Proc. London Math. Soc. (3) 97 (2008) 477–513 © 2008 London Mathematical Society
doi:10.1112/plms/pdn010

Decomposition of weighted Triebel–Lizorkin and Besov spaces on the ball

G. Kyriazis, P. Petrushev and Yuan Xu

Abstract

Weighted Triebel–Lizorkin and Besov spaces on the unit ball B^d in $\mathbb{R}^d$ with weights $w_\mu(x) = (1-|x|^2)^{\mu-1/2}$, $\mu \geqslant 0$, are introduced and explored. A decomposition scheme is developed in terms of almost exponentially localized polynomial elements (needlets) $\{\varphi_\xi\}$, $\{\psi_\xi\}$ and it is shown that the membership of a distribution to the weighted Triebel–Lizorkin or Besov spaces can be determined by the size of the needlet coefficients $\{\langle f, \varphi_\xi\rangle\}$ in appropriate sequence spaces.

1. *Introduction*

Localized bases and frames allow to decompose functions and distributions in terms of building blocks of simple nature and have numerous advantages over other means of representation. In particular, they enable one to encode smoothness and other norms in terms of the coefficients of the decompositions. Meyer's wavelets [**10**] and the φ-transform of Frazier and Jawerth [**5**–**7**] provide such building blocks for decomposition of Triebel–Lizorkin and Besov spaces in the classical case on $\mathbb{R}^d$.

The aim of this article is to develop similar tools for decomposition of weighted Triebel–Lizorkin and Besov spaces on the unit ball B^d in $\mathbb{R}^d$ $(d > 1)$ with weights

$$w_\mu(x) := (1-|x|^2)^{\mu-1/2}, \quad \mu \geqslant 0,$$

where $|x|$ is the Euclidean norm of $x \in B^d$. These include $L_p(B^d, w_\mu)$, the Hardy spaces $H_p(B^d, w_\mu)$, and weighted Sobolev spaces. For our purpose, we develop localized frames which can be viewed as an analog of the φ-transform of Frazier and Jawerth on B^d.

For the construction of our frame elements, we shall use orthogonal polynomials in the weighted space $L_2(w_\mu) := L_2(B^d, w_\mu)$. Denote by Π_n the space of all algebraic polynomials of degree n in d variables and by V_n the subspace of all polynomials of degree n which are orthogonal to lower-degree polynomials in $L_2(w_\mu)$. These are eigenspaces of the differential operator

$$D_\mu := -\Delta + \langle x, \nabla\rangle^2 + (2\mu + d - 1)\langle x, \nabla\rangle. \tag{1.1}$$

More precisely (see, for example, [**3**]),

$$D_\mu P = n(n + d + 2\mu - 1)P \quad \text{for } P \in V_n. \tag{1.2}$$

We have the orthogonal polynomial decomposition

$$L_2(w_\mu) = \bigoplus_{n=0}^{\infty} V_n, \quad V_n \subset \Pi_n. \tag{1.3}$$

Received 15 February 2007; revised 22 January 2008; published online 2 April 2008.

2000 *Mathematics Subject Classification* 41A25, 42B35, 42C15.

The second author has been supported by NSF grant DMS-0709046 and the third author by NSF grant DMS-0604056.

Note that $\dim V_n = \binom{n+d-1}{n} \sim n^{d-1}$. As is shown in [**19**] the orthogonal projector $\mathrm{Proj}_n : L_2(w_\mu) \mapsto V_n$ can be written as

$$(\mathrm{Proj}_n f)(x) = \int_{B^d} f(y)\mathsf{P}_n(x,y)w_\mu(y)dy, \tag{1.4}$$

where, for $\mu > 0$, the kernel $\mathsf{P}_n(x,y)$ has the representation

$$\mathsf{P}_n(x,y) = b_d^\mu b_1^{\mu-(1/2)} \frac{n+\lambda}{\lambda} \int_{-1}^{1} C_n^\lambda \left(\langle x,y\rangle + u\sqrt{1-|x|^2}\sqrt{1-|y|^2}\right)(1-u^2)^{\mu-1}du. \tag{1.5}$$

Here $\langle x,y\rangle$ is the Euclidean inner product in $\mathbb{R}^d$, C_n^λ is the nth degree Gegenbauer polynomial,

$$\lambda = \mu + \frac{d-1}{2}, \tag{1.6}$$

and the constants b_d^μ, $b_1^{\mu-(1/2)}$ are defined by $(b_d^\gamma)^{-1} := \int_{B^d}(1-|x|^2)^{\gamma-1/2}\,dx$. For a representation of $\mathsf{P}_n(x,y)$ in the limiting case $\mu = 0$; see [**19**, (3.8)] or [**14**, (4.2)].

Evidently,

$$K_n(x,y) := \sum_{j=0}^{n} \mathsf{P}_j(x,y) \tag{1.7}$$

is the kernel of the orthogonal projector of $L_2(w_\mu)$ onto the space $\bigoplus_{\nu=0}^n V_\nu$.

A key role in this study will play the fact (established in [**14**]) that if the coefficients on the right hand side in (1.7) are 'smoothed out' by sampling a compactly supported C^∞ function, then the resulting kernel has nearly exponential localization around the main diagonal $y = x$ in $B^d \times B^d$. More precisely, let

$$L_n(x,y) := \sum_{j=0}^{\infty} \widehat{a}\Big(\frac{j}{n}\Big)\mathsf{P}_j(x,y), \tag{1.8}$$

where the 'smoothing' function $\widehat{a}$ is admissible in the sense of the following definition.

Definition 1.1. A function $\widehat{a} \in C^\infty[0,\infty)$ is called admissible of type (a) if $\operatorname{supp}\widehat{a} \subset [0,2]$ and $\widehat{a}(t) = 1$ on $[0,1]$, and of type (b) if $\operatorname{supp}\widehat{a} \subset [1/2,2]$.

We introduce the distance

$$d(x,y) := \arccos\left\{\langle x,y\rangle + \sqrt{1-|x|^2}\sqrt{1-|y|^2}\right\} \quad \text{on } B^d \tag{1.9}$$

and set

$$W_\mu(n;x) := \left(\sqrt{1-|x|^2} + n^{-1}\right)^{2\mu}, \quad x \in B^d. \tag{1.10}$$

One of our main results in [**14**, Theorem 4.2] asserts that for any $k > 0$ there exists a constant $c_k > 0$ depending only on k, d, μ, and $\widehat{a}$ such that

$$|L_n(x,y)| \leqslant c_k \frac{n^d}{\sqrt{W_\mu(n;x)}\sqrt{W_\mu(n;y)}(1+n\,d(x,y))^k}, \quad x,y \in B^d. \tag{1.11}$$

The kernels L_n are our main ingredient in constructing *analysis* and *synthesis* needlet systems $\{\varphi_\xi\}_{\xi\in\mathcal{X}}$ and $\{\psi_\xi\}_{\xi\in\mathcal{X}}$ here, indexed by a multilevel set $\mathcal{X} = \bigcup_{j=0}^\infty \mathcal{X}_j$ (see § 3). This is a pair of dual frames with elements that have nearly exponential localization on B^d and provide representation of every distribution f on B^d:

$$f = \sum_{\xi\in\mathcal{X}} \langle f,\varphi_\xi\rangle\psi_\xi. \tag{1.12}$$

The superb localization of the frame elements prompted us to term them *needlets*.

Our main interest lies with distributions in the weighted Triebel–Lizorkin (F-spaces) and Besov spaces (B-spaces) on B^d. These spaces are naturally defined via orthogonal decompositions (see [**15**, **18**] for the general idea). To be specific, let

$$\Phi_0(x,y) := 1 \quad \text{and} \quad \Phi_j(x,y) := \sum_{\nu=0}^{\infty} \widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)\mathsf{P}_\nu(x,y), \;\; j \geqslant 1,$$

where $\mathsf{P}(\cdot,\cdot)$ is from (1.5) and $\widehat{a}$ is admissible of type (b) (see Definition 1.1) such that $|\widehat{a}| > 0$ on $[3/5, 5/3]$.

The F-space $F^{s\rho}_{pq}$ with $s, \rho \in \mathbb{R}$, $0 < p < \infty$, $0 < q \leqslant \infty$, is defined (§ 4) as the space of all distributions f on B^d such that

$$\|f\|_{F^{s\rho}_{pq}} := \Bigg\| \Bigg(\sum_{j=0}^{\infty} (2^{sj} W_\mu(2^j;\cdot)^{-\rho/d} |\Phi_j * f(\cdot)|)^q \Bigg)^{1/q} \Bigg\|_{L_p(w_\mu)} < \infty, \tag{1.13}$$

where $\Phi_j * f(x) := \langle f, \overline{\Phi(x,\cdot)}\rangle$ (see Definition 2.7).

The corresponding scales of weighted Besov spaces $B^{s\rho}_{pq}$ with $s, \rho \in \mathbb{R}$, $0 < p, q \leqslant \infty$, are defined (§ 5) via the (quasi-)norms

$$\|f\|_{B^{s\rho}_{pq}} := \Bigg(\sum_{j=0}^{\infty} \Big(2^{sj} \|W_\mu(2^j;\cdot)^{-\rho/d} \Phi_j * f(\cdot)\|_{L_p(w_\mu)} \Big)^q \Bigg)^{1/q}. \tag{1.14}$$

Unlike in the classical case on $\mathbb{R}^d$, we have introduced an additional parameter ρ, which allows considering different scales of Triebel–Lizorkin and Besov spaces. To us most natural are the spaces

$$F^s_{pq} := F^{ss}_{pq} \quad \text{and} \quad B^s_{pq} := B^{ss}_{pq}, \tag{1.15}$$

which embed correctly with respect to the smoothness parameter s (see § 4). A 'classical' choice would be to consider the spaces F^{s0}_{pq} and B^{s0}_{pq}, where the weight $W_\mu(2^j;\cdot)$ is excluded from (1.13) and (1.14). The introduction of the parameter ρ enables us to treat these spaces simultaneously.

One of the main results of this paper is the characterization of the F-spaces in terms of the size of the needlet coefficients in the decomposition (1.12), namely,

$$\|f\|_{F^{s\rho}_{pq}} \sim \Bigg\| \Bigg(\sum_{j=0}^{\infty} 2^{sjq} \sum_{\xi \in \mathcal{X}_j} |\langle f, \varphi_\xi\rangle| W_\mu(2^j;\xi)^{-\rho/d} |\psi_\xi(\cdot)|^q \Bigg)^{1/q} \Bigg\|_{L_p(w_\mu)}.$$

Similarly for the Besov spaces $B^{s\rho}_{pq}$ we have the characterization

$$\|f\|_{B^{s\rho}_{pq}} \sim \Bigg(\sum_{j=0}^{\infty} 2^{sjq} \Bigg[\sum_{\xi \in \mathcal{X}_j} \Big(W_\mu(2^j;\xi)^{-\rho/d} \|\langle f, \varphi_\xi\rangle \psi_\xi\|_{L_p(w_\mu)} \Big)^p \Bigg]^{q/p} \Bigg)^{1/q}.$$

Further, the weighted Besov spaces are applied to nonlinear n-term approximation from needlets on B^d.

This is a follow-up paper of [**14**], where the localization (1.11) is established and the construction and basic properties of a single system of needlets are given. Our development here is a part of a broader undertaking for needlet characterization of Triebel–Lizorkin and Besov spaces on nonclassical domains, including the multidimensional unit sphere [**11**, **12**], ball, and cube (interval [**9**, **13**]) with weights. The results in this paper generalize the results in the univariate case from [**9**] (with $\alpha = \beta$), where needlet characterizations of F- and B-spaces on the interval are obtained.

The organization of the paper is the following. In § 2 the needed results from [**14**] and some background material are given, including localized polynomial kernels, the maximal operator, distributions on B^d, and cubature formula on B^d. The definition and some basic properties of needlets are given in § 3. In § 4 the weighted Triebel–Lizorlin space on B^d are introduced and characterized via needlets, while the weighted Besov spaces are explored in § 5. In § 6 Besov spaces are applied to nonlinear n-term approximation from needlets; § 7 contains the proofs of various lemmas from previous sections.

Throughout the paper we use the following notation:

$$\|f\|_p := \left(\int_{B^d} |f(x)|^p w_\mu(x)dx\right)^{1/p}, \quad 0 < p < \infty, \qquad \|f\|_\infty := \text{ess} \sup_{x\in B^d} |f(x)|.$$

For a measurable set $E \subset B^d$, $|E|$ denotes the Lebesgue measure of E, $m(E) := \int_E w_\mu(x)dx$, $\mathbb{1}_E$ is the characteristic function of E, and $\tilde{\mathbb{1}}_E := m(E)^{-1/2}\mathbb{1}_E$ is the $L_2(w_\mu)$ normalized characteristic function of E. Positive constants are denoted by c, $c_1, c_*, \ldots$ and they may vary at every occurrence; $A \sim B$ means $c_1 A \leqslant B \leqslant c_2 A$.

2. *Preliminaries*

2.1. *Localized polynomial kernels on B^d*

The polynomial kernels $L_n(x, y)$ introduced in (1.8) will be our main vehicle in developing needlet systems. Here we give some additional properties of these kernels.

We have

$$\|L_n(x,\cdot)\|_p \leqslant c\Big(\frac{n^d}{W_\mu(n;x)}\Big)^{1-1/p}, \quad x \in B^d, \ 0 < p \leqslant \infty. \tag{2.1}$$

This estimate is an immediate consequence of (1.11) and the following lemma (see [**14**, Lemma 4.6]), which will be instrumental in several proofs below.

Lemma 2.1. *If $\sigma > d/p + 2\mu|1/p - 1/2|, \mu \geqslant 0, 0 < p < \infty$, then*

$$\int_{B^d} \frac{w_\mu(y)dy}{W_\mu(n;y)^{p/2}(1+nd(x,y))^{\sigma p}} \leqslant c\, n^{-d} W_\mu(n;x)^{1-p/2}. \tag{2.2}$$

We now establish a matching lower bound estimate.

Theorem 2.2. *Let $\widehat{a}$ be admissible and let $|\widehat{a}(t)| \geqslant c_* > 0$ for $t \in [3/5, 5/3]$. Then for $0 < p \leqslant \infty$ and $n \geqslant 2$*

$$\|L_n(x,\cdot)\|_p \geqslant c\Big(\frac{n^d}{W_\mu(n;x)}\Big)^{1-1/p}, \quad x \in B^d. \tag{2.3}$$

Here the constant $c > 0$ depends only on d, μ, p, and c_*.

The proof of this theorem is given in § 1.7.

The kernels $L_n(x, y)$ are in a sense Lip 1 functions in both variables with respect to the distance $d(\cdot,\cdot)$ from (1.9). Let $\xi, y \in B^d$ and $c^* > 0$, $n \geqslant 1$. Then for all $x, z \in B_\xi(c^* n^{-1})$ and an arbitrary k, we have

$$|L_n(x,y) - L_n(\xi,y)| \leqslant c_k \frac{n^{d+1} d(x,\xi)}{\sqrt{W_\mu(n;y)}\sqrt{W_\mu(n;z)}(1+nd(y,z))^k}, \tag{2.4}$$

where c_k depends only on k, μ, d, $\widehat{a}$, and c^* (see [**14**, Proposition 4.7]).

We shall also need the following inequality from [**14**, Lemma 4.1]:

$$\left|\sqrt{1-|x|^2} - \sqrt{1-|y|^2}\right| \leqslant \sqrt{2}\, d(x,y), \quad x,y \in B^d, \tag{2.5}$$

which yields

$$W_\mu(n;x) \leqslant 2^\mu W_\mu(n;y)(1+nd(x,y))^{2\mu}, \quad x,y \in B^d. \tag{2.6}$$

2.2. *Reproducing polynomial kernels and applications*

To simplify our notation we introduce the following nonstandard 'convolution'. For functions $\Phi : B^d \times B^d \to \mathbb{C}$ and $f : B^d \to \mathbb{C}$, we write

$$\Phi * f(x) := \int_{B^d} \Phi(x,y) f(y) w_\mu(y)\, dy. \tag{2.7}$$

We denote by $E_n(f)_p$ the best approximation of $f \in L_p(w_\mu)$ from Π_n, that is,

$$E_n(f)_p := \inf_{g\in\Pi_n} \|f-g\|_p. \tag{2.8}$$

Lemma 2.3. *Let L_n be the kernel from (1.8), with $\widehat{a}$ admissible of type* (a). *Then*

(i) *$L_n * g = g$ for $g \in \Pi_n$, that is, L_n is a reproducing kernel for Π_n, and*

(ii) *for any $f \in L_p(w_\mu)$, $1 \leqslant p \leqslant \infty$, we have $L_n * f \in \Pi_{2n}$,*

$$\|L_n * f\|_p \leqslant c\|f\|_p, \quad \text{and} \quad \|f - L_n * f\|_p \leqslant cE_n(f)_p. \tag{2.9}$$

This lemma follows readily by the definition of L_n (see also Definition 1.1) and (2.1) (see [**14**, Proposition 4.8]).

Lemma 2.3(i) and (2.1) are instrumental in relating weighted norms of polynomials.

Proposition 2.4. *For $0 < q \leqslant p \leqslant \infty$ and $g \in \Pi_n$, $n \geqslant 1$,*

$$\|g\|_p \leqslant cn^{(d+2\mu)(1/q-1/p)}\|g\|_q, \tag{2.10}$$

and for any $\gamma \in \mathbb{R}$

$$\|W_\mu(n;\cdot)^\gamma g(\cdot)\|_p \leqslant cn^{d(1/q-1/p)}\|W_\mu(n;\cdot)^{\gamma+1/p-1/q} g(\cdot)\|_q. \tag{2.11}$$

The proof of this proposition is quite similar to the proof of [**9**, Proposition 2.6]; for completeness it is given in § 7.1.

2.3. *Maximal operator*

We denote by $B_\xi(r)$ the ball centred at $\xi \in B^d$ of radius $r > 0$ with respect to the distance $d(\cdot,\cdot)$ on B^d, that is,

$$B_\xi(r) = \{x \in B^d : d(x,\xi) < r\}. \tag{2.12}$$

It is straightforward to show that (see [**14**, Lemma 5.3])

$$|B_\xi(r)| \sim r^d\sqrt{1-|\xi|^2} \tag{2.13}$$

and

$$m(B_\xi(r)) := \int_{B_\xi(r)} w_\mu(x)\,dx \sim r^d(r+\sqrt{1-|\xi|^2})^{2\mu} \sim r^d(r+d(\xi,\partial B^d))^{2\mu}, \tag{2.14}$$

where ∂B^d is the boundary of B^d, that is, the unit sphere in $\mathbb{R}^d$.

The maximal operator $\mathcal{M}_t$ $(t>0)$ is defined by

$$\mathcal{M}_t f(x) := \sup_{B\ni x}\left(\frac{1}{m(B)}\int_B |f(y)|^t w_\mu(y)\,dy\right)^{1/t}, \quad x\in B^d, \tag{2.15}$$

where the sup is over all the balls $B\subset B^d$ (with respect to $d(\cdot,\cdot)$) containing x.

It follows by (2.14) that the measure $m(E) := \int_E w_\mu(x)\,dx$ is a doubling measure on B^d, that is, for $\xi\in B^d$ and $0<r\leqslant\pi$

$$m(B_\xi(2r)) \leqslant c\,m(B_\xi(r)). \tag{2.16}$$

Consequently, the general theory of maximal operators applies and the Fefferman–Stein vector-valued maximal inequality is valid (see [**16**]). If $0<p<\infty, 0<q\leqslant\infty$, and $0<t<\min\{p,q\}$ then for any sequence of functions $\{f_\nu\}_\nu$ on B^d

$$\left\|\left(\sum_{\nu=1}^\infty |\mathcal{M}_t f_\nu(\cdot)|^q\right)^{1/q}\right\|_p \leqslant c\left\|\left(\sum_{\nu=1}^\infty |f_\nu(\cdot)|^q\right)^{1/q}\right\|_p. \tag{2.17}$$

We need to estimate $\mathcal{M}_t\mathbb{1}_B$ for an arbitrary ball $B\subset B^d$.

