On the Construction of Sparse Tensor Product Spaces

Michael Griebel, Helmut Harbrecht

no. 497

ON THE CONSTRUCTION OF SPARSE TENSOR PRODUCT SPACES

MICHAEL GRIEBEL AND HELMUT HARBRECHT

ABSTRACT. Let $\Omega_1 \subset \mathbb{R}^{n_1}$ and $\Omega_2 \subset \mathbb{R}^{n_2}$ be two given domains and consider on each domain a multiscale sequence of ansatz spaces of polynomial exactness r_1 and r_2, respectively. In this paper, we study the optimal construction of sparse tensor products made from these spaces. In particular, we derive the resulting cost complexities to approximate functions with anisotropic and isotropic smoothness on the tensor product domain $\Omega_1 \times \Omega_2$. Numerical results validate our theoretical findings.

1. INTRODUCTION

Many problems in science and engineering lead to problems which are defined on the tensor product of two domains $\Omega_1 \subset \mathbb{R}^{n_1}$ and $\Omega_2 \subset \mathbb{R}^{n_2}$. Examples arise from the second moment analysis of partial differential domains with stochastic input parameters [14, 15, 19, 22], two-scale homogenization [4, 16], radiosity models and radiative transfer [24], or space-time discretizations of parabolic problems [13, 23].

A straightforward discretization uses tensor products of all basis functions from suitable finite dimensional ansatz spaces $V_J^{(1)}$ and $V_J^{(2)}$ which are defined on each domain separately. This leads to the *full tensor product space* $V_J^{(1)} \otimes V_J^{(2)}$. However, in general, the full tensor product space contains too many degrees of freedom such that desirable realistic simulations are still beyond current computing capacities. For this reason, the efficient discretization of functions on product domains is an important task in numerical analysis and scientific computing.

In the present paper, we focus on the construction of *sparse tensor product spaces*, also known as sparse grids [2, 25]. Starting point are multilevel decompositions of the ansatz spaces

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

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

$$\widehat{V}_J^{reg} = \bigoplus_{j_1+j_2 \leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)},$$

see e.g. [2, 11, 12, 25]. 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. The regular sparse grid is optimal with respect to the L^2-norm if both domains have the *same* dimension and are equipped with the *same* type of functions.

Nevertheless, to the best of our knowledge, it has not been systematically studied in the literature 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. Here, the following questions arise: Should the degrees of freedom or the approximation power of the univariate ansatz spaces be equilibrated? Or is it preferable to construct the sparse tensor product space such that an equilibrated cost-benefit rate (see [2]) is guaranteed?

In the present paper, we will answer these questions for the case that the approximation error is measured in the $L^2(\Omega_1 \times \Omega_2)$-norm. Then, the sparse tensor product spaces contain all products $W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}$ with indices (j_1, j_2) contained inside a specific triangular subset of the j_1-j_2 plane with base dependent on n_1, n_2, r_1, r_2. Here, besides the approximation error of functions from Sobolev spaces of dominating mixed smoothness, we also study the approximation error of functions from isotropic Sobolev spaces.

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 turns out that the sparse tensor product space, which equilibrates the number of degrees of freedom (see Section 3), is superior to all the other sparse tensor product spaces under consideration.

To keep the discussion simple, we restrict ourselves to two-fold tensor product domains $\Omega_1 \times \Omega_2$ in this paper. We emphasize here that our results can easily be generalized to arbitrary tensor product domains $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_m$, compare also [2, 10, 18, 20, 21].

The remainder of this paper is organized as follows. In Section 2, we specify the requirements of the multiscale hierarchies on each subdomain. Then, in Section 3, we construct general sparse tensor product spaces. In Section 4, we study their properties. Section 5 is dedicated to the comparison of the cost complexities for the approximation of functions with anisotropic and isotropic smoothness. In Section 6, we provide the results of our numerical experiments. They are in good agreement with the presented theory.

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 denote 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 THE SUBDOMAINS

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

$$(2.1) \qquad\qquad V_0 \subset V_1 \subset \cdots \subset V_j \subset \cdots \subset L^2(\Omega),$$

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

$$(2.2) \qquad\qquad L^2(\Omega) = \overline{\bigcup_{j \in \mathbb{N}_0} V_j}, \qquad V_0 = \bigcap_{j \in \mathbb{N}_0} V_j.$$

Since we will use the spaces V_j for the approximation of functions, we assume that the approximation property

$$(2.3) \qquad\qquad \inf_{v_j \in V_j} \|u - v_j\|_{L^2(\Omega)} \lesssim h_j^s \|u\|_{H^s(\Omega)}, \quad u \in H^s(\Omega),$$

holds for $0 \le s \le r$ 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 Ω. Note that the integer r 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.,

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

Recursively we obtain

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

and thus, with

$$\Psi_J := \bigcup_{j=0}^{J} \Psi_j, \quad \Psi_0 := \Phi_0,$$

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 biorthogonal, or dual, wavelet basis $\widetilde{\Psi} = \bigcup_{j \geq 0} \widetilde{\Psi}_j = \{\widetilde{\psi}_{j,k} : k \in \nabla_j\}$ which defines a dual multiscale analysis, see e.g. [5] for further details. In particular, each function $f \in L^2(\Omega)$ admits the unique representation

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

With the definition of the projections

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

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

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

In particular, for any $f \in H^s(\Omega)$, the approximation property (2.3) induces the estimate

(2.5)
$$\|Q_j f\|_{L^2(\Omega)} \lesssim 2^{-js} \|f\|_{H^s(\Omega)}, \quad 0 \leq s \leq r.$$

3. Sparse tensor product spaces

Consider now two domains $\Omega_1 \subset \mathbb{R}^{n_1}$ and $\Omega_2 \subset \mathbb{R}^{n_2}$ with $n_1, n_2 \in \mathbb{N}$. We aim at the approximation of functions in $L^2(\Omega_1 \times \Omega_2)$. To this end, we assume individually for each subdomain Ω_i, $i = 1, 2$, the multiscale analyses

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

with associated complementary spaces

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

Furthermore, let us denote the polynomial exactnesses of the spaces $V_j^{(1)}$ and $V_j^{(2)}$ by r_1 and r_2, respectively.

In this paper, we study the special sparse tensor product space[1]

(3.6)
$$\widehat{V}_J^{\sigma} := \bigoplus_{j_1 \sigma + \frac{j_2}{\sigma} \leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}$$

---

[1]Here and in the following, 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.
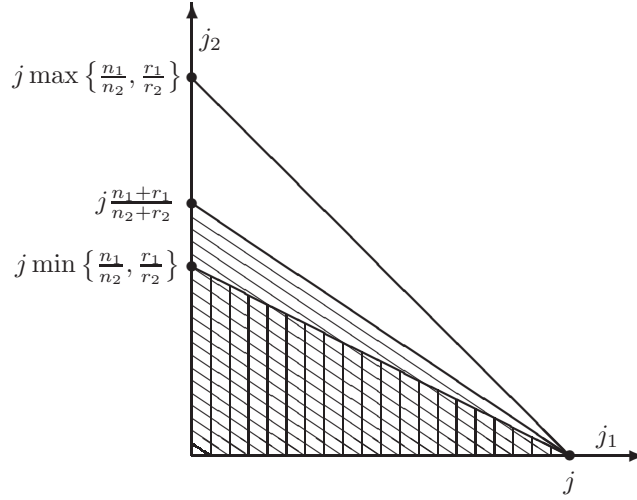
FIGURE 3.1. The index sets (j_1, j_2) of the spaces $\widehat{V}_J^\sigma$ with the same intersection point j on the j_1-axis for the three particular choices of σ.

for an *arbitrary* parameter $\sigma > 0$. In particular, the index pairs $(j_1, j_2) \in \mathbb{N}_0 \times \mathbb{N}_0$ of the included tensor product spaces $W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}$ satisfy the relations

$$0 \le j_1 \le \frac{1}{\sigma} J - \frac{1}{\sigma^2} j_2, \quad 0 \le j_2 \le \sigma J - \sigma^2 j_1.$$

Note that this ansatz would be not optimal if we would be interested in approximation errors with respect to norms other than the $L^2(\Omega_1 \times \Omega_2)$-norm. We would then no longer obtain spaces with triangular index set but so-called *generalized sparse grid spaces* [2, 8, 9, 11] with more complicatedly shaped index sets which exhibit no longer a linear relation between j_1 and j_2. For example, the choice of the $H^1(\Omega_1 \times \Omega_2)$-seminorm would lead to the so-called *energy based sparse grids*, see [2]. Results for other norms and smoothness classes can be found in [11].

