

# A Note on the Construction of $L$ -Fold Sparse Tensor Product Spaces

Michael Griebel · Helmut Harbrecht

Received: 29 April 2012 / Revised: 19 August 2012 / Accepted: 3 September 2012 /  
Published online: 24 January 2013  
© Springer Science+Business Media New York 2013

**Abstract** In the present paper, we consider the construction of general sparse tensor product spaces in arbitrary space dimensions when the single subdomains are of different dimensionality and the associated ansatz spaces possess different approximation properties. Our theory extends the results from Griebel and Harbrecht (Math. Comput., 2013) for the construction of two-fold sparse tensor product space to arbitrary  $L$ -fold sparse tensor product spaces.

**Keywords** High-dimensional problems · Sparse grids · Sparse tensor product spaces · Tensor product domains of different dimensions · Sparse tensor product of ansatz spaces with different approximation power · Optimal construction of sparse grids · Rate of approximation

**Mathematics Subject Classification (2000)** 41A17 · 41A25 · 41A30 · 41A65

## 1 Introduction

Many problems in science and engineering lead to problems which are defined on the tensor product of  $L$  domains  $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_L$ , where  $\Omega_i \in \mathbb{R}^{n_i}$ . Already for the simple situation of  $L = 2$ , there exists a large number of problems. This includes, for instance, radiosity models and radiative transfer [25], where  $\Omega_1$  denotes the spatial

---

Communicated by Wolfgang Dahmen.

M. Griebel  
Institut für Numerische Simulation, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany  
e-mail: [griebel@ins.uni-bonn.de](mailto:griebel@ins.uni-bonn.de)

H. Harbrecht (✉)  
Mathematisches Institut, Universität Basel, Rheinsprung 21, 4051 Basel, Switzerland  
e-mail: [helmut.harbrecht@unibas.ch](mailto:helmut.harbrecht@unibas.ch)

three-dimensional domain of a geometric object under consideration and  $\Omega_2$  is the sphere  $\mathbb{S}^2$ . Then there are the parabolic problems, where  $\Omega_1$  is the time interval and  $\Omega_2$  is the spatial domain [5, 13, 15, 22], or various phase space problems, e.g., the Boltzmann equation, kinetic equations, or the Langevin equation [2], where both  $\Omega_1$  and  $\Omega_2$  are three-dimensional cubes or the full three-dimensional real space.

Even more applications can be found in case of a general  $L > 2$ . For example, non-Newtonian flow can be modeled by a coupled system which consists of the Navier Stokes equation for the flow in a three-dimensional geometry described by  $\Omega_1$  and of the Fokker–Planck equation in the  $3(L - 1)$ -dimensional configuration space on  $\Omega_2 \times \Omega_3 \times \dots \times \Omega_L$ , where each domain  $\Omega_i$  ( $i \geq 2$ ) is a sphere. Here  $L$  denotes the number of atoms in a chain-like molecule which constitutes the non-Newtonian behavior of the flow; for details, see [3, 20, 21]. Also the  $L$ -th moment of linear elliptic boundary value problems with stochastic source terms are known to satisfy a deterministic partial differential equation with the  $L$ -fold tensor product of the elliptic operator on the  $L$ -fold product of the physical domain [23, 24]. This approach extends to stochastic diffusion problems and to more general partial differential equations with stochastic coefficient functions or with stochastic domains [16–18]. Another example is multiscale homogenization. After unfolding [1, 6, 7], it gives rise to the product of the macroscopic physical domain with the  $L - 1$  periodic microscopic domains which correspond to the  $L - 1$  different microscales [19]. Finally, we find the product of  $L$  domains in quantum mechanics for, e.g., the electronic Schrödinger equation where each electron has its associated three-dimensional domain, see, e.g., [10, 26, 27].

A naive, conventional discretization would use tensor products of all basis functions from suitable finite dimensional ansatz spaces  $V_J^{(i)}$ ,  $i = 1, 2, \dots, L$  which are defined on each domain separately. This leads to the *full tensor product space*  $V_J^{(1)} \otimes V_J^{(2)} \otimes \dots \otimes V_J^{(L)}$ . However, in general, the full tensor product space contains way too many degrees of freedom such that desirable realistic simulations are beyond current computing capacities. This relates to the well-known *curse of dimension* which states that the number of degrees of freedom for an approximation necessary to obtain a prescribed accuracy grows exponentially with the dimension.

To overcome the curse of dimension, we will focus in this paper on the construction of *sparse tensor product spaces*, also known as sparse grids [4, 28]. The starting point are multilevel decompositions of the ansatz spaces

$$V_J^{(i)} = W_0^{(i)} \oplus W_1^{(i)} \oplus \dots \oplus W_J^{(i)}, \quad i = 1, 2, \dots, L,$$

which can be constructed via hierarchical bases, interpolants, or wavelet-like bases. From this, the *regular sparse tensor product space* is defined according to

$$\widehat{V}_J^{\text{reg}} = \bigoplus_{j_1 + j_2 + \dots + j_L \leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)} \otimes \dots \otimes W_{j_L}^{(L)}, \quad (1.1)$$

see, e.g., [4, 12, 14, 28]. Its approximation power is nearly as good as that of the corresponding full tensor product space if the functions to be approximated provide additional smoothness in terms of bounded mixed derivatives. Note here that the regular sparse tensor product space (1.1) is optimal if the approximation error is measured in the  $L^2$ -norm and the properties of the ansatz spaces  $V_J^{(i)}$  and  $W_{j_i}^{(i)}$ , respectively,

do not depend in  $i$ . This means that all the domains  $\Omega_i$  have the *same* dimension and all the spaces  $V_J^{(i)}$  and  $W_{j_i}^{(i)}$ , respectively, are equipped with the *same* type of ansatz functions.

To also cover the more general cases, we will introduce in the following the special sparse tensor product space

$$\widehat{V}_J^\alpha := \bigoplus_{\alpha^T \mathbf{j} \leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)} \otimes \dots \otimes W_{j_L}^{(L)}$$

for an *arbitrary* vector  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_L) > \mathbf{0}$  and  $\mathbf{j} = (j_1, j_2, \dots, j_L) \in \mathbb{N}_0^L$ , and we will discuss various choices of  $\alpha$  for respective general situations of local dimensions and regularities of the  $L$  domains.

For the case  $L = 2$ , we already systematically studied in [11] what the most efficient construction of sparse tensor product spaces is if the spatial dimension of the underlying domains or the polynomial exactness (and thus the approximation power) of the ansatz spaces differ. In this paper, we will now extend these results to  $L$ -fold sparse tensor product spaces. It will turn out that, in case of smooth functions, there is a whole range of sparse tensor product spaces which possess the same optimal convergence rate. However, in case of functions with limited regular or mixed Sobolev smoothness, it will turn out that the sparse tensor product space, which equilibrates the number of degrees of freedom (see Sect. 3), is superior to all the other sparse tensor product spaces under consideration. We recently learned that the situation  $L \geq 2$  has been considered in [9], too. There, different approximation powers in the coordinate directions were also taken into account but the underlying domains were only of equal dimension  $n_1 = n_2 = \dots = n_L = 1$ .

The remainder of this paper is organized as follows. In Sect. 2, we specify the requirements of the multiscale hierarchies on each subdomain. Then, in Sect. 3, we construct general sparse tensor product spaces. In Sect. 4, we study their properties. Section 5 is dedicated to the comparison of the cost complexities for the approximation of functions of mixed Sobolev smoothness, where also different degrees of smoothness are possible in the different coordinate directions. Finally, in Sect. 6, we give some concluding remarks.