LEMMA 2.5. *Let $\xi\in B^d$ and $0<r\leqslant\pi$. Then for $x\in B^d$*

$$(\mathcal{M}_t\mathbb{1}_{B_\xi(r)})(x) \sim \left(1+\frac{d(\xi,x)}{r}\right)^{-d/t}\left(1+\frac{d(\xi,x)}{r+d(\xi,\partial B^d)}\right)^{-2\mu/t}, \tag{2.18}$$

and hence

$$c'\Big(1+\frac{d(\xi,x)}{r}\Big)^{-(2\mu+d)/t} \leqslant (\mathcal{M}_t\mathbb{1}_{B_\xi(r)})(x) \leqslant c\Big(1+\frac{d(\xi,x)}{r}\Big)^{-d/t}. \tag{2.19}$$

Here the constants depend only on d, μ, and t.

Proof. It is easy to see that

$$(\mathcal{M}_t\mathbb{1}_{B_\xi(r)})(x) = \sup_{B\ni x}\left(\frac{m(B\cap B_\xi(r))}{m(B)}\right)^{1/t}, \quad x\in B^d,$$

where the sup is taken over all the balls $B\subset B^d$ (with respect to $d(\cdot,\cdot)$) containing x. This immediately leads to $(\mathcal{M}_t\mathbb{1}_{B_\xi(r)})(x)\sim 1$ if $d(x,\xi)\leqslant 2r$, and hence (2.18) holds in this case.

Suppose that $d(\xi,x)>2r$. Then evidently

$$(\mathcal{M}_t\mathbb{1}_{B_\xi(r)})(x) \geqslant \left(\frac{m(B_\xi(r))}{m(B_\xi(d(x,\xi)))}\right)^{1/t}.$$

For the other direction, suppose that $B_z(r^*)\subset B^d$ is the smallest ball such that $x\in\overline{B_z(r^*)}$ and $\overline{B_z(r^*)}\cap\overline{B_\xi(r)}\neq\emptyset$. A simple application of the triangle inequality shows that $B_\xi(d(\xi,x))\subset B_z(5r^*)$. Thus using (2.16)

$$(\mathcal{M}_t\mathbb{1}_{B_\xi(r)})(x) \leqslant \left(\frac{m(B_\xi(r))}{m(B_z(r^*))}\right)^{1/t} \leqslant c\left(\frac{m(B_\xi(r))}{m(B_\xi(d(x,\xi)))}\right)^{1/t}.$$

Therefore, using (2.14)

$$(\mathcal{M}_t \mathbb{1}_{B_\xi(r)})(x) \sim \left(\frac{m(B_\xi(r))}{m(B_\xi(d(x,\xi)))}\right)^{1/t} \sim \left(\frac{r^d(r+d(\xi,\partial B^d))^{2\mu}}{d(x,\xi)^d(d(x,\xi)+d(\xi,\partial B^d))^{2\mu}}\right)^{1/t},$$

which implies (2.18) since $d(\xi,x) > 2r$. Estimate (2.19) is immediate from (2.18). □

2.4. *Distributions on B^d*

To define distributions on B^d, we shall use as test functions the set $\mathcal{D} := C^\infty(B^d)$ of all infinitely continuously differentiable complex valued functions on B^d such that

$$\|\phi\|_{W^k_\infty} := \sum_{|\alpha|\leqslant k} \|\partial^\alpha \phi\|_\infty < \infty \quad \text{for } k = 0,1,\dots. \tag{2.20}$$

We assume that the topology in $\mathcal{D}$ is defined by these norms.

Evidently all polynomials belong to $\mathcal{D}$. More importantly, the space $\mathcal{D}$ of test functions ϕ can be completely characterized by their orthogonal polynomial expansions. Denote

$$\mathcal{N}_k(\phi) := \sup_{n\geqslant 0} (n+1)^k \|\operatorname{Proj}_n \phi\|_2. \tag{2.21}$$

Lemma 2.6. (a) *$\phi \in \mathcal{D}$ if and only if $\|\operatorname{Proj}_n \phi\|_2 = \mathcal{O}(n^{-k})$ for all k.*

(b) *For each $\phi \in \mathcal{D}$, $\phi = \sum_{n=0}^\infty \operatorname{Proj}_n \phi$, where the convergence is in the topology of $\mathcal{D}$.*

(c) *The topology in $\mathcal{D}$ can be equivalently defined by the norms $\mathcal{N}_k(\cdot)$, $k = 0,1,\dots$.*

Proof. Let $\phi \in \mathcal{D}$. Assume that $Q_{n-1} \in \Pi_{n-1}$ $(n \geqslant 1)$ is the polynomial of best $L_2(w_\mu)$-approximation to ϕ, that is, $\|\phi - Q_{n-1}\|_2 = E_{n-1}(\phi)_2$. Since $\mathsf{P}_n(x,\cdot)$ is orthogonal to Π_{n-1},

$$|\operatorname{Proj}_n \phi(x)| = |\langle \phi, \mathsf{P}_n(x,\cdot)\rangle| = |\langle \phi - Q_{n-1}, \mathsf{P}_n(x,\cdot)\rangle| \leqslant E_{n-1}(\phi)_2 \mathsf{P}_n(x,x)^{1/2}.$$

By the Jackson type estimate from [20], for any $k \geqslant 1$,

$$E_n(\phi)_2 \leqslant c_k n^{-2k} \|D^k_\mu \phi\|_2 \leqslant c n^{-2k} \|D^k_\mu \phi\|_\infty \leqslant c n^{-2k} \sum_{|\alpha|\leqslant 2k} \|\partial^\alpha \phi\|_\infty = c n^{-2k} \|\phi\|_{W^{2k}_\infty}.$$

Here D_μ is the differential operator from (1.1). It is easy to see that

$$\|\mathsf{P}_n(x,x)^{1/2}\|_2^2 = \binom{n+d-1}{n} \sim n^{d-1}.$$

All of the above leads to

$$\|\operatorname{Proj}_n \phi\|_2 \leqslant c_k n^{-2k+(d-1)/2} \|\phi\|_{W^{2k}_\infty}, \quad n \geqslant 1, \text{ for any } k \geqslant 1.$$

Therefore, for any $m \geqslant 0$

$$\mathcal{N}_m(\phi) \leqslant c\|\phi\|_{W^{2k}_\infty} \quad \text{if } k \geqslant \frac{m}{2} + \frac{(d-1)}{4}.$$

In the other direction, by Markov's inequality (see [8]) and (2.10), it follows that

$$\|\partial^\alpha \operatorname{Proj}_n \phi\|_\infty \leqslant n^{2|\alpha|} \|\operatorname{Proj}_n \phi\|_\infty \leqslant c n^{2|\alpha|+d/2+\mu} \|\operatorname{Proj}_n \phi\|_2.$$

Consequently, if $\|\operatorname{Proj}_n \phi\|_2 = \mathcal{O}(n^{-k})$ for all k, then $\partial^\alpha \phi = \sum_{n=0}^\infty \partial^\alpha \operatorname{Proj}_n \phi$ for all multi-indices α with the series converging uniformly and

$$\|\phi\|_{W^k_\infty} \leqslant c \sum_{|\alpha|\leqslant k} \sum_{n=0}^\infty n^{2|\alpha|+d/2+\mu} \|\operatorname{Proj}_n \phi\|_2 \leqslant c \mathcal{N}_m(\phi), \quad m \geqslant 2k + d/2 + \mu + 2.$$

This completes the proof of the lemma. □

The space $\mathcal{D}' := \mathcal{D}'(B^d)$ of distributions on B^d is defined as the set of all continuous linear functionals on $\mathcal{D}$. The pairing of $f \in \mathcal{D}'$ and $\phi \in \mathcal{D}$ will be denoted by $\langle f, \phi \rangle := f(\overline{\phi})$, which will be shown to be consistent with the inner product $\langle f, g \rangle := \int_{B^d} f(x)\overline{g(x)} w_\mu(x)\, dx$ in $L_2(w_\mu)$.

We now extend the definition of the nonstandard 'convolution' from (2.7) to distributions.

Definition 2.7. Let $f \in \mathcal{D}'$ and assume that $\Phi : B^d \times B^d \mapsto \mathbb{C}$ is such that $\Phi(x, \cdot) \in \mathcal{D}$ for all $x \in B^d$. We define

$$(\Phi * f)(x) := \langle f, \overline{\Phi(x, \cdot)} \rangle,$$

where on the right hand side, f acts on $\overline{\Phi(x, y)}$ as a function of y.

For later use we next record some simple properties of this 'convolution'.

Lemma 2.8. (i) *If $f \in \mathcal{D}'$ and $\Phi(\cdot, \cdot) \in C^\infty(B^d \times B^d)$, then $\Phi * f \in \mathcal{D}$, and in particular $\mathsf{P}_n * f \in V_n$. We define* $\mathrm{Proj}_n f := \mathsf{P}_n * f$.

(ii) *If $f \in \mathcal{D}'$ and $\Phi(\cdot, \cdot) \in C^\infty(B^d \times B^d)$, then*

$$\langle \Phi * f, \phi \rangle = \langle f, \overline{\Phi} * \phi \rangle, \quad \phi \in \mathcal{D}.$$

(iii) *Let $\Phi(\cdot, \cdot), \Psi(\cdot, \cdot) \in C^\infty(B^d \times B^d)$, and $\Phi(x, y) = \Phi(y, x)$ and $\Psi(x, y) = \Psi(y, x)$ for $x, y \in B^d$. Then for any $f \in \mathcal{D}'$ and $x \in B^d$*

$$\Psi * \overline{\Phi} * f(x) = \langle \Psi(x, \cdot), \Phi(\cdot, \cdot) \rangle * f.$$

The proof of this lemma is standard and will be omitted.

We next give the representation of distributions from $\mathcal{D}'$ in terms of orthogonal polynomials on B^d.

Lemma 2.9. (a) *A linear functional $f \in \mathcal{D}'$ if and only if there exists a $k \geqslant 0$ such that*

$$|\langle f, \phi \rangle| \leqslant c_k \mathcal{N}_k(\phi) \quad \text{for all } \phi \in \mathcal{D}. \tag{2.22}$$

Hence, for $f \in \mathcal{D}'$ there exists a $k \geqslant 0$ such that

$$\| \mathrm{Proj}_n f \|_2 = \| \mathsf{P}_n * f \|_2 \leqslant c_k (n+1)^k, \quad n = 0, 1, \ldots. \tag{2.23}$$

(b) *Every $f \in \mathcal{D}'$ has the representation $f = \sum_{n=0}^\infty \mathrm{Proj}_n f$ in distributional sense, that is,*

$$\langle f, \phi \rangle = \sum_{n=0}^\infty \langle \mathrm{Proj}_n f, \phi \rangle = \sum_{n=0}^\infty \langle \mathrm{Proj}_n f, \mathrm{Proj}_n \phi \rangle \quad \text{for all } \phi \in \mathcal{D}, \tag{2.24}$$

where the series converges absolutely.

Proof. (a) This statement follows immediately by the fact that the topology in $\mathcal{D}$ can be defined by the norms $\mathcal{N}_k(\cdot)$ defined in (2.21).

(b) Using Lemma 2.6(b) we obtain for $\phi \in \mathcal{D}$,

$$\langle f, \phi \rangle = \lim_{N \to \infty} \left\langle f, \sum_{n=0}^N \mathrm{Proj}_n \phi \right\rangle = \lim_{N \to \infty} \sum_{n=0}^N \langle f, \mathrm{Proj}_n \phi \rangle = \sum_{n=0}^\infty \langle \mathrm{Proj}_n f, \mathrm{Proj}_n \phi \rangle,$$

where the last equality is justified by using (2.23) and the rapid decay of $\| \mathrm{Proj}_n \phi \|_2$. ☐

2.5. *Cubature formula and subdivision of B^d*

For the construction of our building blocks (needlets), we shall utilize the positive cubature formula given in [**14**]. This formula is based on almost equally distributed knots on B^d with respect to the distance $d(\cdot,\cdot)$.

Definition 2.10. We say that a set $\mathcal{X}_\varepsilon \subset B^d$, along with an associated partition $\mathcal{R}_\varepsilon$ of B^d consisting of measurable subsets of B^d, is a *set of almost uniformly ε-distributed points* on B^d if

(i) $B^d = \bigcup_{R\in\mathcal{R}_\varepsilon} R$ and the sets in $\mathcal{R}_\varepsilon$ do not overlap ($R_1^\circ \cap R_2^\circ = \emptyset$ if $R_1 \neq R_2$);

(ii) for each $R \in \mathcal{R}_\varepsilon$ there is a unique $\xi \in \mathcal{X}_\varepsilon$ such that $B_\xi(c^*\varepsilon) \subset R \subset B_\xi(\varepsilon)$.

Hence $\#\mathcal{X}_\varepsilon = \#\mathcal{R}_\varepsilon \leqslant c^{**}\varepsilon^{-d}$. Here the constant $c^* > 0$, depending only on d, is fixed but sufficiently small, such that the existence of sets of almost uniformly ε-distributed points on B^d is guaranteed (see the next lemma).

Lemma 2.11 ([**14**]). *For a sufficiently small constant $c^* > 0$, depending only on d, and an arbitrary $0 < \varepsilon \leqslant \pi$ there exists a set $\mathcal{X}_\varepsilon \subset B^d$ of almost uniformly ε-distributed points on B^d, where the associated partition $\mathcal{R}_\varepsilon$ of B^d consists of projections of spherical simplices.*

An important element in the construction of needlets will be the cubature formula given in [**14**, Corollary 5.10].

Proposition 2.12. *There exists a constant $c^\diamond > 0$ (depending only on d) and a sequence $\{\mathcal{X}_j\}_{j=0}^\infty$ of almost uniformly ε_j-distributed points on B^d with $\varepsilon_j := c^\diamond 2^{-j}$, and there exist positive coefficients $\{\lambda_\xi\}_{\xi\in\mathcal{X}_j}$ such that the cubature formula*

$$\int_{B^d} f(x)w_\mu(x)\,dx \sim \sum_{\xi\in\mathcal{X}_j} \lambda_\xi f(\xi) \tag{2.25}$$

is exact for all polynomials of degree at most 2^{j+2}. In addition,

$$\lambda_\xi \sim 2^{-jd}W_\mu(2^j;\xi) \sim m(B_\xi(2^{-j})) \tag{2.26}$$

with constants of equivalence depending only on μ and d.

It follows from above that

$$m(R_\xi) \sim 2^{-jd}W_\mu(2^j;\xi) \sim \lambda_\xi, \quad \xi \in \mathcal{X}_j, \tag{2.27}$$

while

$$|R_\xi| \sim 2^{-jd}(\sqrt{1-|\xi|^2} + 2^{-j}), \quad \xi \in \mathcal{X}_j. \tag{2.28}$$

3. *Localized building blocks (needlets) on B^d*

We utilize the ideas from [**9**, **12**] in constructing a pair of sequences of 'analysis' and 'synthesis' needlets on B^d. Let $\widehat{a}, \widehat{b}$ satisfy the conditions

$$\widehat{a}, \widehat{b} \in C^\infty[0,\infty), \quad \operatorname{supp} \widehat{a}, \widehat{b} \subset [1/2, 2], \tag{3.1}$$

$$|\widehat{a}(t)|, |\widehat{b}(t)| > c > 0, \quad \text{if } t \in [3/5, 5/3], \tag{3.2}$$

$$\overline{\widehat{a}(t)}\,\widehat{b}(t) + \overline{\widehat{a}(2t)}\,\widehat{b}(2t) = 1, \quad \text{if } t \in [1/2, 1]. \tag{3.3}$$

Hence,

$$\sum_{\nu=0}^{\infty} \overline{\widehat{a}(2^{-\nu}t)}\,\widehat{b}(2^{-\nu}t) = 1, \quad t \in [1,\infty). \tag{3.4}$$

It is easy to see that if $\widehat{a}$ satisfies (3.1) and (3.2), then there exists a $\widehat{b}$ satisfying (3.1) and (3.2) such that (3.3) is valid (see, for example, [**6**]).

Let $\widehat{a}$, $\widehat{b}$ satisfy (3.1)–(3.3). We define $\Phi_0(x,y) = \Psi_0(x,y) := 1$,

$$\Phi_j(x,y) := \sum_{\nu=0}^{\infty} \widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)\mathsf{P}_\nu(x,y), \quad j \geqslant 1, \tag{3.5}$$

$$\Psi_j(x,y) := \sum_{\nu=0}^{\infty} \widehat{b}\Big(\frac{\nu}{2^{j-1}}\Big)\mathsf{P}_\nu(x,y), \quad j \geqslant 1. \tag{3.6}$$

Assume that $\mathcal{X}_j$ is the set of knots and λ_ξ are the coefficients of the cubature formula (2.25). We define the jth level *needlets* by

$$\varphi_\xi(x) := \lambda_\xi^{1/2}\Phi_j(x,\xi) \quad \text{and} \quad \psi_\xi(x) := \lambda_\xi^{1/2}\Psi_j(x,\xi), \qquad \xi \in \mathcal{X}_j. \tag{3.7}$$

Notice that for $\xi \in \mathcal{X}_1$, we have $\varphi_\xi(x) = \widehat{a}(1)\mathsf{P}_1(x,\xi)$ and $\psi_\xi(x) = \widehat{b}(1)\mathsf{P}_1(x,\xi)$, but $\mathsf{P}_1(\cdot,\xi) \equiv 0$ if and only if $\xi = 0$. Therefore, to prevent $\varphi_\xi \equiv 0$ and $\psi_\xi \equiv 0$ for $\xi \in \mathcal{X}_1$, we assume that $0 \notin \mathcal{X}_1$.

We set $\mathcal{X} := \bigcup_{j=0}^{\infty} \mathcal{X}_j$, where equal points from different levels $\mathcal{X}_j$ are considered as distinct elements of $\mathcal{X}$, such that $\mathcal{X}$ can be used as an index set. We define the *analysis* and *synthesis* needlet systems Φ and Ψ by

$$\Phi := \{\varphi_\xi\}_{\xi\in\mathcal{X}}, \quad \Psi := \{\psi_\xi\}_{\xi\in\mathcal{X}}. \tag{3.8}$$

Estimate (1.11) yields the rapid decay of needlets, namely, for $x \in B^d$

$$|\Phi_j(\xi,x)|, |\Psi_j(\xi,x)| \leqslant \frac{c_k 2^{jd}}{\sqrt{W_\mu(2^j;\xi)}\sqrt{W_\mu(2^j;x)}(1+2^j d(\xi,x))^k} \quad \forall k, \tag{3.9}$$

and hence

$$|\varphi_\xi(x)|, |\psi_\xi(x)| \leqslant \frac{c_k 2^{jd/2}}{\sqrt{W_\mu(2^j;x)}(1+2^j d(\xi,x))^k} \quad \forall k. \tag{3.10}$$

Note that on account of (2.6) x in the term $\sqrt{W_\mu(2^j;x)}$ in (3.10) can be replaced by ξ.

The needlets are Lip 1 functions in the following sense. Let $\xi \in \mathcal{X}_j$, $j \geqslant 0$, $c^* > 0$, and $\omega \in B^d$. Then for each $x \in B_\omega(c^*2^{-j})$

$$|\varphi_\xi(x) - \varphi_\xi(\omega)|, |\psi_\xi(x) - \psi_\xi(\omega)| \leqslant \frac{c_k 2^{j(d/2+1)} d(\omega,x)}{\sqrt{W_\mu(2^j;\xi)}(1+2^j d(\xi,\omega))^k} \quad \forall k. \tag{3.11}$$

This estimate follows readily from (2.4).

