Industrial Mathematics Institute

2009:03

Anisotropic function spaces with applications

S. Dekel and P. Petrushev

IMI
Preprint Series

Department of Mathematics
University of South Carolina

Anisotropic function spaces with applications

Shai Dekel and Pencho Petrushev

Abstract In this survey we review the recently developed theory of anisotropic spaces and representations of functions based on anisotropic multilevel ellipsoid covers (dilations) of $\mathbb{R}^n$. We also exhibit the relations of the ellipsoid cover approach to earlier concepts of anisotropic structures as well as to the framework of general spaces of homogeneous type. A number of open problems are presented and discussed.

1 Introduction

Anisotropic phenomena appear in various contexts in mathematical analysis and its applications. For instance, functions are frequently very smooth on subdomains of $\mathbb{R}^n$ separated by smooth curves or manofolds, where they have jump discontinuities. This sort of singularities reduce significantly the classical smoothness of the functions and create problems when attempting to find sparse representations of them.

One perhaps useful approach to resolving the singularities of functions along smooth curves and manifolds (and more general singular behaviors) is the utilization of the framework of anisotropic multiscale ellipsoid covers (dilations) of $\mathbb{R}^n$ which may change rapidly from point to point at any level and in depth. The second important element of this concept is to use anisotropic ellipsoid covers adaptively by allowing them to adjust to the singularities of the function under question. Other critical issues are related, in particular, to the anisotropic representation of functions and definition and characterization of the respective anisotropic smoothness spaces.

Shai Dekel
GE Healthcare, 27 Hamaskit St., Herzelia 46733, Israel, e-mail: Shai.Dekel@ge.com

Pencho Petrushev
Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA, e-mail: pencho@math.sc.edu

The purpose of this survey paper is to review the main concepts and problems of this relatively new undertaking, documented so far in [11, 12, 14]. Although we will have some answers to reveal to some of the important questions, there will be plenty of open problems presented as well.

This theory has three main components with the first being the structure of the underlying **ellipsoid covers** of $\mathbb{R}^n$. The main role here is played by discrete multi-level ellipsoid covers of $\mathbb{R}^n$ of the form: $\Theta = \cup_{m\in\mathbb{Z}}\Theta_m$, where each Θ_m consists of ellipsoids of volume $\sim 2^{-a_0 j}$ which cover $\mathbb{R}^n$ and any ellipsoids $\theta_1, \theta_2 \in \Theta_m$ with $\theta_1 \cap \theta_2 \neq \emptyset$ have similar shapes and orientations. In depth the behavior of the ellipsoids is similar, namely, if $\theta_1 \in \Theta_m$, $\theta_2 \in \Theta_{m+1}$ and $\theta_1 \cap \theta_2 \neq \emptyset$, then θ_1 and θ_2 are similar in shape and orientation. An important feature of the set of all ellipsoid covers of $\mathbb{R}^n$ is that it is invariant under affine transforms. Another important issue is that any ellipsoid cover of $\mathbb{R}^n$ generates a quasi-distance, which coupled with the Lebesgue measure transforms $\mathbb{R}^n$ into a homogeneous type space. The properties of anisotropic covers are explored in [12]. A short description of them is given in §2, where we also compare ellipsoid covers of $\mathbb{R}^2$ with the so called multilavel strong local regular (SLR) triangulations of $\mathbb{R}^2$, introduced in [20].

The **anisotropic elements (building blocks)** introduced in [12] and the related representations of functions is the second component of our theory. A sequence of bases $\{\Phi_m\}_{m\in\mathbb{Z}}$ is naturally associated to each discrete ellipsoid cover $\Theta = \cup_{m\in\mathbb{Z}}\Theta_m$. Here each Φ_m consists of C^∞ functions which are supported on the ellipsoids in Θ_m, reproduce polynomials of degree $< k$ and are locally linear independent. The key property of these bases is that each Φ_m is a stable basis in L_p for $0 < p \leq \infty$. This allows to define local projectors into the spaces $S_m = \mathrm{span}(\Phi_m)$ which preserve polynomials of degree $< k$. In turn, these maps induce a sequence of two-level-split bases which provide representation of functions and are aligned with the underlying anisotropic structure in $\mathbb{R}^n$. As is shown in [12] these representations also allow to characterize the anisotropic Besov spaces of positive smoothness. The next step is to define smooth (global) duals to $\{\Phi_m\}$ and thereby to construct kernels $\{S_m\}$ which reproduce polynomials of a certain degree in both variables. This enabled us to deploy the machinery of homogeneous spaces to the construction of continuous and discrete anisotropic wavelet frames. All these constructions and results are presented in §3.

The third component of the theory we review here consists of **anisotropic spaces** associated with anisotropic ellipsoid covers of $\mathbb{R}^n$. The anisotropic homogeneous ($\dot{B}^{\alpha}_{pq}(\Theta)$) and inhomogeneous ($B^{\alpha}_{pq}(\Theta)$) Besov spaces (B-spaces) of positive smoothness are developed in [12] and briefly introduced in §4. In the same section we compare them with the anisotropic B-spaces induced by multilevel SLR-triangulations of $\mathbb{R}^2$ and with classical Besov spaces. In §5 we show that, in analogy to the classical case, certain B-spaces naturally occur in nonlinear N-term approximation from the two-level-split bases. In §6 we advance the idea of using adaptively anisotropic B-space for measuring the smoothness of the functions, which is closely related to the problem for sparse representation of functions. The development of anisotropic Triebel-Lizorkin of an arbitrary smoothness is the grand open problem in this theory. The key is to construct anisotropic frames with well localized elements and prescribed vanishing moments which are faithfully aligned with the underlying anisotropic ellipsoid cover.

Candès and Donoho (e.g. [5, 6]) have developed the so called *curvlets*, which provide an alternative scheme for resolving singularities of functions along smooth curves in $\mathbb{R}^2$. The advantage of curvlets over our approach is that the curvlets form a frame, while our scheme is adaptive, and hence curvlets are easier to implement. On the other hand, the curvlet frame is overly redundant. More precisely at every location and scale there are numerous directional elements with various orientations (the number of orientations increases with the scale). Curvlets are purely L_2-creatures which rely on fine cancelations and are unusable for decomposition of functions in L_p, $p \neq 2$.

Yet another approach to resolving singularities of functions along smooth curves is developed in [1, 15] and is based on the so called *Adaptive Geometric Wavelets*. This method is closely related to the schemes employing ellipsoid covers and nested triangulations considered here; it proved to be very effective in image compression.

In the final Section 7 the two-level-split bases and the machinery of Besov spaces are applied in a regular set-up to the development of meshless multilevel Schwarz preconditioners for elliptic boundary value problems. The details of this development are given in [11], which was the starting point of this work.

Throughout we will use $|E|$ to denote the Lebesgue measure of $E \subset \mathbb{R}^n$; we will denote by c, c_1, c_2, etc. positive constants which may vary at every appearance. The equivalence $a \sim b$ means $c_1 a \leq b \leq c_2 a$.

2 Anisotropic multiscale structures on $\mathbb{R}^n$

In this article we are mainly concerned with anisotropic structures on $\mathbb{R}^n$ induced by anisotropic ellipsoid covers (dilations) of $\mathbb{R}^n$ and the related function spaces . For comparison we will also briefly review the anisotropic multilevel nested triangulations of $\mathbb{R}^2$.

2.1 Anisotropic multilevel ellipsoid covers (dilations) of $\mathbb{R}^n$

We denote by $B(x,r)$ the Euclidean ball in $\mathbb{R}^n$ of radius r centered at x. The image of the unit ball $B^* := B(0,1)$ in $\mathbb{R}^n$ via an affine transform will be called an *ellipsoid*.

Definition 1. We call

$$\Theta = \bigcup_{m\in\mathbb{Z}} \Theta_m$$

a *discrete multilevel ellipsoid cover* of $\mathbb{R}^n$ if the following conditions are obeyed, where $a_0, \ldots, a_8$, and N_1 are positive constants:

(a) Every level Θ_m $(m \in \mathbb{Z})$ consists of ellipsoids θ such that

$$a_1 2^{-a_0 m} \le |\theta| \le a_2 2^{-a_0 m} \tag{1}$$

and Θ_m is a cover of $\mathbb{R}^n$, i.e. $\mathbb{R}^n = \bigcup_{\theta\in\Theta_m}\theta$.

(b) For $\theta \in \Theta$ let A_θ be an affine transform of the form

$$A_\theta(x) = M_\theta x + v_\theta, \quad M_\theta \in \mathbb{R}^{n\times n},$$

such that $\theta = A_\theta(B^*)$ and $v_\theta := A(0)$ is the center of θ. We postulate that for any $\theta \in \Theta_m$ and $\theta' \in \Theta_{m+\nu}$ $(m \in \mathbb{Z}, \nu \ge 0)$ with $\theta \cap \theta' \ne \emptyset$, we have

$$a_3 2^{-a_4 \nu} \le 1/\|M_{\theta'}^{-1} M_\theta\|_{\ell_2\to\ell_2} \le \|M_\theta^{-1} M_{\theta'}\|_{\ell_2\to\ell_2} \le a_5 2^{-a_6\nu}. \tag{2}$$

(c) Each ellipsoid $\theta \in \Theta_m$ can be intersected by at most N_1 ellipsoids from Θ_m.

(d) For every $x \in \mathbb{R}^n$ and $m \in \mathbb{Z}$ there exists $\theta \in \Theta_m$ such that $x \in \theta^\diamond$, where $\theta^\diamond$ is the dilated version of θ by a factor of $a_7 < 1$, i.e. $\theta^\diamond = A_\theta(B(0,a_7))$.

(e) If $\theta \cap \eta \ne \emptyset$ with $\theta \in \Theta_m$ and $\eta \in \Theta_m \cup \Theta_{m+1}$, then $|\theta\cap\eta| > a_8|\eta|$.

We will denote by $\mathbf{p}(\Theta) := \{a_0, a_1, \dots, a_8, N_1\}$ the set of all parameters in the above definition.

Several clarifying remarks are in order.

1. It is crucial that the set of all discrete ellipsoid covers of $\mathbb{R}^n$ is invariant under affine transforms. More precisely, the images $A(\theta)$ of all ellipsoids $\theta \in \Theta$ of a given cover Θ of $\mathbb{R}^n$ via an affine transform A of the form $A(x) = Mx + v$ with $|\det M| = 1$ form an ellipsoid cover of $\mathbb{R}^n$ with the same parameters as Θ.
2. Condition (b) above indicates that if $\theta \cap \theta' \ne \emptyset$, then the ellipsoids θ and θ' are similar in shape and orientation when they are from close levels. In particular, if $M := M_\theta^{-1} M_{\theta'}$ and $M = UDV$ is the singular value decomposition of M, where U and V are orthogonal matrices, and $D = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ is diagonal with $\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_n > 0$, then

$$\|M\|_{\ell_2\to\ell_2} = \sigma_1 \quad \text{and} \quad \|M_{\theta'}^{-1} M_\theta\|_{\ell_2\to\ell_2} = \|M^{-1}\|_{\ell_2\to\ell_2} = 1/\sigma_n.$$

   Therefore, condition (b) is equivalently expressed as

$$a_3 2^{-a_4\nu} \le \sigma_n \le \dots \le \sigma_1 \le a_5 2^{-a_6\nu}. \tag{3}$$

   This condition has a clear geometric interpretation: The affine transform A_θ^{-1}, which maps the ellipsoid θ onto the unit ball B^*, maps the ellipsoid θ' onto an ellipsoid with semi-axes $\sigma_1, \sigma_2, \dots, \sigma_n$ satisfying (3).
3. Condition (e) may seem restrictive, but this is not the case. As is shown in [12] if Θ is a discrete ellipsoid cover satisfying conditions $(a)-(d)$ above, then there exists a discrete ellipsoid cover $\widetilde{\Theta}$ of $\mathbb{R}^n$ which obeys conditions $(a)-(e)$ (with possibly different constants a_1 and a_7) obtained by dilating every ellipsoid $\theta \in \Theta$ by a factor r_θ satisfying $(a_7+1)/2 \le r_\theta \le 1$.

Continuous and semi-continuous ellipsoid covers. Discrete ellipsoid covers of $\mathbb{R}^n$ are easy to derive from semi-continuous or continuous covers, which are in general easier to construct.

In the case of a *semi-continuous ellipsoid cover* $\Theta = \cup_{m\in\mathbb{Z}}\Theta_m$, an ellipsoid $\theta(v,m) \in \Theta_m$ is associated to every $v \in \mathbb{R}^n$ and $m \in \mathbb{Z}$ such that

$$a_1 2^{-a_0 m} \le |\theta(v,m)| \le a_2 2^{-a_0 m},$$

which replaces (1) and the respective affine transforms satisfy a condition similar to (2); conditions (c)-(e) are void.

In the case of a *continuous ellipsoid cover* $\Theta := \cup_{t\in\mathbb{R}}\Theta_t$, an ellipsoid $\theta(v,t) \in \Theta_t$ is associated to every $v \in \mathbb{R}^n$ and $t \in \mathbb{R}$ such that

$$a_1 2^{-a_0 t} \le |\theta(v,t)| \le a_2 2^{-a_0 t},$$

i.e. the scale is continuous as well. For more detail and the exact definitions of ellipsoid covers, see §2.2 in [12]

Examples. (i) The one parameter family of diagonal dilation matrices

$$D_t = \operatorname{diag}(2^{-tb_1}, 2^{-tb_2}, \dots, 2^{-tb_n}), \quad b_j > 0,\ j = 1,\dots,n,$$

apparently induces a continuous ellipsoid cover of $\mathbb{R}^n$.

