
Two-Level-Split Decomposition of Anisotropic Besov Spaces

Wolfgang Dahmen · Shai Dekel · Pencho Petrushev

Received: 4 February 2008 / Accepted: 2 March 2009
© Springer Science+Business Media, LLC 2009

Abstract Anisotropic Besov spaces (B-spaces) are developed based on anisotropic multilevel ellipsoid covers (dilations) of \mathbb{R}^n . This extends earlier results on anisotropic Besov spaces. Furthermore, sequences of anisotropic bases are constructed and utilized for two-level-split decompositions of the B-spaces and nonlinear m -term approximation.

Keywords Anisotropic Besov spaces · Multilevel expansions · Nonlinear approximation

Mathematics Subject Classification (2000) Primary 46E35 · Secondary 41A15 · 41A25 · 41A46

Communicated by Ronald A. DeVore.

This work has been supported in part by NSF Grant DMS-0709046 and by the Leibniz Program of the German Research Foundation.

W. Dahmen

Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, 52056 Aachen, Germany

e-mail: dahmen@igpm.rwth-aachen.de

S. Dekel

27 Hamaskit St., Herzelia, Israel

e-mail: shai.dekel@ge.com

P. Petrushev (✉)

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

e-mail: pencho@math.sc.edu

1 Introduction

Anisotropic phenomena appear in various contexts in mathematical analysis and its applications. The formation of shocks results in jump discontinuities of solutions of hyperbolic conservation laws across lower dimensional manifolds, and sharp edges often separate areas of little detail in digital images, to name just two examples. In the first case such jumps cause a serious obstruction to appropriate regularity theorems, since the available regularity scales are either inherently isotropic or coordinate biased or are subject to an uncontrollable restricted regularity range, as will be pointed out below. In the second case the anisotropy makes it more difficult to come up with efficient encoding strategies that are entropy optimal when using cartoons as model classes. The central objective of this paper is to set up a sufficiently flexible framework for *anisotropic* regularity notions that at least bear a potential for dealing with questions of the above type.

Let us briefly relate now the present approach to earlier considerable efforts in extending the theory of classical Besov and Triebel–Lizorkin spaces to non-isotropic settings. One line of extension is concerned with function and distribution spaces on \mathbb{R}^n furnished with anisotropic dilation structures. Various authors have considered dilations of \mathbb{R}^n induced by a one-parameter family of diagonal matrices of the form $\text{diag}(2^{b_1 t}, \dots, 2^{b_n t})$, $b_j > 0$. Calderón and Torchinsky [4, 5] have developed the so-called parabolic Hardy spaces induced by certain one-parameter dilation groups, which were further extended by Folland and Stein [12]. Bownik [1, 2] and Bownik and Ho [3] have considered dilations on \mathbb{R}^n induced by the powers A^j of an expansive matrix A with eigenvalues λ satisfying $|\lambda| > 1$, and applied them to the development of anisotropic Besov and Triebel–Lizorkin spaces.

Yet another line of generalization of the classical spaces originated from the work of Coifman and Weiss [6, 7], where \mathbb{R}^n was replaced by the more general *spaces of homogeneous type*. Spaces on homogeneous spaces have been further developed by several authors, in particular, more recently by Y. Han and his coauthors, see, e.g., [13].

In this article we introduce and explore a new more general setting for anisotropic Besov spaces (homogeneous and inhomogeneous) on \mathbb{R}^n of any positive smoothness that allows a lot more flexibility in dealing with anisotropic features of functions, e.g., singularities along curves and lower dimensional smooth manifolds. Our approach is based on multilevel discrete (or continuous) anisotropic dilations of \mathbb{R}^n , where only mild conditions are imposed on the dilation matrices which still allow them to rapidly change from point to point and in depth, from level to level. These dilations are much more general than the dilations mentioned above which are shift invariant and where the different levels are obtained from powers of the same matrix or from a one-parameter family of matrices. Instead of “dilations” we shall use the more convenient geometric term “ellipsoid cover of \mathbb{R}^n .”

As will be shown, \mathbb{R}^n equipped with a discrete or continuous ellipsoid cover becomes a homogeneous space, and therefore the general theory of Besov spaces on homogeneous spaces applies. However, this theory only allows one to handle Besov and Triebel–Lizorkin spaces of smoothness α with $|\alpha| < \varepsilon$, where $\varepsilon > 0$ is sufficiently small.

Although very general, our setup allows us to overcome this latter principal limitation. For a given discrete multilevel ellipsoid cover Θ of \mathbb{R}^n (see Sect. 2.1), we define the B-spaces $\dot{B}_{pq}^\alpha(\Theta)$ and $B_{pq}^\alpha(\Theta)$ via *local polynomial approximation* on all ellipsoids from Θ . This approach enables us indeed to develop the anisotropic Besov spaces (B-spaces) \dot{B}_{pq}^α and B_{pq}^α for *arbitrarily large* $\alpha > 0$ which is the principal difference between our theory and the general theory of Besov spaces on homogeneous spaces. Our B-spaces can be viewed as a generalization of the classical Besov spaces as well as the anisotropic Besov spaces studied by Bownik [2] with weight 1 (see Sect. 5.3). For the theory of classical Besov spaces we refer the reader to [16, 18, 19].

The highly anisotropic locally and scalewise varying structure of the dilations considered here prevents us from using Fourier analysis techniques; our approach relies on algebraic polynomials instead. A key ingredient in our theory is a certain *two-level-split decomposition* of B-spaces. For any given anisotropic cover Θ we construct a multilevel sequence of bases which provide a means for decomposing the associated B-spaces. The basis elements are compactly supported smooth functions which are well aligned with the anisotropic space dilations induced by Θ . The two-level-split decomposition of functions originates from [9], where it is applied in a regular setup and used for the purpose of preconditioning systems of equations stemming from mesh-free discretizations of elliptic boundary value problems.

Of course, the usefulness of any regularity notion relies on whether it actually conveys sufficient information about the intrinsic complexity of corresponding functions, e.g., expressed in terms of (quasi-)sparse representations with respect to a suitable system of building blocks. Just as one may expect, we show that the B-spaces $\dot{B}_{\tau\tau}^\alpha(\Theta)$ with $1/\tau = \alpha/n + 1/p$ govern the rates of nonlinear m -term L_p -approximation from the two-level-split bases induced by Θ . This suggests the adaptation of Θ to the approximand. Accordingly, we would like to put forward here a new adaptive strategy for measuring the smoothness of functions, namely, by choosing from the collection of all anisotropic B-spaces, induced by an admissible set of ellipsoid covers of \mathbb{R}^n , one that maximizes the regularity index, viz., the rate of best m -term approximation.

To better illustrate our adaptive approach to smoothness, we discuss two examples that are very simple but nevertheless reveal some perhaps unexpected effects. As for the first example, we show in Sect. 7.1 that for a suitable ellipsoid cover Θ , the B-space smoothness of the characteristic function of the unit ball $B(0, 1) \subset \mathbb{R}^2$ in $\dot{B}_{\tau\tau}^\alpha(\Theta)$ is essentially $4/p$, while in the corresponding (classical isotropic) Besov spaces it is $2/p$. More strikingly, in the adaptive B-space scales $\dot{B}_{\tau\tau}^\alpha(\Theta)$, the smoothness of the characteristic function of any square $Q \subset \mathbb{R}^2$ is arbitrarily high, i.e., can be any $\alpha > 0$, while in the corresponding isotropic Besov spaces it is essentially $2/p$ (see Sect. 7.2). However, it is important to note that the cover Θ needed to describe that level of smoothness depends on α , a fact that has some interesting consequences worth a brief discussion (see Sect. 7.3). As a consequence, the rates of adaptive m -term approximation for these model functions from the two-level-split bases are better than, e.g., the rates of m -term approximation from wavelets.

Let us finally mention that more special anisotropic B-spaces of technically different nature but somewhat similar in spirit have been introduced and used in [10, 14] (see also [8, 15, 17]). These are B-spaces generated by multilevel *nested* triangulations and are defined via local *piecewise polynomial* approximation. They are specif-

ically designed for the purpose of nonlinear m -term spline approximation, and consequently are much more specialized than the B-spaces of this paper. Moreover, it is important to note that insisting on hierarchies of *nested* partitions entails a distinct lack of flexibility compared to the present approach since, for instance, the successors of a triangle with a sharp angle will always exhibit angles at least that sharp. Yet another type of anisotropic Besov spaces has been considered in [11].

The outline of the paper is the following. In Sect. 2 we introduce and give some basic properties of the continuous and discrete multilevel covers of \mathbb{R}^n . In Sect. 3 we construct the bases induced by discrete multilevel covers of \mathbb{R}^n and establish their stability. In Sect. 4 we utilize the bases from Sect. 3 for the construction of two-level-split bases and decompositions. In Sect. 5 we define the homogeneous and inhomogeneous B-spaces induced by multilevel covers of \mathbb{R}^n and establish the equivalence of various B-norms. Section 6 is devoted to nonlinear m -term approximation from two-level-split bases. In Sect. 7 we show that the characteristic functions of the unit ball and any square in \mathbb{R}^n have higher adaptive B-space smoothness than their regular Besov space smoothness. In Appendix we give the proof of the equivalence of several B-space norms.

Throughout the paper we use the following notation: $\|f\|_p := \|f\|_{L_p(\mathbb{R}^n)}$; for a measurable set $E \subset \mathbb{R}^n$, $|E|$ denotes the Lebesgue measure of E , $\mathbb{1}_E$ is the characteristic function of E ; \mathcal{P}_k denotes the set of all algebraic polynomials in n variables of total degree $\leq k - 1$. The Schwartz class of rapidly decaying C^∞ functions on \mathbb{R}^n is denoted by \mathcal{S} and its dual by \mathcal{S}' . Positive constants are denoted by c, c_1, c_*, \dots and they may vary at every occurrence; $A \sim B$ means $c_1 A \leq B \leq c_2 A$ for some unspecified but fixed constants c_1, c_2 .

2 Anisotropic Multilevel Ellipsoid Covers of \mathbb{R}^n

We begin with some preliminaries. The image of the unit ball $B^* := B(0, 1)$ in \mathbb{R}^n via an affine transform will be called an *ellipsoid*. For a given ellipsoid $\theta \subset \mathbb{R}^n$, we let A_θ be an affine transform such that $\theta = A_\theta(B^*)$. Denoting by $v_\theta := A_\theta(0)$ the “center” of θ , we have

$$A_\theta(x) = M_\theta x + v_\theta, \tag{2.1}$$

where M_θ is a nonsingular $n \times n$ matrix.

It will be convenient for us to use the singular value decomposition of matrices, that is, any nonsingular real matrix $M \in \mathbb{R}^{n \times n}$ can be represented in the form $M = UDV$, where the matrices U and V are $n \times n$ orthogonal matrices and D is diagonal and $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. It is easy to see that $\sigma_1^2 \geq \dots \geq \sigma_n^2$ are the eigenvalues of $M^T M$ and $\sigma_1^{-2} \leq \dots \leq \sigma_n^{-2}$ are the eigenvalues of $(M^{-1})^T M^{-1}$. Hence

$$\|M\|_{\ell_2 \rightarrow \ell_2} = \sigma_1 \quad \text{and} \quad \|M^{-1}\|_{\ell_2 \rightarrow \ell_2} = 1/\sigma_n. \tag{2.2}$$

These norms have a clear geometric meaning. Thus if M_θ is as in (2.1), then $\text{diam } \theta = 2\|M_\theta\|_{\ell_2 \rightarrow \ell_2}$.

2.1 Discrete Ellipsoid Covers (Dilations) of \mathbb{R}^n

The notion of a *discrete multilevel ellipsoid cover* of \mathbb{R}^n we introduce and explore here will play a central role in this study.

Definition 2.1 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, a_1, \dots, a_8 , and N_1 are positive constants:

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

$$a_1 2^{-a_0 m} \leq |\theta| \leq a_2 2^{-a_0 m}$$

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

(b) For each $\theta \in \Theta$ let A_θ be an affine transform, associated with θ , 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 $\theta \in \Theta_m$ ($m \in \mathbb{Z}$) and $\theta' \in \Theta_{m+v}$ ($v \geq 0$) with $\theta \cap \theta' \neq \emptyset$, we have

$$a_3 2^{-a_4 v} \leq 1 / \|M_{\theta'}^{-1} M_\theta\|_{\ell_2 \rightarrow \ell_2} \leq \|M_\theta^{-1} M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq a_5 2^{-a_6 v}. \tag{2.3}$$

- (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 θ^\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 $\theta \cap \eta \neq \emptyset$ with $\theta \in \Theta_m$ and $\eta \in \Theta_m \cup \Theta_{m+1}$, then $|\theta \cap \eta| > a_8 |\eta|$.

Remarks 1. Any discrete multilevel ellipsoid cover Θ of \mathbb{R}^n depends on a collection of parameters (which are not necessarily independent) that contain all constants appearing in conditions (a)–(e) of Definition 2.1. We shall denote this set of parameters by $\mathbf{p}(\Theta)$.

2. It is important to note that the set of all 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 the parameters of Θ . If $|\det M| \neq 1$, then only the constants a_1 and a_2 in condition (a) should be changed accordingly.

3. Condition (b) indicates that if $\theta \cap \theta' \neq \emptyset$, then the ellipsoids θ and θ' cannot change uncontrollably in shape and orientation when they are from close levels. More precisely, set $M := M_{\theta'}^{-1} M_\theta$ and let $M = UDV$ be the singular value decomposition of M , where U and V are orthogonal matrices, and $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ is diagonal with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. As in (2.2),

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

Therefore, condition (b) is equivalently expressed as

$$a_3 2^{-a_4 \nu} \leq \sigma_n \leq \dots \leq \sigma_1 \leq a_5 2^{-a_6 \nu}. \tag{2.4}$$

This condition also 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 (2.4).

4. Evidently the sign of ν in condition (b) can be reversed. Namely, condition (b) is equivalent to the following condition:

(b') If $\theta \in \Theta_m$ and $\theta' \in \Theta_{m-\nu}$ ($\nu \geq 0$), with $\theta \cap \theta' \neq \emptyset$, then

$$(1/a_5) 2^{a_6 \nu} \leq 1 / \|M_{\theta'}^{-1} M_\theta\|_{\ell_2 \rightarrow \ell_2} \leq \|M_\theta^{-1} M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq (1/a_3) 2^{a_4 \nu}. \tag{2.5}$$

Therefore, if $M := M_\theta^{-1} M_{\theta'} = UDV$ with $D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ as above, then (2.5) is equivalent to

$$(1/a_5) 2^{a_6 \nu} \leq \sigma_n \leq \dots \leq \sigma_1 \leq (1/a_3) 2^{a_4 \nu}. \tag{2.6}$$

5. We need to interrelate the semi-axes of intersecting ellipsoids from Θ . For a given $\theta \in \Theta$, set

$$\sigma_{\max}(\theta) := \|M_\theta\|_{\ell_2 \rightarrow \ell_2} \quad \text{and} \quad \sigma_{\min}(\theta) := \|M_\theta^{-1}\|_{\ell_2 \rightarrow \ell_2}^{-1}. \tag{2.7}$$

These are the maximum and minimum semi-axes of the ellipsoid θ .

Lemma 2.2 *If $\theta \in \Theta_m$, $\theta' \in \Theta_{m+\nu}$, $\nu \geq 0$, and $\theta \cap \theta' \neq \emptyset$, then*

$$a_3 2^{-a_4 \nu} \sigma_{\max}(\theta) \leq \sigma_{\max}(\theta') \leq a_5 2^{-a_6 \nu} \sigma_{\max}(\theta) \tag{2.8}$$

and

$$a_3 2^{-a_4 \nu} \sigma_{\min}(\theta) \leq \sigma_{\min}(\theta') \leq a_5 2^{-a_6 \nu} \sigma_{\min}(\theta). \tag{2.9}$$

Proof Let us abbreviate $\|\cdot\| := \|\cdot\|_{\ell_2 \rightarrow \ell_2}$. By property (b) of Θ we have

$$\|M_{\theta'}\| \leq \|M_\theta\| \cdot \|M_\theta^{-1} M_{\theta'}\| \leq a_5 2^{-a_6 \nu} \|M_\theta\|,$$

and by (b'),

$$\|M_\theta\| \leq \|M_{\theta'}\| \cdot \|M_{\theta'}^{-1} M_\theta\| \leq (1/a_3) 2^{a_4 \nu} \|M_{\theta'}\|,$$

which yield (2.8). One similarly proves (2.9). □

6. Condition (b) can be replaced by the following slightly stronger but more convenient condition:

(b'') For any $\theta, \theta' \in \Theta_m$ ($m \in \mathbb{Z}$) with $\theta \cap \theta' \neq \emptyset$,

$$a_3 \leq 1 / \|M_{\theta'}^{-1} M_\theta\|_{\ell_2 \rightarrow \ell_2} \leq \|M_\theta^{-1} M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq a_5, \quad a_3 \leq 1 \leq a_5, \tag{2.10}$$

and for any $\theta \in \Theta_m$ and $\theta' \in \Theta_{m+1}$ ($m \in \mathbb{Z}$) with $\theta \cap \theta' \neq \emptyset$,

$$2^{-a_4} \leq 1/\|M_{\theta'}^{-1}M_\theta\|_{\ell_2 \rightarrow \ell_2} \leq \|M_\theta^{-1}M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq 2^{-a_6}. \tag{2.11}$$

It is readily seen that this condition implies condition (b).

7. By property (b') (which is equivalent to (b)) it follows that the ellipsoids of Θ_m expand to \mathbb{R}^n as $m \rightarrow -\infty$; more precisely, if $(\theta_m)_{m \leq 0}$ is a sequence of ellipsoids $\theta_m \in \Theta_m$ which contain a fixed point $x \in \mathbb{R}^n$, then $\bigcup_{m \leq 0} \theta_m = \mathbb{R}^n$. In the other direction, property (b) implies that for any compact $K \subset \mathbb{R}^n$,

$$\max\{\text{diam } \theta : \theta \in \Theta_m, \theta \cap K \neq \emptyset\} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

8. Property (d) indicates that every point $x \in \mathbb{R}^n$ is “well covered” by at least one ellipsoid from every level Θ_m .

9. Condition (e) may seem restrictive, but the next observation shows that this is not the case.

Lemma 2.3 *Suppose Θ is a discrete multilevel ellipsoid cover of \mathbb{R}^n satisfying conditions (a)–(d) above. Then there exists a discrete multilevel ellipsoid cover $\tilde{\Theta}$ of \mathbb{R}^n with properties (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 $(a_7 + 1)/2 \leq r_\theta \leq 1$.*

Proof The idea is that slight dilations of each $\theta \in \Theta$, one at a time, allows one to realize in addition (e). In fact, evidently, there exist constants N_0 and N_2 depending on the parameters of Θ such that each ellipsoid $\theta \in \Theta_m$ ($m \in \mathbb{Z}$) can be intersected by at most N_2 ellipsoids from Θ_{m+1} and by at most N_0 ellipsoids from Θ_{m-1} . Now set $N := 2(N_0 + N_1 + N_2) + 1$, $b := (a_7 + 1)/2$, and $\delta := (1 - b)/N$.