Here, we restrict ourselves to the $L^2(\Omega_1 \times \Omega_2)$-norm but consider Ω_1 and Ω_2 with n_1 and n_2 in general not equal and possibly equipped with different polynomial ansatz spaces of exactness r_1 and r_2, respectively.

Reasonable choices of the parameter σ could then be as follows.

- We may equilibrate the degrees of freedom in all tensor product spaces $W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}$ whose indices (j_1, j_2) satisfy $j_1 \sigma + j_2/\sigma = J$. This choice leads to $\sigma = \sqrt{n_1/n_2}$.
- The sparse tensor product space $\widehat{V}_J^\sigma$ (3.6) can be rewritten as

$$\widehat{V}_J^\sigma = \sum_{j_1\sigma + j_2/\sigma = J} V_{j_1}^{(1)} \otimes V_{j_2}^{(2)}.$$

  Then, it can be seen easily that the choice $\sigma := \sqrt{r_1/r_2}$ equilibrates the approximation power of the contained tensor product spaces $V_{j_1}^{(1)} \otimes V_{j_2}^{(2)}$.
- Following the idea of an *equilibrated cost-benefit rate* (see [2]), we get the condition

$$2^{j_1(n_1+r_1)} \cdot 2^{j_2(n_2+r_2)} \overset{!}{=} 2^{J \cdot const}.$$

  Then, by choosing as $const = \sqrt{(n_1 + r_1)(n_2 + r_2)}$ we arrive at $\sigma = \sqrt{\frac{n_1+r_1}{n_2+r_2}}$.

In Figure 3.1, we display the index sets (j_1, j_2) which belong to the associated sparse tensor product spaces $\widehat{V}_J^\sigma$ for these three cases of σ.

4. Properties of the sparse tensor product spaces

In the following section, we present results for the approximation of functions in the sparse tensor product spaces $\widehat{V}_J^\sigma$ with arbitrary $\sigma > 0$. First, we count the degrees of freedom which are contained in these spaces.

Theorem 4.1. *The dimension of the sparse tensor product space*

$$\widehat{V}_J^\sigma = \sum_{\sigma j_1 + j_2/\sigma \le J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}$$

is essentially $\mathcal{O}(2^{J \max\{n_1/\sigma, n_2\sigma\}})$. More precisely, it holds

$$(4.7) \qquad \dim \widehat{V}_J^\sigma \lesssim \begin{cases} 2^{J \max\{n_1/\sigma, n_2\sigma\}}, & \text{if } n_1/\sigma \ne n_2\sigma, \\ 2^{Jn_2\sigma} J, & \text{if } n_1/\sigma = n_2\sigma. \end{cases}$$

Proof. Since $\dim\left(W_{j_1}^{(1)} \otimes W_{j_2}^{(2)}\right) \sim 2^{j_1 n_1 + j_2 n_2}$, it holds

$$\dim \widehat{V}_J^\sigma \sim \sum_{j_1=0}^{J/\sigma} \sum_{j_2=0}^{\sigma J - \sigma^2 j_1} 2^{j_1 n_1 + j_2 n_2} \lesssim 2^{Jn_2\sigma} \sum_{j_1=0}^{J/\sigma} 2^{j_1(n_1 - n_2\sigma^2)}.$$

Now, if $n_1/\sigma < n_2\sigma$, the exponent in the sum is always negative. This implies

$$(4.8) \qquad \dim \widehat{V}_J^\sigma \lesssim 2^{Jn_2\sigma}.$$

If $n_1/\sigma > n_2\sigma$, the exponent is always positive and it follows that

$$(4.9) \qquad \dim \widehat{V}_J^\sigma \lesssim 2^{Jn_2\sigma} 2^{J/\sigma(n_1 - n_2\sigma^2)} = 2^{Jn_1/\sigma}.$$

In the case of $n_1/\sigma = n_2\sigma$ we obtain

$$(4.10) \qquad \dim \widehat{V}_J^\sigma \lesssim 2^{Jn_2\sigma} \sum_{j_1=0}^{J/\sigma} 1 \lesssim 2^{Jn_2\sigma} \frac{J}{\sigma}.$$

The combination of (4.8)–(4.10) yields the desired result. $\qquad\qquad\qquad\qquad\square$

Remark 4.2. *(i) Note that the constant in estimate (4.7) depends on the particular choice of σ.*

(ii) Estimate (4.7) is sharp since it holds $\dim W_{J/\sigma}^{(1)} \sim 2^{Jn_1/\sigma}$ whereas $\dim W_{\sigma J}^{(2)} \sim 2^{Jn_2\sigma}$. Thus, we also have the lower bound $2^{J \max\{n_1/\sigma, n_2\sigma\}}$.

(iii) The full tensor product space $V_{J/\sigma}^{(1)} \otimes V_{J\sigma}^{(2)}$ consists of $2^{J(n_1/\sigma + n_2\sigma)}$ degrees of freedom.

In view of just optimal cost we would get $n_1/\sigma = n_2\sigma$, that is $\sigma = \sqrt{n_1/n_2}$, from balancing the terms in $\max\{n_1/\sigma, n_2\sigma\}$. But this choice of $\widehat{V}_J^\sigma$ needs not to give us necessarily the best rate of approximation yet. Therefore, our next result is concerned with the rate of approximation in the sparse tensor spaces $\widehat{V}_J^\sigma$. To this end, for $s_1, s_2 \ge 0$, we introduce the anisotropic Sobolev spaces

$$H_{mix}^{s_1,s_2}(\Omega_1 \times \Omega_2) := H^{s_1}(\Omega_1) \otimes H^{s_2}(\Omega_2).$$

Obviously, the highest possible rate of convergence is attained in the space $H_{mix}^{r_1,r_2}(\Omega_1 \times \Omega_2)$. Therefore, in the following theorem, we restrict ourselves without loss of generality to $s_1 \le r_1$ and $s_2 \le r_2$.

Theorem 4.3. *Let $0 < s_1 \le r_1$, $0 < s_2 \le r_2$ and $f \in H_{mix}^{s_1,s_2}(\Omega_1 \times \Omega_2)$. Then, the approximation*

$$(4.11) \qquad \widehat{f}_J = \sum_{j_1\sigma + \frac{j_2}{\sigma} \le J} \left(Q_{j_1}^{(1)} \otimes Q_{j_2}^{(2)}\right) f \in \widehat{V}_J^\sigma$$

satisfies

$$(4.12) \qquad \|f - \widehat{f}_J\|_{L^2(\Omega)} \lesssim \begin{cases} 2^{-J\min\{s_1/\sigma, s_2\sigma\}} \|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}, & \text{if } s_1/\sigma \neq s_2\sigma, \\ 2^{-Js_1/\sigma}\sqrt{J}\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}, & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

Proof. It follows by standard tensor product arguments from (2.5) that

$$\|f - \widehat{f}_J\|^2_{L^2(\Omega)} \lesssim \sum_{j_1\sigma + \frac{j_2}{\sigma} > J} \left\|\left(Q^{(1)}_{j_1} \otimes Q^{(2)}_{j_2}\right)f\right\|^2_{L^2(\Omega)}$$

$$\lesssim \sum_{j_1\sigma + \frac{j_2}{\sigma} > J} 2^{-2(s_1 j_1 + s_2 j_2)}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}.$$

Now, we split the index set

$$I := \left\{(j_1, j_2) : j_1\sigma + \frac{j_2}{\sigma} > J\right\}$$

into two disjoint sets $I = I_1 \cup I_2$ given by

$$I_1 := \left\{(j_1, j_2) : 0 \leq j_1 \leq \frac{J}{\sigma}, \ J\sigma - j_1\sigma^2 < j_2\right\},$$

$$I_2 := \left\{(j_1, j_2) : \frac{J}{\sigma} < j_1, 0 \leq j_2\right\}.$$

Thus, we get

$$\|f - \widehat{f}_J\|^2_{L^2(\Omega)} \lesssim \left\{\sum_{j_1=0}^{J/\sigma} \sum_{j_2 = J\sigma - j_1\sigma^2 + 1}^{\infty} 2^{-2(s_1 j_1 + s_2 j_2)} + \sum_{j_1 = J/\sigma + 1}^{\infty} \sum_{j_2 = 0}^{\infty} 2^{-2(s_1 j_1 + s_2 j_2)}\right\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$$

$$\lesssim \left\{\sum_{j_1=0}^{J/\sigma} 2^{-2j_1(s_1 - s_2\sigma^2) - 2Js_2\sigma} + \sum_{j_1 = J/\sigma+1}^{\infty} 2^{-2s_1 j_1}\right\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$$

$$\lesssim \left\{\sum_{j_1=0}^{J/\sigma} 2^{-2j_1(s_1 - s_2\sigma^2) - 2Js_2\sigma} + 2^{-2J\frac{s_1}{\sigma}}\right\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$$

$$= 2^{-2Js_2\sigma}\left\{\sum_{j_1=0}^{J/\sigma} 2^{-2j_1\sigma(\frac{s_1}{\sigma} - s_2\sigma)} + 2^{-2J(\frac{s_1}{\sigma} - s_2\sigma)}\right\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}.$$

To obtain an estimate for this expression, we again distinguish three cases.