We shall need estimates of the norms of the needlets. By (2.1), (2.3), and since $0 \notin \mathcal{X}_1$, we have for $0 < p \leqslant \infty$,

$$\|\varphi_\xi\|_p \sim \|\psi_\xi\|_p \sim \|\tilde{\mathbb{1}}_{R_\xi}\|_p \sim \Big(\frac{2^{jd}}{W_\mu(2^j;\xi)}\Big)^{1/2-1/p}, \quad \xi \in \mathcal{X}_j. \tag{3.12}$$

Furthermore, there exist constants $c^*, c > 0$ such that

$$\|\varphi_\xi\|_{L_\infty(B_\xi(c^*2^{-j}))},\ \|\psi_\xi\|_{L_\infty(B_\xi(c^*2^{-j}))} \geqslant c\Big(\frac{2^{jd}}{W_\mu(2^j;\xi)}\Big)^{1/2}, \quad \xi \in \mathcal{X}_j. \tag{3.13}$$

The proof of (3.13) is given in § 7. Notice that if $\widehat{a}$, $\widehat{b}$ are real valued, then Lemma 7.1 below yields

$$|\varphi_\xi(\xi)|, |\psi_\xi(\xi)| \geqslant c\left(\frac{2^{jd}}{W_\mu(2^j;\xi)}\right)^{1/2}, \quad \xi \in \mathcal{X}_j.$$

Our first step in implementing needlets is to establish needlet decompositions of $\mathcal{D}'$ and $L_p(w_\mu)$.

PROPOSITION 3.1. (a) *For any* $f \in \mathcal{D}'$,

$$f = \sum_{j=0}^{\infty} \Psi_j * \overline{\Phi}_j * f \quad \text{in } \mathcal{D}' \tag{3.14}$$

and

$$f = \sum_{\xi\in\mathcal{X}} \langle f, \varphi_\xi\rangle \psi_\xi \quad \text{in } \mathcal{D}'. \tag{3.15}$$

(b) *For* $f \in L_p(w_\mu)$, $1 \leqslant p \leqslant \infty$, *(3.14) and (3.15) hold in* $L_p(w_\mu)$. *Moreover, if* $1 < p < \infty$, *then the convergence in (3.14) and (3.15) is unconditional.*

Proof. By Definition 2.7 and (3.5) we have, for $f \in \mathcal{D}'$,

$$\overline{\Phi} * f = \sum_{\nu=0}^{2^j} \overline{\widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)} \mathsf{P}_\nu * f \tag{3.16}$$

and using Lemma 2.8 and that $\mathsf{P}_\nu * \mathsf{P}_\nu(\cdot, y) = \mathsf{P}_\nu(\cdot, y)$

$$\Psi * \overline{\Phi} * f = \sum_{\nu=0}^{2^j} \overline{\widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)}\widehat{b}\Big(\frac{\nu}{2^{j-1}}\Big) \mathsf{P}_\nu * f. \tag{3.17}$$

Then (3.14) follows from the above, (3.4), and Lemma 2.9.

Note that $\Psi_j(x, y)\overline{\Phi(y, z)}$ belongs to $\Pi_{2^{j+1}-1}$ as a function of y and, therefore, employing the cubature formula from Proposition 2.12 we obtain

$$\Psi_j * \overline{\Phi_j(\cdot, z)} = \int_{B^d} \Psi_j(x, y)\overline{\Phi(y, z)} w_\mu(y) dy = \sum_{\xi\in\mathcal{X}_j} \lambda_\xi \Psi_j(x, \xi)\overline{\Phi(\xi, z)} = \sum_{\xi\in\mathcal{X}_j} \psi_\xi(x)\overline{\varphi_\xi(z)},$$

which leads to

$$\Psi_j * \overline{\Phi}_j * f = \sum_{\xi\in\mathcal{X}_j} \langle f, \varphi_\xi\rangle \psi_\xi.$$

Combining this with (3.14) yields (3.15).

The convergence of (3.14) and (3.15) in $L_p(w_\mu)$ for $f \in L_p(w_\mu)$ follows in a similar fashion (see also [**9**, Proposion 3.1]). The unconditional convergence in $L_p(w_\mu)$, $1 < p < \infty$, follows by Theorem 4.5 and Proposition 4.12 below. □

4. *Weighted Triebel–Lizorkin spaces on* B^d

Following the general idea of using spectral decompositions (see, for example, [**15**, **18**]), we next employ orthogonal polynomials to introduce weighted Triebel–Lizorkin spaces on B^d. To this end, we define a sequence of kernels $\{\Phi_j\}$ by

$$\Phi_0(x,y) := 1 \quad \text{and} \quad \Phi_j(x,y) := \sum_{\nu=0}^{\infty} \widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)\mathsf{P}_\nu(x,y), \;\; j \geqslant 1, \tag{4.1}$$

where $\{\mathsf{P}_\nu(x,y)\}$ are from (1.4) and (1.5) and $\widehat{a}$ obeys the conditions

$$\widehat{a} \in C^\infty[0,\infty), \quad \operatorname{supp} \widehat{a} \subset [1/2, 2], \tag{4.2}$$

$$|\widehat{a}(t)| > c > 0 \quad \text{if } t \in [3/5, 5/3]. \tag{4.3}$$

Definition 4.1. Let $s, \rho \in \mathbb{R}$, $0 < p < \infty$, and $0 < q \leqslant \infty$. Then the weighted Triebel–Lizorkin space $F_{pq}^{s\rho} := F_{pq}^{s\rho}(w_\mu)$ is defined as the set of all $f \in \mathcal{D}'$ such that

$$\|f\|_{F_{pq}^{s\rho}} := \left\| \left(\sum_{j=0}^{\infty} \left[2^{sj} W_\mu(2^j;\cdot)^{-\rho/d} |\Phi_j * f(\cdot)| \right]^q \right)^{1/q} \right\|_p < \infty \tag{4.4}$$

with the usual modification when $q = \infty$.

Observe that the above definition is independent of the choice of $\widehat{a}$ as long as $\widehat{a}$ satisfy (4.2) and (4.3) (see Theorem 4.5 below).

Proposition 4.2. *For all $s, \rho \in \mathbb{R}$, $0 < p < \infty$, and $0 < q \leqslant \infty$, $F_{pq}^{s\rho}$ is a quasi-Banach space which is continuously embedded in $\mathcal{D}'$.*

Proof. The completeness of the space $F_{pq}^{s\rho}$ follows easily (see, for example, [**18**, p. 49]) by the continuous embedding of $F_{pq}^{s\rho}$ in $\mathcal{D}'$, which we establish next.

Let $\{\Phi_j\}$ be the kernels from the definition of $F_{pq}^{s\rho}$ with $\widehat{a}$ obeying (4.2) and (4.3) that are the same as (3.1) and (3.2). As already indicated there exists a function $\widehat{b}$ satisfying (3.1) and (3.3). We use this function to define $\{\Psi_j\}$ as in (3.6). Then by Proposition 3.1 $f = \sum_{j=0}^{\infty} \Psi_j * \Phi_j * f$ in $\mathcal{D}'$, and hence

$$\langle f, \phi\rangle = \sum_{j=0}^{\infty} \langle \Psi_j * \overline{\Phi}_j * f, \phi\rangle, \quad \phi \in \mathcal{D}.$$

We now employ (3.16) and (3.17) and the Cauchy–Schwarz inequality to obtain, for $j \geqslant 2$,

$$\begin{aligned} |\langle \Psi_j * \overline{\Phi}_j * f, \phi\rangle|^2 &= \left| \sum_{\nu=2^{j-2}+1}^{2^j} \overline{\widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)}\widehat{b}\Big(\frac{\nu}{2^{j-1}}\Big) \langle \operatorname{Proj}_\nu f, \operatorname{Proj}_\nu \phi\rangle \right|^2 \\ &\leqslant \sum_{\nu=2^{j-2}+1}^{2^j} \left|\widehat{a}\Big(\frac{\nu}{2^{j-1}}\Big)\right|^2 \|\operatorname{Proj}_\nu f\|_2^2 \sum_{\nu=2^{j-2}+1}^{2^j} \left|\widehat{b}\Big(\frac{\nu}{2^{j-1}}\Big)\right|^2 \|\operatorname{Proj}_\nu \phi\|_2^2 \\ &\leqslant 2^j \|\Phi_j * f\|_2^2 \max_{2^{j-2} < \nu \leqslant 2^j} \|\operatorname{Proj}_\nu \phi\|_2^2. \end{aligned}$$

Using inequality (2.10) we obtain

$$\|\Phi_j * f\|_2 \leqslant c 2^{j(d+2\mu)/p} \|\Phi_j * f\|_p \leqslant c 2^{j((d+2\mu)/p + 2\mu|\rho|/d - s)} \|2^{sj} W_\mu(2^j;\cdot)^{-\rho/d} \Phi_j * f(\cdot)\|_p.$$

From the above estimates we infer

$$|\langle \Psi_j * \overline{\Phi}_j * f, \phi\rangle| \leqslant c 2^{-j} \|f\|_{F_{pq}^{s\rho}} 2^{jk} \max_{2^{j-2} < \nu \leqslant 2^j} \|\operatorname{Proj}_\nu f\|_2 \leqslant c 2^{-j} \|f\|_{F_{pq}^{s\rho}} \mathcal{N}_k(\phi)$$

for $k \geqslant (d+2\mu)/p + 2\mu|\rho|/d + 3/2 - s$. A similar estimate trivially holds for $j = 0, 1$. Summing up we obtain

$$|\langle f, \phi\rangle| \leqslant c\|f\|_{F^{s\rho}_{pq}}\mathcal{N}_k(\phi),$$

which completes the proof. □

As a companion to $F^{s\rho}_{pq}$, we now introduce the sequence spaces $f^{s\rho}_{pq}$. Here we assume that $\{\mathcal{X}_j\}_{j=0}^{\infty}$ is a sequence of almost uniformly ε_j-distributed points on B^d ($\varepsilon_j := c^\diamond 2^{-j}$) with associated neighborhoods $\{R_\xi\}_{\xi\in\mathcal{X}_j}$, given by Proposition 2.12. Just as in the definition of needlets in § 3, we set $\mathcal{X} := \bigcup_{j\geqslant 0}\mathcal{X}_j$.

Definition 4.3. Suppose that $s, \rho \in \mathbb{R}$, $0 < p < \infty$, and $0 < q \leqslant \infty$. Then $f^{s\rho}_{pq}$ is defined as the space of all complex-valued sequences $h := \{h_\xi\}_{\xi\in\mathcal{X}}$ such that

$$\|h\|_{f^{s\rho}_{pq}} := \left\| \left(\sum_{j=0}^{\infty} 2^{sjq} \sum_{\xi\in\mathcal{X}_j} [|h_\xi| W_\mu(2^j;\xi)^{-\rho/d}\tilde{\mathbb{1}}_{R_\xi}(\cdot)]^q \right)^{1/q} \right\|_p < \infty \tag{4.5}$$

with the usual modification for $q = \infty$. Recall that $\tilde{\mathbb{1}}_{R_\xi} := m(R_\xi)^{-1/2}\mathbb{1}_{R_\xi}$.

In analogy to the classical case on $\mathbb{R}^d$ we introduce 'analysis' and 'synthesis' operators by

$$S_\varphi : f \longrightarrow \{\langle f, \varphi_\xi\rangle\}_{\xi\in\mathcal{X}} \quad \text{and} \quad T_\psi : \{h_\xi\}_{\xi\in\mathcal{X}} \longrightarrow \sum_{\xi\in\mathcal{X}} h_\xi\psi_\xi. \tag{4.6}$$

We next show that the operator T_ψ is well defined on $f^{s\rho}_{pq}$.

Lemma 4.4. *Let $s, \rho \in \mathbb{R}$, $0 < p < \infty$, and $0 < q \leqslant \infty$. Then for any $h \in f^{s\rho}_{pq}$, $T_\psi h := \sum_{\xi\in\mathcal{X}} h_\xi\psi_\xi$ converges in $\mathcal{D}'$. Moreover, the operator $T_\psi : f^{s\rho}_{pq} \to \mathcal{D}'$ is continuous, that is, there exist constants $k > 0$ and $c > 0$ such that*

$$|\langle T_\psi h, \phi\rangle| \leqslant c\mathcal{N}_k(\phi)\|h\|_{f^{s\rho}_{pq}}, \quad h \in f^{s\rho}_{pq}, \ \ \phi \in \mathcal{D}. \tag{4.7}$$

Proof. Let $h \in f^{s\rho}_{pq}$. Then by the definition of $f^{s\rho}_{pq}$ it follows that

$$2^{js}|h_\xi| W_\mu(2^j;\xi)^{-\rho/d}\|\tilde{\mathbb{1}}_{R_\xi}(\cdot)\|_p \leqslant \|h\|_{f^{s\rho}_{pq}} \quad \text{for } \xi \in \mathcal{X}_j,\ j \geqslant 0.$$

Now, using (2.27),

$$\|\tilde{\mathbb{1}}_{R_\xi}\|_p = m(R_\xi)^{1/p-1/2} \sim [2^{-jd}W_\mu(2^j,\xi)]^{1/p-1/2} \quad \text{for } \xi \in \mathcal{X}_j$$

and since $2^{-2\mu j} \leqslant W_\mu(2^j,\xi) \leqslant 2^{2\mu}$ it follows that

$$|h_\xi| \leqslant c2^{-j(s+d(1/2-1/p))}W_\mu(2^j;\xi)^{\rho/d-1/p+1/2}\|h\|_{f^{s\rho}_{pq}} \leqslant c2^{j\gamma}\|h\|_{f^{s\rho}_{pq}}, \quad \xi \in \mathcal{X}_j, \tag{4.8}$$

where

$$\gamma := d|\frac{s}{d} - \frac{1}{p} + \frac{1}{2}| + 2\mu|\frac{\rho}{d} - \frac{1}{p} + \frac{1}{2}|.$$

On the other hand, by Lemma 2.6, $\phi = \sum_{n=0}^{\infty}\operatorname{Proj}_n\phi$ in $\mathcal{D}$ for $\phi \in \mathcal{D}$, and for $\xi \in \mathcal{X}_j$

$$\psi_\xi(x) := \lambda_\xi^{1/2}\Psi_j(x,\xi) = \lambda_\xi^{1/2}\sum_{2^{j-2}<\nu<2^j}\widehat{b}\Big(\frac{\nu}{2^{j-1}}\Big)\mathsf{P}_\nu(x,\xi), \quad \lambda_\xi \sim 2^{-jd}W_\mu(2^j,\xi).$$

Consequently,

$$\langle \psi_\xi, \phi\rangle = \lambda_\xi^{1/2} \sum_{2^{j-2}<\nu<2^j} \widehat{b}\Big(\frac{\nu}{2^{j-1}}\Big) \operatorname{Proj}_\nu \overline{\phi}$$

and hence

$$|\langle \psi_\xi, \phi\rangle| \leqslant c2^{-jd/2} W_\mu(2^j,\xi)^{1/2} \sum_{2^{j-2}<\nu<2^j} \|\operatorname{Proj}_\nu \phi\|_\infty.$$

Since $\operatorname{Proj}_\nu \phi \in \Pi_\nu$, by Proposition 2.4 $\|\operatorname{Proj}_\nu \phi\|_\infty \leqslant c\nu^{(d+2\mu)/2}\|\operatorname{Proj}_\nu \phi\|_2$. Therefore,

$$|\langle \psi_\xi, \phi\rangle| \leqslant c2^{j\mu} \sum_{2^{j-2}<\nu<2^j} \|\operatorname{Proj}_\nu \phi\|_2.$$

Combining this with (4.8) and using that $\#\mathcal{X}_j \leqslant c2^{jd}$, we obtain, for $\phi \in \mathcal{D}$

$$\begin{aligned}\sum_{\xi\in\mathcal{X}} |h_\xi||\langle \psi_\xi, \phi\rangle| &\leqslant \sum_{j=0}^\infty \sum_{\xi\in\mathcal{X}_j} |h_\xi||\langle \psi_\xi, \phi\rangle| \\ &\leqslant c\|h\|_{f_{pq}^{s\rho}} \sum_{j=0}^\infty (\#\mathcal{X}_j) 2^{j(\gamma+\mu)} \sum_{2^{j-2}<\nu<2^j} \|\operatorname{Proj}_\nu \phi\|_2 \qquad (4.9)\\ &\leqslant c\|h\|_{f_{pq}^{s\rho}} \sup_{\nu\geqslant 0}(\nu+1)^k \|\operatorname{Proj}_\nu \phi\|_2 \sum_{j=0}^\infty 2^{j(\gamma+\mu+d+1-k)} \\ &\leqslant c\|h\|_{f_{pq}^{s\rho}} \mathcal{N}_k(\phi),\end{aligned}$$

where $k := [\gamma]+\mu+d+4 > \gamma+\mu+d+3$, which makes the series above convergent. Consequently, the series $T_\psi h = \sum_{\xi\in\mathcal{X}} h_\xi\psi_\xi$ converges in $\mathcal{D}'$. We define $T_\psi h$ via

$$\langle T_\psi h, \phi\rangle := \sum_{\xi\in\mathcal{X}} h_\xi \langle \psi_\xi, \phi\rangle$$

for all $\phi \in \mathcal{D}$. Estimate (4.7) follows from (4.9). □

We now give our main result on weighted Triebel–Lizorkin spaces.

THEOREM 4.5. *Let $s, \rho \in \mathbb{R}$, $0 < p < \infty$, and $0 < q \leqslant \infty$. The operators $S_\varphi : F_{pq}^{s\rho} \to f_{pq}^{s\rho}$ and $T_\psi : f_{pq}^{s\rho} \to F_{pq}^{s\rho}$ are bounded and $T_\psi \circ S_\varphi = \mathrm{Id}$ on $F_{pq}^{s\rho}$. Consequently, $f \in F_{pq}^{s\rho}$ if and only if $\{\langle f, \varphi_\xi\rangle\}_{\xi\in\mathcal{X}} \in f_{pq}^{s\rho}$. Furthermore,*

$$\|f\|_{F_{pq}^{s\rho}} \sim \|\{\langle f,\varphi_\xi\rangle\}\|_{f_{pq}^{s\rho}} \sim \left\| \left(\sum_{j=0}^\infty 2^{sjq} \sum_{\xi\in\mathcal{X}_j} [|\langle f,\varphi_\xi\rangle| W_\mu(2^j;\xi)^{-\rho/d} |\psi_\xi(\cdot)|]^q \right)^{1/q} \right\|_p. \qquad (4.10)$$

In addition, the definition of $F_{pq}^{s\rho}$ is independent of the particular selection of $\widehat{a}$ satisfying (4.2) and (4.3).

The proof of this theorem relies on several lemmas and its proofs are given in § 7.2. In the following, we assume that $\{\Phi_j\}$ are from the definition of weighted Triebel–Lizorkin spaces, while $\{\varphi_\xi\}_{\xi\in\mathcal{X}}$ and $\{\psi_\xi\}_{\xi\in\mathcal{X}}$ are needlet systems defined as in (3.7) with no connection between the functions $\widehat{a}$s from (4.1) and (3.5).

LEMMA 4.6. *For any $k>0$ there exists a constant $c_k>0$ such that*

$$|\Phi_j * \psi_\xi(x)| \leqslant c_k \frac{2^{jd/2}}{\sqrt{W_\mu(2^j;x)}(1+2^j d(x,\xi))^k}, \quad \xi \in \mathcal{X}_\nu, \ \ j-1 \leqslant \nu \leqslant j+1, \tag{4.11}$$

and $\Phi_j * \psi_\xi \equiv 0$ for $\xi \in \mathcal{X}_\nu$ if $\nu \geqslant j+2$ or $\nu \leqslant j-2$. Here $\mathcal{X}_\nu := \emptyset$ if $\nu < 0$.