For a fixed $\theta \in \Theta_m$, set $\theta_j := A_\theta(B(0, b + j\delta))$, $j = 0, 1, \dots, N$. So, θ_j is the dilated version of θ by a factor of $b + j(1 - b)/N = (1 - \frac{j}{N})b + \frac{j}{N}$. Now let $S_j := \theta_j \setminus \theta_{j-1}$, $j = 1, 2, \dots, N$, be the shells induced by the ellipsoids θ_j . Also, set $S_0 := \theta_0$ and $S_{N+1} := \mathbb{R}^n \setminus \theta$. We shall call a shell S_j ($1 \leq j \leq N$) a “bad shell” if there is an ellipsoid $\eta \in \Theta_{m-1} \cup \Theta_m \cup \Theta_{m+1}$, $\eta \neq \theta$, such that $\eta \cap S_j \neq \emptyset$ and either $\eta \cap S_{j-1} = \emptyset$ or $\eta \cap S_{j+1} = \emptyset$. It is readily seen that a single ellipsoid η can spoil at most two shells S_j and hence at most $2(N_0 + N_1 + N_2)$ shells can be spoiled by ellipsoids from $\Theta_{m-1} \cup \Theta_m \cup \Theta_{m+1}$. Since there are $N = 2(N_0 + N_1 + N_2) + 1$ shells at least one shell, say, S_{j_0} ($1 \leq j_0 \leq N$), will be “good.” Then it is easy to see that if $\eta \cap \theta_{j_0} \neq \emptyset$ for some $\eta \in \Theta_{m-1} \cup \Theta_m \cup \Theta_{m+1}$, then $|\theta_{j_0} \cap \eta| \geq a_8|\eta|$ for some constant $a_8 > 0$ depending only on the parameters of Θ . Thus the dilated version θ_{j_0} of θ has the desired property.

One completes the proof of this proposition inductively by processing as above all ellipsoids from Θ ordered in a sequence. The rule is that once an ellipsoid from Θ has been processed it will never be touched again. □

Examples A first class of examples is based on the so-called Strong Locally Regular (SLR) multilevel triangulations that have been introduced in [14]. They provide another means for constructing discrete ellipsoid covers of \mathbb{R}^2 . Given an SLR-triangulation \mathcal{T} , one considers for each triangle $\Delta \in \mathcal{T}$ 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.

It will be shown in Lemma 2.9 how to derive discrete multilevel covers with the above properties from so-called continuous or semi-continuous ellipsoid covers that will be introduced next and are easier to construct. Any example of such a cover therefore provides an example for the discrete setting as well.

2.2 Continuous Covers and Quasi-Distance

As mentioned before, the essential developments of this paper are based on the above discrete multilevel covers. Nevertheless, we pause to introduce next two variants, namely continuous and semi-continuous covers, for the following reasons. First, while on the one hand these variants are easier to construct, as reflected by several examples given below, we shall give, on the other hand, a general mechanism that generates from any continuous cover a discrete one satisfying properties (a)–(e). In this sense these relaxed notions simplify the realization of concrete examples of discrete covers. Second, they will later be used for the regularity analysis of the examples in Sect. 7. Moreover, we shall see that (all versions of) covers give rise to *quasi-distances*, which allows us to relate the present results to previously known results concerning Besov spaces on homogeneous spaces.

Definition 2.4 We say that

$$\Theta := \bigcup_{t \in \mathbb{R}} \Theta_t$$

is a *continuous multilevel ellipsoid cover* of \mathbb{R}^n or a *continuous system of dilations* of \mathbb{R}^n if it satisfies the following conditions, where a_0, \dots, a_6 are positive constants:

- (a) For every $v \in \mathbb{R}^n$ and $t \in \mathbb{R}$ there exists an ellipsoid $\theta(v, t) \in \Theta_t$ and an affine transform $A_{v,t}$ such that

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

$\theta(v, t) = A_{v,t}(B^*)$, and $A_{v,t}$ is of the form $A_{v,t}(x) = M_{v,t}x + v$, where $M_{v,t}$ is a nonsingular $n \times n$ matrix.

- (b) For any $v, y \in \mathbb{R}^n, t \in \mathbb{R}$, and $s > 0$, if $\theta(v, t) \cap \theta(y, t + s) \neq \emptyset$, then

$$a_3 2^{-a_4 s} \leq 1 / \|M_{y,t+s}^{-1} M_{v,t}\|_{\ell_2 \rightarrow \ell_2} \leq \|M_{v,t}^{-1} M_{y,t+s}\|_{\ell_2 \rightarrow \ell_2} \leq a_5 2^{-a_6 s}. \quad (2.12)$$

For continuous covers also the scale is allowed to vary continuously. The following variant takes us halfway towards the fully discrete setting introduced in the previous setting.

Definition 2.5 We call

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

a *multilevel semi-continuous ellipsoid cover* of \mathbb{R}^n if the following conditions are obeyed, where a_0, \dots, a_6 are positive constants:

- (a) For every $v \in \mathbb{R}^n$ and $m \in \mathbb{Z}$ there exists an ellipsoid $\theta(v, m) \in \Theta_m$ and an affine transform $A_{v,m}$ such that

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

$\theta(v, m) = A_{v,m}(B^*)$, and $A_{v,m}$ is of the form $A_{v,m}(x) = M_{v,m}x + v$, where $M_{v,m}$ is a nonsingular $n \times n$ matrix.

- (b) For any $v, y \in \mathbb{R}^n, m \in \mathbb{Z}$, and $\nu > 0$, if $\theta(v, m) \cap \theta(y, m + \nu) \neq \emptyset$, then

$$a_3 2^{-a_4 \nu} \leq 1 / \|M_{y,m+\nu}^{-1} M_{v,m}\|_{\ell_2 \rightarrow \ell_2} \leq \|M_{v,m}^{-1} M_{y,m+\nu}\|_{\ell_2 \rightarrow \ell_2} \leq a_5 2^{-a_6 \nu}. \tag{2.13}$$

It is readily seen that any continuous ellipsoid cover Θ of \mathbb{R}^n induces a semi-continuous ellipsoid cover by defining Θ_m as Θ_t with $t = m$.

Examples (i) The regular cover of \mathbb{R}^n consisting of all balls in \mathbb{R}^n is the simplest example of a continuous ellipsoid (ball) cover of \mathbb{R}^n .

- (ii) The one-parameter family of diagonal dilation matrices

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

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

(iii) In [4, 5] Calderón and Torchinsky developed the so-called parabolic Hardy spaces generated by continuous dilation matrices associated with a continuous semigroup $A_t, t > 0, A_{st} := A_t A_s$ of affine transforms on \mathbb{R}^n leaving the origin fixed and satisfying

$$t^\alpha \leq \|A_t\|_{\ell_2 \rightarrow \ell_2} \leq t^\beta, \quad t \geq 1,$$

where $1 \leq \alpha \leq \beta < \infty$. One can easily see that any such semigroup of matrices gives rise to a continuous ellipsoid cover (dilation) of \mathbb{R}^n .

(iv) Consider an arbitrary $n \times n$ real matrix A 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) in the sense of Definition 2.5. These particular kinds of dilations are used in [1–3] for the development of anisotropic Hardy, Besov, and Triebel–Lizorkin spaces.

(v) In Sect. 7, we construct two nontrivial continuous covers of \mathbb{R}^2 . These are examples of anisotropic ellipsoid covers where the ellipsoids change rapidly from point to point and in depth.

In the following we first show that a quasi-distance is naturally associated with any continuous or semi-continuous or discrete ellipsoid covers of \mathbb{R}^n . Second, we prove that \mathbb{R}^n equipped with a distance of this sort and the Lebesgue measure becomes a space of homogeneous type. Thus the anisotropic Besov spaces we develop in this paper are in particular spaces on homogeneous type spaces.

We begin by recalling the definition of a quasi-distance on a set $X \neq \emptyset$: A map $\rho : X \times X \rightarrow [0, \infty)$ is called a *quasi-distance* on X if for any $x, y, z \in X$:

- (i) $\rho(x, y) = 0 \iff x = y,$
- (ii) $\rho(y, x) = \rho(x, y),$
- (iii) $\rho(x, z) \leq K(\rho(x, y) + \rho(y, z))$ with $K \geq 1.$

Definition 2.6 Assuming that Θ is a continuous or semi-continuous or discrete ellipsoid cover of $\mathbb{R}^n,$ we define

$$\rho(x, y) := \min\{|\theta| : \theta \in \Theta \text{ and } x, y \in \theta\}, \quad x, y \in \mathbb{R}^n. \tag{2.14}$$

Note that in order to cover the standard Euclidean distance we should take the n th root of $\rho(x, y)$ defined above. But since this would only change the constant K in (iii) while adding to technicalities, we prefer to work with the above version.

Proposition 2.7 For any continuous, semi-continuous or discrete cover Θ of \mathbb{R}^n the map $\rho : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, \infty)$ defined above is a quasi-distance on $\mathbb{R}^n.$

We shall prove this proposition only in the case when Θ is a discrete cover of $\mathbb{R}^n.$ The proof when Θ is a continuous or semi-continuous cover is similar or even easier. The claim will follow from the following lemma:

Lemma 2.8 There is an integer $s_* \geq 1$ depending only on $\mathbf{p}(\Theta)$ such that for any ellipsoids $\theta \in \Theta_m$ ($m \in \mathbb{Z}$) and $\theta' \in \Theta_{m+v}$ ($v \geq 0$) with $\theta \cap \theta' \neq \emptyset$ and any $j \geq s_*$ there exists an ellipsoid $\eta_j \in \Theta_{m-j}$ such that $\theta \cup \theta' \subset \eta_j.$

Proof Let $\omega' := A_\theta^{-1}(\theta'),$ and recall that $A_\theta^{-1}(\theta) = B^*.$ By property (b) of Θ it follows that

$$\text{diam } \omega' = 2\|M_\theta^{-1}M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq 2a_5 2^{-a_0 v} \leq 2a_5,$$

and hence

$$A_\theta^{-1}(\theta \cup \theta') = B^* \cup \omega' \subset B(0, 1 + 2a_5). \tag{2.15}$$

By property (d) of Θ for any $j \geq 1$ there exists $\theta_j \in \Theta_{m-j}$ such that $v_\theta \in \theta_j^\diamond,$ where $\theta_j^\diamond := A_{\theta_j}(B(0, a_7))$ is the dilated θ_j by a factor of $a_7 < 1.$ Set $\omega_j := A_\theta^{-1}(\theta_j),$ and let ω_j^\diamond be the dilated ω_j by a factor of $a_7.$ Also, denote by $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ the semi-axes of the ellipsoid $\omega_j.$ By property (b') (see (2.6)), which is equivalent to property (b), it follows that

$$\sigma_n \geq (1/a_5)2^{a_6 j}. \tag{2.16}$$

On the other hand, by a simple geometric property of ellipsoids,

$$\text{dist}(\omega_j^\diamond, \partial\omega_j) = (1 - a_7)\sigma_n,$$

where $\text{dist}(E_1, E_2)$ denotes the Euclidean distance between the sets $E_1, E_2 \subset \mathbb{R}^n.$ From this and (2.16) it follows that

$$\text{dist}(\omega_j^\diamond, \partial\omega_j) \geq \frac{1 - a_7}{a_5} 2^{a_6 j}. \tag{2.17}$$

Now choose $s_* \geq 1$ so that

$$\frac{1 - a_7}{a_5} 2^{a_6 s_*} \geq 1 + 2a_5. \tag{2.18}$$

Let $j \geq s_*$. Since $v_\theta \in \theta_j^\diamond$, then $0 \in \omega_j^\diamond$, and using (2.15), (2.17) and (2.18) we infer $A_\theta^{-1}(\theta \cup \theta') \subset \omega_j$, which implies $\theta \cup \theta' \subset A_\theta(\omega_j) =: \eta_j$. This completes the proof. \square

Proof of Proposition 2.7 Remark 7 gives property (i). Property (ii) is obvious by the definition of $\rho(\cdot, \cdot)$. To show (iii), let $x, y, z \in \mathbb{R}^n$, and assume that $\rho(x, z) = |\theta|$, $x, z \in \theta$, and $\rho(z, y) = |\theta'|$, $z, y \in \theta'$, where $\theta \in \Theta_m$ and $\theta' \in \Theta_{m+\nu}$. Without loss of generality we may assume that $\nu \geq 0$. We now apply Lemma 2.8 to conclude that there exists an ellipsoid $\eta \in \Theta_{m-s_*}$ such that $\theta \cup \theta' \subset \eta$, and hence

$$\rho(x, y) \leq |\eta| \leq a_2 2^{-a_0(m-s_*)} \leq (a_2/a_1) 2^{a_0 s_*} (|\theta| + |\theta'|) \leq K(\rho(x, z) + \rho(z, y)). \square$$

The following observation justifies exploiting the greater comfort of continuous or semi-continuous covers:

Lemma 2.9 *For every continuous or semi-continuous ellipsoid cover Θ of \mathbb{R}^n there is a discrete ellipsoid cover $\widehat{\Theta}$ such that each ellipsoid $\theta \in \widehat{\Theta}$ is obtained from an ellipsoid in Θ by dilation by a factor r_θ obeying $(a_7 + 1)/2 \leq r_\theta \leq 1$.*

Furthermore, the quasi-distances $\rho(\cdot, \cdot)$ and $\widehat{\rho}(\cdot, \cdot)$ induced by Θ and $\widehat{\Theta}$ are equivalent, i.e., $\rho(x, y) \sim \widehat{\rho}(x, y)$ for $x, y \in \mathbb{R}^n$.

Proof We may assume that Θ is a semi-continuous ellipsoid cover of \mathbb{R}^n , for otherwise we construct one from the given continuous cover.

We first construct for every $m \in \mathbb{Z}$ a countable set $\widehat{\Theta}_m \subset \Theta_m$ satisfying conditions (c)–(d) of Definition 2.1. This can be done, e.g., in two steps as follows: We first choose countably many ellipsoids from Θ_m so that condition (d) is fulfilled, and then inductively remove from this collection one-by-one all ellipsoids which do not destroy condition (d). After that condition (c) will be automatically fulfilled with some constant N_1 because of condition (b) on Θ . Conditions (a)–(b) on $\widehat{\Theta}_m$ will be inherited from Θ_m .

Secondly, Lemma 2.3 enables us to correct $\{\widehat{\Theta}_m\}$ so that condition (e) is obeyed as well.

The equivalence of $\rho(\cdot, \cdot)$ and $\widehat{\rho}(\cdot, \cdot)$ is an easy consequence of the above construction. \square

Proposition 2.10 *Let Θ be a continuous or semi-continuous or discrete ellipsoid cover of \mathbb{R}^n , and set $B_\rho(x, r) := \{y \in \mathbb{R}^n : \rho(x, y) < r\}$. Then*

$$|B_\rho(x, r)| \sim r, \quad x \in \mathbb{R}^n, r > 0, \tag{2.19}$$

where the constants of equivalence depend only on $\mathbf{p}(\Theta)$. Therefore, the Lebesgue measure on \mathbb{R}^n is a doubling measure with respect to the quasi-distance $\rho(\cdot, \cdot)$, i.e.,

$$|B_\rho(x, 2r)| \leq c |B_\rho(x, r)|, \quad x \in \mathbb{R}^n, r > 0. \tag{2.20}$$

In other words, \mathbb{R}^n equipped with the distance $\rho(\cdot, \cdot)$ defined by Definition 2.14 and the Lebesgue measure is a space of homogeneous type.

Proof We only consider the case when $\rho(\cdot, \cdot)$ is generated by a discrete ellipsoid cover of \mathbb{R}^n . We first prove that $|B_\rho(x, r)| \geq cr$. Choose m so that $a_2 2^{-a_0 m} < r \leq a_2 2^{-a_0(m-1)}$. There exists $\theta \in \Theta_m$ such that $x \in \theta$, and hence by property (a) of Θ ,

$$a_1 2^{-a_0 m} \leq |\theta| \leq a_2 2^{-a_0 m} < r.$$

From this and the definition of $\rho(\cdot, \cdot)$ it follows that $\theta \subset B_\rho(x, r)$, and hence

$$|B_\rho(x, r)| \geq |\theta| \geq a_1 2^{-a_0 m} \geq (a_1/a_2) 2^{-a_0} r = cr.$$

In the other direction, note first that

$$B_\rho(x, r) = \bigcup_{\theta \in \Theta: x \in \theta, |\theta| < r} \theta.$$

Suppose $\theta \in \Theta_m$ is of minimum level such that $x \in \theta$ and $|\theta| < r$. Let $\theta' \in \Theta$ be any ellipsoid such that $x \in \theta'$ and $|\theta'| < r$. Then $\theta' \in \Theta_{m+l}$ for some $l \geq 0$. Setting $\omega' := A_\theta^{-1}(\theta')$, one has, by property (b) of Θ ,

$$\text{diam } \omega' = 2 \|M_\theta^{-1} M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq 2a_5.$$

Therefore, $A_\theta^{-1}(B_\rho(x, r)) \subset B(0, 2a_5)$, and hence $|A_\theta^{-1}(B_\rho(x, r))| \leq c$, $c = c(n, a_5)$. This leads to

$$|B_\rho(x, r)| \leq c |\det M_\theta| = c |\theta| < cr. \quad \square$$

3 Construction of a Multilevel System of Bases

Given a discrete ellipsoid cover Θ , we shall 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 . Second, we shall utilize the bases $\{\Phi_m\}$ for the introduction of maps which locally preserve polynomials.

3.1 Coloring the Ellipsoids in Θ

We begin by splitting the ellipsoid cover Θ into no more than $2N_1$ disjoint subsets (colors) $\{\Theta^\nu\}_{\nu=1}^{2N_1}$ 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. Indeed, using property (c) of Θ it is easy to color any level Θ_m by using no more than N_1 colors. So, we use at most N_1 colors to color the ellipsoids in $\{\Theta_{2j}\}_{j \in \mathbb{Z}}$ and, further, at most N_1 colors to color the ellipsoids in $\{\Theta_{2j+1}\}_{j \in \mathbb{Z}}$.

Thus, we may assume that we have the following disjoint splitting:

$$\Theta = \bigcup_{\nu=1}^{2N_1} \Theta^\nu \quad \text{and} \quad \Theta_{2j} = \bigcup_{\nu=1}^{N_1} \Theta_{2j}^\nu, \quad \Theta_{2j+1} = \bigcup_{\nu=N_1+1}^{2N_1} \Theta_{2j+1}^\nu, \quad j \in \mathbb{Z}, \quad (3.1)$$

where if $\theta' \in \Theta_{m_1}^{\nu_1}$, $\theta'' \in \Theta_{m_2}^{\nu_2}$ with $|m_1 - m_2| \leq 1$, and $\theta' \cap \theta'' \neq \emptyset$, then $\nu_1 \neq \nu_2$.

3.2 Definition of Single-Level Bases

We first introduce $2N_1$ smooth piecewise polynomial bumps associated with the colors from above. For fixed positive integers L and k ($L \geq k$) we define

$$\phi_\nu(x) := (1 - |x|^2)_+^{L+\nu k}, \quad \nu = 1, 2, \dots, 2N_1, \quad x_+ := \max\{x, 0\}. \tag{3.2}$$

Notice that $\phi_\nu \in C^{L+\nu k-1} \subset C^L$.

Remark 3.1 The bumps ϕ_ν can be modified to be C^∞ functions. To this end, let $h \in C^\infty(\mathbb{R}^n)$ be such that $\text{supp } h = \overline{B(0, 1)}$, $h \geq 0$, and $\int_{\mathbb{R}^n} h = 1$. Denote $h_\delta(x) := \delta^{-n} h(\delta^{-1}x)$. Then for $0 < \delta < 1$ the bumps $\phi_\nu^* := \phi_\nu * h_\delta$ apparently have the properties: $\phi_\nu^* \in C^\infty$, ϕ_ν^* is a polynomial of degree exactly $2(L + \nu k)$ on $B(0, 1 - \delta)$, and $\text{supp } \phi_\nu^* = \overline{B(0, 1 + \delta)}$. Now the bumps $\{\phi_\nu^*\}$, dilated by a factor of $1 + \delta$ with δ sufficiently small (depending on the parameters of Θ), can be successfully used in place of $\{\phi_\nu\}$.