Throughout this paper, the notion “essential” in the context of complexity estimates means “up to logarithmic terms.” Moreover, to avoid the repeated use of generic but unspecified constants, we signify by  $C \lesssim D$  that  $C$  is bounded by a multiple of  $D$  independently of parameters which  $C$  and  $D$  may depend on. Obviously,  $C \gtrsim D$  is defined as  $D \lesssim C$ , and  $C \sim D$  as  $C \lesssim D$  and  $C \gtrsim D$ .

## 2 Approximation on a Subdomain

Let  $\Omega \subset \mathbb{R}^d$  be a sufficiently smooth, bounded domain. We consider a nested sequence of finite dimensional spaces

$$V_0 \subset V_1 \subset \dots \subset V_j \subset \dots \subset L^2(\Omega), \tag{2.1}$$

each of which consists of piecewise polynomial ansatz functions  $V_j = \text{span}\{\varphi_{j,k} : k \in \Delta_j\}$  ( $\Delta_j$  denotes a suitable index set) such that  $\dim V_j \sim 2^{jn}$  and

$$L^2(\Omega) = \overline{\bigcup_{j \in \mathbb{N}_0} V_j}. \tag{2.2}$$

We will use the spaces  $V_j$  for the approximation of functions. To this end, we assume that the approximation property

$$\inf_{v_j \in V_j} \|u - v_j\|_{H^q(\Omega)} \lesssim h_j^{s-q} \|u\|_{H^s(\Omega)}, \quad u \in H^s(\Omega), \tag{2.3}$$

holds for  $q < \gamma$ ,  $q \leq s \leq r$  uniformly in  $j$ . Here we set  $h_j := 2^{-j}$ , i.e.,  $h_j$  corresponds to the width of the mesh associated with the subspace  $V_j$  on  $\Omega$ . The parameter  $\gamma > 0$  refers to the *regularity* of the functions which are contained in  $V_j$ , i.e.,

$$\gamma := \sup\{s \in \mathbb{R} : V_j \subset H^s(\Omega)\}.$$

The integer  $r > 0$  refers to the *polynomial exactness*, that is, the maximal order of polynomials which are locally contained in the space  $V_j$ .

We now introduce a wavelet basis associated with the multiscale analysis (2.1) and (2.2) as follows. The wavelets  $\Psi_j := \{\psi_{j,k} : k \in \nabla_j\}$ , where  $\nabla_j := \Delta_j \setminus \Delta_{j-1}$ , are the bases of the complementary spaces  $W_j$  of  $V_{j-1}$  in  $V_j$ , i.e.,

$$V_j = V_{j-1} \oplus W_j, \quad V_{j-1} \cap W_j = \{0\}, \quad W_j = \text{span}\{\Psi_j\}.$$

Recursively, we obtain

$$V_J = \bigoplus_{j=0}^J W_j, \quad W_0 := V_0,$$

and thus, with

$$\Psi_J := \bigcup_{j=0}^J \Psi_j, \quad \Psi_0 := \text{span}\{\varphi_{0,k} : k \in \Delta_0\},$$

we get a wavelet basis in  $V_J$ . A final requirement is that the infinite collection  $\Psi := \bigcup_{j \geq 0} \Psi_j$  forms a Riesz basis of  $L^2(\Omega)$ . Then there exists also a dual (or biorthogonal) wavelet basis  $\tilde{\Psi} = \bigcup_{j \geq 0} \tilde{\Psi}_j = \{\tilde{\psi}_{j,k} : k \in \nabla_j\}$  which defines a dual multiscale analysis of regularity  $\tilde{\gamma} > 0$ , see, e.g., [8] for further details. In particular, each function  $f \in L^2(\Omega)$  admits the unique representation

$$f = \sum_{j=0}^{\infty} \sum_{k \in \nabla_j} (f, \tilde{\psi}_{j,k})_{L^2(\Omega)} \psi_{j,k}. \tag{2.4}$$

With the definition of the projections

$$Q_j : L^2(\Omega) \rightarrow W_j, \quad Q_j f = \sum_{k \in \nabla_j} (f, \tilde{\psi}_{j,k})_{L^2(\Omega)} \psi_{j,k},$$

the atomic decomposition (2.4) gives rise to the multilevel decomposition

$$f = \sum_{j=0}^{\infty} Q_j f.$$

Since for all  $-\gamma < q < \tilde{\gamma}$  also a (properly scaled version of the) wavelet basis  $\Psi$  is a Riesz basis of  $H^q(\Omega)$ , we especially have

$$\|f\|_{H^q(\Omega)}^2 \sim \sum_{j=0}^{\infty} \|Q_j f\|_{H^q(\Omega)}^2, \quad -\gamma < q < \tilde{\gamma}, \tag{2.5}$$

see [8]. Finally, for any  $f \in H^s(\Omega)$ , the approximation property (2.3) induces the estimate

$$\|Q_j f\|_{H^q(\Omega)} \lesssim 2^{-j(s-q)} \|f\|_{H^s(\Omega)}, \quad q \leq s \leq r. \tag{2.6}$$

### 3 Sparse Tensor Product Spaces

Consider now  $L$  domains  $\Omega_i \subset \mathbb{R}^{n_i}$  with  $n_i \in \mathbb{N}$  for all  $i = 1, 2, \dots, L$ . We aim for the approximation of functions in anisotropic Sobolev spaces

$$\mathbf{H}^s(\boldsymbol{\Omega}) := H^{s_1}(\Omega_1) \otimes H^{s_2}(\Omega_2) \otimes \dots \otimes H^{s_L}(\Omega_L),$$

where  $\boldsymbol{\Omega} := \Omega_1 \times \Omega_2 \times \dots \times \Omega_L$ . To this end, we assume to individually have for each subdomain  $\Omega_i, i = 1, 2, \dots, L$ , the multiscale analysis

$$V_0^{(i)} \subset V_1^{(i)} \subset V_2^{(i)} \subset \dots \subset L^2(\Omega_i), \quad V_0^{(i)} = \text{span}\{\Phi_j^{(i)}\}, \quad i = 1, 2, \dots, L,$$

with associated complementary spaces

$$V_j^{(i)} = V_{j-1}^{(i)} \oplus W_j^{(i)}, \quad V_{j-1}^{(i)} \cap W_j^{(i)} = \{0\}, \quad W_j^{(i)} = \text{span}\{\Psi_j^{(i)}\}.$$

Furthermore, for all  $i = 1, 2, \dots, L$ , let us denote the polynomial exactnesses of the spaces  $V_j^{(i)}$  by  $r_i$ , and the regularity of the primal and dual wavelet bases by  $\gamma_i$  and  $\tilde{\gamma}_i$ , respectively.

In this paper, we study the approximation of functions in the special sparse tensor product space<sup>1</sup>

$$\widehat{\mathbf{V}}_{\mathbf{j}}^{\boldsymbol{\alpha}} := \bigoplus_{\boldsymbol{\alpha}^T \mathbf{j} \leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)} \otimes \dots \otimes W_{j_L}^{(L)} \tag{3.1}$$

for an arbitrary vector  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_L) > \mathbf{0}$  and  $\mathbf{j} = (j_1, j_2, \dots, j_L) \in \mathbb{N}_0^L$ . Here and subsequently, inequalities for vectors are to be understood componentwise, i.e.,

<sup>1</sup>Here and subsequently, the summation limits are in general no natural numbers and must of course be rounded properly. We leave this to the reader to avoid cumbersome floor/ceil-notations.

