Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Best N Term Approximation Spaces for Sparse Grids

P.-A. Nitsche

Research Report No. 2003-11
August 2003

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Best N Term Approximation Spaces for Sparse Grids

P.-A. Nitsche

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2003-11 August 2003

Abstract

We consider best N term approximation using anisotropic tensor product wavelet bases ('sparse grids'). We introduce a tensor product structure $\otimes_q$ on certain quasi-Banach spaces of ℓ_q-type, $q < 1$. We prove, that the approximation spaces $A^\alpha_q(L_2)$ and $A^\alpha_q(H^1)$ equal tensor products of Besov spaces $B^\alpha_q(L_q)$, e. g.

$$A^\alpha_q(L_2([0,1]^d)) = B^\alpha_q(L_q([0,1])) \otimes_q \ldots \otimes_q B^\alpha_q(L_q([0,1])).$$

1 Introduction

In this paper, we investigate the approximation spaces for best N term approximation using tensor product bases of one-dimensional hierarchical wavelet bases. These tensor product bases are anisotropic, i. e. high frequencies in some directions can be tensorized with low frequencies in other directions. Subsets of these bases span the so-called *sparse grid spaces*, which have been successfully applied to the numerical treatment of elliptic PDEs, to data mining or to high-dimensional integration.
Given a basis φ_k of a function space V with norm $\|\cdot\|_V$, a *best N term approximation* g_N to a function $f \in V$ realizes the following infimum

$$\inf\{\|f-g\|_V \,:\, g = \sum_{i=1}^{N} c_i \varphi_{k_i}\}.$$

The space, in which the approximation is seeked, is the nonlinear manifold consisting of all linear combinations of the given basis with at most N terms; for this reason best N term approximation is often called *nonlinear approximation*.

Best N term approximation is an important theoretical tool in the mathematical treatment of *adaptive numerical approximation*, since it yields upper bounds: if the sequence (g_N) of best N term approximations converges at a certain rate α, no adaptive scheme (using this particular basis) can do better. On the other hand, one should strive to construct adaptive approximation schemes which reproduce the rates achieved by best N term approximation.

Best N term approximation has been successfully employed in the mathematical analysis of adaptive wavelet methods for elliptic PDEs, see e. g. [5], [2], [6]. In [5], the authors construct an adaptive scheme for isotropically supported multi-dimensional wavelet bases which produces an approximation to a solution to an elliptic PDE at the 'optimal rate'. The term 'optimal rate' refers to the rate of best N term approximation using this particular basis. For instance, the alg orithm given in [5] yields an approximation rate in the H^1 norm of $1/2$ for solutions to smooth uniformly elliptic PDEs on two-dimensional polygonal domains using isotropically supported piecewise bilinear ansatz functions; the rate $1/2$ is optimal with respect to this isotropically supported wavelet basis.

However, it has been numerically observed (e. g. in the group of Zenger, see e. g. [1]), that adaptive approximation using *sparse grids spaces* can approximate certain classes of singularities arising in elliptic PDEs due to polyhedral domains at substantially higher rates. For the elliptic PDE from above on a two-dimensional polygonal domain, this yields approximation rates in the H^1 norm of $1-\varepsilon$ for arbitrarily small $\varepsilon > 0$, using piecewise bilinear ansatz functions as well.

This has been rigorously proved in an a priori wavelet context in [15]: solutions to elliptic PDEs (with smooth data) in polyhedral domains (in dimensions 2 and 3) can be approximated by sparse grid wavelet spaces (appropriately refined towards the singular support) built from biorthogonal spline wavelets of local polynomial degree p at any rate $< p+1$ with respect to the L_2-norm and at any rate $< p$ with respect to the H^1-norm.

It has already been known for several years, that sparse grid approximation can overcome the so-called curse of dimensionality (i. e. the exponential dependence of the approximation complexity on the dimension) for sufficiently regular functions, see e. g. [4], [13], [19]. For instance, an approximation with sparse grids built from ansatz functions which are piecewise polynomials of degree p gives a convergence rate in L_2 of $p+1$ (up to logarithmic terms), if the function to be approximated belongs to $H^{p+1,\dots,p+1}([0,1]^d)$ (Sobolev space of highest mixed derivatives on the d-dimensional cube $[0,1]^d$), whereas a full grid approximation would only give a rate of $(p+1)/d$ (but a regularity of $H^{p+1}([0,1]^d)$ would suffice).
Now it has been seen for the particular class of elliptic singularities, that when it comes to adaptive approximation, no additional assumptions on the equations (apart from the problem geometry) are necessary to give the substantially higher sparse grid rates.

The basis dependence of optimal approximation rates together with the numerical experience and our analysis of sparse approximation of singularities described above motivated the investigation of best N term approximation spaces for sparse grid wavelet spaces. To establish a characterization of sparse best N term approximation spaces in terms of classical smoothness spaces of Besov type is the purpose of the present paper.

We characterize the class of functions which can be approximated in L_2 or H^1 by anisotropic tensor product bases ('sparse grid bases') at a rate α. We prove that this class is a tensor product of appropriate one-dimensional Besov spaces. The spaces in question cease to be Banach spaces but are only quasi-Banach spaces satisfying only a generalized triangle inequality. A large part of this paper is therefore devoted to the construction of a tensor product structure on this type of quasi-Banach sp ace.

For the sparse grid spaces built from the Haar system, Oswald has already considered best N term approximation as well as approximability of certain singularity functions, see [16].

Acknowledgement: The author would like to thank Christoph Schwab for introducing him to this subject and lots of helpful discussions, as well as Radu-Alexandru Todor, who contributed the proof of lemma 1.

2 Notions from Approximation Theory

We assume familiarity with the basic concepts of linear and nonlinear approximation theory. An excellent survey on this topic is [11], see [18] as well. We only give here the notation used throughout this paper.

Let H be a separable Hilbert space with norm $\|\cdot\|_H$, and let $\{\psi_k : k \in I\} \subset H$ for some index set I be a basis of H, i.e. the completion of the linear span of $\{\psi_k : k \in I\}$ (with respect to $\|\cdot\|_H$) equals H:

$$\overline{\operatorname{span}\{\psi_k : k \in I\}} = H.$$

We denote the nonlinear manifolds, from which approximation takes place, by

$$\Sigma_N := \Big\{ \sum_{k\in\Lambda} c_k \psi_k : \Lambda \subset I, \#\Lambda \le N \Big\}.$$

The space Σ_N consists of all linear combinations of functions from the set $\{\psi_k : k \in I\}$ with at most N terms.

For a function $f \in H$, the approximation error $\sigma_N(f)_H$ is defined by

$$\sigma_N(f)_H := \inf_{S\in\Sigma_N} \|f - S\|_H.$$

For real $\alpha > 0$ and $0 < q < \infty$, the approximation space $A_q^\alpha(H)$ is defined by

$$A_q^\alpha(H) := \{f \in H : |f|_{A_q^\alpha(H)} < \infty\},$$

where

$$|f|_{A_q^\alpha(H)} := \left(\sum_{N\in\mathbb{N}} (N^\alpha \, \sigma_N(f)_H)^q \frac{1}{N} \right)^{1/q}.$$

We set $\|f\|_{A_q^\alpha(H)} = |f|_{A_q^\alpha(H)} + \|f\|_H$.

There hold the inclusions

$$A_q^\alpha(H) \subset A_p^\alpha(H), \quad 0 < q < p < \infty,$$

however, all the spaces $A_q^\alpha(H)$ correspond to an asymptotic decrease in error like $\mathcal{O}(N^{-\alpha})$.

Note, that the approximation spaces $A_q^\alpha(H)$ depend implicitely (but decisively) on the chosen basis $\{\psi_k \,:\, k \in I\}$.

One of the basic tasks in approximation theory is to characterize the approximation spaces $A_q^\alpha(H)$, ideally, by classical spaces like C^k, Sobolev or Besov spaces. One possibility is to prove the so called Jackson and Bernstein inequalities for some appropriate second space X:

Jackson inequality: $\sigma_N(f)_H \leq C\, N^{-r}\, \|f\|_X$ for all $f \in X$ and $N \in \mathbb{N}$.