(ii) Suppose A is an $n\times n$ real matrix with eigenvalues λ satisfying $|\lambda| > 1$. Then it is easy to see that the affine transforms $A_{v,m}(x) := A^{-m}x + v$, $v \in \mathbb{R}^n$, $m \in \mathbb{Z}$, define a semi-continuous ellipsoid cover (dilations) of $\mathbb{R}^n$. This particular kind of dilations are used in [2, 3, 4] for the development of anisotropic Hardy, Besov, and Triebel-Lizorkin spaces.

(iii) The continuous covers used in Section 6 (see also §7 in [12]) are nontrivial examples of anisotropic ellipsoid covers of $\mathbb{R}^2$, where the ellipsoids change rapidly from point to point and in depth.

The point is that, on the one hand, continuous and semi-continuous covers are easier to construct and, on the other, given a semi-continuous or continuous cover one can construct a discrete ellipsoid cover with essentially the same (equivalent) metric (see [12]).

Quasi-distance. A quasi-distance is naturally associated with any discrete, semi-continuous or continuous ellipsoid covers of $\mathbb{R}^n$. Recall that a *quasi-distance* on a set $X \neq \emptyset$ is a map $\rho : X\times X \to [0,\infty)$ satisfying the conditions:

(a) $\rho(x,y) = 0 \Longleftrightarrow x = y$,
(b) $\rho(y,x) = \rho(x,y)$,
(c) $\rho(x,z) \le \kappa(\rho(x,y) + \rho(y,z))$.

Here $\kappa \ge 1$ is a constant.

Definition 2. Assuming that Θ is a continuous, semi-continuous or discrete ellipsoid cover of $\mathbb{R}^n$, we define $\rho : \mathbb{R}^n \times \mathbb{R}^n \to [0,\infty)$ by

$$\rho(x,y) := \inf\{|\theta| : \theta \in \Theta \text{ and } x,y \in \theta\}. \tag{4}$$

Proposition 1. [12] *For any ellipsoid cover Θ of $\mathbb{R}^n$ the map $\rho : \mathbb{R}^n \times \mathbb{R}^n \to [0,\infty)$ defined above is a quasi-distance on $\mathbb{R}^n$.*

Spaces of homogeneous type were first introduced in [8] (see also [9, 16]) as a means for extending the Calderón-Zygmund theory of singular integral operators to more general settings. Let X be a topological space endowed with a Borel measure μ and a quasi-distance $\rho(\cdot,\cdot)$. Assume that the balls $B_\rho(x,r) := \{y \in X : \rho(x,y) < r\}$, $x \in X$, $r > 0$, form a basis for the topology in X. The space (X,ρ,μ) is said to be of *homogenous type* if there exists a constant $A > 0$ such that for all $x \in X$ and $r > 0$,

$$\mu(B_\rho(x,2r)) \le A\mu(B_\rho(x,r)). \tag{5}$$

If (5) holds then μ is said to be a *doubling measure* [25, Chapter 1, 1.1]. A space of homogeneous type is said to be *normal*, if uniformly $\mu(B(x,r)) \sim r$.

Suppose Θ is an ellipsoid cover of $\mathbb{R}^n$ and let $\rho(\cdot,\cdot)$ be the associated quasi-distance, defined in (4). Denote $B_\rho(x,r) := \{y \in \mathbb{R}^n : \rho(x,y) < r\}$ for $x \in \mathbb{R}^n$, $r > 0$. As is shown in [12] there exist ellipsoids $\theta', \theta'' \in \Theta$ such that $\theta' \subset B_\rho(x,r) \subset \theta''$ and $|\theta'| \sim |B_\rho(x,r)| \sim |\theta''| \sim r$. Consequently, $\mathbb{R}^n$ equipped with the distance $\rho(\cdot,\cdot)$ and the Lebesgue measure, i.e. $(\mathbb{R}^n,\rho,dx)$ is a homogeneous type space. Therefore, the machinery of spaces of homogeneous type can be employed to our purposes here.

2.2 *Comparison of ellipsoid covers with nested triangulations in* $\mathbb{R}^2$

An alternative way of introducing anisotropic structures in $\mathbb{R}^2$ is through multilevel nested triangulations. The strong locally regular (SLR) triangulations, introduced in [20], provide a structure compatible with ellipsoid covers. We next recall briefly the definition of SLR-triangulations.

We call $\mathscr{T} = \bigcup_{m\in\mathbb{Z}} \mathscr{T}_m$ an SLR-triangulation of $\mathbb{R}^2$ with levels $\{\mathscr{T}_m\}$ if the following conditions are obeyed:

(a) Every level $\mathscr{T}_m$ consists of closed triangles with disjoint interiors which cover $\mathbb{R}^2$ and there are no hanging vertices.

(b) $\mathscr{T}_{m+1}$ is a refinement of $\mathscr{T}_m$ ($m \in \mathbb{Z}$) and each triangle $\triangle \in \mathscr{T}_m$ is subdivided and has uniformly bounded number of children in $\mathscr{T}_{m+1}$.

(c) For each $\triangle \in \mathscr{T}$ let $A_\triangle$ be an affine transform of the form

$$A_\triangle(x) = M_\triangle x + v_\triangle, \quad M_\triangle \in \mathbb{R}^{n\times n},$$

such that $\triangle = A_\triangle(\triangle^*)$, where $\triangle^*$ is an equilateral reference triangle. Now the condition is that for any $\triangle \in \mathscr{T}_m$ and $\triangle' \in \mathscr{T}_m \cup \mathscr{T}_{m+1}$ such that $\triangle' \cap \triangle \ne \emptyset$ one has

$$c_1 \le 1/\|M_{\triangle'}^{-1} M_\triangle\|_{\ell_2\to\ell_2} \le \|M_\triangle^{-1} M_{\triangle'}\|_{\ell_2\to\ell_2} \le c_2. \tag{6}$$

In [20] condition (c) is formulated in an equivalent form via a minimum angle condition.

Note that the multilevel SLR-triangulations provide a means for constructing discrete ellipsoid covers of $\mathbb{R}^2$. Given an SLR-triangulation $\mathscr{T}$ one considers for each triangle $\triangle \in \mathscr{T}$ the minimum area circumscribed ellipse. Then one dilates the resulting ellipses by a sufficiently large factor > 1 to obtain a discrete ellipse cover of $\mathbb{R}^2$.

The main advantage of ellipse covers over SLR-triangulations is that the latter are nested which makes them less flexible and harder to construct. On the other hand, as shown in [13] in presence of an SLR-triangulation it is easier to construct building blocks consisting of piecewise polynomials. Also the respective generalized Besov spaces and nonlinear approximation theory are easier to develop. We will be more specific about these issues later on.

3 Building blocks

The construction of simple elements (building blocks) which allow to represent the functions and characterize the norms of the spaces of interest is imperative for our theory. Here we first define a sequence of of bases consisting of C^∞ functions supported on the ellipsoids of the underlying anisotropic ellipsoid cover. Secondly, we construct compactly supported duals which generate local projectrors and two-level-split bases. Thirdly, we develop smooth global duals which provide polynomial reproducing kernels that we utilize to the construction of anisotropic frames.

3.1 Construction of a multilevel system of bases

Given a discrete ellipsoid cover Θ of $\mathbb{R}^n$, we first construct for each level $m \in \mathbb{Z}$ a stable basis Φ_m whose elements are smooth functions supported on the ellipsoids of Θ_m. The procedure begins by first *coloring the ellipsoids in* Θ. It is easy to see that Θ can be split into no more than $2N_1$ disjoint subsets (colors) $\{\Theta^\ell\}_{\ell=1}^{2N_1}$ so that for any $m \in \mathbb{Z}$ neither two ellipsoids $\theta', \theta'' \in \Theta_m \cup \Theta_{m+1}$ with $\theta' \cap \theta'' \neq \emptyset$ are of the same color.

Our second step is to construct locally independent piecewise polynomial bumps. For fixed positive integers M and k ($M \geq k$) we define

$$\tilde{\phi}_\ell(x) := (1 - |x|^2)_+^{M+\ell k}, \quad \ell = 1, 2, \ldots, 2N_1.$$

Notice that $\tilde{\phi}_\ell$, $\ell = 1, \ldots, 2N_1$, being of different degrees are linearly independent on any ball contained in $B^* = B(0,1)$.

The next step is to smooth out each $\tilde{\phi}_\ell$ by convolving it with a compactly supported C^∞ function. Namely, let $h \in C^\infty(\mathbb{R}^n)$ be such that $\operatorname{supp} h = \overline{B^*}$, $h \geq 0$, and

$\int_{\mathbb{R}^n} h = 1$. Denote $h_\delta(x) := \delta^{-n} h(\delta^{-1}x)$. Then for $0 < \delta < 1$ (we choose δ sufficiently small) the bump

$$\phi_\ell^* := \tilde{\phi}_\ell * h_\delta$$

belongs to C^∞, ϕ_ℓ^* is a polynomial of degree exactly $2(M + \ell k)$ on $B(0, 1-\delta)$ and $\operatorname{supp} \phi_\ell^* = \overline{B(0,1+\delta)}$. Now we define $\phi_\ell(x) := \phi_\ell^*((1+\delta)x)$.

For any $\theta \in \Theta$ we let A_θ denote the affine transform from Definition 1 such that $A_\theta(B^*) = \theta$ and set

$$\phi_\theta := \phi_\ell \circ A_\theta^{-1} \quad \text{for} \quad \theta \in \Theta^\ell,\ 1 \le \ell \le 2N_1.$$

We introduce an mth level partition of unity by defining for each $\theta \in \Theta_m$

$$\varphi_\theta := \frac{\phi_\theta}{\sum_{\theta' \in \Theta_m} \phi_{\theta'}}. \tag{7}$$

By property (d) of ellipsoids covers it follows that $\sum_{\theta \in \Theta_m} \varphi_\theta(x) = 1$ for $x \in \mathbb{R}^n$.

Let

$$\{P_\beta : |\beta| \le k-1\}, \quad \text{where } \deg P_\beta = |\beta|, \tag{8}$$

be an orthonormal basis in $L_2(B^*)$ for the space $\mathscr{P}_k$ of all polynomials in n variables of total degree $k-1$. For each $\theta \in \Theta$ and $|\beta| < k$ we define

$$P_{\theta,\beta} := |\theta|^{-1/2} P_\beta \circ A_\theta^{-1} \quad \text{and} \quad g_{\theta,\beta} := \varphi_\theta P_{\theta,\beta}. \tag{9}$$

To simplify our notation, we denote

$$\Lambda_m := \{\lambda := (\theta,\beta) : \theta \in \Theta_m, |\beta| < k\} \quad \text{and} \quad g_\lambda := g_{\theta,\beta}, \quad \lambda = (\theta,\beta). \tag{10}$$

Also, for $\lambda = (\theta,\beta)$ we will denote by θ_λ and β_λ the components of λ.

Now we define the mth level basis Φ_m by

$$\Phi_m := \{g_\lambda : \lambda \in \Lambda_m\} \quad \text{and set} \quad S_m := \operatorname{span}(\Phi_m), \tag{11}$$

where "span" means "closed span".

By the definition of $\{g_\lambda\}$ it readily follows that $\mathscr{P}_k \subset S_m$. More importantly, Φ_m is locally linearly independent and L_p-stable, which will be recorded in the next theorem.

Theorem 1. *Any function $f \in S_m$ has a unique representation*

$$f(x) = \sum_{\lambda \in \Lambda_m} \langle f, \tilde{g}_\lambda \rangle g_\lambda(x), \tag{12}$$

where for every $x \in \mathbb{R}^n$ the sum is finite and the functions $\{\tilde{g}_\lambda\}$ have the following properties: $\operatorname{supp}(\tilde{g}_\lambda) \subset \theta_\lambda$, $\|\tilde{g}_{\theta,\beta}\|_p \sim |\theta|^{1/p-1/2}$ *and the biorthogonal relation $\langle g_{\lambda'}, \tilde{g}_\lambda \rangle = \delta_{\lambda',\lambda}$ holds. Moreover, for any $f \in S_m \cap L_p$, $0 < p \le \infty$, such that $f = \sum_{\lambda \in \Lambda_m} c_\lambda g_\lambda$ we have*

$$\|f\|_p \sim \Big(\sum_{\lambda\in\Lambda_m} \|c_\lambda g_\lambda\|_p^p\Big)^{1/p} \tag{13}$$

with the obvious modification when $p = \infty$.

The proof of this theorem is based on the local linear independence of the functions $\{g_\lambda : \lambda \in \Lambda_m\}$ and uses a compactness argument, see [12] for the details.

We will denote $\widetilde{\Phi}_m := \{\tilde{g}_\lambda : \lambda \in \Lambda_m\}$.

3.2 Compactly supported duals and local projectors

Our next step is to introduce simple operators which map $L_p^{\rm loc}$ into S_m and locally preserve the polynomials $P \in \mathscr{P}_k$ with $\mathscr{P}_k$ being the set of all polynomials of degree $< k$. These operators will give us a vehicle for developing a decomposition scheme which allows to characterize the anisotropic Besov norms induced by ellipsoid covers of $\mathbb{R}^n$.

Using the bases $\{\Phi_m\}$ and their duals $\{\widetilde{\Phi}_m\}$ from Theorem 1 we introduce projectors Q_m mapping $L_p^{\rm loc}$ $(1 \le p \le \infty)$ onto the spaces S_m defined by

$$Q_m f := \int_{\mathbb{R}^n} Q_m(x,y) f(y)dy \quad \text{with} \quad Q_m(x,y) := \sum_{\lambda\in\Lambda_m} \tilde{g}_\lambda(y) g_\lambda(x). \tag{14}$$

Evidently, Q_m is a linear operator which maps $L_p^{\rm loc}$ into S_m and preserves locally all polynomials from $\mathscr{P}_k$. To be more specific, setting

$$\theta^* := \cup\{\theta' \in \Theta_m : \theta \cap \theta' \neq \emptyset\} \quad \text{for } \theta \in \Theta_m, \tag{15}$$

it is easy to see that if $f|_{\theta^*} = P|_{\theta^*}$ with $P \in \mathscr{P}_k$, then $Q_m f|_\theta = P|_\theta$.

