases}$$

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

$$\begin{aligned} &\left| \sum_{\mathbf{m} \in \mathbb{Z}^n} \hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^1} \frac{\beta(\sigma(\mathbf{j})+\mathbf{m}2^1)}{\beta(\sigma(\mathbf{j})+\mathbf{m}2^1)} \right|^2 \\ &\leq \left(\sum_{\mathbf{m} \in \mathbb{Z}^n} \left| \hat{f}_{\sigma(\mathbf{j})+\mathbf{m}2^1} \beta(\sigma(\mathbf{j})+\mathbf{m}2^1) \right|^2 \right) \left(\sum_{\mathbf{m} \in \mathbb{Z}^n} \beta(\sigma(\mathbf{j})+\mathbf{m}2^1)^{-2} \right) \end{aligned} \quad (15)$$

and hence it holds

$$\begin{aligned} \left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}}^2 &\leq \sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{v} \in \{0,1\}^n, \mathbf{1} - \mathbf{v} \geq \mathbf{0}} \left(\sum_{\mathbf{m} \in \mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}}}|^2 |\beta(\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}})|^2 \right) \times \\ &\quad \times \left(\sum_{\mathbf{m} \in \mathbb{Z}^n} \beta(\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}})^{-2} \right) \tilde{\beta}(\sigma(\mu_{\mathbf{v}}^{\mathbf{l}}(\mathbf{j})))^2. \end{aligned}$$

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

$$\tilde{\beta}(\sigma(\mu_{\mathbf{v}}^{\mathbf{l}}(\mathbf{j})))^2 \sum_{\mathbf{m} \in \mathbb{Z}^n} |\beta(\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}})|^{-2} \leq C^2 g(\mathbf{l})^2 \quad (16)$$

for all $\mathbf{j} \in \mathcal{J}_1$ and $\mathbf{v} \in \{0,1\}^n$, $\mathbf{1} - \mathbf{v} \geq \mathbf{0}$, with a constant C independent of \mathbf{j} and \mathbf{v} . Then, with $|\{0,1\}^n| = 2^n$, we have

$$\left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}} \leq 2^n C g(\mathbf{l}) \left(\sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{m} \in \mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}}}|^2 \beta(\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}})^2 \right)^{\frac{1}{2}},$$

which results with relation (14) in

$$\sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} \left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}} \leq 2^n C \sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} g(\mathbf{l}) \left(\sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{m} \in \mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}}}|^2 \beta(\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}})^2 \right)^{\frac{1}{2}}.$$

With Hölder's inequality we obtain

$$\begin{aligned} \sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} \left\| \sum_{\mathbf{j} \in \mathcal{J}_1} \check{f}_{\mathbf{j}} \psi_{\mathbf{j}} \right\|_{\mathcal{H}_{\tilde{\beta}}} &\leq 2^n C \left(\sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} g(\mathbf{l})^2 \right)^{\frac{1}{2}} \times \\ &\quad \times \left(\sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} \sum_{\mathbf{j} \in \mathcal{J}_1} \sum_{\mathbf{m} \in \mathbb{Z}^n} |\hat{f}_{\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}}}|^2 \beta(\sigma(\mathbf{j}) + \mathbf{m}2^{\mathbf{l}})^2 \right)^{\frac{1}{2}}. \end{aligned}$$

This leads finally to the estimate

$$\|f - I_{\mathcal{J}} f\|_{\mathcal{H}_{\tilde{\beta}}} \lesssim \left(\sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}} g(\mathbf{l})^2 \right)^{\frac{1}{2}} \|f\|_{\mathcal{H}_{\tilde{\beta}}}. \quad (17)$$

