Constructive Approximation (2021) 53:39–83
https://doi.org/10.1007/s00365-019-09490-1

CONSTRUCTIVE APPROXIMATION

Product Besov and Triebel–Lizorkin Spaces with Application to Nonlinear Approximation

Athanasios G. Georgiadis[1] · George Kyriazis[1] · Pencho Petrushev[2]

Received: 26 February 2019 / Revised: 7 October 2019 / Accepted: 22 October 2019 /
Published online: 23 December 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

The Littlewood–Paley theory of homogeneous product Besov and Triebel–Lizorkin spaces is developed in the spirit of the φ-transform of Frazier and Jawerth. This includes the frame characterization of the product Besov and Triebel–Lizorkin spaces and the development of almost diagonal operators on these spaces. The almost diagonal operators are used to obtain product wavelet decomposition of the product Besov and Triebel–Lizorkin spaces. The main application of this theory is to nonlinear m-term approximation from product wavelets in L^p and Hardy spaces. Sharp Jackson and Bernstein estimates are obtained in terms of product Besov spaces.

Keywords Product spaces · Besov spaces · Triebel–Lizorkin spaces · φ-Transform · Wavelets · Nonlinear approximation · Jackson estimate · Bernstein estimate

Mathematics Subject Classification 42B25 · 42B35 · 46F10 · 4117

Communicated by Wolfgang Dahmen.

Pencho Petrushev has been supported by NSF Grant DMS-1714369.

✉ George Kyriazis
kyriazis@ucy.ac.cy

Athanasios G. Georgiadis
georgiades.athanasios@ucy.ac.cy

Pencho Petrushev
pencho@math.sc.edu

1 Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus

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

1 Introduction

The purpose of this article is to develop nonlinear approximation from product wavelets in L^p and Hardy (H^p) spaces. Homogeneous product Besov and Triebel–Lizorkin spaces naturally appear in this theory. The inhomogeneous product Besov and Triebel–Lizorkin spaces are part of the biparameter or multiparameter harmonic analysis and have been developed for quite sometime. For this theory we refer the reader to [23,24,29] and the references therein. It has been developed analogously to the classical (one parameter) Besov and Triebel–Lizorkin spaces, see [22,26–28]. The theory of homogeneous product Besov and Triebel–Lizorkin spaces, however, seems underdeveloped.

In this paper we develop rapidly the Littlewood–Paley theory of homogeneous product Besov spaces $\dot{B}^{\mathbf{s}}_{pq}$ and Triebel–Lizorkin spaces $\dot{F}^{\mathbf{s}}_{pq}$ on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, n_1, n_2 \geq 1$, where the smoothness parameter $\mathbf{s}$ is a vector in $\mathbb{R}^2$ and $0 < p, q \leq \infty$ are as in the classical one parameter case. To this end we introduce product frames and utilize them for characterization of the product Besov and Triebel–Lizorkin spaces in the spirit of the φ-transform of Frazier and Jawerth [11–14]. Almost diagonal operators are also introduced, and their boundedness on the respective product Besov and Triebel–Lizorkin sequence spaces is obtained. In turn, the almost diagonal operators are used for establishing the product wavelet characterization of the product spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ that is also a focal point in this study.

One should not think that all results for product spaces can be obtained by iterating ideas from the one parameter setting. The product Hardy spaces provide a typical example where one parameter ideas do not work. There are a number of papers on this subject that reinforce this claim. For example the atomic decomposition of the product Hardy spaces H^p takes more complicated form than in the classical one parameter theory. We refer the reader to [1,2,9,10,16,18,19] and the references therein for the theory of product Hardy spaces.

The main focus of this article is on nonlinear m-term approximation from product wavelets and product frames in $L^p(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, $1 < p < \infty$, and in the product Hardy spaces $H^p(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, $0 < p \leq 1$. The product wavelets used here are constructed based on the univariate Meyer's wavelets, but our theory is not restricted to such wavelets only. Product Besov spaces are naturally involved in the approximation process. More explicitly these are the Besov spaces $\dot{B}^{\mathbf{s}}_{\tau\tau} := \dot{B}^{(s_1,s_2)}_{\tau\tau}$ with smoothness $s_1 := n_1(1/\tau - 1/p)$ and $s_2 := n_2(1/\tau - 1/p)$, where $1/\tau := \alpha + 1/p$ and $\alpha > 0$ is the parameter that determines the rate of approximation; α can be arbitrarily large. Jackson and Bernstein estimates for product wavelet nonlinear approximation are established that allow almost complete characterization of the rates of approximation. To be more specific, denote by $\sigma_m(f)_p$ the best m-term approximation of f from product wavelets in L^p if $1 < p < \infty$ or in the Hardy space H^p if $0 < p \leq 1$ on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Our Jackson estimates (Theorem 7.2) assert that for any $f \in \dot{B}^{\mathbf{s}}_{\tau\tau}$, we have

$$\sigma_m(f)_p \leq cm^{-\alpha}\|f\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}, \quad p \leq 2, \quad \text{and}$$
$$\sigma_m(f)_p \leq cm^{-\alpha}(\ln m)^{1/2-1/p}\|f\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}, \quad p > 2.$$

Almost matching companion Bernstein estimates (Theorem 7.3) are also established. Both estimates above are sharp, but there is a logarithmic factor in the estimate on the right that is a reflection of the fact that the "essential supports" of a (logarithmic) number of wavelets of the same size overlap at any point. The establishment of these results (in particular, Lemma 7.9) exhibits the difficulties one has to overcome when working in the product space setting.

Similar results are valid in the inhomogeneous setting and involve the respective inhomogeneous Besov spaces. But to keep the size of this article limited we will not elaborate on these sorts of results.

We would like to point out that the linear approximation from product wavelets has been studied in [7].

The organization of the paper is as follows. In Sect. 2 we introduce our notation and collect all basic facts that are needed for the development of our theory. The homogeneous product Besov and Triebel–Lizorkin spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ are defined in Sect. 3, and some of their basic features are discussed. In Sect. 4, we establish the frame characterization of the product spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ in the spirit of the φ-transform of Frazier and Jawerth. Almost diagonal operators are introduced in Sect. 5 and their boundedness on product Besov and Triebel–Lizorkin sequence spaces $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}$ and $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}$ is established. Product wavelets are introduced in Sect. 6, and the wavelet characterization of the product spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ is established. The nonlinear approximation theory from product wavelets and frames is developed in Sect. 7. Section 8 is an appendix where the proofs of some claims from previous sections are placed.

2 Preliminaries

In this section we present the background material we need for the development of the homogeneous product Besov and Triebel–Lizorkin spaces.

2.1 Notation

The action will be in the product space $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $n_1, n_2 \in \mathbb{N}$. We first introduce some convenient notation in the single parameter case of $\mathbb{R}^n$. For any $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ and $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ ($\mathbb{N}_0 := \mathbb{N} \cup \{0\}$), we write $x^\alpha := (x^1)^{\alpha_1} \cdots (x^n)^{\alpha_n}$, $|\alpha| := \alpha_1 + \cdots + \alpha_n$, and $\partial^\alpha := \big(\frac{\partial}{\partial x^1}\big)^{\alpha_1} \dots \big(\frac{\partial}{\partial x^n}\big)^{\alpha_n}$. The inner product of $x, y \in \mathbb{R}^n$ will be denoted by $x \cdot y$. However, $|x|$ will stand for the ℓ^∞ norm of x, i.e. $|x| := \max\{|x^j|: j = 1, \dots, n\}$. The Fourier transform $\widehat{f}$ of a function f on $\mathbb{R}^n$ is defined by $\widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi} \mathrm{d}x$.

We will use the notation $\mathbf{x} := (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $x_i := (x_i^1, \dots, x_i^{n_i}) \in \mathbb{R}^{n_i}$, $i = 1, 2$. The ℓ^∞ norm of $x_i \in \mathbb{R}^{n_i}$ will be denoted by $|x_i|$ and $|\mathbf{x}| := \max\{|x_1|, |x_2|\}$. Also, $\mathbf{x} \cdot \mathbf{y}$ will stand for the inner product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. For a set $A \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ or $A \subset \mathbb{R}^{n_i}$, $i = 1, 2$, we denote by $|A|$ its Lebesgue measure and $\mathbb{1}_A$ will stand for its characteristic function. We will denote by $\|\cdot\|_p = \|\cdot\|_{L^p}$ the L^p-norm on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.

Positive constants will be denoted by c, c_1, etc., and they may vary at every occurrence. The notation $a \sim b$ will stand for $c_1 \le a/b \le c_2$.

Remark 2.1 A word of caution is in order. As was indicated above, we denote by $|x_i|$ or $|\mathbf{x}|$ the ℓ^∞-norm of $x_i \in \mathbb{R}^{n_i}$ $(i = 1, 2)$ or $\mathbf{x} \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. We use these norms throughout because they are well aligned with the product structure of the spaces in this article.

2.2 The Classes $\mathcal{S}_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ and $\mathcal{S}'_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$

In the single parameter case on $\mathbb{R}^n$, the Schwartz class $\mathcal{S}(\mathbb{R}^n)$ is defined as the set of all $C^\infty(\mathbb{R}^n)$ rapidly decreasing functions with topology induced by the family of semi-norms $\|\cdot\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta(\cdot)(x)|$, $\alpha, \beta \in \mathbb{N}_0^n$. Then the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ is its topological dual.

Let now $\mathcal{S} := \mathcal{S}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ be the Schwartz space on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. It is easily seen that if $\phi \in \mathcal{S}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, then for any $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $M_1, M_2 > 0$, and $\alpha \in \mathbb{N}_0^{n_1+n_2}$, there exists a constant $c = c(N_1, N_2, \alpha) > 0$ such that

$$|\partial^\alpha \phi(\mathbf{x})| \le c \prod_{i=1}^{2} \left(1 + |x_i|\right)^{-M_i}. \tag{2.1}$$

The dual of $\mathcal{S}$ is the space of tempered distributions $\mathcal{S}' := \mathcal{S}'(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$.

We define $\mathcal{S}_\infty := \mathcal{S}_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ as the subspace of $\mathcal{S}$ consisting of all functions $\phi \in \mathcal{S}$ that in addition satisfy

$$\int_{\mathbb{R}^{n_i}} x_i^{\nu_i} \phi(x_1, x_2) \mathrm{d}x_i = 0, \quad \forall \nu_i \in \mathbb{N}_0^{n_i}, \ i = 1, 2. \tag{2.2}$$

Note that $\mathcal{S}_\infty$ is a closed subspace of $\mathcal{S}$ and therefore complete.

Since the Fourier transform is a continuous linear transformation from $\mathcal{S}$ onto $\mathcal{S}$, it is not hard to see that the topology in $\mathcal{S}_\infty$ can be generated by the family of semi-norms

$$\|\phi\|_M^* := \sup_{|\alpha|\le M} \sup_{\boldsymbol{\xi}\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}} |\partial^\alpha \widehat{\phi}(\boldsymbol{\xi})| \prod_{i=1}^{2} \left(|\xi_i|^M + |\xi_i|^{-M}\right), \quad M \in \mathbb{N}_0.$$

Therefore,

$$\mathcal{S}_\infty = \left\{\phi \in \mathcal{S} \colon \|\phi\|_M^* < \infty, \ \forall\, M \in \mathbb{N}_0\right\}.$$

The space $\mathcal{S}'_\infty := \mathcal{S}'_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ is the dual of $\mathcal{S}_\infty$ (with the week-$*$ topology); that is, $f \in \mathcal{S}'_\infty$ if and only if there exist constants $c > 0$ and $M \in \mathbb{N}_0$ such that

$$|\langle f, \phi\rangle| \le c\|\phi\|_M^*, \quad \forall \phi \in \mathcal{S}_\infty,$$

where $\langle f, \phi\rangle := f(\bar{\phi})$.

2.3 Calderón Reproducing Formula on $\mathcal{S}'_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$

Here we generalize the classical Calderón formula in the case of product spaces.

We begin by recalling the relevant decomposition results in the single parameter case. We consider two decomposition identities. For the first we assume that $\varphi \in \mathcal{S}(\mathbb{R}^n)$ is such that $\widehat{\varphi}$ is compactly supported, bounded away from the origin, and

$$\sum_{\nu\in\mathbb{Z}} \widehat{\varphi}(2^{-\nu}\xi) = 1, \quad \xi \in \mathbb{R}^n \backslash \{0\}. \tag{2.3}$$

Then for any $f \in \mathcal{S}_\infty(\mathbb{R}^n)$ (or $f \in \mathcal{S}'_\infty(\mathbb{R}^n)$),

$$f = \sum_{\nu\in\mathbb{Z}} \varphi_\nu * f, \tag{2.4}$$

where the convergence is in $\mathcal{S}$ (or in $\mathcal{S}'_\infty$) and we used the notation

$$f_\nu(x) := 2^{\nu n} f(2^\nu x)$$

for any function (or distribution) defined on $\mathbb{R}^n$.

The second version of Calderon's formula relies on a pair of functions φ, ψ that satisfy the conditions:

$$\begin{array}{ll} (i) & \varphi, \psi \in \mathcal{S}(\mathbb{R}^n), \\ (ii) & \operatorname{supp} \widehat{\varphi}, \widehat{\psi} \subset \{\xi \in \mathbb{R}^n \colon 2^{-1} \le |\xi| \le 2\}, \\ (iii) & |\widehat{\varphi}(\xi)| \ge c > 0 \ \text{ if } \ 2^{-3/4} \le |\xi| \le 2^{3/4}, \end{array} \tag{2.5}$$

and in addition,

$$\sum_{j\in\mathbb{Z}} \overline{\widehat{\varphi}(2^{-j}\xi)} \widehat{\psi}(2^{-j}\xi) = 1, \quad \xi \in \mathbb{R}^n \backslash \{0\}. \tag{2.6}$$

Then for any $f \in \mathcal{S}_\infty(\mathbb{R}^n)$ (or $f \in \mathcal{S}'_\infty(\mathbb{R}^n)$),

$$f = \sum_{j\in\mathbb{Z}} \tilde{\varphi}_j * \psi_j * f \quad \text{in } \mathcal{S} \text{ (or in } \mathcal{S}'_\infty), \tag{2.7}$$

where $\tilde{\varphi}_j(x) = \overline{\varphi_j(-x)}$.

The existence of pairs of functions φ, ψ obeying (2.5)–(2.6) is well known. In fact, for any φ satisfying conditions (2.5) there is ψ satisfying (2.5) such that (2.6) is valid. For details we refer the reader to [13,14]. Also, observe that (2.4) and (2.7) hold for any $f \in L^2(\mathbb{R}^n)$ with convergence in L^2.

To extend Calderon's formulas to the biparameter case we need some additional notation: For any function (or distribution) f defined on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \mathbb{Z}^2$, we define

$$f_{\boldsymbol{\nu}}(x_1, x_2) := 2^{\boldsymbol{\nu}\cdot\mathbf{n}} f\big(2^{\nu_1}x_1, 2^{\nu_2}x_2\big), \quad \boldsymbol{\nu}\cdot\mathbf{n} = \nu_1 n_1 + \nu_2 n_2. \tag{2.8}$$

Proposition 2.2 (a) *Let φ^1, φ^2 satisfy (2.3), (2.5) with $n = n_1, n_2$, respectively. Set $\varphi := \varphi^1 \otimes \varphi^2$. Then for any $f \in \mathcal{S}_\infty$ (or $f \in \mathcal{S}'_\infty$),*

$$f = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \varphi_{\boldsymbol{\nu}} * f \tag{2.9}$$

with convergence in $\mathcal{S}$ (or in $\mathcal{S}'_\infty$).

(b) *Let φ^1, ψ^1 and φ^2, ψ^2 be two pairs of functions satisfying (2.5)–(2.6) with $n = n_1, n_2$, respectively. Set $\varphi := \varphi^1 \otimes \varphi^2$ and $\psi := \psi^1 \otimes \psi^2$. Then for any $f \in \mathcal{S}_\infty$ (or $f \in \mathcal{S}'_\infty$),*

$$f = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \tilde{\psi}_{\boldsymbol{\nu}} * \varphi_{\boldsymbol{\nu}} * f \tag{2.10}$$

with convergence in $\mathcal{S}$ (or in $\mathcal{S}'_\infty$).

Furthermore, both (2.9) and (2.10) are valid for any $f \in L^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ with convergence in L^2.

For the proof of this proposition in the more general case of spaces with anisotropic dilations, see Lemma 3.15 and Proposition 3.16 in [1].

2.4 Construction of Frames on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$

We next introduce the product analog of the Frazier–Jawerth frames (the φ-transform). We denote by $\mathcal{D}^i$ the set of all dyadic cubes in $\mathbb{R}^{n_i}$ $(i = 1, 2)$ and by $\mathcal{D}^i_\nu$ $(\nu \in \mathbb{Z})$ the set of all cubes $I \in \mathcal{D}^i$ of side-length $\ell(I) = 2^{-\nu}$. For any $I_i \in \mathcal{D}^i$ we denote by x_{I_i} its lower-left corner.

Dyadic Rectangles Let $\mathcal{R}$ be the set of all dyadic rectangles in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Namely, this is the set of all rectangles $\mathbf{I}$ of the form $\mathbf{I} = I_1 \times I_2$ with $I_1 \in \mathcal{D}^1$, $I_2 \in \mathcal{D}^2$, and we set $\mathbf{x}_\mathbf{I} := (x_{I_1}, x_{I_2})$. Further, $\mathcal{R}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \mathbb{Z}^2$, will stand for the set of all rectangles $\mathbf{I} = I_1 \times I_2 \in \mathcal{R}$ such that $I_i \in \mathcal{D}^i_{\nu_i}$.

Given $f \in \mathcal{S}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ and $\mathbf{I} \in \mathcal{R}$, we set

$$f_\mathbf{I}(\mathbf{x}) := |\mathbf{I}|^{-1/2} f\Big(\frac{x_1 - x_{I_1}}{\ell(I_1)}, \frac{x_2 - x_{I_2}}{\ell(I_2)}\Big), \quad \mathbf{x} = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}. \tag{2.11}$$

As usual (2.11) is extended by duality to the corresponding spaces of tempered distributions.

With $\varphi := \varphi^1 \otimes \varphi^2$ and $\psi := \psi^1 \otimes \psi^2$ just as in Proposition 2.2 (b), we consider the systems

$$\{\varphi_\mathbf{I}\}_{\mathbf{I} \in \mathcal{R}} \quad \text{and} \quad \{\psi_\mathbf{I}\}_{\mathbf{I} \in \mathcal{R}}. \tag{2.12}$$

Proposition 2.3 *For any $f \in \mathcal{S}_\infty$ (or $f \in \mathcal{S}'_\infty$) we have*

$$f = \sum_{\boldsymbol{I} \in \mathcal{R}} \langle f, \varphi_{\boldsymbol{I}} \rangle \psi_{\boldsymbol{I}} \tag{2.13}$$

with convergence in $\mathcal{S}$ *(or in* $\mathcal{S}'_\infty$*). Also, representation* (2.13) *is valid for each* $f \in L^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ *with convergence in* L^2*; i.e.,* $\{\varphi_{\boldsymbol{I}}\}_{\boldsymbol{I}\in\mathcal{R}}$, $\{\psi_{\boldsymbol{I}}\}_{\boldsymbol{I}\in\mathcal{R}}$ *is a pair of dual frames.*

The decomposition (2.13) follows by Proposition 2.2 just as in the one parameter case, see [13].