$\alpha > \mathbf{0}$  means  $\alpha_i > 0$  for all  $i = 1, 2, \dots, L$ . Moreover, we set  $\mathbf{r} = (r_1, r_2, \dots, r_L)$ ,  $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_L)$ , and  $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_L)$ .

Now we can formulate the following problem: *How must  $\alpha > \mathbf{0}$  be chosen such that, for given parameters  $-\tilde{\boldsymbol{\gamma}} < \mathbf{q} < \boldsymbol{\gamma}$  and  $\mathbf{q} < \mathbf{s} \leq \mathbf{r}$ , a function  $f \in \mathbf{H}^{\mathbf{s}}(\boldsymbol{\Omega})$  can be best approximated in  $\widehat{\mathbf{V}}_J^\alpha$  in the  $\mathbf{H}^{\mathbf{q}}(\boldsymbol{\Omega})$ -norm?*

Natural choices of the parameter  $\alpha > \mathbf{0}$  are:

- (i) To equilibrate the accuracy in the involved univariate spaces  $V_{J/\alpha_i}^{(i)}$ ,  $i = 1, 2, \dots, L$ , we obtain the condition

$$2^{-J(r_1-q_1)/\alpha_1} = 2^{-J(r_2-q_2)/\alpha_2} = \dots = 2^{-J(r_L-q_L)/\alpha_L}.$$

This means that we have to choose  $\alpha_i = r_i - q_i$  for all  $i = 1, 2, \dots, L$ .

- (ii) To equilibrate the number of degrees of freedom in the involved univariate spaces  $V_{J/\alpha_i}^{(i)}$ ,  $i = 1, 2, \dots, L$ , we obtain the condition

$$2^{Jn_1/\alpha_1} = 2^{Jn_2/\alpha_2} = \dots = 2^{Jn_L/\alpha_L}.$$

This condition is satisfied if  $\alpha_i = n_i$  for all  $i = 1, 2, \dots, L$ .

- (iii) Following the idea of an *equilibrated cost-benefit rate* (see [4]), we get the condition

$$2^{j_1(n_1+r_1-q_1)} \cdot 2^{j_2(n_2+r_2-q_2)} \dots 2^{j_L(n_L+r_L-q_L)} = 2^{J \cdot \text{const.}} \quad \text{for all } \boldsymbol{\alpha}^T \mathbf{j} = J.$$

For  $\text{const.} = 1$ , we find  $\alpha_i = n_i + r_i - q_i$  for all  $i = 1, 2, \dots, L$ .

In the next section, we will compute for these and other choices the cost and the error of the approximation in the related sparse tensor product spaces.

### 4 Properties of the Sparse Tensor Product Spaces

To compute the convergence rate of functions in the sparse tensor product spaces  $\widehat{\mathbf{V}}_J^\alpha$  with arbitrary  $\alpha > \mathbf{0}$ , we first count the degrees of freedom which are contained in these spaces.

**Theorem 4.1** *For any  $\alpha > \mathbf{0}$ , the dimension of the sparse tensor product space*

$$\widehat{\mathbf{V}}_J^\alpha = \bigoplus_{\boldsymbol{\alpha}^T \mathbf{j} \leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)} \otimes \dots \otimes W_{j_L}^{(L)}$$

*is proportional to  $2^{J \max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}} J^{R-1}$ , where  $R$  counts how often the maximum is attained.*

*Proof* We shall use induction to prove this theorem. Consider first the case  $L = 1$ . Due to  $\dim W_{j_1}^{(1)} \sim 2^{j_1 n_1}$ , it holds that

$$\dim \widehat{\mathbf{V}}_J^\alpha \sim \sum_{j_1=0}^{J/\alpha_1} 2^{j_1 n_1} \lesssim 2^{Jn_1/\alpha_1}.$$

Since  $R = 1$ , the desired result follows immediately.

We now show the induction step  $L \mapsto L + 1$ . To this end, we denote the  $(L + 1)$ -fold sparse tensor product space by

$$\widehat{V}_J^{(\alpha, \alpha_{L+1})} = \bigoplus_{\alpha^T \mathbf{j} + \alpha_{L+1} j_{L+1} \leq J} W_{j_1}^{(1)} \otimes \dots \otimes W_{j_L}^{(L)} \otimes W_{j_{L+1}}^{(L+1)}, \tag{4.1}$$

where, without loss of generality<sup>2</sup>, we will assume

$$\frac{n_1}{\alpha_1} = \dots = \frac{n_R}{\alpha_R} > \frac{n_{R+1}}{\alpha_{R+1}} \geq \frac{n_{R+2}}{\alpha_{R+2}} \geq \dots \geq \frac{n_L}{\alpha_L} \quad \text{and} \quad \frac{n_L}{\alpha_L} \geq \frac{n_{L+1}}{\alpha_{L+1}}. \tag{4.2}$$

Obviously, there holds the identity

$$\begin{aligned} \widehat{V}_J^{(\alpha, \alpha_{L+1})} &= \bigoplus_{j_L=0}^{J/\alpha_{L+1}} \left[ \bigoplus_{\alpha^T \mathbf{j} \leq J - \alpha_{L+1} j_{L+1}} W_{j_1}^{(1)} \otimes \dots \otimes W_{j_L}^{(L)} \right] \otimes W_{j_{L+1}}^{(L+1)} \\ &= \bigoplus_{j_{L+1}=0}^{J/\alpha_{L+1}} \widehat{V}_{J - \alpha_{L+1} j_{L+1}}^\alpha \otimes W_{j_{L+1}}^{(L+1)}. \end{aligned} \tag{4.3}$$

According to our induction hypothesis, the sparse tensor product space  $\widehat{V}_{J - \alpha_{L+1} j_{L+1}}^\alpha$  consists of  $\mathcal{O}(2^{(J - \alpha_{L+1} j_{L+1})n_1/\alpha_1} (J - \alpha_{L+1} j_{L+1})^{R-1})$  degrees of freedom. Hence, since  $\dim W_{j_{L+1}}^{(L+1)} \sim 2^{j_{L+1}n_{L+1}}$ , we get the estimate

$$\begin{aligned} \dim \widehat{V}_J^{(\alpha, \alpha_{L+1})} &\lesssim \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} 2^{j_{L+1}n_{L+1}} 2^{(J - \alpha_{L+1} j_{L+1})n_1/\alpha_1} (J - \alpha_{L+1} j_{L+1})^{R-1} \\ &\lesssim 2^{Jn_1/\alpha_1} J^{R-1} \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} 2^{j_{L+1}(n_{L+1} - n_1\alpha_{L+1}/\alpha_1)}. \end{aligned}$$

Due to the ordering (4.2), we have to distinguish two cases, namely  $n_1/\alpha_1 > n_{L+1}/\alpha_{L+1}$  and  $n_1/\alpha_1 = n_{L+1}/\alpha_{L+1}$ . In the case  $n_1/\alpha_1 > n_{L+1}/\alpha_{L+1}$ , which always happens for  $R < L$  and might happen if  $R = L$ , the exponent in the last sum is always negative, which immediately implies

$$\dim \widehat{V}_J^{(\alpha, \alpha_{L+1})} \lesssim 2^{Jn_1/\alpha_1} J^{R-1}. \tag{4.4}$$

In the case  $n_1/\alpha_1 = n_{L+1}/\alpha_{L+1}$ , which can only happen if  $R = L$ , the exponent in the sum is zero, and it follows that

$$\dim \widehat{V}_J^{(\alpha, \alpha_{L+1})} \lesssim 2^{Jn_1/\alpha_1} J^{L-1} \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} 1 \lesssim 2^{Jn_1/\alpha_1} J^L. \tag{4.5}$$

<sup>2</sup>Otherwise, we apply the induction to an appropriate permutation of the spatial dimensions.