Let us now specifically consider the error in the \mathcal{H}^s -norm for approximating $f \in \mathcal{H}_{\text{mix}}^{l,r}$ in the sparse grid space $V_{\mathcal{J}_L}$ by interpolation. To this end, we specify

$$\beta(\mathbf{k}) := \lambda_{\text{iso}}(\mathbf{k})^r \lambda_{\text{mix}}(\mathbf{k})^l \quad \text{and} \quad \tilde{\beta}(\mathbf{k}) := \lambda_{\text{iso}}(\mathbf{k})^s, \quad (18)$$

where

$$\lambda_{\text{iso}}(\mathbf{k}) := 1 + |\mathbf{k}|_{\infty} \quad \text{and} \quad \lambda_{\text{mix}}(\mathbf{k}) := \prod_{d=1}^n (1 + |k_d|). \quad (19)$$

Furthermore, let us recall the following upper bound: For $L \in \mathbb{N}_0$, $T < 1$, $s < t$ and $t \geq 0$ it holds

$$\sum_{\mathbf{l} \in \mathbb{N}_0^n \setminus \mathcal{J}_L^T} 2^{-t\|\mathbf{l}\|_1 + s\|\mathbf{l}\|_{\infty}} \lesssim \begin{cases} 2^{-((t-s) + (Tt-s)\frac{n-1}{n-T})L} L^{n-1} & \text{for } T \geq \frac{s}{t}, \\ 2^{-(t-s)L} & \text{for } T < \frac{s}{t}. \end{cases} \quad (20)$$

A proof is for example given in [18], Theorem 4.

Now, we can give the following lemma:

Lemma 1. *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 $f \in \mathcal{H}_{\text{mix}}^{t,r}$ with a pointwise convergent Fourier series. Then it holds*

$$\|f - I_{\mathcal{J}_L^T} f\|_{\mathcal{H}^s} \lesssim \begin{cases} 2^{-((t-(s-r)) + (Tt-(s-r))\frac{n-1}{n-T})L} L^{\frac{n-1}{2}} \|f\|_{\mathcal{H}_{\text{mix}}^{t,r}} & \text{for } T \geq \frac{s-r}{t}, \\ 2^{-(t-(s-r))L} \|f\|_{\mathcal{H}_{\text{mix}}^{t,r}} & \text{for } T < \frac{s-r}{t}. \end{cases} \quad (21)$$

Proof. For $\mathbf{j} \in \mathcal{J}_1$ and $\mathbf{v} \in \{0, 1\}^n$ with $\mathbf{1} - \mathbf{v} \geq \mathbf{0}$ it follows the relation

$$\begin{aligned} & \sum_{\mathbf{m} \in \mathbb{Z}^n} \prod_{d=1}^n \left(1 + |\sigma(j_d) + m_d 2^{l_d}|\right)^{-2t} (1 + |\sigma(\mathbf{j}) + \mathbf{m} 2^{\mathbf{l}}|_{\infty})^{-2r} \\ & \lesssim \sum_{\mathbf{m} \in \mathbb{Z}^n} \prod_{d=1}^n \left(2^{l_d} (1 + |m_d|)\right)^{-2t} \left(2^{|\mathbf{l}|_{\infty}} (1 + |\mathbf{m}|_{\infty})\right)^{-2r} \\ & \lesssim 2^{-2t\|\mathbf{l}\|_1} 2^{-2r\|\mathbf{l}\|_{\infty}} \sum_{\mathbf{m} \in \mathbb{Z}^n} \prod_{d=1}^n (1 + |m_d|)^{-2t} (1 + |\mathbf{m}|_{\infty})^{-2r}. \end{aligned} \quad (22)$$

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 (22), we obtain

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

According to the estimates (16) and (17) with the definitions (18) and (19), this yields

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

and, with relation (20), we obtain the upper estimate

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

which gives the desired result. \square

Note again that our previous article [9] contained a weaker version of Lemma 1, which just involved the logarithmic term L^{n-1} instead of $L^{(n-1)/2}$ and which was restricted to the case $r = 0$.