We next give some properties of the frame elements. Clearly, φ and ψ obey (2.1), and hence for any $N_1, N_2 > 0$ there exists a constant $c > 0$ such that

$$|\varphi_{\mathbf{I}}(\mathbf{x})|, |\psi_{\mathbf{I}}(\mathbf{x})| \le c2^{\boldsymbol{\nu}\cdot\mathbf{n}/2}\prod_{i=1}^{2}(1+2^{\nu_i}|x_i - x_{I_i}|)^{-N_i}, \quad \mathbf{x} \in \mathbb{R}^{n_1}\times\mathbb{R}^{n_2},\ \mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}. \tag{2.14}$$

Also,

$$\int_{\mathbb{R}^{n_i}} x_i^\alpha \varphi_{\mathbf{I}}(x_1, x_2)\mathrm{d}x_i = 0, \quad \forall \alpha \in \mathbb{N}_0^{n_i},\ i = 1, 2. \tag{2.15}$$

Furthermore, if $\boldsymbol{\nu} \in \mathbb{Z}^2$ and $\mathbf{I} \in \mathcal{R}_{\boldsymbol{\nu}}$, then

$$\operatorname{supp}\widehat{\varphi_{\boldsymbol{\nu}}}, \operatorname{supp}\widehat{\varphi_{\mathbf{I}}} \subset \left\{\boldsymbol{\xi} \in \mathbb{R}^{n_1}\times\mathbb{R}^{n_2} \colon 2^{\nu_i - 1} \le |\xi_i| \le 2^{\nu_i+1},\ i = 1, 2\right\}, \tag{2.16}$$

and the same is true for $\operatorname{supp}\widehat{\psi_{\boldsymbol{\nu}}}$ and $\operatorname{supp}\widehat{\psi_{\mathbf{I}}}$.

It is sometimes beneficial to work with a single frame, i.e. to have $\varphi_{\mathbf{I}} = \psi_{\mathbf{I}}$. It is easy to construct (see, e.g., [13]) real-valued functions θ^1, θ^2 that satisfy conditions (2.5) on $\mathbb{R}^{n_1}, \mathbb{R}^{n_2}$ and such that

$$\sum_{j\in\mathbb{Z}} |\widehat{\theta^i}(2^{-j}\xi_i)|^2 = 1, \quad \xi_i \in \mathbb{R}^{n_i},\ i = 1, 2.$$

Set $\theta := \theta^1 \otimes \theta^2$. Then $\{\theta_{\mathbf{I}}\}_{\mathbf{I}\in\mathcal{R}}$ is a (tight) frame for $L^2(\mathbb{R}^{n_1}\times\mathbb{R}^{n_2})$, and for any $f \in \mathcal{S}_\infty$ (or $f \in \mathcal{S}'_\infty$) we have $f = \sum_{\mathbf{I}\in\mathcal{R}}\langle f, \theta_{\mathbf{I}}\rangle\theta_{\mathbf{I}}$.

2.5 The Strong Maximal Operator

The strong maximal operator is defined by

$$M_s f(\mathbf{x}) = \sup_{\mathbf{x}\in\mathbf{I}} \frac{1}{|\mathbf{I}|}\int_{\mathbf{I}} |f(\mathbf{y})|\mathrm{d}\mathbf{y}, \quad \mathbf{x} \in \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}, \tag{2.17}$$

where the sup is taken over all rectangles $\mathbf{I} = I_1 \times I_2 \subset \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ with sides parallel to the coordinate axes. It will be convenient to us to use the following modification of the strong maximal operator:

$$\mathcal{M}_t f(\mathbf{x}) := \left(M_s|f|^t(\mathbf{x})\right)^{1/t} = \sup_{\mathbf{x}\in\mathbf{I}}\left(\frac{1}{|\mathbf{I}|}\int_{\mathbf{I}}|f(\mathbf{y})|^t\mathrm{d}\mathbf{y}\right)^{1/t}, \quad t > 0. \tag{2.18}$$

The following version of the Fefferman–Stein *vector-valued maximal inequality* (see [25]) follows readily by applying the single parameter one twice: If $0 < p < \infty$,

$0 < q \le \infty$, and $0 < t < \min\{p, q\}$, then for any sequence of functions $\{f_{\boldsymbol{\nu}}\}$ on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$,

$$\Big\| \Big(\sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} (\mathcal{M}_t f_{\boldsymbol{\nu}})^q \Big)^{1/q} \Big\|_p \le c \Big\| \Big(\sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} |f_{\boldsymbol{\nu}}|^q \Big)^{1/q} \Big\|_p. \tag{2.19}$$

We will also need the following estimate on a *Peetre-type* maximal function:

Lemma 2.4 *Let $t, b_1, b_2 > 0$. Then there exists a constant $c = c(t, b_1, b_2) > 0$ such that for any function $f \in \mathcal{S}'(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ with* $\operatorname{supp} \widehat{f} \subset [-b_1, b_1]^{n_1} \times [-b_2, b_2]^{n_2}$ *we have*

$$\sup_{\boldsymbol{y} \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}} \frac{|f(\boldsymbol{y})|}{\prod_{i=1}^{2} \big(1 + b_i |x_i - y_i|\big)^{n_i/t}} \le c \mathcal{M}_t f(\boldsymbol{x}), \quad \boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}. \tag{2.20}$$

The above inequality is well known, see, e.g., [24, Theorem 1.6.4].

3 Homogeneous Product Besov and Triebel–Lizorkin Spaces

In this section we introduce the homogeneous product Besov and Triebel–Lizorkin spaces and list some of their basic properties.

Let $\varphi := \varphi^1 \otimes \varphi^2$, where each of the functions φ^1, φ^2 satisfies conditions (2.5).

Definition 3.1 (i) Let $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$ and $0 < p, q \le \infty$. The homogeneous product Besov space $\dot{B}^{\mathbf{s}}_{pq} := \dot{B}^{\mathbf{s}}_{pq}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ is defined as the set of all $f \in \mathcal{S}'_\infty$ such that

$$\|f\|_{\dot{B}^{\mathbf{s}}_{pq}} := \|f\|_{\dot{B}^{\mathbf{s}}_{pq}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})} := \Big(\sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \big(2^{\mathbf{s} \cdot \boldsymbol{\nu}} \|\varphi_{\boldsymbol{\nu}} * f\|_p \big)^q \Big)^{1/q} < \infty. \tag{3.1}$$

(ii) For $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$, $0 < q \le \infty$, and $0 < p < \infty$, the homogeneous product Triebel–Lizorkin space $\dot{F}^{\mathbf{s}}_{pq} := \dot{F}^{\mathbf{s}}_{pq}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ is defined as the set of all $f \in \mathcal{S}'_\infty$ such that

$$\|f\|_{\dot{F}^{\mathbf{s}}_{pq}} := \|f\|_{\dot{F}^{\mathbf{s}}_{pq}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})} := \Big\| \Big(\sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \big(2^{\mathbf{s} \cdot \boldsymbol{\nu}} |\varphi_{\boldsymbol{\nu}} * f| \big)^q \Big)^{1/q} \Big\|_p < \infty. \tag{3.2}$$

As usual the ℓ^q-norm above is replaced by the sup-norm if $q = \infty$.

Several remarks are in order.

(1) The definitions of the spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ are independent of the particular selection of the functions φ^1, φ^2, satisfying conditions (2.5). This follows similarly to the wavelet characterization of the spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ in Theorem 6.4 below.
(2) The homogeneous product Besov spaces $\dot{B}^{\mathbf{s}}_{pq}$ and Triebel–Lizorkin spaces $\dot{B}^{\mathbf{s}}_{pq}$ are continuously embedded in $\mathcal{S}'_\infty$.

(3) The homogeneous product Besov spaces $\dot{B}^{\mathbf{s}}_{pq}$ and Triebel–Lizorkin spaces $\dot{B}^{\mathbf{s}}_{pq}$ are (quasi-)Banach spaces.
(4) The space $\mathcal{S}_\infty$ is dense in $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ whenever $q < \infty$.
(5) The norms in the spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ do not recognize algebraic polynomials; that is, $\|f + P\|_{\dot{B}^{\mathbf{s}}_{pq}} = \|f\|_{\dot{B}^{\mathbf{s}}_{pq}}$ and $\|f + P\|_{\dot{F}^{\mathbf{s}}_{pq}} = \|f\|_{\dot{F}^{\mathbf{s}}_{pq}}$ for any polynomial P on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Therefore, the spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ consist of equivalent classes (modulo polynomials).

4 Frame Decomposition of Product Besov and Triebel–Lizorkin Spaces

The frame decomposition of the homogeneous product Besov and Triebel–Lizorkin spaces is a central component of our theory. We develop it analogously to the single parameter φ-transform of Frazier and Jawerth [11–14].

The discrete product Besov and Triebel–Lizorkin spaces will be spaces of sequences $\{a_{\mathbf{I}}\}_{\mathbf{I}\in\mathcal{R}}$ of complex numbers indexed by the set $\mathcal{R}$ of all dyadic rectangles in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Recall that $\mathcal{R}_{\boldsymbol{\nu}}$ stands for the set of all dyadic rectangles $\mathbf{I} \in \mathcal{R}$ such that $\mathbf{I} = I_1 \times I_2$ with $I_i \in \mathcal{D}^i_{\nu_i}$, $i = 1, 2$.

Definition 4.1 Let $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$ and $0 < q \le \infty$.

(i) The product Besov sequence space $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq} := \dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$, $0 < p \le \infty$, is defined as the set of all complex-valued sequences $a = \{a_{\mathbf{I}}\}_{\mathbf{I}\in\mathcal{R}}$ such that

$$\|a\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}} := \Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2} \Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}} \Big(2^{\mathbf{s}\cdot\boldsymbol{\nu}} |\mathbf{I}|^{1/p-1/2} |a_{\mathbf{I}}| \Big)^p \Big)^{q/p} \Big)^{1/q} < \infty. \tag{4.1}$$

(ii) The product Triebel–Lizorkin sequence space $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq} := \dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$, $0 < p < \infty$, is defined as the set of all complex-valued sequences $a = \{a_{\mathbf{I}}\}_{\mathbf{I}\in\mathcal{R}}$ such that

$$\|a\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}} := \|a\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})} := \Big\| \Big(\sum_{\mathbf{I}\in\mathcal{R}} \big(2^{\mathbf{s}\cdot\mathbf{r}_{\mathbf{I}}} |a_{\mathbf{I}}| \tilde{\mathbb{1}}_{\mathbf{I}}(\cdot) \big)^q \Big)^{1/q} \Big\|_p < \infty, \tag{4.2}$$

where $\mathbf{r}_{\mathbf{I}} := \boldsymbol{\nu}$ if $\mathbf{I} \in \mathcal{R}_{\boldsymbol{\nu}}$ and $\tilde{\mathbb{1}}_{\mathbf{I}} := |\mathbf{I}|^{-1/2}\mathbb{1}_{\mathbf{I}}$ with $\mathbb{1}_{\mathbf{I}}$ being the characteristic function of the rectangle $\mathbf{I}$. Above, the standard modification is used when $q = \infty$.

In our further development we will use the "analysis" and "synthesis" operators defined by

$$S_\varphi : f \to \{\langle f, \varphi_{\mathbf{I}}\rangle\}_{\mathbf{I}\in\mathcal{R}} \quad \text{and} \quad T_\psi : \{a_{\mathbf{I}}\}_{\mathbf{I}\in\mathcal{R}} \to \sum_{\mathbf{I}\in\mathcal{R}} a_{\mathbf{I}}\psi_{\mathbf{I}}. \tag{4.3}$$

One of the central assertions in this theory is the following:

Theorem 4.2 *Let* $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$ *and* $0 < q \le \infty$.

(*i*) *If* $0 < p \le \infty$, *then the operators* $S_\varphi\colon \dot{B}^{\mathbf{s}}_{pq} \to \dot{\mathfrak{b}}^{\mathbf{s}}_{pq}$ *and* $T_\psi : \dot{\mathfrak{b}}^{\mathbf{s}}_{pq} \to \dot{B}^{\mathbf{s}}_{pq}$ *are bounded, and* $T_\psi \circ S_\varphi$ *is the identity on* $\dot{B}^{\mathbf{s}}_{pq}$. *In particular,*

$$\|f\|_{\dot{B}^{\mathbf{s}}_{pq}} \sim \|\{\langle f, \varphi_{\boldsymbol{I}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}}, \quad \forall f \in \dot{B}^{\mathbf{s}}_{pq}.$$

(*ii*) *If* $0 < p < \infty$, *then the operators* $S_\varphi : \dot{F}^{\mathbf{s}}_{pq} \to \dot{\mathfrak{f}}^{\mathbf{s}}_{pq}$ *and* $T_\psi : \dot{\mathfrak{f}}^{\mathbf{s}}_{pq} \to \dot{F}^{\mathbf{s}}_{pq}$ *are bounded, and* $T_\psi \circ S_\varphi$ *is the identity on* $\dot{F}^{\mathbf{s}}_{pq}$. *In particular,*

$$\|f\|_{\dot{F}^{\mathbf{s}}_{pq}} \sim \|\{\langle f, \varphi_{\boldsymbol{I}}\rangle\}\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}}, \quad \forall f \in \dot{F}^{\mathbf{s}}_{pq}.$$

Theorem 4.2 is the analogue in the product space setting of the results of Frazier and Jawerth on the φ-transform from [11–13]. Its proof depends on the following:

Lemma 4.3 *Let* $0 < t \le 1$, $\tau_i > n_i/t$, $i = 1, 2$, *and* $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{Z}^2$. *Then for any sequence* $a = \{a_{\boldsymbol{I}}\}_{\boldsymbol{I}\in\mathcal{R}}$ *of complex numbers and any* $\boldsymbol{J} \in \mathcal{R}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \mathbb{Z}^2$,

$$\sum_{\boldsymbol{I}\in\mathcal{R}_{\boldsymbol{\mu}}} |a_{\boldsymbol{I}}| \prod_{i=1}^{2} \big(1 + 2^{\min\{\mu_i,\nu_i\}}|x_{I_i} - x_{J_i}|\big)^{-\tau_i} \le c \prod_{i=1}^{2} \max\big\{1, 2^{(\mu_i-\nu_i)n_i/t}\big\} \mathcal{M}_t\Big(\sum_{\boldsymbol{I}\in\mathcal{R}_{\boldsymbol{\mu}}} |a_{\boldsymbol{I}}| \mathbb{1}_{\boldsymbol{I}}\Big)(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \boldsymbol{J}. \tag{4.4}$$

We defer the proof of this lemma to the appendix.

Proof of Theorem 4.2 We will only prove the result for Triebel–Lizorkin spaces. The proof of the result for Besov spaces is similar and we omit it. Assume $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$, $0 < p < \infty$, and $0 < q < \infty$. The case $q = \infty$ is similar; we omit it.

(a) We next prove the boundedness of the synthesis operator T_ψ. We first consider the case of *finitely supported sequences*. Let $a = \{a_{\mathbf{I}}\} \in \dot{\mathfrak{f}}^{\mathbf{s}}_{pq}$ be a finitely supported sequence. We define $f := T_\psi a = \sum_{\mathbf{I}} a_{\mathbf{I}}\psi_R$. Let $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \mathbb{Z}^2$. From (2.13) and (2.16) it follows that

$$\varphi_{\boldsymbol{\nu}} * f(\mathbf{x}) = \sum_{\mu_1=\nu_1-1}^{\nu_1+1} \sum_{\mu_2=\nu_2-1}^{\nu_2+1} \sum_{\mathbf{I}\in\mathcal{R}_\mu} a_{\mathbf{I}}\varphi_{\boldsymbol{\nu}} * \psi_{\mathbf{I}}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}. \tag{4.5}$$

Assume that $\tau_i > n_i/t$, $i = 1, 2$, and $0 < t < \min\{1, p, q\}$. Because $|\mu_i - \nu_i| \le 1$, (2.1) and (2.14) imply

$$|\varphi_{\boldsymbol{\nu}} * \psi_{\mathbf{I}}(\mathbf{x})| \le \int_{\mathbb{R}^{n_1+n_2}} |\varphi_{\boldsymbol{\nu}}(\mathbf{x} - \mathbf{y})||\psi_{\mathbf{I}}(\mathbf{y})| d\mathbf{y} \le c \prod_{i=1}^{2} 2^{3\mu_i n_i/2} \int_{\mathbb{R}^{n_i}} \big(1 + 2^{\mu_i}|x_i - y_i|\big)^{-\tau_i} \big(1 + 2^{\mu_i}|y_i - x_{I_i}|\big)^{-\tau_i} dy_i.$$

As is well known, if $\tau_i > n_i$, then

$$\int_{\mathbb{R}^{n_i}} 2^{\mu_i n_i}\big(1+2^{\mu_i}|x_i-y_i|\big)^{-\tau_i}\big(1+2^{\mu_i}|y_i-x_{I_i}|\big)^{-\tau_i}\,\mathrm{d}y_i \le c\big(1+2^{\mu_i}|x_i-x_{I_i}|\big)^{-\tau_i}.$$

Therefore,

$$|\varphi_{\boldsymbol{\nu}} * \psi_{\mathbf{I}}(\mathbf{x})| \le c\prod_{i=1}^{2} 2^{\mu_i n_i/2}\big(1+2^{\mu_i}|x_i-x_{I_i}|\big)^{-\tau_i}. \tag{4.6}$$

By (4.5) and (4.6) we obtain

$$2^{\mathbf{s}\cdot\boldsymbol{\nu}}|\varphi_{\boldsymbol{\nu}} * f(\mathbf{x})| \le c\sum_{\mu_1=\nu_1-1}^{\nu_1+1}\sum_{\mu_2=\nu_2-1}^{\nu_2+1}\sum_{\mathbf{I}\in\mathcal{R}_\mu}|a_{\mathbf{I}}|2^{\mathbf{s}\cdot\mathbf{r}_{\mathbf{I}}}|\mathbf{I}|^{-1/2}\prod_{i=1}^{2}\big(1+2^{\mu_i}|x_i-x_{I_i}|\big)^{-\tau_i},$$

where $\mathbf{r}_{\mathbf{I}} := \boldsymbol{\mu} = (\mu_1,\mu_2)$ when $\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}$.