For any $\theta \in \Theta$ we let A_θ denote the affine transform from Definition 2.1 such that $A_\theta(B^*) = \theta$ (recall $B^* := B(0, 1)$) and set

$$\phi_\theta := \phi_\nu \circ A_\theta^{-1} \quad \text{if } \theta \in \Theta^\nu, \quad 1 \leq \nu \leq 2N_1. \tag{3.3}$$

In a standard way we introduce m th level partitions of unity by defining for any $\theta \in \Theta_m$

$$\varphi_\theta := \frac{\phi_\theta}{\sum_{\theta' \in \Theta_m} \phi_{\theta'}}. \tag{3.4}$$

By the properties of the ellipsoids from Θ it follows that

$$\sum_{\theta \in \Theta_m} \varphi_\theta(x) = 1 \quad \text{and} \quad 0 < c_1 \leq \sum_{\theta \in \Theta_m} \phi_\theta(x) \leq c_2 \quad \text{for } x \in \mathbb{R}^n. \tag{3.5}$$

Also, for any multi-index $\beta = (\beta_1, \dots, \beta_n)$ with $|\beta| = \beta_1 + \dots + \beta_n \leq L$, and $\theta, \theta' \in \Theta_m$ such that $\theta \cap \theta' \neq \emptyset$,

$$\|\partial^\beta(\varphi_\theta \circ A_{\theta'})\|_\infty \leq c, \tag{3.6}$$

where c depends only on L and $\mathbf{p}(\Theta)$.

Let

$$\{P_\beta : |\beta| \leq k - 1\}, \quad \text{where } \deg P_\beta = |\beta|, \tag{3.7}$$

be an orthonormal basis in $L_2(B^*)$ for the space \mathcal{P}_k of all polynomials in n variables of total degree $k - 1$. Since $\|P_\beta\|_{L_2(B^*)} = 1$, then $\|P_\beta\|_{L_\infty(B^*)} \sim 1$, and hence

$$\|P_\beta \phi_\nu\|_{L_2(B^*)} \sim \|P_\beta \phi_\nu\|_{L_\infty(B^*)} \sim 1. \tag{3.8}$$

For any $\theta \in \Theta$ and $|\beta| < k$ we define

$$P_{\theta, \beta} := |\theta|^{-1/2} P_\beta \circ A_\theta^{-1} \tag{3.9}$$

and

$$g_{\theta,\beta} := \varphi_\theta P_{\theta,\beta}. \tag{3.10}$$

Notice that $\|g_{\theta,\beta}\|_2 \sim 1$, and in general $\|g_{\theta,\beta}\|_p \sim |\theta|^{1/p-1/2}$ ($0 < p \leq \infty$); also $g_{\theta,\beta} \in C^L$. Now we define the m th level basis Φ_m by

$$\Phi_m := \{g_{\theta,\beta} : \theta \in \Theta_m, |\beta| \leq k - 1\} \tag{3.11}$$

and set

$$S_m := \text{span}(\Phi_m), \tag{3.12}$$

i.e., S_m is the set of all functions f on \mathbb{R}^n of the form

$$f(x) = \sum_{\theta \in \Theta_m, |\beta| < k} c_{\theta,\beta} g_{\theta,\beta}(x), \quad x \in \mathbb{R}^n, \tag{3.13}$$

where $\{c_{\theta,\beta}\}$ is an arbitrary collection of real numbers. Notice that the sum in (3.13) is finite for any $x \in \mathbb{R}^n$ and hence well defined.

Remark 3.2 (a) By the definition of $g_{\theta,\beta}$ it readily follows that $\mathcal{P}_k \subset S_m$.

(b) Φ_m is linearly independent, i.e., if $\sum_{\theta \in \Theta_m, |\beta| < k} c_{\theta,\beta} g_{\theta,\beta} = 0$ a.e., then $c_{\theta,\beta} = 0 \forall \theta \in \Theta_m, \forall |\beta| < k$. More importantly, Φ_m is in a sense locally linearly independent and L_p stable, which will be established in the next theorem.

Theorem 3.3 *Any function $f \in S_m$ has a unique representation*

$$f(x) = \sum_{\theta \in \Theta_m, |\beta| < k} \langle f, \tilde{g}_{\beta,\theta} \rangle g_{\theta,\beta}(x), \tag{3.14}$$

where for every $x \in \mathbb{R}^n$ the sum is finite and the functions $\tilde{g}_{\beta,\theta}$ have the following properties: For every $\theta \in \Theta_m$ there exists an ellipsoid $B_\theta := A_\theta(B_\theta^*) \subset \theta$ for some ball $B_\theta^* \subset B^*$ with $|B_\theta| \sim |\theta|$ such that for $0 < p \leq \infty$,

$$\langle g_{\theta',\beta'}, \tilde{g}_{\theta,\beta} \rangle = \delta_{\theta,\theta'} \delta_{\beta,\beta'}, \quad \text{supp } \tilde{g}_{\theta,\beta} \subset \overline{B_\theta}, \quad \|\tilde{g}_{\theta,\beta}\|_p \sim |\theta|^{1/p-1/2}. \tag{3.15}$$

Moreover, if $f \in S_m \cap L_p$, $0 < p \leq \infty$, and $f = \sum_{\theta \in \Theta_m, |\beta| < k} c_{\theta,\beta} g_{\theta,\beta}$, then

$$\|f\|_p \sim \left(\sum_{\theta \in \Theta_m, |\beta| < k} \|c_{\theta,\beta} g_{\theta,\beta}\|_p^p \right)^{1/p}, \tag{3.16}$$

with the obvious modification when $p = \infty$. Here all constants of equivalence depend only on $\mathbf{p}(\Theta)$, L , p , and k .

Proof We first construct the balls $B_\theta^* \subset B^*$. Fix $\theta \in \Theta_m$ ($m \in \mathbb{Z}$) and let \mathcal{X}_θ be the set of all $\theta' \in \Theta_m$ such that $\theta' \cap \theta \neq \emptyset$. Define

$$\mathcal{X}_\theta^* := \{A_\theta^{-1}(\theta') : \theta' \in \mathcal{X}_\theta\}.$$

By the properties of Θ , for any $\eta \in \mathcal{X}_\theta^*$ we have $|\eta \cap B^*| \geq a_8 |B^*| = c > 0$, and there exist balls B_1, B_2 such that $B_1 \subset \eta \subset B_2$ and $|B_1| \sim |B_2| \sim 1$. From this and because $\#\mathcal{X}_\theta^* \leq N_1$ it follows that there exists a ball $B_\theta^* \subset B^*$ such that $|B_\theta^*| \sim 1$, and for each $\eta \in \mathcal{X}_\theta^*$ either $B_\theta^* \subset \eta$ or $B_\theta^* \cap \eta = \emptyset$. Define $B_\theta := A_\theta(B_\theta^*)$.

Denote by \mathcal{Y}_θ the set of all $\theta' \in \mathcal{X}_\theta$ such that $B_\theta \subset \theta'$, and set

$$\mathcal{F}_\theta := \{g_{\theta',\beta}^\diamond := g_{\theta',\beta} \mathbb{1}_{B_\theta} : \theta' \in \mathcal{Y}_\theta, |\beta| < k\}.$$

It is an important observation that the set of functions \mathcal{F}_θ is linearly independent. Indeed, since all ellipsoids $\theta' \in \mathcal{Y}_\theta$ contain B_θ , every two of them have distinct colors. Thus, if $\theta' \in \mathcal{Y}_\theta$ and $\theta' \in \Theta_m^{v_j}$ for some $1 \leq v_j \leq 2N_1$, then $\phi_{\theta'} P_{\theta',\beta}$ is a polynomial of degree exactly $L + v_j k + |\beta'|$ on B_θ and $L + v_j k \leq L + v_j k + |\beta'| < L + v_j(k + 1)$. Consequently, the functions $\{\phi_{\theta'} P_{\theta',\beta} \mathbb{1}_{B_\theta} : \theta' \in \mathcal{Y}_\theta, |\beta| < k\}$ are linearly independent on B_θ and hence \mathcal{F}_θ is linearly independent.

Define $g_{\theta',\beta}^* := |\theta|^{1/2} g_{\theta',\beta}^\diamond \circ A_\theta$. Notice that $\text{supp } g_{\theta',\beta}^* = \overline{B_\theta^*}$ and $\|g_{\theta',\beta}^*\|_2 \sim \|g_{\theta',\beta}^\diamond\|_\infty \sim 1$. Let

$$\mathcal{F}_\theta^* := \{g_{\theta',\beta}^* : \theta' \in \mathcal{Y}_\theta, |\beta| < k\} \quad \text{and} \quad \Lambda_\theta := \{\lambda := (\theta', \beta) : \theta' \in \mathcal{Y}_\theta, |\beta| < k\}.$$

As \mathcal{F}_θ is linearly independent, \mathcal{F}_θ^* is linearly independent as well. Consequently, the Gram matrix

$$G_\theta := \left((g_{\theta',\beta'}^*, g_{\theta'',\beta''}^*) \right)_{(\theta',\beta'),(\theta'',\beta'') \in \Lambda_\theta}$$

is nonsingular, and hence its inverse

$$G_\theta^{-1} =: (R_{(\theta',\beta'),(\theta'',\beta'')})_{(\theta',\beta'),(\theta'',\beta'') \in \Lambda_\theta}$$

exists.

We next show that the functions

$$\tilde{g}_{\theta,\beta} := \sum_{(\theta',\beta') \in \Lambda_\theta} R_{(\theta,\beta),(\theta',\beta')} g_{\theta',\beta'}^\diamond \tag{3.17}$$

form a dual system to Φ_m . Indeed, if $\theta' \in \Theta_m$ and $\theta' \notin \mathcal{Y}_\theta$, then $\theta' \cap \theta = \emptyset$, and hence $\langle g_{\theta',\beta'}, \tilde{g}_{\theta,\beta} \rangle = 0$. If $\theta' \in \mathcal{Y}_\theta$ and $|\beta'| < k$, then

$$\begin{aligned} \langle g_{\theta',\beta'}, \tilde{g}_{\theta,\beta} \rangle &= |\theta| \langle g_{\theta',\beta'} \circ A_\theta, \tilde{g}_{\theta,\beta} \circ A_\theta \rangle = \sum_{(\theta'',\beta'') \in \Lambda_\theta} R_{(\theta,\beta),(\theta'',\beta'')} \langle g_{\theta',\beta'}^*, g_{\theta'',\beta''}^* \rangle \\ &= (G_\theta^{-1} G_\theta)_{(\theta,\beta),(\theta',\beta')} = \delta_{(\theta,\beta),(\theta',\beta')} \end{aligned}$$

as claimed.

Our next and most important step is to show that

$$|R_{(\theta',\beta'),(\theta'',\beta'')}| \leq c \quad \forall (\theta', \beta'), (\theta'', \beta'') \in \Lambda_\theta, \tag{3.18}$$

where $c > 0$ depends only on $\mathbf{p}(\Theta)$, L , and k . We shall use a compactness argument.

It is readily seen that the set \mathcal{F}_θ^* is a particular case of the general case when we have a collection of linearly independent functions

$$\mathcal{F} = \left\{ f_{j,\beta} := \frac{(\phi_{v_j} P_\beta) \circ L_j}{\sum_{j=1}^J \phi_{v_j} \circ L_j} \cdot \mathbb{1}_{B_0} : j = 1, 2, \dots, J, |\beta| < k \right\},$$

where $B_0 \subset B^*$ is a ball with $|B_0| \geq c_1 > 0$, the indices

$$1 \leq v_1 < v_2 < \dots < v_J \leq N_1$$

are fixed, ϕ_v are from (3.2), and P_β are as in (3.7) with the normalization from (3.8), i.e., $\|P_\beta\|_2 = 1$, which implies $\|\phi_v P_\beta\|_\infty \sim 1$. We also assume that L_j , $j = 1, 2, \dots, J$, are affine transforms of the form $L_j(x) = M_j x + v_j$ satisfying the following conditions:

(i) $M_j = U_j D_j V_j$, where U_j and V_j are orthogonal $n \times n$ matrices,

$$D_j = \text{diag}(\sigma_1^j, \sigma_2^j, \dots, \sigma_n^j)$$

$$\text{with } 0 < c_2 \leq \min_\ell \sigma_\ell^j \leq \max_\ell \sigma_\ell^j \leq c_3, \text{ and } |v_j| \leq c_4;$$

- (ii) $0 < c_5 \leq \sum_{j=1}^J (\phi_{v_j} \circ L_j)(x) \leq c_6$ for $x \in B^*$;
- (iii) $L_j(B_0) \subset B^*$.

Let $\Lambda := \{\lambda := (j, \beta) : j = 1, 2, \dots, J, |\beta| < k\}$. Since F is linearly independent, the Gram matrix $G := (\langle f_\lambda, f_{\lambda'} \rangle)_{\lambda, \lambda' \in \Lambda}$ is nonsingular, and hence $G^{-1} =: (R_{\lambda, \lambda'})_{\lambda, \lambda' \in \Lambda}$ exists.

Each of the affine transforms L_j depends on parameters from a subset, say, K , of the set $\mathbb{R}^{n \times n} \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \times \mathbb{R}^n$. The set of all orthogonal $n \times n$ matrices is a compact subset of $\mathbb{R}^{n \times n}$. Hence the parameters of all affine transforms L_j satisfying condition (i) belong to a compact subset, say, K_1 of $\mathbb{R}^{n \times n} \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \times \mathbb{R}^n$. On the other hand, condition (iii) on L_j can be expressed in the form

$$\max_{|x-x_0| \leq a} |M_j x + v_j| \leq 1,$$

where x_0 and a ($a \sim 1$) are the center and radius of B_0 . Therefore, conditions (ii) and (iii) define K as a closed subset of the compact K_1 , and hence K is compact.

The entries of G and $\det G$ apparently depend continuously on the parameters of the affine transforms L_j , $j = 1, 2, \dots, J$, and since K is compact,

$$|\langle f_\lambda, f_{\lambda'} \rangle| \leq c_7 \quad \forall \lambda, \lambda' \in \Lambda \quad \text{and} \quad 0 < c_8 \leq \det G \leq c_9.$$

From this it follows that

$$|R_{\lambda, \lambda'}| \leq c_{10} \quad \forall \lambda, \lambda' \in \Lambda, \tag{3.19}$$

where c_{10} as well as c_7 , c_8 , and c_9 depends only on c_1, \dots, c_6 , $\mathbf{p}(\Theta)$, L , and k . Finally, using that there are only finitely many possibilities for the indices $1 \leq v_1 <$

$v_2 < \dots < v_J \leq N_1$ we conclude that the constant c_{10} in estimate (3.19) can be selected independently of these indices.

Applying the above claim to the specific case at hand it follows that estimate (3.18) holds. Then (3.15) follows by (3.17) and (3.18).

The stability properties (3.16) follow by a standard argument from properties (3.15) of the dual. □

Let us now introduce the more compact notation

$$\Lambda_m := \{ \lambda := (\theta, \beta) : \theta \in \Theta_m, |\beta| < k \}, \tag{3.20}$$

and if $\lambda := (\theta, \beta)$ we shall denote by θ_λ and β_λ the components of λ .

We denote by $\tilde{\Phi}_m := \{ \tilde{g}_\lambda : \lambda \in \Lambda_m \}$ the dual system introduced in Theorem 3.3, i.e., $\langle g_{\lambda'}, \tilde{g}_\lambda \rangle = \delta_{\lambda, \lambda'}$ for $\lambda, \lambda' \in \Lambda_m$.

3.3 Local Projectors Onto Polynomials

The anisotropic regularity notions we are aiming at will rely on appropriate multilevel decompositions of functions based on a given ellipsoid cover. These decompositions, in turn, will be obtained with the aid of appropriate operators which map L_p^{loc} into S_m and locally preserve \mathcal{P}_k and hence provide good local approximation.

Recall first the definition of local and global moduli of smoothness that are standard means to describe the quality of approximation. The usual forward differences of f in direction $h \in \mathbb{R}^n$ on a set $E \subset \mathbb{R}^n$ are defined by $\Delta_h f(x) := \Delta_h^1 f(x) := f(x+h) - f(x)$ if the line segment $[x, x+h]$ is contained in E and by $\Delta_h f(x) := 0$ otherwise. Inductively, we set

$$\Delta_h^k f(x) := \Delta_h(\Delta_h^{k-1} f(x)) \quad \text{if } [x, x+kh] \subset E \quad \text{and} \quad \Delta_h^k f(x) := 0 \text{ otherwise.}$$

Then the k th L_p -moduli of smoothness on θ and \mathbb{R}^n are defined by

$$\omega_k(f, \theta)_p := \sup_{h \in \mathbb{R}^n} \|\Delta_h^k f\|_{L_p(\theta)} \quad \text{and} \quad \omega_k(f, t)_p := \sup_{|h| \leq t} \|\Delta_h^k f\|_p, \quad t > 0. \tag{3.21}$$

The form of the operators will differ somewhat for $p \geq 1$ and $p < 1$.

(a) *The case $1 \leq p \leq \infty$.* There are actually a number of ways to construct suitable operators. While the basic principles are known, some of the estimates to be used are less familiar. Since these notions are essential for what follows, we sketch the arguments for the reader's convenience. A first natural idea is to use the bases $\{\Phi_m\}$ and their duals $\{\tilde{\Phi}_m\}$ to introduce projectors mapping L_p^{loc} onto the spaces S_m by

$$Q_m f := \sum_{\lambda \in \Lambda_m} \langle f, \tilde{g}_\lambda \rangle g_\lambda, \tag{3.22}$$

with \tilde{g}_λ for $\lambda = (\theta, \beta)$ from (3.15).

Alternatively, simpler local projectors onto polynomials are obtained as follows: As in (3.7), suppose $\{P_\beta : |\beta| \leq k-1\}$ is an arbitrary (but fixed) orthonormal basis for \mathcal{P}_k in $L_2(B^*)$. Then for any $\theta \in \Theta$,

$$\{P_{\theta, \beta} := |\theta|^{-1/2} P_\beta \circ A_\theta^{-1} : |\beta| < k\} \tag{3.23}$$

is an orthonormal basis for \mathcal{P}_k in $L_2(\theta)$. Using again our compact notation from (3.10), we can write $g_{\theta,\beta} := \varphi_\theta P_{\theta,\beta}$, $P_\lambda := P_{\theta,\beta}$, and $g_\lambda := g_{\theta,\beta}$, $\lambda := (\theta, \beta)$. We define

$$T_m f := \sum_{\theta \in \Theta_m} \sum_{|\beta| \leq k-1} \langle f, P_{\theta,\beta} \mathbb{1}_{\theta_\lambda} \rangle g_{\theta,\beta} = \sum_{\lambda \in \Lambda_m} \langle f, P_\lambda \mathbb{1}_{\theta_\lambda} \rangle g_\lambda, \tag{3.24}$$

where Λ_m is given by (3.20).