LEMMA 4.7. *For any $t>0$ and $\xi \in \mathcal{X}_j$, $j \geqslant 0$,*

$$|\varphi_\xi(x)|, |\psi_\xi(x)| \leqslant c(\mathcal{M}_t \tilde{\mathbb{1}}_{R_\xi})(x), \quad x \in B^d, \tag{4.12}$$

and

$$\tilde{\mathbb{1}}_{R_\xi}(x) \leqslant c(\mathcal{M}_t \varphi_\xi)(x), c(\mathcal{M}_t \psi_\xi)(x), \quad x \in B^d. \tag{4.13}$$

DEFINITION 4.8. For any set of complex numbers $\{h_\xi\}_{\xi \in \mathcal{X}_j}$ $(j \geqslant 0)$ we define

$$h_\xi^* := \sum_{\eta \in \mathcal{X}_j} \frac{|h_\eta|}{(1+2^j d(\eta,\xi))^\sigma}, \quad \xi \in \mathcal{X}_j, \tag{4.14}$$

where $\sigma > 1$ is a sufficiently large constant that will be selected later on.

LEMMA 4.9. *Let $P \in \Pi_{2^j}$, $j \geqslant 0$, and denote $a_\xi := \max_{x \in R_\xi} |P(x)|$ for $\xi \in \mathcal{X}_j$. There exists an $r \geqslant 1$, depending only on σ, μ, and d such that if*

$$b_\xi := \max\left\{ \min_{x \in R_\eta} |P(x)| : \eta \in \mathcal{X}_{j+r}, R_\xi \cap R_\eta \neq \emptyset \right\}, \quad \xi \in \mathcal{X}_j,$$

then

$$a_\xi^* \leqslant c b_\xi^* \tag{4.15}$$

with constant independent of P, j, and ξ.

LEMMA 4.10. *Assume that $t>0$, $\gamma \in \mathbb{R}$, and let $\{b_\xi\}_{\xi \in \mathcal{X}_j}$ $(j \geqslant 0)$ be a set of complex numbers. Also, let σ in the definition (4.14) of b_ξ^* obey $\sigma > d + (d+2\mu)/t + 2\mu|\gamma|$. Then for any $\xi \in \mathcal{X}_j$*

$$b_\xi^* W_\mu(2^j;\xi)^\gamma \mathbb{1}_{R_\xi}(x) \leqslant c\mathcal{M}_t \left(\sum_{\eta \in \mathcal{X}_j} |b_\eta| W_\mu(2^j;\eta)^\gamma \mathbb{1}_{R_\eta}(\cdot) \right)(x), \quad x \in R_\xi. \tag{4.16}$$

Proof of Theorem 4.5. Choose $0 < t < \min\{p,q\}$ and let σ in Definition 4.8 obey $\sigma > d + (d+2\mu)/t + 2\mu|\rho|/d$. Now, choose $k \geqslant \sigma + 2\mu|\rho|/d$. Observe first that the right-hand side equivalence in (4.10) follows immediately from Lemma 4.7 and the maximal inequality (2.17).

Let $\{\Phi_j\}$ be a sequences of kernels as in the definition of weighted Triebel–Lizorkin spaces, that is, Φ_j is defined by (4.1) with $\widehat{a}$ satisfying (4.2) and (4.3), the same as (3.1)–(3.2). As already mentioned, there exists a function $\widehat{b}$ satisfying (3.1) and (3.2) such that (3.3) holds. Let Ψ_j be defined by (3.6) with this $\widehat{b}$. In addition, let $\{\varphi_\xi\}_{\xi \in \mathcal{X}}$ and $\{\psi_\xi\}_{\xi \in \mathcal{X}}$ be the associated needlet systems defined as in (3.7) using these $\widehat{a}$ and $\widehat{b}$.

Exactly in the same way, let $\{\widetilde{\Phi}_j\}$ and $\{\widetilde{\Psi}_j\}$ be two sequences of kernels defined as above using completely different functions $\widehat{a}$ and $\widehat{b}$. Also, assume that $\{\widetilde{\varphi}_\xi\}$, $\{\widetilde{\psi}_\xi\}$ are the associated needlet systems, defined as in (3.5)–(3.7). As a result, we have two completely different systems of kernels and associated needlet systems.

Let us first prove the boundedness of the operator $T_{\widetilde{\psi}}: f_{pq}^{s\rho} \to F_{pq}^{s\rho}$, defined similarly as in (4.6) with $\{\psi_\xi\}$ replaced by $\{\widetilde{\psi}_\xi\}$. Here we assume that the space $F_{pq}^{s\rho}$ is defined by $\{\Phi_j\}$. Let $h \in f_{pq}^{s\rho}$ and denote $f := \sum_\xi h_\xi \widetilde{\psi}_\xi$, which is well defined according to Lemma 4.6. Using Lemma 4.6 we have, for $x \in B^d$,

$$|\Phi_j * f(x)| = \left| \sum_{\xi \in \mathcal{X}} h_\xi \Phi_j * \widetilde{\psi}_\xi(x) \right| \leqslant \sum_{j-1 \leqslant \nu \leqslant j+1} \sum_{\xi \in \mathcal{X}_\nu} |h_\xi| |\Phi_j * \widetilde{\psi}_\xi(x)|$$
$$\leqslant c 2^{jd/2} \sum_{j-1 \leqslant \nu \leqslant j+1} \sum_{\xi \in \mathcal{X}_\nu} \frac{|h_\xi|}{\sqrt{W_\mu(2^\nu; x)}(1 + 2^\nu d(\xi, x))^k}.$$

For $\eta \in \mathcal{X}_j$, denote $\Gamma_\eta := \{w \in \mathcal{X}_{j-1} \cup \mathcal{X}_j \cup \mathcal{X}_{j+1} : R_w \cap R_\eta \neq \emptyset\}$. Here $\mathcal{X}_{-1} := \emptyset$. First note that $\#\Gamma_\eta \leqslant c$. Secondly, for $x \in R_\eta$ and $w \in \Gamma_\eta$, we have $d(x, w) \leqslant c 2^{-j}$ and using inequality (2.6)

$$W_\mu(2^j; x)^{-\rho/d} \leqslant c W_\mu(2^j; w)^{-\rho/d} \leqslant c W_\mu(2^j; \xi)^{-\rho/d} (1 + 2^j d(\xi, \omega))^{2\mu|\rho|/d}.$$

We use the above estimates to obtain, for $x \in R_\eta$,

$$W_\mu(2^j; x)^{-\rho/d} |\Phi_j * f(x)|$$
$$\leqslant c 2^{jd/2} \sum_{j-1 \leqslant \nu \leqslant j+1} \sum_{\omega \in \Gamma_\eta \cap \mathcal{X}_\nu} \sum_{\xi \in \mathcal{X}_\nu} \frac{|h_\xi| W_\mu(2^j; \xi)^{-\rho/d} \mathbb{1}_{R_\omega}(x)}{\sqrt{W_\mu(2^\nu; \omega)}(1 + 2^\nu d(\xi, \omega))^{k - 2\mu|\rho|/d}}$$
$$\leqslant c 2^{jd/2} \sum_{\omega \in \Gamma_\eta} \frac{H_\omega^* \mathbb{1}_{R_\omega}(x)}{\sqrt{W_\mu(2^j; \omega)}} \leqslant c \sum_{\omega \in \Gamma_\eta} H_\omega^* \tilde{\mathbb{1}}_{R_\omega}(x),$$

where $H_\omega := h_\omega W_\mu(2^j; \omega)^{-\rho/d}$. Here we used that $k - 2\mu|\rho|/d \geqslant \sigma$ and (2.27). We insert the above in (4.4) and use Lemma 4.10 (with $\gamma = 0$) and the maximal inequality (2.17) to obtain

$$\begin{aligned} \|f\|_{F_{pq}^{s\rho}} &\leqslant c \left\| \left(\sum_{j=0}^\infty \left[2^{sj} \sum_{\eta \in \mathcal{X}_j} \sum_{\omega \in \Gamma_\eta} H_\omega^* \tilde{\mathbb{1}}_{R_\omega}(\cdot) \right]^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty \left[2^{sj} \sum_{\xi \in \mathcal{X}_j} H_\xi^* \tilde{\mathbb{1}}_{R_\xi}(\cdot) \right]^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty \left[\mathcal{M}_t \left(\sum_{\xi \in \mathcal{X}_j} 2^{sj} |H_\xi| \tilde{\mathbb{1}}_{R_\xi} \right)(\cdot) \right]^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty \left[\sum_{\xi \in \mathcal{X}_j} 2^{sj} |H_\xi| \tilde{\mathbb{1}}_{R_\xi}(\cdot) \right]^q \right)^{1/q} \right\|_p \leqslant c \|\{h_\xi\}\|_{f_{pq}^{s\rho}}, \end{aligned} \tag{4.17}$$

where in the second inequality above we used that $\#\Gamma_\eta \leqslant c$. Hence, the operator $T_{\widetilde{\psi}}: f_{pq}^{s\rho} \to F_{pq}^{s\rho}$ is bounded.

Assume now that the space $F_{pq}^{s\rho}$ is defined in terms of $\{\overline{\Phi}_j\}$ in place of $\{\Phi_j\}$. Using this definition we shall prove the boundedness of the operator $S_\varphi : F_{pq}^{s\rho} \to f_{pq}^{s\rho}$.

Let $f \in F_{pq}^{s\rho}$. Then $\overline{\Phi}_j * f \in \Pi_{2^j}$. For $\xi \in \mathcal{X}_j$, we define

$$a_\xi := \max_{x \in R_\xi} |\overline{\Phi}_j * f(x)|, \quad b_\xi := \max \left\{ \min_{x \in R_\eta} |\overline{\Phi}_j * f(x)| : \eta \in \mathcal{X}_{j+r}, R_\xi \cap R_\eta \neq \emptyset \right\},$$

where $r \geqslant 1$ is from Lemma 4.9. Then by the same lemma $a^*_\xi \sim b^*_\xi$. Therefore, using (2.27),

$$|\langle f, \varphi_\xi \rangle| = \lambda_\xi^{1/2} |\overline{\Phi}_j * f(\xi)| \leqslant cm(R_\xi)^{1/2} a_\xi \leqslant cm(R_\xi)^{1/2} a^*_\xi \leqslant cm(R_\xi)^{1/2} b^*_\xi.$$

From this, recalling that $\tilde{\mathbb{1}}_{R_\xi} := m(R_\xi)^{-1/2} \mathbb{1}_{R_\xi}$, we obtain

$$\begin{aligned} \|\{\langle f, \varphi_\xi \rangle\}\|_{f^{s\rho}_{pq}} &= \left\| \left(\sum_{j=0}^\infty 2^{jsq} \sum_{\xi \in \mathcal{X}_j} [|\langle f, \varphi_\xi \rangle| W_\mu(2^j;\xi)^{-\rho/d} \tilde{\mathbb{1}}_{R_\xi}(\cdot)]^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty 2^{jsq} \sum_{\xi \in \mathcal{X}_j} [b^*_\xi W_\mu(2^j;\xi)^{-\rho/d} \mathbb{1}_{R_\xi}(\cdot)]^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty 2^{jsq} \left[\mathcal{M}_t \left(\sum_{\xi \in \mathcal{X}_j} b_\xi W_\mu(2^j;\xi)^{-\rho/d} \mathbb{1}_{R_\xi}(\cdot) \right)(\cdot) \right]^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty 2^{jsq} \left[\sum_{\xi \in \mathcal{X}_j} b_\xi W_\mu(2^j;\xi)^{-\rho/d} \mathbb{1}_{R_\xi}(\cdot) \right]^q \right)^{1/q} \right\|_p . \end{aligned} \tag{4.18}$$

Here for the second inequality above we used Lemma 4.10 and for the third one the maximal inequality (2.17).

Denote $m_\eta := \min_{x \in R_\eta} |\overline{\Phi}_j * f(x)|$ for $\eta \in \mathcal{X}_{j+r}$ and

$$\mathcal{X}_{j+r}(\xi) := \{w \in \mathcal{X}_{j+r} : R_w \cap R_\xi \neq \emptyset\} \quad \text{for } \xi \in \mathcal{X}_j.$$

Evidently $\#\mathcal{X}_{j+r}(\xi) \leqslant c(r,d)$. Further, for $w, \eta \in \mathcal{X}_{j+r}(\xi)$ we have $d(w,\eta) \leqslant c2^{-j}$, and hence

$$m_w \leqslant c \frac{m_w}{(1 + 2^{j+r} d(w,\eta))^\sigma} \leqslant c m^*_\eta, \quad c = c(r,\sigma,d).$$

Therefore, for any $\eta \in \mathcal{X}_{j+r}(\xi)$, $b_\xi = \max_{w \in \mathcal{X}_{j+r}(\xi)} m_w \leqslant cm^*_\eta$, and hence

$$b_\xi \mathbb{1}_{R_\xi} \leqslant \sum_{\eta \in \mathcal{X}_{j+r}(\xi)} m^*_\eta \mathbb{1}_{R_\eta}. \tag{4.19}$$

Clearly, $W_\mu(2^j;\xi) \sim W_\mu(2^{j+r};\eta)$ for $\eta \in \mathcal{X}_{j+r}(\xi)$. This along with (4.19) leads to

$$b_\xi W_\mu(2^j;\xi)^{-\rho/d} \mathbb{1}_{R_\xi} \leqslant c \sum_{\eta \in \mathcal{X}_{j+r}(\xi)} m^*_\eta W_\mu(2^{j+r};\eta)^{-\rho/d} \mathbb{1}_{R_\eta}. \tag{4.20}$$

Using this estimate in (4.18), we obtain

$$\begin{aligned} \|\{\langle f, \varphi_\xi \rangle\}\|_{f^{s\rho}_{pq}} &\leqslant c \left\| \left(\sum_{j=0}^\infty 2^{jsq} \left(\sum_{\eta \in \mathcal{X}_{j+r}} m^*_\eta W_\mu(2^{j+r};\eta)^{-\rho/d} \mathbb{1}_{R_\eta}(\cdot) \right)^q \right)^{1/q} \right\|_p \\ &\leqslant c \left\| \left(\sum_{j=0}^\infty 2^{jsq} \left[\mathcal{M}_t \left(\sum_{\eta \in \mathcal{X}_{j+r}} m_\eta W_\mu(2^{j+r};\eta)^{-\rho/d} \mathbb{1}_{R_\eta} \right)(\cdot) \right]^q \right)^{1/q} \right\|_p \end{aligned}$$

$$\leqslant c \left\| \left(\sum_{j=0}^{\infty} \left(2^{js} \sum_{\eta \in \mathcal{X}_{j+r}} m_\eta W_\mu(2^{j+r};\eta)^{-\rho/d} \mathbb{1}_{R_\eta}(\cdot) \right)^q \right)^{1/q} \right\|_p$$

$$\leqslant c \left\| \left(\sum_{j=0}^{\infty} (2^{js} W_\mu(2^j;\cdot)^{-\rho/d} |\overline{\Phi}_j * f(\cdot)|)^q \right)^{1/q} \right\|_p = c\|f\|_{F_{pq}^{s\rho}}.$$

Here, for the first inequality, we used that $\#\mathcal{X}_{j+r}(\xi) \leqslant c$, for the second inequality we used Lemma 4.10, and for third one the maximal inequality (2.17). We also used that

$$W_\mu(2^{j+r};\eta) \sim W_\mu(2^j;x) \quad \text{if } x \in R_\eta,\ \eta \in \mathcal{X}_{j+r}.$$

Thus the boundedness of $S_\varphi : F_{pq}^{s\rho} \to f_{pq}^{s\rho}$ is established.

The identity $T_\psi \circ S_\varphi = \mathrm{Id}$ follows by Proposition 3.1.

It remains to show that $F_{pq}^{s\rho}$ is independent of the particular selection of $\widehat{a}$ in the definition of $\{\Phi_j\}$. Denote by $\|\cdot\|_{F_{pq}^{s\rho}(\Phi)}$ the F-norm defined by $\{\Phi_j\}$. Then by the above proof it follows that

$$\|f\|_{F_{pq}^{s\rho}(\Phi)} \leqslant c\|\{\langle f, \widetilde{\varphi}_\xi\rangle\}\|_{f_{pq}^{s\rho}} \quad \text{and} \quad \|\{\langle f, \varphi_\xi\rangle\}\|_{f_{pq}^{s\rho}} \leqslant c\|f\|_{F_{pq}^{s\rho}(\overline{\Phi})},$$

and hence

$$\|f\|_{F_{pq}^{s\rho}(\Phi)} \leqslant c\|\{\langle f, \widetilde{\varphi}_\xi\rangle\}\|_{f_{pq}^{s\rho}} \leqslant c\|f\|_{F_{pq}^{s\rho}(\overline{\widetilde{\Phi}})}.$$

Now the desired independence follows by interchanging the roles of $\{\Phi_j\}$,$\{\widetilde{\Phi}_j\}$, and their complex conjugates. □

In a sense, the spaces F_{pq}^{ss} are more natural than the spaces $F_{pq}^{s\rho}$ with $\rho \neq s$ since they embed 'correctly' with respect to the smoothness index s.

Proposition 4.11. *Let $0 < p < p_1 < \infty$, $0 < q, q_1 \leqslant \infty$, and $-\infty < s_1 < s < \infty$. Then we have the continuous embedding*

$$F_{pq}^{ss} \subset F_{p_1 q_1}^{s_1 s_1} \quad \text{if } \frac{s}{d} - \frac{1}{p} = \frac{s_1}{d} - \frac{1}{p_1}. \tag{4.21}$$

The proof of this embedding result uses the idea of the proof in the classical case on $\mathbb{R}^n$ given, for example, in [**18**, p. 129], but is more involved. We place it in § 7.2.

Finally, we would like to link the weighted Triebel–Lizorkin spaces $F_{pq}^{s\rho}$ to $L_p(w_\mu)$ and weighted potential space (generalized weighted Sobolev space) on B^d.

We define the weighted potential space $H_p^s := H_p^s(w_\mu)$, $s > 0$, $1 \leqslant p \leqslant \infty$, on B^d as the set of all $f \in \mathcal{D}'$ such that

$$\|f\|_{H_p^s} := \left\| \sum_{n=0}^{\infty} (n+1)^s \operatorname{Proj}_n f \right\|_p < \infty, \tag{4.22}$$

where $\operatorname{Proj}_n f := \mathsf{P}_n * f$.

We have the following identification of certain weighted Triebel–Lizorkin spaces.

Proposition 4.12. *We have*

$$F_{p2}^{s0} \sim H_p^s, \quad s > 0,\ 1 < p < \infty,$$

and

$$F_{p2}^{00} \sim L_p(w_\mu), \quad 1 < p < \infty,$$

with equivalent norms. Consequently, for any $f \in L_p(w_\mu)$, $1 < p < \infty$,

$$\|f\|_p \sim \left\| \left(\sum_{j=0}^{\infty} \sum_{\xi \in \mathcal{X}_j} (|\langle f, \varphi_\xi \rangle||\psi_\xi(\cdot)|)^2 \right)^{1/2} \right\|_p .$$

The proof of this proposition uses the multipliers from [**2**, Theorem 5.2] and can be carried out exactly as in the case of spherical harmonic expansions in [**12**, Proposition 4.3]. We omit it.

5. *Weighted Besov spaces on* B^d

For the definition of weighted Besov spaces on B^d, we use the sequence of kernels $\{\Phi_j\}$ defined in (4.1) with $\widehat{a}$ obeying (4.2) and (4.3) (see [**15**, **18**] for the general idea of using orthogonal or spectral decompositions for introducing Besov spaces).

Definition 5.1. Let $s, \rho \in \mathbb{R}$ and $0 < p, q \leqslant \infty$. The weighted Besov space $B_{pq}^{s\rho} := B_{pq}^{s\rho}(w_\mu)$ is defined as the set of all $f \in \mathcal{D}'$ such that

$$\|f\|_{B_{pq}^{s\rho}} := \left(\sum_{j=0}^{\infty} \left(2^{sj} \|W_\mu(2^j;\cdot)^{-\rho/d} \Phi_j * f(\cdot)\|_p \right)^q \right)^{1/q} < \infty, \tag{5.1}$$

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

Observe that as in the case of weighted Triebel–Lizorkin spaces the above definition is independent of the particular choice of $\widehat{a}$ obeying (4.2) and (4.3) (see Theorem 5.4). Also, as for $F_{pq}^{s\rho}$ the Besov space $B_{pq}^{s\rho}$ is a quasi-Banach space which is continuously embedded in $\mathcal{D}'$. We skip the details.