4 Discussion and Concluding Remarks

Now, we want to compare the error bounds for standard information and linear information. To this end, for $\mathbf{l} \in \mathbb{N}_0^n$, let us first define the approximation operator $Q_{\mathbf{l}}$ with respect to the \mathcal{L}_2 -norm by $Q_{\mathbf{l}} := Q_{l_1} \otimes \dots \otimes Q_{l_n} : \mathcal{L}_2(\mathbb{T}^n) \rightarrow V_{\mathbf{l}}$, where

$$Q_{\mathbf{l}} : \mathcal{L}_2(\mathbb{T}) \rightarrow V_{\mathbf{l}} : f \mapsto \sum_{0 \leq j \leq 2^{\mathbf{l}} - 1} \hat{f}_{\sigma(\mathbf{j})} \phi_{\mathbf{j}}.$$

Then, for any admissible index set \mathcal{J} , we define the general sparse grid approximation operator $Q_{\mathcal{J}} : \mathcal{L}_2(\mathbb{T}^n) \rightarrow V_{\mathcal{J}}$ by

$$Q_{\mathcal{J}} f := \sum_{\mathbf{l} \in \mathcal{J}} \sum_{\mathbf{j} \in \mathcal{J}_{\mathbf{l}}} \hat{f}_{\sigma(\mathbf{j})} \omega_{\sigma(\mathbf{j})}. \quad (23)$$

Now, we can get from [9] the following upper estimate associated to linear information for the special case of the error measured in the \mathcal{H}^s -norm: For $L \in \mathbb{N}_0$, $T < 1$, $s < t + r$, $t \geq 0$ and $f \in \mathcal{H}_{\text{mix}}^{t,r}(\mathbb{T}^n)$ it holds

$$\inf_{\tilde{f} \in V_{\mathcal{J}_L^T}} \|f - \tilde{f}\|_{\mathcal{H}^s} \leq \|f - Q_{\mathcal{J}_L^T} f\|_{\mathcal{H}^s} \lesssim \begin{cases} 2^{L((s-r)-t + (Tt-(s-r))\frac{n-1}{n-T})} \|f\|_{\mathcal{H}_{\text{mix}}^{t,r}} & \text{for } T \geq \frac{s-r}{t}, \\ 2^{L((s-r)-t)} \|f\|_{\mathcal{H}_{\text{mix}}^{t,r}} & \text{for } T \leq \frac{s-r}{t}, \end{cases} \quad (24)$$

with the approximation operator $Q_{\mathcal{J}_L^T}$ as defined via (23) with the index set \mathcal{J}_L^T from (2).

We see that, for general sparse grids, there is a difference in the error behavior between the best approximation (24) and the approximation by interpolation (21)

in the situation $T \geq \frac{s-r}{t}$. Indeed, in the \mathcal{H}^s -norm error estimate for the interpolant resulting from Lemma 1 with $t + \frac{r}{n} > \frac{1}{2}$, $t > 0$, $s - r \geq 0$ and $T \geq \frac{s-r}{t}$, there is the logarithmic factor $L^{(n-1)/2}$ present. In contrast, for the best linear approximation error in the \mathcal{H}^s -norm, there is no such logarithmic term involved, compare relation (24) with $s - r \geq 0$, $t > 0$ and $T \geq \frac{s-r}{t}$. For the special example of regular sparse grids, i.e. $T = 0$, see also Table 1.

Table 1: Convergence behavior in case of regular sparse grid spaces, i.e. $T = 0$, in situation $T \geq \frac{s-r}{t}$ for best linear approximation and for interpolation, i.e. $Q_{\mathcal{J}_L^0}$ and $I_{\mathcal{J}_L^0}$, respectively. Note that, for reasons of simplicity, we restricted ourselves to just the case $s = r$ here.

