Wavelet characterizations for anisotropic Besov spaces with $0 < p < 1$

Gustavo Garrigós* Reinhard Hochmuth Anita Tabacco

August 10, 2003

Abstract

We present a wavelet characterization of anisotropic Besov spaces $B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)$, valid for the whole range $0 < p, q < \infty$, and in terms of MRA's with dilation adapted to the anisotropy of the space. Our proofs combine classical techniques based on Bernstein and Jackson-type inequalities, and non-linear methods for the cases $p < 1$. Among the consequences of our results, we characterize $B^{\boldsymbol{\alpha}}_{p,q}$ as a linear approximation space, and derive embeddings and interpolation formulas for $B^{\boldsymbol{\alpha}}_{p,q}$, which seem new in the literature when $p < 1$.

1 Introduction

The characterization of Besov norms as weighted sums of wavelet coefficients has important applications in data compression, non-linear approximation and the numerical resolution of elliptic PDE's (see, e.g., the survey papers [6, 7, 4]). In recent years, an increasing interest in non isotropic models and semi-elliptic equations has turned the attention to the more general class of *anisotropic Besov spaces* $B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n_+$ and $0 < p, q < \infty$ [15, 16, 19]. Wavelet-based techniques for such problems depend on a corresponding characterization of $B^{\boldsymbol{\alpha}}_{p,q}$ in terms of compactly supported wavelets, which only recently has been introduced by the authors in two different and independent frameworks† [14, 13].

There were however some limitations in these works that we try to overcome here. The results in [14, 15], for instance, provide characterizations of $B^{\boldsymbol{\alpha}}_{p,q}$ with no

2000 Math Subject Classification: 42B35, 42C40, 41A17.

Keywords: Jackson and Bernstein inequalities, multilevel decomposition, approximation and interpolation spaces.

*First author supported by *Programa Ramón y Cajal* and the TMR Network "Harmonic Analysis 1998-2002". The authors thank an anonymous referee whose careful reading and suggestions led to a much improved version of this paper.

†Some earlier results with not compactly supported wavelets can be found in [1].

restrictions in the indices $\boldsymbol{\alpha}, p, q$, but only for wavelet bases with a very special tensor product construction and for function spaces defined in the unit cube $[0,1]^2$. On the other hand, the results in [13] were carried out with a very general MRA setting in $\mathbb{R}^d$ which allows the use of non-separable wavelets, but with very specific techniques of linear operators which only produced characterizations of $B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)$ for indices $p, q \geq 1$. These two works, therefore, left open the study of the cases $0 < p < 1$ with a general MRA setting of non-separable wavelets, which we think may play a useful role in certain applications.

The goal of the present paper is precisely to fill this gap by presenting a uniform approach to $B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)$ valid in the whole range $0 < p, q < \infty$, and which takes into account the general MRA setting introduced in [13]. This requires a new strategy compared with [13], where classical methods based on Bernstein and Jackson inequalities must be combined with the non-linear techniques introduced by DeVore-Popov in the isotropic situation ([10], see also [3]). In fact, in this paper we shall focus more in the quasi-Banach setting: $0 < p < 1$, where non-linear methods appear naturally, and refer to [13] for a complete treatment of the linear methods in the Banach situation $p \geq 1$.

To be more precise, let us fix throughout the paper an anisotropy $\boldsymbol{a} = (a_1, \ldots, a_n)$, with all $a_1, \ldots, a_n > 0$ and normalized by $(\frac{1}{a_1} + \cdots + \frac{1}{a_n})/n = 1$. The wavelet system we are looking for will be dilated by a matrix M, which we shall choose to be "compatible" with $\boldsymbol{a}$. Namely, we let

$$M = \operatorname{diag}\,(\lambda^{\frac{1}{a_1}}, \ldots, \lambda^{\frac{1}{a_n}}), \quad \text{for some } \lambda > 1, \tag{1.1}$$

so that we have "the correct homogeneity" over Besov seminorms: $|\det M|^{\frac{1}{p}}\,|f(M\,\cdot)|_{B^{\alpha\boldsymbol{a}}_{p,q}} = \lambda^{\alpha}|f|_{B^{\alpha\boldsymbol{a}}_{p,q}}$.

Our wavelet approach is based on *multiresolution analysis (MRA)*. We consider a fixed *scaling function* $\varphi \in L^2(\mathbb{R}^n)$ (which we assume compactly supported) and define, for each $0 < p < \infty$, the multiresolution spaces

$$V_j^{(p)} = \overline{\operatorname{span}}_{L^p}\left\{\varphi^{(p)}_{j,\gamma} \equiv |\det M|^{j/p}\varphi(M^j\cdot - \gamma)\right\}_{\gamma\in\mathbb{Z}^n}, \quad j \in \mathbb{Z}. \tag{1.2}$$

As it is well-known, these spaces are nested $V_j^{(p)} \subset V_{j+1}^{(p)} \subset \ldots$, so that one obtains a *multilevel decomposition* for every $f \in L^p(\mathbb{R}^n)$:

$$f = P_0^{(p)} f + \sum_{j=0}^{\infty} Q_j^{(p)} f, \quad f \in L^p(\mathbb{R}^n), \tag{1.3}$$

where $P_j^{(p)}$ are suitable "projections" of L^p onto $V_j^{(p)}$ (which we shall define below) and $Q_j^{(p)}$ are "details" given by $Q_j^{(p)} := P_{j+1}^{(p)} - P_j^{(p)}$. The first result in this paper can then be stated as follows.

THEOREM 1.4 *Let $0 < p, q < \infty$ and $\alpha > 0$. Then, for an MRA as above with a compactly supported and sufficiently smooth scaling function we have:*

$$B^{\alpha\boldsymbol{a}}_{p,q}(\mathbb{R}^n) = \Big\{ f \in L^p(\mathbb{R}^n) \ : \ \sum_{j=0}^{\infty} \lambda^{\alpha j q} \|Q_j^{(p)} f\|_p^q < \infty \Big\}.$$

Moreover, the following equivalence of (quasi)-seminorms holds:

$$\|f\|_{B^{\alpha\boldsymbol{a}}_{p,q}} \sim \|P_0^{(p)} f\|_p + \Big[\sum_{j=0}^{\infty} \lambda^{\alpha j q} \|Q_j^{(p)} f\|_p^q \Big]^{\frac{1}{q}} \quad \textit{and} \quad |f|_{B^{\alpha\boldsymbol{a}}_{p,q}} \sim \Big[\sum_{j=-\infty}^{\infty} \lambda^{\alpha j q} \|Q_j^{(p)} f\|_p^q \Big]^{\frac{1}{q}}.$$

The previous result can also be translated into the language of approximation theory. In fact, if we denote the *error of approximation of $f \in L^p(\mathbb{R}^n)$ to $V_j^{(p)}$* by

$$E_j^{(p)}(f) := \inf_{g \in V_j^{(p)}} \|f - g\|_p, \quad j \in \mathbb{Z},$$

and define the approximation space of order (α, q) by

$$A_q^\alpha(L^p) = \Big\{ f \in L^p \ : \ \|f\|_p + \Big[\sum_{j=0}^{\infty} \lambda^{\alpha j q} \|E_j^{(p)} f\|_p^q \Big]^{\frac{1}{q}} < \infty \Big\},$$

then a reformulation of Theorem 1.4 will show that $A_q^\alpha(L^p) = B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)$ with equivalent norms (see Theorem 5.1 below). We must say that it is precisely this application which will motivate our definition of the projector $P_j^{(p)} f$ for $p < 1$, as an *almost best approximation* of f to $V_j^{(p)}$. Dealing with such projectors will require some extra care, since typically they are non-linear and of course they will not be continuous.

Before going into our second result, it is worth rephrasing our previous theorem when $p \geq 1$. In this case, if M is integer-valued and $m = |\det M|$, then it is well-known that there exists a *family of compactly supported wavelets* $\{\psi_1, \ldots, \psi_{m-1}\}$ so that $Q_j^{(p)}$ is a *linear projection* from L^p onto the subspace

$$W_j^{(p)} = \overline{\text{span}}_{L^p} \Big\{ \psi^{(p)}_{\ell;j,\gamma} \equiv |\det M|^{j/p} \psi_\ell(M^j \cdot - \gamma) \ : \ \gamma \in \mathbb{Z}^n, \, \ell = 1, \ldots, m-1 \Big\}$$

for every $j \in \mathbb{Z}$. Moreover, the multilevel decomposition in (1.3) takes the form

$$f = \sum_{\gamma \in \mathbb{Z}^n} c_\gamma \varphi_\gamma + \sum_{\ell=1}^{m-1} \sum_{j=0}^{\infty} \sum_{\gamma \in \mathbb{Z}^n} d_{\ell;j,\gamma} \psi^{(p)}_{\ell;j,\gamma}, \quad f \in L^p(\mathbb{R}^n) \tag{1.5}$$

and since the generators of $V_j^{(p)}$ and $W_j^{(p)}$ are *p-stable bases* we have

$$\|P_0^{(p)} f\|_p^p \sim \sum_{\gamma \in \mathbb{Z}^n} |c_\gamma|^p \quad \text{and} \quad \|Q_j^{(p)} f\|_p^p \sim \sum_{\ell=1}^{m-1} \sum_{\gamma \in \mathbb{Z}^n} |d_{\ell;j,\gamma}|^p, \quad j \in \mathbb{Z}.$$

Finally, the coefficients can be expressed in terms of a *dual basis* $\{\tilde{\varphi}_\gamma, \tilde{\psi}_{\ell;j,\gamma}\}_{\gamma,\ell,j}$ by

$$c_\gamma = \langle f, \tilde{\varphi}_\gamma \rangle \quad \text{and} \quad d_{\ell;j,\gamma} = \langle f, \tilde{\psi}^{(p')}_{\ell;j,\gamma} \rangle, \quad \gamma \in \mathbb{Z}^n, \ j \in \mathbb{Z}, \ \ell = 1, \ldots, m-1.$$

If we insert this in Theorem 1.4 we obtain, when $p \geq 1$, a third and much more useful characterization of Besov spaces in terms of size of the wavelet coefficients (compared with the size of "details" and approximation errors above). These are the results obtained in [13].

Now observe that there is no hope that such reasoning can be applied when $p < 1$, since the non-linearity of $P_j^{(p)}$ makes the space Im $Q_j^{(p)}$ typically much larger than $W_j^{(p)}$. Our second main theorem gives a substitute for this situation and shows that, in a certain range of indices, the Besov spaces $B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)$ with $\tau < 1$ can be embedded in L^p for some $p > 1$, and have corresponding characterizations in terms of weighted sums of wavelet coefficients.