In the case $s_1/\sigma < s_2\sigma$, we find

$$\|f - \widehat{f}_J\|^2_{L^2(\Omega)} \lesssim 2^{-2Js_2\sigma}\left\{2^{-2J(\frac{s_1}{\sigma} - s_2\sigma)} + 2^{-2J(\frac{s_1}{\sigma} - s_2\sigma)}\right\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$$

$$\lesssim 2^{-2J\frac{s_1}{\sigma}}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}.$$

If $s_1/\sigma > s_2\sigma$, we conclude that

$$\|f - \widehat{f}_J\|^2_{L^2(\Omega)} \lesssim 2^{-2Js_2\sigma}\{1 + 1\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$$

$$\lesssim 2^{-2Js_2\sigma}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}.$$

Finally, if $s_1/\sigma = s_2\sigma$, we have

$$\|f - \widehat{f}_J\|^2_{L^2(\Omega)} \lesssim 2^{-2Js_2\sigma}\left\{\sum_{j_1=0}^{J/\sigma} 1 + 1\right\}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$$

$$\lesssim 2^{-2Js_2\sigma}\frac{J}{\sigma}\|f\|^2_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Remark 4.4. *(i) The constant in estimate* (4.12) *depends again on the particular choice of σ.*

(ii) As before, one readily verifies that the estimate (4.12) *is essentially sharp. Nevertheless, if $s_1 < r_1$ and $s_2 < r_2$, then the factor $\sqrt{J}$ for the case $s_1/\sigma = s_2\sigma$ in* (4.12) *can be removed by using more sophisticated estimates, see* [12, 19].

(iii) We essentially obtain the same error estimate in the full tensor product space $V_{J/\sigma}^{(1)} \otimes V_{J\sigma}^{(2)}$.

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

Corollary 4.5. *Let $0 < s_1 \le r_1$, $0 < s_2 \le r_2$ and $f \in H_{mix}^{s_1,s_2}(\Omega_1 \times \Omega_2)$. Furthermore, denote by $N := \dim \widehat{V}_J^\sigma$ the number of degrees of freedom in the sparse tensor product space $\widehat{V}_J^\sigma$ and set*

$$\alpha := \frac{\min\{s_1/\sigma, s_2\sigma\}}{\max\{n_1/\sigma, n_2\sigma\}}.$$

If $n_1/\sigma \ne n_2\sigma$, then the approximation (4.11) *in $\widehat{V}_J^\sigma$ produces the following convergence rate in terms of the degrees of freedom N:*

$$\|f - \widehat{f}_J\|_{L^2(\Omega)} \lesssim \begin{cases} N^{-\alpha}\|f\|_{H_{mix}^{s_1,s_2}(\Omega_1\times\Omega_2)}, & \text{if } s_1/\sigma \ne s_2\sigma, \\ N^{-\alpha}\sqrt{\log N}\|f\|_{H_{mix}^{s_1,s_2}(\Omega_1\times\Omega_2)}, & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

If $n_1/\sigma = n_2\sigma$, we have

$$\|f - \widehat{f}_J\|_{L^2(\Omega)} \lesssim \begin{cases} N^{-\alpha}(\log N)^\alpha\|f\|_{H_{mix}^{s_1,s_2}(\Omega_1\times\Omega_2)}, & \text{if } s_1/\sigma \ne s_2\sigma, \\ N^{-\alpha}(\log N)^{\alpha+\frac{1}{2}}\|f\|_{H_{mix}^{s_1,s_2}(\Omega_1\times\Omega_2)}, & \text{if } s_1/\sigma = s_2\sigma. \end{cases}$$

Proof. If $n_1/\sigma \ne n_2\sigma$ we have $N \sim 2^{J\max\{n_1/\sigma,n_2\sigma\}}$ due to (4.7). Hence, it holds

$$N^{-\alpha} = N^{-\frac{\min\{s_1/\sigma,s_2\sigma\}}{\max\{n_1/\sigma,n_2\sigma\}}} \sim 2^{-J\min\{s_1/\sigma,s_2\sigma\}}$$

which, together with (4.12), yields the first error estimate.

If $n_1/\sigma = n_2\sigma$, then the sparse tensor product space $\widehat{V}_J^\sigma$ contains $N \sim 2^{Jn_2\sigma} = 2^{J\min\{n_1/\sigma,n_2\sigma\}}J$ degrees of freedom. Consequently, by noting that $J \lesssim \log N$, we obtain from (4.12) the estimate

$$\inf_{\widehat{f}_J\in\widehat{V}_J^\sigma} \|f - \widehat{f}_J\|_{L^2(\Omega)} \lesssim \left(\frac{N}{\log N}\right)^{-\alpha}\|f\|_{H_{mix}^{s_1,s_2}(\Omega_1\times\Omega_2)}$$

provided that $s_1/\sigma \ne s_2\sigma$.

In case of $s_1/\sigma = s_2\sigma$, the additional factor $\sqrt{\log J} \sim \sqrt{\log(N/J)} \lesssim \sqrt{\log(N)}$ needs to be inserted as a multiplicative factor. This completes the proof. $\qquad\qquad$ $\square$

5. DISCUSSION OF THE RESULTS

5.1. Maximal regularity.
In Corollary 4.5 we computed the cost complexities when approximating functions from $H_{mix}^{s_1,s_2}(\Omega_1 \times \Omega_2)$ in the sparse tensor product spaces $\widehat{V}_J^\sigma$ with arbitrary $\sigma > 0$ and for $0 < s_1 \le r_1$ and $0 < s_2 \le r_2$. Now, we derive the cost complexity needed for the representation of a given function $f \in H_{mix}^{r_1,r_2}(\Omega_1 \times \Omega_2)$ with maximal regularity, i.e., for $s_1 = r_1$ and $s_2 = r_2$. Note that related results have been computed in [3, 13] for the particular situation of space-time discretizations of parabolic problems (i.e., $n_1 = 1$, $n_2 \ge 1$).

In the following lemma we exclude the case $\frac{r_1}{n_1} = \frac{r_2}{n_2}$ to ensure $\frac{n_1}{\sigma} \ne n_2\sigma$ and $\frac{r_1}{\sigma} \ne r_2\sigma$.

Lemma 5.1. *Assume that $\frac{r_1}{n_1} \neq \frac{r_2}{n_2}$ and let σ be of the form*

$$(5.13) \qquad \sigma = \sqrt{\frac{(1-\lambda)n_1 + \lambda r_1}{(1-\lambda)n_2 + \lambda r_2}}, \qquad \lambda \in (0,1).$$

Then, the cost complexity to approximate a function $f \in H_{mix}^{r_1,r_2}(\Omega_1 \times \Omega_2)$ in the sparse tensor product spaces $\widehat{V}_J^\sigma$ is

$$(5.14) \qquad \mathcal{O}\left(N^{-\min\left\{\frac{r_1}{n_1},\frac{r_2}{n_2}\right\}}\right).$$

Proof. Assume that $\frac{r_1}{n_1} < \frac{r_2}{n_2}$. Then, due to the inequalities

$$r_1\big((1-\lambda)n_2 + \lambda r_2\big) < r_2\big((1-\lambda)n_1 + \lambda r_1\big),$$
$$n_1\big((1-\lambda)n_2 + \lambda r_2\big) > n_2\big((1-\lambda)n_1 + \lambda r_1\big),$$

it follows that

$$\min\left\{\frac{r_1}{\sigma}, r_2\sigma\right\} = \frac{r_1}{\sigma} \quad \text{and} \quad \max\left\{\frac{n_1}{\sigma}, n_2\sigma\right\} = \frac{n_1}{\sigma}.$$

Analogously, in the case $\frac{r_1}{n_1} > \frac{r_2}{n_2}$, it holds

$$\min\left\{\frac{r_1}{\sigma}, r_2\sigma\right\} = r_2\sigma \quad \text{and} \quad \max\left\{\frac{n_1}{\sigma}, n_2\sigma\right\} = n_2\sigma.$$

Consequently, from Corollary 4.5, for $s_1 = r_1$ and $s_2 = r_2$, the cost complexity is

$$\mathcal{O}\left(N^{-\frac{\min\{r_1/\sigma, r_2\sigma\}}{\max\{n_1/\sigma, n_2\sigma\}}}\right) = \mathcal{O}\left(N^{-\min\left\{\frac{r_1}{n_1},\frac{r_2}{n_2}\right\}}\right).$$

$\square$

Remark 5.2. *(i) For the left limit $\lambda = 0$ it holds $\sigma = \sqrt{n_1/n_2}$ and we have $n_1/\sigma = n_2\sigma = \sqrt{n_1 n_2}$. Thus, the sparse tensor product space $\widehat{V}_J^\sigma$ contains $\mathcal{O}(2^{J\sqrt{n_1 n_2}}J)$ degrees of freedom due to Theorem 4.1. Hence, the logarithmic term $(\log N)^{\min\{\frac{r_1}{n_1},\frac{r_2}{n_2}\}}$ has to be inserted as multiplicative factor into (5.14).*

(ii) For the right limit $\lambda = 1$ we have $\sigma = \sqrt{r_1/r_2}$ and thus $r_1/\sigma = r_2\sigma = \sqrt{r_1 r_2}$. Then, with Theorem 4.3, we achieve only an accuracy of order $\mathcal{O}(2^{-J\sqrt{r_1 r_2}}\sqrt{J})$ in the sparse tensor product space $\widehat{V}_J^\sigma$. Thus, the logarithmic term $\sqrt{\log(N)}$ has to be inserted as multiplicative factor into (5.14).