Another simple operator with similar properties is given in [12].

Evidently, the operators $\{Q_m\}$ from above are no longer usable, when working in L_p with $p < 1$. In this case, for a given ellipsoid $\theta \in \Theta$, we let $T_{\theta,p} : L_p(\theta) \to \mathscr{P}_k|_\theta$ be a projector such that

$$\|f - T_{\theta,p} f\|_{L_p(\theta)} \le cE_k(f,\theta)_p, \quad f \in L_p(\theta), \tag{16}$$

where $E_k(f,\theta)_p := \inf_{P\in\mathscr{P}_k} \|f - P\|_{L_p(\theta)}$. Thus $T_{\theta,p} f$ is simply a near best approximation to f from $\mathscr{P}_k$ in $L_p(\theta)$, and hence $T_{\theta,p}$ can be realized as a linear projector onto $\mathscr{P}_k|_\theta$ if $p \ge 1$ by using, say, the Averaged Taylor polynomials, see e.g. [13]. Of course, $T_{\theta,p}$ is a nonlinear operator if $p < 1$.

We now define the operator $T_{m,p} : L_p^{\rm loc} \to S_m$ by

$$T_{m,p} f := \sum_{\theta\in\Theta_m} \varphi_\theta T_{\theta,p} f. \tag{17}$$

Evidently, the operator $T_{m,p}$ $(0 < p \le \infty)$ is a local projector onto $\mathscr{P}_k$ (nonlinear if $p < 1$) just like Q_m. Since $T_{m,p} f \in S_m$, it can be represented in terms of the basis functions g_λ as

$$T_{m,p} f = \sum_{\theta \in \Theta_m} \sum_{|\beta|<k} b_{\theta,\beta}(f) g_{\theta,\beta} = \sum_{\lambda \in \Lambda_m} b_\lambda(f) g_\lambda, \tag{18}$$

where $b_\lambda(f) := \langle T_{m,p} f, \tilde{g}_\lambda \rangle$ depends nonlinearly on f if $p < 1$.

In summary, if $\hat{T}_m := Q_m$ or $\hat{T}_m := T_{m,p}$, then

$$\hat{T}_m f = \sum_{\lambda \in \Lambda_m} b_\lambda(f) g_\lambda, \ \text{ where } \ b_\lambda(f) = \begin{cases} \langle f, \tilde{g}_\lambda \rangle & \text{if } \hat{T}_m = Q_m, \\ \langle T_{m,p} f, \tilde{g}_\lambda \rangle & \text{if } \hat{T}_m = T_{m,p}. \end{cases} \tag{19}$$

We now recall briefly the definition of local and global moduli of smoothness that are standard means for describing the quality of approximation. The forward differences of a function f on a set $E \subset \mathbb{R}^n$ in direction $h \in \mathbb{R}^n$ are defined by

$$\Delta_h^k f(x) := \sum_{j=0}^{k} (-1)^{k+j} \binom{k}{j} f(x + jh) \quad \text{if } [x, x + kh] \subset E$$

and $\Delta_h^k f(x) := 0$ otherwise. Then the kth L_p-moduli of smoothness on E and $\mathbb{R}^n$ are defined by

$$\omega_k(f, E)_p := \sup_{h \in \mathbb{R}^n} \|\Delta_h^k f\|_{L_p(E)} \quad \text{and} \quad \omega_k(f, t)_p := \sup_{|h| \le t} \|\Delta_h^k f\|_p, \quad t > 0. \tag{20}$$

We next give the most important properties of the operators Q_m and $T_{m,p}$ from above.

Proposition 2. *Suppose $\hat{T}_m$ is any of the operators Q_m or $T_{m,p}$ if $1 \le p \le \infty$, and $\hat{T}_m := T_{m,p}$ if $0 < p < 1$. Then for $f \in L_p^{\text{loc}}$ and $\theta \in \Theta_m$ $(m \in \mathbb{Z})$*

$$\|f - \hat{T}_m f\|_{L_p(\theta)} \le c \sum_{\theta' \in \Theta_m :\, \theta' \cap \theta \ne \emptyset} \omega_k(f, \theta')_p,$$

and $\|f - \hat{T}_m f\|_{L_p(K)} \to 0$ as $m \to \infty$ for any compact $K \subset \mathbb{R}^n$.

Furthermore, if $f \in L_p$ $(L_\infty := C_0)$, then $\|f - \hat{T}_m f\|_p \to 0$ as $m \to \infty$.

3.3 Two-level-split bases

Assume that T_m $(m \in \mathbb{Z})$ is one of the operators Q_m or $T_{m,p}$ if $p \ge 1$, and $T_m := T_{m,p}$ if $p < 1$, defined in §3.2. These operators and the bases $\{\Phi_m\}_{m \in \mathbb{Z}}$ from (11) will be used to define two-level-split bases which will play an important role in what follows.

We will make use of the following representation of consecutive level polynomial bases, defined in (9):

$$P_{\theta,\alpha} =: \sum_{|\beta|<k} C^{\theta,\eta}_{\alpha,\beta} P_{\eta,\beta}, \quad \theta \in \Theta_m, \quad \eta \in \Theta_{m+1}. \tag{21}$$

Then since $\sum_{\eta\in\Theta_{m+1}} \varphi_\eta = 1$, we have

$$P_{\theta,\alpha} = \sum_{\eta\in\Theta_{m+1}:\theta\cap\eta\neq\emptyset} \sum_{|\beta|<k} C^{\theta,\eta}_{\alpha,\beta} P_{\eta,\beta}\varphi_\eta \quad \text{on } \theta.$$

This yields

$$\begin{aligned} T_{m+1}f - T_m f &= \sum_{\eta\in\Theta_{m+1}} \sum_{|\beta|<k} b_{\eta,\beta}(f) P_{\eta,\beta}\varphi_\eta - \sum_{\theta\in\Theta_m} \sum_{|\alpha|<k} b_{\theta,\alpha}(f) P_{\theta,\alpha}\varphi_\theta \\ &= \sum_{\theta\in\Theta_m} \varphi_\theta \sum_{\eta\in\Theta_{m+1}} \sum_{|\beta|<k} b_{\eta,\beta}(f) P_{\eta,\beta}\varphi_\eta \\ &\quad - \sum_{\theta\in\Theta_m} \sum_{|\alpha|<k} b_{\theta,\alpha}(f) \sum_{\theta\cap\eta\neq\emptyset} \sum_{|\beta|<k} m^{\theta,\eta}_{\alpha,\beta} P_{\eta,\beta}\varphi_\theta\varphi_\eta \\ &= \sum_{\eta\in\Theta_{m+1}} \sum_{\theta\in\Theta_m:\theta\cap\eta\neq\emptyset} \sum_{|\beta|<k} \Big\{ b_{\eta,\beta}(f) - \sum_{|\alpha|<k} m^{\theta,\eta}_{\alpha,\beta} b_{\theta,\alpha}(f) \Big\} P_{\eta,\beta}\varphi_\eta\varphi_\theta, \end{aligned} \tag{22}$$

where $b_\lambda(f)$ are given by (19) and depends on the choice of T_m. Thus, setting

$$\mathscr{V}_m := \{\nu = (\eta,\theta,\beta) : \eta \in \Theta_{m+1}, \theta \in \Theta_m, \theta\cap\eta \neq 0, |\beta| < k\}, \quad m \in \mathbb{Z}, \tag{23}$$

the building blocks in (22) have the form

$$F_\nu := P_{\eta,\beta}\varphi_\eta\varphi_\theta, \quad \nu = (\eta,\theta,\beta) \in \mathscr{V}_m, \tag{24}$$

where $P_{\eta,\beta}$ are defined in (9) and φ_η, φ_θ are from (7). We define

$$\mathscr{F}_m := \{F_\nu : \nu \in \mathscr{V}_m\} \quad \text{and} \quad W_m := \operatorname{span}(\mathscr{F}_m), \quad m \in \mathbb{Z}. \tag{25}$$

Note that $F_\nu \in C^\infty$, $\operatorname{supp} F_\nu = \overline{\theta\cap\eta}$ if $\nu = (\eta,\theta,\beta)$ and $\|F_\nu\|_2 \sim 1$.

One uses the argument of the proof of Theorem 1 (see [12]) to establish the stability of the two-level-split bases:

Theorem 2. *Any $f \in W_m$ has a unique representation*

$$f = \sum_{\nu\in\mathscr{V}_m} c_\nu(f) F_\nu, \tag{26}$$

where the dual functionals $c_\nu(\cdot)$ are of the following form: For each $\nu \in \mathscr{V}_m$, $\nu = (\eta,\theta,\beta)$, there is an ellipsoid $B_\nu \subset \theta\cap\eta$ with $|B_\nu| \sim |\eta|$ and $B_\nu = A_\eta(B^*_\nu)$ for some ball $B^*_\nu \subset B^*$ such that $c_\nu(f) = \langle f, \tilde{F}_\nu\rangle$, where $\operatorname{supp}\tilde{F}_\nu \subset \overline{B}_\nu$, $\|\tilde{F}_\nu\|_p \sim |\eta|^{1/p-1/2}$.

Moreover, if $f \in W_m$ and $f = \sum_{\nu\in\mathscr{V}_m} a_\nu F_\nu$, then

$$\|f\|_p \sim \Big(\sum_{\nu\in\mathscr{V}_m} \|a_\nu F_\nu\|_p^p\Big)^{1/p}, \quad 0<p\le\infty, \tag{27}$$

with the obvious modification when $p=\infty$.

Using the results from §3.2 one easily derives multilevel decompositions of functions using the two-level-split bases from above.

Theorem 3. *For any* $f\in L_p^{\mathrm{loc}}(\mathbb{R}^n)$, $0<p\le\infty$,

$$f = T_0 f + \sum_{m\ge 0}(T_{m+1}f - T_m f) = \sum_{m\ge -1}\sum_{\nu\in\mathscr{V}_m} d_\nu(f)F_\nu, \tag{28}$$

where the convergence is in $L_p(K)$ *for all compacta* $K\subset\mathbb{R}^n$. *Here for* $m\ge 0$

$$d_\nu(f) = b_{\eta,\beta}(f) - \sum_{|\alpha|<k} C_{\alpha,\beta}^{\theta,\eta} b_{\theta,\beta}(f), \quad \nu := (\eta,\theta,\beta) \tag{29}$$

with $C_{\alpha,\beta}^{\theta,\eta}$ *from* (21), *while* $\mathscr{V}_{-1} := \Lambda_0$, $F_\lambda := g_\lambda$ *and* $d_\lambda(f) := b_\lambda(f)$ *if* $\lambda\in\mathscr{V}_{-1}$.

Moreover, if $f\in L_p$ $(L_\infty := C_0)$, *then* (28) *as well as*

$$f = \sum_{m\in\mathbb{Z}} (T_{m+1} - T_m) f \tag{30}$$

hold in L_p.

3.4 Global duals and polynomial reproducing kernels

A substantial drawback of the operators Q_m and $T_{m,p}$ considered in §§3.2-3.3 is that their transposed operators do not reproduce polynomials. For instance, it is easy to see that for the operator Q_m from (14) we have $Q_m P(x) = \int_{\mathbb{R}^n} Q_m(x,y)P(y)dy$ $\forall P\in\mathscr{P}_k$, however, $Q_m P(y) = \int_{\mathbb{R}^n} Q_m(x,y)P(x)dx$ is no longer true for $P\in\mathscr{P}_k$. Consequently, these operators do not fit in the general framework of approximation to the identity operators in homogeneous spaces, which allows to construct anisotropic wavelet frames (see e.g. [16]). This problem is fixed in [14] by introducing smooth duals to the bases $\{g_\lambda\}_{\lambda\in\Lambda_m}$, which we describe next.

As in [14] to simplify our set-up we will assume for the rest of this section that in the definition of ellipsoid covers of $\mathbb{R}^n$ we have $a_0 = 0$ (see Definition 1). Also to make our presentation more compatible with [14] we will slightly change our notation assuming that all operators of interest reproduce polynomials of degree $<r$ instead of degree $<k$.

The next step is to introduce an appropriate generalization to higher orders of the *approximation to the identity* definition given in [16]. To this end we first have to establish some convenient notation. Let $K(x,y)$ be a smooth kernel. For $x,y\in\mathbb{R}^n$ the Taylor representation of $K(x,y)$ centered at x with y fixed can be written in the form

$$K(z,y) = T_{r-1,x}(K(\cdot,y))(z) + R_{r,x}(K(\cdot,y))(z), \quad z \in \mathbb{R}^n, \tag{31}$$

where T_{r-1} is the Taylor polynomial of degree $r-1$ and $R_{r,x}$ is the rth order Taylor remainder.