Bernstein inequality: $\|S\|_X \leq C\, N^r\, \|S\|_H$ for all $S \in \Sigma_N$ and $N \in \mathbb{N}$.

Then one can characterize the approximation spaces as interpolation spaces (which are usually better understood), theorem 1 in [11]:

Theorem 1. *If the Jackson and Bernstein inequalities are valid for some appropriate space X, then for each $0 < \alpha < r$ and $0 < q < \infty$ the following relation holds between approximation and interpolation spaces:*

$$A_q^\alpha(H) = (H, X)_{\alpha/r,q}$$

(equivalent norms).

Here, $(H, X)_{\theta,q}$ are the so called real interpolation spaces, which consist of all functions f, for which

$$\left(\int_0^\infty (t^{-\theta}\, K(f,t))^q \frac{dt}{t} \right)^{1/q}$$

is finite, and

$$K(f,t) = \inf_{g \in X} \|f - g\|_H + t|g|_X$$

is the K functional.
For more details on interpolation spaces, see e.g. [3] or [17].

In the following, we will encounter Besov spaces, which play an important rôle in best N term approximation. We briefly recall their definition and basic relations. Let $\alpha > 0$ and $0 < p, q < \infty$. Let r be the smallest integer larger than α. Then a function f is in the Besov space $B_q^\alpha(L_p(\Omega))$ if

$$|f|_{B_q^\alpha(L_p(\Omega))} = \left(\int_0^\infty (t^{-\alpha}\, \omega_r(f,t)_p)^q \frac{dt}{t} \right)^{1/q} < \infty.$$

We set $\|f\|_{B_q^\alpha(L_p(\Omega))} = |f|_{B_q^\alpha(L_p(\Omega))} + \|f\|_{L_p(\Omega)}$.
Here, $\omega_r(f,t)_p = \sup_{|h|\leq t} \|\Delta_h^r(f,\,\cdot\,)\|_{L_p(\Omega)}$ is the r-th order modulus of continuity and Δ_h^r is the r-th power of the difference operator $\Delta_h f(x) = f(x+h) - f(x)$.

We have

$$(L_p(\Omega), W^r(L_p(\Omega)))_{\theta,q} = B_q^{\theta r}(L_p(\Omega)), \quad 0 < \theta < 1, 0 < q < \infty,$$

where $W^r(L_p(\Omega))$ is the Sobolev space of order r built on $L_p(\Omega)$.

We will almost exclusively be interested in the scale of Besov spaces $B_q^\alpha(L_q(\Omega))$ with $q^{-1} = \alpha + 1/2$. Interpolation on this scale yields again a space from this scale:

$$(L_2(\Omega), B_q^\alpha(L_q(\Omega))_{\theta,s} = B_s^{\theta\alpha}(L_s(\Omega)), \quad \text{if } s^{-1} = \theta\alpha + 1/2. \tag{1}$$

We will further need a result on best N term approximation in $\ell_2(\mathbb{N}^k)$ (theorem 4 in [11]):

Theorem 2. *For best N term approximation in $\ell_2(\mathbb{N}^k)$, a vector c is in $A_q^\alpha(\ell_2(\mathbb{N}^k))$ if and only if c is in the Lorentz sequence space $\ell_{\tau,q}$ with $\tau^{-1} = \alpha + 1/2$.*

The only Lorentz sequence spaces we will need are the spaces $\ell_{q,q}$ which coincide with ℓ_q.

3 Sparse Grid Basis

In this section we describe a class of bases for which our best N term approximation result will hold. These bases are tensor products of one-dimensional wavelet bases. The domain Ω under consideration will be the unit cube $[0,1]^d$.
Let $\{\psi_{jk} : j \in \mathbb{N}, k = 1, \ldots \Delta_j\}$, $\Delta_j \lesssim 2^j$, be a basis for $L^2([0,1])$ such that the following norm equivalences hold:

$$\begin{gathered} \|\sum_{j,k} c_{jk}\psi_{jk}\|_{B_q^\alpha(L_q([0,1]))} \sim \|(2^{j(\alpha - \frac{1}{q} + \frac{1}{2})} c_{jk})\|_{\ell_q(\mathbb{N})}, \\ 0 < \alpha < \alpha_0, \quad q \text{ in an open interval around } (\alpha + 1/2)^{-1} \end{gathered} \tag{2}$$

and

(3) $$\|\sum_{j,k} c_{jk}\psi_{jk}\|_{L_2([0,1])} \sim \|(c_{jk})\|_{\ell_2(\mathbb{N})},$$

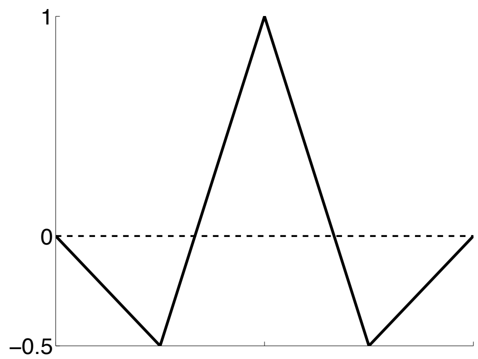
We will call such a basis for simplicity an α_0-basis.

Any wavelet system, for which (i) $\psi_{jk} \in B_q^\beta(L_q([0,1]))$ for some $\beta > \alpha$, (ii) ψ_{jk} has r vanishing moments with $r > \alpha$ and (iii) $r > 1/q - 1$ if $q < 1$, satisfies the norm equivalence

$$\|\sum_{j,k} c_{jk}\psi_{jk}\|_{B_q^\alpha(L_q([0,1]))} \sim \|(2^{j(\alpha-\frac{1}{q}+\frac{1}{2})}c_{jk})\|_{\ell_q(\mathbb{N})};$$

see e. g. [11] or [7].

So for instance the wavelet systems constructed in [9] or [10], which are biorthogonal spline wavelets of local degree p with $p+1$ vanishing moments are an α_0-basis for $\alpha_0 = p+1$. The following picture shows the interior wavelet of the simplest type $p = 1$. This wavelet hat two vanishing moments ($r = 2$), and it belongs to $B_q^\alpha(L_q)$ for (α, q) with $\alpha < 1 + 1/q$. Hence, the corresponding wavelet system constitutes an α_0-ba sis with $\alpha_0 = 2$.



In the case that α and q are related by $q^{-1} = \alpha + 1/2$, the norm equivalence (2) simply reads

(4) $$\|\sum_{j,k} c_{jk}\psi_{jk}\|_{B_q^\alpha(L_q([0,1]))} \sim \|(c_{jk})\|_{\ell_q(\mathbb{N})},$$
$$0 < \alpha < \alpha_0, \quad q = (\alpha + 1/2)^{-1}.$$

Note, that for $\alpha > 1/2$, the according value of q is between 0 and 1. In this case we have to deal with the quasi-Banach spaces ℓ_q (for which the triangle inequality fails to hold).

By tensorization of the one-dimensional basis $\{\psi_{jk}\}$ we get a basis for $L^2([0,1]^d)$:

$$\{\psi_{j_1 k_1 \dots j_d k_d} \, : \, j_i \in \mathbb{N}, k_i = 1, \dots \Delta_{j_i}\},$$

$$\psi_{j_1 k_1 \dots j_d k_d}(x) := \psi_{j_1 k_1} \otimes \dots \otimes \psi_{j_d k_d}(x) = \psi_{j_1 k_1}(x_1) \cdot \dots \cdot \psi_{j_d k_d}(x_d).$$

This basis is by construction anisotropic, that is, low frequencies in some directions can be paired with high frequencies in other directions. Certain finite subsets of this anisotropic tensor product basis are the so called sparse grid spaces

$$\hat{V}_L := \{\psi_{j_1 k_1 \dots j_d k_d} \, : \, j_1 + \dots + j_d \leq L, k_i = 1, \dots, \Delta_{j_i}\}.$$

We aim at describing the approximation spaces for best N term approximation in the multi-dimensional case by tensorization of the one-dimensional spaces. For this, we first have to declare a tensor product structure on pairs of the involved spaces $B_q^\alpha(L_q)$ and ℓ_q. This is done in the following section.