Fix $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ and $\boldsymbol{\mu} = (\mu_1,\mu_2)\in\mathbb{Z}^2$. Evidently, there exists a unique dyadic rectangle $\mathbf{J} := \mathbf{J}(\mathbf{x},\boldsymbol{\mu}) := J_1\times J_2 \in\mathcal{R}_{\boldsymbol{\mu}}$ that contains $\mathbf{x}$. Clearly,

$$1+2^{\mu_i}|x_i-x_{I_i}| \le 1+2^{\mu_i}(|x_i-x_{J_i}|+|x_{J_i}-x_{I_i}|) \le 2\big(1+2^{\mu_i}|x_{J_i}-x_{I_i}|\big),$$

and, therefore,

$$\begin{aligned}&\sum_{\mathbf{I}\in\mathcal{R}_\mu}|a_{\mathbf{I}}|2^{\mathbf{s}\cdot\mathbf{r}_{\mathbf{I}}}|\mathbf{I}|^{-1/2}\prod_{i=1}^{2}\big(1+2^{\mu_i}|x_{J_i}-x_{I_i}|\big)^{-\tau_i}\\ &\quad=\sum_{\mathbf{I}\in\mathcal{R}_\mu}|b_{\mathbf{I}}|\prod_{i=1}^{2}\big(1+2^{\mu_i}|x_{J_i}-x_{I_i}|\big)^{-\tau_i} \le c\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}|b_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)(\mathbf{x}),\end{aligned}$$

where $b_{\mathbf{I}} := |a_{\mathbf{I}}|2^{\mathbf{s}\cdot\mathbf{r}_{\mathbf{I}}}|\mathbf{I}|^{-1/2}$ and we applied Lemma 4.3. Putting the above together we arrive at

$$2^{\mathbf{s}\cdot\boldsymbol{\nu}}|\varphi_{\boldsymbol{\nu}} * f(\mathbf{x})| \le c\sum_{\mu_1=\nu_1-1}^{\nu_1+1}\sum_{\mu_2=\nu_2-1}^{\nu_2+1}\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}|b_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)(\mathbf{x}). \tag{4.7}$$

We use (4.7) in the definition of $\|f\|_{\dot{F}^{\mathbf{s}}_{pq}}$ and the maximal inequality (2.19) to obtain

$$\begin{aligned}\|f\|_{\dot{F}^{\mathbf{s}}_{pq}} &= \Big\|\Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2}\big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}|\varphi_{\boldsymbol{\nu}}*f|\big)^q\Big)^{1/q}\Big\|_p \\ &\le c\Big\|\Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2}\Big(\sum_{\mu_1=\nu_1-1}^{\nu_1+1}\sum_{\mu_2=\nu_2-1}^{\nu_2+1}\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}|b_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)\Big)^q\Big)^{1/q}\Big\|_p \\ &\le c\Big\|\Big(\sum_{\boldsymbol{\mu}\in\mathbb{Z}^2}\Big(\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}|b_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)\Big)^q\Big)^{1/q}\Big\|_p \\ &\le c\Big\|\Big(\sum_{\boldsymbol{\mu}\in\mathbb{Z}^2}\Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}|b_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)^q\Big)^{1/q}\Big\|_p \\ &\le c\Big\|\Big(\sum_{\boldsymbol{\mu}\in\mathbb{Z}^2}\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}2^{\boldsymbol{\nu}\cdot\mathbf{s}}|a_{\mathbf{I}}|\tilde{\mathbb{1}}_{\mathbf{I}}\Big)^q\Big)^{1/q}\Big\|_p \\ &= c\|a\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}}.\end{aligned}$$

Consequently, $\|T_\psi a\|_{\dot{F}^{\mathbf{s}}_{pq}} \le c\|a\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}}$ for all finitely supported sequences $a = \{a_{\mathbf{I}}\}$. But, it is easy to show that finitely supported sequences are dense in $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}$ $(q < \infty)$, and the boundedness of T_ψ in the general case follows by a limiting argument.

(b) We now prove the boundedness of the "analysis" operator S_φ. Let $f \in \dot{F}^{\mathbf{s}}_{pq}$. Define $\tilde{\varphi}(\mathbf{x}) := \overline{\varphi(-\mathbf{x})}$, $\mathbf{x} \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. For any $\mathbf{I} = I_1 \times I_2 \in \mathcal{R}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \mathbb{Z}^2$, we have using (2.8) and (2.11),

$$|\langle f, \varphi_{\mathbf{I}}\rangle| = 2^{-\boldsymbol{\nu}\cdot\mathbf{n}/2}|f * \tilde{\varphi}_{\boldsymbol{\nu}}(\mathbf{x}_{\mathbf{I}})|.$$

This and the fact that $1 + 2^{\nu_i}|x_{I_i} - x_i| \le 2$, $\forall x_i \in I_i$, yield

$$\begin{aligned}\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}}\big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}|\langle f,\varphi_{\mathbf{I}}\rangle|\tilde{\mathbb{1}}_{\mathbf{I}}(\mathbf{x})\big)^q &= \sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}}2^{\mathbf{s}\cdot\boldsymbol{\nu}q}|f*\tilde{\varphi}_{\boldsymbol{\nu}}(x_{\mathbf{I}})|^q\mathbb{1}_{\mathbf{I}}(\mathbf{x}) \\ &\le c\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}}2^{\mathbf{s}\cdot\boldsymbol{\nu}q}|f*\tilde{\varphi}_{\boldsymbol{\nu}}(x_{\mathbf{I}})|^q\prod_{i=1}^{2}\big(1+2^{\nu_i}|x_{I_i}-x_i|\big)^{-n_iq/t}\mathbb{1}_{\mathbf{I}}(\mathbf{x}) \\ &\le c\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}}\Big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}\sup_{\mathbf{y}\in\mathbf{I}}|f*\tilde{\varphi}_{\boldsymbol{\nu}}(\mathbf{y})|\prod_{i=1}^{2}\big(1+2^{\nu_i}|y_i-x_i|\big)^{-n_i/t}\Big)^q\mathbb{1}_{\mathbf{I}}(\mathbf{x}),\end{aligned}$$

where $0 < t < \min\{p, q\}$. From (2.16) it follows that

$$\operatorname{supp}\widehat{f*\tilde{\varphi}_{\boldsymbol{\nu}}} = \operatorname{supp}(\widehat{f}\cdot\widehat{\tilde{\varphi}_{\boldsymbol{\nu}}}) \subset [-2^{\nu_1+1}, 2^{\nu_1+1}]^{n_1} \times [-2^{\nu_2+1}, 2^{\nu_2+1}]^{n_2}.$$

We now invoke the maximal inequality (2.20) and obtain

$$\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}} \big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}|\langle f,\varphi_{\mathbf{I}}\rangle|\tilde{\mathbb{1}}_{\mathbf{I}}(\mathbf{x})\big)^q \le c2^{\mathbf{s}\cdot\boldsymbol{\nu}q}\big[\mathcal{M}_t\big(f*\tilde{\varphi}_{\boldsymbol{\nu}}\big)(\mathbf{x})\big]^q, \quad \forall \mathbf{x}\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}.$$

We use this and the maximal inequality (2.19) to obtain

$$\begin{aligned}\|\{\langle f,\varphi_{\mathbf{I}}\rangle\}\|_{\dot{\mathfrak{f}}^{s}_{pq}} &= \Big\|\Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2}\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}}\big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}|\langle f,\varphi_{\mathbf{I}}\rangle|\tilde{\mathbb{1}}_{\mathbf{I}}\big)^q\Big)^{1/q}\Big\|_p\\ &\le c\Big\|\Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2}\big(\mathcal{M}_t\big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}f*\tilde{\varphi}_{\boldsymbol{\nu}}\big)\big)^q\Big)^{1/q}\Big\|_p\\ &\le c\Big\|\Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2}\big(2^{\mathbf{s}\cdot\boldsymbol{\nu}}|f*\tilde{\varphi}_{\boldsymbol{\nu}}|\big)^q\Big)^{1/q}\Big\|_p\\ &\le c\|f\|_{\dot{F}^{\mathbf{s}}_{pq}}.\end{aligned}$$

Here for the last inequality we also used the fact that the definition of the norm in $\dot{F}^{\mathbf{s}}_{pq}$ is independent of the particular selection of the function φ in (3.2). Therefore, the "analysis" operator S_φ is bounded.

The fact that $T_\psi \circ S_\varphi$ is the identity on $\dot{F}^{\mathbf{s}}_{pq}$ follows immediately from (2.13). □

5 Almost Diagonal Operators

Almost diagonal operators acting on Besov or Triebel–Lizorkin sequence spaces are an important tool in dealing with these spaces. In this section we develop almost diagonal operators in the product framework. We use them in the next section to establish wavelet characterization of the product Besov and Triebel–Lizorkin spaces.

We will use the notation

$$\ell_i(\mathbf{I}) := 2^{-\nu_i} \quad \text{for } \mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}},\ \boldsymbol{\nu}=(\nu_1,\nu_2)\in\mathbb{Z}^2,\ i=1,2.$$

Also, let $J_i := n_i/\min\{1,p\}$ in the case of the space $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$, and $J_i := n_i/\min\{1,p,q\}$ in the case of $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$, $i=1,2$.

Definition 5.1 Assume $\mathbf{s}=(s_1,s_2)\in\mathbb{R}^2$, $0<q\le\infty$, and let $0<p\le\infty$ in the case of the space $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$ and $0<p<\infty$ in the case of $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$. A linear operator A with matrix $\{a_{IJ}\}_{I,J\in\mathcal{R}}$ is called almost diagonal on $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$ or on $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$ if there exists $\varepsilon>0$ such that

$$\|A\|_\varepsilon := \sup_{\mathbf{I},\mathbf{J}} |a_{\mathbf{IJ}}|/\omega_{\mathbf{IJ}}(\varepsilon) < \infty,$$

where

$$\omega_{\mathbf{IJ}}(\varepsilon) := \prod_{i=1}^{2} \left(\frac{\ell_i(\mathbf{I})}{\ell_i(\mathbf{J})}\right)^{s_i} \left(1 + \frac{|x_{I_i} - x_{J_i}|}{\max\{\ell_i(\mathbf{I}), \ell_i(\mathbf{J})\}}\right)^{-J_i-\varepsilon}$$
$$\times \min\left\{\left(\frac{\ell_i(\mathbf{I})}{\ell_i(\mathbf{J})}\right)^{(n_i+\varepsilon)/2}, \left(\frac{\ell_i(\mathbf{J})}{\ell_i(\mathbf{I})}\right)^{(n_i+\varepsilon)/2+J_i-n_i}\right\}.$$

We next establish the boundedness of almost diagonal operators on $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$ and $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$.

Theorem 5.2 *Let* $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$, $0 < q \le \infty$.

(a) *If A is an almost diagonal operator on* $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$, $0 < p \le \infty$, *then for any sequence* $h \in \dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$ *we have*
$$\|Ah\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})} \le c\|h\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})}. \tag{5.1}$$

(b) *If A is an almost diagonal operator on* $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$, $0 < p < \infty$, *then for any sequence* $h \in \dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$ *we have*
$$\|Ah\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})} \le c\|h\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})}. \tag{5.2}$$

The constant $c > 0$ *in* (5.1) *and* (5.2) *above is independent of the sequence* h.

For the proof of Theorem 5.2 we will need the following well-known Hardy inequalities:

Lemma 5.3 (a) *Let* $\gamma, q > 0$. *There exist a constant* $c = c(\gamma, q) > 0$ *such that for any sequence of non-negative numbers* $\{a_m\}_{m\in\mathbb{Z}}$,

$$\sum_{j\in\mathbb{Z}} \left(\sum_{m\ge j} 2^{-(m-j)\gamma} a_m\right)^q \le c \sum_{m\in\mathbb{Z}} a_m^q \tag{5.3}$$

and
$$\sum_{j\in\mathbb{Z}} \left(\sum_{m\le j} 2^{-(j-m)\gamma} a_m\right)^q \le c \sum_{m\in\mathbb{Z}} a_m^q. \tag{5.4}$$

(b) *Let* $\gamma_1, \gamma_2, q > 0$. *There exists a constant* $c = c(\gamma_1, \gamma_2, q) > 0$ *such that for any sequence* $\{d_{\boldsymbol{\mu}}\}_{\boldsymbol{\mu}\in\mathbb{Z}^2}$ *of non-negative numbers,*

$$\sum_{\nu_1\in\mathbb{Z}} \sum_{\nu_2\in\mathbb{Z}} \left(\sum_{\mu_1\ge\nu_1} \sum_{\mu_2<\nu_2} 2^{-(\mu_1-\nu_1)\gamma_1} 2^{-(\nu_2-\mu_2)\gamma_2} d_{\boldsymbol{\mu}}\right)^q \le c \sum_{\mu_1\in\mathbb{Z}} \sum_{\mu_2\in\mathbb{Z}} d_{\boldsymbol{\mu}}^q. \tag{5.5}$$

Inequality (5.5) follows by a simple combination of inequalities (5.3) and (5.4).

Proof of Theorem 5.2 We will only prove the boundedness of the almost diagonal operators on the Triebel–Lizorkin sequence spaces $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$. The proof in the case of the Besov space $\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}(\mathcal{R})$ is similar and will be omitted.

Let $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$ and $0 < p < \infty$. Assume $q < \infty$; the case $q = \infty$ is easier and we omit it.

Let A be a linear operator with matrix $\{a_{\mathbf{IJ}}\}_{\mathbf{I},\mathbf{J}\in\mathcal{R}}$ that is almost diagonal on $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$. We have to show that estimate (5.2) is valid. Given $\mathbf{I} \in \mathcal{R}$ we split $\mathcal{R}$ into four subsets: $\mathcal{R} = \cup_{i=1}^4 \mathcal{R}^i_{\mathbf{I}}$, where

$$\begin{aligned}
\mathcal{R}^1_{\mathbf{I}} &:= \{\mathbf{J} \in \mathcal{R}\colon \ell_1(\mathbf{J}) \le \ell_1(\mathbf{I}),\ \ell_2(\mathbf{J}) \le \ell_2(\mathbf{I})\},\\
\mathcal{R}^2_{\mathbf{I}} &:= \{\mathbf{J} \in \mathcal{R}\colon \ell_1(\mathbf{J}) \le \ell_1(\mathbf{I}),\ \ell_2(\mathbf{J}) > \ell_2(\mathbf{I})\},\\
\mathcal{R}^3_{\mathbf{I}} &:= \{\mathbf{J} \in \mathcal{R}\colon \ell_1(\mathbf{J}) > \ell_1(\mathbf{I}),\ \ell_2(\mathbf{J}) \le \ell_2(\mathbf{I})\},\\
\mathcal{R}^4_{\mathbf{I}} &:= \{\mathbf{J} \in \mathcal{R}\colon \ell_1(\mathbf{J}) > \ell_1(\mathbf{I}),\ \ell_2(\mathbf{J}) > \ell_2(\mathbf{I})\}.
\end{aligned}$$

Then

$$|(Ah)_{\mathbf{I}}| \le \sum_{\mathbf{J}\in\mathcal{R}} |a_{\mathbf{IJ}}||h_{\mathbf{J}}| = \sum_{i=1}^4 \sum_{\mathbf{J}\in\mathcal{R}^i_{\mathbf{I}}} |a_{\mathbf{IJ}}||h_{\mathbf{J}}|.$$

Applying the (quasi-)norm in $\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})$ (see (4.2)) we obtain

$$\|Ah\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}(\mathcal{R})} \le c \sum_{i=1}^4 \Big\| \Big(\sum_{\mathbf{I}\in\mathcal{R}} \Big(2^{\mathbf{s}\cdot\mathbf{r}_{\mathbf{I}}} \sum_{\mathbf{J}\in\mathcal{R}^i_{\mathbf{I}}} |a_{\mathbf{IJ}}||h_{\mathbf{J}}|\tilde{\mathbb{1}}_{\mathbf{I}} \Big)^q \Big)^{1/q} \Big\|_p =: c \sum_{i=1}^4 N_i.$$

Recall that $\mathbf{r}_{\mathbf{I}} := \boldsymbol{\nu}$ if $\mathbf{I} \in \mathcal{R}_{\boldsymbol{\nu}}$.

We will only estimate N_2; the estimation of N_1, N_3, N_4 is carried out along the same lines. Observe that if $\mathbf{I} \in \mathcal{R}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} = (\nu_1, \nu_2) \in \mathbb{Z}^2$, and $\mathbf{J} \in \mathcal{R}^2_{\mathbf{I}}$, then $\mathbf{J} \in \mathcal{R}_{\boldsymbol{\mu}}$, for some $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{Z}^2$ such that $\mu_1 \ge \nu_1$ and $\mu_2 < \nu_2$. Therefore,

$$N_2 = \Big\| \Big(\sum_{\nu_1\in\mathbb{Z}} \sum_{\nu_2\in\mathbb{Z}} \sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}} \Big(\sum_{\mu_1\ge\nu_1} \sum_{\mu_2<\nu_2} \sum_{\mathbf{J}\in\mathcal{R}_{\boldsymbol{\mu}}} |a_{\mathbf{IJ}}||h_{\mathbf{J}}|b_{\mathbf{I}}(\cdot) \Big)^q \Big)^{1/q} \Big\|_p, \tag{5.6}$$

where $b_{\mathbf{I}}(\mathbf{x}) := 2^{\mathbf{s}\cdot\mathbf{r}_{\mathbf{I}}} \tilde{\mathbb{1}}_{\mathbf{I}}(\mathbf{x})$.

Since A is almost diagonal, there exist constants $c, \varepsilon > 0$ such that

$$\begin{aligned}
|a_{\mathbf{IJ}}| &\le c 2^{(\mu_1-\nu_1)\left(s_1 - \mathcal{J}_1 + \frac{n_1}{2} - \frac{\varepsilon}{2}\right)} 2^{(\mu_2-\nu_2)\left(s_2 + \frac{n_2}{2} + \frac{\varepsilon}{2}\right)}\\
&\quad \times \left(1 + 2^{\nu_1}\left|x_{I_1} - x_{J_1}\right|\right)^{-\mathcal{J}_1-\varepsilon} \left(1 + 2^{\mu_2}\left|x_{I_2} - x_{J_2}\right|\right)^{-\mathcal{J}_2-\varepsilon},
\end{aligned}$$

and therefore

$$
\begin{aligned}
|a_{\mathbf{IJ}}||h_{\mathbf{J}}|b_{\mathbf{I}}(\mathbf{x}) &\le c2^{\mu\cdot s}2^{-\mu_1(\mathcal{J}_1-\frac{n_1}{2}+\frac{\varepsilon}{2})}2^{\nu_1(\mathcal{J}_1+\frac{\varepsilon}{2})}2^{\mu_2(n_2+\varepsilon)/2}2^{-\varepsilon\nu_2/2}|h_{\mathbf{J}}|\mathbb{1}_{\mathbf{I}}(\mathbf{x})\\
&\quad\times\left(1+2^{\nu_1}\left|x_{I_1}-x_{J_1}\right|\right)^{-\mathcal{J}_1-\varepsilon}\left(1+2^{\mu_2}\left|x_{I_2}-x_{J_2}\right|\right)^{-\mathcal{J}_2-\varepsilon}.
\end{aligned}
$$

We choose $t>0$ so that

$$
\frac{1}{t}=\frac{1}{\min\{1,p,q\}}+\frac{\varepsilon}{4\max\{n_1,n_2\}}.
$$

Then $0<t<\min\{1,p,q\}$ and $\frac{n_i}{t}\le\mathcal{J}_i+\frac{\varepsilon}{4}<\mathcal{J}_i+\varepsilon$. Applying Lemma 4.3 we get

$$
\begin{aligned}
&\sum_{\mathbf{J}\in\mathcal{R}_\mu}|h_{\mathbf{J}}|\prod_{i=1}^{2}\left(1+2^{\min\{\mu_i,\nu_i\}}|x_{J_i}-x_{I_i}|\right)^{-\mathcal{J}_i-\varepsilon}\\
&\quad\le c2^{(\mu_1-\nu_1)n_1/t}\mathcal{M}_t\Big(\sum_{\mathbf{J}\in\mathcal{R}_\mu}|h_{\mathbf{J}}|\mathbb{1}_{\mathbf{J}}\Big)(x)\le c2^{(\mu_1-\nu_1)(\mathcal{J}_1+\varepsilon/4)}\mathcal{M}_t\Big(\sum_{\mathbf{J}\in\mathcal{R}_\mu}|h_{\mathbf{J}}|\mathbb{1}_{\mathbf{J}}\Big)(\mathbf{x}),
\end{aligned}
$$

which in turn leads to

$$
\begin{aligned}
&\sum_{\mathbf{J}\in\mathcal{R}_\mu}|a_{\mathbf{IJ}}||h_{\mathbf{J}}|b_{\mathbf{I}}(\mathbf{x})\\
&\quad\le c2^{\mu\cdot s}2^{\mu_1n_1/2}2^{-(\mu_1-\nu_1)\varepsilon/4}2^{\mu_2(n_2+\varepsilon)/2}2^{-\varepsilon\nu_2/2}\mathcal{M}_t\Big(\sum_{\mathbf{J}\in\mathcal{R}_\mu}|h_{\mathbf{J}}|\mathbb{1}_{\mathbf{J}}\Big)(\mathbf{x})\\
&\quad=c2^{-(\mu_1-\nu_1)\varepsilon/4}2^{-(\nu_2-\mu_2)\varepsilon/2}\mathcal{M}_t\Big(\sum_{\mathbf{J}\in\mathcal{R}_\mu}|h_{\mathbf{J}}|b_{\mathbf{J}}\Big)(\mathbf{x}).
\end{aligned}
$$