The combination of (4.4) and (4.5) yields the desired result in case of  $L + 1$ . This completes the proof.  $\square$

*Remark 4.2* The statement of Theorem 4.1 is sharp, which can be seen as follows. The sparse tensor product space  $\widehat{\mathbf{V}}_J^\alpha$  contains the space  $W_0^{(1)} \otimes \dots \otimes W_0^{(k-1)} \otimes W_{J/\alpha_k}^{(k)} \otimes W_0^{(k+1)} \otimes \dots \otimes W_0^{(L)}$  for all  $1 \leq k \leq L$  whose number of degrees of freedom is proportional to  $2^{Jn_k/\alpha_k}$  (i.e.,  $2^{Jn_k/\alpha_k}$  bounds its number of degrees of freedom from below and above). Thus,  $2^{J \max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}}$  is obviously a lower bound for the number of degrees of freedom of the sparse tensor product space  $\widehat{\mathbf{V}}_J^\alpha$ . If the maximum is attained  $R$  times, for example, if  $n_1/\alpha_1 = n_2/\alpha_2 = \dots = n_R/\alpha_R$ , then there are asymptotically  $J^{R-1}$  further combinations of  $(j_1, j_2, \dots, j_R)$  such that  $\sum_{k=1}^R \alpha_k j_k = J$ . The numbers of degrees of freedom of the associated spaces  $W_{j_1}^{(1)} \otimes \dots \otimes W_{j_R}^{(R)} \otimes W_0^{(R+1)} \otimes \dots \otimes W_0^{(L)}$  are also proportional to  $2^{J \max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}}$ . This yields the logarithmic factor  $J^{R-1}$  in the estimate as well as its sharpness.

Next we consider the approximation power in the sparse tensor product spaces  $\widehat{\mathbf{V}}_J^\alpha$ . Obviously, the highest possible rate of convergence is attained in the space  $\mathbf{H}^{\mathbf{r}}(\Omega)$ , where  $\mathbf{r} = (r_1, r_2, \dots, r_L)$  is the vector of polynomial exactness of the underlying full tensor product spaces. Therefore, in the following theorem, we restrict ourselves without loss of generality to  $\mathbf{s} \leq \mathbf{r}$ .

**Theorem 4.3** *Let  $-\tilde{\gamma} < \mathbf{q} < \boldsymbol{\gamma}$  and  $\mathbf{q} < \mathbf{s} \leq \mathbf{r}$  and  $f \in \mathbf{H}^{\mathbf{s}}(\Omega)$ . Then, the sparse grid projector*

$$\widehat{\mathbf{Q}}_J^\alpha : \mathbf{H}^{\mathbf{q}}(\Omega) \rightarrow \widehat{\mathbf{V}}_J^\alpha, \quad \widehat{\mathbf{Q}}_J^\alpha f = \sum_{\alpha^T \mathbf{j} \leq J} (Q_{j_1}^{(1)} \otimes Q_{j_2}^{(2)} \otimes \dots \otimes Q_{j_L}^{(L)}) f \tag{4.6}$$

satisfies

$$\|(I - \widehat{\mathbf{Q}}_J^\alpha) f\|_{\mathbf{H}^{\mathbf{q}}(\Omega)} \lesssim 2^{-J \min\{\frac{s_1-q_1}{\alpha_1}, \frac{s_2-q_2}{\alpha_2}, \dots, \frac{s_L-q_L}{\alpha_L}\}} J^{(P-1)/2} \|f\|_{\mathbf{H}^{\mathbf{s}}(\Omega)}. \tag{4.7}$$

Here,  $P$  counts how often the minimum is attained in the exponent.

*Proof* We show the assertion again by induction over  $L$ . For the case  $L = 1$ , the tensor product domain  $\Omega$  is a single domain  $\Omega_1$  and the projector  $\widehat{\mathbf{Q}}_J^\alpha$  onto the sparse grid is simply  $\sum_{j_1=0}^{J/\alpha_1} Q_{j_1}^{(1)}$ . Hence, the assertion follows immediately from (2.5) and (2.6), since

$$\begin{aligned} \|(I - \widehat{\mathbf{Q}}_J^{\alpha_1}) f\|_{H^{q_1}(\Omega_1)}^2 &\lesssim \sum_{j_1=J/\alpha_1+1}^{\infty} \|Q_{j_1}^{(1)} f\|_{H^{q_1}(\Omega_1)}^2 \\ &\lesssim \sum_{j_1=J/\alpha_1+1}^{\infty} 2^{-2j_1(s_1-q_1)} \|f\|_{H^{s_1}(\Omega_1)}^2 \\ &\lesssim 2^{-2J(s_1-q_1)/\alpha_1} \|f\|_{H^{s_1}(\Omega_1)}^2 \end{aligned}$$

for all  $f \in H^{s_1}(\Omega_1)$ .



We now assume that the assertion is shown for the product domain  $\Omega = \Omega_1 \times \dots \times \Omega_L$  and the sparse grid projector  $\widehat{\mathbf{Q}}_J^\alpha$  given by (4.6). This means the error estimate (4.7) is valid for any function from  $\mathbf{H}^s(\Omega)$ . To prove the induction step  $L \mapsto L + 1$ , we shall assume without loss of generality<sup>3</sup> that

$$\frac{s_1 - q_1}{\alpha_1} = \dots = \frac{s_P - q_P}{\alpha_P} < \frac{s_{P+1} - q_{P+1}}{\alpha_{P+1}} \leq \frac{s_{P+2} - q_{P+2}}{\alpha_{P+2}} \leq \dots \leq \frac{s_L - q_L}{\alpha_L}$$

and

$$\frac{s_L - q_L}{\alpha_L} \leq \frac{s_{L+1} - q_{L+1}}{\alpha_{L+1}}.$$

The induction step then reads as follows. For any given function

$$\begin{aligned} f \in \mathbf{H}^{(s, s_{L+1})}(\Omega \times \Omega_{L+1}) &:= \mathbf{H}^s(\Omega) \otimes H^{s_{L+1}}(\Omega_{L+1}) \\ &= H^{s_1}(\Omega_1) \otimes \dots \otimes H^{s_{L+1}}(\Omega_{L+1}), \end{aligned}$$

we shall show that the sparse grid projection

$$\widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})} = \sum_{\alpha^T \mathbf{j} + \alpha_{L+1} j_{L+1} \leq J} \mathcal{Q}_{j_1}^{(1)} \otimes \dots \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)} : \mathbf{H}^{(q, q_{L+1})}(\Omega \times \Omega_{L+1}) \rightarrow \widehat{\mathbf{V}}_J^{(\alpha, \alpha_{L+1})}$$

satisfies the error estimate

$$\begin{aligned} &\| (I - \widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\Omega \times \Omega_{L+1})} \\ &\lesssim 2^{-J \min\{\frac{s_1 - q_1}{\alpha_1}, \frac{s_2 - q_2}{\alpha_2}, \dots, \frac{s_{L+1} - q_{L+1}}{\alpha_{L+1}}\}} J^{(P^* - 1)/2} \| f \|_{\mathbf{H}^{(s, s_{L+1})}(\Omega \times \Omega_{L+1})}. \end{aligned}$$

Here, the sparse tensor product space  $\widehat{\mathbf{V}}_J^{(\alpha, \alpha_{L+1})}$  is defined in (4.1) and  $P^* := P + 1$  if  $(s_1 - q_1)/\alpha_1 = (s_{L+1} - q_{L+1})/\alpha_{L+1}$  and  $P^* := P$  otherwise.