4 A Tensor Product Structure on certain Quasi-Banach Spaces

For the later development, we need a tensor product structure on pairs of $B_q^\alpha(L_q)$ and ℓ_q, respectively. If $q \geq 1$, we are in the realm of Banach spaces and such a structure is well known. We state this result for the spaces $\ell_q(\mathbb{N}^k)$ and $\ell_q(\mathbb{N})$:

Theorem 3. *Let $q \geq 1$. Then there is a tensor norm $\| \cdot \|_q$ on the algebraic tensor product*

$$\ell_q(\mathbb{N}^k) \otimes \ell_q(\mathbb{N}).$$

We denote the completion of $\ell_q(\mathbb{N}^k) \otimes \ell_q(\mathbb{N})$ with respect to this norm the tensor product $\ell_q(\mathbb{N}^k) \otimes_q \ell_q(\mathbb{N})$. This space is a Banach space, and the following isometry holds

$$\ell_q(\mathbb{N}^k) \otimes_q \ell_q(\mathbb{N}) \cong \ell_q(\mathbb{N}^{k+1}).$$

For a proof and more details, see e. g. [14].

However, in the case $q < 1$, such a structure seems not to be known. We will give a tensor product structure for a special class of quasi-Banach spaces including ℓ_q and $B_q^\alpha(L_q)$. First, we have to introduce some notation.

Definition 1. *Let X be a linear space. A function $\|\cdot\|_X : X \to \mathbb{R}$ is a q-quasi norm, if the following three properties hold:*
(a) $\|x\|_X = 0$ if and only if $x = 0$,
(b) $\|\alpha x\|_X = |\alpha|\,\|x\|_X$ for $x \in X$ and $\alpha \in \mathbb{R}$, and
(c) $\|x+y\|_X^q \leq \|x\|_X^q + \|y\|_X^q$ for $x, y \in X$. If X is complete with respect to the quasi metric induced by $\|\cdot\|_X$, we call X a q-quasi Banach space.

Examples: The spaces $\ell_q(\mathbb{N}^d)$, $L_q(\Omega)$ and $B_q^\alpha(L_q(\Omega))$ are q-quasi Banach spaces.

Remark 1. *A q-quasi Banach space possesses not necessarily a non-trivial topological dual. For instance, the topological dual of $L_q(\Omega)$ is trivial. However, the topological dual of $\ell_q(\mathbb{N}^d)$ is $\ell_\infty(\mathbb{N}^d)$.*

Definition 2. *A q-quasi Banach space X admits a q-estimate, if there is a set of functions $\{f_i\}$ spanning X, such that*

$$\|\sum_i c_i f_i\|_X \sim \|(c_i)\|_{\ell_q(\mathbb{N})}$$

for all convergent series $\sum_i c_i f_i$.

Examples: The spaces $B_q^\alpha(L_q([0,1]))$ admit q-estimates.

We further need some more ℓ_q-type spaces:
(a) $\ell_q(\mathbb{N}^d)$ is the space of all d-multi-indexed sequences

$$x = (x^{i_1\cdots i_d})_{(i_1,\ldots,i_d)\in\mathbb{N}^d},$$

for which

$$\|x\|_{\ell_q(\mathbb{N}^d)} := \left(\sum_{i_1,\ldots,i_d=1}^{\infty} |x^{i_1\cdots i_d}|^q\right)^{1/q}$$

is finite.
(b) $\ell_q(n)$ is the space of all n-vectors

$$x = (x^1, \ldots, x^n),$$

for which

$$\|x\|_{\ell_q(n)} := \left(\sum_{i=1}^{n} |x^i|^q \right)^{1/q}$$

is finite.
(c) $\ell_q(n,X)$ is the space of all X-valued n-vectors (X a q-quasi Banach space)

$$x = (x^1, \dots, x^n), \quad x^i \in X,$$

for which

$$\|x\|_{\ell_q(n,X)} := \left(\sum_{i=1}^{n} \|x^i\|_{X)}^q \right)^{1/q}$$

is finite.
(d) $\ell_q(\mathbb{N}^n, X)$ is the space of all X-valued n-multi-indexed sequences (X a q-quasi Banach space)

$$x = (x^{i_1 \cdots i_n})_{(i_1,\dots,i_n)\in\mathbb{N}^n}, \quad x^{i_1,\dots,i_n} \in X,$$

for which

$$\|x\|_{\ell_q(\mathbb{N}^n,X)} := \left(\sum_{i_1,\dots,i_n=1}^{\infty} \|x^{i_1 \cdots i_n}\|_X^q \right)^{1/q}$$

is finite.

Now we have the notation to declare a topological tensor product structure on the algebraic tensor products between spaces of ℓ_q- or $B_q^\alpha(L_q)$-type. This is done in the following theorem:

Theorem 4. *Let X and Y be q-quasi Banach spaces, $0 < q < 1$, admitting a q-estimate. Then the function*

$$\|\cdot\| : X \otimes Y \to \mathbb{R},$$

$$\|z\| := \inf_{z=\sum_{i=1}^{n} x_i \otimes y_i} \|(x_i)\|_{\ell_q(n,X)} \sup_{\|\lambda\|_{\ell_q(n)} \le 1} \|\sum_{i=1}^{n} \lambda_i \, y_i\|_Y,$$

on the algebraic tensor product of X and Y is a q-quasi norm.
We denote the completion of the algebraic tensor product $X \otimes Y$ under the induced quasi metric by $X \otimes_q Y$ and call it the q-tensor product of X and Y.
In case of $X = \ell_q(\mathbb{N}^d))$ and $Y = \ell_q(\mathbb{N})$, the function $\|\cdot\|$ is a cross norm, i.e. for simple tensors $x \otimes y$ it holds $\|x \otimes y\| = \|x\|_{\ell_q(\mathbb{N}^d)} \|y\|_{\ell_q(\mathbb{N})}$. Furthermore, we have an isometric isomorphism

$$\ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}) \cong \ell_q(\mathbb{N}^{d+1}).$$

We will need the following lemma:

Lemma 1. *Let X be a q-quasi Banach space admitting a q-estimate. Then, given x, x_1, $\ldots$, $x_n \in X$, there exists a functional ϕ (not necessarily continuous on all of X), such that*

$$\phi(x) = \|x\|_X,$$
$$\text{and } \sum_{i=1}^{n} |\phi(x)|^q \leq C \sum_{i=1}^{n} \|x_i\|_X^q,$$

where C depends only on the space X but not on the vectors x, x_i.
In case of $X = \ell_q(\mathbb{N}^d)$, the constant C can be chosen to be 1.