Putting all of the above together we obtain

$$
\begin{aligned}
&\sum_{\mathbf{I}\in\mathcal{R}_\nu}\Big(\sum_{\mu_1\ge\nu_1}\sum_{\mu_2<\nu_2}\sum_{\mathbf{J}\in\mathcal{R}_\mu}|a_{\mathbf{IJ}}||h_{\mathbf{J}}|b_{\mathbf{I}}(\mathbf{x})\Big)^q\\
&\quad\le c\Big(\sum_{\mu_1\ge\nu_1}\sum_{\mu_2<\nu_2}2^{-(\mu_1-\nu_1)\varepsilon/4}2^{-(\nu_2-\mu_2)\varepsilon/2}\mathcal{M}_t\Big(\sum_{\mathbf{J}\in\mathcal{R}_\mu}|h_{\mathbf{J}}|b_{\mathbf{J}}\Big)(\mathbf{x})\Big)^q.
\end{aligned}
$$

From this and (5.6), and using the discrete Hardy-type inequality (5.5) and the maximal inequality (2.19), we obtain

$$N_2 \le c\Big\| \Big(\sum_{\nu_1\in\mathbb{Z}} \sum_{\nu_2\in\mathbb{Z}} \Big(\sum_{\mu_1\ge\nu_1} \sum_{\mu_2<\nu_2} 2^{-(\mu_1-\nu_1)\varepsilon/4} 2^{-(\nu_2-\mu_2)\varepsilon/2} \mathcal{M}_t \Big(\sum_{\mathbf{J}\in\mathcal{R}_{\boldsymbol{\mu}}} |h_{\mathbf{J}}| b_{\mathbf{J}} \Big)\Big)^q \Big)^{1/q} \Big\|_p$$
$$\le c\Big\| \Big(\sum_{\mu_1\in\mathbb{Z}} \sum_{\mu_2\in\mathbb{Z}} \Big(\mathcal{M}_t \Big(\sum_{\mathbf{J}\in\mathcal{R}_\mu} 2^{\mathbf{s}\cdot\mathbf{r_J}} |h_{\mathbf{J}}| \tilde{\mathbb{1}}_{\mathbf{J}} \Big)\Big)^q \Big)^{1/q} \Big\|_p$$
$$\le c\Big\| \Big(\sum_{\mu_1\in\mathbb{Z}} \sum_{\mu_2\in\mathbb{Z}} \Big(\sum_{\mathbf{J}\in\mathcal{R}_{\boldsymbol{\mu}}} 2^{\mathbf{s}\cdot\mathbf{r_J}} |h_{\mathbf{J}}| \tilde{\mathbb{1}}_{\mathbf{J}} \Big)^q \Big)^{1/q} \Big\|_p$$
$$= c\|h\|_{\mathfrak{f}^{\mathbf{s}}_{pq}}.$$

The proof is complete. □

6 Product Wavelet Bases

In this section we introduce product wavelets on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and show that they can be used for characterization of the product Besov and Triebel–Lizorkin spaces. The product wavelets will be defined as products of two families of regular tensor-product wavelets on $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$. For simplicity our construction of product wavelets will be based on the orthogonal univariate Mayer's wavelets. However, any other wavelet bases can be used instead as long as the wavelets are sufficiently smooth and have sufficiently fast decay and sufficiently many vanishing moments.

6.1 Regular Tensor-Product Wavelets on $\mathbb{R}^n$

We will use notation similar to the notation from Sect. 2.4. We denote by $\mathcal{D}$ the set of all dyadic cubes in $\mathbb{R}^n$ and by $\mathcal{D}_j$ the set of all cubes $I \in \mathcal{D}$ of side-length $\ell(I) = 2^{-j}$. For any $I \in \mathcal{D}$ we denote by x_I its lower-left corner and by $|I|$ its volume. Also, for any function g on $\mathbb{R}^n$ we define

$$g_I(x) := |I|^{-1/2} g\Big(\frac{x - x_I}{\ell(I)}\Big), \quad I \in \mathcal{D}. \tag{6.1}$$

We assume that φ is the scaling function and ψ is the associated wavelet in Meyer's wavelet system [20]. We also assume both φ and ψ normalized in $L^2(\mathbb{R})$, that is, $\|\varphi\|_{L^2} = \|\psi\|_{L^2} = 1$. Therefore, $\{2^{j/2}\psi(2^j x - k)\}_{j\in\mathbb{Z},k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(\mathbb{R})$. Define $\psi^0 := \varphi$ and $\psi^1 := \psi$. Let E be the set of all nonzero vertices of $[0, 1]^n$. Set

$$\psi^e(x) := \prod_{j=1}^n \psi^{e^j}(x^j), \quad e = (e^1, \dots, e^n) \in E. \tag{6.2}$$

Then the set

$$\left\{\psi_I^e : I \in \mathcal{D}, e \in E\right\}$$

is the regular tensor-product wavelet basis on $\mathbb{R}^n$, see, e.g., [30].

The most important properties of the wavelets $\{\psi_I^e\}$ can be summarized as follows: Each ψ_I^e is in $\mathcal{S}(\mathbb{R}^n)$,

$$\operatorname{supp} \widehat{\psi_I^e} \subset [-b2^j, b2^j]^n, \quad I \in \mathcal{D}_j, \quad \text{for some constant } b > 0, \tag{6.3}$$

and for any constants $M, K > 0$,

$$|\partial^\alpha \psi_I^e(x)| \le c 2^{j(n/2+|\alpha|)}(1 + 2^j|x - x_I|)^{-M}, \quad I \in \mathcal{D}_j, \ |\alpha| \le K, \tag{6.4}$$

and

$$\int_{\mathbb{R}^n} x^\alpha \psi_I^e(x) \mathrm{d}x = 0, \quad |\alpha| \le K. \tag{6.5}$$

Note that from $\|\psi_I^e\|_{L^2} = 1$ and (6.4), it follows that

$$\|\psi_I^e\|_{L^p} \sim |I|^{1/p-1/2}, \quad 0 < p \le \infty. \tag{6.6}$$

Remark 6.1 Other wavelets can be used in place of Meyer's wavelets, where the conditions $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$ and (6.3) are relaxed, but some limited smoothness is assumed, and conditions (6.4)–(6.5) are satisfied with the constant $M, K < \infty$ fixed. For example, the compactly supported orthogonal Daubechies wavelets [5] or bi-orthogonal wavelets [3] can be used. Then the theory that follows can be developed in full but with limited smoothness of the product Besov and Triebel–Lizorkin spaces and limited rates of approximation. We will not elaborate on these aspects of the theory.

6.2 Definition of Product Wavelets on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$

We now use the regular tensor-product wavelets from above to define product wavelets on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $n_1, n_2 \in \mathbb{N}$.

We will use the notation $\mathbf{x} := (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $x_i := (x_i^1, \dots, x_i^{n_i}) \in \mathbb{R}^{n_i}$, $i = 1, 2$. We consider two regular tensor-product wavelet bases on $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$:

$$\psi_I^{e_i}(x_i), \quad I \in \mathcal{D}^i, \ e_i \in E^i,$$

where we used (6.1) and

$$\psi^{e_i}(x_i) := \prod_{j=1}^{n_i} \psi^{e_i^j}(x_i^j), \quad e_i = (e_i^1, \dots, e_i^{n_i}) \in E^i, \quad i = 1, 2.$$

Here $\mathcal{D}^i$ is the set of all dyadic cubes in $\mathbb{R}^{n_i}$ and E^i is the set of all nonzero vertices of the cube $[0, 1]^{n_i}$.

Recall that $\mathcal{R}$ is the set of all dyadic rectangles in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. We set $E := E^1 \times E^2$. Now, the product wavelets are defined by

$$\psi_{\mathbf{I}}^{\mathbf{e}}(\mathbf{x}) := \psi_{I_1}^{e_1}(x_1)\psi_{I_2}^{e_2}(x_2), \quad \mathbf{x} = (x_1, x_2),\ \mathbf{I} = I_1 \times I_2 \in \mathcal{R},\ \mathbf{e} = (e_1, e_2) \in E.$$

Proposition 6.2 *The family*

$$\mathcal{W} := \{\psi_{\boldsymbol{I}}^{\boldsymbol{e}} : \boldsymbol{I} \in \mathcal{R}, \boldsymbol{e} \in E\}$$

is an orthonormal basis for $L^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$. Furthermore, for any $f \in \mathcal{S}_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ (or $f \in \mathcal{S}'_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$), we have

$$f = \sum_{\boldsymbol{I} \in \mathcal{R}, \boldsymbol{e} \in E} \langle f, \psi_{\boldsymbol{I}}^{\boldsymbol{e}} \rangle \psi_{\boldsymbol{I}}^{\boldsymbol{e}}, \tag{6.7}$$

where the convergence is in $\mathcal{S}$ (or in $\mathcal{S}'_\infty$).

This proposition relies on the following:

Lemma 6.3 *For any $M, K > 0$ there exists a constant $c > 0$ such that for all $f \in \mathcal{S}_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, $\boldsymbol{\nu} \in \mathbb{Z}^2$, $\boldsymbol{I} = I_1 \times I_2 \in \mathcal{R}_{\boldsymbol{\nu}}$, and $\boldsymbol{e} \in E$, we have*

$$|\langle f, \psi_{\boldsymbol{I}}^{e} \rangle| \le \frac{c2^{-K(|\nu_1|+|\nu_2|)}}{(1 + 2^{\nu_1 \wedge 0}|x_{I_1}|)^M (1 + 2^{\nu_2 \wedge 0}|x_{I_2}|)^M}, \quad \nu_1 \in \mathbb{Z},\ \nu_2 \in \mathbb{Z}, \tag{6.8}$$

where we used the notation: $a \wedge b := \min\{a, b\}$.

The proof of this lemma is in the appendix.

Proof of Proposition 6.2 By the definition of the product wavelets $\{\psi_{\mathbf{I}}^{\mathbf{e}}\}$ it is obvious that $\mathcal{W}$ is an orthonormal sequence in $L^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$. Further, as is well known (see, e.g., [30, §5.1]), $\{\psi_{I_i}^{e_i} : I_i \in \mathcal{D}^i, e \in E^i\}$ is an orthonormal basis for $\mathbb{R}^{n_i}$, $i = 1, 2$, and $L^2(\mathbb{R}^{n_1}) \otimes L^2(\mathbb{R}^{n_2}) = L^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$. Consequently, $\mathcal{W}$ is an orthonormal basis for $L^2(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$.

Let $f \in \mathcal{S}_\infty(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$. To show that (6.7) holds with convergence in $\mathcal{S}$, it suffices to show that for any multi indices $\alpha, \beta \in \mathbb{N}_0^{n_1} \times \mathbb{N}_0^{n_2}$,

$$\sum_{\mathbf{I} \in \mathcal{R}, \mathbf{e} \in E} \sup_{\mathbf{x}} |\langle f, \psi_{\mathbf{I}}^{\mathbf{e}} \rangle||\mathbf{x}^\alpha||\partial^\beta \psi_{\mathbf{I}}^{\mathbf{e}}(\mathbf{x})| < \infty.$$

Assume $|\alpha| \le \ell$, $|\beta| \le m$ for some $\ell, m \ge 0$. Consider the case when $\mathbf{I} \in \mathcal{R}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} = (\nu_1, \nu_2)$ with $\nu_1 \ge 0$, $\nu_2 < 0$. Then using (6.4) and (6.8), we obtain

$$|\langle f, \psi_{\mathbf{I}}^{\mathbf{e}} \rangle||\mathbf{x}^\alpha||\partial^\beta \psi_{\mathbf{I}}^{\mathbf{e}}(\mathbf{x})| \le \frac{c2^{-K(|\nu_1|+|\nu_2|)}(1 + |\mathbf{x}|)^\ell 2^{\nu_1(n_1/2+m)+\nu_2(n_2/2+m)}}{(1 + |x_{I_1}|)^M (1 + 2^{\nu_2}|x_{I_2}|)^M (1 + 2^{\nu_1}|x_1 - x_{I_1}|)^M (1 + 2^{\nu_2}|x_2 - x_{I_2}|)^M}.$$

Here $M, K > 0$ can be arbitrarily large. We choose $K := m + \ell + 3n_1/2 + n_2 + 1$ and $M := K + n_1 + n_2 + 1$. Clearly, because $\nu_1 \ge 0$, $\nu_2 < 0$, we have

$$\frac{(1+|x_1|)^\ell}{(1+|x_{I_1}|)^\ell(1+2^{\nu_1}|x_1-x_{I_1}|)^\ell} \le 1, \quad \frac{(1+|x_2|)^\ell}{(1+2^{\nu_2}|x_{I_2}|)^\ell(1+2^{\nu_2}|x_2-x_{I_2}|)^\ell} \le 2^{-\nu_2\ell},$$

and hence

$$|\langle f, \psi_{\mathbf{I}}^{\mathbf{e}}\rangle||\mathbf{x}^\alpha||\partial^\beta\psi_{\mathbf{I}}^{\mathbf{e}}(\mathbf{x})| \le \frac{c2^{-N(|\nu_1|+|\nu_2|)}}{(1+|x_{I_1}|)^{n_1+1}(1+2^{\nu_2}|x_{I_2}|)^{n_2+1}},$$

where $N := n_1 + n_2 + 1$. It easy to see that for any $\mathbf{I} \in \mathcal{R}_{\boldsymbol{\nu}}$,

$$\begin{aligned}\frac{1}{(1+|x_{I_1}|)^{n_1+1}(1+2^{\nu_2}|x_{I_2}|)^{n_2+1}} &\le \frac{2^{\nu_1(n_1+1)}}{(1+2^{\nu_1}|x_{I_1}|)^{n_1+1}(1+2^{\nu_2}|x_{I_2}|)^{n_2+1}}\\ &\le c2^{\nu_1(n_1+1)}\int_{\mathbf{I}}\frac{2^{\nu_1 n_1+\nu_2 n_2}}{(1+2^{\nu_1}|x_1|)^{n_1+1}(1+2^{\nu_2}|x_2|)^{n_2+1}}d\mathbf{x},\end{aligned}$$

implying

$$\begin{aligned}&\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}}}\frac{1}{(1+|x_{I_1}|)^{n_1+1}(1+2^{\nu_2}|x_{I_2}|)^{n_2+1}}\\ &\quad\le c2^{\nu_1(n_1+1)}\int_{\mathbb{R}^{n_1+n_2}}\frac{2^{\nu_1 n_1+\nu_2 n_2}}{\prod_{i=1}^2(1+2^{\nu_i}|x_i|)^{n_i+1}}d\mathbf{x} \le c2^{\nu_1(n_1+1)}.\end{aligned}$$

Therefore,

$$\sum_{\nu_1\ge 0}\sum_{\nu_2<0}\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\nu}},\mathbf{e}\in E}|\langle f, \psi_{\mathbf{I}}^{\mathbf{e}}\rangle||\mathbf{x}^\alpha||\partial^\beta\psi_{\mathbf{I}}^{\mathbf{e}}(\mathbf{x})| \le c\sum_{\nu_1\ge 0}\sum_{\nu_2<0}2^{-N(|\nu_1|+|\nu_2|)}2^{\nu_1(n_1+1)} < \infty.$$

One similarly shows that this inequality is valid when the summation $\sum_{\nu_1\ge 0}\sum_{\nu_2<0}$ is replaced by $\sum_{\nu_1\ge 0}\sum_{\nu_2\ge 0}$ or $\sum_{\nu_1<0}\sum_{\nu_2\ge 0}$ or $\sum_{\nu_1<0}\sum_{\nu_2<0}$. Consequently, the series in (6.7) converges in $\mathcal{S}$, and because $\mathcal{W}$ is a basis for L^2, it follows that this series converges to f.

By duality it follows that (6.7) is valid for any $f \in \mathcal{S}'_\infty$ with convergence in $\mathcal{S}'_\infty$. □

6.3 Characterization of Product Besov and Triebel–Lizorkin Spaces

In this section we show that the wavelet basis

$$\mathcal{W} = \left\{\psi_{\mathbf{I}}^{\mathbf{e}} : \mathbf{I} = I_1 \times I_2 \in \mathcal{R},\ e = (e_1, e_2) \in E\right\},$$

defined above, can be used for decomposition of the product Besov and Triebel–Lizorkin spaces.

Theorem 6.4 *Let* $\mathbf{s} = (s_1, s_2) \in \mathbb{R}^2$, $0 < q \le \infty$, *and* $f \in \mathcal{S}'_\infty$.

(i) *If* $0 < p \le \infty$, *then* $f \in \dot{B}^{\mathbf{s}}_{pq}$ *if and only if* $\{\langle f, \psi^{\boldsymbol{e}}_{\boldsymbol{I}}\rangle\} \in \dot{\mathfrak{b}}^{\mathbf{s}}_{pq}$, $\boldsymbol{e} \in E$, *and*

$$\|f\|_{\dot{B}^{\mathbf{s}}_{pq}} \sim \sum_{e \in E} \|\{\langle f, \psi^{\boldsymbol{e}}_{\boldsymbol{I}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}}. \tag{6.9}$$

(ii) *If* $0 < p < \infty$, *then* $f \in \dot{F}^{\mathbf{s}}_{pq}$ *if and only if* $\{\langle f, \psi^{\boldsymbol{e}}_{\boldsymbol{I}}\rangle\} \in \dot{\mathfrak{f}}^{\mathbf{s}}_{pq}$, $\boldsymbol{e} \in E$, *and*

$$\|f\|_{\dot{F}^{\mathbf{s}}_{pq}} \sim \sum_{e \in E} \|\{\langle f, \psi^{\boldsymbol{e}}_{\boldsymbol{I}}\rangle\}\|_{\dot{\mathfrak{f}}^{\mathbf{s}}_{pq}}. \tag{6.10}$$

In addition, if $f \in \dot{B}^{\mathbf{s}}_{pq}$ *or* $f \in \dot{F}^{\mathbf{s}}_{pq}$ *with* $p, q \ne \infty$, *then*

$$f = \sum_{\boldsymbol{I} \in \mathcal{R}, e \in E} \langle f, \psi^{\boldsymbol{e}}_{\boldsymbol{I}}\rangle \psi^{\boldsymbol{e}}_{\boldsymbol{I}}, \tag{6.11}$$

where the convergence is unconditional in the norm of $\dot{B}^{\mathbf{s}}_{pq}$ *or* $\dot{F}^{\mathbf{s}}_{pq}$, *respectively.*

Proof We will utilize the frame $\{\theta_{\mathbf{I}}\}$ from Sect. 2.4 (this is the case when $\varphi_{\mathbf{I}} = \psi_{\mathbf{I}} = \theta_{\mathbf{I}}$) and the frame characterization of the spaces $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ from Theorem 4.2 via $\{\theta_{\mathbf{I}}\}$. The almost diagonal operators from Theorem 5.2 will also play an important role. We will carry out the proof for Besov spaces only; the proof for Triebel–Lizorkin spaces is the same.