Analogously to (4.3), we find the identity

$$\begin{aligned} \widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})} &= \sum_{\alpha^T \mathbf{j} + \alpha_{L+1} j_{L+1} \leq J} \mathcal{Q}_{j_1}^{(1)} \otimes \dots \otimes \mathcal{Q}_{j_L}^{(L)} \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)} \\ &= \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} \widehat{\mathbf{Q}}_{J - \alpha_{L+1} j_{L+1}}^\alpha \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)}. \end{aligned}$$

In particular, it holds that

$$I - \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} \widehat{\mathbf{Q}}_{J - \alpha_{L+1} j_{L+1}}^\alpha \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)}$$

<sup>3</sup>Otherwise, we apply the induction to an appropriate permutation of the spatial dimensions.

$$\begin{aligned}
 &= \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} (I - \widehat{\mathbf{Q}}_{J-\alpha_{L+1}j_{L+1}}^\alpha) \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)} + I \otimes \left( I - \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} \mathcal{Q}_{j_{L+1}}^{(L+1)} \right) \\
 &= \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} (I - \widehat{\mathbf{Q}}_{J-\alpha_{L+1}j_{L+1}}^\alpha) \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)} + I \otimes \sum_{j_{L+1}=J/\alpha_{L+1}+1}^\infty \mathcal{Q}_{j_{L+1}}^{(L+1)}.
 \end{aligned}$$

Thus, in view of (2.5) and (2.6), since  $f \in \mathbf{H}^{(s, s_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})$ , we find

$$\begin{aligned}
 &\| (I - \widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\lesssim \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} \| ((I - \widehat{\mathbf{Q}}_{J-\alpha_{L+1}j_{L+1}}^\alpha) \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\quad + \sum_{j_{L+1}=J/\alpha_{L+1}+1}^\infty \| (I \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2.
 \end{aligned}$$

The second norm on the right-hand side can be estimated as

$$\begin{aligned}
 \| (I \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 &\lesssim 2^{-2j_{L+1}(s_{L+1}-q_{L+1})} \| f \|_{\mathbf{H}^{(q, s_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\lesssim 2^{-2j_{L+1}(s_{L+1}-q_{L+1})} \| f \|_{\mathbf{H}^{(s, s_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2.
 \end{aligned}$$

The induction hypotheses (4.7) implies

$$\begin{aligned}
 &\| ((I - \widehat{\mathbf{Q}}_{J-\alpha_{L+1}j_{L+1}}^\alpha) \otimes I) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\lesssim 2^{-2(J-\alpha_{L+1}j_{L+1})(s_1-q_1)/\alpha_1} (J - \alpha_{L+1}j_{L+1})^{P-1} \| f \|_{\mathbf{H}^{(s, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2.
 \end{aligned}$$

By employing in addition (2.6), the first norm on the right-hand side of the above expression can be estimated according to

$$\begin{aligned}
 &\| ((I - \widehat{\mathbf{Q}}_{J-\alpha_{L+1}j_{L+1}}^\alpha) \otimes \mathcal{Q}_{j_{L+1}}^{(L+1)}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\lesssim 2^{-2(J-\alpha_{L+1}j_{L+1})(s_1-q_1)/\alpha_1} (J - \alpha_{L+1}j_{L+1})^{P-1} 2^{-2j_{L+1}(s_{L+1}-q_{L+1})} \\
 &\quad \times \| f \|_{\mathbf{H}^{(s, s_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\lesssim 2^{-2J(s_1-q_1)/\alpha_1} J^{P-1} 2^{-2j_{L+1}((s_{L+1}-q_{L+1})-(s_1-q_1)\alpha_{L+1}/\alpha_1)} \| f \|_{\mathbf{H}^{(s, s_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2.
 \end{aligned}$$

Altogether, we thus get

$$\begin{aligned}
 &\| (I - \widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})}) f \|_{\mathbf{H}^{(q, q_{L+1})}(\mathfrak{Q} \times \Omega_{L+1})}^2 \\
 &\lesssim \left[ 2^{-2J(s_1-q_1)/\alpha_1} J^{P-1} \sum_{j_{L+1}=0}^{J/\alpha_{L+1}} 2^{-2j_{L+1}((s_{L+1}-q_{L+1})-(s_1-q_1)\alpha_{L+1}/\alpha_1)} \right.
 \end{aligned}$$

$$+ \sum_{j_{L+1}=J/\alpha_{L+1}+1}^{\infty} 2^{-2j_{L+1}(s_{L+1}-q_{L+1})} \left] \|f\|_{\mathbf{H}^{(s,s_{L+1})}(\Omega \times \Omega_{L+1})}^2.$$

If  $(s_1 - q_1)/\alpha_1 < (s_{L+1} - q_{L+1})/\alpha_{L+1}$ , the exponent in the first sum is always negative, and we obtain

$$\begin{aligned} & \| (I - \widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})}) f \|_{\mathbf{H}^{(\mathbf{q}, q_{L+1})}(\Omega \times \Omega_{L+1})}^2 \\ & \lesssim (2^{-2J(s_1 - q_1)/\alpha_1} J^{P-1} + 2^{-2J(s_{L+1} - q_{L+1})/\alpha_{L+1}}) \|f\|_{\mathbf{H}^{(s,s_{L+1})}(\Omega \times \Omega_{L+1})}^2 \\ & \lesssim 2^{-2J(s_1 - q_1)/\alpha_1} J^{P-1} \|f\|_{\mathbf{H}^{(s,s_{L+1})}(\Omega \times \Omega_{L+1})}^2. \end{aligned}$$

If  $(s_1 - q_1)/\alpha_1 = (s_{L+1} - q_{L+1})/\alpha_{L+1}$ , which may only happen if  $P = L$ , the exponent in the first sum is always 0, which, due to  $J/\alpha_{L+1} + 1$  terms in the sum, leads to

$$\begin{aligned} & \| (I - \widehat{\mathbf{Q}}_J^{(\alpha, \alpha_{L+1})}) f \|_{\mathbf{H}^{(\mathbf{q}, q_{L+1})}(\Omega \times \Omega_{L+1})}^2 \\ & \lesssim (2^{2J(s_1 - q_1)/\alpha_1} J^L + 2^{-2J(s_{L+1} - q_{L+1})/\alpha_{L+1}}) \|f\|_{\mathbf{H}^{(s,s_{L+1})}(\Omega \times \Omega_{L+1})}^2 \\ & \lesssim 2^{-2J(s_1 - q_1)/\alpha_1} J^L \|f\|_{\mathbf{H}^{(s,s_{L+1})}(\Omega \times \Omega_{L+1})}^2. \end{aligned}$$

This completes the proof. □

*Remark 4.4* (i) The constant in (4.7) depends on the particular choice of  $\alpha$  (and of course also of  $L, \mathbf{q}$ , and  $\mathbf{r}$ ).