In the particular case of spaces of homogeneous type generated by an anisotropic ellipsoid cover of $\mathbb{R}^n$ with a qusi-distance $\rho(\cdot,\cdot)$ we will need the notation

$$\mu(x,y,d) := \begin{cases} \mu_0 & \text{if} \quad \rho(x,y) < d, \\ \mu_1 & \text{if} \quad \rho(x,y) \geq d. \end{cases} \tag{32}$$

Definition 3. Let $(\mathbb{R}^n, \rho, dx)$ be a normal space of homogeneous type. A sequence of kernel operators $\{S_m\}_{m\in\mathbb{Z}}$, formally defined by $S_m(f)(x) := \int_{\mathbb{R}^n} S_m(x,y) f(y) dy$, is an *approximation to the identity of order* (μ,δ,r), where $\mu = (\mu_0,\mu_1)$, $0 < \mu_0 \leq \mu_1 \leq 1$, $\delta > 0$, $r \in \mathbb{N}$, with respect to $\rho(\cdot,\cdot)$, if for some constant $c > 0$ the following conditions are satisfied:

(i) $|S_m(x,y)| \leq c \frac{2^{-m\delta}}{(2^{-m}+\rho(x,y))^{1+\delta}}, \quad \forall x,y \in \mathbb{R}^n.$

(ii) For $1 \leq k \leq r$ and all $x,y,z \in \mathbb{R}^n$,

$$\begin{aligned} |R_{k,x}(S_m(\cdot,y))(z)| &\leq c\rho(x,z)^{\mu(x,z,2^{-m})k} \\ &\quad \times \left(\frac{2^{-m\delta}}{(2^{-m}+\rho(x,y))^{1+\delta+\mu(x,z,2^{-m})k}} + \frac{2^{-m\delta}}{(2^{-m}+\rho(y,z))^{1+\delta+\mu(x,z,2^{-m})k}} \right), \end{aligned}$$

$$\begin{aligned} |R_{k,y}(S_m(x,\cdot))(z)| &\leq c\rho(y,z)^{\mu(y,z,2^{-m})k} \\ &\quad \times \left(\frac{2^{-m\delta}}{(2^{-m}+\rho(x,y))^{1+\delta+\mu(y,z,2^{-m})k}} + \frac{2^{-m\delta}}{(2^{-m}+\rho(x,z))^{1+\delta+\mu(y,z,2^{-m})k}} \right), \end{aligned}$$

(iii) For $1 \leq k \leq r$ and all $x,x',y,y' \in \mathbb{R}^n$

$$\begin{aligned} &|R_{k,y}(R_{k,x}(S_m(\cdot,\cdot))(x'))(y')|, |R_{k,x}(R_{k,y}(S_m(\cdot,\cdot))(y'))(x')| \\ &\leq c\rho(x,x')^{\mu(x,x',2^{-m})k}\rho(y,y')^{\mu(y,y',2^{-m})k} \\ &\times \Bigg(\frac{2^{-m\delta}}{(2^{-m}+\rho(x,y))^{1+\delta+\mu(x,x',2^{-m})k+\mu(y,y',2^{-m})k}} \\ &\qquad + \frac{2^{-m\delta}}{(2^{-m}+\rho(x,y'))^{1+\delta+\mu(x,x',2^{-m})k+\mu(y,y',2^{-m})k}} \\ &+ \frac{2^{-m\delta}}{(2^{-m}+\rho(x',y))^{1+\delta+\mu(x,x',2^{-m})k+\mu(y,y',2^{-m})k}} \\ &\qquad + \frac{2^{-m\delta}}{(2^{-m}+\rho(x',y'))^{1+\delta+\mu(x,x',2^{-m})k+\mu(y,y',2^{-m})k}} \Bigg) \end{aligned}$$

[To clarify our notation, denote $g_m(x,x',y) := R_{k,x}\left(S_m(\cdot,y)\right)(x')$, then for fixed $x,x' \in \mathbb{R}^n$, $R_{k,y}\left(R_{k,x}\left(S_m(\cdot,\cdot)\right)(x')\right)(y') := R_{k,y}\left(g_m(x,x',\cdot)\right)(y')$].

(iv) $P(x) = \int_{\mathbb{R}^n} S_m(x,y)P(y)dy$ and $P(y) = \int_{\mathbb{R}^n} S_m(x,y)P(x)dx$ for all $P \in \mathscr{P}_r$.

Note that the definition of an approximation of the identity given in [16] corresponds to the case $0 < \delta < r = 1$.

To construct well localized kernels $S_m(x,y)$ which reproduce polynomials we need to construct an appropriate dual basis to Φ_m. Let G_m be the Gram matrix

$$G_m := \left[A_{\lambda,\lambda'}\right]_{\lambda,\lambda'\in\Lambda_m}, \quad A_{\lambda,\lambda'} := \langle g_\lambda, g_{\lambda'}\rangle := \int_{\mathbb{R}^n} g_\lambda g_{\lambda'}.$$

By Theorem 1, for any sequence $t = (t_\lambda)_{\lambda\in\Lambda_m}$ in $l_2(\Lambda_m)$ we have

$$c_1\|t\|_{l_2} \le \langle G_m t, t\rangle = \Big\|\sum_{\lambda\in\Lambda_m} t_\lambda g_\lambda\Big\|_2 \le c_2 \le \|t\|_{l_2},$$

where the constants $c_1, c_2 > 0$ are independent on t and m. Therefore, the operator $G_m : l_2 \to l_2$ with matrix G_m is symmetric, positive and $c_1 I \le G_m \le c_2 I$. Hence, G_m^{-1} exists and $c_2^{-1} I \le G_m^{-1} \le c_1^{-1} I$. Denote by $G_m^{-1} =: \left[B_{\lambda,\lambda'}\right]_{\lambda,\lambda'\in\Lambda_m}$ the matrix of the operator G_m^{-1}.

The next lemma shows that the entries of G_m^{-1} decay away from its main diagonal at sub-exponential rate.

Lemma 1. [14] *There exist constants* $0 < q_*, \gamma < 1$ *and* $c > 0$ *depending only on* $\mathbf{p}(\Theta)$ *and* r *such that for any entry* $B_{\lambda,\lambda'}$ *of* G_m^{-1} $(\lambda,\lambda' \in \Lambda_m)$ *and points* $x \in \theta_\lambda$, $y \in \theta_{\lambda'}$

$$|B_{\lambda,\lambda'}| \le c q_*^{(2^m\rho(x,y))^\gamma}. \tag{33}$$

Definition of smooth duals. We define new duals by

$$\tilde{\tilde{g}}_\lambda := \sum_{\lambda'\in\Lambda_m} B_{\lambda,\lambda'} g_{\lambda'}, \quad \lambda \in \Lambda_m, \tag{34}$$

and set $\tilde{\tilde{\Phi}}_m := \{\tilde{\tilde{g}}_\lambda\}_{\lambda\in\Lambda_m}$. For $\lambda \in \Lambda_m$, let x_0 be any point in θ_λ. Combining (33) and (34) it follows that

$$|\tilde{\tilde{g}}_\lambda(x)| \le c2^{-m/2} \sum_{x\in\theta_{\lambda'}} |B_{\lambda,\lambda'}| \le c2^{-m/2} q_*^{(2^m\rho(x,x_0))^\gamma}. \tag{35}$$

Therefore, each $\tilde{\tilde{g}}_\lambda$ has sub-exponential decay with respect to the quasi-distance induced by Θ. Also, it is easy to verify the biorthogonality relation, namely,

$$\langle g_\lambda, \tilde{\tilde{g}}_{\lambda'}\rangle = \sum_{\lambda''\in\Lambda_m} B_{\lambda',\lambda''}\langle g_\lambda, g_{\lambda''}\rangle = \left(G_m^{-1} G_m\right)_{\lambda',\lambda} = \delta_{\lambda,\lambda'}.$$

We use the bases Φ_m and $\tilde{\tilde{\Phi}}_m$ to introduce an approximation to the identity determined by the operators $\{S_m\}_{m\in\mathbb{Z}}$ with kernels

$$S_m(x,y) := \sum_{\lambda\in\Lambda_m} g_\lambda(x)\tilde{\tilde{g}}_\lambda(y). \tag{36}$$

In the next theorem we record the fact that these kernels define the desired approximation to the identity.

Theorem 4. [14] *For a discrete ellipsoid cover Θ, the kernels from (36) define an approximation to the identity with respect to the quasi-distance $\rho(\cdot,\cdot)$ induced by Θ. Here the vector μ can be defined as $\mu := (a_6,a_4)$, the parameter δ can be selected arbitrarily large and the parameter r is the degree of the polynomials used in the construction of the local ellipsoid "bumps" in §3.1.*

3.5 Construction of anisotropic wavelet frames

Wavelet operators. Let $\{S_m\}_{m\in\mathbb{Z}}$ be an approximation to the identity of order (μ,δ,r). Then evidently the kernels of the *wavelet operators* $D_m := S_{m+1} - S_m$ satisfy conditins (i)-(iii) in Definition 3, while the polynomial reproduction condition (iv) is replaced by the following *zero moment* condition

$$\int_{\mathbb{R}^n} D_m(x,y)P(y)dy = 0, \ \int_{\mathbb{R}^n} D_m(x,y)P(x)dx = 0 \ \ \forall P\in\mathscr{P}_r. \tag{37}$$

The next lemma shows that any two wavelet operators (kernels) from different scales are "almost orthogonal".

Lemma 2. [14] *Suppose two kernel operators $\{D^1_m\}_{m\in\mathbb{Z}}$ and $\{D^2_m\}_{m\in\mathbb{Z}}$ satisfy (37) for some $r\geq 1$ and conditions (i)-(ii) of an approximation to the identity of order (μ,δ,r) for some $\delta\geq\mu_1 r$. Then*

$$|D^1_k D^2_l(x,y)| \leq c2^{-|k-l|\mu_0 r}\frac{2^{-\min\{k,l\}\delta}}{\left(2^{-\min\{k,l\}}+\rho(x,y)\right)^{1+\delta}}, \quad k,l\in\mathbb{Z}. \tag{38}$$

Dual wavelet operators. In this section we leverage significantly on the results of Han and Sawyer [19] (see also [16]) concerning the Calderón reproducing formula in spaces of homogeneous type and adapt them to our specific setting. We begin with the definitions for anisotropic test functions and molecules.

Definition 4. Let $\rho(\cdot,\cdot)$ be a quasi-distance on $\mathbb{R}^n$. A function $f\in C(\mathbb{R}^n)$ is said to be in the *anisotropic test function space* $\mathscr{M}(\varepsilon,\delta,x_0,t)$, $0<\varepsilon,\delta\leq 1$, $x_0\in\mathbb{R}^n$, $t\in\mathbb{R}$, if there exists a constant $C>0$ such that

(i) $|f(x)|\leq C\frac{2^{-t\delta}}{(2^{-t}+\rho(x,x_0))^{1+\delta}}$ $\forall x\in\mathbb{R}^n$.

(ii) $|f(x)-f(y)|\leq C\rho(x,y)^\varepsilon\frac{2^{-t\delta}}{(2^{-t}+\rho(x,x_0))^{1+\delta+\varepsilon}}$ for all $x,y\in\mathbb{R}^n$, where $\rho(x,y) \le \frac{1}{2\kappa}(2^{-t} + \rho(x,x_0))$ with κ the constant of the quasi-distance (see §2.1).

One can easily show that $\mathscr{M}(\varepsilon,\delta,x_0,t)$ is a Banach space with norm $\|f\|_{\mathscr{M}}$ defined as the infimum of all constants C such that (i)-(ii) are valid. We also denote $\mathscr{M}(\varepsilon,\delta) := \mathscr{M}(\varepsilon,\delta,0,0)$.

Definition 5. The set of *molecules* $\mathscr{M}_0(\varepsilon,\delta,x_0,t)$ is defined as the set of all anisotropic test functions $f \in \mathscr{M}(\varepsilon,\delta,x_0,t)$ such that $\int_{\mathbb{R}^n} f(y)\,dy = 0$.

We denote by $\mathscr{M}_0(\varepsilon,\delta)$ the subspace of all molecules in $\mathscr{M}(\varepsilon,\delta)$.

For some $\gamma > \varepsilon$, let $\overset{\circ}{\mathscr{M}}(\varepsilon,\delta)$ be the closure of $\mathscr{M}(\gamma,\delta)$ in the norm of $\mathscr{M}(\varepsilon,\delta)$. Then, we define $\overset{\circ}{\mathscr{M}}{}'(\varepsilon,\delta)$ as the dual of $\overset{\circ}{\mathscr{M}}(\varepsilon,\delta)$.

We are now prepared to state the Calderón reproducing formula which implies the existence of dual wavelet operators.

Theorem 5. [Continuous Calderón reproducing formula] *Suppose $(\mathbb{R}^n,\rho,dx)$ is a normal space of homogeneous type and let $\{S_m\}_{m\in\mathbb{Z}}$ be an approximation to the identity of order (μ,δ,r) with respect to $\rho(\cdot,\cdot)$. Set $D_m := S_{m+1} - S_m$ for $m \in \mathbb{Z}$. Then there exist linear operators $\{\tilde{D}_m\}_{m\in\mathbb{Z}}$ and $\{\hat{D}_m\}_{m\in\mathbb{Z}}$ such that for any $f \in \mathscr{M}_0(\varepsilon,\gamma)$, $0 < \varepsilon,\gamma < \mu_0$,*

$$f = \sum_{m\in\mathbb{Z}} \tilde{D}_m D_m(f) = \sum_{m\in\mathbb{Z}} D_m \hat{D}_m(f), \tag{39}$$

where the series converge in the norm of $\mathscr{M}(\varepsilon',\gamma')$, $\varepsilon' < \varepsilon$, $\gamma' < \gamma$, and in $L_p(\mathbb{R}^n)$, $1 < p < \infty$. Furthermore, for any $\varepsilon < \mu_0$, the kernels of $\{\tilde{D}_m\}$ and $\{\hat{D}_m\}$ satisfy conditions (i)-(iii) of an approximation to the identity of order $(\mu,\varepsilon,1)$ (with constants depending on ε) and the r-th zero moments condition (37).