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

$$\theta^* := \bigcup \{ \theta' \in \Theta_m : \theta \cap \theta' \neq \emptyset \} \quad \text{for } \theta \in \Theta_m, \tag{3.25}$$

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

(b) *The case $0 < p < 1$.* Apparently the above operators are no longer usable when working in L_p with $p < 1$. Hence we need to modify them. In fact, the following construction would cover the full range $0 < p \leq \infty$. For $0 < p \leq \infty$ and a given ellipsoid $\theta \in \Theta$, we let $T_{\theta,p} : L_p(\theta) \rightarrow \mathcal{P}_k|_\theta$ be a projector such that

$$\|f - T_{\theta,p} f\|_{L_p(\theta)} \leq c \omega_k(f, \theta)_p, \quad f \in L_p(\theta), \tag{3.26}$$

where $c > 0$ depends only on $\mathbf{p}(\Theta)$, L , k , and p . Note that $T_{\theta,p} f$ can simply be defined as the best (or near best) approximation to f from \mathcal{P}_k in $L_p(\theta)$. Then (3.26) is a consequence of Whitney’s theorem. Also, $T_{\theta,p}$ can be realized as a linear projector onto $\mathcal{P}_k|_\theta$ if $p \geq 1$ by using, say, the Averaged Taylor polynomials, see e.g. [10]. Of course, $T_{\theta,p}$ will be a nonlinear operator if $p < 1$.

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

$$T_{m,p} f := \sum_{\theta \in \Theta_m} \varphi_\theta(x) T_{\theta,p} f. \tag{3.27}$$

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

$$T_{m,p} f := \sum_{\theta \in \Theta_m} \sum_{|\beta| < k} b_{\theta,\beta}(f) g_{\theta,\beta} = \sum_{\lambda \in \Lambda_m} b_\lambda(f) g_\lambda, \tag{3.28}$$

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

In summary, any $\hat{T}_m \in \{Q_m, T_m, T_{m,p}\}$ defined by (3.22), (3.24), or (3.28), has the representation

$$\hat{T}_m f = \sum_{\lambda \in \Lambda_m} b_\lambda(f) g_\lambda, \quad \text{where } b_\lambda(f) = \begin{cases} \langle f, \tilde{g}_\lambda \rangle & \text{if } \hat{T}_m = Q_m, \\ \langle f, P_\lambda \mathbb{1}_{\theta_\lambda} \rangle & \text{if } \hat{T}_m = T_m, \\ \langle T_{m,p} f, \tilde{g}_\lambda \rangle & \text{if } \hat{T}_m = T_{m,p}. \end{cases} \tag{3.29}$$

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

Lemma 3.4 Assume that $0 < p \leq \infty$, and let \hat{T}_m be the operator Q_m from (3.22) or T_m from (3.24) or $T_{m,p}$ from (3.27) if $1 \leq p \leq \infty$, and $\hat{T}_m := T_{m,p}$ if $0 < p < 1$. Then for $f \in L_p^{\text{loc}}$ and $\theta \in \Theta_m$ ($m \in \mathbb{Z}$),

$$\|\hat{T}_m f\|_{L_p(\theta)} \leq c \|f\|_{L_p(\theta^*)}, \tag{3.30}$$

where θ^* is from (3.25), and

$$\|f - \hat{T}_m f\|_{L_p(\theta)} \leq c \sum_{\theta' \in \Theta_m: \theta' \cap \theta \neq \emptyset} \omega_k(f, \theta')_p. \tag{3.31}$$

Furthermore, if $f \in L_p^{\text{loc}}$,

$$\|f - \hat{T}_m f\|_{L_p(K)} \rightarrow 0 \quad \text{as } m \rightarrow \infty \text{ for any compact } K \subset \mathbb{R}^n, \tag{3.32}$$

and if $f \in L_p$ ($L_\infty := C_0$),

$$\|f - \hat{T}_m f\|_p \rightarrow 0 \quad \text{as } m \rightarrow \infty. \tag{3.33}$$

Proof We first prove (3.30) in the case $\hat{T}_m = Q_m$ and $1 \leq p \leq \infty$ (the proof in the other cases is similar). By (3.22) and (3.15) it follows that

$$\begin{aligned} \|Q_m f\|_{L_p(\theta)} &\leq \sum_{\lambda \in \Lambda_m: \theta_\lambda \cap \theta \neq \emptyset} |\langle f, \tilde{g}_\lambda \rangle| \|g_\lambda\|_p \\ &\leq \sum_{\lambda \in \Lambda_m: \theta_\lambda \cap \theta \neq \emptyset} \|f\|_{L_p(\theta_\lambda)} \|\tilde{g}_\lambda\|_{p'} \|g_\lambda\|_p \leq c \sum_{\lambda \in \Lambda_m: \theta_\lambda \cap \theta \neq \emptyset} \|f\|_{L_p(\theta_\lambda)} \\ &\leq c \|f\|_{L_p(\theta^*)} \quad (1/p + 1/p' = 1) \end{aligned}$$

as claimed.

The proof of (3.31) relies on a Whitney-type estimate that will be given first. For a given set $E \subset \mathbb{R}^n$, denote by $E_k(f, E)_p$ the best $L_p(E)$ approximation of f from \mathcal{P}_k , i.e.,

$$E_k(f, E)_p := \inf_{P \in \mathcal{P}_k} \|f - P\|_{L_p(E)}. \tag{3.34}$$

We claim now that for $0 < p \leq \infty$ and any $\theta \in \Theta_m$,

$$E_k(f, \theta^*)_p \leq c_p \sum_{\theta' \in \mathcal{X}_\theta} \omega_k(f, \theta')_p, \tag{3.35}$$

where θ^* is defined in (3.25) and $\mathcal{X}_\theta := \{\theta' \in \Theta_m : \theta' \cap \theta \neq \emptyset\}$. Indeed, by Whitney’s estimate for a ball it follows by an affine transform that $E_k(f, \theta)_p \leq c \omega_k(f, \theta)_p$ for any ellipsoid θ . For any $\theta' \in \mathcal{X}_\theta$, let $P_{\theta'} \in \mathcal{P}_k$ be such that $\|f - P_{\theta'}\|_{L_p(\theta')} = E_k(f, \theta')_p$. Then using condition (e) on discrete covers we get

$$\begin{aligned} \|P_{\theta'} - P_\theta\|_{L_p(\theta')} &\leq c \|P_{\theta'} - P_\theta\|_{L_p(\theta' \cap \theta)} \leq c \|f - P_{\theta'}\|_{L_p(\theta' \cap \theta)} + c \|f - P_\theta\|_{L_p(\theta' \cap \theta)} \\ &\leq c \|f - P_{\theta'}\|_{L_p(\theta')} + c \|f - P_\theta\|_{L_p(\theta)} \leq c \omega_k(f, \theta')_p + c \omega_k(f, \theta)_p. \end{aligned}$$

By condition (c) on discrete covers, we know that $\#\mathcal{X}_\theta \leq N_1$. From this and the preceding estimate we conclude that

$$\begin{aligned} \|f - P_\theta\|_{L_p(\theta^*)} &\leq c \sum_{\theta' \in \mathcal{X}_\theta} \|f - P_{\theta'}\|_{L_p(\theta')} \\ &\leq c \sum_{\theta' \in \mathcal{X}_\theta} \|P_{\theta'} - P_\theta\|_{L_p(\theta')} + c \sum_{\theta' \in \mathcal{X}_\theta} \|f - P_{\theta'}\|_{L_p(\theta')} \\ &\leq c \sum_{\theta' \in \mathcal{X}_\theta} \omega_k(f, \theta')_p, \end{aligned} \tag{3.36}$$

which yields (3.35).

To prove (3.31) when \hat{T}_m is Q_m from (3.22) or T_m from (3.24) and $1 \leq p \leq \infty$, we pick $P^* \in \mathcal{P}_k$ so that $\|f - P^*\|_{L_p(\theta^*)}$ realizes the estimate in (3.35), i.e.,

$$\|f - P^*\|_{L_p(\theta^*)} \leq c \sum_{\theta' \in \mathcal{X}_\theta} \omega_k(f, \theta')_p. \tag{3.37}$$

Then using that $\hat{T}_m(P^*) = P^*$ and (3.30) we get

$$\|f - \hat{T}_m f\|_{L_p(\theta)} \leq \|f - P^*\|_{L_p(\theta)} + \|\hat{T}_m(P^* - f)\|_{L_p(\theta)} \leq c\|f - P^*\|_{L_p(\theta^*)},$$

and (3.31) follows.

We now establish (3.31) when $\hat{T}_m := T_{m,p}$. We pick $P^* \in \mathcal{P}_k$ as above, and using that $T_{m,p}P^* = P^*$ and (3.27) we get

$$P^* - T_{m,p}f = T_{m,p}P^* - T_{m,p}f = \sum_{\theta \in \Theta_m} \varphi_\theta(P^* - T_{\theta,p}f).$$

Hence

$$\begin{aligned} \|f - T_{m,p}f\|_{L_p(\theta)} &\leq c\|f - P^*\|_{L_p(\theta)} + c\|P^* - T_{m,p}f\|_{L_p(\theta)} \\ &\leq c\|f - P^*\|_{L_p(\theta)} + c \sum_{\theta' \in \Theta_m: \theta' \cap \theta \neq \emptyset} \|P^* - T_{\theta',p}f\|_{L_p(\theta')} \\ &\leq c\|f - P^*\|_{L_p(\theta)} \\ &\quad + c \sum_{\theta' \in \Theta_m: \theta' \cap \theta \neq \emptyset} (\|f - P^*\|_{L_p(\theta')} + \|f - T_{\theta',p}f\|_{L_p(\theta')}), \end{aligned}$$

which combined with (3.26) and (3.37) yields (3.31).

As was already mentioned in Sect. 2.1 (Remark 7), for any compact $K \subset \mathbb{R}^n$,

$$\max\{\text{diam } \theta : \theta \in \Theta_m, \theta \cap K \neq \emptyset\} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

This and (3.31) readily imply (3.32), which leads to (3.33). □

4 Two-Level-Split Bases

In this section we assume that T_j ($j \in \mathbb{Z}$) is one of the operators Q_j or T_j or $T_{j,p}$, if $p \geq 1$, and $T_j := T_{j,p}$, if $p < 1$, defined in Sect. 3.3. We shall use these operators and the bases $\{\Phi_j\}_{j \in \mathbb{Z}}$ from (3.11) to define two-level-split bases which will play an important role in our further development. Further, we shall make use of the following representation of consecutive level polynomial bases, defined in (3.9):

$$P_{\theta,\alpha} = \sum_{|\beta| < k} m_{\alpha,\beta}^{\theta,\eta} P_{\eta,\beta}, \quad \theta \in \Theta_j, \eta \in \Theta_{j+1}. \tag{4.1}$$

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

$$P_{\theta,\alpha} = \sum_{\eta \in \Theta_{j+1} : \theta \cap \eta \neq \emptyset} \sum_{|\beta| < k} m_{\alpha,\beta}^{\theta,\eta} P_{\eta,\beta} \varphi_\eta \quad \text{on } \theta. \tag{4.2}$$

This yields

$$\begin{aligned} T_{j+1}f - T_j f &= \sum_{\eta \in \Theta_{j+1}} \sum_{|\beta| < k} b_{\eta,\beta}(f) P_{\eta,\beta} \varphi_\eta - \sum_{\theta \in \Theta_j} \sum_{|\alpha| < k} b_{\theta,\alpha}(f) P_{\theta,\alpha} \varphi_\theta \\ &= \sum_{\theta \in \Theta_j} \varphi_\theta \sum_{\eta \in \Theta_{j+1}} \sum_{|\beta| < k} b_{\eta,\beta}(f) P_{\eta,\beta} \varphi_\eta \\ &\quad - \sum_{\theta \in \Theta_j} \sum_{|\alpha| < k} b_{\theta,\alpha}(f) \sum_{\theta \cap \eta \neq \emptyset} \sum_{|\beta| < k} m_{\alpha,\beta}^{\theta,\eta} P_{\eta,\beta} \varphi_\theta \varphi_\eta \\ &= \sum_{\eta \in \Theta_{j+1}} \sum_{\theta \in \Theta_j : \theta \cap \eta \neq \emptyset} \sum_{|\beta| < k} \left\{ b_{\eta,\beta}(f) - \sum_{|\alpha| < k} m_{\alpha,\beta}^{\theta,\eta} b_{\theta,\alpha}(f) \right\} \\ &\quad \times P_{\eta,\beta} \varphi_\eta \varphi_\theta, \end{aligned} \tag{4.3}$$

where $b_\lambda(f)$ is given by (3.29), depending of the choice of T_m . Thus, setting

$$\mathcal{M}_j := \{ \mu = (\eta, \theta, \beta) : \eta \in \Theta_{j+1}, \theta \in \Theta_j, \theta \cap \eta \neq \emptyset, |\beta| < k \}, \quad j \in \mathbb{Z},$$

the building blocks in (4.3) have the form

$$F_\mu := P_{\eta,\beta} \varphi_\eta \varphi_\theta, \quad \mu = (\eta, \theta, \beta) \in \mathcal{M}_j, \tag{4.4}$$

where $P_{\eta,\beta}$ are defined in (3.9) and $\varphi_\eta, \varphi_\theta$ are from (3.4). We define

$$\mathcal{F}_j := \{ F_\mu : \mu \in \mathcal{M}_j \} \quad \text{and} \quad W_j := \text{span } \mathcal{F}_j, \quad j \in \mathbb{Z}. \tag{4.5}$$

To estimate the L_p -norms of the F_μ we shall make use of the following fact:

Lemma 4.1 *The coefficients in (4.1) are uniformly bounded, that is,*

$$|m_{\alpha,\beta}^{\theta,\eta}| \leq c < \infty, \tag{4.6}$$

where c depend only on $\mathbf{p}(\Theta)$, L , and k .

Proof Using (3.9), identity (4.1) can be rewritten as

$$P_\alpha \circ A_\theta^{-1} \circ A_\eta = \sum_{|\beta| < k} m_{\alpha,\beta}^{\theta,\eta} P_\beta,$$

and since $\{P_\beta\}$ is an orthonormal basis for \mathcal{P}_k in $L_2(B^*)$, we obtain

$$|m_{\alpha,\beta}^{\theta,\eta}| = \left| \int_{B^*} (P_\alpha \circ A_\theta^{-1} \circ A_\eta)(y) P_\beta(y) dy \right| \leq c,$$

where we have used condition (b) on Θ (see Definition 2.1). □

Note that $\text{supp } F_\mu = \overline{\theta \cap \eta}$ if $\mu = (\eta, \theta, \beta)$, and $\|F_\mu\|_2 \sim 1$, which implies $\|F_\mu\|_\infty \sim |\theta \cap \eta|^{-1/2} \sim |\eta|^{-1/2}$ and

$$\|F_\mu\|_p \sim |\eta|^{1/p-1/2}, \quad 0 < p \leq \infty. \tag{4.7}$$

The main result of this section now reads as follows.

Theorem 4.2 *Any $f \in W_j$ has a unique representation*

$$f = \sum_{\mu \in \mathcal{M}_j} c_\mu(f) F_\mu, \tag{4.8}$$

where the dual functionals $c_\mu(\cdot)$ are of the following form: For each $\mu \in \mathcal{M}_j$, $\mu = (\eta, \theta, \beta)$, there is an ellipsoid $B_\mu \subset \theta \cap \eta$ with $|B_\mu| \sim |\eta|$ and $B_\mu = A_\eta(B_\mu^*)$ for some ball $B_\mu^* \subset B^*$ such that

$$c_\mu(f) = \langle f, \tilde{F}_\mu \rangle, \quad \text{where } \text{supp } \tilde{F}_\mu \subset \overline{B}_\mu, \quad \|\tilde{F}_\mu\|_p \sim |\eta|^{1/p-1/2}. \tag{4.9}$$

Moreover, if $f \in W_j$ and $f = \sum_{\mu \in \mathcal{M}_j} a_\mu F_\mu$, then

$$\|f\|_p \sim \left(\sum_{\mu \in \mathcal{M}_j} \|a_\mu F_\mu\|_p^p \right)^{1/p}, \quad 0 < p \leq \infty, \tag{4.10}$$

with the obvious modification when $p = \infty$. Here all constants depend only on $\mathbf{p}(\Theta)$, L , and k .

The proof of this theorem is a mere repetition of the proof of Theorem 3.3 and is omitted. We observe, however, that the ‘‘coloring’’ scheme from Sect. 3.1 and the definition of the single-level bases $\{\Phi_m\}$ in Sect. 3.2 were specifically designed to ensure the linear independence of the two-level-split bases.

We are now prepared to derive multi-level decompositions of functions using the two-level-split bases from above.

Theorem 4.3 *For any $f \in L_p^{\text{loc}}$, $0 < p \leq \infty$,*

$$f = T_0 f + \sum_{j \geq 0} (T_{j+1} f - T_j f) = \sum_{j \geq -1} \sum_{\mu \in \mathcal{M}_j} d_\mu(f) F_\mu, \tag{4.11}$$

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

$$d_\mu(f) = b_{\eta,\beta}(f) - \sum_{|\alpha| < k} m_{\alpha,\beta}^{\theta,\eta} b_{\theta,\beta}(f), \quad \mu := (\eta, \theta, \beta), \tag{4.12}$$

with $m_{\alpha,\beta}^{\theta,\eta}$ from (4.1), while $\mathcal{M}_{-1} := \Lambda_0$, $F_\lambda := g_\lambda$, and $d_\lambda(f) := b_\lambda(f)$ if $\lambda \in \mathcal{M}_{-1}$. Moreover, if $f \in L_p$ ($L_\infty := C_0$), then (4.11) as well as

$$f = \sum_{j \in \mathbb{Z}} (T_{j+1} - T_j) f \tag{4.13}$$

hold in L_p .

Proof The relation (4.12) has been derived in (4.3) and the convergence of the left-hand side identity in (4.11) in $L_p(K)$, and in L_p follows from Lemma 3.4. To prove the convergence of (4.13), one has to show that $\|T_j f\|_p \rightarrow 0$ when $j \rightarrow -\infty$. Since this part of the claim will not be essential for the subsequent developments, we only sketch the argument. Recall from Remark 7 in Sect. 2.1 that each $\theta \in \Theta_j$ expands to \mathbb{R}^n as $j \rightarrow -\infty$. This implies that the basis elements F_μ , $\mu \in \mathcal{M}_j$, become flatter and flatter as $j \rightarrow -\infty$, which easily leads to the desired convergence (also see the proof of Remark 5.1 below). □

5 Anisotropic Besov Spaces (B-spaces)

In this section we introduce anisotropic Besov spaces induced by discrete ellipsoid covers of \mathbb{R}^n and show that they can be characterized by the two-level-split bases from the previous section. Our approach to B-spaces is based on local polynomial approximation on the ellipsoids of the underlying discrete ellipsoid cover Θ of \mathbb{R}^n . Recall that the classical Besov spaces can also be defined via local polynomial approximation (local oscillations), see, e.g., [19]. Furthermore, the anisotropic Besov space from [10, 14] are defined via local piecewise polynomial approximation. We maintain that local approximations rather than global means 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.

5.1 Homogeneous B-spaces

We now introduce the homogeneous B-spaces $\dot{B}_{pq}^\alpha(\Theta)$ induced by an arbitrary discrete ellipsoid cover Θ of \mathbb{R}^n . In the definition of the B-spaces $\dot{B}_{pq}^\alpha(\Theta)$ there is a hidden parameter k which we choose to be the smallest integer satisfying the condition

$$k > \frac{a_0}{a_6} \cdot \frac{\alpha}{n}. \tag{5.1}$$

This will guarantee the equivalence of the norms in $\dot{B}_{pq}^\alpha(\Theta)$ introduced below. Here a_0 and a_6 are two of the parameters of Θ (see Definition 2.1). It can be shown that the norms in $\dot{B}_{pq}^\alpha(\Theta)$ with different k 's satisfying (5.1) are equivalent.