(ii) For any  $1 \leq k \leq L$ , the sparse tensor product space  $\widehat{\mathbf{V}}_J^\alpha$  contains the full tensor product space  $V_0^{(1)} \otimes \dots \otimes V_0^{(k-1)} \otimes V_{J/\alpha_k}^{(k)} \otimes V_0^{(k+1)} \otimes \dots \otimes V_0^{(L)}$ . But it does not contain the space  $V_0^{(1)} \otimes \dots \otimes V_0^{(k-1)} \otimes V_{J/\alpha_{k+1}}^{(k)} \otimes V_0^{(k+1)} \otimes \dots \otimes V_0^{(L)}$ . Therefore, the convergence rate cannot be higher than  $2^{-J(s_k - q_k)/\alpha_k}$ . Repeating this argument for all  $1 \leq k \leq L$  and taking the minimum, we obtain the lower bound  $2^{-J \min\{\frac{s_1 - q_1}{\alpha_1}, \frac{s_2 - q_2}{\alpha_2}, \dots, \frac{s_L - q_L}{\alpha_L}\}}$ . Thus, we conclude that the error estimate (4.7) is essentially sharp.

(iii) Note that for  $\mathbf{s} \neq \mathbf{r}$ , the logarithmic factor  $J^{(R-1)/2}$  in (4.7) can be reduced or even removed by using more sophisticated estimates, see [14].

By combining Theorems 4.1 and 4.3, we can express the convergence rate in terms of the number of degrees of freedom  $N := \dim \widehat{\mathbf{V}}_J^\alpha$ . This gives us the cost complexity of the approximation of functions in the sparse tensor product spaces  $\widehat{\mathbf{V}}_J^\alpha$ .

**Corollary 4.5** *Let  $-\tilde{\gamma} < \mathbf{q} < \boldsymbol{\gamma}$ ,  $\mathbf{q} < \mathbf{s} \leq \mathbf{r}$ , and  $f \in \mathbf{H}^{\mathbf{s}}(\Omega)$ . Furthermore, denote by  $N := \dim \widehat{\mathbf{V}}_J^\alpha$  the number of degrees of freedom in the sparse tensor product space  $\widehat{\mathbf{V}}_J^\alpha$ , and set*

$$\beta := \frac{\min\{(s_1 - q_1)/\alpha_1, (s_2 - q_2)/\alpha_2, \dots, (s_L - q_L)/\alpha_L\}}{\max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}}.$$

Assume that the minimum in the numerator is attained  $P$  times and the maximum in the denominator is attained  $R$  times. Then, the approximation (4.6) in  $\widehat{\mathbf{V}}_J^\alpha$  possesses the following convergence rate in terms of the degrees of freedom  $N$ :

$$\|f - \widehat{f}_J\|_{\mathbf{H}^q(\Omega)} \lesssim N^{-\beta} (\log N)^{(P-1)/2+\beta(R-1)} \|f\|_{\mathbf{H}^s(\Omega)}. \tag{4.8}$$

*Proof* First assume  $R = 1$ . We then have  $N \sim 2^J \max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}$  due to Theorem 4.1. Hence, it holds that

$$N^{-\beta} = N^{-\frac{\min\{(s_1-q_1)/\alpha_1, (s_2-q_2)/\alpha_2, \dots, (s_L-q_L)/\alpha_L\}}{\max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}}} \sim 2^{-J \min\{\frac{s_1-q_1}{\alpha_1}, \frac{s_2-q_2}{\alpha_2}, \dots, \frac{s_L-q_L}{\alpha_L}\}},$$

which, together with (4.7), yields the desired error estimate (4.8) for  $R = 1$ .

If  $R > 1$ , then the sparse tensor product space  $\widehat{\mathbf{V}}_J^\alpha$  contains

$$N \sim 2^J \max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\} J^{R-1}$$

degrees of freedom. Thus, we now have the relation

$$\left(\frac{N}{J^{R-1}}\right)^{-\beta} \sim 2^{-J \min\{\frac{s_1-q_1}{\alpha_1}, \frac{s_2-q_2}{\alpha_2}, \dots, \frac{s_L-q_L}{\alpha_L}\}}.$$

Consequently, by taking the logarithm on both sides, we get  $\log(N/J^{R-1}) \sim J$ , and since  $\log(N/J^{R-1}) \leq \log N$ , we obtain from (4.7) the estimate

$$\|f - \widehat{f}_J\|_{L^2(\Omega)} \lesssim \left(\frac{N}{(\log N)^{R-1}}\right)^{-\beta} (\log N)^{(P-1)/2} \|f\|_{\mathbf{H}^s(\Omega)},$$

i.e., the desired error estimate (4.8) for  $R > 1$ . This completes the proof. □

### 5 Discussion of the Results

We now discuss the best choice of the parameter  $\alpha > \mathbf{0}$ . To this end, we intend to approximate a function  $f \in \mathbf{H}^s(\Omega)$  of maximal smoothness, i.e., we shall apply Corollary 4.5 in the case  $\mathbf{s} = \mathbf{r}$ . Our first question addresses the highest possible convergence rate  $\beta^*$  which can be achieved in the sparse tensor product spaces  $\widehat{\mathbf{V}}_J^\alpha$ .

**Lemma 5.1** *It holds that*

$$\begin{aligned} \beta &= \frac{\min\{(r_1 - q_1)/\alpha_1, (r_2 - q_2)/\alpha_2, \dots, (r_L - q_L)/\alpha_L\}}{\max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}} \\ &\leq \beta^* := \min\left\{\frac{r_1 - q_1}{n_1}, \frac{r_2 - q_2}{n_2}, \dots, \frac{r_L - q_L}{n_L}\right\} \end{aligned}$$

for all  $\alpha > \mathbf{0}$ .

*Proof* Choose  $\alpha > \mathbf{0}$  arbitrarily but fixed and let  $k, \ell \in \{1, 2, \dots, L\}$  be such that

$$\frac{r_k - q_k}{\alpha_k} = \min \left\{ \frac{r_1 - q_1}{\alpha_1}, \frac{r_2 - q_2}{\alpha_2}, \dots, \frac{r_L - q_L}{\alpha_L} \right\}$$

and

$$\frac{n_\ell}{\alpha_\ell} = \max \left\{ \frac{n_1}{\alpha_1}, \frac{n_2}{\alpha_2}, \dots, \frac{n_L}{\alpha_L} \right\}.$$

It thus holds that  $(r_k - q_k)/\alpha_k \leq (r_i - q_i)/\alpha_i$  and  $n_\ell/\alpha_\ell \geq n_i/\alpha_i$  for  $i = 1, 2, \dots, L$ . In particular, we find

$$\beta = \frac{r_k - q_k}{\alpha_k} \cdot \frac{\alpha_\ell}{n_\ell} \leq \frac{r_i - q_i}{\alpha_i} \cdot \frac{\alpha_i}{n_i} = \frac{r_i - q_i}{n_i} \quad \text{for all } i = 1, 2, \dots, L,$$

which immediately implies the assertion

$$\beta \leq \min \left\{ \frac{r_1 - q_1}{n_1}, \frac{r_2 - q_2}{n_2}, \dots, \frac{r_L - q_L}{n_L} \right\}.$$

□

This lemma thus gives an upper bound  $\beta^*$  for the convergence rate  $\beta$  in Corollary 4.5. We now show that this rate is essentially achieved for the canonical choices (i)–(iii) of  $\alpha$  of Sect. 3.