By a duality argument we obtain

Corollary 1. *Under the hypothesis of Theorem 5 for any $f \in \overset{\circ}{\mathscr{M}}{}'(\varepsilon,\delta)$ the series in (39) converges in $\overset{\circ}{\mathscr{M}}{}'(\varepsilon_*,\delta_*)$ with $\varepsilon < \varepsilon_* < \mu_0$, $\gamma < \gamma_* < \mu_0$.*

We next sketch the proof of Theorem 5. The method of proof is essentially similar to the method used in [19]. We use Coifman's idea to write the identity operator I as

$$I = \sum_k D_k = \sum_k D_k \sum_l D_l = \sum_{k,l} D_k D_l.$$

For an integer $N > 0$ we introduce the operator $D_m^N := \sum_{|j|\le N} D_{m+j}$ and define the operators T_N and R_N by

$$I = \sum_{k,l} D_k D_l = \sum_{k\in\mathbb{Z}} D_k^N D_k + \sum_{|j|>N} \sum_{k\in\mathbb{Z}} D_{k+j} D_k =: T_N + R_N.$$

Let $0 < \varepsilon,\gamma < \mu_0$. We claim that R_N is bounded on $\mathscr{M}_0(\varepsilon,\gamma,x_0,t)$ for any $x_0 \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Moreover, there exist constants $\tau > 0$ and $c > 0$ such that

$$\|R_N f\|_{\mathscr{M}_0(\varepsilon,\gamma,x_0,t)} \le c2^{-N\tau}\|f\|_{\mathscr{M}_0(\varepsilon,\gamma,x_0,t)} \quad \text{for} \quad f \in \mathscr{M}_0(\varepsilon,\gamma,x_0,t). \tag{40}$$

Assume the claim for a moment. Choosing N so that $c2^{-N\tau} < 1$, then (40) implies that the operator T_N^{-1} exists and is bounded on $\mathscr{M}_0(\varepsilon,\gamma,x_0,t)$. Thus, we obtain

$$I = T_N^{-1}T_N = \sum_m \left(T_N^{-1}D_m^N\right)D_m = \sum_m \tilde{D}_m D_m,$$

where $\tilde{D}_m := T_N^{-1}D_m^N$. The regularity conditions on the kernels $\{D_m\}$ and (37) imply that for any fixed N and $y \in \mathbb{R}^n$ the function $D_m^N(\cdot,y)$ is in $\mathscr{M}_0(\mu_0,\delta)$. This immediately implies that $\tilde{D}_m(\cdot,y) = T_N^{-1}D_m^N(\cdot,y)$ is in $\mathscr{M}_0(\varepsilon,\gamma)$ for any $0 < \varepsilon,\gamma < \mu_0$. Similarly, we can write

$$I = T_N T_N^{-1} = (\sum_m D_m^N D_m)T_N^{-1} = \sum_m D_m D_m^N T_N^{-1} = \sum_m D_m \hat{D}_m,$$

where $\hat{D}_m := D_m^N T_N^{-1}$. By the same token, for any fixed N and $x \in \mathbb{R}^n$, the function $\hat{D}_m(x,\cdot)$ is in $\mathscr{M}_0(\varepsilon,\gamma)$ for any $0 < \varepsilon,\gamma < \mu_0$ and the proof is complete.

Discussion. In the proof of Theorem 5 we applied tools from the general theory of spaces of homogeneous type to construct dual wavelet operators. Although the kernels of the dual operators $\{\tilde{D}_m\}$ and $\{\hat{D}_m\}$ have the same vanishing moments as $\{D_m\}$, we only claim very "modest" regularity and decay on them. For example, in Theorem 5 we claim that for any $0 < \gamma < \mu_0$, there exists a constant $c > 0$ such that

$$|\tilde{D}_m(x,y)|, |\hat{D}_m(x,y)| \le \frac{c2^{-m\gamma}}{(2^{-m} + \rho(x,y))^{1+\gamma}}.$$

At the same time, the construction of the anisotropic approximation of the identity over an ellipsoid cover in §3.4 (see Theorem 4) produces wavelet kernels $\{D_m\}$ such that for any $\delta > 0$

$$|D_m(x,y)| \le \frac{c2^{-m\delta}}{(2^{-m} + \rho(x,y))^{1+\delta}}, \quad c = c(\delta).$$

It is an *open problem* to define higher order anisotropic test function spaces and prove that the operators $R_N := \sum_{|j|>N}\sum_{k\in\mathbb{Z}} D_{k+j}D_k$ are bounded on these higher order spaces as in (40).

Applying the Calderón reproducing formula we obtain the following Littlewood-Paley type result (see [16]).

Proposition 3. *Suppose $\{S_m\}_{m\in\mathbb{Z}}$ is an anisotropic approximation of the identity and let $D_m = S_{m+1} - S_m$, $m \in \mathbb{Z}$. Then for any $f \in L_p(\mathbb{R}^n)$, $1 < p < \infty$, we have*

$$\|f\|_p \sim \left\|\left(\sum_m |D_m(f)(\cdot)|^2\right)^{1/2}\right\|_p.$$

3.6 Discrete wavelet frames

Here we describe briefly the construction of wavelet frames using the discrete Calderón reproducing formula, which in turn is obtained by "sampling" the continuous Calderón reproducing formula (see e.g. [16, 14]). We first introduce the following sampling process.

Definition 6. Let $\rho(\cdot,\cdot)$ be a quasi-distance on $\mathbb{R}^n$. We call a set of closed domains $\Omega_{m,k} \subset \mathbb{R}^n$, $m \in \mathbb{Z}$, $k \in I_m$, and points $y_{m,k} \in \Omega_{m,k}$, a *sampling set* if the following conditions are satisfied:

(a) For each $m \in \mathbb{Z}$, the sets $\Omega_{m,k}$, $k \in I_m$, have disjoint interiors.
(b) $\mathbb{R}^n = \cup_{k \in I_m} \Omega_{m,k}$ for $m \in \mathbb{Z}$.
(c) Each set $\Omega_{m,k}$ satisfies $\Omega_{m,k} \subset B_\rho(x_{m,k}, c2^{-m})$ for some point $x_{m,k} \in \mathbb{R}^n$ ($c > 0$ is a constant).
(d) There exists a constant $c' > 0$ such that for any $m \in \mathbb{Z}$ and $k \in I_m$, we have $\rho(y_{m,k}, y_{m,k'}) > c'2^{-m}$ for all $k' \in I_m$, $k' \neq k$, except perhaps for a set of uniformly bounded number of points.

In the next theorem we present the discrete Calderón reproducing formula.

Theorem 6. [14] *Let $\{S_m\}_{m\in\mathbb{Z}}$ be an anisotropic approximation to the identity of order (μ,δ,r) with respect to the quasi-distance induced by an ellipsoid cover Θ of $\mathbb{R}^n$. Denote $D_m := S_{m+1} - S_m$ and let $\{\Omega_{m,k}\}$ and $\{y_{m,k}\}$ with $y_{m,k} \in \Omega_{m,k}$ be a sampling set for Θ. Then there exist $N > 0$ and linear operators $\{\hat{E}_m\}$ such that for any $f \in \mathscr{M}_0(\varepsilon,\gamma)$, $0 < \varepsilon,\gamma < \mu_0$,*

$$f = \sum_{m\in\mathbb{Z}} \sum_{k\in I_{m+N}} |\Omega_{m+N,k}| \hat{E}_m(f)(y_{m+N,k}) D_m(\cdot, y_{m+N,k}), \tag{41}$$

where the convergence is in $\mathscr{M}(\varepsilon',\gamma')$, $\varepsilon' < \varepsilon$, $\gamma' < \gamma$, and in $L_p(\mathbb{R}^n)$, $1 < p < \infty$. Furthermore, the kernels of $\{\hat{E}_m\}$ satisfy conditions (i)-(iii) of anisotropic approximations to the identity of order $(\mu,\varepsilon,1)$ for any $\varepsilon < \mu_0$ (with constants depending on ε) and the rth degree zero moments condition (37).

The proof of this theorem follows in the footsteps of the proof in the general case of homogeneous spaces (see e.g. [16]).

Definition of anisotropic wavet frames. We denote briefly $K_m := I_{m+N}$ and define the functions $\{\psi_{m,k}\}$ by

$$\psi_{m,k}(x) := |\Omega_{m+N,k}|^{1/2} D_m(x, y_{m+N,k})$$

and the functionals $\{\tilde{\psi}_{m,k}\}$ by

$$\tilde{\psi}_{m,k}(x) := |\Omega_{m+N,k}|^{1/2} \hat{E}_m(y_{m+N,k}, x), \quad m \in \mathbb{Z}, \ k \in K_m.$$

Then (41) takes the form

$$f = \sum_m \sum_{k \in K_m} \langle f, \tilde{\psi}_{m,k} \rangle \psi_{m,k}. \tag{42}$$

The next theorem shows that $\{\psi_{m,k}\}$, $\{\tilde{\psi}_{m,k}\}$ is a pair of dual frames.

Theorem 7. [14] *Let $\{S_m\}_{m\in\mathbb{Z}}$ be an anisotropic approximation to the identity of order (μ,δ,r). Denote $D_m := S_{m+1} - S_m$ and let $\{\Omega_{m,k}\}$ and $\{y_{m,k}\}$, $y_{m,k} \in \Omega_{m,k}$ be a sampling set for Θ. Then there exist constants $0 < A \le B < \infty$ such that for any $f \in L_2(\mathbb{R}^n)$*

$$A\|f\|_2^2 \le \sum_m \sum_{k \in K_m} |\langle f, \tilde{\psi}_{m,k} \rangle|^2 \le B\|f\|_2^2. \tag{43}$$

3.7 Two-level-split frames

We now use the two-level-split construction from §3.3 and the smooth duals $\{\tilde{\tilde{g}}_\lambda\}$ from §3.4 to derive a useful representation for the wavelet kernels $D_m(x,y)$.

For $\lambda = (\theta,\beta)$ we denote $\tilde{\tilde{g}}_{\theta,\beta} := \tilde{\tilde{g}}_\lambda$, where $\tilde{\tilde{g}}_\lambda$ is defined in (34). Then the kernel $S_m(x,y)$, defined in (36), has the representation

$$S_m(x,y) = \sum_{\theta \in \Theta_m} \sum_{|\beta|<r} \tilde{\tilde{g}}_{\theta,\beta}(y) P_{\theta,\beta} \varphi_\theta(x).$$

Now precisely as in §3.3 we get

$$\begin{aligned} D_m(x,y) &:= S_{m+1}(x,y) - S_m(x,y) \\ &= \sum_{\eta \in \Theta_{m+1}} \sum_{\theta \in \Theta_m : \theta \cap \eta \neq \emptyset} \sum_{|\beta|<r} \Big\{ \tilde{\tilde{g}}_{\eta,\beta}(y) - \sum_{|\alpha|<r} C^{\theta,\eta}_{\alpha,\beta} \tilde{\tilde{g}}_{\theta,\alpha}(y) \Big\} P_{\eta,\beta}(x) \varphi_\eta(x) \varphi_\theta(x), \end{aligned}$$

The new dual functions $\tilde{\tilde{F}}_\nu$, $\nu = (\eta,\theta,\beta) \in \mathscr{V}_m$, are defined by

$$\tilde{\tilde{F}}_\nu = \tilde{\tilde{F}}_{\eta,\theta,\beta} := \tilde{\tilde{g}}_{\eta,\beta} - \sum_{|\alpha|<r} C^{\theta,\eta}_{\alpha,\beta} \tilde{\tilde{g}}_{\theta,\alpha}. \tag{44}$$

Thus we arrive at the following representation

$$D_m(x,y) = \sum_{\nu \in \mathscr{V}_m} \tilde{\tilde{F}}_\nu(y) F_\nu(x).$$

Observe that since each $\theta \in \Theta_m$ is intersected by finitely many ellipsoids from Θ_{m+1} it follows by (35) that the duals $\{\tilde{\tilde{F}}_\nu\}$ have sub-exponential localization as the duals $\{\tilde{\tilde{g}}_\lambda\}$. Also, Theorem 2 and Proposition 3 imply that $\{F_\nu\}$, $\{\tilde{\tilde{F}}_\nu\}$ is a pair of dual frames.

Proposition 4. *For any $f \in L_2(\mathbb{R}^n)$*

$$\|f\|_2 \sim \Big(\sum_m \|D_m(f)\|_2^2\Big)^{1/2} \sim \Big(\sum_\nu \langle f, \tilde{\tilde{F}}_\nu\rangle^2\Big)^{1/2}.$$

4 Anisotropic Besov spaces (B-spaces)

In this section we review the anisotropic Besov spaces of positive smoothness induced by discrete ellipsoid covers of $\mathbb{R}^n$, introduced in [12], and compare them with the B-spaces based on anisotropic nested triangulations of $\mathbb{R}^2$ from [13, 20]. We will be mainly interested in the homogeneous versions of these spaces.

4.1 B-spaces induced by anisotropic covers of $\mathbb{R}^n$

Assuming that Θ is discrete ellipsoid cover of $\mathbb{R}^n$ (see Definition 1) we will define the homogeneous B-spaces $\dot{B}^\alpha_{pq}(\Theta)$ of positive smoothness $\alpha > 0$. In this definition there is a hidden parameter k which we choose to be the smallest integer satisfying the condition

$$k > \frac{a_0}{a_6} \cdot \frac{\alpha}{n}. \tag{45}$$

This will guarantee the equivalence of the norms in $\dot{B}^\alpha_{pq}(\Theta)$ introduced below. Here a_0 and a_6 are the constants from Definition 1, §2.1.