Definition of $\dot{B}_{pq}^\alpha(\Theta)$ via Local Moduli of Smoothness For a given discrete ellipsoid cover Θ of \mathbb{R}^n and $\alpha > 0$, $0 < p, q \leq \infty$, we define the space $\dot{B}_{pq}^\alpha(\Theta)$ as the set of all functions $f \in L_p^{\text{loc}}$ such that

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)} := \left(\sum_{j \in \mathbb{Z}} \left(\sum_{\theta \in \Theta_j} |\theta|^{-\alpha p/n} \omega_k(f, \theta)_p^p \right)^{q/p} \right)^{1/q} < \infty, \tag{5.2}$$

where $\omega_k(f, \theta)_p$ is the k th local modulus of smoothness of f (see (3.21)).

As in (3.34), let $E_k(f, \theta)_p$ denote the best L_p approximation of f on θ from \mathcal{P}_k . By Whitney’s theorem, $E_k(f, \theta)_p \sim \omega_k(f, \theta)_p$, and hence replacing $\omega_k(f, \theta)_p$ in (5.2) by $E_k(f, \theta)_p$ will result in an equivalent norm.

The definition of $\dot{B}_{pq}^\alpha(\Theta)$ needs some *additional clarification*. Evidently the norm in $\dot{B}_{pq}^\alpha(\Theta)$ is not a true norm because $\|P\|_{\dot{B}_{pq}^\alpha(\Theta)} = 0$ for $P \in \mathcal{P}_k$. Hence $\dot{B}_{pq}^\alpha(\Theta)$ is a quotient space modulo \mathcal{P}_k . We shall use the operators Q_j , T_j , and $T_{j,p}$ from Sect. 3.3 to construct a representer for each $f \in \dot{B}_{pq}^\alpha(\Theta)$. As before, let T_j ($j \in \mathbb{Z}$) be one of the operators Q_j or T_j or $T_{j,p}$ if $p \geq 1$, and $T_j := T_{j,p}$ if $p < 1$. We define

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^T := \left(\sum_{j \in \mathbb{Z}} (2^{a_0 j \alpha/n} \|(T_{j+1} - T_j)f\|_p)^q \right)^{1/q}. \tag{5.3}$$

On account of property (c) of ellipsoid covers, Lemma 3.4 yields

$$\|f - T_j f\|_p \leq c \left(\sum_{\theta \in \Theta_j} \omega_k(f, \theta)_p^p \right)^{1/p},$$

and using that $\|(T_{j+1} - T_j)f\|_p \leq c\|f - T_{j+1}f\|_p + c\|f - T_jf\|_p$, we obtain

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^T \leq c\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}. \tag{5.4}$$

To describe the nature of $\dot{B}_{pq}^\alpha(\Theta)$ we have to distinguish two basic cases.

Case 1: $0 < \alpha < n/p$ or $\alpha = n/p$ and $q \leq 1$.

We address first the scale components $(T_{j+1} - T_j)f$ for negative j .

Remark 5.1 Whenever $f \in \dot{B}_{pq}^\alpha(\Theta)$ for $0 < \alpha < n/p$ or $\alpha = n/p$ and $q \leq 1$, there exists $P \in \mathcal{P}_k$ such that the series

$$\sum_{j=-\infty}^{-1} (T_{j+1} - T_j)f = T_0 f - P \tag{5.5}$$

converges uniformly on \mathbb{R}^n .

Proof Assume first that $0 < \alpha < n/p$. By (5.3)–(5.4) it follows that

$$\|(T_{j+1} - T_j)f\|_p \leq c2^{-a_0 j \alpha/n} \|f\|_{\dot{B}_{pq}^\alpha(\Theta)}.$$

Since $(T_{j+1} - T_j)f \in W_j$ as in Theorem 4.3, we have $(T_{j+1} - T_j)f = \sum_{\mu \in \mathcal{M}_j} d_\mu F_\mu$. For $\mu = (\eta, \theta, \beta)$ denote η by η_μ . Now using Theorem 4.2 and (4.7),

$$\begin{aligned} & \| (T_{j+1} - T_j)f \|_\infty \\ & \leq c \sup_{\mu \in \mathcal{M}_j} \|d_\mu F_\mu\|_\infty \leq c \sup_{\mu \in \mathcal{M}_j} |\eta_\mu|^{-1/p} \|d_\mu F_\mu\|_p \\ & \leq c 2^{a_0 j/p} \left(\sum_{\mu \in \mathcal{M}_j} \|d_\mu F_\mu\|_p \right)^{1/p} \leq c 2^{a_0 j/p} \| (T_{j+1} - T_j)f \|_p. \end{aligned} \tag{5.6}$$

Therefore,

$$\| (T_{j+1} - T_j)f \|_\infty \leq c 2^{a_0 j(1/p - \alpha/n)} \|f\|_{\dot{B}_{pq}^\alpha(\Theta)},$$

which implies

$$\sum_{j=-\infty}^{-1} \| (T_{j+1} - T_j)f \|_\infty \leq c \|f\|_{\dot{B}_{pq}^\alpha(\Theta)}, \tag{5.7}$$

since by assumption $\alpha < n/p$. Now setting $P := T_0 f - \sum_{j=-\infty}^{-1} (T_{j+1} - T_j)f$, and noting that each $T_j f$ belongs to $C(\mathbb{R}^n)$, we infer from (5.7) that $P \in C(\mathbb{R}^n)$ and

$$T_j f \rightrightarrows P \quad \text{as } j \rightarrow -\infty, \text{ uniformly on } \mathbb{R}^n. \tag{5.8}$$

It is easy to see that $P \in \mathcal{P}_k$. Indeed, since the ellipsoids in Θ expand to \mathbb{R}^n (see Remark 7 after Definition 2.1), any compact $K \subset \mathbb{R}^n$ is contained in a sequence $\{\theta_v^\diamond\}_{v=1}^\infty$ of ellipsoids dilated by a factor $a_7 < 1$ (see property (d) of Θ) such that $\ell(\theta_1) > \ell(\theta_2) > \dots$ and $\bigcup_{v \geq 0} \theta_v = \mathbb{R}^n$, where $\ell(\theta) = j$ means $\theta \in \Theta_j$. This implies $\varphi_{\theta_v} \rightrightarrows 1$ as $v \rightarrow \infty$ on K . Now taking into account the definition of T_j this leads to $T_j f \rightrightarrows P_K$ as $j \rightarrow -\infty$ for some $P_K \in \mathcal{P}_k$ on all compact sets K . Therefore, (5.8) holds for some $P \in \mathcal{P}_k$, and (5.5) follows.

If $\alpha = n/p$ and $q \leq 1$, we argue in a similar fashion that

$$\begin{aligned} \left(\sum_{j=-\infty}^{-1} \| (T_{j+1} - T_j)f \|_\infty \right)^q & \leq \sum_{j=-\infty}^{-1} \| (T_{j+1} - T_j)f \|_\infty^q \\ & \leq \sum_{j=-\infty}^{-1} 2^{a_0 j q \alpha/n} \| (T_{j+1} - T_j)f \|_p^q \leq c \|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^q, \end{aligned}$$

which also yields (5.7) and allows us to continue as before. □

We now turn to the series $\sum_{j=0}^\infty (T_{j+1} - T_j)f$.

Remark 5.2 If $f \in \dot{B}_{pq}^\alpha(\Theta)$, $\alpha > 0$, $0 < p, q \leq \infty$, then

$$\sum_{j=0}^\infty (T_{j+1} - T_j)f = f - T_0 f \quad \text{in } L_p. \tag{5.9}$$

Proof First let $p \geq 1$. Since by (5.4)

$$\sum_{j=0}^{\infty} (2^{a_0 j \alpha/n} \|(T_{j+1} - T_j)f\|_p)^q \leq c \|f\|_{\dot{B}_{pq}^\alpha}^q, \tag{5.10}$$

we have $\|(T_{j+1} - T_j)f\|_p \leq c 2^{-a_0 j \alpha/n} \|f\|_{\dot{B}_{pq}^\alpha}$, and therefore

$$\sum_{j=0}^{\infty} \|(T_{j+1} - T_j)f\|_p \leq c \|f\|_{\dot{B}_{pq}^\alpha}. \tag{5.11}$$

One readily infers from (5.11) that the partial sums $\sum_{j=0}^J (T_{j+1} - T_j)f$ form a Cauchy sequence in L_p . By Lemma 3.4 the limit must agree with $f - T_0 f$, which confirms (5.10) for $p \geq 1$.

For $p < 1$ we shall first show that

$$\sum_{j=0}^{\infty} \|(T_{j+1} - T_j)f\|_p^p \leq c \|f\|_{\dot{B}_{pq}^\alpha}^p. \tag{5.12}$$

In fact, when $q \leq p$ we deduce from (5.10) that

$$\sum_{j=0}^{\infty} 2^{a_0 j \alpha/n} \|(T_{j+1} - T_j)f\|_p \leq c \|f\|_{\dot{B}_{pq}^\alpha},$$

and therefore $\|(T_{j+1} - T_j)f\|_p^p \leq c 2^{-a_0 j p \alpha/n} \|f\|_{\dot{B}_{pq}^\alpha}^p$, which yields (5.12) for $q \leq p$.

When $q > p$ we apply Hölder’s inequality to

$$\sum_{j=0}^{\infty} \|(T_{j+1} - T_j)f\|_p^p = \sum_{j=0}^{\infty} 2^{-a_0 j \alpha p/n} (2^{a_0 j \alpha/n} \|(T_{j+1} - T_j)f\|_p)^p$$

with $\tilde{p} := q/p > 1$ to conclude (5.12) in this case as well. Now since

$$\left\| \sum_{j=J}^{J+1} (T_{j+1} - T_j)f \right\|_p^p \leq \sum_{j=J}^{J+1} \|(T_{j+1} - T_j)f\|_p^p \rightarrow 0 \quad \text{as } J \rightarrow \infty,$$

and using again that $T_j f \rightarrow f$ in $L_p(K)$ on each compact $K \subset \mathbb{R}^n$ (see Lemma 3.4), we arrive at (5.9) also in the case $p < 1$. □

Remarks 5.1–5.2 reveal the structure of the functions in $\dot{B}_{pq}^\alpha(\Theta)$ under the above circumstances.

Remark 5.3 Let $0 < \alpha < n/p$ or $\alpha = n/p$ and $q \leq 1$. Then for any $f \in \dot{B}_{pq}^\alpha(\Theta)$ there exists $P \in \mathcal{P}_k$ such that $f = \sum_{j \in \mathbb{Z}} (T_{j+1} - T_j)f + P$ in $L_p(K)$ for all compact

sets $K \subset \mathbb{R}^n$. Thus, one can define $\dot{B}_{pq}^\alpha(\Theta)$, in this case, as the set of all functions $f \in L_p^{\text{loc}}$ such that $\|f\|_{\dot{B}_{pq}^\alpha} < \infty$ and

$$f = \sum_{j \in \mathbb{Z}} (T_{j+1} - T_j) f \quad \text{in } L_p(K) \text{ for all compact sets } K \subset \mathbb{R}^n. \tag{5.13}$$

In addition, by (5.7) and (5.12) it follows that

$$\|f\|_{L_p(K)} \leq c(1 + |K|^{1/p}) \|f\|_{\dot{B}_{pq}^\alpha} \quad \text{for all compact sets } K \subset \mathbb{R}^n. \tag{5.14}$$

In order to refine the above observations, we consider the following two subcases corresponding to lower, and respectively higher, smoothness α .

Case 1 (a): $0 < \alpha < n(1/p - 1)$ or $0 < \alpha = n(1/p - 1)$ and $q > 1$.

In this case we necessarily have $0 < p < 1$. Moreover, if we were to deal with the classical Besov spaces, the above condition on α means that embedding in \mathcal{S}' fails. Thus, we are content with the convergence from above in this case.

Alternatively we have:

Case 1 (b): $n(1/p - 1) < \alpha < n/p$ or $\alpha = n(1/p - 1)$ and $q \leq 1$ or $\alpha = n/p$ and $q \leq 1$.

For classical Besov spaces this is known to imply that \dot{B}_{pq}^α is embedded in \mathcal{S}' , which persists to hold here as well, as will be shown next. The main observation can be stated as follows:

Remark 5.4 Let $\alpha > n(1/p - 1)$ or $\alpha = n(1/p - 1)$ and $q \leq 1$. Then for any $f \in \dot{B}_{pq}^\alpha(\Theta)$,

$$\sum_{j=0}^{\infty} (T_{j+1} - T_j) f = f - T_0 f \quad \text{in } \mathcal{S}'. \tag{5.15}$$

Proof From Remark 5.2 we already know that $\sum_{j=0}^{\infty} (T_{j+1} - T_j) f$ converges in L_p and thus in \mathcal{S}' whenever $p \geq 1$, i.e., (5.15) holds for $p \geq 1$.

Suppose now that $p < 1$. Since $(T_{j+1} - T_j) f \in W_j$ similarly as in (5.6),

$$\begin{aligned} \|(T_{j+1} - T_j) f\|_1 &\leq \sum_{\mu \in \mathcal{M}_j} \|d_\mu F_\mu\|_1 \leq c \sum_{\mu \in \mathcal{M}_j} |\eta_\mu|^{1-1/p} \|d_\mu F_\mu\|_p \\ &\leq c 2^{-a_0 j(1-1/p)} \left(\sum_{\mu \in \mathcal{M}_j} \|d_\mu F_\mu\|_p^p \right)^{1/p} \\ &\leq c 2^{-a_0 j(1-1/p)} \|(T_{j+1} - T_j) f\|_p. \end{aligned}$$

Thus, on account of (5.10), we obtain

$$\sum_{j=0}^{\infty} (2^{a_0 j[\alpha/n - (1/p - 1)]} \|(T_{j+1} - T_j) f\|_p)^q \leq c \|f\|_{\dot{B}_{pq}^\alpha}^q,$$

which implies

$$\sum_{j=0}^{\infty} \|(T_{j+1} - T_j)f\|_1 \leq c \|f\|_{\dot{B}_{pq}^\alpha}.$$

By the argument in the proof of Remark 5.2, see (5.11), we conclude that

$$\sum_{j=0}^{\infty} (T_{j+1} - T_j)f = f - T_0f \quad \text{in } L_1 \text{ and hence in } \mathcal{S}', \tag{5.16}$$

which completes the proof. □

In summary, from Remark 5.1 and 5.4 we can state the following:

Remark 5.5 Let $\alpha > 0$, and suppose $n(1/p - 1) < \alpha < n/p$ or $\alpha = n(1/p - 1)$ and $q \leq 1$ or $\alpha = n/p$ and $q \leq 1$. Then, for any $f \in \dot{B}_{pq}^\alpha(\Theta)$ there exists $P \in \mathcal{P}_k$ such that $f = \sum_{j \in \mathbb{Z}} (T_{j+1} - T_j)f + P$ converges in \mathcal{S}' . Thus, the space $\dot{B}_{pq}^\alpha(\Theta)$ can be defined as the set of all regular tempered distributions f such that $\|f\|_{\dot{B}_{pq}^\alpha(\Theta)} < \infty$ and

$$f = \sum_{j \in \mathbb{Z}} (T_{j+1} - T_j)f \quad \text{in } \mathcal{S}'. \tag{5.17}$$

Also, it follows from the above that

$$|\langle f, \phi \rangle| \leq c \|f\|_{\dot{B}_{pq}^\alpha(\Theta)} \|\phi\|_\infty, \quad \forall \phi \in \mathcal{S},$$

i.e., $\dot{B}_{pq}^\alpha(\Theta)$ is continuously embedded in \mathcal{S}' .

We now turn to:

Case 2: $\alpha > n/p$ or $\alpha = n/p$ and $q > 1$.

As in Case 1 (b), since $\alpha > n(1/p - 1)$, we have

$$\sum_{j=0}^{\infty} (T_{j+1} - T_j)f = f - T_0f \quad \text{in } \mathcal{S}'. \tag{5.18}$$

But now dealing with the series $\sum_{j=-\infty}^{-1} (T_{j+1} - T_j)f$ is more complicated because it is not convergent. As in the classical case we need to renormalize this series.

Lemma 5.6 *There exist polynomials $P \in \mathcal{P}_k$ and $P_j \in \mathcal{P}_k, j = -1, -2, \dots$, such that*

$$T_0f = P + \lim_{m \rightarrow -\infty} \sum_{j=m}^{-1} (T_{j+1} - T_j)f - P_j \quad \text{in } \mathcal{S}' \tag{5.19}$$

and also uniformly on every compact set $K \subset \mathbb{R}^n$. Moreover, P_j ($j \leq -1$) can be defined as the $k - 1$ degree Taylor polynomial of the function $(T_{j+1} - T_j)f$ based at 0.

Proof From (5.3)–(5.4) it follows that

$$\sum_{j=-\infty}^{-1} (2^{a_0 j \alpha/n} \|(T_{j+1} - T_j)f\|_p)^q \leq c \|f\|_{\dot{B}_{pq}^s(\Theta)}^q =: \mathcal{N}^q. \tag{5.20}$$

Hence

$$\|(T_{j+1} - T_j)f\|_p \leq 2^{-a_0 j \alpha/n} \mathcal{N},$$

and as in (5.6),

$$\|(T_{j+1} - T_j)f\|_\infty \leq c 2^{-a_0 j (\alpha/n - 1/p)} \mathcal{N}, \quad j \leq -1. \tag{5.21}$$

Let P_j be the $k - 1$ degree Taylor polynomial of $(T_{j+1} - T_j)f$ based at $x = 0$. We need to estimate $|(T_{j+1} - T_j)f(x) - P_j(x)|$ for all $x \in \mathbb{R}^n$. To this end we next estimate $|\partial^\beta (T_{j+1} - T_j)f(x)|$ whenever $|\beta| = k$.

As before, we can write $(T_{j+1} - T_j)f = \sum_{\mu \in \mathcal{M}_j} d_\mu F_\mu$. From (3.6) it readily follows that if $\mu = (\eta, \theta, \beta)$, then

$$\|\partial^\beta F_\mu\|_\infty \leq c (\|A_\eta^{-1}\|_{\ell_2 \rightarrow \ell_2})^{|\beta|} \|F_\mu\|_\infty \leq c [\sigma_{\min}(\eta)]^{-k} \|F_\mu\|_\infty.$$

Fix $x \in \mathbb{R}^n$, and let $\eta \in \Theta_{j+1}$ be such that $x \in \eta$. From above, Theorem 4.2, and the properties of Θ , we infer

$$\begin{aligned} |\partial^\beta (T_{j+1} - T_j)f(x)| &\leq \sum_{\mu \in \mathcal{M}_j: x \in \eta_\mu} |d_\mu \partial^\beta F_\mu(x)| \\ &\leq c [\sigma_{\min}(\eta)]^{-k} \sum_{\mu \in \mathcal{M}_j: x \in \eta_\mu} \|d_\mu F_\mu\|_\infty \\ &\leq c [\sigma_{\min}(\eta)]^{-k} \|(T_{j+1} - T_j)f\|_\infty. \end{aligned} \tag{5.22}$$

Let $\theta_0 \in \Theta_0$ be such that $0 \in \theta_0$, and denote briefly $\sigma_0 := \sigma_{\min}(\theta_0)$. Also, let $\theta' \in \Theta$ be an ellipsoid of maximum level $\leq j$ such that $0, x \in \theta'$. Then $\theta' \in \Theta_{j-r}$ for some $r \geq 0$.