The following estimate on inner products will play an important role: For any $M, K > 0$ there exists a constant $c > 0$ such that for all $\mathbf{I}, \mathbf{J} \in \mathcal{R}$, $\mathbf{I} = I_1 \times I_2$, $\mathbf{J} = J_1 \times J_2$, and $\mathbf{e} = (e_1, e_2) \in E$, we have

$$|\langle \theta_{\mathbf{I}}, \psi^{\mathbf{e}}_{\mathbf{J}}\rangle| \le c \prod_{i=1}^{2} \min\left\{\frac{\ell(I_i)}{\ell(J_i)}, \frac{\ell(J_i)}{\ell(I_i)}\right\}^K \left(1 + \frac{|x_{I_i} - x_{J_i}|}{\max\{\ell(I_i), \ell(J_i)\}}\right)^{-M}. \tag{6.12}$$

This inequality follows at once from

$$|\langle \theta_{I_i}, \psi^{\mathbf{e}}_{J_i}\rangle| \le c \min\left\{\frac{\ell(I_i)}{\ell(J_i)}, \frac{\ell(J_i)}{\ell(I_i)}\right\}^K \left(1 + \frac{|x_{I_i} - x_{J_i}|}{\max\{\ell(I_i), \ell(J_i)\}}\right)^{-M}, \quad i = 1, 2.$$

This estimate is well known and due to the infinite smoothness, fast decay, and vanishing moments of θ_{I_i} and $\psi^{\mathbf{e}}_{J_i}$. It can be derived from [13, Lemma B.1], see also [17, Lemma 2.1]. In essence its proof is contained in the proof of the more complicated Lemma 6.3 above.

Assume that $f \in \dot{B}^{\mathbf{s}}_{pq}$. Using Proposition 2.3 we have $f = \sum_{\mathbf{J} \in \mathcal{R}} \langle f, \theta_{\mathbf{J}}\rangle \theta_{\mathbf{J}}$ with convergence in $\mathcal{S}'_\infty$ and hence for any $\mathbf{I} \in \mathcal{R}$, $\mathbf{e} \in E$,

$$\langle f, \psi^{\mathbf{e}}_{\mathbf{I}}\rangle = \sum_{\mathbf{J} \in \mathcal{R}} \langle f, \theta_{\mathbf{J}}\rangle \overline{\langle \psi^{\mathbf{e}}_{\mathbf{I}}, \theta_{\mathbf{J}}\rangle} = \sum_{\mathbf{J} \in \mathcal{R}} a^{\mathbf{e}}_{\mathbf{IJ}} \langle f, \theta_{\mathbf{J}}\rangle, \qquad a^{\mathbf{e}}_{\mathbf{IJ}} := \overline{\langle \psi^{\mathbf{e}}_{\mathbf{I}}, \theta_{\mathbf{J}}\rangle}.$$

From (6.12) and Theorem 5.2 it readily follows that the operator $A^{\mathbf{e}}$ with matrix $\{a^{\mathbf{e}}_{\mathbf{IJ}}\}_{\mathbf{I},\mathbf{J}\in\mathcal{R}}$ is almost diagonal on $\dot{\mathfrak{b}}^{s}_{pq}$ and hence it is bounded. We use this and Theorem 4.2 with $\{\theta_{\mathbf{I}}\}$ instead of $\{\varphi_{\mathbf{I}}\}$ to obtain

$$\|\{\langle f,\psi^{\mathbf{e}}_{\mathbf{I}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}} = \|A^{\mathbf{e}}\{\langle f,\theta_{\mathbf{J}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}} \le c\|\{\langle f,\theta_{\mathbf{J}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}} \le c\|f\|_{\dot{B}^{\mathbf{s}}_{pq}}, \quad \mathbf{e}\in E. \tag{6.13}$$

Hence, $\{\langle f,\psi^{\mathbf{e}}_{\mathbf{I}}\rangle\}\in\dot{\mathfrak{b}}^{s}_{pq}$, $\forall\mathbf{e}\in E$.

For the other direction, assume that $f\in\mathcal{S}'_\infty$ and $\{\langle f,\psi^{\mathbf{e}}_{\mathbf{I}}\rangle\}\in\dot{\mathfrak{b}}^{s}_{pq}$, $\mathbf{e}\in E$. Appealing to Proposition 6.2, we have $f=\sum_{\mathbf{I}\in\mathcal{R},\mathbf{e}\in E}\langle f,\psi^{\mathbf{e}}_{\mathbf{I}}\rangle\psi^{\mathbf{e}}_{\mathbf{I}}$ with convergence in $\mathcal{S}'_\infty$, and hence

$$\langle f,\theta_{\mathbf{I}}\rangle = \sum_{\mathbf{J}\in\mathcal{R},\mathbf{e}\in E}\langle f,\psi^{\mathbf{e}}_{\mathbf{J}}\rangle\overline{\langle\theta_{\mathbf{I}},\psi^{\mathbf{e}}_{\mathbf{J}}\rangle} = \sum_{\mathbf{J}\in\mathcal{R},\mathbf{e}\in E} b^{\mathbf{e}}_{\mathbf{IJ}}\langle f,\psi^{\mathbf{e}}_{\mathbf{J}}\rangle, \qquad b^{\mathbf{e}}_{\mathbf{IJ}} := \overline{\langle\theta_{\mathbf{I}},\psi^{\mathbf{e}}_{\mathbf{J}}\rangle}.$$

Let $B^{\mathbf{e}}$ be the operator with matrix $\{b^{\mathbf{e}}_{\mathbf{IJ}}\}_{\mathbf{I},\mathbf{J}\in\mathcal{R}}$, $\mathbf{e}\in E$. As above from (6.12) and Theorem 5.2, it follows that $B^{\mathbf{e}}$ is an almost diagonal operator on $\dot{\mathfrak{b}}^{s}_{pq}$ and hence it is bounded. This along with Theorem 4.2 implies

$$\|f\|_{\dot{B}^{\mathbf{s}}_{pq}} \le c\|\{\langle f,\theta_{\mathbf{I}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}} \le c\sum_{\mathbf{e}\in E}\|B^{\mathbf{e}}\{\langle f,\psi^{\mathbf{e}}_{\mathbf{J}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}} \le c\sum_{\mathbf{e}\in E}\|\{\langle f,\psi^{\mathbf{e}}_{\mathbf{J}}\rangle\}\|_{\dot{\mathfrak{b}}^{\mathbf{s}}_{pq}}. \tag{6.14}$$

Therefore, $f\in\dot{B}^{\mathbf{s}}_{pq}$. The equivalence (6.9) follows by (6.13) and (6.14).

The unconditional convergence in (6.11) follows readily by the wavelet characterization of the norms in $\dot{B}^{\mathbf{s}}_{pq}$ and $\dot{F}^{\mathbf{s}}_{pq}$ from above. □

6.4 Product Hardy Spaces

As elsewhere in harmonic analysis and approximation theory, it is natural to work in Hardy spaces H^p rather than in L^p when $0<p\le 1$. The theory, of product Hardy spaces H^p, $0<p\le 1$, was initiated by Gundy and Stein [16] and has attracted considerable attention. We refer the reader to [1,2,9,10,19] and the references therein for more information on product Hardy spaces.

The product Hardy spaces $H^p=H^p(\mathbb{R}^{n_1}\times\mathbb{R}^{n_2})$, $0<p\le 1$, are usually defined via the Lusin-area function, but there is also a Littlewood–Paley characterization of these spaces as well as characterization via the φ-transform (see [1,19]). The Littlewood–Paley characterization of the product Hardy spaces H^p, $0<p\le 1$, simply asserts that $H^p=\dot{F}^{0}_{p2}(\mathbb{R}^{n_1}\times\mathbb{R}^{n_2})$ with equivalent norms. We will take this as a definition for product Hardy space H^p and set (see Definition 3.1)

$$\|f\|_{H^p} := \Big\|\Big(\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2}|\varphi_{\boldsymbol{\nu}}*f|^2\Big)^{1/2}\Big\|_p. \tag{6.15}$$

However, this needs some further clarification because the space $\dot{F}^{0}_{p2}(\mathbb{R}^{n_1}\times\mathbb{R}^{n_2})$ consists of equivalence classes modulo polynomials.

Proposition 6.5 *Let* $f \in \dot{F}^0_{p2}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, $0 < p < \infty$.

(a) *Let* $\varphi = \varphi^1 \otimes \varphi^2$, *where* φ^1, φ^2 *satisfy* (2.3), (2.5) *with* $n = n_1, n_2$, *respectively. Then there exists* $f_0 \in \mathcal{S}'$ *in the equivalent class determined by* f *such that*

$$f_0 = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \varphi_{\boldsymbol{\nu}} * f, \quad \text{with convergence in } \mathcal{S}'. \tag{6.16}$$

Moreover, $f_0 \in \mathcal{S}'$ *is independent of the specific selection of* φ.

(b) *Let* φ^1, ψ^1 *and* φ^2, ψ^2 *be two pairs of functions satisfying* (2.5)–(2.6) *with* $n = n_1, n_2$, *respectively. Set* $\varphi := \varphi^1 \otimes \varphi^2$ *and* $\psi := \psi^1 \otimes \psi^2$. *Then*

$$f_0 = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \tilde{\psi}_{\boldsymbol{\nu}} * \varphi_{\boldsymbol{\nu}} * f, \quad \text{with convergence in } \mathcal{S}'.$$

(c) *Let* $\{\psi^{\boldsymbol{e}}_{\boldsymbol{I}} : (\boldsymbol{I}, \boldsymbol{e}) \in \mathcal{R} \times E\}$ *be the wavelet basis defined in Sect.* 6.2. *Then*

$$f_0 = \sum_{\boldsymbol{I} \in \mathcal{R}, \boldsymbol{e} \in E} \langle f, \psi^{\boldsymbol{e}}_{\boldsymbol{I}} \rangle \psi^{\boldsymbol{e}}_{\boldsymbol{I}}, \quad \text{with convergence in } \mathcal{S}'.$$

This claim is analogous to [13, Remark B.4]. We include its proof in the appendix.

Convention From now on we will identify $f \in H^p(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) = \dot{F}^0_{p2}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, $0 < p < \infty$, with its *canonical representative*

$$\sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \varphi_{\boldsymbol{\nu}} * f = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^2} \tilde{\psi}_{\boldsymbol{\nu}} * \varphi_{\boldsymbol{\nu}} * f = \sum_{\mathbf{I} \in \mathcal{R}, \mathbf{e} \in E} \langle f, \psi^{\mathbf{e}}_{\mathbf{I}} \rangle \psi^{\mathbf{e}}_{\mathbf{I}},$$

see Proposition 6.5.

Observe that $f \in H^p(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$, $1 < p < \infty$, if and only if $f \in L^p(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$ and $\|f\|_{L^p} \sim \|f\|_{H^p}$, see [1].

We will need the φ-transform (Theorem 4.2) and wavelet (Theorem 6.4) characterizations of $H^p(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}) = \dot{F}^0_{p2}(\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})$:

$$\|f\|_{H^p} \sim \Big\| \Big(\sum_{\mathbf{I} \in \mathcal{R}} \big[|\langle f, \varphi_{\mathbf{I}} \rangle| \tilde{\mathbb{1}}_{\mathbf{I}}(\cdot) \big]^2 \Big)^{1/2} \Big\|_p \sim \Big\| \Big(\sum_{\mathbf{I} \in \mathcal{R}, \mathbf{e} \in E} \big[|\langle f, \psi^{\mathbf{e}}_{\mathbf{I}} \rangle| \tilde{\mathbb{1}}_{\mathbf{I}}(\cdot) \big]^2 \Big)^{1/2} \Big\|_p.$$

7 Nonlinear *m*-Term Approximation from Product Wavelets

Here we consider nonlinear m-term approximation from the product wavelet basis $\{\psi^{\mathbf{e}}_{\mathbf{I}}\}_{\mathbf{I} \in \mathcal{R}, \mathbf{e} \in E}$ defined in Sect. 6.2 in L^p, $1 < p < \infty$, or H^p, $0 < p \le 1$. Denote by Σ_m the set of all functions on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ of the form

$$g = \sum_{(\mathbf{I}, \mathbf{e}) \in \Lambda_m} a_{\mathbf{I}\mathbf{e}} \psi^{\mathbf{e}}_{\mathbf{I}},$$

where $\Lambda_m \subset \mathcal{R} \times E$, $\#\Lambda_m \le m$, and Λ_m is allowed to vary with g. We define

$$\sigma_m(f) := \inf_{g \in \Sigma_m} \|f - g\|_p, \tag{7.1}$$

where $\|\cdot\|_p = \|\cdot\|_{L^p}$ if $1 < p < \infty$ and $\|\cdot\|_p = \|\cdot\|_{H^p}$ if $0 < p \le 1$.

In what follows we assume that

$$0 < p < \infty, \quad \alpha > 0, \quad \frac{1}{\tau} := \alpha + \frac{1}{p}, \quad s_1 := n_1\alpha, \quad s_2 := n_2\alpha, \quad \mathbf{s} := (s_1, s_2). \tag{7.2}$$

Thus $s_1 := n_1\big(\frac{1}{\tau} - \frac{1}{p}\big)$ and $s_2 := n_2\big(\frac{1}{\tau} - \frac{1}{p}\big)$. The Besov spaces

$$\dot{B}^{\mathbf{s}}_{\tau\tau} := \dot{B}^{(s_1,s_2)}_{\tau\tau}$$

will play an important role here. Our goal is to establish a sharp Jackson estimate for $\sigma_m(f)$ and companion Bernstein estimate in terms of the Besov spaces $\dot{B}^{\mathbf{s}}_{\tau\tau}$.

Observe that from Theorem 6.4 and the fact that $\|\psi^{\mathbf{e}}_{\mathbf{I}}\|_{L^q} \sim |\mathbf{I}|^{\frac{1}{q}-\frac{1}{2}}$, $0 < q \le \infty$, [a consequence of (6.6)], it follows that for any $f \in \dot{B}^{\mathbf{s}}_{\tau\tau}$,

$$\|f\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \sim \Big(\sum_{\mathbf{I}\in\mathcal{R}, \mathbf{e}\in E} \|\langle f, \psi^{\mathbf{e}}_{\mathbf{I}}\rangle \psi^{\mathbf{e}}_{\mathbf{I}}\|^{\tau}_{L^p}\Big)^{1/\tau}. \tag{7.3}$$

The embedding of $\dot{B}^{\mathbf{s}}_{\tau\tau}$ into L^p or H^p will play a critical role.

Proposition 7.1 *Let $f \in \dot{B}^{s}_{\tau\tau}$, where $\mathbf{s}$ and τ are as in* (7.2).

(a) *If $1 < p < \infty$, then $f \in L^p$ and*

$$\|f\|_{L^p} \le c\|f\|_{\dot{B}^{s}_{\tau\tau}}. \tag{7.4}$$

(b) *If $0 < p \le 1$, then $f \in H^p$ and*

$$\|f\|_{H^p} \le c\|f\|_{\dot{B}^{s}_{\tau\tau}}. \tag{7.5}$$

Above, $c > 0$ is a constant independent of f.

We now come to the main results in this section.

Theorem 7.2 (Jackson estimate) *If $f \in \dot{B}^{s}_{\tau\tau}$, then for $m \ge 2$,*

$$\sigma_m(f)_p \le cm^{-\alpha}\|f\|_{\dot{B}^{s}_{\tau\tau}}, \quad 0 < p \le 2, \tag{7.6}$$

$$\sigma_m(f)_p \le cm^{-\alpha}(\ln m)^{1/2-1/p}\|f\|_{\dot{B}^{s}_{\tau\tau}}, \quad 2 < p < \infty, \tag{7.7}$$

where the constant $c > 0$ is independent of m.

Theorem 7.3 (Bernstein estimate) *If* $g \in \Sigma_m$*,* $m \ge 2$*, then*

$$\|g\|_{\dot{B}^s_{\tau\tau}} \le cm^\alpha \big(\ln m\big)^{1/p-1/2} \|g\|_p, \quad 0 < p < 2, \tag{7.8}$$

$$\|g\|_{\dot{B}^s_{\tau\tau}} \le cm^\alpha \|g\|_p, \quad 2 \le p < \infty, \tag{7.9}$$

where the constant $c > 0$ *is independent of* m.

Before proving these theorems we derive direct and inverse estimates that follow from the above Jackson and Bernstein estimates. Denote by $K(f,t)$ the K-functional associated with L^p and $\dot{B}^{\mathbf{s}}_{\tau\tau}$ if $1 < p < \infty$ or H^p and $\dot{B}^{\mathbf{s}}_{\tau\tau}$ if $0 < p \le 1$; namely (see, e.g., [8]), for $f \in L^p$, $1 < p < \infty$,

$$K(f,t) = K(f,t; L^p, \dot{B}^{\mathbf{s}}_{\tau\tau}) := \inf_{g \in \dot{B}^{\mathbf{s}}_{\tau\tau}} \big(\|f-g\|_p + t\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}\big), \quad t > 0,$$

and L^p above is replaced by H^p whenever $f \in H^p$, $0 < p \le 1$.

Theorem 7.4 (Direct estimate) *If* $f \in L^p$*,* $1 < p < \infty$*, or* $f \in H^p$*,* $0 < p \le 1$*, then*

$$\sigma_m(f)_p \le cK(f, m^{-\alpha}), \quad 0 < p \le 2,$$
$$\sigma_m(f)_p \le cK\big(f, m^{-\alpha}(\ln m)^{1/2-1/p}\big), \quad 2 < p < \infty, \ m \ge 2.$$

Theorem 7.5 (Inverse estimate) *If* $f \in L^p$*,* $1 < p < \infty$*, or* $f \in H^p$*,* $0 < p \le 1$*, then*

$$K(f, m^{-\alpha}) \le cm^{-\alpha}(\ln m)^{1/p-1/2}\Big[\Big(\sum_{k=1}^m \frac{1}{k}(k^\alpha \sigma_k(f)_p)^\mu\Big)^{1/\mu} + \|f\|_p\Big], \ 0 < p < 2,$$

$$K(f, m^{-\alpha}) \le cm^{-\alpha}\Big[\Big(\sum_{k=1}^m \frac{1}{k}(k^\alpha \sigma_k(f)_p)^\mu\Big)^{1/\mu} + \|f\|_p\Big], \ 2 \le p < \infty, \ m \ge 2.$$

Here $\mu := \min\{\tau, 1\}$.

The proofs of Theorems 7.4, 7.5 are standard and will be omitted, see, e.g., [8, Chapter 7, Theorem 5.1].

Corollary 7.6 *Let* $f \in L^p$*,* $1 < p < \infty$*, or* $f \in H^p$*,* $0 < p \le 1$*, and* $0 < \gamma < \alpha$*. Then:* (*a*) *If* $0 < p \le 2$*, then*

$$K(f, t^\alpha) = O(t^\gamma) \ \ \textit{implies} \ \ \sigma_m(f)_p = O(m^{-\gamma}),$$

and

$$\sigma_m(f)_p = O(m^{-\gamma}) \ \ \textit{implies} \ \ K(f, t^\alpha) = O\big(t^\gamma(\ln 1/t)^\beta\big), \quad \beta := 1/p - 1/2.$$

(*b*) *If* $2 < p < \infty$*, then*

$$K(f, t^\alpha) = O(t^\gamma) \ \ \textit{implies} \ \ \sigma_m(f)_p = O\big(m^{-\gamma}(\ln m)^\beta\big), \quad \beta := (1/2 - 1/p)\gamma/\alpha,$$

and

$$\sigma_m(f)_p = O(m^{-\gamma}) \quad \textit{implies} \quad K(f, t^\alpha) = O(t^\gamma).$$

The next proposition shows that the Jackson and Bernstein estimates from Theorems 7.2, 7.3 are sharp. In particular, the logarithmic terms in (7.7) and (7.8) cannot be removed and they are of the correct form.