Definition of $\dot{B}^\alpha_{pq}(\Theta)$ via local moduli of smoothness. For $\alpha > 0$ and $0 < p, q \le \infty$ the space $\dot{B}^\alpha_{pq}(\Theta)$ is defined as the set of all functions $f \in L_p^{\mathrm{loc}}$ such that

$$\|f\|_{\dot{B}^\alpha_{pq}(\Theta)} := \Big(\sum_{m\in\mathbb{Z}} \Big(\sum_{\theta\in\Theta_m} |\theta|^{-\alpha p/n}\, \omega_k(f,\theta)_p^p\Big)^{q/p}\Big)^{1/q} < \infty, \tag{46}$$

where $\omega_k(f,\theta)_p$ is the kth local modulus of smoothness of f (see (20)).

This definition needs some additional clarification. Observe that $\|P\|_{\dot{B}^\alpha_{pq}(\Theta)} = 0$ for $P \in \mathscr{P}_k$ and hence the norm in $\dot{B}^\alpha_{pq}(\Theta)$ is a semi-(quasi-)norm and $\dot{B}^\alpha_{pq}(\Theta)$ is a quotient space modulo $\mathscr{P}_k$. We will use the operators Q_m and $T_{m,p}$ from §3.2 to construct a meaningful representation of each $f \in \dot{B}^\alpha_{pq}(\Theta)$. Let T_m ($m \in \mathbb{Z}$) be one of the operators Q_m or $T_{m,p}$ if $p \ge 1$, and $T_m := T_{m,p}$ if $p < 1$. We define

$$\|f\|^T_{\dot{B}^\alpha_{pq}(\Theta)} := \Big(\sum_{m\in\mathbb{Z}} \Big(2^{a_0 m\alpha/n} \|(T_{m+1} - T_m)f\|_p\Big)^q\Big)^{1/q}. \tag{47}$$

Proposition 2 and property (c) of ellipsoid covers imply

$$\|f - T_m f\|_p \le c\Big(\sum_{\theta\in\Theta_m} \omega_k(f,\theta)_p^p\Big)^{1/p}$$

and since $\|(T_{m+1}-T_m)f\|_p \le c\|f-T_{m+1}f\|_p + c\|f-T_m f\|_p$, we get

$$\|f\|^T_{\dot{B}^\alpha_{pq}(\Theta)} \le c\|f\|_{\dot{B}^\alpha_{pq}(\Theta)}. \tag{48}$$

For more precise description of $\dot{B}^\alpha_{pq}(\Theta)$ we have to distinguish between two basic cases.

Case 1: $0<\alpha<n/p$ or $\alpha = n/p$ and $q \le 1$. Then as is shown in [12] for any $f \in \dot{B}^\alpha_{pq}(\Theta)$ there exists a polynomial $P \in \mathscr{P}_k$ such that

$$f = \sum_{m\in\mathbb{Z}} (T_{m+1}-T_m)f + P \quad \text{ in } \; L_p(K) \tag{49}$$

for all compact sets $K \subset \mathbb{R}^n$.

Case 2: $\alpha > n/p$ or $\alpha = n/p$ and $q>1$. Now the space $\dot{B}^\alpha_{pq}(\Theta)$ can be viewed as the set of all regular tempered distributions f such that $\|f\|_{\dot{B}^\alpha_{pq}(\Theta)} < \infty$ and

$$f = \sum_{m\in\mathbb{Z}} (T_{m+1}-T_m)f,$$

where the convergence is in $\mathscr{S}'/\mathscr{P}_k$. This means that there exist polynomials $P \in \mathscr{P}_k$ and $P_m \in \mathscr{P}_k$, $m\in\mathbb{Z}$, such that

$$f = P + \lim_{j\to-\infty} \sum_{m=j}^{\infty} (T_{m+1}-T_m)f + P_m \quad \text{in} \quad \mathscr{S}'.$$

In addition, $\dot{B}^\alpha_{pq}(\Theta)$ is continuously embedded in $\mathscr{S}'$.

Other norms in $\dot{B}^\alpha_{pq}(\Theta)$. The good understanding of the B-spaces depends on having several equivalent norms in $\dot{B}^\alpha_{pq}(\Theta)$. Note that if $\{d_\nu(f)\}$ are defined from $(T_{m+1}-T_m)f = \sum_{\nu\in\mathscr{V}_m} d_\nu(f)F_\nu$, then using Theorem 2

$$\|f\|^T_{\dot{B}^\alpha_{pq}(\Theta)} \sim \Big(\sum_{m\in\mathbb{Z}}\Big(\sum_{\nu\in\mathscr{V}_m}(|\eta_\nu|^{-\alpha/n}\|d_\nu(f)F_\nu\|_p)^p\Big)^{q/p}\Big)^{1/q}. \tag{50}$$

Observe that the above equivalence holds if $d_\nu(f)$ are replaced by $\langle f, \tilde{\tilde{F}}_\nu\rangle$ due to the sub-exponential localization of the duals $\{\tilde{\tilde{F}}_\nu\}$.

Also, we define

$$\|f\|^A_{\dot{B}^\alpha_{pq}(\Theta)} := \inf_{f=\sum_{\nu\in\mathscr{V}} a_\nu F_\nu} \Big(\sum_{m\in\mathbb{Z}}\Big(\sum_{\nu\in\mathscr{V}_m}(|\eta_\nu|^{-\alpha/n}\|a_\nu F_\nu\|_p)^p\Big)^{q/p}\Big)^{1/q}. \tag{51}$$

Here the infimum is taken over all representations $f = \sum_{\nu\in\mathscr{V}} a_\nu F_\nu$, where the convergence is to be understood as described in Cases 1-2 above.

In the next theorem we record the equivalence of the above norms.

Theorem 8. [12] *If $\alpha > 0$, $0 < p,q \le \infty$, and condition (45) is satisfied, then the norms $\|\cdot\|_{\dot{B}^\alpha_{pq}(\Theta)}$, $\|\cdot\|^T_{\dot{B}^\alpha_{pq}(\Theta)}$, and $\|\cdot\|^A_{\dot{B}^\alpha_{pq}(\Theta)}$ are equivalent.*

The embedding of $\dot{B}^\alpha_{pq}$ in $\mathscr{S}'$ or (49) readily imply the completeness of $\dot{B}^\alpha_{pq}(\Theta)$.

Inhomogeneous B-spaces. Sometimes it is more convenient to use the inhomogeneous versions $B^\alpha_{pq}(\Theta^+)$ of the B-spaces induced by anisotropic ellipsoid covers of $\mathbb{R}^n$, which are simpler than the homogeneous counterparts $\dot{B}^\alpha_{pq}(\Theta)$.

For the definition of the inhomogeneous spaces $B^\alpha_{pq}(\Theta^+)$ one only needs ellipsoid covers with levels $m = 0, 1, \ldots$, i.e. covers of the form

$$\Theta^+ := \bigcup_{m=0}^{\infty} \Theta_m.$$

The space $B^\alpha_{pq}(\Theta^+)$, $\alpha > 0$, $0 < p,q \le \infty$, is defined as the set of all functions $f \in L_p(\mathbb{R}^n)$ such that

$$|f|_{B^\alpha_{pq}(\Theta^+)} := \Big(\sum_{m\ge 0}\Big(\sum_{\theta\in\Theta_m}(|\theta|^{-\alpha p/n}\,\omega_k(f,\theta)_p^p\Big)^{q/p}\Big)^{1/q} < \infty, \tag{52}$$

where $\omega_k(f,\theta)_p$ is the kth local modulus of smoothness of f in $L_p(\theta)$.

The (quasi-)norm in $B^\alpha_{pq}(\Theta^+)$ is defined by

$$\|f\|_{B^\alpha_{pq}(\Theta^+)} := \|f\|_p + |f|_{B^\alpha_{pq}(\Theta^+)}.$$

Other equivalent norms in $B^\alpha_{pq}(\Theta^+)$ can be defined similarly as for the homogeneous B-spaces from above. In particular, using the notation from from Theorem 3 one has

$$\|f\|_{B^\alpha_{pq}(\Theta^+)} \sim \Big(\sum_{m\ge -1}\Big(\sum_{\nu\in\mathscr{V}_m}(|\eta_\nu|^{-\alpha/n}\|d_\nu(f)F_\nu\|_p)^p\Big)^{q/p}\Big)^{1/q}. \tag{53}$$

For more details about anisotropic B-spaces induced by ellipsoid covers and proofs we refer the reader to [12].

4.2 B-spaces induced by nested multilevel triangulations of $\mathbb{R}^2$

We first recall briefly some basic definitions and facts from [20, 13].

Spline multiresolution analysis (MRA). Let $\mathscr{T} = \bigcup_{m\in\mathbb{Z}} \mathscr{T}_m$ be an SLR-triangulation of $\mathbb{R}^2$ (see §2.2). Denote by V_m the set of all vertices of triangles from $\mathscr{T}_m$.

For $r \ge 0$ and $k \ge 2$, we denote by $S^{k,r}_m = S^{k,r}(\mathscr{T}_m)$ the set of all r times differentiable piecewise polynomial functions of degree $< k$ over $\mathscr{T}_m$, i.e. $s \in S^{k,r}_m$ if $s \in C^r(\mathbb{R}^2)$ and $s = \sum_{\triangle\in\mathscr{T}_m} \mathbb{1}_\triangle \cdot P_\triangle$ with $P_\triangle \in \mathscr{P}_k$.

It will be convenient to denote, for any vertex $v \in V_m$, by $\mathrm{Star}^1(v)$ the union of all triangles $\triangle \in \mathscr{T}_m$ attached to v. Inductively for $\ell \geq 2$, we define $\mathrm{Star}^\ell(v)$ as the union of $\mathrm{Star}^{\ell-1}(v)$ and the stars of all vertices of $\mathrm{Star}^{\ell-1}(v)$.

We assume that for each $m \in \mathbb{Z}$ there exists a subspace S_m of $S_m^{k,r}$ and a family $\Phi_m = \{\varphi_\theta : \theta \in \Theta_m\} \subset S_m$ satisfying the following conditions:

(a) $S_m \subset S_{m+1}$ and $\mathscr{P}_{\tilde{k}} \subset S_m$, for some $1 \leq \tilde{k} \leq k$,

(b) Φ_m is a stable basis for S_m in L_p $(1 \leq p \leq \infty)$,

(c) For every $\theta \in \Theta_m$ there is a vertex $v_\theta \in V_m$ such that φ_θ and its dual are supported on $\mathrm{Star}^\ell(v_\theta)$, where $\ell \geq 1$ is a constant independent of θ and m.

We denote $\Phi := \bigcup_{m\in\mathbb{Z}} \Phi_m$ and $\Theta := \bigcup_{m\in\mathbb{Z}} \Theta_m$.

A simple example of spline MRA is the sequence $\{S_m\}_{m\in\mathbb{Z}}$ of all continuous piecewise linear functions ($r = 0$, $k = 2$) on the levels $\{\mathscr{T}_m\}_{m\in\mathbb{Z}}$ of a given SLR-triangulation $\mathscr{T}$ of $\mathbb{R}^2$. A basis for each space S_m is given by the set Φ_m of the Courant elements φ_θ, supported on the cells θ of $\mathscr{T}_m$ (θ is the union of all triangles of $\mathscr{T}_m$ attached to a vertex, say, v_θ). The function φ_θ takes value 1 at v_θ and 0 at all other vertices.

A concrete construction of a spline MRA for an arbitrary SLR-triangulation $\mathscr{T}$ is given in [13], where $S_m = S_m^{k,r} = S^{k,r}(\mathscr{T}_m)$ for given $r \geq 1$ and $k > 4r+1$.

Local spline approximation. For $\triangle \in \mathscr{T}_m$ we set

$$\Omega_\triangle^\ell := \cup\{\mathrm{Star}^\ell(v) : v \in V_m,\ \triangle \subset \mathrm{Star}^\ell(v)\}.$$

We now let $\mathbb{S}_\triangle(f)_p$ denote the error of L_p-approximation from S_m on $\Omega_\triangle^\ell$, i.e.

$$\mathbb{S}_\triangle(f)_p := \inf_{s\in S_m} \|f - s\|_{L_p(\Omega_\triangle^\ell)}. \tag{54}$$

Definition of $\dot{\mathscr{B}}_{pq}^\alpha(\Phi)$. Given a spline MRA $\{S_m\}_{m\in\mathbb{Z}}$ over an SLR-tiangulation $\mathscr{T}$ of $\mathbb{R}^2$ and an associated family of basis functions Φ, as described above, we define the B-space $\dot{\mathscr{B}}_{pq}^\alpha(\Phi)$, $\alpha > 0$, $0 < p,q \leq \infty$, as the set of all $f \in L_p^{\mathrm{loc}}(\mathbb{R}^2)$ such that

$$\|f\|_{\dot{\mathscr{B}}_{pq}^\alpha(\Phi)} := \left(\sum_{m\in\mathbb{Z}} \left[2^{m\alpha}\left(\sum_{\triangle\in\mathscr{T},\, 2^{-m}\leq|\triangle|<2^{-m+1}} \mathbb{S}_\triangle(f)_p^p\right)^{1/p}\right]^q\right)^{1/q} < \infty \tag{55}$$

with the ℓ_q-norm replaced by the sup-norm if $q = \infty$.