(iii) If $\frac{r_1}{n_1} = \frac{r_2}{n_2}$, we have $n_1/\sigma = n_2\sigma$ and $r_1/\sigma = r_2\sigma$ for all $\lambda = [0,1]$. Therefore, both types of logarithmic terms have to be inserted as multiplicative factors into the cost complexity estimate (5.14) in this situation.

One readily verifies that $\sigma = \sigma(\lambda)$ from (5.13) satisfies the inequality

$$(5.15) \qquad \min\left\{\frac{r_1}{r_2}, \frac{n_1}{n_2}\right\} \leq \sigma^2 \leq \max\left\{\frac{r_1}{r_2}, \frac{n_1}{n_2}\right\} \quad \text{for all } \lambda \in [0,1].$$

In particular, since the function $\sigma = \sigma(\lambda)$ is continuous in λ, each value between the upper and lower bound is admitted. Via Lemma 5.1 and Remark 5.2 we covered all such choices of σ.

In the following lemma we consider the remaining choices of σ, i.e., the case that the parameter σ lies outside of the interval given by (5.15). We will prove that the convergence rate is then substantially smaller than that in (5.14). This result shows that all interesting parameters σ are of the form (5.13). We therefore can restrict ourselves to such choices of σ (cf. Figure 5.2). Note finally that the interval in (5.15) degenerates to the single point $\sigma^2 = \frac{n_1}{n_2}$ in the case $\frac{r_1}{n_1} = \frac{r_2}{n_2}$.

FIGURE 5.2. The optimal convergence rate is achieved for all σ^2 which are in the interval $\left(\min\left\{\frac{r_1}{r_2},\frac{n_1}{n_2}\right\},\max\left\{\frac{r_1}{r_2},\frac{n_1}{n_2}\right\}\right)$. If σ^2 is an endpoint of this interval, logarithmic factors appear. The convergence rate is substantially smaller outside the interval.

Lemma 5.3. *Let σ be such that*

$$\sigma^2 < \min\left\{\frac{r_1}{r_2},\frac{n_1}{n_2}\right\} \quad or \quad \max\left\{\frac{r_1}{r_2},\frac{n_1}{n_2}\right\} < \sigma^2.$$

Then, the convergence rate is substantially smaller than that in (5.14).

Proof. Under the assumption $\sigma^2 < \min\left\{\frac{r_1}{r_2},\frac{n_1}{n_2}\right\}$ it follows that

$$\min\left\{\frac{r_1}{r_2},\sigma^2\right\} = \sigma^2 \quad \text{and} \quad \max\left\{\frac{n_1}{n_2},\sigma^2\right\} = \frac{n_1}{n_2}.$$

Consequently, according to Corollary 4.5, the cost complexity is essentially, i.e., up to $\log N$-terms,

$$\mathcal{O}\left(N^{-\frac{\min\left\{\frac{r_1}{\sigma},r_2\sigma\right\}}{\max\left\{\frac{n_1}{\sigma},n_2\sigma\right\}}}\right) = \mathcal{O}\left(N^{-\frac{\frac{r_2}{\sigma}\min\left\{\frac{r_1}{r_2},\sigma^2\right\}}{\frac{n_2}{\sigma}\max\left\{\frac{n_1}{n_2},\sigma^2\right\}}}\right) = \mathcal{O}\left(N^{-\frac{r_2\sigma^2}{n_1}}\right).$$

We therefore have to show that

$$\frac{r_2\sigma^2}{n_1} < \min\left\{\frac{r_1}{n_1},\frac{r_2}{n_2}\right\}.$$

But this inequality follows immediately from

$$\frac{r_2\sigma^2}{n_1} < \frac{r_1}{n_1} \Leftrightarrow \sigma^2 < \frac{r_1}{r_2} \quad \text{and} \quad \frac{r_2\sigma^2}{n_1} < \frac{r_2}{n_2} \Leftrightarrow \sigma^2 < \frac{n_1}{n_2}.$$

In complete analogy the assertion is shown for the case $\max\left\{\frac{r_1}{r_2},\frac{n_1}{n_2}\right\} < \sigma^2$ where the cost complexity is essentially $\mathcal{O}\left(N^{-r_1/(n_2\sigma^2)}\right)$. $\square$

5.2. **Approximation of functions with anisotropic mixed Sobolev smoothness.** If we consider the approximation of a function $f \in H_{mix}^{s_1,s_2}(\Omega_1 \times \Omega_2)$ with *arbitrary* $0 < s_1 \le r_1$ and $0 < s_2 \le r_2$, then we get results which are completely analogous to the last subsection. The only difference is that r_1 has to be replaced by s_1 and r_2 has to be replaced by s_2. This means, for all σ which are contained in the interval

(5.16) $$\min\left\{\frac{s_1}{s_2},\frac{n_1}{n_2}\right\} \le \sigma^2 \le \max\left\{\frac{s_1}{s_2},\frac{n_1}{n_2}\right\},$$

the error $\|f - \widehat{f_J}\|_{L^2(\Omega_1 \times \Omega_2)}$ behaves essentially like

$$\mathcal{O}\left(N^{-\min\left\{\frac{s_1}{n_1},\frac{s_2}{n_2}\right\}}\right).$$

For all σ which are outside the above interval, the convergence rate is substantially smaller. Obviously, equilibrating the number of unknowns is now the best choice since $\sigma = \sqrt{n_1/n_2}$ is the only point which is always contained in the optimality interval given by (5.16).

5.3. **Approximation of functions with isotropic Sobolev smoothness.** In the following discussion, we consider the approximation of an isotropic function $f \in H^p(\Omega_1 \times \Omega_2)$. If $p \geq r_1 + r_2$ we fall into the maximal regularity case, discussed in Subsection 5.1, since $H_{mix}^{r_1,r_2}(\Omega_1 \times \Omega_2) \subset H^{r_1+r_2}(\Omega_1 \times \Omega_2)$.

We therefore may assume that $p < r_1 + r_2$. The highest rate of convergence is then achieved if the minimum of s_1/σ and $s_2\sigma$ constrained to $s_1 + s_2 = p$ is maximal. Since the maximum is attained if $s_1/\sigma = s_2\sigma$, we arrive at the choice

$$(5.17) \qquad s_1 := \min\left\{\frac{p\sigma^2}{\sigma^2+1}, r_1\right\}, \quad s_2 := \min\left\{\frac{p}{\sigma^2+1}, r_2\right\}.$$

We then have the following rates which depend on σ:

Theorem 5.4. *Let $f \in H^p(\Omega_1 \times \Omega_2)$ be a given function and let σ be as in (5.13). Choose s_1, s_2 as in (5.17). Then, as long as*

$$(5.18) \qquad \frac{p\sigma}{1+\sigma^2} \geq \min\left\{\frac{r_1}{\sigma}, r_2\sigma\right\},$$

the function f can be approximated in $\widehat{V}_J^\sigma$ with the maximal convergence rate as specified in Lemma 5.1 and Remark 5.2. Otherwise, we get essentially, i.e., up to $\log N$-terms, the reduced cost complexity $\mathcal{O}(N^{-\beta})$ with

$$\beta = \frac{p}{n_1}\frac{(1-\lambda)n_1 + \lambda r_1}{(1-\lambda)(n_1+n_2) + \lambda(r_1+r_2)} \quad \text{if } \frac{r_1}{n_1} \leq \frac{r_2}{n_2}$$

and

$$\beta = \frac{p}{n_2}\frac{(1-\lambda)n_2 + \lambda r_2}{(1-\lambda)(n_1+n_2) + \lambda(r_1+r_2)} \quad \text{if } \frac{r_1}{n_1} > \frac{r_2}{n_2}.$$