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

Next, we cast the estimates on the degrees of freedom and on the associated error of approximation by interpolation into a form which measures the error with respect to the number M of the involved degrees of freedom. To this end, let us first recall from [9] the following: Let $L \in \mathbb{N}_0$ and $T < 1$. Then, the number of the degrees of freedom of the general sparse grid spaces $V_{\mathcal{J}_L^T}$ with respect to the discretization parameter L is

$$M := |V_{\mathcal{J}_L^T}| \lesssim \sum_{\mathbf{l} \in \mathcal{J}_L^T} 2^{|\mathbf{l}|_1} \lesssim \begin{cases} 2^L & \text{for } 0 < T < 1, \\ 2^L L^{n-1} & \text{for } T = 0, \\ 2^{L \frac{T-1}{T/n-1}} & \text{for } T < 0, \\ 2^{Ln} & \text{for } T = -\infty. \end{cases} \quad (25)$$

Now, we restrict ourselves to the special case of regular sparse grids, i.e. $T = 0$, and $s - r = 0$. Here, a simple consequence of Lemma 1 and relation (25) is that for $L \in \mathbb{N}_0$, $t > \frac{1}{2}$ and $f \in \mathcal{H}_{\text{mix}}^{t,s}$ with a pointwise convergent Fourier series, there holds the relation

$$\|f - I_{\mathcal{J}_L^0} f\|_{\mathcal{H}^s} \lesssim M^{-t} L^{(t+1/2)(n-1)} \|f\|_{\mathcal{H}_{\text{mix}}^{t,s}} \lesssim M^{-t} \log(M)^{(t+1/2)(n-1)} \|f\|_{\mathcal{H}_{\text{mix}}^{t,s}}, \quad (26)$$

whereas, for linear information, we only have

$$\|f - Q_{\mathcal{J}_L^0} f\|_{\mathcal{H}^s} \lesssim M^{-t} \|f\|_{\mathcal{H}_{\text{mix}}^{t,s}} \lesssim M^{-t} \log(M)^{t(n-1)} \|f\|_{\mathcal{H}_{\text{mix}}^{t,s}}.$$

The question, if this gap between the sampling (standard information) and the approximation numbers (linear information) exists or if it can be closed, is for general sampling point distributions an open problem, see [15, 25, 17] for the case $s = 0$ and compare also the recent improvement in [21]. However, in the special case of sampling on Smolyak grids (i.e. for regular sparse grids with $T = 0$) and for the pairing $(\mathcal{L}_2, \mathcal{H}_{\text{mix}}^{t,0})$, i.e. with $s = 0$ in (26), it follows from the lower bound given in [4, 5] that estimate (26) is indeed asymptotically sharp. Hence the corresponding sampling number and approximation number are *not equal* in this case, compare also Table 1 and see [5, 17] for a further more detailed discussion.

In contrast, for general sparse grid spaces in the situation $T < \frac{s-r}{t}$, the error of the best linear approximation and the error by interpolation behave asymptotically equal and hence there is *no gap* in these situations, c.f. Table 2. Note that, compared to the case $T = 0$ given in Table 2 with $T < \frac{s-r}{t}$, Table 1 covers the case $T = 0$ in the situation $T \geq \frac{s-r}{t}$.

Table 2: Convergence behavior in case of general sparse grid spaces with $T < 1$, $s < t + r$ and $T < \frac{s-r}{t}$ for best linear approximation and for interpolation, i.e. $Q_{\mathcal{J}_L^T}$ and $I_{\mathcal{J}_L^T}$, respectively. For a shorter notation we set $\alpha := t + (r - s)$.