Theorem 1.6 *Let $0 < \tau < \infty$, $n(\frac{1}{\tau} - 1)_+ < \alpha < \frac{n}{\tau}$ and $p > 1$ given by $\frac{1}{\tau} = \frac{\alpha}{n} + \frac{1}{p}$. Then, $B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n) \hookrightarrow L^p$, and for an MRA as above we have*

$$B^{\alpha\boldsymbol{a}}_{\tau,\tau} = \{f \in L^p(\mathbb{R}^n) \ : \ \sum_{\gamma \in \mathbb{Z}^n} |\langle f, \tilde{\varphi}_\gamma \rangle|^\tau + \sum_{\ell=1}^{m-1} \sum_{j=0}^{\infty} \sum_{\gamma \in \mathbb{Z}^n} |\langle f, \tilde{\psi}^{(p')}_{\ell;j,\gamma} \rangle|^\tau < \infty\}.$$

Moreover, the following equivalence of quasi-seminorms holds:

$$\|f\|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \sim [\sum_{\gamma \in \mathbb{Z}^n} |\langle f, \tilde{\varphi}_\gamma \rangle|^\tau]^{\frac{1}{\tau}} + [\sum_{\ell=1}^{m-1} \sum_{j=0}^{\infty} \sum_{\gamma \in \mathbb{Z}^n} |\langle f, \tilde{\psi}^{(p')}_{\ell;j,\gamma} \rangle|^\tau]^{\frac{1}{\tau}}$$

$$|f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \sim \Big[\sum_{\ell=1}^{m-1} \sum_{j \in \mathbb{Z}} \sum_{\gamma \in \mathbb{Z}^n} |\langle f, \tilde{\psi}^{(p')}_{\ell;j,\gamma} \rangle|^\tau\Big]^{\frac{1}{\tau}}.$$

As far as we know, a complete proof of the previous result, even in the isotropic situation, has only recently appeared in [4, Th. 30.7], while previous works just focused in the domain $[0,1]^n$ (see[10]), or in *homogeneous* Besov spaces [8]. We also refer to [4] for modern applications of such theorems to problems involving thresholding.

Finally, we mention that among the consequences of our theorems there some interpolation and embedding results for anisotropic Besov spaces which seem to be new in the literature when $p < 1$. These are presented in §5, complementing those obtained for the unit cube in [15] and the isotropic case in [10]. We also make special reference to the preprint [19], discovered after the write-up of this manuscript, and which contains related results in non-linear approximation and interpolation of anisotropic Besov spaces.

2 Background on anisotropic Besov spaces

Fix $\boldsymbol{a} = (a_1, \ldots, a_n)$, with $a_j > 0$ and $\frac{1}{a_1} + \cdots + \frac{1}{a_n} = n$. We denote by

$$|h|_{\boldsymbol{a}} = \sum_{j=1}^{n} |h_j|^{a_j}, \quad h = (h_1, \ldots, h_n) \in \mathbb{R}^n$$

the *anisotropic pseudo-distance* related to $\boldsymbol{a}$. Observe that $|Mh|_{\boldsymbol{a}} = \lambda |h|_{\boldsymbol{a}}$. We use the standard notation $\Delta_h^k f$ for the k^{th}-iteration of the difference operator: $\Delta_h f(x) = f(x+h) - f(x)$, $k \geq 1$, $x, h \in \mathbb{R}^n$. Also, if $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ is the canonical basis of $\mathbb{R}^n$, we write

$$\Delta_h^{\boldsymbol{\ell}} f(x) = (\Delta_{h_1 \boldsymbol{e}_1}^{\ell_1} \circ \ldots \circ \Delta_{h_n \boldsymbol{e}_n}^{\ell_n}) f(x), \quad x \in \mathbb{R}^n,$$

for the *mixed difference* of order $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n) \in \mathbb{N}^n$ and step $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$.

Let $0 < \alpha, p, q < \infty$ and write $\boldsymbol{\alpha} = \alpha \boldsymbol{a} = (\alpha_1, \ldots, \alpha_n)$. Consider the smallest positive integer $L > \max\{\alpha_1, \ldots, \alpha_n\}$. We say that a function $f \in L^p(\mathbb{R}^n)$ belongs to the *anisotropic Besov space* $B_{p,q}^{\boldsymbol{\alpha}}(\mathbb{R}^n)$ whenever

$$|f|_{B_{p,q}^{\boldsymbol{\alpha}}} := \sum_{|\boldsymbol{\ell}|=L} \left[\int_{\mathbb{R}^n} \Big(|h|_{\boldsymbol{a}}^{-\alpha} \|\Delta_h^{\boldsymbol{\ell}} f\|_p \Big)^q \frac{dh}{|h|_{\boldsymbol{a}}^n} \right]^{1/q} < \infty. \tag{2.1}$$

We denote the corresponding Besov quasi-norm by $\|f\|_{B_{p,q}^{\boldsymbol{\alpha}}} = \|f\|_p + |f|_{B_{p,q}^{\boldsymbol{\alpha}}}$.

We observe that our definition of seminorm in (2.1) is equivalent to other expressions which may appear in the literature, with $\Delta_h^{\boldsymbol{\ell}} f$ replaced by the *total difference operator* $\Delta_h^L f$ or by the *partial differences* $\Delta_{t\boldsymbol{e}_i}^{k_i} f$, for $k_i > \alpha_i$ (see e.g. [23, §2.3.4], or [22, §1.4] where proofs of the equivalences are given, valid also for $p < 1$). In this paper we shall stick to the mixed differences in (2.1), and some equivalent expressions using the *mixed modulus of continuity*

$$\omega_p^{\boldsymbol{\ell}}(f; t) = \sup_{|h|_{\boldsymbol{a}} \leq t} \|\Delta_h^{\boldsymbol{\ell}} f\|_p, \quad t > 0.$$

Proposition 2.2 *The following quasi-seminorms are equivalent:*

$$|f|_{B_{p,q}^{\boldsymbol{\alpha}}} \sim \sum_{|\boldsymbol{\ell}|=L} \left[\int_0^\infty \Big(t^{-\alpha} \omega_p^{\boldsymbol{\ell}}(f; t) \Big)^q \frac{dt}{t} \right]^{1/q} \sim \sum_{|\boldsymbol{\ell}|=L} \left[\sum_{j \in \mathbb{Z}} \Big(\lambda^{j\alpha} \omega_p^{\boldsymbol{\ell}}(f; \lambda^{-j}) \Big)^q \right]^{1/q}. \tag{2.3}$$

We sketch the proof of this proposition, since we could not find a precise reference in the anisotropic setting; the ideas, however, are more or less standard. We just need to show the first equivalence, since the last sum is just a discretization of the middle integral following from the monotonicity of the modulus of continuity. For the first assertion, the key is the equivalence between *averaged modulus of continuity*, which we state separately below.

Proposition 2.4 *Let $0 < p, q < \infty$ and $\boldsymbol{\ell} = (\ell_1, \dots, \ell_n)$ with $\ell_j \geq 1$. Then, there exists $C = C(p, q, \boldsymbol{\ell}, \boldsymbol{a}) > 0$ and $c = c(\boldsymbol{\ell}, \boldsymbol{a}) \geq 1$ such that*

$$\omega_p^{\boldsymbol{\ell}}(f; t) \leq C \left(\frac{1}{t^n} \int_{|h|_{\boldsymbol{a}} \leq ct} \|\Delta_h^{\boldsymbol{\ell}} f\|_p^q \, dh \right)^{\frac{1}{q}}, \quad \forall\, t > 0. \tag{2.5}$$

Proof: To prove (2.5) we need an identity involving mixed difference operators with variable increments: for all $x, h, s \in \mathbb{R}^n$, and $\boldsymbol{m} \geq \mathbf{1}$

$$\Delta_h^{\boldsymbol{m}} f(x) = \sum_{\mathbf{1} \leq \mathbf{k} \leq \boldsymbol{m}} (-1)^{|\mathbf{k}|} \binom{\boldsymbol{m}}{\mathbf{k}} \left[\sum_{\mathbf{0} \leq \boldsymbol{\delta} \leq \mathbf{1}} (-1)^{|\boldsymbol{\delta}|} \Delta_{\boldsymbol{\delta} h + \mathbf{k} s}^{\boldsymbol{m}} f(x + (\mathbf{1} - \boldsymbol{\delta}) \mathbf{k} h) \right]. \tag{2.6}$$

Above we use common notation for multi-indices: $|\boldsymbol{m}| = m_1 + \ldots + m_n$, $\binom{\boldsymbol{m}}{\mathbf{k}} = \binom{m_1}{k_1} \ldots \binom{m_n}{k_n}$, $\boldsymbol{m}h = (m_1 h_1, \dots, m_n h_n)$ and $\boldsymbol{m} \geq \mathbf{1}$ if $m_j \geq 1$ for all $j = 1, \dots, n$. The identity (2.6) is an anisotropic version of the 1-dimensional formula

$$\Delta_h^m f(x) = \sum_{k=1}^m (-1)^k \binom{m}{k} \left[\Delta_{ks}^m f(x + kh) - \Delta_{h+ks}^m f(x) \right], \quad x, h, s \in \mathbb{R}$$

(see [9, p. 184]). We could not find any reference for (2.6), but once the formula is set, the reader can easily prove it by induction using the 1-dimensional case. To pass from (2.6) to (2.5), one first takes L^p-norms in both sides of (2.6) and raises to the q^{th} power using the quasi-triangular inequality. Next, integrate the resulting expressions over $|s|_{\boldsymbol{a}} \leq t$ and enlarge appropriately the region of integration on the right, so that after taking the sup over h one concludes with (2.5).

□

Remark 2.7 As a final remark we should warn the reader that, when $p < 1$, our spaces *do not coincide in general* with the Besov spaces defined with Littlewood-Paley theory by Triebel et al. In particular, when $\alpha < n(\frac{1}{p} - 1)_+$ our spaces $B_{p,q}^{\boldsymbol{\alpha}}(\mathbb{R}^n)$ contain functions which are not distributions, while Triebel's spaces may contain distributions which are not functions (see e.g. [21]). The interest of functional definitions (rather than distributional ones) seem to be motivated by various applications in approximation theory (see [10]).

3 Multiresolution and non-linear projectors

We recall briefly the multiresolution setting introduced in [13]. Given the dilation matrix M in (1.1), we fix a *pair of biorthogonal M-scaling functions $\{\varphi, \tilde{\varphi}\}$, which we assume bounded and compactly supported.* These are crucial assumptions in applications, so we do not consider here more general settings where the scaling functions have a different decay. We also recall from §3.3 in [13] that the theory of multiresolution analyses is naturally limited to matrices M *with integer entries*, a constraint which together with (1.1) implies that the anisotropies must be of the form $(\frac{1}{a_1},\ldots,\frac{1}{a_n})\in\mathbb{R}_+\log\mathbb{Z}^n_+$. From now on we shall only consider such anisotropies, although this constraint will not be used anywhere else in the paper (apart from the existence of $\{\varphi,\tilde\varphi\}$). We refer to §7.2 below for alternative ideas in other situations.

We shall denote by $\{V_j^{(p)}\}_{j\in\mathbb{Z}}$ the multiresolution spaces generated by φ as in (1.2). Observe that $f\in V_j^{(p)}$ if and only if $T^{-j}f\in V_0^{(p)}$, where $T=T_{(p)}$ denotes the isometry in $L^p(\mathbb{R}^n)$ given by $Tf(x)=|\det M|^{1/p}f(Mx)$.