We next introduce the sequence spaces $b_{pq}^{s\rho}$ associated to the weighted Besov spaces $B_{pq}^{s\rho}$. To this end, we assume that $\{\mathcal{X}_j\}_{j=0}^{\infty}$ is a sequence of almost uniformly ε_j-distributed points on B^d ($\varepsilon_j := c^\diamond 2^{-j}$) with associated neighborhoods $\{R_\xi\}_{\xi \in \mathcal{X}_j}$, given by Proposition 2.12. As before we set $\mathcal{X} := \bigcup_{j \geqslant 0} \mathcal{X}_j$.

Definition 5.2. Let $s, \rho \in \mathbb{R}$ and $0 < p, q \leqslant \infty$. Then $b_{pq}^{s\rho}$ is defined to be the space of all complex-valued sequences $h := \{h_\xi\}_{\xi \in \mathcal{X}}$ such that

$$\|h\|_{b_{pq}^{s\rho}} := \left(\sum_{j=0}^{\infty} 2^{j(s-d/p+d/2)q} \left[\sum_{\xi \in \mathcal{X}_j} \left(W_\mu(2^j;\xi)^{-\rho/d+1/p-1/2} |h_\xi| \right)^p \right]^{q/p} \right)^{1/q} \tag{5.2}$$

is finite, with the usual modification for $p = \infty$ or $q = \infty$.

We shall employ again the analysis and synthesis operators S_φ and T_ψ defined in (4.6). This lemma guarantees that the operator T_ψ is well defined on $b_{pq}^{s\rho}$.

Lemma 5.3. *Let $s,\rho\in\mathbb{R}$ and $0<p,q\leqslant\infty$. Then for any $h\in b^{s\rho}_{pq}$, $T_\psi h:=\sum_{\xi\in\mathcal{X}}h_\xi\psi_\xi$ converges in $\mathcal{D}'$. Moreover, the operator $T_\psi : b^{s\rho}_{pq}\to\mathcal{D}'$ is continuous.*

The proof of this lemma is quite similar to the proof of Lemma 4.4 and will be omitted. Our main result in this section is the following characterization of weighted Besov spaces.

Theorem 5.4. *Let $s,\rho\in\mathbb{R}$ and $0<p,q\leqslant\infty$. Then the operators $S_\varphi : B^{s\rho}_{pq}\to b^{s\rho}_{pq}$ and $T_\psi : b^{s\rho}_{pq}\to B^{s\rho}_{pq}$ are bounded and $T_\psi\circ S_\varphi=\mathrm{Id}$ on $B^{s\rho}_{pq}$. Consequently, for $f\in\mathcal{D}'$ we have $f\in B^{s\rho}_{pq}$ if and only if $\{\langle f,\varphi_\xi\rangle\}_{\xi\in\mathcal{X}}\in b^{s\rho}_{pq}$. Moreover,*

$$\|f\|_{B^{s\rho}_{pq}}\sim\|\{\langle f,\varphi_\xi\rangle\}\|_{b^{s\rho}_{pq}}\sim\left(\sum_{j=0}^{\infty}2^{sjq}\left[\sum_{\xi\in\mathcal{X}_j}\Big(W_\mu(2^j;\xi)^{-\rho/d}\|\langle f,\varphi_\xi\rangle\psi_\xi\|_p\Big)^p\right]^{q/p}\right)^{1/q}. \tag{5.3}$$

In addition, the definition of $B^{s\rho}_{pq}$ is independent of the particular selection of $\widehat{a}$ satisfying (4.2) and (4.3).

For the proof of this theorem we shall utilize some of the lemmas from § 4 as well as the following additional lemma and its proof is given in § 7.2.

Lemma 5.5. *Let $0<p\leqslant\infty$ and $\gamma\in\mathbb{R}$. Then for any $P\in\Pi_{2^j}$, $j\geqslant 0$,*

$$\left(\sum_{\xi\in\mathcal{X}_j}W_\mu(2^j;\xi)^\gamma\max_{x\in R_\xi}|P(x)|^p m(R_\xi)\right)^{1/p}\leqslant c\|W_\mu(2^j;\cdot)^\gamma P(\cdot)\|_p. \tag{5.4}$$

Proof of Theorem 5.4. We first note that the right-hand side of (5.3) follows immediately from (3.12). Just as in the proof of Theorem 4.5, we assume that $\{\Phi_j\}$ are kernels defined by (4.1), with $\widehat{a}$ satisfying (4.2) and (4.3). Next, suppose that $\{\Psi_j\}$ are defined by (3.6) with $\widehat{b}$ obeying (3.1) and (3.3). Also, let $\{\varphi_\xi\}_{\xi\in\mathcal{X}}$ and $\{\psi_\xi\}_{\xi\in\mathcal{X}}$ be the associated needlet systems defined as in (3.7). Further, assume that $\{\widetilde{\Phi}_j\}$, $\{\widetilde{\Psi}_j\}$, $\{\widetilde{\varphi}_\xi\}$, $\{\widetilde{\psi}_\xi\}$ is a second (completely different) set of kernels and needlets.

Our first step is to prove the boundedness of the operator $T_{\widetilde{\psi}} : b^{s\rho}_{pq}\to B^{s\rho}_{pq}$ defined as in (4.6) with $\{\psi_\xi\}$ replaced by $\{\widetilde{\psi}_\xi\}$; we assume that $B^{s\rho}_{pq}$ is defined by $\{\Phi_j\}$.

Pick $0<t<\min\{p,1\}$ and $k\geqslant 2\mu|\rho|/d+\mu+(2\mu+d)/t$. Suppose that $h\in b^{s\rho}_{pq}$ and let $f:=\sum_{\xi\in\mathcal{X}}h_\xi\widetilde{\psi}_\xi$, which is well defined on account of Lemma 5.3. Similarly as in the proof of Theorem 4.5, we use Lemmas 2.5 and 4.6, and (2.6) to obtain

$$\begin{aligned}
W_\mu(2^j;x)^{-\rho/d}|\Phi_j*f(x)| &\leqslant c\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|W_\mu(2^j;x)^{-\rho/d}|\Phi_j*\widetilde{\psi}_\xi(x)|\\
&\leqslant c\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|\frac{2^{jd/2}W_\mu(2^j;x)^{-\rho/d}}{\sqrt{W_\mu(2^j;x)}(1+2^jd(\xi,x))^k}\\
&\leqslant c2^{jd/2}\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|\frac{W_\mu(2^j;\xi)^{-\rho/d-1/2}}{(1+2^jd(\xi,x))^{k-2\mu|\rho|/d-\mu}}\\
&\leqslant c2^{jd/2}\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|W_\mu(2^j;\xi)^{-\rho/d-1/2}\mathcal{M}_t(\mathbb{1}_{R_\xi})(x),
\end{aligned}$$

where $\mathcal{X}_{-1} := \emptyset$ and in the fourth inequality we used that $k \geqslant 2\mu|\rho|/d + \mu + (2\mu + d)/t$. Now employing the maximal inequality (2.17) we obtain

$$\begin{aligned}
\|W_\mu(2^j;\cdot)^{-\rho/d}\Phi_j * f(\cdot)\|_p &\leqslant c2^{jd/2}\Bigg\|\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|W_\mu(2^j;\xi)^{-\rho/d-1/2}\mathcal{M}_t(\mathbb{1}_{R_\xi})(\cdot)\Bigg\|_p\\
&\leqslant c2^{jd/2}\Bigg\|\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|W_\mu(2^j;\xi)^{-\rho/d-1/2}\mathbb{1}_{R_\xi}(\cdot)\Bigg\|_p\\
&\leqslant c2^{jd(1/2-1/p)}\Bigg(\sum_{j-1\leqslant\nu\leqslant j+1}\sum_{\xi\in\mathcal{X}_\nu}|h_\xi|^pW_\mu(2^j;\xi)^{-(\rho/d-1/p+1/2)p}\Bigg)^{1/p}.
\end{aligned}$$

Using this in Definition 5.1, we obtain $\|f\|_{B^{s\rho}_{pq}} \leqslant c\|\{h_\xi\}\|_{b^{s\rho}_{pq}}$. Hence the operator $T_{\widetilde{\psi}} : b^{s\rho}_{pq} \to B^{s\rho}_{pq}$ is bounded.

We next prove the boundedness of the operator $S_\varphi : B^{s\rho}_{pq} \to b^{s\rho}_{pq}$, assuming that the space $B^{s\rho}_{pq}$ is defined in terms of $\{\overline{\Phi}_j\}$ in place of $\{\Phi_j\}$. Observe first that

$$|\langle f,\varphi_\xi\rangle| \sim m(R_\xi)^{1/2}|\overline{\Phi}_j * f(\xi)| \sim 2^{-jd/2}W_\mu(2^j;\xi)^{1/2}|\overline{\Phi}_j * f(\xi)|, \quad \xi\in\mathcal{X}_j.$$

Since $\overline{\Phi}_j * f \in \Pi_{2^j}$, Lemma 5.5 implies that

$$\begin{aligned}
\sum_{\xi\in\mathcal{X}_j}\Big(W_\mu(2^j;\xi)^{-\rho/d+1/p-1/2}|\langle f,\varphi_\xi\rangle|\Big)^p &\leqslant c2^{-jd(p/2-1)}\sum_{\xi\in\mathcal{X}_j}W_\mu(2^j;\xi)^{-\rho p/d}|\overline{\Phi}_j * f(\xi)|^p m(R_\xi)\\
&\leqslant c2^{-jd(p/2-1)}\|W_\mu(2^j;\xi)^{-\rho/d}\overline{\Phi}_j * f\|_p^p.
\end{aligned}$$

This at once yields $\|\{\langle f,\varphi\rangle\}\|_{b^{s\rho}_{pq}} \leqslant c\|f\|_{B^{s\rho}_{pq}}$.

The identity $T_\psi \circ S_\varphi = \mathrm{Id}$ follows by Proposition 3.1.

The independence of $B^{s\rho}_{pq}$ of the particular selection of $\widehat{a}$ in the definition of $\{\Phi_j\}$ follows from above exactly as in the Triebel–Lizorkin case (see the proof of Theorem 4.5). □

The parameter ρ in the definition of the Besov spaces $B^{s\rho}_{pq}$ allows to consider different scales of spaces. A 'classical' choice of ρ would be $\rho = 0$. However, we maintain that most natural are the spaces B^{ss}_{pq} ($\rho = s$). The main advantages of the spaces B^{ss}_{pq} over $B^{s\rho}_{pq}$ with $\rho \neq s$ are that, first, they embed 'correctly' with respect to the smoothness index s, and secondly, the right smoothness spaces in nonlinear n-term weighted approximation from needles are defined in terms of spaces B^{ss}_{pq} (see § 6 below).

Proposition 5.6. *Let $0 < p \leqslant p_1 \leqslant \infty$, $0 < q \leqslant q_1 \leqslant \infty$, and $-\infty < s_1 \leqslant s < \infty$. Then we have the continuous embedding*

$$B^{ss}_{pq} \subset B^{s_1s_1}_{p_1q_1} \quad \text{if } \frac{s}{d} - \frac{1}{p} = \frac{s_1}{d} - \frac{1}{p_1}. \tag{5.5}$$

Proof. With Φ_j from Definition 5.1 we have $\Phi_j * f \in \Pi_{2^{j+1}}$ and applying Proposition 2.4 we obtain

$$\|W_\mu(2^j;\cdot)^{-s_1/d}\Phi_j * f(\cdot)\|_{p_1} \leqslant c2^{jd(1/p-1/p_1)}\|W_\mu(2^j;\cdot)^{-s/d}\Phi_j * f(\cdot)\|_p,$$

where we used that $s/d - 1/p = s_1/d - 1/p_1$. This leads immediately to (5.5). □

We finally want to link the weighted Besov spaces to best polynomial approximation in $L_p(w_\mu)$. As in (2.8), let $E_n(f)_p$ denote the best approximation of $f \in L_p(w_\mu)$ from Π_n. Denote by A^s_{pq} the approximation space of all functions $f \in L_p(w_\mu)$ such that

$$\|f\|_{A^s_{pq}} := \|f\|_p + \left(\sum_{j=0}^{\infty}(2^{sj}E_{2^j}(f)_p)^q\right)^{1/q} < \infty. \tag{5.6}$$

Proposition 5.7. *Let $s > 0$, $1 \leqslant p \leqslant \infty$, and $0 < q \leqslant \infty$. Then $f \in B^{s0}_{pq}$ if and only if $f \in A^s_{pq}$. Moreover,*

$$\|f\|_{B^{s0}_{pq}} \sim \|f\|_{A^s_{pq}}. \tag{5.7}$$

The proof of this proposition is similar to the proof [**12**, Proposition 5.3] and [**9**, Proposition 6.2]. We omit it.

6. *Application of weighted Besov spaces to nonlinear approximation*

Let us consider a nonlinear n-term approximation for a needlet system $\{\psi_\eta\}_{\eta\in\mathcal{X}}$ defined as in (3.5)–(3.8) with $\widehat{b} = \widehat{a}$, $\widehat{a} \geqslant 0$. Thus $\varphi_\eta = \psi_\eta$ are real valued. Then by Proposition 3.1, for any $f \in L_p(w_\mu)$, $1 \leqslant p \leqslant \infty$,

$$f = \sum_{\xi\in\mathcal{X}}\langle f, \psi_\xi\rangle\psi_\xi \quad \text{in } L_p(w_\mu).$$

Let Σ_n be the nonlinear set of all functions g of the form

$$g = \sum_{\xi\in\Lambda} a_\xi\psi_\xi,$$

where $\Lambda \subset \mathcal{X}$, $\#\Lambda \leqslant n$, and Λ may vary with g. Denote by $\sigma_n(f)_p$ the error of best $L_p(w_\mu)$-approximation to $f \in L_p(w_\mu)$ from Σ_n, that is,

$$\sigma_n(f)_p := \inf_{g\in\Sigma_n}\|f - g\|_p.$$

We consider approximation in $L_p(w_\mu)$, $0 < p < \infty$. Suppose that $s > 0$ and let $1/\tau := s/d + 1/p$. Denote briefly

$$B^s_\tau := B^{ss}_{\tau\tau}.$$

From Theorem 5.4 and (3.12) one derives the following representation of the norm in B^s_τ:

$$\|f\|_{B^s_\tau} \sim \left(\sum_{\xi\in\mathcal{X}}\|\langle f, \psi_\xi\rangle\psi_\xi\|_p^\tau\right)^{1/\tau}. \tag{6.1}$$

The following embedding result shows the importance of the spaces B^s_τ in nonlinear approximation from needlets.

Proposition 6.1. *If $f \in B^s_\tau$, then f can be identified as a function $f \in L_p(w_\mu)$ and*

$$\|f\|_p \leqslant c\left\|\sum_{\xi\in\mathcal{X}}|\langle f, \psi_\xi\rangle\psi_\xi(\cdot)|\right\|_p \leqslant c\|f\|_{B^s_\tau}. \tag{6.2}$$

We now give the main result of this section.

THEOREM 6.2 ([Jackson estimate]). *If $f \in B^s_\tau$, then*

$$\sigma_n(f)_p \leqslant cn^{-s/d}\|f\|_{B^s_\tau}, \quad n \geqslant 1. \tag{6.3}$$

The proof of Proposition 6.1 and Theorem 6.2 relies on the following lemma.

LEMMA 6.3. *Let $F = \sum_{\xi\in\mathcal{Y}_n} a_\xi\psi_\xi$, where $\mathcal{Y}_n \subset \mathcal{X}$ and $\#\mathcal{Y}_n \leqslant n$. Suppose that $\|a_\xi\psi_\xi\|_p \leqslant A$ for $\xi \in \mathcal{Y}_n$, where $0 < p < \infty$. Then $\|F\|_p \leqslant cAn^{1/p}$.*

Proof. This lemma is trivial when $0 < p \leqslant 1$. Suppose that $1 < p < \infty$. Fix $0 < t < 1$. By Theorem 4.7, for any $\xi \in \mathcal{X}$ we have $|\psi_\xi(x)| \leqslant c(\mathcal{M}_t\tilde{\mathbb{1}}_{R_\xi})(x)$ for $x \in B^d$, and applying the maximal inequality (2.17) we infer that

$$\|F\|_p \leqslant c\left\|\sum_{\xi\in\mathcal{Y}_n} \mathcal{M}_t(a_\xi\tilde{\mathbb{1}}_{R_\xi})\right\|_p \leqslant c\left\|\sum_{\xi\in\mathcal{Y}_n} |a_\xi|\tilde{\mathbb{1}}_{R_\xi}\right\|_p.$$

From $\|a_\xi\psi_\xi\|_p \leqslant A$, (3.12), and (2.27) it follows that $|a_\xi| \leqslant cAm(R_\xi)^{1/2-1/p}$, and hence

$$\|F\|_p \leqslant c\left\|\sum_{\xi\in\mathcal{Y}_n} m(R_\xi)^{-1/p}\mathbb{1}_{R_\xi}\right\|_p. \tag{6.4}$$

For $\xi \in \mathcal{X}$ denote by $\mathcal{X}_\xi$ the set of all $\eta \in \mathcal{X}$ such that $R_\eta \cap R_\xi \neq \emptyset$ and $\ell(\eta) \leqslant \ell(\xi)$, where $\ell(\eta)$, $\ell(\xi)$ are the levels of η, ξ in $\mathcal{X}$ (for example, $\ell(\xi) = j$ if $\xi \in \mathcal{X}_j$).

Let $\xi \in \mathcal{X}_j$ and $\eta \in \mathcal{X}_\xi \cap \mathcal{X}_\nu$ ($\nu \leqslant j$). Since $R_\eta \cap R_\xi \neq \emptyset$ it follows that $d(\xi,\eta) \leqslant c2^{-\nu}$ (see Definition 2.10). This combined with inequality (2.6) leads to

$$W_\mu(2^j;\xi) \leqslant W_\mu(2^\nu;\xi) \leqslant 2^\mu W_\mu(2^\nu;\eta)(1 + 2^\nu d(\xi,\eta))^{2\mu} \leqslant cW_\mu(2^\nu;\eta).$$

Therefore, using (2.27) we obtain $m(R_\xi)/m(R_\eta) \leqslant c2^{-d(j-\nu)}$ and hence

$$\sum_{\eta\in\mathcal{X}_\xi} (m(R_\xi)/m(R_\eta))^{1/p} \leqslant c < \infty. \tag{6.5}$$

Define $U(x) := \min\{m(R_\xi) : \xi \in \mathcal{Y}_n, x \in R_\xi\}$ for $x \in E := \bigcup_{\xi\in\mathcal{Y}_n} R_\xi$. By (6.5) it follows that

$$\sum_{\xi\in\mathcal{Y}_n} m(R_\xi)^{-1/p}\mathbb{1}_{R_\xi}(x) \leqslant cU(x)^{-1/p}, \quad x \in E.$$

We use this in (6.4) to obtain

$$\begin{aligned}\|F\|_p &\leqslant c\|U^{-1/p}\|_p = c\left(\int_E U^{-1}(x)w_\mu(x)dx\right)^{1/p}\\ &\leqslant cA\left(\sum_{\xi\in\mathcal{Y}_n} m(R_\xi)^{-1}\int_{B^d} \mathbb{1}_{R_\xi}(x)w_\mu(x)dx\right)^{1/p} = cA(\#\mathcal{Y}_n)^{1/p} \leqslant cAn^{1/p}.\end{aligned}$$

□

For the proof of Proposition 6.1 and Theorem 6.2 one proceeds exactly as in the proof of [**12**, Proposition 6.1 and Theorem 6.2], using Lemma 6.3. We omit the further details.