	$f \in$	\mathcal{H}^s -error	dof M	convergence
$Q_{\mathcal{J}_L^T}, 0 < T < \frac{s-r}{t}$	$\mathcal{H}_{\text{mix}}^{t,r}, t > 0$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^L)$	$\mathcal{O}(M^{-\alpha})$
$I_{\mathcal{J}_L^T}, 0 < T < \frac{s-r}{t}$	$\mathcal{H}_{\text{mix}}^{t,r}, t + \frac{r}{n} > \frac{1}{2}$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^L)$	$\mathcal{O}(M^{-\alpha})$
$Q_{\mathcal{J}_L^T}, T = 0$	$\mathcal{H}_{\text{mix}}^{t,r}, t > 0$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^L L^{n-1})$	$\mathcal{O}(M^{-\alpha} \log(M)^{\alpha(n-1)})$
$I_{\mathcal{J}_L^T}, T = 0$	$\mathcal{H}_{\text{mix}}^{t,r}, t + \frac{r}{n} > \frac{1}{2}$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^L L^{n-1})$	$\mathcal{O}(M^{-\alpha} \log(M)^{\alpha(n-1)})$
$Q_{\mathcal{J}_L^T}, -\infty < T < 0$	$\mathcal{H}_{\text{mix}}^{t,r}, t > 0$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^{L \frac{T-1}{T/n-1}})$	$\mathcal{O}(M^{-\alpha \frac{T/n-1}{T-1}})$
$I_{\mathcal{J}_L^T}, -\infty < T < 0$	$\mathcal{H}_{\text{mix}}^{t,r}, t + \frac{r}{n} > \frac{1}{2}$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^{L \frac{T-1}{T/n-1}})$	$\mathcal{O}(M^{-\alpha \frac{T/n-1}{T-1}})$
$Q_{\mathcal{J}_L^{-\infty}}$	$\mathcal{H}_{\text{mix}}^{t,r}, t \geq 0$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^{Ln})$	$\mathcal{O}(M^{-\frac{\alpha}{n}})$
$I_{\mathcal{J}_L^{-\infty}}$	$\mathcal{H}_{\text{mix}}^{t,r}, t + \frac{r}{n} > \frac{1}{2}$	$\mathcal{O}(2^{-\alpha L})$	$\mathcal{O}(2^{Ln})$	$\mathcal{O}(M^{-\frac{\alpha}{n}})$

Note furthermore that, in the case $T \geq \frac{s-r}{t} > 0$, we still see the additional factor $L^{\frac{n-1}{2}}$, i.e. a gap, for the error estimate and we do not know a lower bound like for $T = 0$ and $s = 0$ as in [4, 5], i.e. we do not know if our estimate is sharp in this situation. But this is not the interesting case anyway, since we have a dimension-dependent error rate and dimension-dependent convergence complexity.

Altogether our theory shows that, in many cases, general sparse grids are a suitable approach to avoid the curse of dimensionality at least to a certain level of extent. Especially, in case of \mathcal{H}^s -error measurement of an interpolate $I_{\mathcal{J}_L^T} f$ for $f \in \mathcal{H}_{\text{mix}}^{t,r}$ with $t + \frac{r}{n} > \frac{1}{2}$, $s < t + r$ and $0 < T < \frac{s-r}{t}$, the rate of the interpolation error with

respect to the involved degrees of freedom is independent of the dimension and hence the curse of dimensionality is completely avoided, see also the first row in Table 2, albeit it may still be present in the big O constants. Moreover, our numerical results given in [9] are consistent with the derived theory.

Let us furthermore note that we only discussed the sparse grid interpolation operator $I_{\mathcal{J}_L^T}$ for the periodic case. However, all the related estimates can be transferred in a straightforward way to other spectral transforms like e.g. discrete cosine, sine, Chebyshev, Legendre, generalized Hermite, Jacobi transform and Laguerre transforms, which allow to treat non-periodic cases as well. Here, just the cost for the underlying one-dimensional transform may differ, e.g. in the worst case $O(2^{2l})$ for a naive polynomial transform might be involved in a general situation compared to $O(l2^l)$ for the fast Fourier transform in the simple periodic setting. Indeed, our implemented software library HCFFFT (www.hcfft.org) allows in particular to deal with discrete cosine, sine, Chebyshev, Legendre, generalized Hermite, Jacobi transform and Laguerre transforms, and also their mixtures, compare [29]. Moreover, discretization spaces associated with arbitrary admissible index sets and also spaces with finite-order weights [8, 10] and dimension-adaptive methods [1, 7, 16] can be treated analogously.