We next show that the systems $\{\lambda^{jd/p}\varphi(M^j\cdot-\gamma)\}_{\gamma\in\mathbb{Z}^n}$ are p-stable bases of $V_j^{(p)}$, which of course suffices to do for $j=0$. We present a different argument than Cohen et al. (see [4, §3.7]), where only tensor product scaling functions could be handled. These ideas will also be crucial later, when dealing with the non-linear projectors $P_j^{(p)}$.

Proposition 3.1 *Let $\varphi,\tilde\varphi$ be bounded and compactly supported biorthogonal scaling functions. Then $\{\varphi_\gamma\}_{\gamma\in\mathbb{Z}^n}$ is a p-stable basis of $V_0^{(p)}$, meaning that every $f\in V_0^{(p)}$ can be written uniquely as $f=\sum_{\gamma\in\mathbb{Z}^n}c_\gamma\varphi_\gamma$ for some $\{c_\gamma\}\in\ell^p(\mathbb{Z}^n)$, and moreover*

$$\|\sum_{\gamma\in\mathbb{Z}^n}c_\gamma\varphi_\gamma\|_{L^p(\mathbb{R}^n)}\sim\Big[\sum_{\gamma\in\mathbb{Z}^n}|c_\gamma|^p\Big]^{\frac{1}{p}}.$$

Proof: The proof for $1\le p<\infty$ is standard (see the Appendix in [13]). We assume therefore $0<p<1$. Then, for every sequence $\{c_k\}_{k\in\mathbb{Z}^n}\in\ell^p(\mathbb{Z}^n)$, we have

$$\int_{\mathbb{R}^n}|\sum_{k\in\mathbb{Z}^n}c_k\varphi(x-k)|^p\,dx\le\sum_{k\in\mathbb{Z}^n}|c_k|^p\,\|\varphi\|^p_{L^p}.$$

Thus, the series $\sum_{k\in\mathbb{Z}^n}c_k\varphi_k$ converges in L^p and defines a function in $V_0^{(p)}$. Conversely, let $f=\sum_{k\in\mathbb{Z}^n}c_k\varphi_k\in V_0^{(p)}$, for a finite sequence of scalars $\{c_k\}$. Then, the biorthogonality gives $c_k=\langle f,\tilde\varphi_k\rangle$, and therefore

$$|c_k|\le\int_{\mathbb{R}^n}|f|\,|\tilde\varphi_k|\;\le\;\|\tilde\varphi\|_\infty\int_{\widetilde J_k}|f|,$$

where we denote by $\widetilde J_k$ the *compact set* $\widetilde J_k=\operatorname{Supp}\tilde\varphi_k=\widetilde J_0+k$, $k\in\mathbb{Z}^n$. Now, the key step is the equality of the finite-dimensional subspaces of $L^p(\widetilde J_k)$

$$V_0^{(p)}|_{\widetilde J_k}:=\Big\{g\chi_{\widetilde J_k}\;:\;g\in V_0^{(p)}\Big\}=\operatorname{span}\Big\{\varphi(\cdot-\ell)\chi_{\widetilde J_k}\;:\;\ell\in\mathbb{Z}^n\Big\}.\qquad(3.2)$$

We conclude that $V_0^{(p)}|_{\widetilde J_k}$ is *independent of* p, and by the equivalence of all quasinorms in finite dimensional spaces we also have the reverse Hölder estimate

$$\int_{\widetilde J_k}|g|\le c\left(\int_{\widetilde J_k}|g|^p\right)^{\frac{1}{p}},\quad\forall\,g\in V_0^{(p)},\qquad(3.3)$$

with a constant $c>0$ independent of g, and by translation-invariance also independent of k. Thus,

$$\sum_{k\in\mathbb{Z}^n}|c_k|^p \leq \|\tilde{\varphi}\|_\infty^p \sum_{k\in\mathbb{Z}^n}\Big(\int_{\tilde{J}_k}|f|\Big)^p \leq c^p \sum_{k\in\mathbb{Z}^n}\int_{\tilde{J}_k}|f|^p \lesssim \int_{\mathbb{R}^n}|f|^p,$$

at least for $f\in\operatorname{span}\{\varphi_k\}$. A density argument then establishes the proposition. $\square$

Let us now pass to the construction of the projectors onto $V_j^{(p)}$. When $p\geq 1$, one uses the standard definition in terms of φ and $\tilde{\varphi}$:

$$P_j = P_j^{(p)}\colon L^p(\mathbb{R}^n)\longrightarrow V_j^{(p)}, \quad f\longmapsto P_jf=\sum_{k\in\mathbb{Z}^n}\langle f,\tilde{\varphi}_{j,k}^{(p')}\rangle\varphi_{j,k}^{(p)}.$$

We recall that these are bounded linear projections satisfying:

(1) $\|f-P_jf\|_p\lesssim E_j^{(p)}(f)=\inf_{g\in V_j^{(p)}}\|f-g\|_p$;

(2) $\|f-P_jf\|_p\to 0$, as $j\to+\infty$.

(3) $\|P_jf\|_p\to 0$, as $j\to-\infty$ (only when $1<p<\infty$).

(4) $f=P_0f+\sum_{j=0}^\infty Q_jf \quad \Big(=\sum_{j\in\mathbb{Z}}Q_jf, \quad \text{if } p>1\Big), \quad \forall\, f\in L^p(\mathbb{R}^n)$,

where in the last formula $Q_j=P_{j+1}-P_j$, and the convergence is in the L^p-norm.

For the cases $0<p<1$, the previous definition of P_j is no longer valid since functions in $L^p(\mathbb{R}^n)$ may not be locally integrable. We proceed instead with a different definition based on non-linear approximation (see the abstract framework in [3, 10]). More precisely, consider first the space of "piecewise $V_0^{(p)}$-functions":

$$\mathcal{V}_0^{(p)} := \Big\{g\in L^p_{loc}(\mathbb{R}^n) \quad | \quad \forall\,\ell\in\mathbb{Z}^n,\ \exists\, f^\ell\in V_0^{(p)}\ :\ g|_{I_\ell}=f^\ell|_{I_\ell}\Big\},$$

where we denote $I_\ell=[0,1]^n+\ell$, $\ell\in\mathbb{Z}^n$. Note that $\mathcal{V}_0^{(p)}$ is *a priori* much larger than $V_0^{(p)}$. Next, we define a mapping assigning to every $f\in L^p_{loc}(\mathbb{R}^n)$ a function $R_0^{(p)}f\in\mathcal{V}_0^{(p)}$ so that

$$\|f-R_0^{(p)}f\|_{L^p(I_\ell)} = \inf_{g\in V_0^{(p)}}\|f-g\|_{L^p(I_\ell)}, \quad \forall\ \ell\in\mathbb{Z}^n. \tag{3.4}$$

That is, in each cube I_ℓ, $(R_0^{(p)}f)|_{I_\ell}$ is a *best approximation to* $f|_{I_\ell}$ *in* $V_0^{(p)}|_{I_\ell}$, which exists because the latter is a finite dimensional (hence closed) subspace of $L^p(I_\ell)$ (see

(3.2)). Since minimizers may be hard to compute, there is no loss if we take instead *near best approximations*, that is if we define $R_0^{(p)} f \in \mathcal{V}_0^{(p)}$ so that

$$\|f - R_0^{(p)} f\|_{L^p(I_\ell)} \leq 2 \inf_{g \in V_0^{(p)}} \|f - g\|_{L^p(I_\ell)}, \quad \forall \ \ell \in \mathbb{Z}^n. \tag{3.5}$$

In either case we observe $R_0^{(p)}$ is not necessarily continuous, nor even linear. From (3.5), however, it easily follows that $R_0^{(p)} f = f$ if $f \in \mathcal{V}_0^{(p)}$, and for all $f \in L^p(\mathbb{R}^n)$

$$\|R_0^{(p)} f\|^p_{L^p(\mathbb{R}^n)} \leq \|f\|^p_{L^p(\mathbb{R}^n)} + \sum_{\ell \in \mathbb{Z}^n} \|f - R_0^{(p)} f\|^p_{L^p(I_\ell)} \leq 3 \, \|f\|^p_{L^p(\mathbb{R}^n)}. \tag{3.6}$$

Finally, we define $P_0^{(p)}$ by composition of $R_0^{(p)}$ with $P_0^{(1)}$:

$$f \in L^p(\mathbb{R}^n) \longmapsto P_0^{(p)} f := P_0^{(1)}(R_0^{(p)} f) = \sum_{k \in \mathbb{Z}^n} \langle R_0^{(p)} f, \tilde{\varphi}_k \rangle \varphi_k.$$

Proposition 3.7 *For every $0 < p \leq 1$, the (non-linear) operator $P_0^{(p)}$ is a bounded projection from $L^p(\mathbb{R}^n)$ onto $V_0^{(p)}$, that is*

$$P_0^{(p)} f = f \quad \forall \ f \in V_0^{(p)}, \quad \textit{and} \quad \|P_0^{(p)} f\|_p \leq C \, \|f\|_p, \quad f \in L^p(\mathbb{R}^n). \tag{3.8}$$

Moreover, there exists $C > 0$ such that

$$\|f - P_0^{(p)} f\|_p \leq C \inf_{g \in V_0^{(p)}} \|f - g\|_p, \quad \forall \ f \in L^p(\mathbb{R}^n). \tag{3.9}$$

Proof: For the first inequality, the same argument we gave in the proof of Proposition 3.1 shows that

$$\|P_0^{(1)} g\|_p \lesssim \Big(\sum_{k \in \mathbb{Z}^n} |\langle g, \tilde{\varphi}_k \rangle|^p \Big)^{\frac{1}{p}} \lesssim \|g\|_p \quad \forall \ g \in \mathcal{V}_0^{(p)}, \tag{3.10}$$

and thus using (3.6) we easily obtain (3.8). The left hand identity of (3.8) also follows from Proposition 3.1, since by p-stability every $f \in V_0^{(p)}$ can be written as $f = \sum_{k \in \mathbb{Z}^n} \langle f, \tilde{\varphi}_k \rangle \varphi_k$ and we also have $R_0^{(p)} f = f$. For the last assertion, we take any $g \in \operatorname{span}\{\varphi_k\}$, and use the previous results and the quasi-triangular inequality to obtain

$$\begin{aligned}
\|f - P_0^{(p)} f\|_p &\lesssim \|f - g\|_p + \|g - P_0^{(p)} f\|_p \\
&= \|f - g\|_p + \|P_0^{(1)}(g - R_0^{(p)} f)\|_p \\
\text{by (3.10)} \quad &\lesssim \|f - g\|_p + \|g - R_0^{(p)} f\|_p \\
&\lesssim \|f - g\|_p + \|f - R_0^{(p)} f\|_p \\
\text{by (3.5)} \quad &\lesssim \|f - g\|_p + \Big(\sum_{\ell \in \mathbb{Z}^n} \|f - g\|^p_{L^p(I_\ell)} \Big)^{\frac{1}{p}} = 2 \|f - g\|_p.
\end{aligned}$$

Taking the infimum over all $g \in \operatorname{span}\{\varphi_k\}$, which is dense in $V_0^{(p)}$, we obtain (3.9). $\square$

Remark 3.11 It is worth noting that for all $0 < \rho < p$, the (non-linear) operator $P_0^{(\rho)}$ also defines a bounded projection from $L^p(\mathbb{R}^n)$ onto $V_0^{(p)}$. This is essentially due to the fact that $(R_0^{(\rho)} f)|_{I_\ell}$ is a near best approximation to $f|_{I_\ell}$ in $V_0^{(p)}|_{I_\ell}$ (with a constant only depending on p, ρ and n), as it can be shown using reverse Hölder inequalities in $V_0^{(p)}|_{I_\ell} = V^{(\rho)}|_{I_\ell}$. Moreover, mimicking the same steps as in Proposition 3.7 one can also obtain (3.9) with $P_0^{(p)}$ replaced by $P_0^{(\rho)}$. We leave details to the reader.