Remark 2. *The existence of an algebraic functional ϕ as in lemma 1 can be regarded as kind of a substitute for the Hahn-Banach extension theorem. It would be interesting to know, if lemma 1 can be derived solely from the properties of a q-quasi norm without the assumption of admitting a q-estimate (and therefore boiling down the case to an ℓ_q-problem).*

Proof of lemma 1. We begin with the case $X = \ell_q(\mathbb{N}^d)$. We identify the topological dual of $\ell_q(\mathbb{N}^d)$ in the standard way with $\ell_\infty(\mathbb{N}^d)$.
Let $e_{j_1\cdots j_d}$ be the evaluation functional of the $(j_1, \ldots, j_d)$-th component:

$$e_{j_1\cdots j_d}(y) = y^{j_1\cdots j_d}, \quad y = (y^{i_1\cdots i_d})_{(i_1,\ldots,i_d)\in\mathbb{N}^d} \in \ell_q(\mathbb{N}^d).$$

Now set

$$C_{j_1\cdots j_d} := \left(\frac{\sum_{i=1}^{n} \|x_i\|_{\ell_q(\mathbb{N}^d)}^q}{\sum_{i=1}^{n} |x_i^{j_1\cdots j_d}|^q} \right)^{1/q}$$

and consider the set of functionals

$$A = \{\phi_{j_1\cdots j_d} = (\operatorname{sgn} x^{j_1\cdots j_d})\, C_{j_1\cdots j_d}\, e_{j_1\cdots j_d} \,:\, (j_1, \ldots, j_d) \in \mathbb{N}^d\} \subset (\ell_q(\mathbb{N}^d))'.$$

Note, that

$$\phi \in A \quad \Rightarrow \quad \sum_{i=1}^{n} |\phi(x_i)|^q = \sum_{i=1}^{n} \|x_i\|_{\ell_q(\mathbb{N}^d)}^q.$$

We claim: there is $\phi \in A$ such that $\phi(x) \geq \|x\|_{\ell_q(\mathbb{N}^d)}$. Downscaling then yields the assertion of the Lemma.
Observe that

$$\sum_{j_1,\dots,j_d=1}^{\infty} C_j^{-q} = \sum_{j_1,\dots j_d=1}^{\infty} \frac{\sum_{i=1}^n |x_i^{j_1\cdots j_d}|^q}{\sum_{i=1}^n \|x_i\|_{\ell_q(\mathbb{N}^d)}^q} = 1.$$

Now assume that there is no $\phi \in A$ with $\phi(x) \geq \|x\|_{\ell_q(\mathbb{N}^d)}$. This implies

$$C_{j_1\cdots j_d} < \frac{\|x\|_{\ell_q(\mathbb{N}^d)}}{|x^{j_1\cdots j_d}|} \quad \text{for all } (j_1,\dots j_d) \in \mathbb{N}^d.$$

Summing up gives

$$\sum_{j_1,\dots,j_d=1}^{\infty} C_j^{-q} > 1;$$

this is a contradiction.

In the general case, we use the q-estimate to transfer the problem to ℓ_q: write $x = \sum_k c_k f_k$ and $x_i = \sum_k c_k^i f_k$. Then choose $\tilde{\phi} \in \ell_q(\mathbb{N})'$ such that $\tilde{\phi}((c_k)) = \|(c_k)\|_{\ell_q(\mathbb{N})}$ and $\sum_{i=1}^n |\tilde{\phi}((c_k^i)_k)|^q \leq \sum_{i=1}^n \|(c_k^i)_k\|_{\ell_q(\mathbb{N})}^q$. Employing the q-estimate and rescaling of $\tilde{\phi}$ yields the assertion with C the product of the lower and upper constant in the q-estimate. $\square$

Proof of theorem 4. We first note, that by definition of the (algebraic) tensor product, we have

$$\sum_{i=1}^{n} x_i \otimes y_i = \sum_{i=1}^{\bar{n}} \bar{x}_i \otimes \bar{y}_i, \quad x_i, \bar{x}_i \in X, y_i, \bar{y}_i \in Y,$$

if and only if

$$\text{for all } \phi \in X^* : \sum_{i=1}^{n} \phi(x_i) y_i = \sum_{i=1}^{\bar{n}} \phi(\bar{x}_i) \bar{y}_i.$$

Here we denote by X^* the (non-trivial) algebraic dual of the q-quasi Banach space X.

We begin with the first part of the statement of theorem 4.
We have to show properties (a), (b), (c) of the definition of a q-quasi norm.
Property (b), i. e. $\|\alpha z\| = |\alpha| \, \|z\|$ for $z \in X \otimes Y$, $\alpha \in \mathbb{R}$, readily follows, since the involved quasi norms are homogeneous.
Property (a), i. e. $\|z\| = 0$ if and only if $z = 0$:
If $z = 0 = 0 \otimes 0$, we have $\|z\| = 0$. To the contrary, assume $z \neq 0$. Let $z = \sum_{i=1}^{\bar{n}} \bar{x}_i \otimes \bar{y}_i$ be a

representation of z with linearly independent sets $(\bar{x}_i)$, $(\bar{y}_i)$. (Such a linearly independent representation always exists for every algebraic tensor product.) Fix one of the pairs $(\bar{x}_i, \bar{y}_i)$, say $(\bar{x}_1, \bar{y}_1)$, and set

$$\varepsilon := \min\{\|y\|_Y \ : \ y = \|\bar{x}_1\|_X \, \bar{y}_1 + \sum_{i=2}^{\bar{n}} \alpha_i \bar{y}_i\} > 0.$$

Now choose according to lemma 1 a functional ϕ with

$$\phi(\bar{x}_1) = \|\bar{x}_1\|_X \quad \text{and} \quad \sum_{i=1}^{\bar{n}} |\phi(\bar{x}_i)|^q \le C \sum_{i=1}^{\bar{n}} \|\bar{x}_i\|_X^q.$$

Let an arbitrary representation $\sum_{i=1}^n x_i \otimes y_i$ of z be given. Then we estimate

$$\|\sum_{i=1}^n \phi(x_i) y_i\|_Y^q \le \sum_{i=1}^n |\phi(x_i)|^q \, \|y_i\|_Y^q \le C \sum_{i=1}^n \|x_i\|_X^q \sup_{i=1,\ldots,n} \|y_i\|_Y^q$$

$$\le C \, \|(x_i)\|_{\ell_q(n,X)}^q \sup_{\|\lambda\|_{\ell_q(n)} \le 1} \|\sum_{i=1}^n \lambda_i y_i\|_Y^q.$$

On the other hand, we have

$$\|\sum_{i=1}^n \phi(x_i) y_i\|_Y^q = \|\sum_{i=1}^{\bar{n}} \phi(\bar{x}_i) \bar{y}_i\|_X^q$$

$$= \Big\| \|\bar{x}_1\|_X \bar{y}_1 + \sum_{i=2}^{\bar{n}} \phi(\bar{x}_i) \bar{y}_i \Big\|_X^q \ge \varepsilon^q.$$

Combining this gives

$$\varepsilon \, C^{-1/q} \le \|(x_i)\|_{\ell_q(n,X)} \sup_{\|\lambda\|_{\ell_q(n)} \le 1} \|\sum_{i=1}^n \lambda_i y_i\|_Y$$

for every representation of z. Taking the infimum over all representations yields $\|z\| \ge \varepsilon \, C^{-1/q} > 0$.

Property (c), i. e. the generalized triangle inequality $\|z + w\|^q \le \|z\|^q + \|w\|^q$:
We show first

(5) $$\|\sum_{i=1}^n \mu_i y_i\|_Y^q \le \|\mu\|_{\ell_q(n)}^q \sup_{\|\lambda\|_{\ell_q(n)} \le 1} \|\sum_{i=1}^n \lambda_i y_i\|_Y^q.$$

For this purpose set $\theta := \|\mu\|_{\ell_q(n)}^{-1}$, such that $\|\theta\mu\|_{\ell_q(n)} = 1$. Then we have

$$\|\sum_{i=1}^n \mu_i y_i\|_Y^q = \theta^{-q} \|\sum_{i=1}^n \theta\mu_i y_i\|_Y^q = \|\mu\|_{\ell_q(n)}^q \|\sum_{i=1}^n \theta\mu_i y_i\|_Y^q$$

$$\leq \|\mu\|_{\ell_q(n)}^q \sup_{\|\lambda\|_{\ell_q(n)}\leq 1} \|\sum_{i=1}^n \lambda_i y_i\|_Y^q.$$

Next we show, that for $m < n$

$$\sup_{\|\lambda\|_{\ell_q(m)}\leq 1} \|\sum_{i=1}^m \lambda_i y_i\|_Y^q \leq 1 \quad \text{and} \quad \sup_{\|\lambda\|_{\ell_q(n-m)}\leq 1} \|\sum_{i=m+1}^n \lambda_{i-m} y_i\|_Y^q \leq 1 \tag{6}$$

implies

$$\sup_{\|\lambda\|_{\ell_q(n)}\leq 1} \|\sum_{i=1}^n \lambda_i y_i\|_Y^q \leq 1. \tag{7}$$