Proposition 7.7 *The Jackson estimates (7.6)–(7.7) as well as the Bernstein estimates (7.8)–(7.9) are sharp in the following sense: Let $p, \alpha, \tau, s_1, s_2$ be as in (7.2). Then for any $m \geq 2$,*

$$\sup_{\|f\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}=1} \sigma_m(f)_p \geq cm^{-\alpha}, \quad 0 < p \leq 2, \tag{7.10}$$

$$\sup_{\|f\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}=1} \sigma_m(f)_p \geq cm^{-\alpha}(\ln m)^{1/2-1/p}, \quad 2 < p < \infty, \tag{7.11}$$

$$\sup_{g\in\Sigma_m, \|g\|_p=1} \|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \geq cm^{\alpha}(\ln m)^{1/p-1/2}, \quad 0 < p \leq 2, \tag{7.12}$$

$$\sup_{g\in\Sigma_m, \|g\|_p=1} \|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \geq cm^{\alpha}, \quad 2 \leq p < \infty. \tag{7.13}$$

Above, the constant $c > 0$ is independent of m.

The proof of this proposition is deferred to the appendix.

Remark 7.8 (a) Note that the parameter $\alpha > 0$ above can be arbitrarily large due to the fact that the product wavelets that we work with are based on Meyer's wavelets and characterize the spaces $\dot{B}^{\mathbf{s}}_{pq}$, $\dot{F}^{\mathbf{s}}_{pq}$ in the complete range of the parameters s, p, q. The above approximation results can be obtained for product wavelets based on, e.g., compactly supported orthogonal Daubechies wavelets [5] or bi-orthogonal wavelets [3], but with a limited range for α depending on their smoothness, decay, and number of vanishing moments.

(b) In the Jackson estimate (7.7) and Bernstein estimate (7.8) there are logarithmic factors $(\ln m)^{1/2-1/p}$ and $(\ln m)^{1/p-1/2}$ that prevent them from perfectly matching their respective counterparts. However, as is shown in Proposition 7.7, they cannot be removed and are of the right form. These logarithmic factors are due to the fact that the "essential supports" of a (logarithmic) number of product wavelets of the same size overlap at any point.

The proofs of Proposition 7.1 and Theorems 7.2, 7.3 rely on the following lemma, where as before $\mathcal{R}$ stands for the set of all dyadic rectangles in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.

Lemma 7.9 *If $F = \sum_{\boldsymbol{I}\in\Upsilon_m} |\boldsymbol{I}|^{-1/p}\mathbb{1}_{\boldsymbol{I}}$, where $\Upsilon_m \subset \mathcal{R}$, $\#\Upsilon_m \leq m$, $m \geq 2$, and $1 < p < \infty$, then*

$$\|F\|_p \leq cm^{1/p}(\ln m)^{1-1/p}, \tag{7.14}$$

where the constant $c > 0$ depends only on p. Furthermore, this estimate is sharp.

Evidently, $\|F\|_p \leq cm^{1/p}$ in the case $0 < p \leq 1$.

The proof of Lemma 7.9 depends on the following:

Lemma 7.10 *If* $1 < p < \infty$ *and* $N \in \mathbb{N}$, *then*

$$(6p)^{-p}\Big(\sum_{j=1}^{N} 2^{j/p}x_j\Big)^p \le \sum_{j=1}^{N} 2^j x_j, \quad 0 \le x_j \le 1. \tag{7.15}$$

Proof Consider the function

$$h(x) := A\sum_{j=1}^{N} 2^j x_j - \Big(\sum_{j=1}^{N} 2^{j/p}x_j\Big)^p, \quad x = (x_1, \dots, x_N), \quad A := (6p)^p. \tag{7.16}$$

We have to show that $h(x) \ge 0$ on the set $[0, 1]^N$. We may assume that $N \ge 2$. Clearly,

$$\frac{\partial h}{\partial x_j}(x) = A2^j - p2^{j/p}\left(\sum_{\nu=1}^{N} 2^{\nu/p}x_\nu\right)^{p-1},$$

and hence

$$\frac{\partial h}{\partial x_j}(x) = 0, \ \ j = 1, \dots, N$$

if and only if

$$\sum_{\nu=1}^{N} 2^{\nu/p}x_\nu = \big(A2^{j(1-1/p)}/p\big)^{1/(p-1)}, \ \ j = 1, \dots, N.$$

Evidently, this system for $x_1, \dots, x_N$ has no solutions if $N \ge 2$. Therefore, the function h has no critical points on $(0, 1)^N$, and hence $\min_{x\in[0,1]^N} h(x)$ is attained on the boundary of $[0, 1]^N$. Clearly, the boundary of $[0, 1]^N$ is contained in the union of sets Ω_ℓ of the form

$$\Omega_\ell := \big\{x \in \mathbb{R}^N \colon 0 \le x_{j_k} \le 1, \ k = 1, \dots, \ell, \ \text{and } x_j = 0 \text{ or } x_j = 1, \ \ j \ne j_k\big\}$$

for some set of indices $\{j_r\}_{r=1}^{\ell}$, $1 \le j_1 < \cdots < j_\ell \le N$, with $\ell \in \{1, \dots, N\}$.

Consider $h(x)$ on Ω_ℓ in the case when $\ell \ge 2$. Then in the definition of $h(x)$ in (7.16) $x_j = 0$ or $x_j = 1$ for $j \ne j_k$. Just as above we conclude that the system $\frac{\partial h}{\partial x_{j_k}}(x) = 0$, $k = 1, \dots, \ell$, ($\ell \ge 2$) has no solutions in the interior of Ω_ℓ, and hence $\min_{x\in\Omega_\ell} h(x)$ is attained on the boundary of Ω_ℓ if $\ell \ge 2$.

Consequently, it suffices to show that $h(x) \ge 0$ on any set Ω_1 of the form

$$\Omega_1 := \big\{x \in \mathbb{R}^N \colon 0 \le x_\ell \le 1 \text{ for some } \ell \in \{1, \dots, N\}, \ \text{and } x_j = 0 \text{ or } 1, \ \ j \ne \ell\big\}. \tag{7.17}$$

It is readily seen that $\min_{x\in\Omega_1} h(x) \ge 0$ if in (7.17) $x_j = 0$ for all $j \ne \ell$.

Assume that Ω_1 is the set of all $x \in \mathbb{R}^N$ such that $0 \le x_\ell \le 1$ for some $1 \le \ell \le N$, $x_{j_r} = 1$ for $1 \le j_1 < \cdots < j_\mu \le N$, $j_r \ne \ell$, and $x_j = 0$ for $j \ne j_r$ and $j \ne \ell$; $\mu \ge 1$. We use that $(a+b)^p \le 2^p(a^p + b^p)$, $a, b \ge 0$, to obtain for $x \in \Omega_1$,

$$\left(\sum_{j=1}^{N} 2^{j/p} x_j\right)^p \le 2^p \left[(2^{\ell/p} x_\ell)^p + \left(\sum_{r=1}^{\mu} 2^{j_r/p}\right)^p\right] \le 2^p \big(2^\ell x_\ell + (3p)^p 2^{j_\mu}\big). \tag{7.18}$$

Here we used that $0 \le x_\ell \le 1$ and

$$\left(\sum_{r=1}^{\mu} 2^{j_r/p}\right)^p \le \left(\sum_{j=1}^{j_\mu} 2^{j/p}\right)^p = \left(\frac{2^{j_\mu+1} - 1}{2^{1/p} - 1}\right)^p \le (3p)^p 2^{j_\mu}.$$

From (7.18) it follows that

$$\left(\sum_{j=1}^{N} 2^{j/p} x_j\right)^p \le (6p)^p \sum_{j=1}^{N} 2^j x_j, \quad x \in \Omega_1,$$

implying $\min_{x \in \Omega_1} h(x) \ge 0$. Consequently, $\min_{x \in [0,1]^N} h(x) \ge 0$, which implies (7.15). The proof of the lemma is complete. □

Proof of Lemma 7.9 Assume $1 < p < \infty$. Clearly, Lemma 7.9 is invariant under dyadic dilations, and hence we may assume that $|\mathbf{I}| \le 1$ for all $\mathbf{I} \in \Upsilon_m$. Assume $|\mathbf{I}| \ge 2^{-N}$, $\forall \mathbf{I} \in \Upsilon_m$, for some $N \in \mathbb{N}$. Define $\mathfrak{R}_j := \{\mathbf{I} \in \mathcal{R}: |\mathbf{I}| = 2^{-j}\}$.

We denote by $\mathcal{A}^1$ and $\mathcal{A}^2$ the sets of all dyadic cubes $I_1 \in \mathcal{D}^1$, $I_2 \in \mathcal{D}^2$ such that $\mathbf{I} = I_1 \times I_2 \in \Upsilon_m$. Further, we denote by $\mathcal{B}^i$ $(i = 1, 2)$ the collection of all nonempty sets $\Omega_i \subset \mathbb{R}^{n_i}$ of the form

$$\Omega_i = I_i \backslash \cup \big\{J_i : J_i \in \mathcal{A}^i, J_i \subset I_i\big\}, \quad I_i \in \mathcal{A}^i. \tag{7.19}$$

Thus each set $\Omega_i \in \mathcal{B}^i$, $\Omega_i \ne \emptyset$, is obtained by subtracting from a dyadic cube $I_i \in \mathcal{A}^i$ all smaller dyadic cubes from $\mathcal{A}^i$ that are contained in I_i. It is readily seen that $\mathcal{B}^i$ consists of disjoint sets and $\#\mathcal{B}^i \le m$, $i = 1, 2$.

Now, denote by $\mathcal{X}_j$ the collection of all sets $\Omega \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ of the form

$$\Omega = \Omega_1 \times \Omega_2, \quad \Omega_i \in \mathcal{B}^i, \; i = 1, 2, \quad \Omega_1 \times \Omega_2 \subset I_1 \times I_2 \in \mathfrak{R}_j,$$

where I_1, I_2 are the cubes from the definition of Ω_i in (7.19). We set $\mathcal{X} := \cup_{j=0}^{N} \mathcal{X}_j$. It is easy to see that $\mathcal{X}$ consists of sets with disjoint interiors,

$$\cup_{\mathbf{I} \in \Upsilon_m} \mathbf{I} \subset \cup_{\Omega \in \mathcal{X}} \Omega, \quad \text{and} \quad \#\mathcal{X} \le m^2. \tag{7.20}$$

Also, observe that for any $\mathbf{I} \in \Upsilon_m$ and $\Omega \in \mathcal{X}$, either $\Omega \subset \mathbf{I}$ or $\Omega \cap \mathbf{I} = \emptyset$.

Clearly, the function F can be represented in the form

$$F(x) = \sum_{\Omega\in\mathcal{X}} \mathbb{1}_\Omega(x)\big(k_0(\Omega) + 2^{1/p}k_1(\Omega) + \cdots + 2^{N/p}k_N(\Omega)\big), \tag{7.21}$$

where $k_j(\Omega)$ is the number of rectangles $\mathbf{I} \in \Upsilon_m \cap \mathfrak{R}_j$ ($|\mathbf{I}|^{-1/p} = 2^{j/p}$) that contain Ω ($0 \le k_j \le m$).

Define $m_j := \#\mathcal{X}_j$. It is readily seen that

$$\sum_{\Omega\in\mathcal{X}} |\Omega| k_j(\Omega) = \sum_{\mathbf{I}\in\Upsilon_m\cap\mathfrak{R}_j} |\mathbf{I}| = m_j 2^{-j}, \quad j = 1, \dots, N. \tag{7.22}$$

We claim that

$$|\Omega| \le 2 \cdot 2^{-j-k_j(\Omega)}, \quad \forall \Omega \in \mathcal{X}_j. \tag{7.23}$$

Indeed, let $\Omega = \Omega_1 \times \Omega_2 \in \mathcal{X}_j$ ($j \ge 0$), and assume $k_j := k_j(\Omega) \ge 2$. Then Ω is contained in k_j distinct dyadic rectangles $\mathbf{I}^1, \dots, \mathbf{I}^{k_j}$ from $\mathfrak{R}_j$; i.e., $\mathbf{I}^\nu \in \mathcal{R}$ and $|\mathbf{I}^\nu| = 2^{-j}$, $\nu = 1, \dots, k_j$. We may assume that the dyadic rectangles $\mathbf{I}^\nu = I_1^\nu \times I_2^\nu$ are indexed so that $\Omega_1 \subset I_1^1 \subset I_1^2 \subset \cdots \subset I_1^{k_j}$. Since these are nested dyadic cubes in $\mathbb{R}^{n_1}$, we get

$$|\Omega_1| \le |I_1^1| \le \frac{|I_1^{k_j}|}{2^{n_1(k_j-1)}}, \quad \text{implying} \quad |\Omega| = |\Omega_1||\Omega_2| \le \frac{|I_1^{k_j}||I_2^{k_j}|}{2^{n_1(k_j-1)}} = \frac{1}{2^j 2^{n_1(k_j-1)}},$$

where we used that $\Omega_2 \subset I_2^{k_j}$. Because $n_1 \ge 1$, the above estimate implies (7.23).

Let $K := (2p+2)\log_2 m$. From (7.21) it follows that

$$\begin{aligned} \|F\|_p^p &= \sum_{\Omega\in\mathcal{X}_m} |\Omega| \left(\sum_{j=0}^{N} 2^{j/p}k_j(\Omega)\right)^p \le 2^p \sum_{\Omega\in\mathcal{X}_m} |\Omega| \left(\sum_{k_j(\Omega)\le K} 2^{j/p}k_j(\Omega)\right)^p \\ &\quad + 2^p \sum_{\Omega\in\mathcal{X}_m} |\Omega| \left(\sum_{k_j(\Omega)>K} 2^{j/p}k_j(\Omega)\right)^p =: Q_1 + Q_2. \end{aligned} \tag{7.24}$$

We next estimate Q_2. Using (7.20), (7.23), and that $k_j(\Omega) \le m$, we obtain

$$Q_2 \le 2^{p+1} \sum_{\Omega\in\mathcal{X}_m} \left(\sum_{k_j(\Omega)>K} \frac{2^{j/p}k_j(\Omega)}{2^{j/p+k_j(\Omega)/p}}\right)^p \le 2^{p+1} m^2 \Big(\frac{m^2}{2^{K/p}}\Big)^p = 2^{p+1}. \tag{7.25}$$

To estimate Q_1 we set $x_j := k_j/K$, $k_j := k_j(\Omega)$, and use Lemma 7.10 to obtain

$$\left(\sum_{k_j\le K} 2^{j/p}k_j\right)^p = K^p\left(\sum_{k_j\le K} 2^{j/p}x_j\right)^p \le (6p)^p K^p \sum_{k_j\le K} 2^j x_j = (6p)^p K^{p-1}\sum_{k_j\le K} 2^j k_j.$$

Here we used that $0\le x_j\le 1$ and $p>1$. The above and (7.22) lead to

$$Q_1 \le (12p)^p K^{p-1}\sum_{\Omega\in\mathcal{X}_m}|\Omega|\sum_{j=0}^N 2^j k_j(\Omega) = (12p)^p K^{p-1}\sum_{j=0}^N\sum_{\Omega\in\mathcal{X}_m}|\Omega|2^j k_j(\Omega)$$
$$\le (12p)^p K^{p-1}\sum_{j=0}^N m_j = c(p)m(\log_2 m)^{p-1}.$$

This combined with (7.24) and (7.25) yields (7.14).

It remains to show that estimate (7.14) is sharp. For simplicity, we consider the case when $n_1=n_2=1$. Fix $N\in\mathbb{N}$ sufficiently large and let $\mathcal{Y}_j$, $j=0,\dots,N$, be the set of all dyadic rectangles $\mathbf{I}=I_1\times I_2\subset[0,1]^2$ such that $|I_1|=2^{-j}$ and $|I_2|=2^{-N+j}$; hence $|\mathbf{I}|=2^{-N}$. Consider the function

$$F(x) := \sum_{j=0}^N\sum_{\mathbf{I}\in\mathcal{Y}_j}|\mathbf{I}|^{-1/p}\mathbb{1}_{\mathbf{I}}(x),\quad 1<p<\infty.$$

Clearly, $\#\mathcal{Y}_j = 2^N$ and hence $m := \#\big(\cup_{j=0}^N\mathcal{Y}_j\big) = 2^N(N+1)$. Furthermore, $F(x) = 2^{N/p}(N+1)\mathbb{1}_{[0,1)^2}$, implying

$$\|F\|_p = 2^{N/p}(N+1) = m^{1/p}(N+1)^{1-1/p}\sim m^{1/p}(\ln m)^{1-1/p}.$$

Therefore, estimate (7.14) is sharp. The proof of Lemma 7.9 is complete. □

Lemma 7.11 *If* $F=\sum_{(\boldsymbol{I},\boldsymbol{e})\in\mathcal{A}_m} a_{I e}\psi_{\boldsymbol{I}}^{\boldsymbol{e}}$, *where* $\mathcal{A}_m\subset\mathcal{R}\times E$, $\#\mathcal{A}_m\le m$, $m\in\mathbb{N}$, *and* $\|a_{\boldsymbol{Ie}}\psi_{\boldsymbol{I}}^{\boldsymbol{e}}\|_{L^p}\le A$ *for all* $(\boldsymbol{I},\boldsymbol{e})\in\mathcal{A}_m$, $0<p<\infty$, *then*

$$\|F\|_p \le cm^{1/p}A,\quad 0<p\le 2, \tag{7.26}$$
$$\|F\|_p \le cm^{1/p}(\ln m)^{1/2-1/p}A,\quad 2<p<\infty,\ m\ge 2. \tag{7.27}$$

Proof The proof of inequality (7.26) is just like the proof of [4, Lemma 4.1], and the proof of (7.27) is carried out along the same lines but uses Lemma 7.9. For completeness we next give the details.

From $\|\psi_{\mathbf{I}}^{\mathbf{e}}\|_{L^p}\sim|\mathbf{I}|^{1/p-1/2}$ (see (6.6)) and the condition $\|a_{\mathbf{Ie}}\psi_{\mathbf{I}}^{\mathbf{e}}\|_{L^p}\le A$, it follows that $|a_{\mathbf{Ie}}|\le cA|\mathbf{I}|^{1/2-1/p}$. Define $\mathcal{B}_m := \{\mathbf{I}\in\mathcal{R}\colon(\mathbf{I},\mathbf{e})\in\mathcal{A}_m\}$ and observe that

$\#\mathcal{B}_m \le \#\mathcal{A}_m \le 2^{n_1+n_2}\#\mathcal{B}_m$. We now use that $\|F\|_p \sim \|F\|_{\dot{F}^0_{p2}}$ and the wavelet characterization of $\dot{F}^0_{p2}$ to obtain

$$\begin{aligned}\|F\|_p &\le c\Big\|\Big(\sum_{(\mathbf{I},\mathbf{e})\in\mathcal{A}_m}\big[|a_{\mathbf{Ie}}||\mathbf{I}|^{-1/2}\mathbb{1}_{\mathbf{I}}(\cdot)\big]^2\Big)^{1/2}\Big\|_{L^p}\\ &\le cA\Big\|\Big(\sum_{\mathbf{I}\in\mathcal{B}_m}\big[|\mathbf{I}|^{-1/p}\mathbb{1}_{\mathbf{I}}(\cdot)\big]^2\Big)^{1/2}\Big\|_{L^p}. \qquad (7.28)\end{aligned}$$

(a) Let $0 < p \le 2$. Then from the above,

$$\begin{aligned}\|F\|_p &\le cA\Big\|\Big(\sum_{\mathbf{I}\in\mathcal{B}_m}\big[|\mathbf{I}|^{-1/p}\mathbb{1}_{\mathbf{I}}(\cdot)\big]^p\Big)^{1/p}\Big\|_{L^p}\\ &= cA\Big\|\sum_{\mathbf{I}\in\mathcal{B}_m}|\mathbf{I}|^{-1}\mathbb{1}_{\mathbf{I}}(\cdot)\Big\|_{L^1}^{1/p} = cA\big(\#\mathcal{B}_m\big)^{1/p} \le cAm^{1/p},\end{aligned}$$

which confirms (7.26).