4.3 Comparison of different B-spaces and Besov spaces

The most substantial distinction between $\dot{B}_{pq}^\alpha(\Theta)$ and $\dot{\mathscr{B}}_{pq}^\alpha(\Phi)$ is that the spaces $\dot{B}_{pq}^\alpha(\Theta)$ are defined via *local polynomial* approximation $\sim \omega_k(f,\theta)_p$, while $\dot{\mathscr{B}}_{pq}^\alpha(\Phi)$ are defined via *local spline* approximation: $\mathbb{S}_\triangle(f)_p$. As a result, loosely speaking the spaces $\dot{B}_{pq}^\alpha(\Theta)$ have larger norms that the spaces $\dot{\mathscr{B}}_{pq}^\alpha(\Phi)$. However, if $\mathbb{S}_\triangle(f)_p$ in (55) is replaced by $\omega_k(f,\Omega_\triangle^1)_p$ then the resulting quantity would be equivalent to

$\|f\|_{\dot{B}^{\alpha}_{pq}(\Theta)}$, where Θ is the ellipse cover of $\mathbb{R}^2$ obtained by dilating the minimum area circumscribed ellipses for all triangle $\triangle \in \mathscr{T}$ as mentioned in §2.2.

Another important distinction between $\dot{B}^{\alpha}_{pq}(\Theta)$ and $\dot{\mathscr{B}}^{\alpha}_{pq}(\Phi)$ is that the underlying multilevel triangulation for the later space is nested, while the ellipsoid cover generating the former is not nested. Therefore, in constructing ellipsoid cover and dealing with B-spaces $\dot{B}^{\alpha}_{pq}(\Theta)$ one has much more freedom.

It is quite easy to show that (see [11]) if Θ is an ellipsoid cover of $\mathbb{R}^n$ consisting of Euclidean balls, then the B-spaces $\dot{B}^{\alpha}_{pq}(\Theta)$ are the same as the respective classical Besov spaces $\dot{B}^{\alpha}_{q}(L_p)$ (with equivalent norms). We maintain that local moduli of smoothness rather than global ones are more natural for the definition of anisotropic (and even classical) Besov spaces of positive smoothness since they more adequately reflect the nature of the spaces. For the theory of (classical) Besov spaces we refer the reader to [23, 26].

As already mentioned the powers A^j of a real $n \times n$ matrix A with eigenvalues λ obeying $|\lambda| > 1$ generate a semi-continuous and hence discrete ellipsoid cover of $\mathbb{R}^n$. It can be shown that for $\alpha > n(1/p-1)_+$ the associated B-spaces $\dot{B}^{\alpha}_{pq}$ are exactly the same (with equivalent norms) as the anisotropic Besov spaces (with weight 1) developed in [3].

As indicated in §2.1, $\mathbb{R}^n$ equipped with the distance $\rho(\cdot,\cdot)$ introduced in Definition 2 and the Lebesgue measure is a space of homogeneous type and hence the general theory of Besov spaces on homogeneous spaces applies (see e.g. [19]). In fact, in the specific setting of this paper the anisotropic Besov spaces given by the general theory are the same as the B-spaces from here for sufficiently small $\alpha > 0$. The main distinction between the two theories is that we can handle B-spaces of an arbitrary smoothness $\alpha > 0$, while the general theory of Besov spaces on homogeneous spaces is only feasible for smoothness α with $|\alpha| < \varepsilon$ for some sufficiently small ε.

5 Nonlinear approximation

One of the main applications of the anisotropic B-spaces is to nonlinear N-term approximation from the two-level-split bases introduced in §3.3, which is the purpose of this section. We will also compare here the two-level-split bases with anisotropic hierarchical spline bases as tools for nonlinear approximation.

The B-spaces of nonlinear approximation. A particular type of B-spaces plays an important role in nonlinear N-term approximation in L_p. Given $0 < p < \infty$ and $\alpha > 0$ let τ be defined by

$$1/\tau = \alpha/n + 1/p, \tag{56}$$

which in the case of classical Besov spaces signifies the critical embedding in L_p. For nonlinear approximation in $L_\infty := C_0$ τ is determined by $1/\tau = \alpha/n$ and necessarily $\alpha \geq 1$ (otherwise the embedding (60) below is not valid).

For a given discrete ellipsoid cover Θ of $\mathbb{R}^n$, the homogeneous B-spaces $\dot{B}^{\alpha}_{\tau}(\Theta) := \dot{B}^{\alpha}_{\tau\tau}(\Theta)$ are of a particular importance in nonlinear approximation from the two-level-split bases. From (46) we have

$$\|f\|_{\dot{B}^{\alpha}_{\tau}(\Theta)} := \Big(\sum_{\theta\in\Theta} |\theta|^{-\alpha\tau/n}\omega_k(f,\theta)^{\tau}_{\tau}\Big)^{1/\tau}. \tag{57}$$

Observe that in general $\tau < 1$, however, just as in [20] it can be showen that for any $0 < q < p$

$$\|f\|_{\dot{B}^{\alpha}_{\tau}(\Theta)} \sim \Big(\sum_{\theta\in\Theta} |\theta|^{(1/p-1/q)\tau}\omega_k(f,\theta)^{\tau}_{q}\Big)^{1/\tau}. \tag{58}$$

This allows to work in L_q with $q \ge 1$ if $p > 1$ instead of L_τ.

The key point here is that the norm in $\dot{B}^{\alpha}_{\tau}(\Theta)$ has the representation

$$\|f\|_{\dot{B}^{\alpha}_{\tau}(\Theta)} \sim \Big(\sum_{\nu\in\mathscr{V}} \|d_\nu(f)F_\nu\|^{\tau}_{p}\Big)^{1/\tau}, \quad \mathscr{V} := \cup_{m\in\mathbb{Z}}\mathscr{V}_m, \tag{59}$$

which implies the embedding of $\dot{B}^{\alpha}_{\tau}(\Theta)$ in L_p: Every $f \in \dot{B}^{\alpha}_{\tau}(\Theta)$ can be identified modulo $\mathscr{P}_k$ as a function in $L_p(\mathbb{R}^n)$ such that

$$\|f\|_p \le c\|f\|_{\dot{B}^{\alpha}_{\tau}(\Theta)}. \tag{60}$$

This identification will always be assumed in what follows. In fact, the above shows that $\dot{B}^{\alpha}_{\tau}(\Theta)$ lies on the Sobolev embedding line.

The situation is quite the same for the inhomogeneous B-spaces $B^{\alpha}_{\tau} := B^{\alpha}_{\tau\tau}(\Theta^+)$ associated with a discrete ellipsoid cover $\Theta^+ = \cup_{m\ge 0}\Theta_m$ of $\mathbb{R}^n$.

Nonlinear N-term approximation from $\mathscr{F} := \cup_{m\in\mathbb{Z}}\mathscr{F}_m = \{F_\nu : \nu \in \mathscr{V}\}$. We let $\mathscr{E}_N$ denote the nonlinear set of all functions g of the form

$$g = \sum_{\nu\in\Gamma_N} a_\nu F_\nu,$$

where $\Gamma_N \subset \mathscr{V}$, $\#\Gamma \le N$, and Γ is allowed to vary with g. Then the error $\sigma_N(f)_p$ of best L_p-approximation of $f \in L_p(\mathbb{R}^n)$ from $\mathscr{E}_N$ is defined by

$$\sigma_N(f)_p := \inf_{g\in\mathscr{E}_N} \|f-g\|_p.$$

Theorem 9 (Jackson estimate). *If $f \in \dot{B}^{\alpha}_{\tau}(\Theta)$, $\alpha > 0$, $0 < p \le \infty$, then*

$$\sigma_N(f)_p \le cN^{-\alpha/n}\|f\|_{\dot{B}^{\alpha}_{\tau}(\Theta)}, \tag{61}$$

where c depends only on α, p, and the parameters of Θ.

When $0 < p < \infty$, estimate (61) follows by the general Theorem 3.4 in [20] and in the case $p = \infty$ its proof can be carried out as the proof of Theorem 3.1 in [21].

In a standard way the Jackson estimate (61) leads to a direct estimate for nonlinear N-term approximation from $\mathscr{F}$ which involves the K-functional between L_p and $\dot{B}^{\alpha}_{\tau}(\Theta)$. It is a challenging *open problem* to prove a companion inverse estimate due to the fact that $\mathscr{F}$ is possibly redundant and nonnested.

Comparison with nonlinear N-term approximation from nested spline bases. Nonlinear N-term approximation in L_p $(0 < p \le \infty)$ from the spline basis elements in $\Phi = \cup_{m\in\mathbb{Z}}\Phi_m$ (see §4.2) has been developed in [20, 13, 21, 10]. In [20, 13] Jackson and Bernstein estimates are established involving the B-spaces $\dot{\mathscr{B}}^{\alpha}_{\tau}(\Phi) := \dot{\mathscr{B}}^{\alpha}_{\tau\tau}(\Phi)$ with norm

$$\|f\|_{\dot{\mathscr{B}}^{\alpha}_{\tau}(\Phi)} := \Big(\sum_{\triangle\in\mathscr{T}} (|\triangle|^{-\alpha}\mathbb{S}_{\triangle}(f)_{\tau})^{\tau}\Big)^{1/\tau}, \tag{62}$$

where $1/\tau := \alpha + 1/p$ for $\alpha > 0$ if $0 < p < \infty$ and $\alpha \ge 1$ if $p = \infty$. Then the standard machinery of Approximation theory is used to characterize the respective approximation spaces as real interpolation spaces between L_p and $\dot{\mathscr{B}}^{\alpha}_{\tau}(\Phi)$.

The most important difference between the nonlinear N-term approximation from $\mathscr{F}$ and Φ is that the spaces $\dot{\mathscr{B}}^{\alpha}_{\tau}(\Phi)$ (defined by local spline approximation) are specifically designed for the purposes of nonlinear spline approximation and allow to characterize the rates of approximation $O(N^{-\beta})$ for all $\beta > 0$, while in the former case β is limited. On the other hand, the spaces $\dot{B}^{\alpha}_{\tau}(\Theta)$ are of more general nature and are direct generalization of Besov spaces. They are much less sensitive to changes in the underlying ellipsoid cover Θ compared to changes in $\dot{\mathscr{B}}^{\alpha}_{\tau}(\Phi)$ when changing the respective triangulation $\mathscr{T}$. In general, the spaces $\dot{B}^{\alpha}_{\tau}(\Theta)$ are better than $\dot{\mathscr{B}}^{\alpha}_{\tau}(\Phi)$ as a tool for measuring the anisotropic features of functions (see below).

6 Measuring smoothness via anisotropic B-spaces

It has always been a question in analysis how to measure the smoothness of a given function, and as a consequence, there is a variety of smoothness space. We next show how the anisotropic B-spaces $\dot{B}^{\alpha}_{\tau}(\Theta)$ can be deployed to measuring the smoothness of functions and how this is related to nonlinear N-term approximation from the two-level-split bases.

We focus on two "simple" examples of discontinuous functions on $\mathbb{R}^2$, namely, $\mathbb{1}_{B(0,1)}$ the characteristic function of the unit disk $B(0,1)$ and $\mathbb{1}_Q$ the characteristic function of a square $Q \subset \mathbb{R}^2$. As shown in [12] each of these functions has higher order smoothness α in $\dot{B}^{\alpha}_{\tau}(\Theta)$ for an appropriately selected ellipse cover Θ compared with its (classical) Besov space smoothness. Moreover, their smoothness via suitable covers will be seen to differ substantially.

As in the previous section, for given $0 < p < \infty$ and $\alpha > 0$, let τ be defined by $1/\tau = \alpha/2 + 1/p$.

Theorem 10. [12] (i) *There exists an anisotropic ellipsoid cover* Θ *of* $\mathbb{R}^2$ *such that* $\mathbb{1}_{B(0,1)} \in \dot{B}^{\alpha}_{\tau}(\Theta)$ *for any* $\alpha < 4/p$. *In comparison, in the scale of Besov spaces* $\dot{B}^{\alpha}_{\tau\tau}$ *one has* $\mathbb{1}_{B(0,1)} \in \dot{B}^{\alpha}_{\tau\tau}$ *for* $\alpha < 2/p$. *Here the bounds for* α *are sharp.*

(ii) *For any square* Q *in* $\mathbb{R}^2$ *and any* $\alpha > 0$ *there exists an anisotropic ellipsoid cover* Θ *of* $\mathbb{R}^2$ *such that* $\mathbb{1}_Q \in \dot{B}^{\alpha}_{\tau}(\Theta)$, *while in the scale of Besov spaces* $\dot{B}^{\alpha}_{\tau\tau}$ *one has only* $\mathbb{1}_Q \in \dot{B}^{\alpha}_{\tau\tau}$ *for* $\alpha < 2/p$ *and this bound for* α *is sharp.*

This theorem coupled with the Jackson estimate (61) leads to the following approximation result.

Corollary 2. [12] (i) *There exists a discrete ellipse cover* Θ *of* $\mathbb{R}^2$ *such that for any* $0 < p < \infty$ *the nonlinear N-term approximation from* $\mathscr{F}_\Theta$ *satisfies*

$$\sigma_N(\mathbb{1}_{B(0,1)})_p \le cN^{-\gamma} \quad \text{for all} \quad \gamma < 2/p.$$

(ii) *For any* $\alpha > 0$ *there exists a discrete ellipse cover* Θ *of* $\mathbb{R}^2$ *such that for any* $0 < p < \infty$ *the nonlinear N-term approximation from* $\mathscr{F}_\Theta$ *satisfies*

$$\sigma_N(\mathbb{1}_Q)_p \le cN^{-\alpha}.$$

For comparison, if $\sigma_m^W(f)_p$ *denotes the best N-term approximation of* f *in* L_p $(p \ge 1)$ *from any reasonable wavelet basis, then for* $E = B(0,1)$ *or* $E = Q$

$$\sigma_N^W(\mathbb{1}_E)_p \le cN^{-\gamma} \quad \text{for all} \quad \gamma < 1/p.$$

All estimates above are sharp.