It is an open problem to prove the companion to (6.3) Bernstein estimate:

$$\|g\|_{B^s_\tau} \leqslant cn^{s/d}\|g\|_p \quad \text{for } g \in \Sigma_n,\ 1 < p < \infty. \tag{6.6}$$

This estimate would allow to characterize the rates of nonlinear n-term approximation from needlet systems in $L_p(w_\mu)$ ($1 < p < \infty$).

7. *Proofs*

7.1. *Proofs for §§ 2 and 3*

Proof of Theorem 2.2. We shall first establish (2.3) for $p = 2$. From the definition of the kernels $\mathsf{P}_n(x,y)$ (see (1.4) and (1.5)) it follows that

$$\int_{B^d} \mathsf{P}_n(x,y)\mathsf{P}_m(x,y) w_\mu(y)\, dy = \delta_{n,m}\mathsf{P}_n(x,x),$$

and hence

$$\int_{B^d} |L_n(x,y)|^2 w_\mu(y) dy = \sum_{k=0}^{2n} \Big|\widehat{a}\Big(\frac{k}{n}\Big)\Big|^2 \mathsf{P}_k(x,x). \tag{7.1}$$

Therefore, for $p = 2$ estimate (2.3) will follow by the following lemma. □

Lemma 7.1. *For any* $\varepsilon > 0$

$$\sum_{j=n}^{n+[\varepsilon dn]} \mathsf{P}_j(x,x) \geqslant \frac{cn^d}{W_\mu(n;x)}, \quad x \in B^d, \ \ n \geqslant 1/\varepsilon, \tag{7.2}$$

where $c > 0$ *depends only on* ε, μ, *and* d.

Proof. Assume that $\mu > 0$. We shall utilize representation (1.5) of $\mathsf{P}_n(x,y)$. The case $\mu = 0$ is easier and will be omitted (in this case one uses representation (4.2) of $\mathsf{P}_n(x,y)$ from [**14**]).

From (1.5) it is obvious that $\mathsf{P}_n(x,x)$ depends only on $|x|$. For the rest of the proof, we denote $\mathsf{P}_{n,d}(r) := \mathsf{P}_n(x,x)$, where $r := |x|$, and $\Lambda_{n,d}(r) := \sum_{j=n}^{n+[\varepsilon dn]} \mathsf{P}_{j,d}(r)$. Summing up the well-known recurrence relation [**17**, (4.7.29)]

$$C_n^\lambda(x) - C_{n-2}^\lambda(x) = \frac{n+\lambda-1}{\lambda-1} C_n^{\lambda-1}(x), \quad \text{where } C_{-1}^\lambda(x) = C_{-2}^\lambda(x) := 0,$$

we obtain

$$C_n^\lambda(x) = \sum_{0\leqslant 2j\leqslant n} \frac{n-2j+\lambda-1}{\lambda-1} C_{n-2j}^{\lambda-1}(x).$$

Combining this with (1.5) we arrive at

$$\mathsf{P}_{n,d}(r) = \frac{b_d^\mu}{b_{d-2}^\mu}\frac{n+\lambda}{\lambda} \sum_{0\leqslant 2j\leqslant n} \mathsf{P}_{n-2j,d-2}(r).$$

Hence

$$\begin{aligned}\Lambda_{n,d}(r) = \sum_{k=n}^{n+[\varepsilon dn]} \mathsf{P}_{k,d}(r) &= \frac{b_d^\mu}{b_{d-2}^\mu} \sum_{k=n}^{n+[\varepsilon dn]} \frac{k+\lambda}{\lambda} \sum_{0\leqslant 2j\leqslant k} \mathsf{P}_{k-2j,d-2}(r)\\ &\geqslant c\,n^2 \sum_{k=n}^{n+[\varepsilon(d-2)n]} \mathsf{P}_{k,d-2}(r) = c\,n^2\Lambda_{n,d-2}(r).\end{aligned}$$

Here $c > 0$ depends only on ε, μ, and d; we used that $n \geqslant 1/\varepsilon$.

Evidently, the above estimate leads to (7.2) using induction on d, provided that we prove (7.2) for $d = 1$ and $d = 2$. However, the case $d = 1$ is already established in [**9**, Proposition 2.4], namely,

$$\Lambda_{n,1}(r) \geqslant \frac{cn}{W_\mu(n;r)}. \tag{7.3}$$

It remains to prove (7.2) in the case $d = 2$. The proof relies on the well-known identity

$$C_n^\lambda(x) = \sum_{0\leqslant 2k\leqslant n} \frac{\Gamma(\mu)(n-2k+\mu)\Gamma(k+\lambda-\mu)\Gamma(n-k+\lambda)}{\Gamma(\lambda)\Gamma(\lambda-\mu)k!\Gamma(n-k+\mu+1)} C_{n-2k}^\mu(x) \tag{7.4}$$

(see [**1**, p. 59]) and the product formula of Gegenbauer polynomials [**4**, Vol I, Paragraph 3.15.1, (20)]:

$$\frac{C_n^\mu(s)C_n^\mu(t)}{C_n^\mu(1)} = b_1^{\mu-1/2}\int_{-1}^1 C_n^\mu\left(st + u\sqrt{1-s^2}\sqrt{1-t^2}\right)(1-u^2)^{\mu-1}\,du. \tag{7.5}$$

Using (7.4) (with $\lambda = \mu + 1/2$) along with (1.5) and then (7.5), we obtain

$$\begin{aligned} \mathsf{P}_{n,2}(r) &= b_2^\mu \frac{n+\mu+1/2}{\mu+1/2} \sum_{0\leqslant 2k\leqslant n} c_{k,n} \frac{n-2k+\mu}{\mu} \frac{[C_{n-2k}^\mu(r)]^2}{C_{n-2k}^\mu(1)} \\ &= \frac{b_2^\mu}{b_1^\mu} \frac{n+\mu+1/2}{\mu+1/2} \sum_{0\leqslant 2k\leqslant n} c_{k,n}\mathsf{P}_{n-2k,1}(r), \end{aligned} \tag{7.6}$$

where

$$c_{k,n} = \frac{\Gamma(\mu+1)\Gamma(k+1/2)\Gamma(n-k+\mu+1/2)}{\Gamma(\mu+1/2)\Gamma(1/2)\Gamma(n-k+\mu+1)k!}.$$

Here we used that the $L_2(w_\mu)$-normalized Gegenbauer polynomial $\widetilde{C}_n^\mu$ can be written in the form $\widetilde{C}_n^\mu(x) = h_n^{-1/2}C_n^\mu(x)$ with $h_n := (b_1^\mu)^{-1}\mu/(n+\mu)C_n^\mu(1)$, which is a matter of simple verification, and hence

$$\mathsf{P}_{n,1}(r) = [\widetilde{C}_n^\mu(r)]^2 = b_1^\mu \frac{n+\mu}{\mu}\frac{[C_n^\mu(r)]^2}{C_n^\mu(1)}.$$

It is straightforward to verify that if $0 \leqslant k \leqslant n/2$, then $c_{k,n} \sim (kn)^{-1/2}$, and hence $c_{k,n} \geqslant cn^{-1}$. Therefore, from (7.6)

$$\begin{aligned} \Lambda_{n,2}(r) &= \sum_{k=n}^{n+[2\varepsilon n]} \mathsf{P}_{k,2}(r) = \frac{b_2^\mu}{b_1^\mu}\sum_{k=n}^{n+[2\varepsilon n]} \frac{k+\mu+1/2}{\mu+1/2}\sum_{0\leqslant 2j\leqslant k} c_{j,k}\mathsf{P}_{k-2j,1}(r) \\ &\geqslant c\sum_{k=n}^{n+[2\varepsilon n]}\sum_{0\leqslant 2j\leqslant k}\mathsf{P}_{k-2j,1}(r) \geqslant c\,n\Lambda_{n,1}(r). \end{aligned}$$

This combined with (7.3) yields (7.2) for $d = 2$. □

We now continue with the proof of Theorem 2.2. Applying (7.2) with $\varepsilon = 2/3d$ yields $\|L_n(x,\cdot)\|_2 \geqslant cn^d W_\mu(n;x)^{-1}$ for $n \geqslant 2d$. If $2 \leqslant n < 2d$, then as in the proof of Lemma 7.1 it follows that

$$\|L_n(x,\cdot)\|_2^{1/2} \geqslant c(\mathsf{P}_n(x,x) + \mathsf{P}_{n+1}(x,x)) \geqslant c(C_n^\mu(|x|) + C_{n+1}^\mu(|x|)) > c > 0$$

for all $x \in B^d$, where we used the fact that the polynomials C_n^μ and C_{n+1}^μ have no common zeros. Taking into account that $W_\mu(n;x) \sim 1$ when $n \leqslant 2d$, the above leads again to $\|L_n(x,\cdot)\|_2 \geqslant cn^d W_\mu(n;x)^{-1}$. This completes the proof of estimate (2.3) for $p = 2$.

Now, one easily derives (2.3) for $p \neq 2$ from the same estimate for $p = 2$ and the upper bound estimate (2.1). Indeed, for $2 < p < \infty$ applying Hölder's inequality we obtain

$$\begin{aligned} \frac{cn^d}{W_\mu(n,x)} &\leqslant \int_{B^d} |L_n(x,y)|^2 w_\mu(y)\,dy \leqslant \|L_n(x,\cdot)\|_p\|L_n(x,\cdot)\|_{p'} \\ &\leqslant c_1\|L_n(x,\cdot)\|_p\Big(\frac{n^d}{W_\mu(n,x)}\Big)^{1-1/p'} \quad \left(\frac{1}{p}+\frac{1}{p'}=1\right), \end{aligned}$$

which implies (2.3). One proceeds similarly whenever $p = \infty$.

If $0 < p < 2$, using (2.3) for $p = 2$ and (2.1) for $p = \infty$, then we obtain

$$\begin{aligned}\frac{cn^d}{W_\mu(n,x)} &\leqslant \int_{B^d} |L_n(x,y)|^2 w_\mu(y)\,dy \leqslant \int_{B^d} |L_n(x,y)|^p w_\mu(y)\,dy \|L_n(x,\cdot)\|_\infty^{2-p} \\ &\leqslant c_1 \int_{B^d} |L_n(x,y)|^p w_\mu(y)\,dy \Big(\frac{n^d}{W_\mu(n,x)}\Big)^{2-p}.\end{aligned}$$

This again leads to (2.3). The proof of Theorem 2.2 is complete.

Proof of Proposition 2.4. Let $g \in \Pi_n$. Assume that $1 < q < \infty$ and let L_n be the kernel from (1.8), with $\widehat{a}$ admissible of type (a). By Lemma 2.3, $g = L_n * g$. We use this, Hölder's inequality, (2.1), and that $W_\mu(n;x) \geqslant n^{-2\mu}$ to obtain

$$|g(x)| \leqslant c\|g\|_q \left(\frac{n^d}{W_\mu(n;x)}\right)^{1/q} \leqslant cn^{(d+2\mu)/q}\|g\|_q, \quad x \in B^d,$$

and hence

$$\|g\|_\infty \leqslant cn^{(d+2\mu)/q}\|g\|_q, \quad 1 < q \leqslant \infty. \tag{7.7}$$

Let $0 < q \leqslant 1$. The above inequality with $q = 2$ yields

$$\|g\|_\infty^2 \leqslant cn^{(d+2\mu)} \int_{B^d} |g(y)|^{2-q}|g(y)|^q w_\mu(y)dy \leqslant cn^{d+2\mu}\|g\|_\infty^{2-q}\|g\|_q^q.$$

Therefore, (7.7) holds for $0 < q \leqslant 1$ as well.

Let $0 < q < p < \infty$. Using (7.7) we have

$$\begin{aligned}\|g\|_p &= \left(\int_{B^d} |g(x)|^{p-q}|g(x)|^q w_\mu(x)\,dx\right)^{1/p} \\ &\leqslant cn^{(d+2\mu)(1/q-1/p)}\|g\|_q^{(p-q)/p}\|g\|_q^{q/p} = cn^{(d+2\mu)(1/q-1/p)}\|g\|_q.\end{aligned}$$

Thus we have proved (2.10).

We next prove (2.11). Assume first that $1 < q < \infty$. Using again that $g = L_n * g$, Hölder's inequality $(1/q + 1/q' = 1)$, and (1.11) we obtain for $x \in B^d$,

$$\begin{aligned}|g(x)| &\leqslant \|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q \left(\int_{B^d} \left|L_n(x,y)W_\mu(n;y)^{-\gamma-1/p+1/q}\right|^{q'} w_\mu(y)\,dy\right)^{1/q'} \\ &\leqslant c\frac{n^d}{W_\mu(n;x)^{1/2}} \left(\int_{B^d} \frac{w_\mu(y)\,dy}{W_\mu(n;y)^{q'/2+\beta}(1+nd(x,y))^\sigma}\right)^{1/q'} \|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q,\end{aligned}$$

where $\beta = q'(\gamma + 1/p - 1/q)$. The last integral can be estimated by using (2.2), yielding

$$|g(x)| \leqslant c\frac{n^{d/q}}{W_\mu(n;x)^{\gamma+1/p}}\|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q.$$

Hence

$$\|W_\mu(n;\cdot)^{\gamma+1/p}g(\cdot)\|_\infty \leqslant cn^{d/q}\|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q, \quad 1 < q \leqslant \infty. \tag{7.8}$$

Let $0 < q \leqslant 1$. Then by (7.8) with $q = 2$ we have

$$\begin{aligned}\|W_\mu(n;\cdot)^{\gamma+1/p}g(\cdot)\|_\infty &\leqslant cn^{d/2}\|W_\mu(n;\cdot)^{\gamma+1/p-1/2}g(\cdot)\|_2 \\ &\leqslant cn^{d/2}\|W_\mu(n;\cdot)^{\gamma+1/p}g(\cdot)\|_\infty^{1-q/2}\|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q^{q/2}.\end{aligned}$$

Therefore, (7.8) holds for $0 < q \leqslant 1$ as well.

Let $p < \infty$. Using (7.8), we have

$$\|W_\mu(n;\cdot)^\gamma g(\cdot)\|_p = \Big(\int_{B^d} |W_\mu(n;x)^\gamma g(x)|^{p-q}\,|W_\mu(n;x)^\gamma g(x)|^q\, w_\mu(x)\,dx\Big)^{1/p}$$
$$\leqslant cn^{d(1/q-1/p)}\|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q^{1-q/p}\Big(\int_{B^d}\frac{|W_\mu(n;x)^\gamma g(x)|^q}{W_\mu(n;x)^{(p-q)/p}}w_\mu(x)\,dx\Big)^{1/p}$$
$$= cn^{d(1/q-1/p)}\|W_\mu(n;\cdot)^{\gamma+1/p-1/q}g(\cdot)\|_q.$$

Hence (2.11) holds for $p < \infty$. If $p = \infty$, then (2.11) follows from (7.8). □

Proof of (3.13). From (3.10) with k sufficiently large ($k > d + 2\mu$ will do) and (3.12), we infer for $0 < r \leqslant \pi$

$$0 < c_1 \leqslant \|\varphi_\xi\|_2$$
$$\leqslant \|\varphi_\xi\|_{L_\infty(B_\xi(r))} m(B_\xi(r)) + c2^{jd}\int_{B^d\setminus B_\xi(r)} \frac{w_\mu(y)}{W_\mu(2^j;y)(1+2^j d(\xi,y))^{2k}}\,dy$$
$$\leqslant \|\varphi_\xi\|_{L_\infty(B_\xi(r))} m(B_\xi(r)) + c\frac{2^{jd}}{(1+2^jr)^k}\int_{B^d} \frac{w_\mu(y)}{W_\mu(2^j;y)(1+2^j d(\xi,y))^{k}}\,dy$$
$$\leqslant \|\varphi_\xi\|_{L_\infty(B_\xi(r))} m(B_\xi(r)) + \frac{c_2}{1+2^jr},$$

where c_2 depends only on k, d, and μ. For the last inequality we used Lemma 2.1 with $p = 2$. Let $r := c^*2^{-j}$, where $c^* > 0$ is selected such that $c_2/(1+2^jr) = c_2/(1+c^*) < c_1/2$. Then from above

$$\|\varphi_\xi\|_{L_\infty(B_\xi(c^*2^{-j}))} \geqslant \frac{c}{m(B_\xi(c^*2^{-j}))} \geqslant c\Big(\frac{2^{jd}}{W_\mu(2^j;\xi)}\Big)^{1/2}.$$

A similar estimate holds for ψ_ξ as well. □

7.2. *Proofs for* §§ 4 *and* 5

Proof of Lemma 4.6. Using the orthogonality of the subspaces $\mathcal{V}_n^d$, we have $\Phi_j * \psi_\xi(x) = 0$ if $\xi \in \mathcal{X}_\nu$ for $\nu \geqslant j+2$ or $\nu \leqslant j-2$.

Let $\xi \in \mathcal{X}_\nu$, $j-1 \leqslant \nu \leqslant j+1$. From the localization of the kernels Φ_j, given in (3.9), and the needlet localization from (3.10) it follows that for any $k > 0$ there is a constant $c_k > 0$ such that

$$|\Phi_j * \psi_\xi(x)| \leqslant c_k\frac{2^{j3d/2}}{\sqrt{W_\mu(2^j;x)}}\int_{B^d}\frac{w_\mu(y)}{W_\mu(2^j;y)(1+2^jd(x,y))^k(1+2^jd(y,\xi))^k}\,dy.$$

Denote

$$\Omega_\xi := \Big\{y \in B^d : d(y,\xi) \geqslant \frac{d(x,\xi)}{2}\Big\} \quad\text{and}\quad \Omega_x := \Big\{y \in B^d : d(x,y) \geqslant \frac{d(x,\xi)}{2}\Big\}.$$

Evidently, $B^d = \Omega_\xi \cup \Omega_x$ and hence

$$|\Phi_j * \psi_\xi(x)| \leqslant c_k\frac{2^{j3d/2}}{\sqrt{W_\mu(2^j;x)}(1+2^jd(x,\xi))^k}\int_{\Omega_\xi}\frac{w_\mu(y)}{W_\mu(2^j;y)(1+2^jd(x,y))^k}\,dy$$
$$+ c_k\frac{2^{j3d/2}}{\sqrt{W_\mu(2^j;x)}(1+2^jd(x,\xi))^k}\int_{\Omega_x}\frac{w_\mu(y)}{W_\mu(2^j;y)(1+2^jd(y,\xi))^k}\,dy$$
$$=: J_1 + J_2.$$

We may assume that $k > d$. Then employing Lemma 2.1 with $p = 2$, we obtain

$$\int_{\Omega_\xi} \frac{w_\mu(y)}{W_\mu(2^j; y)(1 + 2^j d(x, y))^k}\, dy \leqslant \int_{B^d} \frac{w_\mu(y)}{W_\mu(2^j; y)(1 + 2^j d(x, y))^k}\, dy \leqslant c2^{-jd},$$

which yields

$$J_1 \leqslant c\frac{2^{jd/2}}{\sqrt{W_\mu(2^j; x)}(1 + 2^j d(x, \xi))^k}.$$

One similarly estimates J_2. This completes the proof of the lemma. □

Proof of Lemma 4.7. Estimate (4.12) follows readily from the localization of the needlets (see (3.10)) and the lower bound estimate from (2.19) taking into account that $R_\xi \subset B_\xi(c^\diamond 2^{-j})$ for $\xi \in \mathcal{X}_j$.