(b) Let $2 < p < \infty$. From (7.28) it follows that

$$\begin{aligned}\|F\|_p &\le cA\Big\|\Big(\sum_{\mathbf{I}\in\mathcal{B}_m}|\mathbf{I}|^{-2/p}\mathbb{1}_{\mathbf{I}}(\cdot)\Big)^{1/2}\Big\|_{L^p}\\ &= cA\Big(\Big\|\sum_{\mathbf{I}\in\mathcal{B}_m}|\mathbf{I}|^{-2/p}\mathbb{1}_{\mathbf{I}}(\cdot)\Big\|_{L^{p/2}}\Big)^{1/2}\\ &\le cA\big[(\#\mathcal{B}_m)^{2/p}(\ln\#\mathcal{B}_m)^{1-2/p}\big]^{1/2} \le cm^{1/p}(\ln m)^{1/2-1/p}A.\end{aligned}$$

Here for the former inequality we used Lemma 7.9 with p replaced by $p/2 > 1$. Therefore, (7.27) is valid. □

Proof of Proposition 7.1 and Theorem 7.2 This proof uses well-known ideas, see, e.g., [6, Corollary 1, p. 117] or [4, Corollary 4.1] or [21, Theorem 6.2]. Define

$$N(f) := \Big(\sum_{(\mathbf{I},\mathbf{e})\in\mathcal{R}\times E}\|a_{\mathbf{Ie}}\psi^{\mathbf{e}}_{\mathbf{I}}\|_p^\tau\Big)^{1/\tau}, \quad a_{\mathbf{Ie}} := \langle f, \psi^{\mathbf{e}}_{\mathbf{I}}\rangle. \qquad (7.29)$$

We may assume $N(f) > 0$. Further, we introduce the notation

$$\mathcal{X}_r := \Big\{(\mathbf{I},\mathbf{e}) \in \mathcal{R}\times E: 2^{-r}N(f) \le \|a_{\mathbf{Ie}}\psi^{\mathbf{e}}_{\mathbf{I}}\|_p < 2^{-r+1}N(f)\Big\}, \quad r \in \mathbb{N}_0, \qquad (7.30)$$

and set $\mathcal{J}_r := \{\mathbf{I} \in \mathcal{R}: (\mathbf{I},\mathbf{e}) \in \mathcal{X}_r\}$. Observe that $\#\mathcal{J}_r \le \#\mathcal{X}_r \le 2^{n_1+n_2}(\#\mathcal{J}_r)$. Clearly,

$$\cup_{r\le\nu}\mathcal{X}_r := \Big\{(\mathbf{I},\mathbf{e}) \in \mathcal{R}\times E: \|a_{\mathbf{Ie}}\psi^{\mathbf{e}}_{\mathbf{I}}\|_p \ge 2^{-\nu}N(f)\Big\}.$$

From this and (7.29) it follows that

$$\#\mathcal{X}_\nu \le \#(\cup_{r\le\nu}\mathcal{X}_r) \le 2^{\nu\tau}. \tag{7.31}$$

Define

$$G_\nu := \sum_{(\mathbf{I},\mathbf{e})\in\mathcal{X}_\nu} a_{\mathbf{I}\mathbf{e}}\psi_{\mathbf{I}}^{\mathbf{e}}.$$

From (7.30) it follows that $\|a_{\mathbf{I}\mathbf{e}}\psi_{\mathbf{I}}^{\mathbf{e}}\|_p < 2^{-\nu+1}N(f)$ for $(\mathbf{I},\mathbf{e})\in\mathcal{X}_\nu$. We now invoke Lemma 7.11 to obtain in the case $2 < p < \infty$,

$$\|G_\nu\|_{L^p} \le c2^{-\nu}N(f)(\#\mathcal{I}_\nu)^{1/p}\big(\ln(\#\mathcal{I}_\nu)\big)^{1/2-1/p} \le cN(f)2^{-\nu(1-\tau/p)}\nu^{1/2-1/p}. \tag{7.32}$$

In the case $0 < p \le 2$, we appeal again to Lemma 7.11 and obtain

$$\|G_\nu\|_p \le cN(f)2^{-\nu(1-\tau/p)}, \quad 0 < p \le 2. \tag{7.33}$$

Using (7.32) we obtain for $2 < p < \infty$,

$$\begin{aligned}\Big\|\sum_{\nu\ge j}G_\nu\Big\|_{L^p} \le \sum_{\nu\ge j}\|G_\nu\|_{L^p} \le cN(f)\sum_{\nu\ge j}2^{-\nu(1-\tau/p)}\nu^{1/2-1/p}\\ \le cN(f)2^{-j(1-\tau/p)}(j+1)^{1/2-1/p},\end{aligned} \tag{7.34}$$

and similarly from (7.33) it follows that

$$\Big\|\sum_{\nu\ge j}G_\nu\Big\|_{L^p} \le \sum_{\nu\ge j}\|G_\nu\|_{L^p} \le cN(f)2^{-j(1-\tau/p)}, \quad 1 < p \le 2. \tag{7.35}$$

In the case $0 < p \le 1$, we use (7.33) to obtain

$$\Big\|\sum_{\nu\ge j}G_\nu\Big\|_{H^p} \le \Big(\sum_{\nu\ge j}\|G_\nu\|_{H^p}^p\Big)^{1/p} \le cN(f)\Big(\sum_{\nu\ge j}2^{-\nu p(1-\tau/p)}\Big)^{1/p},$$

implying

$$\Big\|\sum_{\nu\ge j}G_\nu\Big\|_{H^p} \le cN(f)2^{-j(1-\tau/p)}, \quad 0 < p \le 1. \tag{7.36}$$

Estimates (7.34)–(7.36) with $j = 0$ readily imply (7.4)–(7.5).

Assume $2 < p < \infty$ and $m \ge 2$. Choose $j \in \mathbb{N}_0$ so that $2^{j\tau} \le m < 2^{(j+1)\tau}$. Define $\mathcal{Y}_j := \cup_{\nu\le j}\mathcal{X}_\nu$. By (7.31) $\#\mathcal{Y}_j \le 2^{j\tau}$ and using (7.34) and the fact that $f = \sum_{(\mathbf{I},\mathbf{e})\in\mathcal{R}\times E} a_{\mathbf{I}\mathbf{e}}\psi_{\mathbf{I}}^{\mathbf{e}}$, where the series converges unconditionally in L^p, we obtain

$$\begin{aligned}\sigma_m(f)_p \le \Big\|f - \sum_{(\mathbf{I},\mathbf{e})\in\mathcal{Y}_j} a_{\mathbf{I}\mathbf{e}}\psi_{\mathbf{I}}^{\mathbf{e}}\Big\|_{L^p} \le \Big\|\sum_{\nu>j}G_\nu\Big\|_{L^p}\\ \le cN(f)2^{-j\tau(1/\tau-1/p)}(j+1)^{1/2-1/p} \le cm^{-\alpha}(\ln m)^{1/2-1/p}\|f\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}},\end{aligned}$$

which confirms (7.7).

In the case when $1 < p \le 2$ or $0 < p \le 1$ exactly as above, we use (7.35) or (7.36) instead of (7.34) to obtain (7.6). □

Proof of Theorem 7.3 This proof uses the idea of the proof of [4, Theorem 4.3]. Let $g = \sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} a_{\mathbf{Ie}}\psi_{\mathbf{I}}^{\mathbf{e}}$, where $\#\Lambda_m \le m$. There are two cases to distinguish.

Case 1 $2 \le p < \infty$. Using (7.3) and Hölder's inequality, we obtain

$$\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \le c\Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} \|a_{\mathbf{Ie}}\psi_{\mathbf{I}}^{\mathbf{e}}\|_p^\tau\Big)^{1/\tau} \le m^{1/\tau-1/p}\Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} \|a_{\mathbf{Ie}}\psi_{\mathbf{I}}^{\mathbf{e}}\|_p^p\Big)^{1/p}.$$

On the other hand, from (6.6), $\|\psi_{\mathbf{I}}^{e}\|_p \sim |\mathbf{I}|^{1/p-1/2}$. Therefore,

$$\begin{aligned}
\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} &\le cm^{1/\tau-1/p}\Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} |a_{\mathbf{Ie}}|^p|\mathbf{I}|^{1-p/2}\Big)^{1/p}\\
&= cm^{1/\tau-1/p}\Big(\int_{\mathbb{R}^{n_1+n_2}} \sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} |a_{\mathbf{Ie}}|^p|\mathbf{I}|^{-p/2}\mathbb{1}_{\mathbf{I}}(\mathbf{x})d\mathbf{x}\Big)^{1/p}\\
&= cm^{1/\tau-1/p}\Big(\int_{\mathbb{R}^{n_1+n_2}} \Big[\Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} |a_{\mathbf{Ie}}|^p|\mathbf{I}|^{-p/2}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\Big)^{1/p}\Big]^p d\mathbf{x}\Big)^{1/p}\\
&\le cm^{1/\tau-1/p}\Big(\int_{\mathbb{R}^{n_1+n_2}} \Big[\Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} \big[|a_{\mathbf{Ie}}||\mathbf{I}|^{-1/2}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\big]^2\Big)^{1/2}\Big]^p d\mathbf{x}\Big)^{1/p}\\
&= cm^{1/\tau-1/p}\|Sg\|_p \le cm^\alpha\|g\|_p.
\end{aligned}$$

Here we used that $p > 2$ and the characterization of $\|g\|_p$ by the square function $Sg = \Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m}\big[|a_{\mathbf{Ie}}||\mathbf{I}|^{-1/2}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\big]^2\Big)^{1/2}$. Thus (7.9) is established.

Case 2 $0 < p < 2$. We set $\mathcal{I}_m := \{\mathbf{I} \in \mathcal{D}\colon (\mathbf{I},\mathbf{e}) \in \Lambda_m\}$. From (7.3) and $\|\psi_{\mathbf{I}}^{e}\|_p \sim |\mathbf{I}|^{1/p-1/2}$, we get

$$\begin{aligned}
\|g\|^\tau_{\dot{B}^{\mathbf{s}}_{\tau\tau}} &\le c\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} \|a_{\mathbf{Ie}}\psi_{\mathbf{I}}^{\mathbf{e}}\|_p^\tau \le c\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} |a_{\mathbf{Ie}}|^\tau|\mathbf{I}|^{-\tau(1/2-1/p)}\\
&= c\int_{\mathbb{R}^{n_1+n_2}} \sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} |a_{\mathbf{Ie}}|^\tau|\mathbf{I}|^{-\tau(1/2-1/p)-1}\mathbb{1}_{\mathbf{I}}(\mathbf{x})d\mathbf{x}\\
&= c\int_{\mathbb{R}^{n_1+n_2}} \sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m} |a_{\mathbf{Ie}}|^\tau|\mathbf{I}|^{-\tau/2}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\cdot|\mathbf{I}|^{\tau/p-1}\mathbb{1}_{\mathbf{I}}(\mathbf{x})d\mathbf{x}.
\end{aligned}$$

We now apply Hölder's inequality $\sum a_j b_j \le \left(\sum |a_j|^q\right)^{1/q}\left(\sum |b_j|^{q'}\right)^{1/q'}$ with $q = 2/\tau$ to obtain

$$\begin{aligned}\|g\|^{\tau}_{\dot{B}^{\mathbf{s}}_{\tau\tau}} &\le c\int_{\mathbb{R}^{n_1+n_2}}\Big(\sum_{(\mathbf{I},\mathbf{e})\in\Lambda_m}\Big[\frac{|a_{\mathbf{I}\mathbf{e}}|\mathbb{1}_{\mathbf{I}}(\mathbf{x})}{|\mathbf{I}|^{1/2}}\Big]^2\Big)^{\tau/2}\Big(\sum_{\mathbf{I}\in\mathfrak{I}_m}\Big(\frac{\mathbb{1}_{\mathbf{I}}(\mathbf{x})}{|\mathbf{I}|^{1-\tau/p}}\Big)^{\frac{1}{1-\tau/2}}\Big)^{1-\tau/2}d\mathbf{x}\\ &= c\int_{\mathbb{R}^{n_1+n_2}}[Sg(\mathbf{x})]^{\tau}\Big(\sum_{\mathbf{I}\in\mathfrak{I}_m}|\mathbf{I}|^{-\frac{1-\tau/p}{1-\tau/2}}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\Big)^{1-\tau/2}d\mathbf{x}.\end{aligned}$$

We next apply Hölder's inequality $\int fg \le \left(\int |f|^q\right)^{1/q}\left(\int |g|^{q'}\right)^{1/q'}$ with $q = p/\tau$ and get

$$\begin{aligned}\|g\|^{\tau}_{\dot{B}^{\mathbf{s}}_{\tau\tau}} &\le c\Big(\int_{\mathbb{R}^{n_1+n_2}}[Sg(\mathbf{x})]^p d\mathbf{x}\Big)^{\frac{\tau}{p}}\Big(\int_{\mathbb{R}^{n_1+n_2}}\Big(\sum_{\mathbf{I}\in\mathfrak{I}_m}|\mathbf{I}|^{-\frac{1-\tau/p}{1-\tau/2}}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\Big)^{\frac{1-\tau/2}{1-\tau/p}}d\mathbf{x}\Big)^{1-\frac{\tau}{p}}\\ &=: c\|Sg\|^{\tau}_p\cdot Q \le c\|g\|^{\tau}_p\cdot Q.\end{aligned}$$

To estimate Q we use Lemma 7.9 with $\frac{1-\tau/2}{1-\tau/p} > 1$ in place of p and obtain

$$\begin{aligned}Q &= \Big(\Big(\int_{\mathbb{R}^{n_1+n_2}}\Big(\sum_{\mathbf{I}\in\mathfrak{I}_m}|\mathbf{I}|^{-\frac{1-\tau/p}{1-\tau/2}}\mathbb{1}_{\mathbf{I}}(\mathbf{x})\Big)^{\frac{1-\tau/2}{1-\tau/p}}d\mathbf{x}\Big)^{\frac{1-\tau/p}{1-\tau/2}}\Big)^{1-\tau/2}\\ &\le c\Big(m^{\frac{1-\tau/p}{1-\tau/2}}(\ln m)^{1-\frac{1-\tau/p}{1-\tau/2}}\Big)^{1-\tau/2} = cm^{1-\tau/p}\big(\ln m\big)^{\tau/p-\tau/2}.\end{aligned}$$

This leads to

$$\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \le cm^{1/\tau-1/p}\big(\ln m\big)^{1/p-1/2}\|g\|_p \le cm^{\alpha}\big(\ln m\big)^{1/p-1/2}\|g\|_p.$$

The proof is complete. □

7.1 Nonlinear *m*-Term Frame Approximation

In this subsection we consider nonlinear m-term approximation from the product frame $\{\theta_{\mathbf{I}}\}_{\mathbf{I}\in\mathcal{R}}$ defined in Sect. 2.4 in L^p, $1 < p < \infty$, or H^p, $0 < p \le 1$. Denote by Ξ_m the set of all functions on $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$ of the form

$$g = \sum_{\mathbf{I}\in\mathcal{Y}_m} a_{\mathbf{I}}\theta_{\mathbf{I}}, \tag{7.37}$$

where $\mathcal{Y}_m \subset \mathcal{R}$, $\#\mathcal{Y}_m \le m$, and $\mathcal{Y}_m$ is allowed to vary with g. We define

$$\mathcal{F}_m(f) := \inf_{g\in\Xi_m}\|f-g\|_p, \tag{7.38}$$

where $\|\cdot\|_p = \|\cdot\|_{L^p}$ if $1 < p < \infty$ and $\|\cdot\|_p = \|\cdot\|_{H^p}$ if $0 < p \le 1$.

Just as in (7.2) we assume that

$$0 < p < \infty, \quad \alpha > 0, \quad \frac{1}{\tau} := \alpha + \frac{1}{p}, \quad s_1 := n_1\alpha, \quad s_2 := n_2\alpha, \quad \mathbf{s} := (s_1, s_2).$$

As in Sect. 7 the Besov space $\dot{B}^{\mathbf{s}}_{\tau\tau}$ naturally appears here.

Theorem 7.12 (Jackson estimate) *If $f \in \dot{B}^{s}_{\tau\tau}$, then*

$$\mathcal{F}_m(f)_p \le cm^{-\alpha}\|f\|_{\dot{B}^{s}_{\tau\tau}}, \quad 0 < p \le 2,$$
$$\mathcal{F}_m(f)_p \le cm^{-\alpha}(\ln m)^{1/2-1/p}\|f\|_{\dot{B}^{s}_{\tau\tau}}, \quad 2 < p < \infty, \; m \ge 2,$$

where $c > 0$ is a constant depending only on p, α, n_1, n_2.

The proof of this theorem is almost identical to the proof of Theorem 7.2 because all ingredients needed for this proof are in place; we omit it.

We conjecture that the analogs of the Bernstein inequalities (7.8) and (7.9) are valid. The main obstacle in proving these inequalities is that unlike the basis $\{\psi^{\mathbf{e}}_{\mathbf{I}}\}$ the frame $\{\theta_{\mathbf{I}}\}$ is redundant and hence the norm $\|g\|_p$ (even when $p = 2$) of g from (7.37) cannot be estimated from below by any reasonable quantity in terms of the coefficients $\{a_{\mathbf{I}}\}$.

8 Appendix

8.1 Proof of Lemma 4.3

We only consider the case when $\nu_1 \le \mu_1$ and $\nu_2 > \mu_2$; the proof in all other cases is similar. Under the hypothesis of the lemma we assume that $x_{\mathbf{J}} = 0$. We next split $\mathcal{R}$ into a disjoint union of subsets. We define

$$\Omega_{\mathbf{k}} := \{\mathbf{I} \in \mathcal{R}_{\boldsymbol{\mu}} \colon 2^{k_1-1-\nu_1} < |x_{I_1}| \le 2^{k_1-\nu_1} \;\text{ and }\; 2^{k_2-1-\mu_2} < |x_{I_2}| \le 2^{k_2-\mu_2}\},$$

if $\mathbf{k} = (k_1, k_2) \in \mathbb{N}^2$ and set

$$\Omega_{(k_1,0)} := \{\mathbf{I} \in \mathcal{R}_{\boldsymbol{\mu}} \colon 2^{k_1-1-\nu_1} < |x_{I_1}| \le 2^{k_1-\nu_1} \;\text{ and }\; |x_{I_2}| \le 2^{-\mu_2}\}, \quad k_1 \in \mathbb{N},$$
$$\Omega_{(0,k_2)} := \{\mathbf{I} \in \mathcal{R}_{\boldsymbol{\mu}} \colon |x_{I_1}| \le 2^{-\nu_1} \;\text{ and }\; 2^{k_2-1-\mu_2} < |x_{I_2}| \le 2^{k_2-\mu_2}\}, \quad k_2 \in \mathbb{N},$$
$$\Omega_{(0,0)} := \{P \in \mathcal{R}_{\boldsymbol{\mu}} \colon |x_{I_1}| \le 2^{-\nu_1} \;\text{ and }\; |x_{I_2}| \le 2^{-\mu_2}\}.$$

Evidently, $\mathcal{R}_{\boldsymbol{\mu}} = \cup_{\mathbf{k}\in\mathbb{N}_0^2}\Omega_{\mathbf{k}}$ and the sets $\Omega_{\mathbf{k}}$, $\mathbf{k} \in \mathbb{N}_0^2$, are disjoint.