Finally, one could repeat these same arguments to define projectors $P_j^{(p)}$ from $L^p(\mathbb{R}^n)$ onto $V_j^{(p)}$, via best approximations to $f|_{I_{j,\ell}}$ in $V_j^{(p)}|_{I_{j,\ell}}$, with $I_{j,\ell} = M^{-j} I_\ell$. This approach is equivalent to just let $P_j^{(p)} := T^j P_0^{(p)} T^{-j}$, $j \in \mathbb{Z}$, which by homogeneity will satisfy analogous properties as $P_0^{(p)}$ in Proposition 3.7.

4 Bernstein and Jackson inequalities

The characterization of Besov spaces is typically obtained from direct and inverse inequalities involving moduli of smoothness. In this section we shall show that such inequalities hold for any of our "admissible" MRA's, provided their corresponding scaling functions are smooth enough. We begin with the inverse inequality, which is also the simplest.

Proposition 4.1 : Bernstein-type inequality.
If $\varphi \in B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)$, then

$$|f|_{B^{\boldsymbol{\alpha}}_{p,q}} \lesssim \lambda^{\alpha j} \, \|f\|_{L^p}, \quad \forall \ f \in V_j^{(p)}. \tag{4.2}$$

Proof: Let us choose an integer $L > \max\{\alpha_1, \ldots, \alpha_n\}$. By scaling it suffices to show (4.2) for $j = 0$. We only include the proof for $0 < p < 1$ (the other case is similar, see [13]). If $f = \sum_{k \in \mathbb{Z}^n} c_k \varphi_k \in V_0^{(p)}$, then for every $h \in \mathbb{R}^n$ the p-stability gives

$$\|\Delta_h^{\boldsymbol{\ell}} f\|_p^p \leq \sum_{k \in \mathbb{Z}^n} |c_k|^p \, \|\Delta_h^{\boldsymbol{\ell}} \varphi_k\|_p^p \lesssim \|f\|_p^p \, \|\Delta_h^{\boldsymbol{\ell}} \varphi\|_p^p.$$

Thus,

$$|f|_{B^{\boldsymbol{\alpha}}_{p,q}} = \sum_{|\boldsymbol{\ell}| = L} \Big\{ \Big[\int_{\mathbb{R}^n} (|h|_{\boldsymbol{a}}^{-\alpha} \|\Delta_h^{\boldsymbol{\ell}} f\|_p)^q \frac{dh}{|h|_{\boldsymbol{a}}^n} \Big]^{1/q} \lesssim \|f\|_p \, |\varphi|_{B^{\boldsymbol{\alpha}}_{p,q}},$$

completing the proof of (4.2). □

For the direct inequality we need an additional assumption on $\{\varphi, \tilde{\varphi}\}$ guaranteing that there is an integer $L > \max\{\alpha_1, \ldots, \alpha_n\}$ so that *the space $V_0^{(p)}$ contains locally the polynomials of total degree $< L$*:

$$\Pi_L = \operatorname{span}\{x_1^{r_1} \cdots x_n^{r_n} \ : \ r_1 + \ldots + r_n < L\}.$$

This means that for every compact set K and $P \in \Pi_L$ there exists $f \in V_0^{(p)}$ such that $f|_K = P|_K$. Such assumption (which does not depend on $0 < p < \infty$) is known to follow from sufficient smoothness of the scaling function (e.g., $\varphi \in H^{L\boldsymbol{\alpha}}(\mathbb{R}^n)$; see [5, Th. 2.1]). For the particular case $L = 1$, in order that V_0 contains locally the constants it suffices with the mild condition $\varphi, \tilde{\varphi} \in L^\infty_c(\mathbb{R}^n)$ (see, e.g, [17, Th. 6.3]).

Theorem 4.3 : Jackson-type inequality.
If $V_0^{(p)}$ contains locally the polynomials of degree $< L$, then

$$\|f - P_j^{(p)} f\|_p \lesssim \lambda^{-\alpha j} \, |f|_{B^{\boldsymbol{\alpha}}_{p,q}}, \quad \forall\ f \in B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n). \tag{4.4}$$

Proof: Again, a scaling argument reduces matters to $j = 0$. We will just prove the case $0 < p < 1$ (the other one being a bit simpler; see [13]). Following a more or less standard procedure, we first need to establish two lemmas:

Lemma 4.5 : Local boundedness.
In the conditions above, there exists a finite set $K_\# \subset \mathbb{Z}^n$ such that, for every $\gamma \in \mathbb{Z}^n$, we have

$$\|P_0^{(p)} f\|_{L^p(I_\gamma)} \lesssim \|f\|_{L^p(\tilde{I}_\gamma)}\,, \quad \forall\ f \in L^p(\mathbb{R}^n), \tag{4.6}$$

where $\tilde{I}_\gamma = \cup_{m \in K_\#} (I_\gamma + m)$.

Lemma 4.7 : Whitney-type inequality.
Let $f \in L^p_{loc}(\mathbb{R}^n)$ and $J = [a_1, b_1] \times \ldots \times [a_n, b_n]$ be a rectangle in $\mathbb{R}^n$. Then, for every $L \geq 1$ there exists a polynomial $Q \in \Pi_L$ such that

$$\|f - Q\|^p_{L^p(J)} \leq C \sum_{|\boldsymbol{\ell}| = L} \frac{1}{|J|} \int_0^{b_1 - a_1} \cdots \int_0^{b_n - a_n} \|\Delta_h^{\boldsymbol{\ell}} f\|^p_{L^p(J)} \, dh, \tag{4.8}$$

where $C = C(n, p, L) > 0$.

The first lemma has a similar proof as in the isotropic case (see [3]). For completeness we include it below in the case $p < 1$.

Proof of Lemma 4.5:
We may assume $\gamma = 0$ and let $I = [0, 1]^n$. By definition of the projector

$$\|P_0^{(p)} f\|^p_{L^p(I)} \leq \sum_m |\langle R_0^{(p)} f, \tilde{\varphi}_m \rangle|^p \int_I |\varphi_m|^p,$$

with the indices restricted to the finite set $K_\sharp = \{m \in \mathbb{Z}^n \,:\, I \cap [\mathrm{Supp}\ \varphi + m] \neq \emptyset\}$. It therefore suffices to estimate one coefficient, which we may take $m = 0$. Now, proceeding as in Proposition 3.1,

$$|\langle R_0^{(p)} f, \tilde{\varphi} \rangle| \leq \|\tilde{\varphi}\|_\infty \int_{\tilde{J}_0} |R_0^{(p)} f| \leq c' \left[\int_{\tilde{J}_0} |R_0^{(p)} f|^p \right]^{\frac{1}{p}},$$

where we have used the fact that $R_0^{(p)} f$ is locally in $V_0^{(p)}$ and the equivalence of quasi-norms in (3.3). The lemma follows now easily as we did (3.6).

□

The second lemma is the crucial point in most Jackson-type inequalities. The proof for $p < 1$ is much more subtle than its counterpart for $p \geq 1$ (see e.g. [13, Lemma 4.3] for a simple proof in this last case). A classical reference is the paper [24], but since our formulation is slightly different we indicate below how to derive it from there.

Proof of Lemma 4.7:

We recall the statement of Theorem 1 in [24]: *if $f \in L^p(J)$, then there exists a polynomial $Q \in \Pi_L$ such that*

$$\|f - Q\|^p_{L^p(J)} \leq C \sum_{|\boldsymbol{\ell}|=L} \sup_{h_j < (b_j - a_j)/(2L)} \|\Delta^{\boldsymbol{\ell}}_h f\|^p_{L^p(J^{\boldsymbol{\ell}}_h)} \tag{4.9}$$

where the rectangle $J^{\boldsymbol{\ell}}_h = J \cap (J - \boldsymbol{\ell} h)$ is domain of definition $\Delta^{\boldsymbol{\ell}}_h f$ (if one looks at f as a function just defined in J). To pass from the right hand side of (4.9) to (4.8) one uses one more time the identity of the difference operators in (2.6). Indeed, taking $L^p(J^{\boldsymbol{\ell}}_h)$-norms in (2.6) and changing variables in the right hand side we obtain the estimate:

$$\|\Delta^{\boldsymbol{\ell}}_h f\|^p_{L^p(J^{\boldsymbol{\ell}}_h)} \leq \sum_{\mathbf{0} \leq \boldsymbol{\delta} \leq \mathbf{1} \leq \mathbf{k} \leq \boldsymbol{\ell}} c_{\mathbf{k},\boldsymbol{\ell}} \|\Delta^{\boldsymbol{\ell}}_{\boldsymbol{\delta} h + \mathbf{k} s} f\|^p_{L^p(J)}.$$

Then, one integrates this inequality in $0 \leq s_j \leq \frac{b_j - a_j}{2L}$, takes the sup over $h_j < \frac{b_j - a_j}{2L}$, and after changing variables concludes easily with (4.8).