Proof. In case of (5.18), the assertion is obvious due to Corollary 4.5. We therefore may assume that

$$\frac{p\sigma}{1+\sigma^2} = \min\left\{\frac{s_1}{\sigma}, s_2\sigma\right\}.$$

In case of $\frac{r_1}{n_1} \leq \frac{r_2}{n_2}$ it follows that $\max\{n_1/\sigma, n_2\sigma\} = n_1/\sigma$ and thus the cost complexity is essentially

$$(5.19) \qquad \mathcal{O}\left(N^{-\frac{\frac{p\sigma}{1+\sigma^2}}{\max\{\frac{n_1}{\sigma}, n_2\sigma\}}}\right) = \mathcal{O}\left(N^{-\frac{p}{n_1}\frac{(1-\lambda)n_1+\lambda r_1}{(1-\lambda)(n_1+n_2)+\lambda(r_1+r_2)}}\right).$$

Vice versa, if $\frac{r_1}{n_1} > \frac{r_2}{n_2}$, it holds that $\max\{n_1/\sigma, n_2\sigma\} = n_2\sigma$ and consequently the cost complexity is essentially

$$(5.20) \qquad \mathcal{O}\left(N^{-\frac{\frac{p\sigma}{1+\sigma^2}}{\max\{\frac{n_1}{\sigma}, n_2\sigma\}}}\right) = \mathcal{O}\left(N^{-\frac{p}{n_2}\frac{(1-\lambda)n_2+\lambda r_2}{(1-\lambda)(n_1+n_2)+\lambda(r_1+r_2)}}\right).$$

This completes the proof. $\qquad\square$

Remark 5.5. *For $\sigma = \sqrt{r_1/r_2}$, that is $\lambda = 1$, the choice of s_1, s_2 from (5.17) leads to $s_1 = r_1 \min\left\{\frac{p}{r_1+r_2}, 1\right\} = r_1\frac{p}{r_1+r_2}$ and $s_2 = r_2 \min\left\{\frac{p}{r_1+r_2}, 1\right\} = r_2\frac{p}{r_1+r_2}$ provided that $p < r_1 + r_2$. Therefore, (5.18) is never satisfied if $p < r_1 + r_2$ and we get always the reduced cost complexity $\mathcal{O}(N^{-\gamma})$ with*

$$\gamma = \frac{\sqrt{r_1 r_2}\frac{p}{r_1+r_2}}{\max\left\{\frac{n_1}{\sigma}, n_2\sigma\right\}} = -\frac{p}{r_1+r_2}\min\left\{\frac{r_1}{n_1}, \frac{r_2}{n_2}\right\}$$

which is equivalent to (5.19) and (5.20), respectively.

The following result shows that, as long as the convergence rate is not maximal, the best setting to approximate an isotropic function is provided by the choice $\sigma = \sqrt{n_1/n_2}$.

Proposition 5.6. *Let $f \in H^p(\Omega_1 \times \Omega_2)$ be a given function with arbitrary $0 < p < r_1 + r_2$ and σ as in (5.13). Then, the approximation power of $\widehat{V}_J^\sigma$ is the better the smaller the parameter λ is. The highest convergence rate results therefore for the choice $\lambda = 0$, i.e., $\sigma = \sqrt{n_1/n_2}$, while the lowest convergence rate results from the choice $\lambda = 1$, i.e., $\sigma = \sqrt{r_1/r_2}$.*

Proof. To prove the assertion it suffices to show that, for fixed $p < r_1 + r_2$, the convergence rates (5.19) and (5.20) decrease as the parameter λ increases.

Assume first that $\frac{r_1}{n_1} < \frac{r_2}{n_2}$. Then, the function

$$g(\lambda) := \frac{(1-\lambda)n_1 + \lambda r_1}{(1-\lambda)(n_1 + n_2) + \lambda(r_1 + r_2)}$$

is monotonically decreasing due to

$$g(\lambda)' = \frac{r_1 n_2 - n_1 r_2}{((1-\lambda)(n_1 + n_2) + \lambda(r_1 + r_2))^2} < 0.$$

Hence, for $\lambda < \mu$, we have

$$\frac{p}{n_1} \frac{(1-\lambda)n_1 + \lambda r_1}{(1-\lambda)(n_1 + n_2) + \lambda(r_1 + r_2)} \geq \frac{p}{n_1} \frac{(1-\mu)n_1 + \mu r_1}{(1-\mu)(n_1 + n_2) + \mu(r_1 + r_2)}$$

which, in view of (5.19), is the first part of the assertion.

Likewise, it holds that

$$h(\lambda) := \frac{(1-\lambda)n_2 + \lambda r_2}{(1-\lambda)(n_1 + n_2) + \lambda(r_1 + r_2)}$$

is monotonically decreasing provided that $\frac{r_1}{n_1} > \frac{r_2}{n_2}$. This gives

$$\frac{p}{n_2} \frac{(1-\lambda)n_2 + \lambda r_2}{(1-\lambda)(n_1 + n_2) + \lambda(r_1 + r_2)} \geq \frac{p}{n_2} \frac{(1-\mu)n_2 + \mu r_2}{(1-\mu)(n_1 + n_2) + \mu(r_1 + r_2)}$$

for $\lambda < \mu$. We then obtain the second part of the assertion with (5.19) and (5.20). $\square$

Remark 5.7. *It is well known that the highest rate, which is achieved by full tensor product spaces, is given by Kolmogorov's n-width, see [17]. If $r_1 = r_2$, it is $N^{-p/(n_1+n_2)}$, i.e., $N \lesssim \varepsilon^{(n_1+n_2)/p}$, for Sobolev balls in H^p provided that $p \leq r_1$. From Theorem 5.4 it follows that the sparse grid space $\widehat{V}_J^\sigma$ with $\sigma = \sqrt{n_1/n_2}$ also achieves this rate up to logarithmic terms. Note finally that in the case $p > r_1$ our construction implicitly exploits the then appearing mixed regularity and thus obtains better rates.*

6. Numerical experiments

We now present the results of our numerical experiments. To this end, we consider smooth and non-smooth model functions. The practically obtained rates are in good agreement with the presented theory.

6.1. **An alternative decomposition of the sparse grid space.** First, we give a hint on an alternative decomposition of $\widehat{V}_J^\sigma$, which allows a simpler implementation in our case of two-fold tensor-product domains. On Ω_1 we use wavelets for the discretization which gives rise to a decomposition

$$(6.21) \qquad V_{J\sigma}^{(1)} = \bigoplus_{j=0}^{J\sigma} W_j^{(1)} \subset L^2\big((0,1)\big),$$

i.e., we have the sequence

$$W_0^{(1)}, \quad W_1^{(1)}, \quad \cdots \quad W_{J\sigma}^{(1)}.$$

On Ω_2 we apply finite elements with sufficient polynomial exactness on a series of hierarchical triangular meshes which stem from uniform refinement. They generate a family of nested single-scale spaces

$$(6.22) \qquad V_0^{(2)} \subset V_1^{(2)} \subset \cdots \subset V_{J/\sigma}^{(2)} \subset L^2\big((0,1)^2\big).$$

We define the sparse grid space $\widehat{V}_J^\sigma$ by

$$(6.23) \qquad \widehat{V}_J^\sigma := \bigoplus_{j_1\sigma+j_2/\sigma=J} W_{j_1}^{(1)} \otimes V_{j_2}^{(2)}.$$

This definition looks different from that in (3.6). But it is indeed equivalent since

$$\widehat{V}_J^\sigma = \bigoplus_{j_1\sigma+j_2/\sigma\leq J} W_{j_1}^{(1)} \otimes W_{j_2}^{(2)} = \bigoplus_{j_1=0}^{J/\sigma} W_{j_1}^{(1)} \otimes \bigg(\underbrace{\bigoplus_{j_2=0}^{J\sigma-j_1\sigma^2} W_{j_2}^{(2)}}_{=V_{J\sigma-j_1\sigma^2}^{(2)}} \bigg) = \bigoplus_{j_1\sigma+j_2/\sigma=J} W_{j_1}^{(1)} \otimes V_{j_2}^{(2)}.$$

This way, standard finite element discretizations for the sequence (6.22) can directly be employed and the need of a wavelet construction on Ω_2 is avoided. This significantly simplifies programming since it allows, at least for problems involving product operators, the reuse of existing FEM-code on Ω_2, makes the use of standard multilevel solvers directly possible and helps to circumvent the difficulties a direct wavelet discretization may pose for e.g. the treatment of boundary conditions there.