We now prove (4.13). By the lower bound estimate (3.13) it follows that there exists a $\omega \in B_\xi(c^* 2^{-j})$ such that

$$|\varphi_\xi(\omega)| \geqslant c\frac{2^{jd/2}}{\sqrt{W_\mu(2^j; \xi)}}. \tag{7.9}$$

Also, by (3.11) it follows that for every $x \in B_\omega(2^{-j})$

$$|\varphi_\xi(\omega) - \varphi_\xi(x)| \leqslant c\frac{2^{j(d/2+1)} d(\omega, x)}{\sqrt{W_\mu(2^j; \xi)}}. \tag{7.10}$$

By (7.9)–(7.10) it follows that for a sufficiently small constant $\hat{c} > 0$

$$|\varphi_\xi(x)| \geqslant |\varphi_\xi(\omega)| - |\varphi_\xi(\omega) - \varphi_\xi(x)| \geqslant c\frac{2^{jd/2}}{\sqrt{W_\mu(2^j; \xi)}} \geqslant c\tilde{\mathbb{1}}_{B_\omega(\hat{c}2^{-j})}(x), \quad x \in B_\omega(\hat{c}2^{-j}),$$

which yields

$$(\mathcal{M}_t\varphi_\xi)(x) \geqslant c(\mathcal{M}_t\tilde{\mathbb{1}}_{B_\omega(\hat{c}2^{-j})})(x) \geqslant c\tilde{\mathbb{1}}_{B_\xi(2^{-j})}(x) \geqslant c\tilde{\mathbb{1}}_{R_\xi}(x), \quad x \in B^d,$$

where in the second inequality we used (2.19).

One similarly shows that $\mathcal{M}_t\psi_\xi \geqslant c\tilde{\mathbb{1}}_{R_\xi}$. □

For the proof of Lemma 4.9 we need a couple of additional lemmas.

Lemma 7.2. *Let $k > d$ and $j \geqslant 0$. Then*

$$\sum_{\xi \in \mathcal{X}_j} \frac{1}{(1 + 2^j d(x, \xi))^k} \leqslant c, \quad x \in B^d, \tag{7.11}$$

and for any $\xi, \eta \in B^d$

$$\sum_{w \in \mathcal{X}_j} \frac{1}{(1 + 2^j d(\xi, w))^k (1 + 2^j d(\eta, w))^k} \leqslant c\frac{1}{(1 + 2^j d(\xi, \eta))^k}. \tag{7.12}$$

Proof. Fix $\xi \in \mathcal{X}_j$. Evidently, $1 + 2^j d(x, \xi) \sim 1 + 2^j d(x, y)$ for $y \in R_\xi$, and by (2.5)

$$|\sqrt{1 - |\xi|^2} - \sqrt{1 - |y|^2}| \leqslant \sqrt{2}\, d(\xi, y) \leqslant c2^{-j}, \quad y \in R_\xi,$$

which implies that

$$|R_\xi| \sim 2^{-jd}(\sqrt{1 - |\xi|^2} + 2^{-j}) \sim 2^{-jd}(\sqrt{1 - |y|^2} + 2^{-j}), \quad y \in R_\xi.$$

We use the above to obtain

$$\sum_{\xi \in \mathcal{X}_j} \frac{1}{(1+2^j d(x,\xi))^k} \leqslant c \sum_{\xi \in \mathcal{X}_j} \frac{1}{|R_\xi|} \int_{R_\xi} \frac{1}{(1+2^j d(x,y))^k}\, dy$$
$$\leqslant c 2^{jd} \int_{B^d} \frac{1}{(\sqrt{1-|y|^2}+2^{-j})(1+2^j d(x,y))^k}\, dy \leqslant c.$$

Here for the last inequality we used Lemma 2.1 with $p = 2$ and $\mu = 1/2$.

For the proof of (7.12), assume that $\xi \neq \eta$ and denote

$$\mathcal{X}_j(\xi) := \left\{ w \in \mathcal{X}_j : d(\xi, w) \geqslant \frac{d(\xi,\eta)}{2} \right\}, \quad \mathcal{X}_j(\eta) := \left\{ w \in \mathcal{X}_j : d(\eta, w) \geqslant \frac{d(\xi,\eta)}{2} \right\}.$$

Then

$$\sum_{w \in \mathcal{X}_j} \frac{1}{(1+2^j d(\xi,w))^k (1+2^j d(\eta,w))^k} \leqslant c \frac{1}{(1+2^j d(\xi,\eta))^k} \sum_{w \in \mathcal{X}_j(\xi)} \frac{1}{(1+2^j d(\eta,w))^k}$$
$$+ c \frac{1}{(1+2^j d(\xi,\eta))^k} \sum_{w \in \mathcal{X}_j(\eta)} \frac{1}{(1+2^j d(\xi,w))^k}$$
$$\leqslant c \frac{1}{(1+2^j d(\xi,\eta))^k} \Big(\sum_{w \in \mathcal{X}_j} \frac{1}{(1+2^j d(\eta,w))^k} + \sum_{w \in \mathcal{X}_j} \frac{1}{(1+2^j d(\xi,w))^k} \Big)$$
$$\leqslant c \frac{1}{(1+2^j d(\xi,\eta))^k},$$

where for the last inequality we used (7.11). □

LEMMA 7.3. *Assume that $P \in \Pi_{2^j}$ $(j \geqslant 0)$, $\xi \in \mathcal{X}_j$, and let $x_1, x_2 \in B^d$ and let $d(x_\nu, \eta) \leqslant \tilde{c} 2^{-j}$, $\nu = 1, 2$. For any $k > 0$*

$$|P(x_1) - P(x_2)| \leqslant c 2^j d(x_1, x_2) \sum_{\xi \in \mathcal{X}_j} \frac{|P(\xi)|}{(1+2^j d(\eta,\xi))^k},$$

where $c > 0$ depends only on d, k, μ, and $\tilde{c}$.

Proof. Fix $P \in \Pi_{2^j}$ and assume that L_{2^j} is the reproducing kernel from Lemma 2.3 with $n = 2^j$. Then, $L_{2^j} * P = P$. Since $L_{2^j}(x, \cdot) P(\cdot) \in \Pi_{2^{j+2}}$, and the cubature formula (2.25) is exact for all polynomials from $\Pi_{2^{j+2}}$ we have

$$P(x) = \int_{B^d} L_{2^j}(x,y) P(y) w_\mu(y)\, dy = \sum_{\xi \in \mathcal{X}_j} \lambda_\xi L_{2^j}(x,\xi) P(\xi), \quad x \in B^d.$$

We use (2.4) to obtain, for $x_1, x_2 \in B^d$ with $d(x_\nu, \eta) \leqslant \tilde{c} 2^{-j}$, $\nu = 1, 2$,

$$|P(x_1) - P(x_2)| = \Big| \int_{B^d} [L_{2^j}(x_1,y) - L_{2^j}(x_2,y)] P(y) w_\mu(y)\, dy \Big|$$
$$\leqslant \sum_{\xi \in \mathcal{X}_j} |\lambda_\xi| |L_{2^j}(x_1,\xi) - L_{2^j}(x_2,\xi)| |P(\xi)|$$
$$\leqslant c 2^j d(x_1,x_2) \sum_{\xi \in \mathcal{X}_j} \left(\frac{W_\mu(2^j;\xi)}{W_\mu(2^j;\eta)} \right)^{1/2} \frac{|P(\eta)|}{(1+2^j d(\xi,\eta))^k}$$
$$\leqslant c 2^j d(x_1,x_2) \sum_{\eta \in \mathcal{X}_j} \frac{|P(\eta)|}{(1+2^j d(\xi,\eta))^{k-2\mu}}.$$

Here we used that $\lambda_\xi \sim 2^{-jd} W_\mu(2^j;\xi)$ and for the last inequality we used (2.6). Taking into account that $k>0$ can be arbitrarily large, the result follows. $\square$

Proof of Lemma 4.9. Let $d_\xi := \max\{|P(x_1)-P(x_2)| : x_1 \in R_\xi, d(x_1,x_2) \leqslant 2^{-j-r}\}$. Obviously $a_\xi \leqslant b_\xi + d_\xi$. Now Lemma 7.3 yields

$$d_\xi \leqslant c2^{-r} \sum_{\eta\in\mathcal{X}_j} \frac{|P(\eta)|}{(1+2^j d(\xi,\eta))^k}, \quad \xi\in\mathcal{X}_j.$$

From the definition of d_ξ^* in (4.14) we infer that

$$d_\xi^* \leqslant c2^{-r} \sum_{w\in\mathcal{X}_j}\sum_{\eta\in\mathcal{X}_j} \frac{|P(\eta)|}{(1+2^j d(w,\eta))^k (1+2^j d(\xi,w))^k} \leqslant c2^{-r} \sum_{\eta\in\mathcal{X}_j} \frac{|P(\eta)|}{(1+2^j d(\eta,\xi))^k} \leqslant c2^{-r} a_\xi^*,$$

where for the second inequality we interchanged the order of summation and used Lemma 7.2. Hence, $a_\xi^* \leqslant b_\xi^* + d_\xi^* \leqslant b_\xi^* + c2^{-r} a_\xi^*$ with $c>0$ independent of r. By selecting r sufficiently large, we obtain $a_\xi^* \leqslant c b_\xi^*$. $\square$

Proof of Lemma 4.10. We first prove Lemma 4.10 in the case $\rho = 0$. We fix $\xi\in\mathcal{X}_j$ and define $S_0 := \{\eta\in\mathcal{X}_j : d(\eta,\xi) \leqslant c^\diamond 2^{-j}\}$ and

$$S_m := \{\eta\in\mathcal{X}_j : c^\diamond 2^{-j+m-1} < d(\eta,\xi) \leqslant c^\diamond 2^{-j+m}\}, \quad m\geqslant 1,$$

where $c^\diamond$ is the constant from Proposition 2.12. By Definition 2.10 it follows that $\#S_m \leqslant c2^{md}$. Let us also set

$$B_m := B_\xi(c^\diamond(2^m+1)2^{-j}), \quad m\geqslant 0.$$

Evidently, $R_\eta \subset B_m$ for $\eta\in S_\nu$, $0\leqslant\nu\leqslant m$. Moreover, if $\eta\in S_m$, then

$$d(\xi,\partial B^d) \leqslant d(\xi,\eta) + d(\eta,\partial B^d) \leqslant c^\diamond 2^{-j+m} + d(\eta,\partial B^d).$$

Hence, using (2.14), we obtain

$$\frac{m(B_m)}{m(R_\eta)} \leqslant 2^{md}\left(\frac{d(\xi,\partial B^d)+2^{-j+m}}{d(\eta,\partial B^d)+2^{-j}}\right)^{2\mu} \leqslant c2^{md}\left(\frac{d(\eta,\partial B^d)+2^{-j+m}}{d(\eta,\partial B^d)+2^{-j}}\right)^{2\mu} \leqslant c2^{m(d+2\mu)}. \tag{7.13}$$

Set $\gamma := \max\{0, 1-1/t\} < 1$. Using Hölder's inequality if $t>1$ and the t-triangle inequality if $0<t\leqslant 1$, we obtain

$$b_\xi^* = \sum_{\eta\in\mathcal{X}_j} \frac{|b_\eta|}{(1+2^j d(\eta,\xi))^\sigma} \leqslant c\sum_{m\geqslant 0} 2^{-m\sigma} \sum_{\eta\in S_m} |b_\eta| \leqslant c\sum_{m\geqslant 0} 2^{-m(\sigma-d\gamma)} \left(\sum_{\eta\in S_m} |b_\eta|^t\right)^{1/t}.$$

We now use (7.13) to obtain, for $x\in R_\xi$,

$$\begin{aligned} b_\xi^* &= c\sum_{m=0}^\infty 2^{-m(\sigma-d)} \left(\int_{B^d}\left[\sum_{\eta\in S_m} |b_\eta| m(R_\eta)^{-1/t} \mathbb{1}_{R_\eta}(y)\right]^t w_\mu(y)\,dx\right)^{1/t} \\ &\leqslant c\sum_{m=0}^\infty 2^{-m(\sigma-d)} \left(\frac{1}{m(B_m)}\int_{B_m}\left[\sum_{\eta\in S_m}\left(\frac{m(B_m)}{m(R_\eta)}\right)^{1/t} |b_\eta| \mathbb{1}_{R_\eta}(y)\right]^t w_\mu(y)\,dy\right)^{1/t} \end{aligned}$$

$$\leqslant c\sum_{m\geqslant 0} 2^{-m(\sigma-d-(d+2\mu)/t)}\left(\frac{1}{m(B_m)}\int_{B_m}\left[\sum_{\eta\in S_m}|b_\eta|\mathbb{1}_{R_\eta}(y)\right]^t w_\mu(y)\,dy\right)^{1/t}$$
$$\leqslant c\mathcal{M}_t\left(\sum_{w\in\mathcal{X}_j}|b_w|\mathbb{1}_{R_\omega}\right)(x),$$

where for the last inequality we used that $\sigma > d + (d+2\mu)/t$.

Consider now the general case. Using (2.6) we have for $\xi\in\mathcal{X}_j$

$$W_\mu(2^j;\xi)^\gamma b_\xi^* \leqslant \sum_{\eta\in\mathcal{X}_j}\frac{W_\mu(2^j;\xi)^\gamma|b_\eta|}{(1+2^j d(\xi,\eta))^\sigma} \leqslant c\sum_{\eta\in\mathcal{X}_j}\frac{W_\mu(2^j;\eta)^\gamma|b_\eta|}{(1+2^j d(\xi,\eta))^{\sigma-2\mu|\gamma|}} \leqslant c\Big(W_\mu(2^j;\xi)^\gamma|b_\xi|\Big)^*,$$

where we used that $\sigma > d+(d+2\mu)/t+2\mu|\gamma|$. Now (4.16) in the general case follows by the same inequality in the case $\rho = 0$ established above. □

Proof of Proposition 4.11. On account of Theorem 4.5, it suffices to show that under the hypothesis of Proposition 4.11 we have the continuous embedding $f^{ss}_{pq}\subset f^{s_1s_1}_{p_1q_1}$. Moreover, we may assume that $q=\infty$ and $0<q_1<1$.

Suppose that $h\in f^{ss}_{p\infty}$ and $\|h\|_{f^{ss}_{p\infty}} = 1$. We need to show that

$$\|h\|_{f^{s_1s_1}_{p_1q_1}} \leqslant c<\infty,\quad \text{where } \frac{s_1}{d}-\frac{1}{p_1}=\frac{s}{d}-\frac{1}{p},\ \ 0<p<p_1<\infty, s_1<s. \tag{7.14}$$

To simplify our notation we let

$$Q_j(x;s,p) := 2^{js}\sum_{\xi\in\mathcal{X}_j}|h_\xi|W_\mu(2^j;\xi)^{-s/d}\tilde{\mathbb{1}}_{R_\xi}(x)w_\mu(x)^{1/p}.$$

Then

$$\|h\|_{f^{ss}_{p\infty}} := \Big\|\sup_{j\geqslant 0} Q_j(\cdot;s,p)\Big\|_{L^p(B^d)} \quad\text{and}\quad \|h\|_{f^{s_1s_1}_{p_1q_1}} := \left\|\left(\sum_{j\geqslant 0}Q_j(\cdot;s_1,p_1)^{q_1}\right)^{1/q_1}\right\|_{L^{p_1}(B^d)}.$$

Here $\|\cdot\|_{L^p(B^d)}$ stands for the L^p-norm on B^d with weight 1.

Denote $E_\ell := \{x\in B^d : 2^{-\ell} < (1-|x|^2)^{1/2}\leqslant 2^{-\ell+1}\}$ and, also

$$\mathcal{N}_\ell(h) := \left\|\left(\sum_{j=\ell}^{\infty}Q_j(\cdot;s_1,p_1)^{q_1}\right)^{1/q_1}\right\|^{p_1}_{L^{p_1}(E_\ell)},$$
$$\widetilde{\mathcal{N}}_\ell(h) := \left\|\left(\sum_{j=0}^{\ell-1}Q_j(\cdot;s_1,p_1)^{q_1}\right)^{1/q_1}\right\|^{p_1}_{L^{p_1}(E_\ell)},\quad\text{and}\quad \mathcal{M}_\ell := \Big\|\sup_{j\geqslant 0}Q_j(\cdot;s,p)\Big\|^p_{L^p(E_\ell)}.$$

Our next aim is to show that

$$\mathcal{N}_\ell(h)\leqslant c\mathcal{M}_\ell(h)\quad\text{and}\quad \widetilde{\mathcal{N}}_\ell(h)\leqslant c\mathcal{M}_\ell(h),\qquad \ell=1,2,\dots. \tag{7.15}$$

Assuming that the above estimates hold we obtain

$$\|h\|^{p_1}_{f^{s_1s_1}_{p_1q_1}} \leqslant c\sum_{\ell=1}^{\infty}(\mathcal{N}_\ell(h)+\widetilde{\mathcal{N}}_\ell(h)) \leqslant c\sum_{\ell=1}^{\infty}\mathcal{M}_\ell(h)\leqslant c\|h\|^p_{f^{ss}_{p\infty}}\leqslant c.$$

Therefore, (7.15) yields (7.14) and, hence, the claimed embedding result.

To prove that $\mathcal{N}_\ell(h) \leqslant c\mathcal{M}_\ell(h)$, we first observe that if $j \geqslant \ell$ and $R_\xi \cap E_\ell \neq \emptyset$, then

$$W_\mu(2^j;\xi) := \left(\sqrt{1-|\xi|^2}+2^{-j}\right)^{2\mu} \sim 2^{-2\mu\ell} = \delta^{2\mu}, \quad \text{where } \delta := 2^{-\ell},$$

and $w_\mu(x) := (1-|x|^2)^{\mu-1/2} \sim \delta^{2\mu-1}$ for $x \in E_\ell$. We use these and the fact that $\tilde{\mathbb{1}}_{R_\xi} := m(R_\xi)^{-1/2}\mathbb{1}_{R_\xi}$, where $m(R_\xi) \sim 2^{-jd}W_\mu(2^j;\xi)$ (see (2.27)), to obtain for $j \geqslant \ell$

$$\begin{aligned} 1 = \|h\|_{f^{ss}_{p\infty}} &\geqslant \|Q_j(\cdot;s,p)\|_{L^p(E_{\ell-1}\cup E_\ell \cup E_{\ell+1})} \\ &\geqslant c2^{j(s-d/p+d/2)}\delta^{2\mu(-s/d+1/p-1/2)}\left(\sum_{\xi\in\mathcal{X}_j, R_\xi\cap E_\ell\neq\emptyset}|h_\xi|^p\right)^{1/p}. \end{aligned}$$

Hence

$$|h_\xi| \leqslant c2^{j(-s+d/p-d/2)}\delta^{2\mu(s/d-1/p+1/2)}, \quad \text{if } \xi\in\mathcal{X}_j,\ j\geqslant \ell, \text{ and } R_\xi\cap E_\ell\neq\emptyset. \tag{7.16}$$

Using that $W_\mu(2^j,\xi) \sim \delta^{2\mu}$ and $w_\mu(x) \sim \delta^{2\mu-1}$ we have for $x \in E_\ell$ and $k \geqslant \ell$

$$\sum_{j=\ell}^k Q_j(x;s_1,p_1)^{q_1} \leqslant c\sum_{j=\ell}^k\left(2^{j(s_1+d/2)}\delta^{2\mu(-s_1/d+1/p_1-1/2)-1/p_1}\sum_{\xi\in\mathcal{X}_j, R_\xi\cap E_\ell\neq\emptyset}|h_\xi|\mathbb{1}_{R_\xi}(x)\right)^{q_1}.$$

Combining this with (7.16) we obtain

$$\sum_{j=\ell}^k Q_j(x;s_1,p_1)^{q_1} \leqslant c\sum_{j=\ell}^k 2^{j(s_1-s+d/p)q_1}\delta^{-q_1/p_1} \leqslant c_*2^{kdq_1/p_1}\delta^{-q_1/p_1}. \tag{7.17}$$

In a similar fashion one has for $x\in E_\ell$ and $k\geqslant \ell-1$

$$\begin{aligned} \sum_{j=k+1}^\infty Q_j(x;s_1,p_1)^{q_1} &\leqslant c\sum_{j=k+1}^\infty 2^{j(s_1-s)q_1}\delta^{(1/p-1/p_1)q_1}Q_j(x;s,p)^{q_1} \\ &\leqslant c2^{k(s_1-s)q_1}\delta^{(1/p-1/p_1)q_1}\sup_{j>k}Q_j(x;s,p)^{q_1}. \end{aligned} \tag{7.18}$$

Recall the well-known representation

$$\|f\|^p_{L^p(E)} = p\int_0^\infty t^p|\{x\in E : |f(x)|>t\}|\frac{dt}{t}.$$

Hence

$$\begin{aligned} \mathcal{N}_\ell(h) &= p_1\int_0^\infty t^{p_1}\left|\left\{x\in E_\ell : \left(\sum_{j=\ell}^\infty Q_j(x;s_1,p_1)^{q_1}\right)^{1/q_1} > t\right\}\right|\frac{dt}{t} \\ &= p_1\int_0^{(2c_*)^{1/q_1}\delta^{-1/p_1}}\ldots + p_1\int_{(2c_*)^{1/q_1}\delta^{-1/p_1}}^\infty \ldots =: I_1+I_2. \end{aligned}$$

To estimate I_1 we use (7.18) with $k=\ell-1$. We obtain

$$\begin{aligned} I_1 &\leqslant c\int_0^{c\delta^{-1/p_1}} t^{p_1}\left|\left\{x\in E_\ell : c\delta^{1/p-1/p_1}\sup_{j>k}Q_j(x;s,p)>t\right\}\right|\frac{dt}{t} \\ &\leqslant c\delta^{(1/p-1/p_1)p_1}\int_0^{c\delta^{-1/p}} u^{p_1}\left|\left\{x\in E_\ell : \sup_{j>k}Q_j(x;s,p)>u\right\}\right|\frac{du}{u} \\ &\leqslant c\delta^{(1/p-1/p_1)p_1}\delta^{-1/p(p_1-p)}\int_0^\infty u^p\left|\left\{x\in E_\ell : \sup_{j>k}Q_j(x;s,p)>u\right\}\right|\frac{du}{u} \\ &\leqslant cM_\ell(h). \end{aligned}$$

Here in the second estimate we used the substitution $t = cu\delta^{1/p-1/p_1}$ and for the third we used that $u^{p_1} \leqslant cu^p\delta^{-1/p(p_1-p)}$.