□

We turn now to the proof of (4.4), which as before it suffices to do for $j = 0$. Let us fix $\gamma \in \mathbb{Z}^n$ and a polynomial Q of degree $< L$. Note that from our assumptions we have $Q \in \mathcal{V}_0^{(p)}$. Then, using Lemma 4.5 and (3.5) we obtain:

$$\begin{aligned} \|f - P_0^{(p)} f\|_{L^p(I_\gamma)} &\lesssim \|f - Q\|_{L^p(I_\gamma)} + \|P_0^{(1)}(Q - R_0^{(p)} f)\|_{L^p(I_\gamma)} \\ &\lesssim \|f - Q\|_{L^p(J_\gamma)} + \|f - R_0^{(p)} f\|_{L^p(J_\gamma)} \lesssim \|f - Q\|_{L^p(J_\gamma)}. \end{aligned}$$

Now, if for each $\gamma \in \mathbb{Z}^n$ we choose a polynomial $Q = Q^\gamma$ as in Lemma 4.7 we obtain

$$\begin{aligned} \|f - P_0^{(p)} f\|^p_{L^p(\mathbb{R}^n)} &= \sum_{\gamma \in \mathbb{Z}^n} \|f - P_0^{(p)} f\|^p_{L^p(I_\gamma)} \lesssim \sum_{\gamma \in \mathbb{Z}^n} \sum_{|\boldsymbol{\ell}|=L} \frac{1}{c^n} \int_{[0,c]^n} \|\Delta^{\boldsymbol{\ell}}_h f\|^p_{L^p(J_\gamma)} \, dh \\ &\lesssim \sum_{|\boldsymbol{\ell}|=L} \frac{1}{c^n} \int_{[0,c]^n} \|\Delta^{\boldsymbol{\ell}}_h f\|^p_{L^p(\mathbb{R}^n)} \, dh \lesssim \sum_{|\boldsymbol{\ell}|=L} \omega^{\boldsymbol{\ell}}_p(f, c')^p. \end{aligned} \tag{4.10}$$

One now uses the equivalence between averaged moduli of continuity in Proposition (2.4) to conclude

$$\|f - P_0^{(p)} f\|_{L^p(\mathbb{R}^n)} \lesssim \sum_{|\boldsymbol{\ell}|=L} \left[\int_{|h|_{\boldsymbol{a}} \le c''} \frac{|h|_{\boldsymbol{a}}^{n+\alpha q}}{|h|_{\boldsymbol{a}}^{n+\alpha q}} \|\Delta_h^{\boldsymbol{\ell}} f\|_{L^p(\mathbb{R}^n)}^q dh \right]^{1/q} \lesssim |f|_{B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)} ,$$

recovering in the last step the definition of seminorm in (2.1).

□

Remark 4.11 It is possible to weaken the assumptions of Theorem 4.3 using the following version of Lemma 4.7: *for all* $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n) \ge 0$*, there exists a polynomial* $Q \in \Pi_{\boldsymbol{\ell}}$ *so that*

$$\inf_{Q \in \Pi_{\boldsymbol{\ell}}} \|f - Q\|_{L^p(J)}^p \lesssim \sum_{j=1}^n \frac{1}{b_j - a_j} \int_0^{b_j - a_j} \|\Delta_{t\boldsymbol{e}_j}^{\ell_j} f\|_{L^p(J)}^p \, dt,$$

where $\Pi_{\boldsymbol{\ell}} = \mathrm{span}\{x_j^{r_j} \,:\, 0 \le r_j < \ell_j,\ j = 1, \ldots, n\}$. A self-contained proof for this result appears in [15, Th. 3.1]. Using this, one can prove (4.4) by just assuming that V_0 contains locally $\Pi_{\boldsymbol{\ell}}$ for some $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n)$ with $\ell_j > \alpha_j$. At this point, however, we do not know of any simple translation of this condition in terms of the smoothness of φ (except in the simpler tensor product case).

As a corollary of Theorem 4.3 with $L = 1$ we obtain a general property of the non-linear projectors defined in §3.

Corollary 4.12 *If* $\varphi, \tilde{\varphi} \in L_c^\infty(\mathbb{R}^n)$*, then the (non-linear) projectors* $P_j^{(p)}$ *satisfy*

$$\lim_{j \to \infty} \|f - P_j^{(p)} f\|_{L^p(\mathbb{R}^n)} = 0, \quad \forall \ f \in L^p(\mathbb{R}^n).$$

Proof: Let $\varepsilon > 0$ and $f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$ be such that $\|f - f_\varepsilon\|_p < \varepsilon$. By the case $L = 1$ in Theorem 4.3 (which we can use since V_0 reproduces the constants when $\varphi, \tilde{\varphi} \in L_c^\infty(\mathbb{R}^n)$; [17, Th. 6.3]), it follows that $\|f_\varepsilon - P_j^{(p)} f_\varepsilon\|_p \lesssim \lambda^{-j\alpha} |f|_{B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)}$, for some $\alpha > 0$. To conclude we need to use (3.9), since $P_j^{(p)}$ is not linear in general:

$$\begin{aligned} \|f - P_j^{(p)} f\|_p &\lesssim \inf_{g \in V_j^{(p)}} \|f - g\|_p \le \|f - P_j^{(p)} f_\varepsilon\|_p \\ &\lesssim \|f - f_\varepsilon\|_p + \|f_\varepsilon - P_j^{(p)} f_\varepsilon\|_p \lesssim \varepsilon + \lambda^{-j\alpha} |f_\varepsilon|_{B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n)} \lesssim 2\varepsilon, \end{aligned}$$

if j is large enough.

□

5 The first theorem and its applications

We have now all the tools to prove the main results in this paper. Throughout this section we assume that the scaling function $\varphi \in B^{\alpha_0 \boldsymbol{a}}_{p,q}(\mathbb{R}^n)$ for some $\alpha_0 > 0$, and $V_0^{(p)}$ contains locally the polynomials of degree $< L$ for a fixed integer $L \geq 1$. Recall from §1 that $E_j^{(p)}(f) := \inf_{g \in V_j^{(p)}} \|f - g\|_p$. Then we prove the following restatement of Theorem 1.4.

Theorem 5.1 *Let $0 < p, q < \infty$. For every $0 < \alpha < \alpha^* := \min\{\alpha_0, \frac{L}{a_1}, \ldots, \frac{L}{a_n}\}$ we have:*

$$\begin{aligned} B^{\alpha \boldsymbol{a}}_{p,q}(\mathbb{R}^n) &= \Big\{ f \in L^p(\mathbb{R}^n) \;:\; \sum_{j=0}^{\infty} \lambda^{\alpha j q} \|Q_j^{(p)} f\|_p^q < \infty \Big\} \\ &= \Big\{ f \in L^p(\mathbb{R}^n) \;:\; \sum_{j=0}^{\infty} \lambda^{\alpha j q} E_j^{(p)}(f)^q < \infty \Big\}. \end{aligned}$$

In this case, we have the equivalence of quasi-seminorms:

$$|f|_{B^{\alpha \boldsymbol{a}}_{p,q}} \sim \Big[\sum_{j \in \mathbb{Z}} (\lambda^{\alpha j} \|Q_j^{(p)} f\|_p)^q \Big]^{\frac{1}{q}} \sim \Big[\sum_{j \in \mathbb{Z}} (\lambda^{\alpha j} E_j^{(p)}(f))^q \Big]^{\frac{1}{q}}.$$

Proof: Through the proof we denote $\mu = \min\{1, p\}$. We shall only show the equivalence involving the $E_j^{(p)}(f)$'s, the other following easily from

$$\|Q_j^{(p)} f\|_p^\mu \leq E_j^{(p)}(f)^\mu + E_{j+1}^{(p)}(f)^\mu \leq 2 \sum_{\ell \geq j} \|Q_\ell^{(p)} f\|_p^\mu, \quad j \in \mathbb{Z},$$

and inequalities of Hardy type (see, e.g., [11, Th. 2.1]). For the inclusion "$\subseteq$", the key step is the estimate

$$E_j^{(p)}(f) \leq \|f - P_j^{(p)} f\|_{L^p(\mathbb{R}^n)} \lesssim \sum_{|\boldsymbol{\ell}| = L} \omega_p^{\boldsymbol{\ell}}(f, c\lambda^{-j}), \quad j \in \mathbb{Z},$$

which we have seen in the proof of Jackson's inequality (see (4.10)). From here, and using the discrete definition of Besov seminorm in (2.3), we conclude easily

$$\Big[\sum_{j \in \mathbb{Z}} (\lambda^{\alpha j} E_j^{(p)}(f))^q \Big]^{\frac{1}{q}} \lesssim |f|_{B^{\alpha \boldsymbol{a}}_{p,q}},$$

For the converse "$\supseteq$", it suffices to show that for any $f \in L^p(\mathbb{R}^n)$ and $J \in \mathbb{N}$

$$\omega_p^{\boldsymbol{\ell}}(f, \lambda^{-J}) \lesssim \lambda^{-J\alpha_0} \Big[\sum_{j=-\infty}^{J} (\lambda^{\alpha_0 j} E_j^{(p)}(f))^\mu \Big]^{\frac{1}{\mu}}, \quad |\boldsymbol{\ell}| = L. \tag{5.2}$$

Indeed, again an argument involving Hardy's inequality will give (for $0 < \alpha < \alpha_0$):

$$\Big[\sum_{j\in\mathbb{Z}}(\lambda^{\alpha j}\omega_p^{\boldsymbol{\ell}}(f,c\lambda^{-J}))^q\Big]^{\frac{1}{q}} \lesssim \Big[\sum_{j\in\mathbb{Z}}(\lambda^{\alpha j}E_j^{(p)}(f))^q\Big]^{\frac{1}{q}}, \tag{5.3}$$

from which we shall conclude the theorem.

To show (5.2), we first write for any pair of integers $J' < J$:

$$\Delta_h^{\boldsymbol{\ell}} f = \Delta_h^{\boldsymbol{\ell}}(f - P_J^{(p)} f) + \Delta_h^{\boldsymbol{\ell}}(P_{J'}^{(p)} f) + \sum_{j=J'}^{J-1} \Delta_h^{\boldsymbol{\ell}}(Q_j^{(p)} f).$$

Since by Bernstein's inequality $\lim_{J'\to-\infty} |P_{J'}^{(p)} f|_{B^{\alpha}_{p,q}} = 0$, we can take pointwise limits and use Fatou's Lemma to obtain

$$\|\Delta_h^{\boldsymbol{\ell}} f\|_p^\mu \leq \|\Delta_h^{\boldsymbol{\ell}}(f - P_J^{(p)} f)\|_p^\mu + \sum_{j=-\infty}^{J-1} \|\Delta_h^{\boldsymbol{\ell}}(Q_j^{(p)} f)\|_p^\mu, \qquad a.e.\, h \in \mathbb{R}^n.$$

Then, the same reasoning as in the proof of Bernstein's inequality gives

$$\begin{aligned} \|\Delta_h^{\boldsymbol{\ell}} f\|_p^\mu &\lesssim \|f - P_J^{(p)} f\|_p^\mu + \sum_{j=-\infty}^{J-1} \|Q_j^{(p)} f\|_p^\mu \, \|\Delta_h^{\boldsymbol{\ell}} \varphi_{j+1,0}\|_p^\mu \\ \text{by (3.9)} \quad &\lesssim E_J^{(p)}(f)^\mu + \sum_{j=-\infty}^{J} E_j^{(p)}(f)^\mu \, \omega_p^{\boldsymbol{\ell}}(\varphi, \lambda^{-(J-j)})^\mu, \quad \text{if } |h|_{\boldsymbol{a}} \leq \lambda^{-J}. \end{aligned}$$

Now using (2.5) one can control the modulus of continuity of φ by:

$$\omega_p^{\boldsymbol{\ell}}(\varphi, \lambda^{-(J-j)}) \lesssim \lambda^{-(J-j)\alpha_0} \, |\varphi|_{B^{\alpha_0 \boldsymbol{a}}_{p,q}}, \quad j \in \mathbb{Z}.$$

Substituting into the sum above and taking the sup over all $|h|_{\boldsymbol{a}} \leq \lambda^{-J}$ we arrive at (5.2). This completes the proof of the theorem.