For this let $\lambda \in \ell_q(n)$ with $\|\lambda\|_{\ell_q(n)} \leq 1$. We write $\lambda = \mu + \sigma$ with

$$\mu = (\lambda_1, \dots, \lambda_m, 0, \dots, 0) \quad \text{and} \quad \sigma = (0, \dots, 0, \lambda_{m+1}, \dots, \lambda_n).$$

Then it holds $\|\mu\|_{\ell_q(n)}^q + \|\sigma\|_{\ell_q(n)}^q = \|\lambda\|_{\ell_q(n)}^q \leq 1$. Using (5), we get

$$\|\sum_{i=1}^n \lambda_i y_i\|_Y^q \leq \|\sum_{i=1}^m \lambda_i y_i\|_Y^q + \|\sum_{i=m+1}^n \lambda_i y_i\|_Y^q$$

$$\leq \|\mu\|_{\ell_q(m)}^q \sup_{\|\kappa\|_{\ell_q(m)}\leq 1} \|\sum_{i=1}^m \kappa_i y_i\|_Y^q + \|\sigma\|_{\ell_q(n-m)}^q \sup_{\|\kappa\|_{\ell_q(n-m)}\leq 1} \|\sum_{i=m+1}^n \kappa_{i-m} y_i\|_Y^q.$$

Using (6) and $\|\mu\|_{\ell_q(n)}^q + \|\sigma\|_{\ell_q(n)}^q = \|\lambda\|_{\ell_q(n)}^q \leq 1$, it follows

$$\|\sum_{i=1}^n \lambda_i y_i\|_Y^q \leq 1.$$

Taking the supremum over all λ with $\|\lambda\|_{\ell_q(n)} \leq 1$ gives (7).

Now let $z, w \in X \otimes Y$, and let $\varepsilon > 0$. Choose a representation $z = \sum_{i=1}^m x_i \otimes y_i$ with

$$\|(x_i)\|_{\ell_q(m,X)} \sup_{\|\lambda\|_{\ell_q(m)}\leq 1} \|\sum_{i=1}^m \lambda_i y_i\|_Y \leq (\|z\|^q + \varepsilon)^{1/q}.$$

Du to the homogeneity of the tensor product, we can assume without loss of generality

$$\|(x_i)\|_{\ell_q(m,X)}^q \leq \|z\|^q + \varepsilon, \qquad \sup_{\|\lambda\|_{\ell_q(m)}\leq 1} \|\sum_{i=1}^m \lambda_i y_i\|_Y^q \leq 1.$$

Analogously, we choose a representation $w = \sum_{i=m+1}^{n} x_i \otimes y_i$ with

$$\|(x_{m+i})\|_{\ell_q(n-m,X)}^q \leq \|w\|^q + \varepsilon, \qquad \sup_{\|\lambda\|_{\ell_q(n-m)}\leq 1} \|\sum_{i=m+1}^{n} \lambda_{i-m} y_i\|_Y^q \leq 1.$$

Then with (7)

$$\begin{aligned}
&\|(x_i)\|_{\ell_q(n,X)}^q \sup_{\|\lambda\|_{\ell_q(n)}\leq 1} \|\sum_{i=1}^{n} \lambda_i y_i\|_Y^q \\
&\leq \left(\|(x_i)\|_{\ell_q(m,X)}^q + \|(x_{m+i})\|_{\ell_q(n-m,X)}^q\right) \cdot 1 \\
&\leq \|z\|^q + \|w\|^q + 2\varepsilon.
\end{aligned}$$

Since $z + w = \sum_{i=1}^{n} x_i \otimes y_i$ is a representation of $z + w$, we infer

$$\|z + w\|^q \leq \|z\|^q + \|w\|^q + 2\varepsilon.$$

Sending $\varepsilon \to 0$, we arrive at the generalized triangle inequality.
This proves the first part of theorem 4.

For the second part, we first show that the q-quasi norm on $\ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N})$ is a crossnorm, i.e. for simple tensors $x \otimes y$ it holds $\|x \otimes y\| = \|x\|_{\ell_q(\mathbb{N}^d)} \|y\|_{\ell_q(\mathbb{N})}$:
Let $x \otimes y = \sum_{i=1}^{n} x_i \otimes y_i$ be any representation of $x \otimes y$. According to lemma 1 choose a functional ϕ such that

$$\phi(x) = \|x\|_{\ell_q(\mathbb{N}^d)} \quad \text{and} \quad \sum_{i=1}^{n} |\phi(x_i)|^q \leq \sum_{i=1}^{n} \|x_i\|_{\ell_q(\mathbb{N}^d)}^q.$$

Now set $M = \left(\sum_{i=1}^{n} |\phi(x_i)|^q\right)^{1/q}$ and $\lambda_i = \phi(x_i)/M$. Then, $\|\lambda\|_{\ell_q(n)} = 1$.
We have

$$\begin{aligned}
\|(x_i)\|_{\ell_q(n,\ell_q(\mathbb{N}^d))} \sup_{\|\mu\|_{\ell_q(n)}=1} \|\sum_{i=1}^{n} \mu_i y_i\|_{\ell_q(\mathbb{N})} &\geq \|(x_i)\|_{\ell_q(n,\ell_q(\mathbb{N}^d))} \|\sum_{i=1}^{n} \lambda_i y_i\|_{\ell_q(\mathbb{N})} \\
&= \left(\sum_{i=1}^{n} \|x_i\|_{\ell_q(\mathbb{N}^d)}^q\right)^{1/q} \cdot \frac{1}{M} \|\sum_{i=1}^{n} \phi(x_i) y_i\|_{\ell_q(\mathbb{N})} \\
&= \left(\frac{\sum_{i=1}^{n} \|x_i\|_{\ell_q(\mathbb{N}^d)}^q}{\sum_{i=1}^{n} |\phi(x_i)|^q}\right)^{1/q} \cdot \|\phi(x) y\|_{\ell_q(\mathbb{N})} \geq \|x\|_{\ell_q(\mathbb{N}^d)} \|y\|_{\ell_q(\mathbb{N})}.
\end{aligned}$$

Taking the infimum over all representations of $x \otimes y$ yields

$$\|x \otimes y\| \geq \|x\|_{\ell_q(\mathbb{N}^d)} \|y\|_{\ell_q(\mathbb{N})}.$$

The reverse inequality $\|x \otimes y\| \leq \|x\|_{\ell_q(\mathbb{N}^d)} \|y\|_{\ell_q(\mathbb{N})}$ is trivial by definition of $\| \cdot \|$.

Finally we show the isometry $\ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}) \cong \ell_q(\mathbb{N}^d, \ell_q(\mathbb{N}))$. The remaining isometry $\ell_q(\mathbb{N}^d, \ell_q(\mathbb{N})) \cong \ell_q(\mathbb{N}^{d+1})$ is standard.