We now estimate I_2. For $t > (2c_*)^{1/q_1}\delta^{-1/p_1}$ we choose $k \geqslant 0$ in (7.17) to be the largest integer such that $c_* 2^{kdq_1/p_1}\delta^{-q_1/p_1} \leqslant t^{q_1}/2$. Suppose that $k \geqslant \ell$. Then

$$\left\{x \in E_\ell : \left(\sum_{j=\ell}^{\infty} Q_j(x; s_1, p_1)^{q_1}\right)^{1/q_1} > t\right\} \subset \left\{x \in E_\ell : \sum_{j=\ell}^{k} Q_j(x; s_1, p_1)^{q_1} > \frac{t^{q_1}}{2}\right\}$$
$$\cup \left\{x \in E_\ell : \sum_{j=k+1}^{\infty} Q_j(x; s_1, p_1)^{q_1} > \frac{t^{q_1}}{2}\right\} =: A_1 \cup A_2.$$

By (7.17) and the selection of k it follows that $A_1 = \emptyset$. Consequently, using (7.18) we have

$$I_2 \leqslant c \int_{(2c_*)^{1/q_1}\delta^{-1/p_1}}^{\infty} t^{p_1} \left|\left\{x \in E_\ell : \delta^{1/p-1/p_1} \sup_{j>k} Q_j(x; s, p) > ct2^{-k(s_1-s)}\right\}\right| \frac{dt}{t}.$$

From the selection of k we have $2^{kd/p_1}\delta^{-1/p_1} \sim t$ and a little algebra shows that

$$t2^{-k(s_1-s)} \sim t^{p_1/p}\delta^{1/p-1/p_1}.$$

Therefore,

$$\begin{aligned} I_2 &\leqslant c \int_0^{\infty} t^{p_1} \left|\left\{x \in E_\ell : \sup_{j>k} Q_j(x; s, p) > c' t^{p_1/p}\right\}\right| \frac{dt}{t} \\ &\leqslant c \int_0^{\infty} u^{p} \left|\left\{x \in E_\ell : \sup_{j\geqslant 0} Q_j(x; s, p) > u\right\}\right| \frac{du}{u} = c\mathcal{M}_\ell(h), \end{aligned}$$

where we used the substitution $u = c' t^{p_1/p}$.

The case when $k < \ell$ is simpler. Then $c_* 2^{\ell dq_1/p_1}\delta^{-q_1/p_1} \geqslant t^{q_1}/2$, and hence $2^{\ell d/p_1}\delta^{-1/p_1} \geqslant ct$. One estimates I_2 as above with k replaced by ℓ. The result is again $I_2 \leqslant c\mathcal{M}_\ell(h)$.

The above estimates for I_1 and I_2 imply that $\mathcal{N}_\ell(h) \leqslant c\mathcal{M}_\ell(h)$.

To prove that $\widetilde{\mathcal{N}}_\ell(h) \leqslant c\mathcal{M}_\ell(h)$ we proceed quite as above. Note first that if $R_\xi \cap E_\ell \neq \emptyset$, $\xi \in \mathcal{X}_j$, $j < \ell$, then

$$W_\mu(2^j; \xi) := (\sqrt{1-|\xi|^2} + 2^{-j})^{2\mu} \sim 2^{-2j\mu},$$

and $w_\mu(x) := (1-|x|^2)^{\mu-1/2} \sim 2^{-\ell(2\mu-1)}$ for $x \in E_\ell$. Hence for $j < \ell$

$$\begin{aligned} 1 = \|h\|_{f^{ss}_{p\infty}} &\geqslant \|Q_j(\cdot; s, p)\|_{L^p(E_{\ell-1}\cup E_\ell \cup E_{\ell+1})} \\ &\geqslant c2^{j(s-d/p+d/2)}2^{-2j\mu(-s/d+1/p-1/2)} \left(\sum_{\xi\in\mathcal{X}_j, R_\xi\cap E_\ell\neq\emptyset} |h_\xi|^p\right)^{1/p}, \end{aligned}$$

which implies that

$$|h_\xi| \leqslant c2^{j(-s+d/p-d/2)}2^{2j\mu(-s/d+1/p-1/2)}, \quad \text{if } \xi \in \mathcal{X}_j,\ j < \ell, \text{and } R_\xi \cap E_\ell \neq \emptyset.$$

Therefore, for $x \in E_\ell$ and $k < \ell$ we have

$$\begin{aligned}\sum_{j=0}^{k} Q_j(x;s_1,p_1)^{q_1} &\leqslant c\sum_{j=0}^{k}\left(2^{j(s_1+d/2)}2^{-2j\mu(-s_1/d-1/2)}2^{-\ell(2\mu-1)1/p_1}\sum_{\xi\in\mathcal{X}_j}|h_\xi|\mathbb{1}_{R_\xi}(x)\right)^{q_1}\\ &\leqslant c\sum_{j=0}^{k}2^{j(s_1+d/2-s+d/p-d/2)q_1}2^{-2j\mu(-s_1/d-1/2+s/d-1/p+1/2)q_1}2^{-\ell(2\mu-1)q_1/p_1}\\ &\leqslant c\sum_{j=0}^{k}2^{j(d+2\mu)q_1/p_1}2^{-\ell(2\mu-1)q_1/p_1}\leqslant c_*2^{k(d+2\mu)q_1/p_1}2^{-\ell(2\mu-1)q_1/p_1}.\end{aligned} \tag{7.19}$$

Assuming that $R_\xi \cap E_\ell \neq \emptyset$, where $\xi \in \mathcal{X}_j$, $j < \ell$, and $x \in E_\ell$, we denote briefly by

$$U(s,p) := W_\mu(2^j;\xi)^{-s/d}w_\mu(x)^{1/p}.$$

It is readily seen that

$$U(s_1,p_1) = U(s,p)2^{-2j\mu(1/p-1/p_1)}2^{-\ell(2\mu-1)(1/p-1/p_1)}.$$

We use this to obtain for $x \in E_\ell$ and $-1 \leqslant k < \ell$

$$\begin{aligned}\sum_{j=k+1}^{\ell} Q_j(x;s_1,p_1)^{q_1} &\leqslant c\sum_{j=k+1}^{\ell}2^{j[s_1-s-2\mu(1/p-1/p_1)]q_1}2^{\ell(2\mu-1)(1/p-1/p_1)q_1}\sup_{j>k}Q_j(x;s,p)^{q_1}\\ &\leqslant c2^{-k(d+2\mu)(1/p-1/p_1)q_1}2^{\ell(2\mu-1)(1/p-1/p_1)q_1}\sup_{j>k}Q_j(x;s,p)^{q_1}.\end{aligned} \tag{7.20}$$

Denote briefly $D := 2^{\ell(2\mu-1)}$. Just as for $\mathcal{N}_\ell(h)$ we have

$$\begin{aligned}\widetilde{\mathcal{N}}_\ell(h) &= p_1\int_0^\infty t^{p_1}\left|\left\{x\in E_\ell : \left(\sum_{j=0}^{\ell-1}Q_j(x;s_1,p_1)^{q_1}\right)^{1/q_1} > t\right\}\right|\frac{dt}{t}\\ &= p_1\int_0^{(2c_*)^{1/q_1}D^{-1/p_1}}\dots + p_1\int_{(2c_*)^{1/q_1}D^{-1/p_1}}^\infty\dots =: I_1 + I_2.\end{aligned}$$

As above we use (7.20) with $k = -1$ to obtain

$$\begin{aligned}I_1 &\leqslant c\int_0^{cD^{-1/p_1}} t^{p_1}\left|\left\{x\in E_\ell : cD^{1/p-1/p_1}\sup_{j\geqslant 0}Q_j(x;s,p) > t\right\}\right|\frac{dt}{t}\\ &\leqslant cD^{(\frac{1}{p}-\frac{1}{p_1})p_1}\int_0^{cD^{-1/p}} u^{p_1}\left|\left\{x\in E_\ell : \sup_{j\geqslant 0}Q_j(x;s,p) > u\right\}\right|\frac{du}{u}\\ &\leqslant cD^{(\frac{1}{p}-\frac{1}{p_1})p_1}D^{-\frac{1}{p}(p_1-p)}\int_0^\infty u^p\left|\left\{x\in E_\ell : \sup_{j\geqslant 0}Q_j(x;s,p) > u\right\}\right|\frac{du}{u}\\ &\leqslant cM_\ell(h).\end{aligned}$$

For $t > (2c_*)^{1/q_1} D^{-1/p_1}$ we select $k \geqslant 0$ in (7.19) to be the largest integer such that $c_* 2^{k(d+2\mu)q_1/p_1} D^{-q_1/p_1} \leqslant t^{q_1}/2$. Suppose that $k < \ell$. Then

$$\left\{ x \in E_\ell : \left(\sum_{j=0}^{\ell} Q_j(x; s_1, p_1)^{q_1} \right)^{1/q_1} > t \right\} \subset \left\{ x \in E_\ell : \sum_{j=0}^{k} Q_j(x; s_1, p_1)^{q_1} > \frac{t^{q_1}}{2} \right\}$$
$$\cup \left\{ x \in E_\ell : \sum_{j=k+1}^{\ell} Q_j(x; s_1, p_1)^{q_1} > \frac{t^{q_1}}{2} \right\} =: A_1 \cup A_2.$$

By (7.19) and the selection of k it follows that $A_1 = \emptyset$. Consequently, using (7.20) we obtain

$$I_2 \leqslant c \int_{(2c_*)^{1/q_1} D^{-1/p_1}}^{\infty} t^{p_1} \left| \left\{ x \in E_\ell : D^{1/p-1/p_1} \sup_{j>k} Q_j(x; s, p) > ct 2^{k(d+2\mu)(1/p-1/p_1)} \right\} \right| \frac{dt}{t}.$$

From the selection of k we have $2^{k(d+2\mu)(1/p-1/p_1)} D^{-1/p_1} \sim t$ and a simple manipulation shows that $t 2^{k(d+2\mu)(1/p-1/p_1)} \sim t^{p_1/p} D^{1/p-1/p_1}$. Consequently,

$$\begin{aligned} I_2 &\leqslant c \int_0^\infty t^{p_1} \left| \left\{ x \in E_\ell : \sup_{j>k} Q_j(x; s, p) > c' t^{p_1/p} \right\} \right| \frac{dt}{t} \\ &\leqslant c \int_0^\infty u^p \left| \left\{ x \in E_\ell : \sup_{j \geqslant 0} Q_j(x; s, p) > u \right\} \right| \frac{du}{u} = c\mathcal{M}_\ell(h). \end{aligned}$$

The case when $k \geqslant \ell$ is trivial since in this case $A_1 = A_2 = \emptyset$ and hence $I_2 = 0$.

The above estimates for I_1 and I_2 yield $\widetilde{\mathcal{N}}_\ell(h) \leqslant c\mathcal{M}_\ell(h)$. Thus (7.15) is established. □

Proof of Lemma 5.5. For any $\xi \in \mathcal{X}_j$, we denote $a_\xi := \max_{x \in R_\xi} |P(x)|$,

$$m_\xi := \min_{x \in R_\xi} |P(x)|, \quad \text{and} \quad b_\xi := \max\{ \min_{x \in R_w} |P(x)| : w \in \mathcal{X}_{j+r}, R_w \cap R_\xi \neq \emptyset \},$$

where $r \geqslant 1$ is the constant from Lemma 4.9.

Choose $0 < t < p$. By Lemma 4.9 we have $a_\xi^* \leqslant c b_\xi^*$. We use this, Lemmas 4.10, and the maximal inequality (2.17) to obtain

$$\begin{aligned} \left(\sum_{\xi \in \mathcal{X}_j} W_\mu(2^j; \xi)^\gamma a_\xi^p m(R_\xi) \right)^{1/p} &= \left\| \sum_{\xi \in \mathcal{X}_j} W_\mu(2^j; \xi)^\gamma a_\xi \mathbb{1}_{R_\xi}(\cdot) \right\|_p \\ \leqslant c \left\| \sum_{\xi \in \mathcal{X}_j} W_\mu(2^j; \xi)^\gamma b_\xi^* \mathbb{1}_{R_\xi}(\cdot) \right\|_p &\leqslant c \left\| \mathcal{M}_t \left(\sum_{\xi \in \mathcal{X}_j} W_\mu(2^j; \xi)^\gamma b_\xi \mathbb{1}_{R_\xi} \right)(\cdot) \right\|_p \\ &\leqslant c \left\| \sum_{\xi \in \mathcal{X}_j} W_\mu(2^j; \xi)^\gamma b_\xi \mathbb{1}_{R_\xi}(\cdot) \right\|_p . \end{aligned} \tag{7.21}$$

Now, exactly as in the proof of Theorem 4.5 (see (4.20)) we have

$$b_\xi W_\mu(2^j; \xi)^\gamma \mathbb{1}_{R_\xi} \leqslant \sum_{\eta \in \mathcal{X}_{j+r}(\xi)} m_\eta^* W_\mu(2^{j+r}; \eta)^\gamma \mathbb{1}_{R_\eta}, \tag{7.22}$$

where $\mathcal{X}_{j+r}(\xi) := \{w \in \mathcal{X}_{j+r} : R_w \cap R_\xi \neq \emptyset\}$. Combining this with (7.21) and using that $\#\mathcal{X}_{j+r}(\xi) \leqslant c$, Lemmas 4.10, and the maximal inequality (2.17), we obtain

$$\begin{aligned}\Big(\sum_{\xi\in\mathcal{X}_j} W_\mu(2^j;\xi)^\gamma a_\xi^p m(R_\xi)\Big)^{1/p} &\leqslant c\Big\|\sum_{\eta\in\mathcal{X}_{j+r}} m_\eta^* W_\mu(2^{j+r};\eta)^\gamma \mathbb{1}_{R_\eta}(\cdot)\Big\|_p\\ &\leqslant c\Big\|\mathcal{M}_t\Big(\sum_{\eta\in\mathcal{X}_{j+r}} m_\eta W_\mu(2^{j+r};\eta)^\gamma \mathbb{1}_{R_\eta}\Big)(\cdot)\,\|_p \leqslant c\|\sum_{\eta\in\mathcal{X}_{j+r}} m_\eta W_\mu(2^{j+r};\eta)^\gamma \mathbb{1}_{R_\eta}(\cdot)\Big\|_p\\ &\leqslant c\|P\|_p.\end{aligned}$$

Here for the fourth inequality we used that $W_\mu(2^{j+r};\eta) \sim W_\mu(2^j;x)$ if $x \in R_\eta$, $\eta \in \mathcal{X}_{j+r}$. $\square$

Acknowledgement. The authors are indebted to the anonymous referee for pointing out a gap in a proof in an earlier version of this paper and valuable suggestions for improvements.

References

1. R. Askey, *Orthogonal polynomials and special functions*, Regional Conference Series in Applied Mathematics 21 (SIAM, Philadelphia, 1975).

2. Feng Dai and Yuan Xu, 'Maximal function and multiplier theorem for weighted space on the unit sphere', *J. Funct. Anal.* 249 (2007) 477–504.

3. C. Dunkl and Yuan Xu, *Orthogonal polynomials of several variables*, Encyclopedia of Mathematics and its Applications 81 (Cambridge University Press, Cambridge, 2001).

4. A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher transcendental functions*, vols I, II (McGraw-Hill, New York, 1953).

5. M. Frazier and B. Jawerth, 'Decomposition of Besov spaces', *Indiana Univ. Math. J.* 34 (1985) 777–799.

6. M. Frazier and B. Jawerth, 'A discrete transform and decompositions of distribution', *J. Funct. Anal.* 93 (1990) 34–170.

7. M. Frazier, B. Jawerth and G. Weiss, *Littlewood–Paley theory and the study of function spaces*, CBMS Regional Conference Series in Mathematics 79 (American Mathematical Society, Providence, RI, 1991).

8. O. D. Kellogg, 'On bounded polynomials of several variables', *Math. Z* 27 (1928) 5564.

9. G. Kyriazis, P. Petrushev and Yuan Xu, 'Jacobi decomposition of weighted Triebel–Lizorkin and Besov spaces', *Studia Math.* (to appear) (ArXiv:math.CA/0610624).

10. Y. Meyer, *Ondelletes et opérateurs I: Ondelletes* (Hermann, Paris, 1990).

11. F. Narcowich, P. Petrushev and J. Ward, 'Localized tight frames on spheres', *SIAM J. Math. Anal.* 38 (2006) 574–592.

12. F. Narcowich, P. Petrushev and J. Ward, 'Decomposition of Besov and Triebel–Lizorkin spaces on the sphere', *J. Funct. Anal.* 238 (2006) 530–564.

13. P. Petrushev and Yuan Xu, 'Localized polynomial frames on the interval with Jacobi weights', *J. Fourier Anal. Appl.* 11 (2005) 557–575.

14. P. Petrushev and Yuan Xu, 'Localized polynomial frames on the ball', *Constr. Approx.* 27 (2008) 121–148.

15. J. Peetre, *New thoughts on Besov spaces*, Duke University Mathematics Series 1 (Durham, NC, 1976).

16. E. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals* (Princeton University Press, Princeton, NJ, 1993).

17. G. Szegő, *Orthogonal polynomials*, 4th edn, American Mathematical Society Colloquium Publications 23 (Providence, RI, 1975).

18. H. Triebel, *Theory of function spaces*, Monographs in Mathematics 78 (Birkhäuser, Basel, 1983).

19. Yuan Xu, 'Summability of Fourier orthogonal series for Jacobi weight on a ball in $\mathbb{R}^d$', *Trans. Amer. Math. Soc.* 351 (1999) 2439–2458.

20. Yuan Xu, 'Weighted approximation of functions on the unit sphere', *Constr. Approx.* 21 (2005) 1–28.

G. Kyriazis
Department of Mathematics and Statistics
University of Cyprus
1678 Nicosia
Cyprus

kyriazis@ucy.ac.cy

P. Petrushev
Department of Mathematics
University of South Carolina
Columbia, SC 29205
USA

pencho@math.sc.edu

Yuan Xu
Department of Mathematics
University of Oregon
Eugene, Oregon 97403
USA

yuan@math.uoregon.edu