6.2. **Smooth functions.** We consider the situation of $\Omega_1 = (0,1)$ and $\Omega_2 = (0,1)^2$, that is $n_1 = 1$ and $n_2 = 2$. On Ω_1 we use piecewise constant (i.e. $r_1 = 1$) or linear (i.e. $r_1 = 2$) wavelets for the discretization which gives rise to a decomposition (6.21). On Ω_2 we apply either piecewise constant (i.e., $r_2 = 1$) or globally continuous, piecewise linear (i.e., $r_2 = 2$) finite elements on a series of hierarchical triangular meshes which stem from uniform refinement. They generate the family of nested single-scale spaces $\{V_j^{(2)}\}$. We then employ the decomposition (6.23) for the approximation in $L^2(\Omega_1 \times \Omega_2)$.

For different choices of r_1, r_2, and σ, we compute the approximation error of the sparse grid interpolant of the C^∞-function

$$f(x,\mathbf{y}) = \sin(y_1) + \sin^2(2\pi x)(y_1 + y_2)y_2.$$

We first apply piecewise constant finite elements for the sequence $\{V_j^{(2)}\}$ on Ω_2 in $\widehat{V}_J^\sigma$. In Figure 6.3, we plot the measured approximation error versus the number of degrees of freedom for the three values $\sigma = \sqrt{r_1/r_2}$, $\sigma = \sqrt{n_1/n_2}$, and $\sigma = \sqrt{(r_1 + n_1)/(r_2 + n_2)}$. On the interval Ω_1 we employ piecewise constant (left plot) and piecewise linear (right plot) wavelets. All the curves are
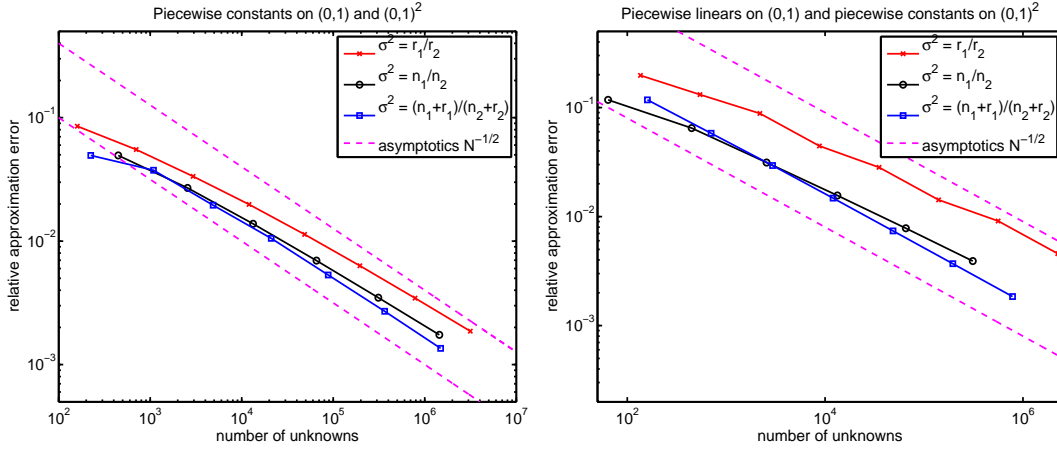
FIGURE 6.3. Cost complexities for approximating a smooth function in case of $n_1 = 1$, $n_2 = 2$ and $r_1 = r_2 = 1$ (left) or $r_1 = 2$, $r_2 = 1$ (right).
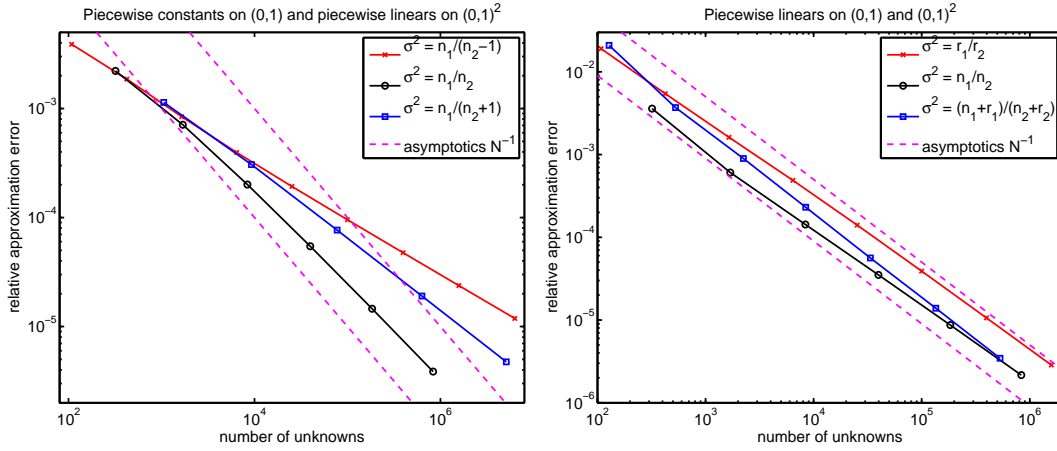

FIGURE 6.4. Cost complexities for approximating a smooth function in case of $n_1 = 1$, $n_2 = 2$ and $r_1 = 1$, $r_2 = 2$ (left) or $r_1 = r_2 = 2$ (right).

quite similar and behave essentially like $N^{-1/2}$ (indicated by the dashed line) as predicted by Lemma 5.1.

The best cost complexity rate is observed for the cost-benefit equilibrated sparse grid space, i.e., for $\sigma = \sqrt{(r_1 + n_1)/(r_2 + n_2)}$, which is in accordance with Remark 5.2. It indeed seems to be linear with $N^{-1/2}$ asymptotics whereas the cost complexity rates for $\sigma = \sqrt{r_1/r_2}$ and $\sigma = \sqrt{n_1/n_2}$ exhibit additional logarithmic factors as stated in Remark 5.2 (ii) and (iii). This holds for both cases, i.e., $r_1 = r_2 = 1$ and $r_1 = 2$, $r_2 = 1$.

Analogous observations are made if we employ piecewise linear wavelets on Ω_1 and globally continuous, piecewise linear finite elements on Ω_2. The results are given in the right plot of Figure 6.4. Here, the predicted asymptotics is N^{-1}, indicated by the dashed line.

If we combine globally continuous, piecewise linear finite elements $\{V_j^{(2)}\}$ on Ω_2 with piecewise constant wavelets on Ω_1, we are in the situation that $\sqrt{n_1/n_2} = \sqrt{r_1/r_2} = \sqrt{(n_1 + r_1)/(n_2 + r_2)}$. Thus, the choice $\sigma = \sqrt{n_1/n_2}$ will lead to the best complexity, i.e., an essential convergence

ansatz functions	$\sigma = \sqrt{\frac{n_1}{n_2}}$	$\sigma = \sqrt{\frac{r_1+n_1}{r_2+n_2}}$	$\sigma = \sqrt{\frac{r_1}{r_2}}$
$r_1 = r_2 = 1$	$N^{-1/2}$	$N^{-9/20}$	$N^{-1/4}$
$r_1 = 1,\ r_2 = 2$	$N^{-1/2}$	$N^{-1/2}$	$N^{-1/2}$
$r_1 = 2,\ r_2 = 1$	$N^{-1/2}$	$N^{-3/8}$	$N^{-1/4}$
$r_1 = r_2 = 2$	$N^{-1/2}$	$N^{-3/7}$	$N^{-3/8}$

TABLE 6.1. The predicted rates of convergence.
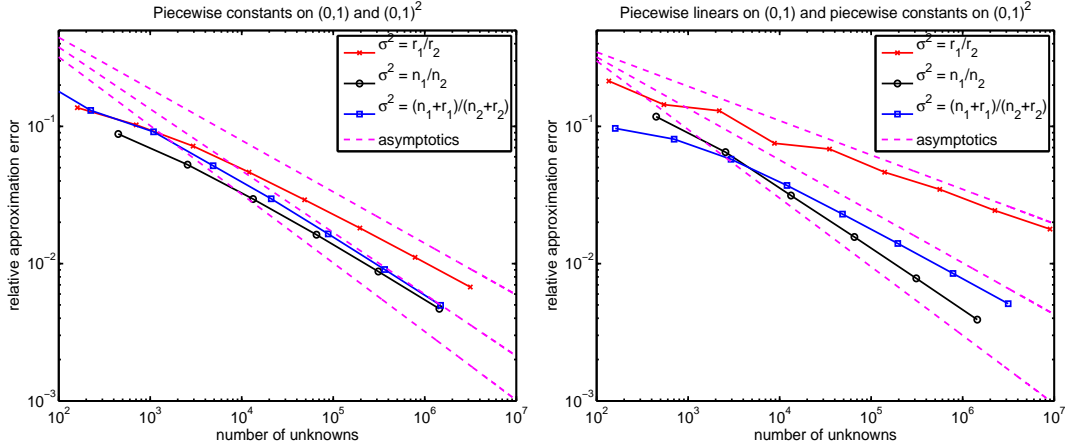
FIGURE 6.5. Cost complexities for approximating a non-smooth function in case of $n_1 = 1$, $n_2 = 2$ and $r_1 = r_2 = 1$ (left) or $r_1 = 2$, $r_2 = 1$ (right).

rate of N^{-1}. This is also seen from the left plot of Figure 6.4 where the approximation error is plotted versus the number of degrees of freedom in case of $\sigma = \sqrt{n_1/(n_2-1)}$, $\sigma = \sqrt{n_1/n_2}$, and $\sigma = \sqrt{n_1/(n_2+1)}$. Indeed, for $\sigma \neq \sqrt{n_1/n_2}$ a substantially smaller convergence rate is observed.