□

Our first corollary gives a rigorous justification of the multilevel decomposition in (1.3) for functions in $B^{\alpha\boldsymbol{a}}_{p,q}(\mathbb{R}^n)$.

Corollary 5.4 *In the conditions of the theorem, if $f \in B^{\alpha\boldsymbol{a}}_{p,q}(\mathbb{R}^n)$ then we can write $f = P_0^{(p)} f + \sum_{j=0}^{\infty} Q_j^{(p)} f$ with convergence in the norm of $B^{\alpha\boldsymbol{a}}_{p,q}$.*

Proof: By Corollary 4.12, we only have to verify that

$$\lim_{j\to\infty} |f - P_j^{(p)} f|_{B^{\alpha\boldsymbol{a}}_{p,q}} = 0, \quad f \in B^{\alpha\boldsymbol{a}}_{p,q}(\mathbb{R}^n).$$

Now, using the elementary fact $E^{(p)}_\ell(f - P^{(p)}_j f) \lesssim E^{(p)}_{\max(j,\ell)}(f)$, it follows from the characterization in Theorem 5.1

$$\begin{aligned} |f - P^{(p)}_j f|^q_{B^{\alpha\boldsymbol{a}}_{p,q}} &\lesssim \sum_{\ell\in\mathbb{Z}} (\lambda^{\ell\alpha} E^{(p)}_\ell (f - P^{(p)}_j f))^q \\ &\lesssim \sum_{\ell\geq j} (\lambda^{\ell\alpha} E^{(p)}_\ell (f))^q \to 0, \quad \text{as } j \to \infty. \end{aligned}$$

□

Theorem 5.1 can also be written as $B^{\alpha\boldsymbol{a}}_{p,q}(\mathbb{R}^n) = A^\alpha_q(L^p(\mathbb{R}^n))$, the latter denoting the space of L^p-approximation obtained from the approximating family $\{V^{(p)}_j\}^\infty_{j=0}$ (in the notation of [11]). This equality and the results in [11, 10] give us the following interpolating properties of anisotropic Besov spaces.

Corollary 5.5 *Let $0 < p, q < \infty$, $\alpha > 0$ and $\boldsymbol{\alpha} = \alpha\boldsymbol{a}$ as in §3. Then, for the real interpolation method we have the identity:*

$$\Big(L^p(\mathbb{R}^n), B^{\boldsymbol{\alpha}}_{p,r}(\mathbb{R}^n)\Big)_{\theta,q} = B^{\theta\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^n), \qquad \forall\, 0 < \theta < 1 \quad \textit{and} \quad 0 < r < \infty.$$

Moreover, if $0 < \alpha_0 < \alpha_1 < \infty$ and we let $\alpha_\theta = (1-\theta)\alpha_0 + \theta\alpha_1$, then we also have

$$\Big(B^{\alpha_0\boldsymbol{a}}_{p,q_0}, B^{\alpha_1\boldsymbol{a}}_{p,q_1}\Big)_{\theta,q} = B^{\alpha_\theta\boldsymbol{a}}_{p,q} \qquad \forall\, 0 < \theta < 1 \quad \textit{and} \quad 0 < q < \infty.$$

Proof: This is a direct consequence of the identity $B^{\alpha\boldsymbol{a}}_{p,q}(\mathbb{R}^n) = A^\alpha_q(L^p(\mathbb{R}^n))$, and Theorems 3.1 and 4.2 in [11] for the approximations spaces $A^\alpha_q(X)$.

□

We conclude with an embedding result announced in the introduction.

Corollary 5.6 *Let $0 < p, q < \infty$, $\alpha > 0$ and*

$$\frac{1}{\tau} = \frac{\alpha}{n} + \frac{1}{p}. \tag{5.7}$$

Then, $B^{\alpha\boldsymbol{a}}_{\tau,p}(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$ and

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{B^{\alpha\boldsymbol{a}}_{\tau,p}(\mathbb{R}^n)}, \quad \forall\, f \in B^{\alpha\boldsymbol{a}}_{\tau,p}(\mathbb{R}^n). \tag{5.8}$$

Proof: The proof relies in a sort of Bernstein inequality:

$$\|f\|_{L^p(\mathbb{R}^n)} \leq C\,\lambda^{j\alpha}\,\|f\|_{L^\tau(\mathbb{R}^n)}, \quad \forall\, f \in V^{(\tau)}_j. \tag{5.9}$$

Indeed, assume for the moment (5.9) and write every $f \in B^{\alpha\boldsymbol{a}}_{\tau,p}(\mathbb{R}^n)$ as $f = P_0^{(\tau)} f + \sum_{j=0}^{\infty} Q_j^{(\tau)} f$. Then, Fatou's lemma and Theorem 5.1 (with $\mu = \min\{1, p\}$) give

$$\begin{aligned} \|f\|_{L^p(\mathbb{R}^n)} &\leq \Big[\|P_0^{(\tau)} f\|_p^\mu + \sum_{j=0}^{\infty} \|Q_j^{(\tau)} f\|_p^\mu \Big]^{\frac{1}{\mu}} \\ &\lesssim \Big[\|P_0^{(\tau)} f\|_\tau^\mu + \sum_{j=0}^{\infty} (\lambda^{j\alpha} \|Q_j^{(\tau)} f\|_\tau)^\mu \Big]^{\frac{1}{\mu}} \lesssim \|f\|_{B^{\alpha\boldsymbol{a}}_{\tau,\mu}}. \end{aligned}$$

Now, if $\mu = p \leq 1$ this is the desired result. If $p > 1$, one must choose $1 \leq p_0 < p < p_1 < \infty$ so that

$$\|f\|_{p_i} \leq C\, \|f\|_{B^{\alpha_i \boldsymbol{a}}_{\tau,1}}, \quad i = 0, 1,$$

where α_i, p_i, τ are related as in (5.7). Using the real interpolation formula in Corollary 5.5 (and interpolation of L^p spaces) one establishes (5.8).

It remains to show (5.9), which by scaling it suffices to do for $j = 0$. Let $f = \sum_k c_k \varphi_k$, for a finite sequence of scalars $\{c_k\}_k$. Then, the same argument we used in (3.3) involving reverse Hölder inequalities gives

$$\|f\|^p_{L^p(\mathbb{R}^n)} \leq \sum_{\gamma \in \mathbb{Z}^n} \int_{I_\gamma} |\sum_k c_k \varphi_k|^p \, dx \lesssim \sum_{\gamma \in \mathbb{Z}^n} \left(\int_{I_\gamma} |\sum_k c_k \varphi_k|^\tau \, dx \right)^{\frac{p}{\tau}} \lesssim \|f\|^p_{L^\tau(\mathbb{R}^n)},$$

where in the last step we must use the inclusion $\ell^1 \hookrightarrow \ell^{\frac{p}{\tau}}$ (since $\tau < p$).

□

Remark 5.10 We shall also need below a local form of the previous embedding: *there is a finite set $K_\# \subset \mathbb{Z}^n$ such that, for every $\gamma \in \mathbb{Z}^n$, we have*

$$\|f\|_{L^p(I_\gamma)} \lesssim \|f\|_{B^{\alpha\boldsymbol{a}}_p(L^\tau(\tilde{I}_\gamma))}, \quad \forall\ f \in B^{\alpha\boldsymbol{a}}_{\tau,p}(\mathbb{R}^n), \tag{5.11}$$

where $\tilde{I}_\gamma = \cup_{m \in K_\#} (I_\gamma + m)$. In the above expression the norm of $B^{\alpha\boldsymbol{a}}_q(X)$ when X is a quasi-normed space, must be interpreted as follows:

$$\|f\|_{B^{\alpha\boldsymbol{a}}_q(X)} = \|f\|_X + \sum_{|\boldsymbol{\ell}|=L} \left[\int_{\mathbb{R}^n} \left(|h|_{\boldsymbol{a}}^{-\alpha} \|\Delta_h^{\boldsymbol{\ell}} f\|_X \right)^q \frac{dh}{|h|^n_{\boldsymbol{a}}} \right]^{1/q}.$$

To obtain (5.11) one can apply the previous corollary to the function Φf, where Φ is smooth and satisfies $\chi_{I_\gamma} \leq \Phi \leq \chi_{2I_\gamma}$. For a sharper version of this result, directly applied to Besov spaces over rectangles, see [15, Th. 5.2].

6 The second theorem

In this section we prove Theorem 1.6, which gives a wavelet decomposition for the spaces $B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)$ when $\alpha > n(\frac{1}{\tau}-1)_+$. Under this assumption, for any $0<\tau<\infty$ we can define the real number $p(\alpha,\tau)>1$ by the identity (5.7), i.e.

$$p(\alpha,\tau) = \left(\frac{1}{\tau}-\frac{\alpha}{n}\right)^{-1} \quad \text{if } \alpha < \tfrac{n}{\tau}, \quad \text{and} \quad p(\alpha,\tau)=\infty \text{ otherwise.} \tag{6.1}$$

It follows from Corollary 5.6, and the trivial embedding $B^{\alpha\boldsymbol{a}}_{\tau,q} \hookrightarrow B^{\alpha_1\boldsymbol{a}}_{\tau,q}$ for $\alpha>\alpha_1$, that $B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$ for every finite p such that $\tau \le p \le p(\alpha,\tau)$. In particular, for those p's which are also >1, every $f\in B^{\alpha\boldsymbol{a}}_{\tau,\tau}$ can be decomposed as:

$$f = \sum_{k\in\mathbb{Z}^n}\langle f,\tilde\varphi_k\rangle\varphi_k + \sum_{\ell=1}^{m-1}\sum_{j=0}^{\infty}[\sum_{k\in\mathbb{Z}^n}\langle f,\tilde\psi^{(p')}_{\ell;j,k}\rangle\psi^{(p)}_{\ell;j,k}] = \sum_{\ell=1}^{m-1}\sum_{j\in\mathbb{Z}}[\sum_{k\in\mathbb{Z}^n}\langle f,\tilde\psi^{(p')}_{\ell;j,k}\rangle\psi^{(p)}_{\ell;j,k}],$$

with convergence of the series at least in $L^p(\mathbb{R}^n)$. Recall that the generating wavelets used in this paper $\psi_\ell,\tilde\psi_\ell$ are compactly supported, since they arise from scaling functions $\varphi,\tilde\varphi$ which are assumed to satisfy this property (see references in [13, §3.2]).

Theorem 1.6 characterizes, among all functions with such representation, those having wavelet coefficients in ℓ^τ. We state below a stronger form of this theorem more akin to the notation in §5, and with a few less restrictions in the indices. As usual, we shall assume that the scaling function $\varphi \in B^{\alpha_0\boldsymbol{a}}_{\tau,\tau}$, and that the space V_0 contains locally the polynomials of degree $<L$.