Consider first the case when $r \geq 1$. By Lemma 2.2 $\sigma_{\min}(\theta') \geq (1/a_5) 2^{-a_6(j-r)} \sigma_0$, and since $\eta \cap \theta' \neq \emptyset$, from this and again by Lemma 2.2, it follows that

$$\sigma_{\min}(\eta) \geq a_3 2^{-a_4 r} \sigma_{\min}(\theta') \geq c 2^{-a_6(j-r)} 2^{-a_4 r} \sigma_0 = c 2^{-(a_4 - a_6)r} 2^{-a_6 j} \sigma_0.$$

From this and (5.21)–(5.22) we infer

$$\begin{aligned} |\partial^\beta (T_{j+1} - T_j)f(x)| &\leq c [\sigma_{\min}(\eta)]^{-k} \|(T_{j+1} - T_j)f\|_\infty \\ &\leq c \sigma_0^{-k} 2^{(a_4 - a_6)rk} 2^{a_6 j k} 2^{-a_0 j (\alpha/n - 1/p)} \mathcal{N} \\ &= c \sigma_0^{-k} 2^{(a_4 - a_6)rk} 2^{j \varepsilon} \mathcal{N}. \end{aligned} \tag{5.23}$$

We need a lower bound for $|x|$ in terms of j and r . By property (d) of Θ there exists an ellipsoid $\theta'' \in \Theta_{j-r+1}$ such that $0 \in \theta''^\circ$, where θ''° is the dilated version of θ'' by a factor of $a_7 < 1$. Since θ' was selected of minimum level so that $0, x \in \theta'$, then $x \notin \theta''$, and hence $|x| \geq (1 - a_7)\sigma_{\min}(\theta'')$. On the other hand, by Lemma 2.2,

$$\sigma_{\min}(\theta'') \geq c\sigma_{\min}(\theta') \geq c\sigma_0 2^{-a_6(j-r)}.$$

Hence $|x| \geq c\sigma_0 2^{-a_6(j-r)} \geq c\sigma_0 2^{a_6 r}$. Combining this with (5.23), we get

$$|\partial^\beta (T_{j+1} - T_j)f(x)| \leq c'(|x| + 1)^\gamma 2^{j\varepsilon} \mathcal{N}, \quad x \in \mathbb{R}^n, \tag{5.24}$$

where $\gamma := \frac{a_4 - a_6}{a_6}k$ and $c' = c\sigma_0^{-ka_4/a_6}$.

The case when $r = 0$ is easier. Then $\sigma_{\min}(\eta) \geq c\sigma_{\min}(\eta_0) \geq c2^{-a_6 j}\sigma_0$, and by (5.22) it follows that (5.23) holds with $r = 0$. Hence (5.24) holds again.

Now, Taylor’s theorem gives

$$\begin{aligned} |(T_{j+1} - T_j)f(x) - P_j(x)| &\leq c \sum_{|\beta|=k} \|\partial^\beta (T_{j+1} - T_j)f\|_{L_\infty(B(0,|x|))} |x|^k \\ &\leq c(|x| + 1)^{\gamma+k} 2^{j\varepsilon} \mathcal{N}. \end{aligned} \tag{5.25}$$

Therefore, for any $R > 0$

$$\sum_{j=-\infty}^{-1} \|(T_{j+1} - T_j)f - P_j\|_{L_\infty(B(0,R))} < \infty. \tag{5.26}$$

Setting

$$P := T_0 f - \sum_{j=-\infty}^{-1} (T_{j+1} - T_j)f - P_j \quad \text{and} \quad Q_j := \sum_{v=j}^{-1} P_v, \quad Q_j \in \mathcal{P}_k,$$

we evidently have $P \in C(\mathbb{R}^n)$. Then by (5.26) $T_j f + Q_j \rightrightarrows P$ as $j \rightarrow -\infty$ uniformly on $B(0, R)$ for all $R > 0$. As in Case 1, one easily shows that $P \in \mathcal{P}_k$.

We now employ (5.25) to find that for any $\phi \in \mathcal{S}$,

$$\begin{aligned} \left| \left\langle \sum_{j=-\infty}^m (T_{j+1} - T_j)f - P_j, \phi \right\rangle \right| &\leq c \sum_{j=-\infty}^m 2^{j\varepsilon} \mathcal{N} | \langle (|\cdot| + 1)^{\gamma+k}, \phi \rangle | \\ &\leq c2^{m\varepsilon} \mathcal{N} \|\phi\|_{\gamma+k+n+1} \rightarrow 0, \quad m \rightarrow -\infty, \end{aligned} \tag{5.27}$$

where $\|\phi\|_s := \|(|x| + 1)^s \phi(x) \|_\infty$. Hence

$$\sum_{j=-\infty}^{-1} (T_{j+1} - T_j)f - P_j = T_0 f - P \quad \text{in } \mathcal{S}' \ (P \in \mathcal{P}_k),$$

which completes the proof of Lemma 5.6. □

In summary, from (5.18) and Lemma 5.6 we conclude the following:

Remark 5.7 Whenever $\alpha > n/p$ or $\alpha = n/p$ and $q > 1$, the space $\dot{B}_{pq}^\alpha(\Theta)$ can be viewed as the set of all regular tempered distributions f such that $\|f\|_{\dot{B}_{pq}^\alpha(\Theta)} < \infty$ and

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

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

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

Moreover, $\dot{B}_{pq}^\alpha(\Theta)$ can be defined as the set of all $f \in S'$ such that $\|f\|_{\dot{B}_{pq}^\alpha(\Theta)} < \infty$ and

$$f = \sum_{j \in \mathbb{Z}} (T_{j+1} - T_j)f - P_j \quad \text{in } S',$$

where the polynomials P_j are defined as in Lemma 5.6 for $j \leq -1$ and $P_j := 0$ if $j \geq 0$. In addition, $\dot{B}_{pq}^\alpha(\Theta)$ is continuously embedded in S' , namely, there exists $s > 0$ such that for any $f \in \dot{B}_{pq}^\alpha(\Theta)$,

$$|\langle f, \phi \rangle| \leq c \|f\|_{\dot{B}_{pq}^\alpha(\Theta)} \|\phi\|_s, \quad \forall \phi \in \mathcal{S}.$$

This last estimate follows from (5.18) and (5.27).

Other Norms in $\dot{B}_{pq}^\alpha(\Theta)$ The understanding of the B-spaces relies to a great extent on having several equivalent norms for $\dot{B}_{pq}^\alpha(\Theta)$ at hand, as will be exemplified later for best m -term approximation with the two-level-split basis functions. Observe first that if

$$(T_{j+1} - T_j)f = \sum_{\mu \in \mathcal{M}_j} d_\mu(f)F_\mu,$$

then, using Theorem 4.2, we can write

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^T \sim \left(\sum_{j \in \mathbb{Z}} \left(\sum_{\mu \in \mathcal{M}_j} (|\eta_\mu|^{-\alpha/n} \|d_\mu(f)F_\mu\|_p)^p \right)^{q/p} \right)^{1/q}. \quad (5.28)$$

We also define

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^A := \inf_{f = \sum_{\mu \in \mathcal{M}} a_\mu F_\mu} \left(\sum_{j \in \mathbb{Z}} \left(\sum_{\mu \in \mathcal{M}_j} (|\eta_\mu|^{-\alpha/n} \|a_\mu F_\mu\|_p)^p \right)^{q/p} \right)^{1/q}. \quad (5.29)$$

Here the infimum is taken over all representations $f = \sum_{\mu \in \mathcal{M}} a_\mu F_\mu$, where the convergence is to be understood as described above.

Central for this part is the equivalence of the above norms.

Theorem 5.8 *If $\alpha > 0, 0 < p, q \leq \infty$, and (5.1) is obeyed, then the norms $\|\cdot\|_{\dot{B}_{pq}^\alpha(\Theta)}$, $\|\cdot\|_{\dot{B}_{pq}^\alpha(\Theta)}^T$, and $\|\cdot\|_{\dot{B}_{pq}^\alpha(\Theta)}^A$ are equivalent.*

We defer the somewhat lengthy proof of this theorem to the appendix.

One uses in a standard way the embedding of \dot{B}_{pq}^α in \mathcal{S}' or (5.14) to show its completeness. We record this in the following proposition:

Proposition 5.9 *For $0 < p, q \leq \infty$ and $\alpha > 0$, the B-space \dot{B}_{pq}^α is a quasi-Banach space which is continuously embedded in \mathcal{S}' if $\alpha > n(1/p - 1)$ or $\alpha = n(1/p - 1)$ and $q \leq 1$.*

5.2 Inhomogeneous B-spaces

Sometimes it is more convenient to use the inhomogeneous version $B_{pq}^\alpha(\Theta^+)$ of the B-spaces induced by anisotropic ellipsoid covers of \mathbb{R}^n . For the inhomogeneous B-spaces we only need ellipsoid covers with levels $j = 0, 1, \dots$, i.e., covers of the form

$$\Theta^+ := \bigcup_{j=0}^{\infty} \Theta_j.$$

In a certain sense, the spaces $B_{pq}^\alpha(\Theta^+)$ are simpler than the homogeneous counterparts $\dot{B}_{pq}^\alpha(\Theta)$, which justifies our fairly brief exposition.

As for the homogeneous B-spaces, there is a hidden index k which is also present in the definition of basis elements. It will always be selected to be the minimum integer such that

$$k > \frac{a_0}{a_6} \cdot \frac{\alpha}{n}, \tag{5.30}$$

where $a_0, a_6 > 0$ are from the definition of discrete ellipsoid covers in Sect. 2.1.

Definition of $B_{pq}^\alpha(\Theta^+)$ via Local Moduli of Smoothness For a given discrete ellipsoid cover Θ^+ of \mathbb{R}^n and $\alpha > 0, 0 < p, q \leq \infty$, we define the space $B_{pq}^\alpha(\Theta^+)$ as the set of all functions $f \in L_p$ such that

$$|f|_{B_{pq}^\alpha(\Theta^+)} := \left(\sum_{j \geq 0} \left(\sum_{\theta \in \Theta_j} (|\theta|^{-\alpha p/n} \omega_k(f, \theta)_p)^q \right)^{1/q} < \infty, \tag{5.31}$$

where $\omega_k(f, \theta)_p$ is the k th local modulus of smoothness of f in $L_p(\theta)$. The norm in B_{pq}^α is defined by

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

Other Norms on $B_{pq}^\alpha(\Theta^+)$ As in the homogeneous case, let T_j ($j \in \mathbb{Z}$) be one of the operators Q_j or $T_{j,p}$, if $p \geq 1$, and $T_j := T_{j,p}$, if $p < 1$, defined in Sect. 3.3.

In view of the representation of functions $f \in L_p$ from Theorem 4.3, we define

$$\|f\|_{B_{pq}^\alpha(\Theta^+)}^T := \|T_0 f\|_p + \left(\sum_{j \geq 0} (2^{a_0 j \alpha/n} \| (T_{j+1} - T_j) f \|_p)^q \right)^{1/q}. \tag{5.32}$$

By Theorem 4.2 and (4.10),

$$\|f\|_{B_{pq}^\alpha(\Theta^+)}^T \sim \left(\sum_{j \geq -1} \left(\sum_{\mu \in \mathcal{M}_j} (|\eta_\mu|^{-\alpha/n} \|d_\mu(f) F_\mu\|_p)^p \right)^{q/p} \right)^{1/q}. \tag{5.33}$$

Here $\mathcal{M}_{-1} := \Lambda_0$, $F_\lambda := g_\lambda$, and $d_\lambda(f) := b_\lambda(f)$ if $\lambda \in \mathcal{M}_{-1}$.

Finally, we define again yet another norm on $B_{pq}^\alpha(\Theta^+)$ via basis elements:

$$\|f\|_{B_{pq}^\alpha(\Theta^+)}^A := \inf_{f = \sum_{\mu \in \mathcal{M}^+} a_\mu F_\mu} \left(\sum_{j \geq -1} \left(\sum_{\mu \in \mathcal{M}_j} (|\eta_\mu|^{-\alpha/n} \|a_\mu F_\mu\|_p)^p \right)^{q/p} \right)^{1/q}, \tag{5.34}$$

where the infimum is taken over all representations $f = \sum_{\mu \in \mathcal{M}^+} a_\mu F_\mu$ in L_p with $\mathcal{M}^+ := \bigcup_{j \geq -1} \mathcal{M}_j$.

It is important that the above norms are equivalent.

Theorem 5.10 *Suppose $\alpha > 0$, $0 < p, q \leq \infty$, and (5.1) holds. Then the norms $\|\cdot\|_{B_{pq}^\alpha(\Theta^+)}$, $\|\cdot\|_{B_{pq}^\alpha(\Theta^+)}^T$, and $\|\cdot\|_{B_{pq}^\alpha(\Theta^+)}^A$ are equivalent.*

The proof of this theorem is quite similar to the proof of Theorem 5.8 and will be omitted.

Similarly to $\dot{B}_{pq}^\alpha(\Theta)$, the inhomogeneous B-space $B_{pq}^\alpha(\Theta^+)$ is a quasi-Banach space. We omit the details.

Remark 5.11 There are continuous and semi-continuous counterparts of the B-spaces $\dot{B}_{pq}^\alpha(\Theta)$ and $B_{pq}^\alpha(\Theta^+)$ defined in this section. For instance, the B-space $\dot{B}_{pq}^\alpha(\Theta)$ can be defined for any continuous ellipsoid cover Θ of \mathbb{R}^n as the set of all functions $f \in L_p^{\text{loc}}$ such that

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)} := \left(\int_0^\infty \left(t^{\alpha/d+1} \int_{\mathbb{R}^n} \omega_k(f, \theta(v, t))_p^p dv \right)^{q/p} dt/t \right)^{1/q} < \infty. \tag{5.35}$$

Here one should assume that $|\theta(v, t)|$ is a measurable function of v, t . We omit the details.

5.3 Interrelation of B-spaces with Besov and Other B-Spaces

It is quite easy to show that if Θ is a regular ellipsoid (ball) cover of \mathbb{R}^n , then the B-spaces $\dot{B}_{pq}^\alpha(\Theta)$ and $B_{pq}^\alpha(\Theta^+)$ considered in this paper are the same as the corresponding classical Besov spaces (see also [9]).

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}_{pq}^α are exactly the same (with equivalent norms) as the anisotropic Besov spaces (with weight 1) developed in [2].

The B-spaces from this paper have a lot in common with the B-spaces introduced in [10, 14]. The main distinctions are that the latter B-spaces are based on multilevel nested triangulations of \mathbb{R}^2 and local piecewise polynomial approximation, while the former are based on ellipsoid covers and local polynomial approximation. Consequently, the B-norms based on nested triangulations can be in general much smaller than the B-norms from this paper.

As indicated in Proposition 2.10, \mathbb{R}^n equipped with the distance introduced in Definition 2.6 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., [13]). 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 .

6 Nonlinear Approximation

In this section, we apply the B-spaces to nonlinear m -term approximation from the two-level-split bases introduced in Sect. 4.

6.1 The B-spaces of Nonlinear Approximation

A specific type of B-spaces will play an important role in nonlinear m -term approximation in L_p . For $0 < p < \infty$ and $\alpha > 0$, let τ be determined by

$$1/\tau = \alpha/n + 1/p, \tag{6.1}$$

which in the case of classical Besov spaces signifies the critical embedding in L_p . For nonlinear approximation in $L_\infty := C_0$ ($p = \infty$) τ is determined by $1/\tau = \alpha/n$, and we shall necessarily assume that $\alpha \geq 1$ (otherwise the embedding (6.5) below is not valid).

For a given discrete ellipsoid cover Θ of \mathbb{R}^n , consider the homogeneous B-space $\dot{B}_\tau^\alpha := \dot{B}_{\tau\tau}^\alpha(\Theta)$. Then

$$\|f\|_{\dot{B}_\tau^\alpha(\Theta)} := \left(\sum_{\theta \in \Theta} |\theta|^{-\alpha\tau/n} \omega_k(f, \theta)_\tau^\tau \right)^{1/\tau}. \tag{6.2}$$

Although in general $\tau < 1$, it can be proved (similarly as in [14]) that for any $0 < q < p$,

$$\|f\|_{\dot{B}_\tau^\alpha(\Theta)} \sim \left(\sum_{\theta \in \Theta} |\theta|^{(1/p-1/q)\tau} \omega_k(f, \theta)_q^\tau \right)^{1/\tau}, \tag{6.3}$$

which allows one to work in L_q with $q \geq 1$ if $p > 1$.

From (5.28) and (4.7) one readily derives this representation of the norm in $\dot{B}_\tau^\alpha(\Theta)$:

$$\|f\|_{\dot{B}_\tau^\alpha(\Theta)} \sim \left(\sum_{\mu \in \mathcal{M}} \|d_\mu(f)F_\mu\|_p^\tau \right)^{1/\tau}. \tag{6.4}$$

The following embedding result indicates why these spaces work in nonlinear approximation. In fact, it shows that $\dot{B}_\tau^\alpha(\Theta)$ lies on the Sobolev embedding line.

Theorem 6.1 *The B-space $\dot{B}_\tau^\alpha(\Theta)$, with p, τ interrelated by (6.1), is continuously embedded in L_p , that is, each $f \in \dot{B}_\tau^\alpha(\Theta)$ can be identified as a function $f \in L_p$ and*

$$\|f\|_p \leq c \|f\|_{\dot{B}_\tau^\alpha(\Theta)}. \tag{6.5}$$

Proof Properties (a) and (c) of discrete ellipsoid covers imply that for any $x \in \mathbb{R}^n$ and $\theta \in \Theta_j$ ($j \in \mathbb{Z}$) such that $x \in \theta$,

$$\sum_{m \leq j} \sum_{\eta \in \Theta_m: x \in \theta} (|\theta|/|\eta|)^{1/p} \leq c, \tag{6.6}$$

where c is independent of x and θ . We now invoke Theorem 3.3 from [14], which shows that (6.6) leads to

$$\left\| \sum_{\mu \in \mathcal{M}} |d_\mu(f)F_\mu(\cdot)| \right\|_p \leq c \left(\sum_{\mu \in \mathcal{M}} \|d_\mu(f)F_\mu\|_p^\tau \right)^{1/\tau}. \quad \square$$

We now consider the inhomogeneous B-space $B_\tau^\alpha := B_{\tau\tau}^\alpha(\Theta^+)$ associated with a discrete ellipsoid cover $\Theta^+ = \bigcup_{j \geq 0} \Theta_j$ of \mathbb{R}^n . The norm in B_τ^α has a representation similar to (6.4), namely,

$$\|f\|_{B_\tau^\alpha(\Theta)} \sim \left(\sum_{\mu \in \mathcal{M}^+} \|d_\mu(f)F_\mu\|_p^\tau \right)^{1/\tau}, \quad \mathcal{M}^+ := \bigcup_{j \geq -1} \mathcal{M}_j, \tag{6.7}$$

and an embedding result similar to Theorem 6.1 holds.