(i) The equilibration of the accuracy in the involved univariate spaces led to the choice  $\alpha_i = r_i - q_i$  for all  $i = 1, 2, \dots, L$ . According to Corollary 4.5 and since we consider  $\mathbf{s} = \mathbf{r}$ , we have here

$$\begin{aligned} \beta &= \frac{\min\{1, 1, \dots, 1\}}{\max\{n_1/(r_1 - q_1), n_2/(r_2 - q_2), \dots, n_L/(r_L - q_L)\}} \\ &= \min \left\{ \frac{r_1 - q_1}{n_1}, \frac{r_2 - q_2}{n_2}, \dots, \frac{r_L - q_L}{n_L} \right\} = \beta^*. \end{aligned}$$

We thus obtain the rate

$$\|f - \widehat{f}_J\|_{\mathbf{H}^q(\Omega)} \lesssim N^{-\beta^*} (\log N)^{(L-1)/2+\beta^*(R-1)} \|f\|_{\mathbf{H}^r(\Omega)}.$$

(ii) The equilibration of the degrees of freedom in the involved univariate spaces corresponds to the choice  $\alpha_i = n_i$  for all  $i = 1, 2, \dots, L$ . This immediately leads to

$$\beta = \frac{\min\{(r_1 - q_1)/n_1, (r_2 - q_2)/n_2, \dots, (r_L - q_L)/n_L\}}{\max\{1, 1, \dots, 1\}} = \beta^*.$$

Hence, the rate of convergence is

$$\|f - \widehat{f}_J\|_{\mathbf{H}^q(\Omega)} \lesssim N^{-\beta^*} (\log N)^{(P-1)/2+\beta^*(L-1)} \|f\|_{\mathbf{H}^r(\Omega)}.$$

(iii) The equilibration of the cost-benefit rate is given by  $\alpha_i = r_i - q_i + n_i$  for all  $i = 1, 2, \dots, L$ . Let  $k \in \{1, 2, \dots, L\}$  such that

$$\frac{r_k - q_k}{n_k + r_k - q_k} = \min \left\{ \frac{r_1 - q_1}{n_1 + r_1 - q_1}, \dots, \frac{r_L - q_L}{n_L + r_L - q_L} \right\}.$$

Then, it follows that

$$\frac{r_k - q_k}{n_k + r_k - q_k} \leq \frac{r_i - q_i}{n_i + r_i - q_i} \quad \text{for all } i = 1, 2, \dots, L, \tag{5.1}$$

which is equivalent to

$$\frac{r_k - q_k}{n_k} \leq \frac{r_i - q_i}{n_i} \quad \text{for all } i = 1, 2, \dots, L.$$

Hence, (5.1) also implies

$$\frac{n_k}{n_k + r_k - q_k} \geq \frac{n_i}{n_i + r_i - q_i} \quad \text{for all } i = 1, 2, \dots, L.$$

It therefore again holds that

$$\beta = \frac{\min\{(r_1 - q_1)/(n_1 + r_1 - q_1), \dots, (r_L - q_L)/(n_L + r_L - q_L)\}}{\max\{n_1/(n_1 + r_1 - q_1), \dots, n_L/(n_L + r_L - q_L)\}} = \beta^*,$$

and Corollary 4.5 yields

$$\|f - \widehat{f}_J\|_{\mathbf{H}^q(\Omega)} \lesssim N^{-\beta^*} (\log N)^{(P-1)/2 + \beta^*(R-1)} \|f\|_{\mathbf{H}^r(\Omega)}.$$

We can generalize these particular examples as follows. Set

$$\alpha_i = \lambda n_i + (1 - \lambda)(r_i - q_i) \quad \text{for all } i = 1, 2, \dots, L, \tag{5.2}$$

where  $\lambda \in [0, 1]$  is an arbitrarily chosen parameter. For  $\lambda = 1$ , we find  $\alpha_i = n_i$ ; for  $\lambda = 0$ , we find  $\alpha_i = r_i - q_i$ ; and for  $\lambda = 1/2$ , we find  $\alpha_i = (n_i + r_i - q_i)/2$ , which yields the same sparse tensor product spaces as the choice  $\alpha_i = n_i + r_i - q_i$ . Thus, the above examples are covered by (5.2).

For the choice (5.2), it holds for a specific  $k$  that

$$\frac{r_k - q_k}{\lambda n_k + (1 - \lambda)(r_k - q_k)} = \min_{i \in \{1, 2, \dots, L\}} \left\{ \frac{r_i - q_i}{\lambda n_i + (1 - \lambda)(r_i - q_i)} \right\}$$

if and only if

$$\frac{n_k}{\lambda n_k + (1 - \lambda)(r_k - q_k)} = \max_{i \in \{1, 2, \dots, L\}} \left\{ \frac{n_i}{\lambda n_i + (1 - \lambda)(r_i - q_i)} \right\}.$$

Consequently, for all  $\lambda \in [0, 1]$ , we find

$$\beta = \frac{\min_{i \in \{1, 2, \dots, L\}} \{(r_i - q_i)/(\lambda n_i + (1 - \lambda)(r_i - q_i))\}}{\max_{i \in \{1, 2, \dots, L\}} \{n_i/(\lambda n_i + (1 - \lambda)(r_i - q_i))\}} = \beta^*.$$

Nevertheless, the logarithmic factors in the convergence rate in (4.8) might differ in the extremal cases  $\lambda = 0$  and  $\lambda = 1$ .

In the case of  $L = 2$ , this construction covers all possible sparse tensor product spaces  $\widehat{\mathbf{V}}_j^\alpha$  which essentially (i.e., except for logarithmic terms) offer the highest possible convergence rate  $\beta^*$ , see [11]. Note however that in the case  $L > 2$ , there might be other choices which also result in optimal sparse tensor product spaces.

**Lemma 5.2** *Let  $k \in \{1, 2, \dots, L\}$  be such that*

$$\frac{r_k - q_k}{n_k} = \min \left\{ \frac{r_1 - q_1}{n_1}, \frac{r_2 - q_2}{n_2}, \dots, \frac{r_L - q_L}{n_L} \right\} = \beta^*.$$

*Then, for all  $\alpha > \mathbf{0}$  such that*

$$\frac{r_k - q_k}{r_i - q_i} \leq \frac{\alpha_k}{\alpha_i} \leq \frac{n_k}{n_i} \quad \text{for all } i = 1, 2, \dots, L, \tag{5.3}$$

*it holds that*

$$\beta = \frac{\min\{(r_1 - q_1)/\alpha_1, (r_2 - q_2)/\alpha_2, \dots, (r_L - q_L)/\alpha_L\}}{\max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}} = \beta^*.$$

*If (5.3) is not satisfied, then it holds that  $\beta < \beta^*$ .*

*Proof* The condition (5.3) immediately implies the equalities

$$\frac{r_k - q_k}{\alpha_k} \leq \frac{r_i - q_i}{\alpha_i} \quad \text{and} \quad \frac{n_i}{\alpha_i} \leq \frac{n_k}{\alpha_k} \quad \text{for all } i = 1, 2, \dots, L.$$