Let us first consider the mapping

$$\Lambda : \ell_q(\mathbb{N}^d) \otimes \ell_q(\mathbb{N}) \to \ell_q(\mathbb{N}^d, \ell_q(\mathbb{N})),$$
$$z = \sum_{i=1}^n x_i \otimes y_i \mapsto \left[f : (j_1, \dots, j_d) \mapsto \sum_{i=1}^n x_i^{j_1 \cdots j_d} y_i \right].$$

The mapping Λ is well defined. To see this, take two representations

$$z = \sum_{i=1}^n x_i \otimes y_i = \sum_{i=1}^{\bar{n}} \bar{x}_i \otimes \bar{y}_i \in \ell_q(\mathbb{N}^d) \otimes \ell_q(\mathbb{N})$$

and denote

$$f : (j_1, \dots, j_d) \mapsto \sum_{i=1}^n x_i^{j_1 \cdots j_d} y_i, \quad \bar{f} : (j_1, \dots, j_d) \mapsto \sum_{i=1}^{\bar{n}} \bar{x}_i^{j_1 \cdots j_d} \bar{y}_i.$$

Let $e_{j_1 \cdots j_d}$ be the evaluation functional of the $(j_1, \dots, j_d)$-th component:

$$e_{j_1 \cdots j_d}(y) = y^{j_1 \cdots j_d}, \quad y = (y^{i_1 \cdots i_d})_{(i_1, \dots, i_d) \in \mathbb{N}^d} \in \ell_q(\mathbb{N}^d).$$

Then we have

$$f((j_1, \dots, j_d)) = \sum_{i=1}^n x_i^{j_1 \cdots j_d} y_i = \sum_{i=1}^n e_{j_1 \cdots j_d}(x_i) y_i$$
$$= \sum_{i=1}^{\bar{n}} e_{j_1 \cdots j_d}(\bar{x}_i) \bar{y}_i = \sum_{i=1}^{\bar{n}} \bar{x}_i^{j_1 \cdots j_d} \bar{y}_i = \bar{f}((j_1, \dots, j_d));$$

hence $f = \bar{f}$.

Now we show, that Λ is of norm 1: Let $z \in \ell_q(\mathbb{N}^d) \otimes \ell_q(\mathbb{N})$ and $f = \Lambda(z)$. For an arbitrary representation $z = \sum_{i=1}^n x_i \otimes y_i$, we have

$$\|f\|^q_{\ell_q(\mathbb{N}^d, \ell_q(\mathbb{N}))} = \sum_{j_1, \dots, j_d}^{\infty} \|\sum_{i=1}^n x_i^{j_1 \cdots j_d} y_i\|^q_{\ell_q(\mathbb{N})} \leq \sum_{j_1, \dots, j_d}^{\infty} \sum_{i=1}^n |x_i^{j_1 \cdots j_d}|^q \|y_i\|^q_{\ell_q(\mathbb{N})}$$

$$= \sum_{i=1}^{n} \|x_i\|^q_{\ell_q(\mathbb{N}^d)} \|y_i\|^q_{\ell_q(\mathbb{N})} \leq \sum_{i=1}^{n} \|x_i\|^q_{\ell_q(\mathbb{N}^d)} \max_{i=1,\ldots,n} \|y_i\|^q_{\ell_q(\mathbb{N})}$$

$$\leq \|(x_i)\|^q_{\ell_q(n,\ell_q(\mathbb{N}^d))} \left(\sup_{\|\lambda\|_{\ell_q(n)}\leq 1} \|\sum_{i=1}^{n} \lambda_i y_i\|_{\ell_q(\mathbb{N})} \right)^q .$$

Taking the infimum over all representations on the right, we arrive at

$$\|f\|^q_{\ell_q(\mathbb{N}^d,\ell_q(\mathbb{N}))} \leq \|z\|^q.$$

Thus, Λ has norm ≤ 1. To see the equality, take a simple tensor $x \otimes y \in \ell_q(\mathbb{N}^d) \otimes \ell_q(\mathbb{N})$. Then we have

$$\|f\|^q_{\ell_q(\mathbb{N}^d,\ell_q(\mathbb{N}))} = \sum_{j_1,\ldots,j_d}^{\infty} \|x^{j_1\cdots j_d} y\|^q_{\ell_q(\mathbb{N})} = \sum_{j_1,\ldots,j_d}^{\infty} |x^{j_1\cdots j_d}|^q \, \|y\|^q_{\ell_q(\mathbb{N})}$$

$$= \|x\|^q_{\ell_q(\mathbb{N}^d)} \|y\|^q_{\ell_q(\mathbb{N})} = \|x \otimes y\|.$$

Thus, Λ has norm 1 and gives by continuous extension to $\ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N})$ an operator

$$\tilde{\Lambda} \,:\, \ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}) \to \ell_q(\mathbb{N}^d, \ell_q(\mathbb{N}))$$

of norm 1.

Next, consider the mapping

$$\tilde{\Gamma} \,:\, \ell_q(\mathbb{N}^d, \ell_q(\mathbb{N})) \to \ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}),$$

which is the continuous extension of the mapping Γ defined on simple functions

$$f \,:\, \{1,\ldots,n\}^d \to \ell_q(\mathbb{N}), \quad (j_1,\ldots,j_d) \mapsto c^{j_1\cdots j_d} \in \ell_q(\mathbb{N}).$$

With such a simple function, we associate

$$z = \Gamma(f) = \sum_{(j_1,\ldots,j_d)\in\{1,\ldots,n\}^d} \underbrace{\chi^{j_1\cdots j_d} \|c^{j_1\cdots j_d}\|_{\ell_q(\mathbb{N})}}_{x^{j_1\cdots j_l}} \otimes \underbrace{\frac{c^{j_1\cdots j_d}}{\|c^{j_1\cdots j_d}\|_{\ell_q(\mathbb{N})}}}_{y^{j_1\cdots j_d}},$$

where $\chi^{j_1\cdots j_d}$ denotes the characteristic function of the multi-index $(j_1,\ldots,j_d)$.

We show, that Γ is of norm 1:

$$\|z\|^q \le \left(\sum_{\{1,\dots,n\}^d} \|c^{j_1\cdots j_d}\|^q_{\ell_q(\mathbb{N})} \right) \underbrace{\sup_{\|\lambda\|_{\ell_q(n^d)}\le 1} \| \sum_{\{1,\dots,n\}^d} \lambda_i \frac{c^{j_1\cdots j_d}}{\|c^{j_1\cdots j_d}\|_{\ell_q(\mathbb{N})}} \|^q}_{\le 1}$$

$$\le \sum_{\{1,\dots,n\}^d} \|c^{j_1\cdots j_d}\|^q_{\ell_q(\mathbb{N})} = \|f\|^q_{\ell_q(\mathbb{N}^d,\ell_q(\mathbb{N}))}.$$

Choosing $f = \chi^{j_1\cdots j_d} c^{j_1\cdots j_d}$ with $c^{j_1\cdots j_d} \in \ell_q(\mathbb{N})$ and norm 1, we see $\|z\| = 1 = \|f\|$. Hence, Γ as well as the continuous extension $\tilde{\Gamma}$ are of norm 1.

Now consider $\tilde{\Gamma}\tilde{\Lambda} : \ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}) \to \ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N})$. This mapping has norm ≤ 1. We show, that for $z = \sum_{i=1}^n x_i y_i$ with $\operatorname{supp}(x_i) \subset \{1,\dots,N\}^d$, this mapping is the identity:

$$\tilde{\Gamma}\tilde{\Lambda}\left(\sum_{i=1}^n x_i y_i\right) = \tilde{\Gamma}\left((j_1,\dots,j_d) \mapsto \sum_{i=1}^n x_i^{j_1\cdots j_d} y_i\right)$$

$$= \sum_{\{1,\dots,N\}^d} \chi^{j_1\cdots j_d} \| \sum_{i=1}^n x_i^{j_1\cdots j_d} y_i \|_{\ell_q(\mathbb{N})} \otimes \frac{\sum_{i=1}^n x_i^{j_1\cdots j_d} y_i}{\|\sum_{i=1}^n x_i^{j_1\cdots j_d} y_i\|_{\ell_q(\mathbb{N})}}$$

$$= \sum_{\{1,\dots,N\}^d} \chi^{j_1\cdots j_d} \otimes \sum_{i=1}^n x_i^{j_1\cdots j_d} y_i = \sum_{i=1}^n x_i \otimes y_i.$$

This establishes $\ell_q(\mathbb{N}^d) \otimes_q \ell_q(\mathbb{N}) \cong \ell_q(\mathbb{N}^d, \ell_q(\mathbb{N}))$. □

5 Approximation in L_2

We are now ready to describe the approximation spaces A_q^α for best N term approximation in $L_2([0,1]^d)$ using the anisotropic ('sparse grid') tensor product basis $\{\psi_{j_1k_1\dots j_dk_d}\}$.