Discussion. As indicated above for appropriate ellipse covers, the B-space smoothness of the characteristic functions of the unit ball and any square in $\mathbb{R}^2$ is higher than their Besov space smoothness. Thus by using adaptive dilations the anisotropic B-spaces are better able to resolve the singularities along smooth or piecewise smooth curves. Consequently, the two-level-split decompositions of these functions are substantially sparser than their wavelet decompositions, which leads to better rates of nonlinear N-term approximation. It might surprise that characteristic functions of polygonal domains have, in a sense, infinite smoothness while those of domains with smooth boundaries have limited regularity. However, the covers that yield higher and higher smoothness in the polygonal case have to become less and less constrained, which means that the parameters in $\mathbf{p}(\Theta)$ are subjected to more and more generous bounds. Keeping these parameters within a compact set would limit the regularity that could be described in this way.

The above two examples illustrate clearly the concept of measuring the smoothness of functions via anisotropic B-space and in particular by the B-spaces of nonlinear approximation $\dot{B}^{\alpha}_{\tau}(\Theta)$. The key idea is to allow the underlying ellipsoid cover to adapt to the given function.

It is a challenging *open problem* to devise a scheme which for a given function f finds an optimal (or near optimal) ellipsoid cover Θ such that f exhibits the highest order α of smoothness in $\dot{B}^{\alpha}_{\tau}(\Theta)$ in the above sense.

7 Application to Preconditioning for Elliptic Boundary Value Problems

In this section we apply the two-level-split bases from §3.3 in a regular set-up to the development of multilevel Schwarz preconditioners for elliptic boundary value problems. We consider the following model problem. Let $a(\cdot,\cdot): V \times V \to \mathbb{R}$ be a symmetric bilinear form on a Hilbert space V with norm $\|\cdot\|_V = \langle \cdot,\cdot \rangle^{1/2}$ that is V-elliptic, i.e. there exist positive constants c_a, C_a such that

$$a(v,v) \geq c_a \|v\|_V^2, \quad |a(v,w)| \leq C_a \|v\|_V \|w\|_V, \quad v,w \in V. \tag{63}$$

The problem is, for a given $f \in V'$ to find $u \in V$ such that

$$a(u,v) = \langle f, v \rangle, \quad \forall\ v \in V. \tag{64}$$

For simplicity we only consider the model case $V = H_0^1(\Omega)$ corresponding to Dirichlet boundary conditions. Higher order problems could be treated in an analogous way. We assume that Ω is a bounded *extension* domain, which means that Ω has a sufficiently regular boundary to permit any element v of any Sobolev or Besov space $X(\Omega)$ over Ω to be extended to $\tilde{v} \in X(\mathbb{R}^n)$, $\tilde{v}|_\Omega = v$, so that $\|v\|_{X(\mathbb{R}^n)} \leq C_X \|v\|_{X(\Omega)}$. This is e.g. the case when the boundary of Ω is piecewise smooth and Ω obeys a uniform cone condition. The homogeneous boundary conditions are supposed to be realized in the trial spaces by suitable polynomial factors in the atoms.

We assume that $\Theta = \cup_{m \geq -1} \Theta_m$ is a regular multilevel cover of $\mathbb{R}^n$ consisting of balls. We will utilize the atoms $\{F_\gamma\}$ defined in §3.3 for $\gamma \in \mathscr{V} = \bigcup_{m=-1}^{\infty} \mathscr{V}_m$, see Theorem 3. For better notation we will index the elements F_γ of the two-level-split bases $\mathscr{F}_m$ by γ instead of v as before.

We will put this in the context of *stable splittings* in the theory of *multilevel Schwarz preconditioners*, see e.g. [22, 27].

Let $V_\gamma := \operatorname{span}(F_\gamma)$, so that $H_0^1(\Omega) := V = \sum_\gamma V_\gamma$. The key fact is that $\{V_\gamma\}_{\gamma \in \mathscr{V}}$ form a *stable splitting* for V:

Theorem 11. *There exist constants $c_V, C_V > 0$ such that for any $v \in V$*

$$c_V \|v\|_V \leq \inf_{v = \sum_\gamma v_\gamma} \left(\sum_{\gamma \in \mathscr{V}} |\eta_\gamma|^{-2/d} \|v_\gamma\|_2^2 \right)^{1/2} \leq C_V \|v\|_V. \tag{65}$$

Moreover, $\{V_\gamma\}_{\gamma \in \mathscr{V}^\ell}$ with $\mathscr{V}^\ell := \bigcup_{m=-1}^{\ell} \mathscr{V}_m$ form a uniformly stable splitting for the spaces $S_m := \operatorname{span}(\Phi_m)$ in the sense of (65) *with the same constants c_V, C_V.*

Using that the norms $a(\cdot,\cdot)^{1/2}$ and $\|\cdot\|_{H^1(\Omega)}$ are equivalent and the well known fact that $\|\cdot\|_{H^1(\Omega)} \sim \|\cdot\|_{B_2^1(L_2(\Omega))}$, estimates (65) are immediate from Theorem 8 taking into account that Besov and B-norms are equivalent in the regular setting. The second part of Theorem 11 follows from the fact that the telescoping expansions underlying the inhomogeneous version of $\|\cdot\|^T_{B^\alpha(\Theta)}$ (see (47) and (53)) terminate without affecting this norm. For more details, see [11].

This allows us to apply the theory of Schwarz methods along the following lines. For $V_0 := S_0 = \mathrm{span}(\Phi_0)$ define $P_{V_0} : V \to V_0$ and $r_{V_0} \in S_0$ by

$$a(P_{V_0}v, F_\gamma) = a(v, F_\gamma), \quad (r_{V_0}, F_\gamma)_{L_2} = \langle f, F_\gamma\rangle, \;\; \gamma \in \mathscr{V}_0 = \Theta_0.$$

Furthermore, we introduce the auxiliary bilinear forms:

$$b_\gamma(v,w) := |\eta_\gamma|^{-2/d}(v,w)_{L_2}, \;\; v,w \in V_\gamma, \;\; \gamma \in \mathscr{V} \setminus \mathscr{V}_0. \tag{66}$$

We now consider the spaces V_γ with norms $\|v\|_{V_\gamma} := (b_\gamma(v,v))^{1/2}$ and define the linear operators $P_{V_\gamma} : V \to V_\gamma$ and $f_\gamma \in V_\gamma$ by

$$\begin{aligned} |\eta_\gamma|^{-2/d}(P_{V_\gamma}v, F_\gamma)_{L_2} &= a(v, F_\gamma), \\ |\eta_\gamma|^{-2/d}(f_\gamma, F_\gamma)_{L_2} &= \langle f, F_\gamma\rangle. \end{aligned} \tag{67}$$

Thus, as usual,

$$P_{V_\gamma}v = a_\gamma(v)F_\gamma, \qquad f_\gamma = r_\gamma(f)F_\gamma, \tag{68}$$

where

$$a_\gamma(v) = \frac{|\eta_\gamma|^{2/d}a(v,F_\gamma)}{\langle F_\gamma, F_\gamma\rangle}, \quad r_\gamma(f) = \frac{|\eta_\gamma|^{2/d}\langle f, F_\gamma\rangle}{\langle F_\gamma, F_\gamma\rangle}. \tag{69}$$

The following theorem now is an immediate consequence of the results in [18, 22].

Theorem 12. *Problem* (64) *is equivalent to the operator equation*

$$P_V u = \overline{f}, \quad \textit{where} \tag{70}$$

$$P_V := P_{V_0} + \sum_{\gamma \in \mathscr{V}\setminus\mathscr{V}_0} P_{V_\gamma}, \quad \overline{f} := r_{V_0} + \sum_{\gamma \in \mathscr{V}\setminus\mathscr{V}_0} f_\gamma.$$

Moreover, the spectral condition number $\kappa(P_V)$ *of the additive Schwarz operator* P_V *satisfies*

$$\kappa(P_V) \le \frac{C_a C_V}{c_a c_V}, \tag{71}$$

where c_a, C_a, c_V, C_V *are the constants from* (63) *and* (65).

Estimate (71) yields that simple iterative schemes, such as Richardson iterations,

$$u^{n+1} = u^n + \alpha(\overline{f} - P_V u^n), \quad n = 0,1,2,\ldots, \tag{72}$$

converge with a fixed error reduction rate per step.

We conclude with a few remarks. First, the operator equation (70) is formulated in the full infinite dimensional space. Alternatively, restricting the summation to a finite subset $\overline{\mathscr{V}}$ of $\mathscr{V}$ (e.g. $\overline{\mathscr{V}} = \mathscr{V}^\ell$), we obtain a finite dimensional discrete problem whose condition fulfills (on account of Theorem 11) the same bound uniformly in the size and choice of $\overline{\mathscr{V}}$. In this sense our preconditioner is asymptotically optimal.

On the other hand, it is conceptually useful to consider the full infinite dimensional problem (70). Then (72) has to be understood as an *idealized* scheme whose numerical implementation requires appropriate *approximate* applications of the (infinite dimensional) operator P_V quite in the spirit of [7]. This can be done by computing in addition to solving the coarse scale problem on $S_0 = V_0$ only finitely many but properly selected components P_{V_γ} each requiring only the solution of a one-dimensional problem. This hints at the adaptive potential of such an approach similar to the developments in [7]. This, in turn, raises the question what accuracy can be achieved at best when using linear combinations of at most N of the atoms. Thus we arrive at the problem for nonlinear N-term approximation from $\{F_\gamma\}$ in H^1.

For more details we refer the reader to [11].

Acknowledgements The second author has been supported by NSF Grant DMS-0709046.

References

1. Alani, A., Averbuch, A., Dekel, S.:, Image coding using geometric wavelets, IEEE Trans. Image Process. **16**, 69–77 (2007)
2. Bownik, M.: Anisotrpopic Hardy spaces and wavelets. Mem. Amer. Math. Soc. **164**, No 781 (2003)
3. Bownik, M.: Atomic and molecular decompositions of anisotrpopic Besov spaces. Math. Z. **250**, 539–571 (2005)
4. Bownik, M., Ho, K.-P.: Atomic and molecular decompositions of anisotrpopic Triebel-Lizorkin spaces. Trans. Amer. Math. Soc. **358**, 1469–1510 (2006)
5. Candès, E., Demanet, L., Donoho, D., Ying, L.: Fast Discrete Curvelet Transforms, Multiscale Model. Simul. **5**, 861-899 (2006)
6. Candès, E., Donoho, D.: Continuous Curvelet Transform: I. Resolution of the Wavefront Set, Appl. Comput. Harmon. Anal. **19**, 162-197 (2005)
7. Cohen, A., Dahmen, W., DeVore, R.: Adaptive wavelet methods II - Beyond the elliptic case. Found. Comput. Math. **2**, 203–245 (2002)
8. Coifman, R., Weiss, G.: Analyse harmonique non-comutative sur certains espaces homogenes. Lecture Notes in Math. **242**, Springer-Verlag (1971)
9. Coifman, R., Weiss, G.: Extensions of Hardy spaces and their use in analysis. Bull. Amer. Math. Soc. **83**, 569–645 (1977)
10. Dahmen, D., Petrushev, P.: "Push-the-Error" algorithm for nonlinear n-term approximation. Constr. Approx. **23**, 261-304 (2006)
11. Dahmen, W., Dekel, S., Petrushev, P.: Multilevel Preconditioning for Partition of Unity Methods - Some Analytic Concepts. Numer. Math. **107**, 503–532 (2007)
12. Dahmen, W., Dekel, S., Petrushev, P.: Two-level-split Decomposition of Anisotropic Besov Spaces. Constr. Approx. (to appear)
13. Davydov, O., Petrushev, P.: Nonlinear approximation from differentiable piecewise polynomials. SIAM J. Math. Anal. **35**, 708–758 (2003)
14. Dekel, D., Han, Y., Petrushev, P.: Anisotropic meshless frames on $\mathbb{R}^n$. J. Fourier Anal. and Appl. (to appear)
15. Dekel, D., Leviatan, D.: Adaptive multivariate approximation using binary space partitions and geometric wavelets, SIAM J. Numer. Anal. **43**, 707–732 (2005)
16. Deng, D., Han, Y.: Harmonic analysis on spaces of homogeneous type. Lecture Notes in Math. **1966** (2009)

17. Folland, G., Stein, E.: Hardy spaces on homogeneous groups. Princeten University Press, N. J. (1982)
18. Griebel, M., Oswald, P.: Remarks on the abstract theory of additive and multiplicative Schwarz methods. Numer. Math. **70**, 163–180 (1995)
19. Han, H., Sawyer, E.: Littlewood-Paley theory on spaces of homogeneous type and classical function spaces. Mem. Amer. Math. Soc. **530** (1994)
20. Karaivanov, K., Petrushev, P.: Nonlinear piecewise polynomial approximation beyond Besov spaces. Appl. Comput. Harmon. Anal. **15**, 177–223 (2003)
21. Karaivanov, K., Petrushev, P., Sharpley, R.C.: Algorithms for nonlinear piecewise polynomial approximation. Trans. Amer. Math. Soc. **355**, 2585–2631 (2003)
22. Oswald, O.: Multilevel Finite Element Approximation. Teubner Skripten zur Numerik, Teubner (1994)
23. Peetre, J.: New thoughts on Besov spaces. Duke Univ. Math. Series, Duke Univ. Durham (1976)
24. Petrushev, P.: Anisotropic spaces and nonlinear n-term spline approximation. Approximation Theory XI: Gatlinburg 2004, 363–394, Mod. Methods Math., Nashboro Press, Brentwood, TN (2005)
25. Stein, E.: Harmonic Analysis: Real-variable methods, Orthogonality and oscillatory integrals. Princeton University Press (1993)
26. Triebel, H.: Theory of Function Spaces. Monographs in Math. vol. 78, Birkhäuser (1983)
27. Xu, J.: Iterative methods by space decomposition and subspace correction, SIAM Review **34**, 581–613 (1992)