6.2 Jackson Estimates for Nonlinear Approximation

Consider nonlinear m -term approximation from $\mathcal{F} := \bigcup_{j \in \mathbb{Z}} \mathcal{F}_j = \{F_\mu : \mu \in \mathcal{M}\}$. We let \mathcal{E}_m denote the nonlinear set of all functions g of the form

$$g = \sum_{\mu \in \Gamma_m} a_\mu F_\mu,$$

where $\Gamma_m \subset \mathcal{M}$, $\#\Gamma \leq m$, and Γ is allowed to vary with g . We denote by $\sigma_m(f)_p$ the error of best L_p -approximation to $f \in L_p(\mathbb{R}^n)$ from \mathcal{E}_m :

$$\sigma_m(f)_p := \inf_{g \in \mathcal{E}_m} \|f - g\|_p.$$

Theorem 6.2 (Jackson estimate) *If $f \in \dot{B}_\tau^\alpha(\Theta)$, $\alpha > 0$, $0 < p \leq \infty$, then*

$$\sigma_m(f)_p \leq cm^{-\alpha/n} \|f\|_{\dot{B}_\tau^\alpha(\Theta)}, \tag{6.8}$$

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

In the case $0 < p < \infty$, estimate (6.8) follows by the general Theorem 3.4 from [14], using (6.6). The proof of (6.8) in the case $p = \infty$ is more involved and can be carried out as the proof of Theorem 3.1 in [15]. We omit it.

The nonlinear m -term approximation $\sigma_m(f)_p$ from $\mathcal{F}^+ := \bigcup_{j \geq -1} \mathcal{F}_j$, where $\mathcal{F}_{-1} := \{g_\lambda : \lambda \in \Lambda_0\}$, can be defined as above by simply replacing \mathcal{M} by \mathcal{M}^+ . Then the Jackson estimate (6.8) holds with $\dot{B}_\tau^\alpha(\Theta)$ replaced by $B_\tau^\alpha(\Theta)$. We omit the details.

The Jackson estimate (6.8) readily yields a direct estimate for nonlinear m -term approximation from \mathcal{F} (or \mathcal{F}^+) by the K -functional between L_p and $\dot{B}_\tau^\alpha(\Theta)$ (or $B_\tau^\alpha(\Theta)$). It is a challenging *open problem* to prove a companion inverse estimate due to the fact that \mathcal{F} is possibly redundant and non-nested.

7 Measuring Smoothness via Anisotropic B-spaces

It is an ancient question in analysis: What is the smoothness of a given function and how should the smoothness be measured? In this section we show how the anisotropic B-spaces $\dot{B}_\tau^\alpha(\Theta)$ developed in the previous sections can be employed for handling this question and how it is related to nonlinear m -term approximation from the two-level-split bases developed here.

We consider 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 $Q \subset \mathbb{R}^2$. Both examples could be generalized at the expense of additional technical efforts. But in the present simple form they already suffice to reveal some interesting and perhaps surprising facts. Indeed, we shall show that each of these functions has higher-order smoothness α in $\dot{B}_\tau^\alpha(\Theta)$ for 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 $\alpha > 0$, τ is defined by $1/\tau = \alpha/2 + 1/p$.

7.1 The Anisotropic B-space Smoothness of $\mathbb{1}_{B(0,1)}$

Theorem 7.1 *There exists an anisotropic ellipsoid cover Θ of \mathbb{R}^2 such that $\mathbb{1}_{B(0,1)} \in \dot{B}_\tau^\alpha(\Theta)$ for any $\alpha < 4/p$. In comparison, in the scale of Besov spaces $\dot{B}_{\tau\tau}^\alpha$ one has $\mathbb{1}_{B(0,1)} \in \dot{B}_{\tau\tau}^\alpha$ for $\alpha < 2/p$. Here the bounds for α are sharp.*

Proof We begin by constructing an appropriate continuous ellipse cover Θ_{con} of \mathbb{R}^2 . For arbitrary $t \leq 0$ and $v \in \mathbb{R}^2$, we define the ellipse $\theta(v, t)$ by

$$\theta(v, t) := 2^{-t/2} B(0, 1) + v,$$

that is the disk of radius $2^{-t/2}$ centered at v .

Let $t > 0$. For any $v = (v_1, 0)$ on the positive x_1 -axis which obeys the condition $|1 - v_1| \leq 2^{-t/3}$, we define $\theta(v, t)$ as the set of all point $x \in \mathbb{R}^2$ such that

$$\frac{(x_1 - v_1)^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} \leq 1,$$

where $\sigma_1 := (|1 - v_1| + 2^{-t/2})2^{-t/6}$ and $\sigma_2 := (|1 - v_1| + 2^{-t/2})^{-1}2^{-5t/6}$.

If $v = (v_1, 0)$, $v_1 \geq 0$, satisfies $|1 - v_1| > 2^{-t/3}$, we set $\theta(v, t) := 2^{-t/2}B(0, 1) + v$.

For any point v which does not lie on the positive x_1 -axis we define $\theta(v, t)$ by a rotation of the ellipse $\theta(|v|, t)$ defined above about the origin which takes $(|v|, 0)$ to v .

We now show that the collection of ellipses \mathcal{E}_{con} defined above is a continuous ellipse cover of \mathbb{R}^2 in the sense of Definition 2.4.

Fix $v = (v_1, 0)$, $v_1 > 0$. Let $t, s > 0$, and assume that $|1 - v_1| \leq 2^{-(t+s)/3}$. (All other cases are similar or trivial.) Denote by $\sigma_1(t)$ the x_1 -semi-axis of $\theta(v, t)$. By the definition we have

$$\frac{\sigma_1(t+s)}{\sigma_1(t)} = \frac{|1 - v_1| + 2^{-(t+s)/2}}{|1 - v_1| + 2^{-t/2}} \cdot 2^{-s/6},$$

which leads to

$$2^{-2s/3} \leq \sigma_1(t+s)/\sigma_1(t) \leq 2^{-s/6}. \tag{7.1}$$

The case when v does not lie on the positive x_1 -axis reduces to the above by rotation.

Now fix $t > 0$, and let $\theta := \theta(v, t)$ and $\theta' := \theta(v', t)$ be such that $\theta \cap \theta' \neq \emptyset$. Assume that $|1 - |v|| \leq 2^{-t/3}$ and $|1 - |v'|| \leq 2^{-t/3}$; the other cases are similar or trivial. Since \mathcal{E}_{con} is rotation invariant, we may assume that $v = (v_1, 0)$, $v_1 > 0$. Denote by σ_1, σ_2 the semi-axes of θ and by σ'_1, σ'_2 ($\sigma'_1 < \sigma'_2$) the semi-axes of θ' . A simple rotation argument gives

$$||v'| - |v|| \leq \sigma'_1 + \sigma_1 = (|1 - |v'|| + |1 - |v|| + 2 \cdot 2^{-t/2})2^{-t/6} \leq 4 \cdot 2^{-t/2},$$

which implies

$$\sigma'_1 \leq (|1 - |v|| + ||v'| - |v|| + 2^{-t/2})2^{-t/6} \leq (|1 - |v|| + 5 \cdot 2^{-t/2})2^{-t/6} \leq 5\sigma_1.$$

Therefore,

$$1/5 \leq \sigma'_1/\sigma_1 \leq 5, \quad \text{and hence} \quad 1/5 \leq \sigma'_2/\sigma_2 \leq 5. \tag{7.2}$$

We may assume that $t \geq 3$ (the case $t < 3$ is trivial). Then $1/2 \leq |v|, |v'| \leq 3/2$. Since $\theta \cap \theta' \neq \emptyset$, the ellipse θ' can be obtained by revolving $\theta(|v'|, t)$ about the origin at an angle γ such that

$$|\gamma| \leq 2\sigma_2 + 2\sigma'_2 \leq 4 \cdot 2^{-t/3}. \tag{7.3}$$

As in Sect. 2, let A_θ be the affine transform which maps $B(0, 1)$ onto θ . Then obviously $A_\theta(x) = M_\theta x + v$, where $M_\theta = \text{diag}(\sigma_1, \sigma_2)$. The affine transform $A_{\theta'}$ mapping $B(0, 1)$ onto θ' is of the form $A_{\theta'}(x) = M_{\theta'}x + v'$, where $M_{\theta'}$ can be represented

as the product of a diagonal and a rotation matrices, namely,

$$M_{\theta'} = A_\gamma \operatorname{diag}(\sigma'_1, \sigma'_2), \quad A_\gamma := \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix}.$$

It is straightforward to show that

$$M_\theta^{-1} M_{\theta'} = \begin{pmatrix} (\sigma'_1/\sigma_1) \cos \gamma & -(\sigma'_2/\sigma_1) \sin \gamma \\ (\sigma'_1/\sigma_2) \sin \gamma & (\sigma'_2/\sigma_2) \cos \gamma \end{pmatrix}.$$

From (7.3) it follows that

$$(\sigma'_2/\sigma_1)|\sin \gamma| \leq 4 \quad \text{and} \quad (\sigma'_1/\sigma_2)|\sin \gamma| \leq 4.$$

This and (7.2) imply that all entries of $M_\theta^{-1} M_{\theta'}$ are bounded in absolute value by 5. Hence

$$\|M_\theta^{-1} M_{\theta'}\|_{\ell_2 \rightarrow \ell_2} \leq c \quad \text{and similarly} \quad \|M_{\theta'}^{-1} M_\theta\|_{\ell_2 \rightarrow \ell_2} \leq c.$$

From this and (7.1) it follows that Θ_{con} is a continuous ellipse cover of \mathbb{R}^2 .

In applying Lemma 2.9 to Θ_{con} , we conclude that there is a discrete ellipse cover Θ of \mathbb{R}^2 such that each ellipse $\theta \in \Theta$ is a dilated version of an ellipse from Θ_{con} by a factor ρ_θ obeying $(a_7 + 1)/2 \leq \rho_\theta \leq 1$. Also, for any $\theta \in \Theta_m$ we have $a_1 2^{-m} \leq |\theta| \leq a_2 2^{-m}$, where a_1, a_2 are positive constants.

It remains to show that $\mathbb{1}_{B(0,1)} \in \dot{B}_\tau^\alpha(\theta)$ for all $\alpha < 4/p$. Denote by Θ'_m the set of all ellipses from Θ which intersect the unit circle S^1 in \mathbb{R}^2 . We need to estimate $\#\Theta'_m$. By condition (c) on discrete ellipsoid covers, only N_1 ellipses from Θ_m may intersect at a time. This and the construction of Θ_{con} and Θ yield

$$\#\Theta'_m \leq c 2^m \left| \bigcup_{\theta \in \Theta'_m} \theta \right| \leq c 2^m 2^{-2m/3} = c 2^{m/3}, \quad m > 0.$$

Evidently, $\#\Theta'_m \leq c$ if $m \leq 0$.

We use these to estimate $\|f\|_{\dot{B}_\tau^\alpha(\Theta)}$, where $f := \mathbb{1}_{B(0,1)}$. Evidently, $\omega_k(f, \theta)_\tau = 0$ if $\theta \in \Theta_m \setminus \Theta'_m$, and hence

$$\begin{aligned} \|f\|_{\dot{B}_\tau^\alpha(\Theta)}^\tau &= \sum_{m \in \mathbb{Z}} \sum_{\theta \in \Theta'_m} |\theta|^{-\alpha\tau/2} \omega_k(f, \theta)_\tau^\tau \\ &\leq c \sum_{m=-\infty}^0 2^{m\alpha\tau/2} + c \sum_{m=1}^\infty (\#\Theta'_m) 2^{-m(1-\alpha\tau/2)} \\ &\leq c + c \sum_{m=1}^\infty 2^{-m(2/3-\alpha\tau/2)} \leq c < \infty, \end{aligned}$$

where we used that $2/3 - \alpha\tau/3 = (2\tau/3)(1/p - \alpha/4) > 0$, since $\alpha < 4/p$. Consequently, $f \in \dot{B}_\tau^\alpha(\Theta)$. Here the bound $4/p$ for α is sharp, since S^1 cannot be covered by $\leq c 2^{m/3}$ ellipsoids of area 2^{-m} whenever $c > 0$ is sufficiently small.

It is well known that $\mathbb{1}_{B(0,1)}$ belongs to the Besov space $B_{\tau\tau}^\alpha$ for $\alpha < 2/p$, and this bound for α is sharp. □

As in Sect. 6.2, denote by $\sigma_m(f)_p$ the best m -term approximation of f from \mathcal{F} in L_p . The following estimate is immediate from Theorem 7.1 and Theorem 6.2:

Corollary 7.2 *Let $0 < p < \infty$. Then*

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

At the same time, if $\sigma_m^W(f)_p$ denotes the best m -term approximation of f in L_p ($p \geq 1$) from any reasonable wavelet basis, then

$$\sigma_m^W(\mathbb{1}_{B(0,1)})_p \leq cm^{-\gamma} \quad \text{for all } \gamma < 1/p.$$

Here both estimate are sharp.

7.2 The Anisotropic B-space Smoothness of $\mathbb{1}_Q$

We assume the notation from Sect. 7.1.

Theorem 7.3 *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 $\mathbb{1}_Q \in \dot{B}_\tau^\alpha$, while in the scale of Besov spaces $\dot{B}_{\tau\tau}^\alpha$ one has only $\mathbb{1}_Q \in \dot{B}_{\tau\tau}^\alpha$ for $\alpha < 2/p$, and this bound for α is sharp.*

Proof Without loss of generality, we may assume that $Q = [-1, 1] \times [0, 2]$. It is convenient to first construct an appropriate continuous ellipse cover and then get the desired discrete ellipse cover upon employing Lemma 2.9.

Our main step is to effectively define the ellipses $\theta(v, t)$ of our continuous cover Θ_{con} with centers v from the triangle $\Delta_0 := [(0, 0), (1, 0), (0, 1)]$ and $t > 0$. Then we use symmetry about the x_2 -axis to define $\theta(v, t)$ for v in the triangle

$$[(-1, 0), (0, 0), (0, 1)].$$

We next apply symmetry about the x_1 -axis to define the ellipses $\theta(v, t)$ for v in the triangle $[(-1, 0), (0, -1), (1, 0)]$. Again, by symmetry about the line $x_2 = -x_1 + 1$, we define $\theta(v, t)$ on the square $[(1, 0), (2, 1), (1, 2), (0, 1)]$. Symmetry about the line $x_2 = x_1 + 1$ enables us to define $\theta(v, t)$ for v in the rectangle

$$[(-1, 0), (1, 2), (0, 3), (-2, 1)].$$

In this way the ellipses $\theta(v, t)$ would be defined with centers v from the square

$$S := [(0, -1), (2, 1), (0, 3), (-2, 1)].$$

Finally, we define the ellipses $\theta(v, t)$ with centers $v \in \mathbb{R}^2 \setminus S$ by

$$\theta(v, t) := 2^{-t/2}B(0, 1) + v.$$

In going further, for $t \leq 0$ we define $\theta(v, t)$ for all centers $v \in \mathbb{R}^2$ by $\theta(v, t) := 2^{-t/2}B(0, 1) + v$.

It remains to define the ellipses $\theta(v, t)$ with centers $v \in \Delta_0$ and $t > 0$. We begin by introducing a parameter δ determined by the condition

$$0 < \delta < \min \left\{ \frac{1}{2}, \frac{4}{2 + p\alpha} \right\}. \tag{7.4}$$

For every $v \in \Delta_0$, $v = (v_1, v_2)$, set $u := u_v := 1 - v_1 - v_2$ and define $\theta(v, t)$ as the set of all $x \in \mathbb{R}^2$ such that

$$\frac{(x_1 - v_1)^2}{\sigma_1^2} + \frac{(x_2 - v_2)^2}{\sigma_2^2} \leq 1,$$

where $\sigma_1 := (u_v 2^{t/2} + 1)^{1-\delta} 2^{-t/2}$ and $\sigma_2 := (u_v 2^{t/2} + 1)^{\delta-1} 2^{-t/2}$. Evidently, $|\theta(v, t)| = \pi \sigma_1 \sigma_2 = \pi 2^{-t}$. Notice also that $\theta(v, t) := 2^{-t/2}B(0, 1) + v$ if $u_v = 0$, which is the same as $v \in [(1, 0), (0, 1)]$.

We next show that the collection of ellipses Θ_{con} defined above is a continuous ellipse cover of \mathbb{R}^2 with parameters depending only on δ .

Fix $v \in \Delta_0$, and for any $t \in \mathbb{R}$ denote by $\sigma_1(t)$ and $\sigma_2(t)$ the semi-axes of $\theta(v, t)$. Then from above we have for $t, s > 0$

$$\frac{\sigma_1(t+s)}{\sigma_1(t)} = \left(\frac{u_v + 2^{-(t+s)/2}}{u_v + 2^{-t/2}} \right)^{1-\delta} 2^{-\delta s/2},$$

which readily implies

$$2^{-s/2} \leq \sigma_1(t+s)/\sigma_1(t) \leq 2^{-\delta s/2}. \tag{7.5}$$

Now, fix $t > 0$ and let $\theta(v, t) \cap \theta(v', t) \neq \emptyset$, $v, v' \in \Delta_0$. Assume $u_{v'} > u_v$. Clearly,

$$\begin{aligned} u_{v'} &\leq u_v + \sigma_1(v) + \sigma_1(v') \\ &\leq u_v + (u_v + 2^{-t/2})^{1-\delta} 2^{-\delta t/2} + (u_{v'} + 2^{-t/2})^{1-\delta} 2^{-\delta t/2} \\ &\leq u_v + 2u_v^{1-\delta} 2^{-\delta t/2} + 2 \cdot 2^{-t/2} + 2u_{v'}^{1-\delta} 2^{-\delta t/2} + 2 \cdot 2^{-t/2}. \end{aligned}$$

If $u_{v'} \geq 4^{1/\delta} 2^{-t/2}$, this leads to

$$u_{v'} \leq 8(u_v + u_v^{1-\delta} 2^{-\delta t/2} + 2^{-t/2}) \leq 16(u_v + 2^{-t/2}),$$

which yields

$$\sigma_1(v')/\sigma_1(v) \leq c, \quad c = c(\delta). \tag{7.6}$$

If $u_{v'} < 4^{1/\delta} 2^{-t/2}$, the same estimate follows immediately with a different constant c .

Estimates (7.5)–(7.6) readily imply that Θ_{con} is a continuous ellipse cover of \mathbb{R}^2 in the sense of Definition 2.4.

By Lemma 2.9, the above cover Θ_{con} induces a discrete ellipse cover, say, Θ of \mathbb{R}^2 , where every $\theta \in \Theta$ is obtained by dilating an ellipse from Θ_{con} (with the same

center) by a factor ρ_θ obeying $(a_1 + 1)/2 \leq \rho_\theta \leq 1$. Also, for any $\theta \in \Theta_m$ we have $a_1 2^{-m} \leq |\theta| \leq a_2 2^{-m}$ for some positive constants a_1, a_2 .

Our next task is to show that the function $\mathbb{1}_Q$, where $Q := [-1, 1] \times [0, 2]$, belongs to the B-space $\dot{B}_\tau^\alpha(\Theta)$. To this end we need an upper bound for the number of all ellipses from Θ_m which intersect the boundary of $[-1, 1] \times [0, 2]$. Denote this set by Θ'_m . By condition (c) on discrete covers and the construction of Θ_{con} and Θ , it follows that for $m > 0$

$$\begin{aligned} \#\Theta'_m &\leq c2^m \left| \bigcup_{\theta \in \Theta'_m} \theta \right| \leq c2^m \int_0^1 \sigma_2(\theta(v_1, 0), 2^{-m}) dv_1 \\ &\leq c2^{m/2} \int_0^1 ((1 - v_1)2^{m/2} + 1)^{\delta-1} dv_1 = c2^{\delta m/2}. \end{aligned}$$

Evidently $\#\Theta_m \leq c$ if $m \leq 0$.