Definition 3. *For $d \ge 2$ and $q^{-1} = \alpha + 1/2$, we define*

$$\hat{B}_q^\alpha([0,1]^d) := B_q^\alpha(L_q([0,1])) \otimes_q \dots \otimes_q B_q^\alpha(L_q([0,1])).$$

Using theorem 4 (respectively theorem 3 for the case $q \ge 1$) and the norm equivalences (4), we get

Lemma 2. *For $\alpha < \alpha_0$ and $q^{-1} = \alpha + 1/2$, we have*

$$\| \sum_{j_i,k_i} c_{j_1k_1,\dots,j_d,k_d} \psi_{j_1k_1,\dots,j_dk_d} \|_{\hat{B}_q^\alpha([0,1]^d)} \sim \|(c_{j_1k_1,\dots,j_dk_d})\|_{\ell_q(\mathbb{N}^d)}.$$

We now prove the Bernstein and Jackson inequalities for the spaces $\hat{B}^\alpha_q([0,1]^d)$.

Lemma 3 (Bernstein inequality)**.** *For all $S \in \Sigma_N$, we have*

$$\|S\|_{\hat{B}^\alpha_q([0,1]^d)} \lesssim N^\alpha \, \|S\|_{L_2([0,1]^d)}.$$

Proof. Let $S = \sum_{i=1}^N c_i \psi_i$, where $\psi_i \in \{\psi_{j_1 k_1} \otimes \cdots \otimes \psi_{j_d k_d}\}$. Using lemma 2 as well as the mutual equivalence between finitely supported ℓ_q norms,

$$\|x\|_{\ell_p(N)} \leq \|x\|_{\ell_q(N)} \leq N^{1/q-1/p} \, \|x\|_{\ell_p(N)}, \quad x \in \mathbb{R}^N, \quad 0 < q \leq p \leq \infty$$

we get (using (3))

$$\|S\|_{\hat{B}^\alpha_q([0,1]^d)} \sim \|(c_i)\|_{\ell_q(N)} \lesssim N^{1/q-1/2} \|(c_i)\|_{\ell_2(\mathbb{N})} \lesssim N^\alpha \, \|S\|_{L_2([0,1]^d)}.$$

□

Lemma 4 (Jackson inequality)**.** *For all $f \in \hat{B}^\alpha_q([0,1]^d)$, we have*

$$\inf_{S \in \Sigma_N} \|f - S\|_{L_2([0,1])} \lesssim N^{-\alpha} \, \|f\|_{\hat{B}^\alpha_q([0,1]^d)}.$$

Proof. Since $f \in \hat{B}^\alpha_q([0,1]^d)$, we have $c \in \ell_q(\mathbb{N}^d)$ for the coefficient vector of the wavelet decomposition. Using (3) and theorem 2, we infer with $\tau = q$ and $\ell_{q,q} = \ell_q$

$$\inf_{S \in \Sigma_N} \|f - S\|_{L_2([0,1])} \lesssim N^{-\alpha} \, \|c\|_{\ell_q(\mathbb{N}^d)} \lesssim N^{-\alpha} \, \|f\|_{\hat{B}^\alpha_q([0,1]^d)}.$$

□

Eventually, we can apply theorem 1 to characterize the approximation spaces:

Theorem 5. *The approximation space $A^\alpha_q(L_2([0,1]^d))$, $q^{-1} = \alpha + 1/2$, $0 < \alpha < \alpha_0$, corresponding to an approximation rate α in $L_2([0,1]^d)$ by best N term approximation using the anisotropic ('sparse grid') tensor product wavelet basis $\{\psi_{j_1 k_1 \ldots j_d k_d}\}$ is given by*

$$A^\alpha_q(L_2([0,1]^d)) = \hat{B}^\alpha_q([0,1]^d).$$

Proof. Theorem 1 together with lemmata 3 and 4 yields

$$A^\alpha_q(L_2([0,1]^d)) = (L_2([0,1]^d), \hat{B}^\beta_s([0,1]^d))_{\alpha/\beta,q}, \quad s^{-1} = \beta + 1/2,$$

for β with $\alpha < \beta < \alpha_0$. From the isomorphism to ℓ_q-spaces and the corresponding interpolation result,

$$(\ell_2, \ell_s)_{\theta,q} = \ell_q,$$

for

$$\theta = \frac{1/q - 1/2}{1/s - 1/2} = \frac{\alpha}{\beta},$$

it follows that

$$(L_2([0,1]^d), \hat{B}_s^\beta([0,1]^d))_{\alpha/\beta,q} = \hat{B}_q^\alpha([0,1]^d).$$

□

6 Approximation in H^1

In this section we treat best N term approximation with respect to the Sobolev $H^1([0,1]^d)$-norm. The methods are similiar to the L^2-case. Therefore, we won't go too much into detail here.

Note, that we can decompose the space $H^1([0,1]^2)$ into an intersection of tensor products of one-dimensional spaces:

$$H^1([0,1]^2) \cong \left(H^1([0,1]) \otimes_2 L_2([0,1])\right) \cap \left(L_2([0,1]) \otimes_2 H^1([0,1])\right),$$

and generally (denoting $L_2([0,1])$ by $H^0([0,1])$),

$$H^1([0,1]^d) \cong \bigcap_{k=1}^d \left(\bigotimes_{i=1}^d H^{\delta_{ik}}([0,1]) \right).$$

Hence, it suffices to treat the case $H^1([0,1]) \otimes_2 L_2([0,1]) \otimes_2 \ldots \otimes_2 L_2([0,1])$.

Renormalizing the wavelet basis ψ_{jk} in the first variable by mulitplying with a factor 2^{-j}, we get a basis normalized in $H^1([0,1])$ satisfying the norm equivalence

$$\| \sum_{j,k} c_{jk}(2^{-j}\psi_{jk}) \|_{H^1([0,1])} \sim \|(c_{jk})\|_{\ell_2(\mathbb{N})},$$

which will be the substitute for (3); see e. g. [9], [10].

Analogously, the Besov norm equivalences (4) read for $\alpha < \alpha_0$

$$\| \sum_{j,k} c_{jk}(2^{-j}\psi_{jk}) \|_{B_q^\alpha(L_q([0,1]))} \sim \|(c_{jk})\|_{\ell_q(\mathbb{N})}, \quad q^{-1} = \alpha - 1/2.$$

The spaces corresponding to the remaining variables keep unchanged. Hence, we have norm equivalences

$$\|\sum_{j_i,k_i} c_{j_1k_1\dots j_dk_d}\psi_{j_1k_1\dots j_dk_d}\|_{H^1([0,1])\otimes_2 L_2([0,1])\otimes_2\dots\otimes_2 L_2([0,1])} \sim \|(c)\|_{\ell_2(\mathbb{N}^d)}$$

as well as

$$\|\sum_{j_i,k_i} c_{j_1k_1\dots j_dk_d}\psi_{j_1k_1\dots j_dk_d}\|_X \sim \|(c)\|_{\ell_q(\mathbb{N}^d)}$$

for the space

$$X = B_q^{1/q+1/2}(L_q([0,1])) \otimes_q B_q^{1/q-1/2}(L_q([0,1])) \otimes_q \dots \otimes_q B_q^{1/q-1/2}(L_q([0,1])).$$

Now the Jackson and Bernstein inequalities imply a characterization of the aproximation spaces by interpolation spaces. We skip further details (intersection etc.) and formulate the result:

Theorem 6. *The approximation space $A_q^\alpha(H^1([0,1]^d))$, $q^{-1} = \alpha + 1/2$, $0 < \alpha < \alpha_0 - 1$, corresponding to an approximation rate α in $H^1([0,1]^d)$ by best N term approximation using the anisotropic ('sparse grid') tensor product wavelet basis $\{\psi_{j_1k_1\dots j_dk_d}\}$ is given by*

$$A_q^\alpha(H^1([0,1]^d)) = \tilde{B}_q^\alpha([0,1]^d),$$

where the space $\tilde{B}_q^\alpha([0,1]^d)$ is defined as

$$\tilde{B}_q^\alpha([0,1]^d) = \bigcap_{k=1}^d \left(\bigotimes_{i=1}^d X_{\delta_{ik}}([0,1]) \right)$$