Theorem 6.2 *Let $0<\tau<\infty$, $n(\frac{1}{\tau}-1)_+<\alpha<\alpha^*$ and $p(\alpha,\tau)$ given by (6.1). Then, for any $p\in(1,\infty)$ such that $\tau\le p\le p(\alpha,\tau)$ we have*

$$B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n) = \Big\{ f\in L^p(\mathbb{R}^n) \;:\; \|P^{(p)}_0 f\|^\tau_\tau + \sum_{j=0}^{\infty}\lambda^{j\tau\alpha}\|Q^{(p)}_j f\|^\tau_\tau < \infty \Big\}$$

In this case, we have the equivalence of quasi-norms:

$$\|f\|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \sim \|P^{(p)}_0 f\|_\tau + [\sum_{j=0}^{\infty}\lambda^{j\tau\alpha}\|Q^{(p)}_j f\|^\tau_\tau]^{\frac{1}{\tau}} \quad \textit{and} \quad |f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \sim \Big[\sum_{j\in\mathbb{Z}}\lambda^{j\tau\alpha}\|Q^{(p)}_j f\|^\tau_\tau\Big]^{\frac{1}{\tau}}.$$

Proof: We first prove the direct inclusion "$\subseteq$". The case $\tau\ge1$ is just a reformulation of Theorem 5.1, since in this case the projectors are linear and $P^{(p)}_j = P^{(1)}_j = P^{(\tau)}_j$. We shall assume, therefore, $\tau<1$, and fix a finite $1<p\le p(\alpha,\tau)$.

Let $f\in B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$. We first claim that $P^{(p)}_0 f \in V^{(\tau)}_0$. Indeed, writing $P^{(p)}_0 f = \sum_k\langle f,\tilde\varphi_k\rangle\varphi_k$, and using the embedding in (5.11) we have

$$|\langle f,\tilde\varphi_k\rangle| \le \Big[\int_{\mathrm{Supp}\tilde\varphi_k}|f|^p\,dx\Big]^{\frac{1}{p}}\|\tilde\varphi\|_{p'} \lesssim \|f\|_{B^{\alpha\boldsymbol{a}}_\tau(L^\tau(J_k))}, \tag{6.3}$$

where $J_k = J + k$ and $J = \text{Supp}\ \tilde{\varphi}$ is a fixed compact set. Then,

$$\sum_{k\in\mathbb{Z}^n} |\langle f, \tilde{\varphi}_k\rangle|^\tau \lesssim \sum_{k\in\mathbb{Z}^n} \|f\|^\tau_{B^{\alpha\boldsymbol{a}}_\tau(L^\tau(J_k))} \lesssim \|f\|^\tau_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)},$$

which by the τ-stability of $\{\varphi_k\}$ implies that $P_0^{(p)} f \in V_0^{(\tau)}$ and

$$\|P_0^{(p)} f\|_\tau \sim \Big[\sum_{k\in\mathbb{Z}^n} |\langle f, \tilde{\varphi}_k\rangle|^\tau\Big]^{\frac{1}{\tau}} \lesssim \|f\|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}}.$$

Next, we show that

$$\Big[\sum_{j\in\mathbb{Z}}\sum_{\ell=1}^{m-1}\sum_{k\in\mathbb{Z}^n} \lambda^{\tau j(\alpha+\frac{n}{2}-\frac{n}{\tau})}\, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau\Big]^{\frac{1}{\tau}} \lesssim |f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}}, \quad \forall\ f \in B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n). \tag{6.4}$$

Since $Q_j^{(p)} f = \sum_{\ell,k} \langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle\, \psi^{(2)}_{\ell;j,k}$ and $\tau < 1$, observe that we also have

$$\sum_{j\in\mathbb{Z}} \lambda^{\alpha j\tau} \|Q_j^{(p)} f\|^\tau_\tau \le \sum_{\ell,j,k} \lambda^{\alpha j\tau}\, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau\, \|\psi^{(2)}_{\ell;j,k}\|^\tau_\tau = \sum_{\ell,j,k} \lambda^{\tau j(\alpha+\frac{n}{2}-\frac{n}{\tau})}\, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau\, \|\psi_\ell\|^\tau_\tau,$$

so that (6.4) will suffice for the direct inclusion of the theorem.

To prove (6.4) we shall use the following estimate, valid for all $\alpha_1 < \alpha^*$ such that $n(\frac{1}{\tau}-1)_+ < \alpha_1 < \frac{n}{\tau}$:

$$\Big[\sum_{\ell,k} \lambda^{\tau j(\frac{n}{2}-\frac{n}{\tau})}\, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau\Big]^{\frac{1}{\tau}} \le C\, \lambda^{-j\alpha_1}\, |f|_{B^{\alpha_1\boldsymbol{a}}_{\tau,\tau}}, \quad j \in \mathbb{Z}. \tag{6.5}$$

Assume for the moment that (6.5) holds and choose $\alpha_1 < \alpha$. Since $\langle \varphi_{j,k'}, \tilde{\psi}_{\ell;j,k}\rangle = 0$ for all ℓ, j, k, k', by definition of $P_j^{(\tau)}$ we must also have

$$\langle P_j^{(\tau)} f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle = 0, \quad \forall\ k \in \mathbb{Z}^n,\ \ j \in \mathbb{Z}, \quad \forall f \in L^\tau(\mathbb{R}^n). \tag{6.6}$$

Combining (6.5) with Theorem 5.1 and using the fact $E_r^{(\tau)}(f - P_j^{(\tau)} f) \lesssim E^{(\tau)}_{\max(r,j)}(f)$, we obtain

$$\begin{aligned}
\sum_{\ell,k} \lambda^{\tau j(\alpha+\frac{n}{2}-\frac{n}{\tau})}\, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau &= \sum_{\ell,k} \lambda^{\tau j(\alpha+\frac{n}{2}-\frac{n}{\tau})}\, |\langle f - P_j^{(\tau)} f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau \\
&\lesssim \lambda^{j\tau(\alpha-\alpha_1)}\, |f - P_j^{(\tau)} f|^\tau_{B^{\alpha_1\boldsymbol{a}}_{\tau,\tau}} \\
&\lesssim \lambda^{j\tau(\alpha-\alpha_1)} \sum_{r=j}^{\infty} \lambda^{r\tau\alpha_1} E_r^{(\tau)}(f)^\tau.
\end{aligned}$$

Thus, summing this estimate in j and using one more time Theorem 5.1 we will have established (6.4):

$$\begin{aligned}\sum_{\ell,j,k} \lambda^{\tau j(\alpha+\frac{n}{2}-\frac{n}{\tau})}\, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau &\lesssim \sum_{r\in\mathbb{Z}} \lambda^{r\tau\alpha_1} E_r^{(\tau)}(f)^\tau \Big(\sum_{j=-\infty}^{r} \lambda^{j\tau(\alpha-\alpha_1)}\Big) \\ &\lesssim \sum_{r\in\mathbb{Z}} \lambda^{r\tau\alpha} E_r^{(\tau)}(f)^\tau \sim |f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}}.\end{aligned}$$

Let us now show (6.5), which by homogeneity we just need to verify for $j=0$. We let $p_1 = p(\alpha_1,\tau) = (\frac{1}{\tau}-\frac{\alpha_1}{n})^{-1}$, which by assumption is a real number in $(1,\infty)$. The proof will be a little more subtle than (6.3) above, since we want seminorms on the right hand side. Observe first that by the local polynomial condition on V_0 the functions $\tilde{\psi}_\ell$ have vanishing moments up to degree L (see e.g. (6.6)). Thus proceeding as in (6.3), we see that for every polynomial Q of degree $< L$

$$\begin{aligned}|\langle f, \tilde{\psi}_{\ell;k}\rangle| &= \Big|\int_{\mathrm{Supp}\tilde{\psi}_{\ell;k}} (f-Q)\, \overline{\tilde{\psi}_{\ell;k}}\, dx\Big| \le \|\tilde{\psi}_\ell\|_{p_1'}\, \|f-Q\|_{L^{p_1}(\mathrm{Supp}\tilde{\psi}_{\ell;k})} \\ &\lesssim \|f-Q\|_{L^\tau(J_k)} + |f-Q|_{B^{\alpha_1\boldsymbol{a}}_\tau(L^\tau(J_k))}.\end{aligned}$$

The difference operator $\Delta_h^{\boldsymbol{\ell}}$ with $|\boldsymbol{\ell}| = L$, annihilates such polynomials, so the second term equals $|f-Q|_{B^{\alpha_1\boldsymbol{a}}_\tau(L^\tau(J_k))} = |f|_{B^{\alpha_1\boldsymbol{a}}_\tau(L^\tau(J_k))}$. For the first term we can choose $Q = Q^k$ as in Whitney's inequality, so that it easily follows:

$$|\langle f, \tilde{\psi}_{\ell;k}\rangle| \lesssim |f|_{B^{\alpha_1\boldsymbol{a}}_\tau(L^\tau(J_k))}.$$

Raising to the power τ and summing in $k\in\mathbb{Z}^n$, we end up with (6.5).

Let us now show the inverse inclusion of the theorem "$\supseteq$". By Theorem 5.1, the case $\tau \ge 1$ only requires to verify that functions on the right hand set belong to $L^\tau(\mathbb{R}^n)$. But this is immediate by Minkowski's inequality since

$$\|f\|_\tau \le \|P_0^{(p)} f\|_\tau + \sum_{j=0}^{\infty} \|Q_j^{(p)} f\|_\tau \lesssim \|P_0^{(p)} f\|_\tau + \Big[\sum_{j=0}^{\infty} \lambda^{j\tau\alpha} \|Q_j^{(p)} f\|_\tau^\tau\Big]^{\frac{1}{\tau}},$$

where the last step follows by Hölder's (since $\tau \ge 1$).

We are thus left with the case $\tau < 1$. By the assumptions on f (and $\lambda > 1$) we also see that

$$\|f\|_\tau^\tau \le \|P_0^{(p)} f\|_\tau^\tau + \sum_{j=0}^{\infty} \|Q_j^{(p)} f\|_\tau^\tau \le \|P_0^{(p)} f\|_\tau^\tau + \sum_{j=0}^{\infty} \lambda^{j\tau\alpha} \|Q_j^{(p)} f\|_\tau^\tau < \infty. \tag{6.7}$$

Next we shall establish a Bernstein type estimate:

$$|g|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \lesssim \lambda^{j\alpha}\, \|g\|_\tau, \quad \forall\, g \in V_j^{(p)}. \tag{6.8}$$

From here we can conclude easily the theorem since

$$|f|^\tau_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \leq |P^{(p)}_0 f|^\tau_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} + \sum_{j=0}^{\infty} |Q^{(p)}_j f|^\tau_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \lesssim \|P^{(p)}_0 f\|^\tau_\tau + \sum_{j=0}^{\infty} \lambda^{j\tau\alpha} \|Q^{(p)}_j f\|^\tau_\tau. \tag{6.9}$$

Moreover, combining (6.4), (6.5) and (6.8) it is not difficult to obtain as well the norm equivalences

$$\|Q^{(p)}_j f\|_\tau \sim \Big[\sum_{\ell,k} \lambda^{\tau j(\frac{n}{2}-\frac{n}{\tau})} \, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau\Big]^{\frac{1}{\tau}} \quad \text{and} \quad \Big[\sum_{j\in\mathbb{Z}} \lambda^{j\tau\alpha} \|Q^{(p)}_j f\|^\tau_\tau\Big]^{\frac{1}{\tau}} \sim |f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}}.$$

To show (6.8) we shall assume $j=0$ and proceed as in Bernstein's inequality:

$$|g|^\tau_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)} \leq |\varphi|^\tau_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}} \sum_{\gamma\in\mathbb{Z}^n} |\langle g, \tilde{\varphi}_\gamma\rangle|^\tau \lesssim \sum_{\gamma\in\mathbb{Z}^n} \int_{\mathrm{Supp}\tilde{\varphi}_\gamma} |g|^\tau \, dx \, \lesssim \|g\|^\tau_{L^\tau(\mathbb{R}^n)},$$

where in the second step we have used the reverse Hölder inequality in $V^{(p)}_0|_{\mathrm{Supp}\ \tilde{\varphi}_\gamma}$. The proof of the theorem is now complete.