We are now prepared to estimate $\|f\|_{\dot{B}_\tau^\alpha(\Theta)}$, where $f := \mathbb{1}_Q$. Evidently, $\omega_k(f, \theta)_\tau = 0$ if $\theta \in \Theta_m \setminus \Theta'_m$. From this and the above estimates of $\#\Theta'_m$, we get

$$\begin{aligned} \|f\|_{\dot{B}_\tau^\alpha(\Theta)}^\tau &= \sum_{m \in \mathbb{Z}} \sum_{\theta \in \Theta'_m} |\theta|^{-\alpha\tau/2} \omega_k(f, \theta)_\tau^\tau \\ &\leq c \sum_{m=-\infty}^0 2^{m\alpha\tau/2} + c \sum_{m=1}^\infty (\#\Theta'_m) 2^{-m(1-\alpha\tau/2)} \\ &\leq c + c \sum_{m=1}^\infty 2^{-m(1-\alpha\tau/2-\delta/2)} \leq c < \infty. \end{aligned}$$

Here we used that $1 - \alpha\tau/2 - \delta/2 = \tau(1/\tau - \alpha/2 - \delta/2\tau) = \tau(1/p - \delta/2\tau) > 0$ by (7.4). Thus, $f \in \dot{B}_\tau^\alpha(\Theta)$.

It is standard and well known that $\mathbb{1}_Q$ belongs to the Besov space $B_{\tau\tau}^\alpha$ if $\alpha < 2/p$ and this bound for α is sharp. □

As above, Theorem 7.1 and Theorem 6.2 lead to the following approximation result:

Corollary 7.4 *For any $\alpha > 0$ there exists a discrete ellipse cover of \mathbb{R}^2 such that for all $0 < p < \infty$*

$$\sigma_m(\mathbb{1}_Q)_p \leq cm^{-\alpha}.$$

In comparison, for wavelet m -term approximation of $\mathbb{1}_Q$ in L_p ($p \geq 1$) one has only

$$\sigma_m^W(\mathbb{1}_Q)_p \leq cm^{-\gamma} \quad \text{for all } \gamma < 1/p,$$

and this estimate is sharp.

7.3 Discussion

One of the basic goals in this paper is to advance regularity notions that better respect anisotropic features of functions. The key idea is to measure the smoothness of functions *adaptively* by means of anisotropic B-spaces, and in particular by the B-spaces of nonlinear approximation $\dot{B}_\tau^\alpha(\Theta)$, by allowing the underlying ellipsoid covers to adapt to the given function. We have considered two model problems which clearly shed some light on this idea. First, as was shown in Sects. 7.1–7.2 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. It is rather obvious that by using adaptive dilations the anisotropic B-spaces are better able to resolve the singularities along smooth or piecewise smooth curves. As a result, the two-level-split decompositions of these functions are sparser than their wavelet decompositions, which leads to better rates of nonlinear m -term approximation. Second, it might be surprising that characteristic functions of polygonal domains have infinite smoothness while those of domains with smooth boundaries have limited regularity. Roughly speaking, the positive curvature in the latter case limits the extent to which ellipses are allowed to stretch. On the other hand, the covers that would 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. Thus, the idea of describing smoothness by maximizing α over covers Θ has to be treated with care. This maximum could be determined by a saturated smoothness as in the first example. Or such a maximum would be achieved only when constraining the parameter sets of the covers to some a-priori chosen compact set which is not an intrinsic property of the underlying function. Finally, 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 $B_\tau^\alpha(\Theta)$ in the above sense.

Appendix

Proof of Theorem 5.8 Evidently, $\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^A \leq \|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^T$, and we have already shown that $\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^T \leq c\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}$.

It remains to prove that

$$\|f\|_{\dot{B}_{pq}^\alpha(\Theta)} \leq c\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}^A. \tag{A.1}$$

We only consider the least favorable case when $1 < p < q < \infty$. Let $f = \sum_{\mu \in \mathcal{M}} a_\mu F_\mu$. For any ellipsoid $\sigma \in \Theta_j$ we have

$$\omega_k(f, \sigma)_p \leq \omega_k\left(\sum_{\mu: \ell(\eta_\mu) < j} a_\mu F_\mu, \sigma\right)_p + c\left\|\sum_{\mu: \ell(\eta_\mu) \geq j} a_\mu F_\mu\right\|_p$$

$$\begin{aligned} &\leq c \sum_{v=1}^{\infty} \left(\sum_{\mu: \eta_{\mu} \in \Theta_{j-v}} |a_{\mu}|^p \omega_k(F_{\mu}, \sigma)_p^p \right)^{1/p} \\ &\quad + c \sum_{v=0}^{\infty} \left(\sum_{\mu: \eta_{\mu} \in \Theta_{j+v}} \|a_{\mu} F_{\mu}\|_p^p \right)^{1/p}. \end{aligned} \tag{A.2}$$

Here $\ell(\eta)$ stands for the level of η in Θ .

We next estimate $\omega_k(F_{\mu}, \sigma)_p$ for any $\mu \in \mathcal{M}$ such that $\eta_{\mu} \in \Theta_{j-v}$, $v \geq 1$. Let $\eta := \eta_{\mu}$, $\theta := \theta_{\mu}$, $\beta := \beta_{\mu}$, and $F := F_{\mu}$. Also let $\sigma^* := A_{\eta}^{-1}\sigma = A_{\eta}^{-1}A_{\sigma}B^*$, $\theta^* := A_{\eta}^{-1}\theta$, and $F^* := F \circ A_{\eta}$. Recall that $F = P_{\eta, \beta} \varphi_{\eta} \varphi_{\theta}$ with $P_{\eta, \beta} := |\eta|^{-1/2} P_{\beta} \circ A_{\eta}^{-1}$. Hence $F^* = |\eta|^{-1/2} P_{\beta}(\varphi_{\eta} \circ A_{\eta})(\varphi_{\theta} \circ A_{\eta})$.

For any $h \in \mathbb{R}^n$ we have with $h^* := A_{\eta}^{-1}h$

$$\begin{aligned} \|\Delta_h^k F\|_{L_p(\sigma)}^p &= |\det M_{\eta}| \|\Delta_{h^*}^k F^*\|_{L_p(\sigma^*)}^p \leq c |\eta| |h^*|^{kp} \left\| \left(\frac{\partial}{\partial h} \right)^k F^* \right\|_{\infty}^p |\sigma^*| \\ &\leq c |\eta| |\sigma^*| (\text{diam } \sigma^*)^{kp} \|F^*\|_{\infty}^p. \end{aligned} \tag{A.3}$$

Here we assumed that $k|h^*| \leq \text{diam } \sigma^*$, since otherwise $\Delta_{h^*}^k F^*(x) \equiv 0$. We also used that $\left\| \left(\frac{\partial}{\partial h} \right)^k F^* \right\|_{\infty} \leq c$ with c a constant depending only on $\mathbf{p}(\Theta)$, k , and L .

Note that $\|F^*\|_{\infty} = \|F\|_{\infty}$, and hence $\|F\|_p^p \sim |\eta| \|F^*\|_{\infty}^p$. Also, we have

$$|\sigma^*| = |\det(M_{\eta}^{-1} M_{\sigma})| = |\det M_{\eta}^{-1}| |\det M_{\sigma}| = |\eta|^{-1} |\sigma|,$$

and $\text{diam } \sigma^* = 2 \|M_{\eta}^{-1} M_{\sigma}\|_{\ell_2 \rightarrow \ell_2} \leq 2a_5 2^{-a_6 v}$ using property (b) of Θ . We use the above observations in (A.3) to obtain

$$\omega_k(F_{\mu}, \sigma)_p^p \leq c |\sigma| |\eta_{\mu}|^{-1} 2^{-a_6 v k p} \|F_{\mu}\|_p^p. \tag{A.4}$$

Combining this with (A.2), we get

$$\begin{aligned} &|\sigma|^{-\alpha p/n} \omega_k(f, \sigma)_p^p \\ &\leq c \left[\sum_{v=1}^{\infty} \left(\sum_{\mu: \eta_{\mu} \in \Theta_{j-v}, \eta_{\mu} \cap \sigma \neq \emptyset} \left(\frac{|\eta_{\mu}|}{|\sigma|} \right)^{\alpha p/n-1} 2^{-a_6 v k p} |\eta_{\mu}|^{-\alpha p/n} \|a_{\mu} F_{\mu}\|_p^p \right)^{1/p} \right]^p \\ &\quad + c \left[\sum_{v=0}^{\infty} \left(\sum_{\mu: \eta_{\mu} \in \Theta_{j+v}, \eta_{\mu} \cap \sigma \neq \emptyset} \left(\frac{|\eta_{\mu}|}{|\sigma|} \right)^{\alpha p/n} |\eta_{\mu}|^{-\alpha p/n} \|a_{\mu} F_{\mu}\|_p^p \right)^{1/p} \right]^p \\ &\leq c \left[\sum_{v=1}^{\infty} \left(\sum_{\mu: \eta_{\mu} \in \Theta_{j-v}, \eta_{\mu} \cap \sigma \neq \emptyset} 2^{-v p (-a_0 \alpha/n + a_6 k) - a_0 v} U_{\mu} \right)^{1/p} \right]^p \\ &\quad + c \left[\sum_{v=0}^{\infty} \left(\sum_{\mu: \eta_{\mu} \in \Theta_{j+v}, \eta_{\mu} \cap \sigma \neq \emptyset} 2^{-a_0 v \alpha p/n} U_{\mu} \right)^{1/p} \right]^p, \end{aligned}$$

where $U_{\mu} := |\eta_{\mu}|^{-\alpha p/n} \|a_{\mu} F_{\mu}\|_p^p$.

Let $\varepsilon := -a_0\alpha/n + a_6k$ and $\delta := a_0\alpha/n > 0$. Notice that $\varepsilon > 0$, since $k > \frac{a_0}{a_6} \cdot \frac{\alpha}{n}$ by (5.1). We use the above in the definition of $\|f\|_{\dot{B}_{pq}^\alpha(\Theta)}$ to obtain

$$\begin{aligned} (\|f\|_{\dot{B}_{pq}^\alpha})^q &\leq c \sum_{j \in \mathbb{Z}} \left\{ \sum_{\sigma \in \Theta_j} \left[\sum_{v=1}^\infty 2^{-\varepsilon v - a_0 v/p} \left(\sum_{\mu: \eta_\mu \in \Theta_{j-v}, \eta_\mu \cap \sigma \neq \emptyset} U_\mu \right)^{1/p} \right]^p \right\}^{q/p} \\ &\quad + c \sum_{j \in \mathbb{Z}} \left\{ \sum_{\sigma \in \Theta_j} \left[\sum_{v=0}^\infty 2^{-\delta v} \left(\sum_{\mu: \eta_\mu \in \Theta_{j+v}, \eta_\mu \cap \sigma \neq \emptyset} U_\mu \right)^{1/p} \right]^p \right\}^{q/p} \\ &=: \Sigma_1 + \Sigma_2. \end{aligned}$$

To estimate Σ_1 we apply Hölder’s inequality and get

$$\begin{aligned} \Sigma_1 &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{\sigma \in \Theta_j} \left(\sum_{v=1}^\infty 2^{-\varepsilon v p/2 - a_0 v} \sum_{\mu: \eta_\mu \in \Theta_{j-v}, \eta_\mu \cap \sigma \neq \emptyset} U_\mu \right) \left(\sum_{v=1}^\infty 2^{-\varepsilon v p'/2} \right)^{p/p'} \right]^{q/p} \\ &\leq c \sum_{j \in \mathbb{Z}} \left(\sum_{\sigma \in \Theta_j} \sum_{v=1}^\infty 2^{-\varepsilon v p/2 - a_0 v} \sum_{\mu: \eta_\mu \in \Theta_{j-v}, \eta_\mu \cap \sigma \neq \emptyset} U_\mu \right)^{q/p} \quad (1/p + 1/p' = 1). \end{aligned}$$

In going further, we switch the order of summation and then apply again Hölder’s inequality to obtain

$$\begin{aligned} \Sigma_1 &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{v=1}^\infty 2^{-\varepsilon v p/2 - a_0 v} \left(\sum_{\mu: \eta_\mu \in \Theta_{j-v}} U_\mu \right) \#\{\sigma \in \Theta_j : \sigma \cap \eta_\mu \neq \emptyset\} \right]^{q/p} \\ &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{v=1}^\infty 2^{-\varepsilon v p/2} \sum_{\mu: \eta_\mu \in \Theta_{j-v}} U_\mu \right]^{q/p} \\ &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{v=1}^\infty 2^{-\varepsilon v q/4} \left(\sum_{\mu: \eta_\mu \in \Theta_{j-v}} U_\mu \right)^{q/p} \right] \left[\sum_{v=1}^\infty 2^{-\varepsilon v p r'/4} \right]^{q/pr'} \quad (r := q/p > 1), \end{aligned}$$

where we used that $\#\{\sigma \in \Theta_j : \sigma \cap \eta_\mu \neq \emptyset\} \leq c2^{a_0 v}$, which is a consequence of properties (a) and (c) of Θ .

Finally, we apply the substitution $l := j - v$ and switch the order of summation above to obtain

$$\begin{aligned} \Sigma_1 &\leq c \sum_{j \in \mathbb{Z}} \sum_{l=-\infty}^{j-1} 2^{-\varepsilon(j-l)q/4} \left(\sum_{\mu: \eta_\mu \in \Theta_l} U_\mu \right)^{q/p} \\ &\leq c \sum_{l \in \mathbb{Z}} \left(\sum_{\mu: \eta_\mu \in \Theta_l} U_\mu \right)^{q/p} \sum_{j=l+1}^\infty 2^{-\varepsilon(j-l)q/4} \leq c \sum_{l \in \mathbb{Z}} \left(\sum_{\mu: \eta_\mu \in \Theta_l} U_\mu \right)^{q/p} \\ &\leq (\|f\|_{\dot{B}_{pq}^\alpha}^A)^q. \end{aligned}$$

We estimate Σ_2 in a similar fashion. Applying Hölder’s inequality we get

$$\begin{aligned} \Sigma_2 &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{\sigma \in \Theta_j} \left(\sum_{v=0}^{\infty} 2^{-\delta v p/2} \sum_{\mu: \eta_\mu \in \Theta_{j+v}, \eta_\mu \cap \sigma \neq \emptyset} U_\mu \right) \left(\sum_{v=1}^{\infty} 2^{-\delta v p'/2} \right)^{p/p'} \right]^{q/p} \\ &\leq c \sum_{j \in \mathbb{Z}} \left(\sum_{\sigma \in \Theta_j} \sum_{v=0}^{\infty} 2^{-\delta v p/2} \sum_{\mu: \eta_\mu \in \Theta_{j+v}, \eta_\mu \cap \sigma \neq \emptyset} U_\mu \right)^{q/p}. \end{aligned}$$

Switching the order of summation and then applying Hölder’s inequality, we obtain

$$\begin{aligned} \Sigma_2 &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{v=0}^{\infty} 2^{-\delta v p/2} \sum_{\mu: \eta_\mu \in \Theta_{j+v}} U_\mu \right]^{q/p} \quad (r := q/p > 1) \\ &\leq c \sum_{j \in \mathbb{Z}} \left[\sum_{v=0}^{\infty} 2^{-\delta v q/4} \left(\sum_{\mu: \eta_\mu \in \Theta_{j+v}} U_\mu \right)^{q/p} \right] \left[\sum_{v=0}^{\infty} 2^{-\delta v p r'/4} \right]^{q/pr'} \\ &\leq c \sum_{j \in \mathbb{Z}} \sum_{l=j}^{\infty} 2^{-\delta(l-j)q/4} \left(\sum_{\mu: \eta_\mu \in \Theta_l} U_\mu \right)^{q/p} \\ &\leq c \sum_{l \in \mathbb{Z}} \left(\sum_{\mu: \eta_\mu \in \Theta_l} U_\mu \right)^{q/p} \sum_{j=-\infty}^l 2^{-\delta(l-j)q/4} \leq c \sum_{l \in \mathbb{Z}} \left(\sum_{\mu: \eta_\mu \in \Theta_l} U_\mu \right)^{q/p} \\ &\leq (\|f\|_{\dot{B}_{pq}^\alpha}^A)^q. \end{aligned}$$

This completes the proof of (A.1). □

References

1. Bownik, M.: Anisotropic Hardy spaces and wavelets. *Mem. Am. Math. Soc.* **164**(781) (2003)
2. Bownik, M.: Atomic and molecular decompositions of anisotropic Besov spaces. *Math. Z.* **250**, 539–571 (2005)
3. Bownik, M., Ho, K.-P.: Atomic and molecular decompositions of anisotropic Triebel–Lizorkin spaces. *Trans. Am. Math. Soc.* **358**, 1469–1510 (2006)
4. Calderón, A., Torchinsky, A.: Parabolic maximal functions associated with a distribution. *Adv. Math.* **16**, 1–64 (1975)
5. Calderón, A., Torchinsky, A.: Parabolic maximal functions associated with a distribution II. *Adv. Math.* **24**, 101–171 (1977)
6. Coifman, R., Weiss, G.: *Analyse harmonique non-comutative sur certains espaces homogenes*. Lecture Notes in Math., vol. 242. Springer, Berlin (1971)
7. Coifman, R., Weiss, G.: Extensions of Hardy spaces and their use in analysis. *Bull. Am. Math. Soc.* **83**, 569–645 (1977)
8. Dahmen, W., Petrushev, P.: “Push-the-Error” algorithm for nonlinear n-term approximation. *Constr. Approx.* **23**, 261–304 (2006)
9. Dahmen, W., Dekel, S., Petrushev, P.: Multilevel preconditioning for partition of unity methods—some analytic concepts. *Numer. Math.* **107**, 503–532 (2007)

10. Davydov, O., Petrushev, P.: Nonlinear approximation from differentiable piecewise polynomials. *SIAM J. Math. Anal.* **35**, 708–758 (2003)
11. Dekel, S., Leviatan, D., Sharir, M.: On bivariate smoothness spaces associated with nonlinear approximation. *Constr. Approx.* **20**, 625–646 (2004)
12. Folland, G., Stein, E.: *Hardy Spaces on Homogeneous Groups*. Princeton University Press, Princeton (1982)
13. Han, Y., Sawyer, E.: Littlewood–Paley theory on spaces of homogeneous type and classical function spaces. *Mem. Am. Math. Soc.* **530** (1994)
14. Karaivanov, B., Petrushev, P.: Nonlinear piecewise polynomial approximation beyond Besov spaces. *Appl. Comput. Harmon. Anal.* **15**, 177–223 (2003)
15. Karaivanov, B., Petrushev, P., Sharpley, R.C.: Algorithms for nonlinear piecewise polynomial approximation. *Trans. Am. Math. Soc.* **355**, 2585–2631 (2003)
16. Peetre, J.: *New Thoughts on Besov Spaces*. Duke Univ. Math. Series. Duke Univ., Durham (1976)
17. Petrushev, P.: Anisotropic spaces and nonlinear n -term spline approximation. In: *Approximation Theory XI: Gatlinburg 2004*. *Mod. Methods Math.*, pp. 363–394. Nashboro Press, Brentwood (2005)
18. Triebel, H.: *Theory of Function Spaces*. *Monographs in Math.*, vol. 78. Birkhäuser, Basel (1983)
19. Triebel, H.: *Theory of Function Spaces II*. *Monographs in Math.*, vol. 84. Birkhäuser, Basel (1992)