(q-tensor product) with

$$X_0([0,1]) = B_q^\alpha(L_q([0,1]))$$

and

$$X_1([0,1]) = B_q^{\alpha+1}(L_q([0,1])).$$

Proof. As before, theorem 1 yields

$$A_q^\alpha(H^1([0,1]^d)) = (H^1([0,1]^d), \tilde{B}_s^\beta([0,1]^d))_{\alpha/\beta,q}, \quad s^{-1} = \beta + 1/2,$$

for β with $\alpha < \beta < \alpha_0 - 1$. From the isomorphism to ℓ_q-spaces and the corresponding interpolation result, it follows that

$$(H^1([0,1]^d), \tilde{B}_s^\beta([0,1]^d))_{\alpha/\beta,q} = \tilde{B}_q^\alpha([0,1]^d).$$

□

Remark 3. *We have shown in [15], that solutions to elliptic PDEs (with smooth data) in polyhedral domains (in dimensions 2 and 3) can be approximated by sparse grid wavelet spaces (appropriately refined towards the singular support) built from biorthogonal spline wavelets of local polynomial degree* p *at any rate* $< p+1$ *with respect to the* L_2*-norm and at any rate* $< p$ *with respect to the* H^1*-norm. This readily implies that corner and edge singularities in solutions to elliptic PDEs in polyhedral doma ins (at least in dimensions 2 and 3) belong to the Besov scales* $\hat{B}^{\alpha}_{q}([0,1]^d)$ *and* $\tilde{B}^{\alpha}_{q}([0,1]^d)$*.*

Remark 4. *The examples of approximation in* $L_2([0,1]^d)$ *and* $H^1([0,1]^d)$ *have been chosen since they are of interest to the numerical analyst. Clearly, any appropriate scale of spaces admitting isomorphisms to* ℓ_q*-spaces gives rise to according approximation results.*

References

[1] S. Achatz: Adaptive finite Dünngitter-Elemente höherer Ordnung für elliptische partielle Differentialgleichungen mit variablen Koeffizienten, Dissertation, TU München, 2003

[2] A. Barinka, T. Barsch, Ph. Charton, A. Cohen, S. Dahlke, W. Dahmen, K. Urban: Adaptive Wavelet Schemes for Elliptic Problems – Implementation and Numerical Experiments; SIAM J. Sci. Comput. 23, no. 3, pp. 910-939 (2001)

[3] C. Bennett, R. Sharpley: Interpolation of Operators; Academic Press, New York (1988)

[4] H. Bungartz, M. Griebel: A Note on the Complexity of Solving Poisson's Equation for Spaces of Bounded Mixed Derivatives; J. Compl. 15, pp. 167-199 (1999)

[5] A. Cohen, W. Dahmen, R. Devore: Adaptive Wavelet Methods for Elliptic Operator Equations: Convergence Rates; Math. Comput. 70, no. 233, pp. 27-75 (2000)

[6] A. Cohen, W. Dahmen, R. Devore: Adaptive Wavelet Schemes for Nonlinear Variational Problems; to appear in SIAM J. Num. Math. (2003)

[7] A. Cohen, R. A. Devore, R. Hochmuth: Restricted Nonlinear Approximation; Constr. Approx. 16, pp. 85-113 (2000)

[8] W. Dahmen: Wavelet Methods for PDEs – some recent developments; J. Comput. Appl. Math. 128, pp. 133-185 (2001)

[9] W. Dahmen, A. Kunoth, K. Urban: Biorthogonal Spline-Wavelets on the Interval – Stability and Moment Conditions; Appl. Comput. Harm. Anal. 6, pp. 132-196 (1999)

[10] W. Dahmen, R. Schneider: Wavelets with Complementary Boundary Conditions – Function Spaces on the Cube; Result. Math. 34, pp. 255-293 (1998)

[11] R. A. Devore: Nonlinear approximation; Acta Numerica 7, pp. 51-150, Cambridge University Press, Cambridge (1998)

[12] R. A. Devore, B. Jawerth, V. Popov: Compression of Wavelet Decompositions; Amer. J. Math. 114, pp. 737-785 (1992)

[13] M. Griebel, S. Knapek: Optimized tensor-product approximation spaces; Constr. Approx. 16 (4), pp. 525-540 (2000)

[14] W. A. Light, E. W. Cheney: Approximation Theory in Tensor Product Spaces; Lecture Notes in Mathematics 1169, Springer-Verlag (1980)

[15] P.-A. Nitsche: Sparse approximation of singularity functions; Research Report No. 2002-18, ETH Zürich (2002)
http://www.sam.math.ethz.ch/reports/details/include.shtml?2002/2002-18.html

[16] P. Oswald: On N-term approximations in the Haar system in H^s-norms; Metric theory of functions and related problems in analysis (Russian), 137–163, Izd. Nauchno-Issled. Aktuarno-Finans. Tsentra (AFTs), Moscow (1999)

[17] J. Peetre: A Theory of Interpolation of Normed Spaces; Course notes, University of Brasilia

[18] V. N. Temlyakov: Nonlinear Methods of Approximation; Found. Comput. Math. 3, pp. 33-107 (2002)

[19] Ch. Zenger: Sparse Grids; in Parallel Algorithms for PDE, Proc. 6th GAMM Seminar, Kiel (ed. W. Hackbusch), Vieweg, Braunschweig, pp. 241-251 (1991)

Research Reports

No.	Authors	Title
03-11	P.-A. Nitsche	Best N Term Approximation Spaces for Sparse Grids
03-10	J.T. Becerra Sagredo	Z-splines: Moment Conserving Cardinal Spline Interpolation of Compact Support for Arbitrarily Spaced Data
03-09	P. Houston, D. Schötzau, Th. Wihler	Energy norm a-posteriori error estimation for mixed discontinuous Galerkin approximations of the Stokes problem
03-08	R. Hiptmair, A. Buffa	Coercive combined field integral equations
03-07	R. Hiptmair, O. Sterz	Current and Voltage Excitations for the Eddy Current Model
03-06	A.-M. Matache, P.-A. Nitsche, C. Schwab	Wavelet Galerkin Pricing of American Options on Lévy Driven Assets
03-05	M. Becheanu, R.A. Todor	On the Set of Diameters of Finite Point-Sets in the Plane
03-04	C. Schwab, R.A. Todor	Sparse finite elements for stochastic elliptic problems - higher order moments
03-03	R. Sperb	Bounds for the first eigenvalue of the elastically supported membrane on convex domains
03-02	F.M. Buchmann	Computing exit times with the Euler scheme
03-01	A. Toselli, X. Vasseur	Domain decomposition preconditioners of Neumann-Neumann type for *hp*-approximations on boundary layer meshes in three dimensions
02-26	M. Savelieva	Theoretical study of axisymmetrical triple flame
02-25	D. Schötzau, C. Schwab, A. Toselli	Mixed *hp*-DGFEM for incompressible flows III: Pressure stabilization
02-24	F.M. Buchmann, W.P. Petersen	A stochastically generated preconditioner for stable matrices
02-23	A.W. Rüegg, A. Schneebeli, R. Lauper	Generalized *hp*-FEM for Lattice Structures
02-22	L. Filippini, A. Toselli	*hp* Finite Element Approximations on Non-Matching Grids for the Stokes Problem
02-21	D. Schötzau, C. Schwab, A. Toselli	Mixed *hp*-DGFEM for incompressible flows II: Geometric edge meshes
02-20	A. Toselli, X. Vasseur	A numerical study on Neumann-Neumann and FETI methods for *hp*-approximations on geometrically refined boundary layer meshes in two dimensions