6.3. **Non-smooth functions.** We now investigate the approximation rates for the function

$$f(x, \mathbf{y}) = |\sin(\pi x) - y_1|.$$

One readily verifies that f is in $H^s(\Omega_1 \times \Omega_2)$ for all $s < 3/2$. We apply the same ansatz functions on Ω_1 and Ω_2 as in the previous subsection and measure the approximation error of the sparse grid interpolant for the three values $\sigma = \sqrt{r_1/r_2}$, $\sigma = \sqrt{n_1/n_2}$, and $\sigma = \sqrt{(r_1 + n_1)/(r_2 + n_2)}$. If we insert the limit case $p := 3/2$, Theorem 5.4 predicts the convergence rates given in Table 6.1.

The errors which were measured in our numerical calculations are plotted in Figures 6.5 and 6.6. The predicted convergence rates from Table 6.1 are visualized there as dashed lines. Figure 6.5 shows the case $r_1 = r_2 = 1$ (left plot) and $r_1 = 2$, $r_2 = 1$ (right plot), Figure 6.6 shows the case $r_1 = 1$, $r_2 = 2$ (left plot) and $r_1 = r_2 = 2$ (right plot). Equilibrating the number of degrees of freedom (that is, the choice $\sigma := \sqrt{n_1/n_2}$) gives in any case the best rate $N^{-1/2}$. Note that this is even the same rate as for smooth functions if we would there apply piecewise constant functions on Ω_1 and Ω_2. As it can be seen in Figures 6.5 and 6.6 this rate is essentially achieved. It is also seen that the choice $\sigma := \sqrt{r_1/r_2}$ gives always a substantially smaller convergence rate. For $\sigma := \sqrt{(r_1 + n_1)/(r_2 + n_2)}$ the rate is somewhere in between. However, its slope cannot be clearly seen from the plots except in the case $r_1 = 2$ and $r_2 = 1$. There, also the related asymptotics $N^{-3/8}$ is substantially different from $N^{-1/2}$.
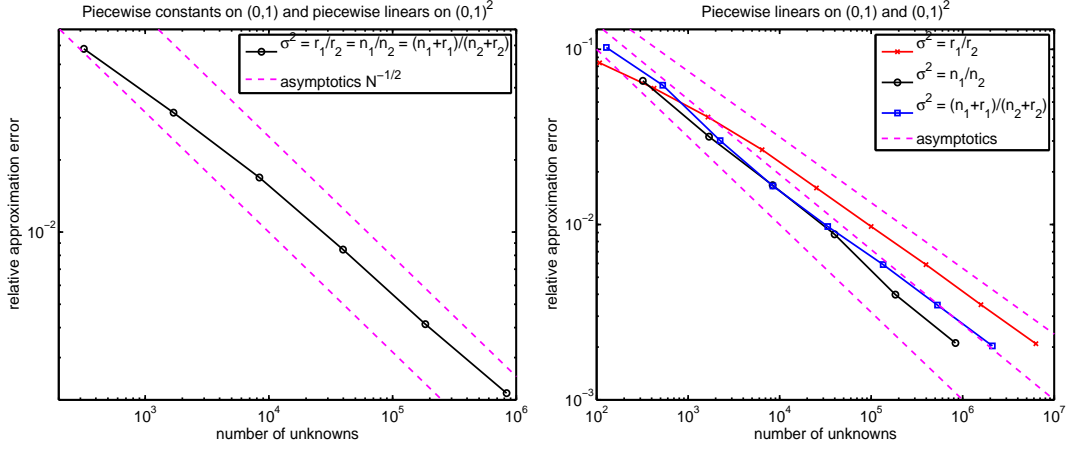
FIGURE 6.6. Cost complexities for approximating a non-smooth function in case of $n_1 = 1$, $n_2 = 2$ and $r_1 = 1$, $r_2 = 2$ (left) or $r_1 = r_2 = 2$ (right).

7. CONCLUDING REMARKS

So far, we only considered the case of the approximation error measured in the $L^2(\Omega_1 \times \Omega_2)$-norm, i.e., the $H_{mix}^{q_1;q_2}(\Omega_1 \times \Omega_2)$-norm with $q_1 = q_2 = 0$. However, certain applications involve other error norms which are of mixed type with more general q_1, q_2. For example, the two-point correlation of solutions to elliptic partial differential equations of second order with stochastic right hand side are in $H_{mix}^{1,1}(\Omega_1 \times \Omega_2)$ (cf. [22]). Moreover, radiative transfer (cf. [24]) and homogenization problems (cf. [16]) lead naturally to the $H_{mix}^{1,0}(\Omega_1 \times \Omega_2)$-norm. Therefore, we comment on the situation where the error is considered in the $H_{mix}^{q_1;q_2}(\Omega_1 \times \Omega_2)$-norm, with $0 \leq q_1 < s_1 \leq r_1, 0 \leq q_2 < s_2 \leq r_2$.[2] Here, we have of course to assume that q_1, q_2 are smaller than the regularities of the associated ansatz functions.

First, in analogy to (2.3), we assume the shifted approximation properties

$$\inf_{v_j \in V_j^{(i)}} \|u - v_j\|_{H^{q_i}(\Omega_i)} \lesssim 2^{-j(s_i - q_i)} \|u\|_{H^{s_i}(\Omega_i)}, \quad u \in H^{s_i}(\Omega_i)$$

to hold for $q_i \leq s_i \leq r_i$ uniformly in j. In particular, for any $f \in H^{s_i}(\Omega_i)$, this induces the shifted estimate

$$(7.24) \qquad \|Q_j^{(i)} f\|_{H^{q_i}(\Omega_i)} \lesssim 2^{-j(s_i - q_i)} \|f\|_{H^{s_i}(\Omega_i)}, \quad q_i \leq s_i \leq r_i.$$

This gives us straightforwardly the following generalization of Theorem 4.3:

Theorem 7.1. *Let $q_1 < s_1 \leq r_1$, $q_2 < s_2 \leq r_2$ and $f \in H_{mix}^{s_1;s_2}(\Omega_1 \times \Omega_2)$. Then, the approximation*

$$\widehat{f}_J = \sum_{j_1 \sigma + \frac{j_2}{\sigma} \leq J} \left(Q_{j_1}^{(1)} \otimes Q_{j_2}^{(2)} \right) f \in \widehat{V}_J^\sigma$$

satisfies

$$\|f - \widehat{f}_J\|_{H_{mix}^{q_1;q_2}(\Omega_1 \times \Omega_2)} \lesssim \begin{cases} 2^{-J \min\{(s_1 - q_1)/\sigma, (s_2 - q_2)\sigma\}} \|f\|_{H_{mix}^{s_1;s_2}(\Omega_1 \times \Omega_2)}, & \text{if } (s_1 - q_1)/\sigma \neq (s_2 - q_2)\sigma, \\ 2^{-J(s_1 - q_1)/\sigma} \sqrt{J} \|f\|_{H_{mix}^{s_1;s_2}(\Omega_1 \times \Omega_2)}, & \text{if } (s_1 - q_1)/\sigma = (s_2 - q_2)\sigma. \end{cases}$$

---

[2]The results can also be carried over to the case of negative q_i with obvious modifications and proper assumptions on the dual system, compare [11].