Hence,

$$\beta = \frac{\min\{(r_1 - q_1)/\alpha_1, (r_2 - q_2)/\alpha_2, \dots, (r_L - q_L)/\alpha_L\}}{\max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}} = \frac{r_k - q_k}{n_k} = \beta^*$$

indeed holds. If (5.3) is not satisfied for some specific  $j \in \{1, 2, \dots, L\}$ , then it follows that

$$\frac{r_j - q_j}{\alpha_j} < \frac{r_k - q_k}{\alpha_k} \quad \text{or} \quad \frac{n_k}{\alpha_k} < \frac{n_j}{\alpha_j}.$$

Therefore, we obtain

$$\beta = \frac{\min\{(r_1 - q_1)/\alpha_1, (r_2 - q_2)/\alpha_2, \dots, (r_L - q_L)/\alpha_L\}}{\max\{n_1/\alpha_1, n_2/\alpha_2, \dots, n_L/\alpha_L\}} \neq \beta^*,$$

which, in view of Lemma 5.1, shows  $\beta < \beta^*$ . □

### 6 Conclusion

The analysis in Sect. 5 reveals that all sparse tensor product spaces  $\widehat{\mathbf{V}}_j^\alpha$  with  $\alpha$  satisfying the criterion (5.3) produce essentially the optimal convergence rate  $\beta^* =$

$\min_{i \in \{1, 2, \dots, L\}} \{(r_i - q_i)/n_i\}$  when a function  $f \in \mathbf{H}^r(\Omega)$  of maximal smoothness is to be approximated. However, if the function to be approximated is only in  $\mathbf{H}^s(\Omega)$  with  $\mathbf{q} < \mathbf{s} < \mathbf{r}$ , then the highest possible convergence rate is

$$\beta = \min_{i \in \{1, 2, \dots, L\}} \{(s_i - q_i)/n_i\}.$$

To achieve this rate, the criterion (5.3) for the choice of  $\alpha > \mathbf{0}$  has to be modified according to

$$\frac{s_k - q_k}{s_i - q_i} \leq \frac{\alpha_k}{\alpha_i} \leq \frac{n_k}{n_i} \quad \text{for all } i = 1, 2, \dots, L.$$

As a consequence, to obtain for all  $\mathbf{q} < \mathbf{s} \leq \mathbf{r}$  the highest possible rate in  $\widehat{\mathbf{V}}_J^\alpha$ , one has to choose  $\alpha_i = n_i$  for all  $i = 1, 2, \dots, L$ , i.e., one has to equilibrate the degrees of freedom in the involved univariate spaces  $V_{J/\alpha_i}^{(i)}$ .

## References

- Allaire, G., Briane, M.: Multiscale convergence and reiterated homogenisation. Proc. R. Soc. Edinb. A **126**(2), 297–342 (1996)
- Balescu, R.: Statistical Dynamics, Matter Out of Equilibrium. Imperial College Press, London (1997)
- Barrett, J.W., Knezevic, D., Süli, E.: Kinetic models of dilute polymers: analysis, approximation and computation. In: 11th School on Mathematical Theory in Fluid Mechanics, Kacov, Czech Republic, 22–29 May 2009. Necas Center for Mathematical Modeling, Prague (2009)
- Bungartz, H.-J., Griebel, M.: Sparse grids. Acta Numer. **13**, 147–269 (2004)
- Chegini, N., Stevenson, R.: Adaptive wavelet schemes for parabolic problems. Sparse matrices and numerical results. SIAM J. Numer. Anal. **49**, 182–212 (2011)
- Cioranescu, D., Damlamian, A., Griso, G.: The periodic unfolding method in homogenization. SIAM J. Math. Anal. **40**(4), 1585–1620 (2008)
- Cioranescu, D., Damlamian, A., Donato, P., Griso, G., Zaki, R.: The periodic unfolding method in domains with holes. SIAM J. Math. Anal. **44**, 718–760 (2012)
- Dahmen, W.: Wavelet and multiscale methods for operator equations. Acta Numer. **6**, 55–228 (1997)
- Dung, D.: Continuous algorithms in  $n$ -term approximation and non-linear width. Constr. Approx. **102**, 217–242 (2000)
- Griebel, M., Hamaekers, J.: Tensor product multiscale many-particle spaces with finite-order weights for the electronic Schrödinger equation. Z. Phys. Chem. **224**, 527–543 (2010)
- Griebel, M., Harbrecht, H.: On the construction of sparse tensor product spaces. Math. Comp. **82**(282), 975–994 (2013)
- Griebel, M., Knapek, S.: Optimized general sparse grid approximation spaces for operator equations. Math. Comput. **78**(268), 2223–2257 (2009)
- Griebel, M., Oeltz, D.: A sparse grid space-time discretization scheme for parabolic problems. Computing **81**(1), 1–34 (2007)
- Griebel, M., Oswald, P., Schiekofer, T.: Sparse grids for boundary integral equations. Numer. Math. **83**(2), 279–312 (1999)
- Griebel, M., Oeltz, D., Vassilevski, P.: Space-time approximation with sparse grids. SIAM J. Sci. Comput. **28**, 701–727 (2006)
- Harbrecht, H., Li, J.: A fast deterministic method for stochastic elliptic interface problems based on low-rank approximation. Research Report No. 2011-24, Seminar für Angewandte Mathematik, ETH, Zürich (2011)
- Harbrecht, H., Schneider, R., Schwab, C.: Sparse second moment analysis for elliptic problems in stochastic domains. Numer. Math. **109**, 167–188 (2008)
- Harbrecht, H., Peters, M., Siebenmorgen, M.: Combination technique based  $k$ -th moment analysis of elliptic problems with random diffusion. Preprint No. 2011-2, Mathematisches Institut, Universität, Basel, Switzerland (2011)



19. Hoang, V.H., Schwab, C.: High-dimensional finite elements for elliptic problems with multiple scales. *Multiscale Model. Simul.* **3**, 168–194 (2005)
20. Le Bris, C., Lelièvre, T.: Multiscale modelling of complex fluids: a mathematical initiation. In: *Multiscale Modeling and Simulation in Science. Lecture Notes in Computational Science and Engineering*, vol. 66, pp. 49–138. Springer, Berlin (2009)
21. Lozinski, A., Owens, R.G., Phillips, T.N.: The Langevin and Fokker–Planck equations in polymer rheology. In: Glowinski, R. (ed.) *Handbook of Numerical Analysis XVI/XVII*. Elsevier/North-Holland, Amsterdam (2010)
22. Schwab, C., Stevenson, R.: Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comput.* **78**(267), 1293–1318 (2009)
23. Schwab, C., Todor, R.-A.: Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.* **95**(4), 707–734 (2003)
24. Schwab, C., Todor, R.-A.: Sparse finite elements for stochastic elliptic problems. Higher order moments. *Computing* **71**, 43–63 (2003)
25. Widmer, G., Hiptmair, R., Schwab, C.: Sparse adaptive finite elements for radiative transfer. *J. Comput. Phys.* **227**(12), 6071–6105 (2008)
26. Yserentant, H.: Sparse grid spaces for the numerical solution of the electronic Schrödinger equation. *Numer. Math.* **101**, 381–389 (2005)
27. Zeiser, A.: Wavelet approximation in weighted Sobolev spaces of mixed order with applications to the electronic Schrödinger equation. *Constr. Approx.* **35**, 293–322 (2012)
28. Zenger, C.: Sparse grids. In: *Parallel Algorithms for Partial Differential Equations*, Kiel, 1990. *Notes Numer. Fluid Mech.*, vol. 31, pp. 241–251. Vieweg, Braunschweig (1991)