□

As a simple corollary we obtain a slight sharpening of (5.8).

Corollary 6.10 *Let $0<\tau\leq 1$, $n(\frac{1}{\tau}-1)<\alpha<\frac{n}{\tau}$ and $p=p(\alpha,\tau)$ given by (6.1). Then,*

$$\|f\|_{L^p(\mathbb{R}^n)} \lesssim |f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)}, \quad \forall\, f\in B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n).$$

Proof: Choose very smooth scaling functions so that $\alpha^*>\alpha$. Since by assumption $f\in B^{\alpha\boldsymbol{a}}_{\tau,\tau}\hookrightarrow L^p$ and $p\in(1,\infty)$, we can write

$$f \;=\; \sum_{j\in\mathbb{Z}} Q^{(p)}_j f \;=\; \sum_{j\in\mathbb{Z}} \Big[\sum_{\ell=1}^{m-1} \sum_{k\in\mathbb{Z}^n} \langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle \psi^{(2)}_{\ell;j,k}\Big]$$

with convergence of the series at least in L^p. Then, using Minkowski's inequality, relation (6.1) and the embedding $\ell^\tau\hookrightarrow\ell^1$ we obtain

$$\|f\|_p \leq \sum_{j\in\mathbb{Z}} \|Q^{(p)}_j f\|_p \lesssim \sum_{\ell,j,k} \lambda^{j(\frac{n}{2}-\frac{n}{p})} \, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle| \lesssim \Big[\sum_{\ell,j,k} \lambda^{\tau j(\alpha+\frac{n}{2}-\frac{n}{\tau})} \, |\langle f, \tilde{\psi}^{(2)}_{\ell;j,k}\rangle|^\tau\Big]^{\frac{1}{\tau}}.$$

By Theorem 6.2, when $f\in B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)$ this last expression is equivalent to $|f|_{B^{\alpha\boldsymbol{a}}_{\tau,\tau}(\mathbb{R}^n)}$, establishing the corollary.

□

7 Final comments and further remarks

We conclude the paper with several remarks about limitations and possible extensions of the results presented above.

7.1 More general wavelet bases

The use of compactly supported wavelets and scaling functions is a typical requirement in problems arising from numerical applications. In our results above this assumption was strongly used, in particular when dealing with the non-linear projectors $P_j^{(p)}$. This is actually a quite common procedure in approximation theory (see e.g. [4, 6, 7]), although it is still fair to ask what happens with wavelet bases having more general decay conditions. We shall not pursue this question here, but we stress that a main difference appears when $p < 1$ and $\alpha < n(\frac{1}{p} - 1)$, since the Besov spaces introduced above *are not distribution spaces*. There is a rich literature on Besov spaces defined from Littlewood-Paley theory (specially in the isotropic setting; see e.g.[20, 2, 18]), but the techniques used in such cases seem too restrictive to be applied here. We think this is an interesting question left to future research.

7.2 Other dilation matrices

As it was pointed out above, the theory of multiresolution analyses has only been developed for dilation matrices M with integer entries. This produces some limitations if one wishes to study anisotropic spaces, since for instance, diagonal integer matrices of the form in (1.1) will only exist when $(\frac{1}{a_1}, \ldots, \frac{1}{a_n}) \in \mathbb{R}_+ \log \mathbb{Z}_+^n$ (see [13, §3.3]). In applications this does not seem a serious restriction, since one can at least cover all the cases $\boldsymbol{a} \in \mathbb{Q}_+^n$. However, if one wishes to analyze more general anisotropic spaces, then the traditional MRA setting must be abandoned.

An alternative approach for function spaces in the cube $[0, 1]^2$ was developed by the second author in [15]. If $\boldsymbol{a} = (a_1, a_2)$ is a general anisotropy in $\mathbb{R}^2$ with $a_1 \geq a_2 > 0$ and $\{V_j\}_{j=0}^{\infty}$ is *1-dimensional dyadic MRA in* $L^2(\mathbb{R})$, then one can consider an adapted subcollection of 2-dimensional tensor-product spaces:

$$\mathcal{V}_{j,\boldsymbol{a}} := V_j \otimes V_{[ja_1/a_2]}, \quad j = 0, 1, 2, \ldots,$$

where $[s]$ denotes the integer part of s. Roughly speaking, the factor a_1/a_2 gives the proportion of "extra details" we need to consider in the x_2-variable everytime we change scale in the x_1-variable. This produces a mutilevel decomposition based on $\{\mathcal{V}_{j,\boldsymbol{a}}\}_{j=0}^{\infty}$ which is well adapted to study anisotropic spaces $B_q^{\alpha\boldsymbol{a}}(L^p[0, 1]^2)$. A similar approach has also been developed more recently in [19].

7.3 Non-linear approximation spaces

Our last remark concerns the characterization of anisotropic Besov spaces as *non-linear approximation spaces.* For this one defines a new approximating family Σ_N, consisting of all $S = \sum_k c_k \varphi_k + \sum_{j=0}^{\infty} \sum_{\ell,k} d_{\ell;jk} \psi_{\ell;j,k}$ with *at most N non-null entries.* If we denote the error of approximation by $\sigma_N(f)_p = \inf_{S \in \Sigma_N} \|f - S\|_p$, then, similarly to the isotropic setting [8], one would like to show

$$B^{\alpha \boldsymbol{a}}_{\tau,\tau} = A^{\alpha}_{\tau}(L^p; \Sigma_N) := \{f \in L^p \ : \ \sum_{j=0}^{\infty} 2^{j\alpha\tau} \sigma_{2^j}(f)^{\tau}_p < \infty\}, \quad \text{if} \quad p = p(\alpha, \tau).$$

Such result (more subtle than the ones in this paper) has been investigated independently in [15] and [19], both with the multilevel decomposition for $[0,1]^2$ described in §7.2. A characterization in the general MRA setting of this paper will be presented separately in [12].

References

[1] M.Z. Berkolaiko and I.Ya. Novikov, "Unconditional bases in spaces of functions of anisotropic smoothness". Proc. Steklov Inst. Math. **204** (1994), 27–41.

[2] G. Bourdaud, "Ondelettes et espaces de Besov", Rev. Mat. Iberoamericana, **11** (1995), 477–512.

[3] C. Canuto and A. Tabacco, "Multilevel decompositions of functional spaces". J. Fourier Anal. and Appl., **3** (1997), 715–742.

[4] A. Cohen, *Wavelet methods in numerical analysis.* In *Handbook for Numerical Analysis*, P.G. Ciarlet and J.L. Lions (eds), Elsevier North Holland, Amsterdam (2002).

[5] A. Cohen, K. Gröchenig and L. Villemoes, "Regularity of multivariate refinable functions". Constr. Approx., **15** (1999), 241–155.

[6] W. Dahmen, *Wavelet and multiscale methods for operator equations.* In *Acta Numerica* **6**, Cambridge University Press (1997), 55-228.

[7] R.A. DeVore, *Nonlinear Approximation.* In *Acta Numerica* **7**, Cambridge University Press (1998), 51-151.

[8] R.A. DeVore, B. Jawerth and V.A. Popov, "Compression of wavelet decompositions". Amer. Jour. Math., **114** (1992), 737–785.

[9] R.A. DeVore and Lorentz, *Constructive Approximation.* Springer-Verlag (1993).

[10] R.A. DeVore and V.A. Popov, "Interpolation of Besov spaces". Trans. Amer. Math. Soc., **305** (1988), 397–414.

[11] R.A. DeVore and V.A. Popov, "Interpolation spaces and non-linear approximation". In *Function spaces and applications*, M. Cwikel et al. eds. Springer Lecture Notes in Math **1302** (1988), 191–205.

[12] G. Garrigós and E. Hernández, "Sharp Jackson and Bernstein inequalities for N-term approximation in sequence spaces with applications". Preprint, 2003. Available at *http://www.uam.es/∼gustavo.garrigos*.

[13] G. Garrigós and A. Tabacco, "Wavelet decompositions of anisotropic Besov spaces". Math. Nachr. **239-240** (2002), 80-102.

[14] R. Hochmuth, "Anisotropic Besov spaces and wavelets", in *Proceedings of the International Wavelet Conference, Tanger*, INRIA report (1998).

[15] R. Hochmuth, "Wavelet characterization for anisotropic Besov spaces". Appl. Comp. Harm. Anal. **12** (2002), 179–208.

[16] R. Hochmuth, "Nonlinear anisotropic boundary problems - Regularity results and multiscale discretization". Nonlinear Anal. **46** (2001), no. 1, Ser. A: Theory Methods, 1–18.

[17] R.Q. Jia, "Approximation properties of multivariate wavelets". Math. Comp. **67** (1998), 647–665.

[18] G. Kyriazis, "Wavelet coefficients measuring smoothness in $H_p(\mathbb{R}^d)$". Appl. Comp. Harm. Anal. **3** (1996), 100–119.

[19] C. Leisner, "Nonlinear Wavelet approximation in anisotropic Besov spaces". Indiana Univ. Math. J. **52** (2003), no. 2, 437–455.

[20] Y. Meyer, *Ondelettes et opérateurs, I: Ondelettes*, Hermann, Paris, (1990)

[21] J. Peetre, "Remarques sur les espaces de Besov. Le cas $0 < p < 1$". C.R. Acad. Sc. Paris A-B **277** (1973), A947-A949.

[22] H. Schmeisser, Anisotropic spaces II, Math. Nachr., **79** (1977), 55–73.

[23] H. Schmeisser and H. Triebel, *Topics in Fourier Analysis and Function spaces*, John Wiley & Sons, 1987.

[24] É. A. Storozhenko and P. Osvald, "Jackson's theorem in the spaces $L^p(\mathbb{R}^k)$, $0 < p < 1$". Siberian Math. J. **19** (1978), 630–640.

Gustavo Garrigós
Departamento de Matemáticas C-XV
Universidad Autónoma de Madrid
Ciudad Universitaria Cantoblanco
28049 Madrid (SPAIN)
gustavo.garrigos@uam.es

Anita Tabacco
Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 Torino (ITALY)
tabacco@polito.it