We believe that our general approach via the functions $\tilde{\beta}$ and β and the pairing $(\mathcal{H}_{\tilde{\beta}}, \mathcal{H}_{\beta})$ is helpful for much more general situations than the one dealt with in this article, as long as the corresponding right hand side of (17) involving the g -based bound of (16) can be handled properly.

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

References

1. 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.
2. H. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:1–123, 2004.
3. D. Düng. B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness. *Journal of Complexity*, 27(6):541–567, 2011.
4. D. Düng. B-spline quasi-interpolation sampling representation and sampling recovery in Sobolev spaces of mixed smoothness. *Acta Mathematica Vietnamica*, 43(1):83–110, 2018.
5. D. Düng, V. Temlyakov, and T. Ullrich. *Hyperbolic Cross Approximation*. Advanced Courses in Mathematics - CRM Barcelona. Birkhäuser/Springer, 2018.
6. D. Düng and T. Ullrich. N-widths and ε -dimensions for high-dimensional approximations. *Foundations of Computational Mathematics*, 13:965–1003, 2013.
7. T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71(1):65–87, 2003.
8. 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.

9. 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.
10. M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *Journal of Complexity*, 26:455–489, 2010.
11. M. Griebel and S. Knapek. Optimized tensor-product approximation spaces. *Constructive Approximation*, 16(4):525–540, 2000.
12. M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Mathematics of Computation*, 78:2223–2257, 2009.
13. K. Hallatschek. Fourier-transform on sparse grids with hierarchical bases. *Numerische Mathematik*, 63(1):83–97, 1992.
14. 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.
15. A. Hinrichs, E. Novak, and J. Vybírál. Linear information versus function evaluations for L_2 -approximation. *Journal of Approximation Theory*, 153(1):97–107, 2008.
16. J. Jakeman and S. Roberts. Local and dimension adaptive stochastic collocation for uncertainty quantification. In J. Garcke and M. Griebel, editors, *Sparse Grids and Applications*, pages 181–203. Springer, 2013.
17. L. Kaemmerer, T. Ullrich, and T. Volkmer. Worst case recovery guarantees for least squares approximation using random samples. *arXiv preprint arXiv:1911.10111*, 2019.
18. S. Knapek. *Approximation und Kompression mit Tensorprodukt-Multiskalenräumen*. Dissertation, University of Bonn, 2000.
19. S. Knapek. Hyperbolic cross approximation of integral operators with smooth kernel. 2000. Technical Report 665, SFB 256, University of Bonn.
20. H. Kreuzler and H. Yserentant. The mixed regularity of electronic wave functions in fractional order and weighted Sobolev spaces. *Numerische Mathematik*, 121(4):781–802, 2012.
21. D. Krieg and M. Ullrich. Function values are enough for L_2 -approximation. *arXiv preprint arXiv:1905.02516*, 2019.
22. T. Kühn, W. Sickel, and T. Ullrich. Approximation of mixed order Sobolev functions on the d -torus: asymptotics, preasymptotics, and d -dependence. *Constructive Approximation*, 42(3):353–398, 2015.
23. F. Kupka. *Sparse Grid Spectral Methods for the Numerical Solution of Partial Differential Equations with Periodic Boundary Conditions*. PhD thesis, University of Wien, 1997.
24. 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.
25. 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.
26. W. Sickel and T. Ullrich. Spline interpolation on sparse grids. *Applicable Analysis*, 90(3-4):337–383, 2011.
27. V. Temlyakov. *Approximation of Periodic Functions*. Nova Science, New York, 1993.
28. H. Triebel. *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, volume 11 of *EMS Tracts in Mathematics*. European Mathematical Society, 2010.
29. V. Velikov. *Fast Sparse Pseudo-spectral Methods for High-dimensional Problems*. Master thesis, Institute for Numerical Simulation, Universität Bonn, 2016.