Proof. The proof follows by standard tensor product arguments from (7.24) in complete analogy to the lines of the proof of Theorem 4.3 with now just s_1, s_2 shifted to $s_1 - q_1$, $s_2 - q_2$ in the decay rates and clauses (but of course not in the smoothness assumption $\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$). $\square$

Also, Corollary 4.5 directly carries over. We obtain:

Corollary 7.2. *Let $q_1 < s_1 \leq r_1$, $q_2 < s_2 \leq r_2$ and $f \in H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)$. Furthermore, denote by $N := \dim \widehat{V}^\sigma_J$ the number of degrees of freedom in the sparse tensor product space $\widehat{V}^\sigma_J$ and set*

$$\alpha := \frac{\min\{(s_1 - q_1)/\sigma, (s_2 - q_2)\sigma\}}{\max\{n_1/\sigma, n_2\sigma\}}.$$

If $n_1/\sigma \neq n_2\sigma$, then the approximation (4.11) in $\widehat{V}^\sigma_J$ produces the following convergence rate in terms of the degrees of freedom N:

$$\|f - \widehat{f}_J\|_{H^{q_1,q_2}_{mix}(\Omega_1 \times \Omega_2)} \lesssim \begin{cases} N^{-\alpha}\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}, & \text{if } (s_1 - q_1)/\sigma \neq (s_2 - q_2)\sigma, \\ N^{-\alpha}\sqrt{\log N}\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}, & \text{if } (s_1 - q_1)/\sigma = (s_2 - q_2)\sigma. \end{cases}$$

If $n_1/\sigma = n_2\sigma$, we have

$$\|f - \widehat{f}_J\|_{H^{q_1,q_2}_{mix}(\Omega_1 \times \Omega_2)} \lesssim \begin{cases} N^{-\alpha}(\log N)^\alpha\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}, & \text{if } (s_1 - q_1)/\sigma \neq (s_2 - q_2)\sigma, \\ N^{-\alpha}(\log N)^{\alpha+\frac{1}{2}}\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}, & \text{if } (s_1 - q_1)/\sigma = (s_2 - q_2)\sigma. \end{cases}$$

Furthermore, the discussion of the results in Subsections 5.1 and 5.2 also carries directly over with the obvious substitutions $\|f - \widehat{f}_J\|_{H^{q_1,q_2}_{mix}(\Omega_1 \times \Omega_2)}$ for $\|f - \widehat{f}_J\|_{L^2(\Omega_1 \times \Omega_2)}$ and $r_1 - q_1, r_2 - q_2$ for r_1, r_2 or $s_1 - q_1, s_2 - q_2$ for s_1, s_2 (in the rates and case clauses but of course not in the smoothness prerequisites $\|f\|_{H^{r_1,r_2}_{mix}(\Omega_1 \times \Omega_2)}$ and $\|f\|_{H^{s_1,s_2}_{mix}(\Omega_1 \times \Omega_2)}$, respectively).

REFERENCES

[1] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Adv. Comput. Math.* **12** (2000) 273–288.

[2] H.J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica* **13** (2004) 147–269.

[3] N. Chegini and R. Stevenson. Adaptive wavelet schemes for parabolic problems. Sparse matrices and numerical results. *SIAM J. Numer. Anal.* **49** (2011) 182–212.

[4] D. Cioranescu, A. Damlamian, and G. Griso. The periodic unfolding method in homogenization. *SIAM J. Appl. Math.* **40** (2008) 1585–1620.

[5] W. Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numerica* **6** (1997) 55–228.

[6] J. Garcke, M. Griebel, and M. Thess. Data mining with sparse grids. *Computing* **67** (2001) 225–253.

[7] J. Garcke, M. Hegland, and O. Nielsen. Parallelisation of sparse grids for large scale data analysis. *ANZIAM J.* **48** (2006) 11–22.

[8] T. Gerstner and M. Griebel. Numerical integration using sparse grids. *Numer. Algorithms* **18** (1998) 209–232.

[9] T. Gerstner and M. Griebel. Dimension-adaptive tensor product quadrature. *Computing* **71** (2003) 65–87.

[10] M. Griebel and J. Hamaekers. Sparse grids for the Schrödinger equation. *M2AN Math. Model. Numer. Anal.* **41** (2007) 215–247.

[11] M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Math. Comput.* **78** (2009) 2223–2257.

[12] M. Griebel, P. Oswald, and T. Schiekofer. Sparse grids for boundary integral equations. *Numer. Math.* **83** (1999), 279–312.

[13] M. Griebel and D. Oeltz. A sparse grid space-time discretization scheme for parabolic problems. *Computing* **81** (2007) 1–34.

[14] H. Harbrecht. A finite element method for elliptic problems with stochastic input data. *Appl. Numer. Math.* **60** (2010) 227–244.

[15] H. Harbrecht, R. Schneider, and C. Schwab. Sparse second moment analysis for elliptic problems in stochastic domains. *Numer. Math.* **109** (2008) 167–188.

[16] V.H. Hoang and C. Schwab. High-dimensional finite elements for elliptic problems with multiple scales. *SIAM Multiscale Model. Simul.* **3** (2005) 168–194.

[17] A. Kolmogorov. Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse. *Annals Math.* **37** (1936) 107–110.

[18] C.C.W. Leentvaar and C.W. Oosterlee. On coordinate transformation and grid stretching for sparse grid pricing of basket options. *J. Comput. Appl. Math.* **222** (2008) 193–209.

[19] T. von Petersdorff and C. Schwab. Sparse wavelet methods for operator equations with stochastic data. *Applications of Mathematics* **51**(2) (2006) 145–180.

[20] D. Pflüger, B. Peherstorfer, and H.-J. Bungartz. Spatially adaptive sparse grids for high-dimensional data-driven problems. *Journal of Complexity* **26** (2010) 508–522.

[21] C. Reisinger and G. Wittum. Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM J. Sci. Comput.* **29** (2007), 440–458.

[22] C. Schwab and R.A. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.* **95** (2003) 707–734.

[23] R. Stevenson and C. Schwab. Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comput.* **78** (2009) 1293–1318.

[24] G. Widmer, R. Hiptmair and C. Schwab. Sparse adaptive finite elements for radiative transfer. *J. Comp. Phys.* **227** (2008) 6071-6105.

[25] C. Zenger. Sparse grids. In *Parallel algorithms for partial differential equations*, Proceedings of the 6th GAMM-Seminar, Kiel/Germany, 1990, Notes Numer. Fluid Mech. 31, 241–251, Vieweg, Braunschweig (1991).

Michael Griebel, Institut für Numerische Simulation, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany

E-mail address: `griebel@ins.uni-bonn.de`

Helmut Harbrecht, Mathematisches Institut, Rheinsprung 21, 4051 Basel, Switzerland

E-mail address: `helmut.harbrecht@unibas.ch`

Verzeichnis der erschienenen Preprints ab No. 475

475. Frehse, Jens; Löbach, Dominique: Improved Lp-Estimates for the Strain Velocities
      in Hardening Problems

476. Kurzke, Matthias; Melcher, Christof; Moser, Roger: Vortex Motion for the Landau-
      Lifshitz-Gilbert Equation with Spin Transfer Torque

477. Arguin, Louis-Pierre; Bovier, Anton; Kistler, Nicola: The Genealogy of Extremal
      Particles of Branching Brownian Motion

478. Bovier, Anton; Gayrard, Véronique: Convergence of Clock Processes in Random
      Environments and Ageing in the p-Spin SK Model

479. Bartels, Sören; Müller, Rüdiger: Error Control for the Approximation of Allen-Cahn
      and Cahn-Hilliard Equations with a Logarithmic Potential

480. Albeverio, Sergio; Kusuoka, Seiichiro: Diffusion Processes in Thin Tubes and their
      Limits on Graphs

481. Arguin, Louis-Pierre; Bovier, Anton; Kistler, Nicola: Poissonian Statistics in the Extremal
      Process of Branching Brownian Motion

482. Albeverio, Sergio; Pratsiovyta, Iryna; Torbin, Grygoriy: On the Probabilistic, Metric and
      Dimensional Theories of the Second Ostrogradsky Expansion

483. Bulíček, Miroslav; Frehse, Jens: C^{α}-Regularity for a Class of Non-Diagonal Elliptic Systems
      with p-Growth

484. Ferrari, Partik L.: From Interacting Particle Systems to Random Matrices

485. Ferrari, Partik L.; Frings, René: On the Partial Connection Between Random Matrices and
      Interacting Particle Systems

486. Scardia, Lucia; Zeppieri, Caterina Ida: Line-Tension Model as the Γ–Limit of a Nonlinear
      Dislocation Energy

487. Bolthausen, Erwin; Kistler, Nicola: A Quenched Large Deviation Principle and a Parisi
      Formula for a Perceptron Version of the Grem

488. Griebel, Michael; Harbrecht, Helmut: Approximation of Two-Variate Functions: Singular
      Value Decomposition Versus Regular Sparse Grids

489. Bartels, Sören; Kruzik, Martin: An Efficient Approach of the Numerical Solution of