Let $\mathbf{I} = I_1 \times I_2 \in \Omega_{\mathbf{k}}$. From the preceding, it readily follows that $1 + 2^{\nu_1}|x_{I_1}| > 2^{k_1-1}$ and $1 + 2^{\mu_2}|x_{I_2}| > 2^{k_2-1}$, and hence

$$\prod_{i=1}^{2}\left(1 + 2^{\min\{\mu_i,\nu_i\}}|x_{I_i} - x_{J_i}|\right)^{-\tau_i} \le c2^{-k_1\tau_1}2^{-k_2\tau_2}.$$

Therefore,

$$\begin{aligned} b_{\mathbf{J}} &:= \sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}} |a_{\mathbf{I}}| \prod_{i=1}^{2} \left(1+2^{\min\{\mu_i,\nu_i\}}|x_{I_i}-x_{J_i}|\right)^{-\tau_i} \\ &\le c \sum_{k_2=0}^{\infty} 2^{-k_2\tau_2} \sum_{k_1=0}^{\infty} 2^{-k_1\tau_1} \sum_{\mathbf{I}\in\Omega_k} |a_{\mathbf{I}}| \\ &\le c \sum_{k_2=0}^{\infty} 2^{-k_2\tau_2} \sum_{k_1=0}^{\infty} 2^{-k_1\tau_1} \left(\sum_{\mathbf{I}\in\Omega_k} |a_{\mathbf{I}}|^t\right)^{1/t}, \end{aligned} \tag{8.1}$$

where for the last inequality we used that $0 < t \le 1$.

The set $\mathcal{R}_{\boldsymbol{\mu}}$ is a disjoint partition of $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and hence we have

$$\sum_{\mathbf{I}\in\Omega_{\mathbf{k}}} |a_{\mathbf{I}}|^t = \int_{\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}} \Big(\sum_{\mathbf{I}\in\Omega_{\mathbf{k}}} |a_{\mathbf{I}}||\mathbf{I}|^{-1/t}\mathbb{1}_{\mathbf{I}}(\mathbf{y})\Big)^t d\mathbf{y}. \tag{8.2}$$

Let $\mathbf{I} \in \Omega_{\mathbf{k}}$ and $\mathbf{y} = (y_1, y_2) \in \mathbf{I}$. Then since $\nu_1 \le \mu_1$ and $\nu_2 > \mu_2$, we have

$$|y_1| \le |y_1 - x_{I_1}| + |x_{I_1}| \le 2^{-\mu_1} + 2^{k_1-\nu_1} < 2 \cdot 2^{k_1-\nu_1}$$

and

$$|y_2| \le |y_2 - x_{I_2}| + |x_{I_2}| \le 2^{-\mu_2} + 2^{k_2-\nu_2} < 2 \cdot 2^{k_2-\mu_2}.$$

Therefore,

$$\cup_{\mathbf{I}\in\Omega_{\mathbf{k}}}\mathbf{I} \subset [-3\cdot 2^{k_1-\nu_1}, 3\cdot 2^{k_1-\nu_1}]^{n_1} \times [-3\cdot 2^{k_2-\mu_2}, 3\cdot 2^{k_2-\mu_2}]^{n_2} =: R.$$

Because $x_{\mathbf{J}} = 0$, $\mathbf{J} \in \mathcal{R}_{\boldsymbol{\nu}}$, and since $\nu_2 > \mu_2$, it follows that $\mathbf{J} \subset R$ and hence $x \in R$. From (8.2) and the definition of the maximal operator $\mathcal{M}_t$ in (2.18), it follows that for any $x \in \mathbf{J} \subset R$,

$$\begin{aligned} \sum_{\mathbf{I}\in\Omega_{\mathbf{k}}} |a_{\mathbf{I}}|^t &= \int_R 2^{\boldsymbol{\mu}\cdot\mathbf{n}} \Big(\sum_{\mathbf{I}\in\Omega_{\mathbf{k}}} |a_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}(\mathbf{y})\Big)^t d\mathbf{y} \\ &\le c|R|2^{\boldsymbol{\mu}\cdot\mathbf{n}}\Big[\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\Omega_{\mathbf{k}}} |a_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)(\mathbf{x})\Big]^t \\ &= c2^{\mathbf{k}\cdot\mathbf{n}}2^{(\mu_1-\nu_1)n_1}\Big[\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\Omega_{\mathbf{k}}} |a_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)(\mathbf{x})\Big]^t. \end{aligned}$$

This coupled with (8.1) yields

$$
\begin{aligned}
b_{\mathbf{J}} &\le c2^{(\mu_1-\nu_1)n_1/t}\sum_{k_2=0}^{\infty}2^{-k_2(\tau_2-n_2/t)}\sum_{k_1=0}^{\infty}2^{-k_1(\tau_1-n_1/t)}\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\Omega_{\mathbf{k}}}|a_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)(\mathbf{x})\\
&\le c2^{(\mu_1-\nu_1)n_1/t}\mathcal{M}_t\Big(\sum_{\mathbf{I}\in\mathcal{R}_{\boldsymbol{\mu}}}|a_{\mathbf{I}}|\mathbb{1}_{\mathbf{I}}\Big)(\mathbf{x}),\quad \mathbf{x}\in\mathbf{J},
\end{aligned}
$$

where we used that $\tau_i > n_i/t$, $i = 1, 2$. This confirms inequality (4.4) in the case under consideration. The proof of Lemma 4.3 is complete. □

8.2 Proof of Lemma 6.3

Let $K \in \mathbb{N}$ and $M \ge K + n_1 + n_2 + 1$. We will consider only the case when $\mathbf{I} = I_1 \times I_2 \in \mathcal{R}_{\boldsymbol{\nu}}$, $\boldsymbol{\nu} = (\nu_1, \nu_2)$ with $\nu_1 \ge 0$ and $\nu_2 < 0$. The proof in the other cases is carried out similarly.

Assume $f \in \mathcal{S}_\infty(\mathbb{R}^{n_1}\times\mathbb{R}^{n_2})$ and consider the Taylor polynomials

$$
\begin{aligned}
P_1(x_1,x_2) &:= \sum_{|\beta_1|\le K-1}\frac{\partial_{x_1}^{\beta_1}f(x_{I_1},x_2)}{(\beta_1)!}(x_1-x_{I_1})^{\beta_1},\\
P_2(x_2) &:= \sum_{|\beta_2|\le K-1}\frac{\partial^{\beta_2}\psi_{I_2}^{e_2}(0_2)}{(\beta_2)!}x_2^{\beta_2},
\end{aligned}
$$

where $\beta_1 \in \mathbb{N}_0^{n_1}$, $\beta_2 \in \mathbb{N}_0^{n_2}$, and 0_2 stands for the origin in $\mathbb{R}^{n_2}$. From (2.2) and (6.5), it follows that

$$
\langle f,\psi_{\mathbf{I}}^{e}\rangle = \int_{\mathbb{R}^{n_1+n_2}}\big[f(x_1,x_2)-P_1(x_1,x_2)\big]\psi_{I_1}^{e_1}(x_1)\big[\psi_{I_2}^{e_2}(x_2)-P_2(x_2)\big]\mathrm{d}x_1\mathrm{d}x_2.
$$

Assume $|x_1 - x_{I_1}| \le 1$. Then by Taylor's theorem and (2.1),

$$
\begin{aligned}
|f(x_1,x_2)-P_1(x_1,x_2)| &\le c|x_1-x_{I_1}|^K\sup_{|z_1-x_{I_1}|\le|x_1-x_{I_1}|}\max_{|\beta_1|=K}|\partial_{x_1}^{\beta_1}f(z_1,x_2)|\\
&\le c|x_1-x_{I_1}|^K\sup_{|z_1-x_{I_1}|\le|x_1-x_{I_1}|}\frac{1}{(1+|z_1|)^M(1+|x_2|)^M}\\
&\le \frac{c|x_1-x_{I_1}|^K}{(1+|x_{I_1}|)^M(1+|x_2|)^M}. \qquad (8.3)
\end{aligned}
$$

Here we used that $|x_{I_1}| \le |z_1| + |z_1 - x_{I_1}| \le |z_1| + |x_1 - x_{I_1}| \le |z_1| + 1$. On the other hand, for $|x_1 - x_{I_1}| \ge 1$, using again (2.1) we get

$$|P_1(x_1, x_2)| \le c \sum_{|\beta_1| \le K-1} |\partial_{x_1}^{\beta_1} f(x_{I_1}, x_2)||x_1 - x_{I_1}|^{|\beta_1|} \le \frac{c|x_1 - x_{I_1}|^K}{(1 + |x_{I_1}|)^M (1 + |x_2|)^M}. \tag{8.4}$$

We have

$$\begin{aligned} |\langle f, \psi_{\mathbf{I}}^e \rangle| \le& \int_{\mathbb{R}^{n_2}} \int_{|x_1 - x_{I_1}| \le 1} |f(x_1, x_2) - P_1(x_1, x_2)||\psi_{I_1}^{e_1}(x_1)| \cdots \mathrm{d}x_1 \mathrm{d}x_2 \\ &+ \int_{\mathbb{R}^{n_2}} \int_{|x_1 - x_{I_1}| \ge 1} (|f(x_1, x_2)| + |P_1(x_1, x_2)|)|\psi_{I_1}^{e_1}(x_1)| \cdots \mathrm{d}x_1 \mathrm{d}x_2. \end{aligned}$$

From this, (8.3), (8.4), and (2.1), i.e., $|f(x_1, x_2)| \le c(1 + |x_1|)^{-M}(1 + |x_2|)^{-M}$, it follows that

$$\begin{aligned} |\langle f, \psi_{\mathbf{I}}^e \rangle| \le& c \left(\int_{|x_1 - x_{I_1}| \ge 1} \frac{|\psi_{I_1}^{e_1}(x_1)|}{(1 + |x_1|)^M} \mathrm{d}x_1 + \int_{\mathbb{R}^{n_1}} \frac{|x_1 - x_{I_1}|^K |\psi_{I_1}^{e_1}(x_1)|}{(1 + |x_{I_1}|)^M} \mathrm{d}x_1 \right) \\ &\times \int_{\mathbb{R}^{n_2}} \frac{|\psi_{I_2}^{e_2}(x_2) - P_2(x_2)|}{(1 + |x_2|)^M} \mathrm{d}x_2 =: c(J_1 + J_2) J_3. \end{aligned}$$

To estimate J_1 we use (6.4) and obtain

$$\begin{aligned} J_1 &\le c \int_{|x_1 - x_{I_1}| \ge 1} \frac{2^{\nu_1 n_1/2}}{(1 + |x_1|)^M (1 + 2^{\nu_1}|x_1 - x_{I_1}|)^M} \mathrm{d}x_1 \\ &\le c \int_{\mathbb{R}^{n_1}} \frac{2^{-\nu_1(M - n_1/2)}}{(1 + |x_1|)^M (1 + |x_1 - x_{I_1}|)^M} \mathrm{d}x_1 \le \frac{c 2^{-\nu_1(M - n_1/2)}}{(1 + |x_{I_1}|)^M}. \end{aligned} \tag{8.5}$$

We now estimate J_2. Using again (6.4) we get

$$\begin{aligned} J_2 &\le \frac{c}{(1 + |x_{I_1}|)^M} \int_{\mathbb{R}^{n_1}} \frac{2^{\nu_1 n_1/2}|x_1 - x_{I_1}|^K}{(1 + 2^{\nu_1}|x_1 - x_{I_1}|)^M} \mathrm{d}x_1 \\ &\le \frac{c 2^{-\nu_1(K + n_1/2)}}{(1 + |x_{I_1}|)^M} \int_{\mathbb{R}^{n_1}} \frac{2^{\nu_1 n_1}}{(1 + 2^{\nu_1}|x_1 - x_{I_1}|)^{M-K}} \mathrm{d}x_1 \le \frac{c 2^{-\nu_1(K + n_1/2)}}{(1 + |x_{I_1}|)^M}, \end{aligned} \tag{8.6}$$

where we used that $M - K > n_1$.

In estimating J_3 we proceed similarly as above. Assume $|x_2| \le 2^{-\nu_2}$. Then by Taylor's theorem and (6.4),

$$|\psi_{I_2}^{e_2}(x_2) - P_2(x_2)| \le c|x_2|^K \sup_{|z_2|\le|x_2|} \max_{|\beta_2|=K} |\partial^{\beta_2}\psi_{I_2}^{e_2}(z_2)|$$
$$\le c|x_2|^K \sup_{|z_2|\le|x_2|} \frac{2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|z_2-x_{I_2}|)^M} \le \frac{c|x_2|^K 2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}. \tag{8.7}$$

Here we used that $|x_{I_2}| \le |z_2 - x_{I_2}| + |z_2| \le |z_2 - x_{I_2}| + |x_2| \le |z_2 - x_{I_2}| + 2^{-\nu_2}$ implying $2^{\nu_2}|x_{I_2}| \le 2^{\nu_2}|z_2 - x_{I_2}| + 1$. Further, for $|x_2| \ge 2^{-\nu_2}$ using (6.4), we have

$$|P_2(x_2)| \le c \sum_{|\beta_2|\le K-1} |\partial^{\beta_2}\psi_{I_2}^{e_2}(0_2)||x_2|^{|\beta_2|}$$
$$\le c \sum_{|\beta_2|\le K-1} \frac{2^{\nu_2(|\beta_2|+n_2/2)}|x_2|^{|\beta_2|}}{(1+2^{\nu_2}|x_{I_2}|)^M} \le \frac{c(|x_2|2^{\nu_2})^K 2^{\nu_2 n_2/2}}{(1+2^{\nu_2}|x_{I_2}|)^M}. \tag{8.8}$$

To estimate J_3 we introduce the sets:

$$U_1 := \{x_2 : |x_2| \le 2^{-\nu_2}\}, \qquad U_2 := \{x_2 : |x_2| \ge 2^{-\nu_2},\ |x_2 - x_{I_2}| \le |x_{I_2}|/2\},$$
$$U_3 := \{x_2 : |x_2| \ge 2^{-\nu_2},\ |x_2 - x_{I_2}| \ge |x_{I_2}|/2\}, \qquad \mathbb{R}^{n_2} = U_1 \cup U_2 \cup U_3.$$

We have

$$J_3 := \int_{\mathbb{R}^{n_2}} \frac{|\psi_{I_2}^{e_2}(x_2) - P_2(x_2)|}{(1+|x_2|)^M} dx_2 = \int_{U_1} + \int_{U_2} + \int_{U_3}.$$

Using (8.7) and that $M - K > n_2$, we get

$$\int_{U_1} \le \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M} \int_{\mathbb{R}^{n_2}} \frac{|x_2|^K}{(1+|x_2|)^M} dx_2 \le \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}. \tag{8.9}$$

Now from (6.4) and (8.8) we infer that

$$\int_{U_2} \le \int_{U_2} \frac{|\psi_{I_2}^{e_2}(x_2)| + |P_2(x_2)|}{(1+|x_2|)^M} dx_2 \le c \int_{U_2} \frac{2^{\nu_2 n_2/2}}{(1+|x_2|)^M(1+2^{\nu_2}|x_2 - x_{I_2}|)^M} dx_2$$
$$+ c \int_{U_2} \frac{c(|x_2|2^{\nu_2})^K 2^{\nu_2 n_2/2}}{(1+|x_2|)^M(1+2^{\nu_2}|x_{I_2}|)^M} dx_2.$$

Note that if $x_2 \in U_2$, then $|x_2| \ge |x_{I_2}| - |x_2 - x_{I_2}| \ge |x_{I_2}|/2$, and because $|x_2| \ge 2^{-\nu_2}$, we have $1 + |x_2| \ge 2^{-\nu_2-1}(1 + 2^{\nu_2}|x_2|) \ge (1/4)2^{-\nu_2}(1 + 2^{\nu_2}|x_{I_2}|)$. We use this in estimating the first integral on the right above. We get

$$\int_{U_2} \le \frac{c2^{\nu_2(M-n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}\int_{\mathbb{R}^{n_2}} \frac{2^{\nu_2 n_2}}{(1+2^{\nu_2}|x_2-x_{I_2}|)^M}\mathrm{d}x_2 + \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}\int_{\mathbb{R}^{n_2}} \frac{|x_2|^K}{(1+|x_2|)^M}\mathrm{d}x_2 \le \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}. \tag{8.10}$$

Here we used that $M-K>n_2$. Again by (6.4) and (8.8) we get

$$\int_{U_3} \le \int_{U_2} \frac{|\psi_{I_2}^{e_2}(x_2)|+|P_2(x_2)|}{(1+|x_2|)^M}\mathrm{d}x_2 \le c\int_{U_3} \frac{2^{\nu_2 n_2/2}}{(1+|x_2|)^M(1+2^{\nu_2}|x_2-x_{I_2}|)^M}\mathrm{d}x_2 + c\int_{U_3} \frac{c(|x_2|2^{\nu_2})^K 2^{\nu_2 n_2/2}}{(1+|x_2|)^M(1+2^{\nu_2}|x_{I_2}|)^M}\mathrm{d}x_2.$$

If $x_2\in U_3$, then $|x_2-x_{I_2}|\ge |x_{I_2}|/2$ and $1+|x_2|>2^{-\nu_2}$. We use these inequalities and the fact that $M-K>n_2$ to obtain

$$\begin{aligned}\int_{U_3} &\le \frac{c2^{\nu_2 n_2/2}}{(1+2^{\nu_2}|x_{I_2}|)^M}\int_{\mathbb{R}^{n_2}} \frac{1}{(1+|x_2|)^M}\mathrm{d}x_2 + \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}\int_{\mathbb{R}^{n_2}} \frac{|x_2|^K}{(1+|x_2|)^M}\mathrm{d}x_2\\ &\le \frac{c2^{\nu_2 n_2/2}}{(1+2^{\nu_2}|x_{I_2}|)^M}\int_{\mathbb{R}^{n_2}} \frac{2^{\nu_2 K}}{(1+|x_2|)^{M-K}}\mathrm{d}x_2 + \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}\\ &\le \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}.\end{aligned}$$

This along with (8.9) and (8.10) yields

$$J_3 \le \frac{c2^{\nu_2(K+n_2/2)}}{(1+2^{\nu_2}|x_{I_2}|)^M}.$$

In turn, this and (8.5)–(8.6) lead to

$$|\langle f,\psi_{\mathbf{I}}^e\rangle| \le \frac{c2^{-\nu_1(K-n_1/2)+\nu_2(K+n_2/2)}}{(1+|x_{I_1}|)^M(1+2^{\nu_2}|x_{I_2}|)^M}.$$

Since $M,K>0$ with $M\ge K+n_1+n_2+1$ can be arbitrarily large, the above estimate implies (6.8). The proof is complete. □

8.3 Proof of Proposition 6.5

We will use ideas from [13, Remerk B.4]. To prove part (a) we first use the product version of the standard Plancherel–Polya–Nikolskii inequality (see, e.g., [24, Theorem 1.6.2]) to obtain

$$\|\varphi_{\boldsymbol{\nu}} * f\|_\infty \le c2^{(\nu_1 n_1+\nu_2 n_2)/p}\|\varphi_\nu * f\|_p \le c2^{(\nu_1 n_1+\nu_2 n_2)/p}\|f\|_{\dot{F}^0_{p2}}. \tag{8.11}$$

Set $b_1 := n_1/2n_2$ and $b_2 := n_2/2n_1$. Define

$$\mathcal{X}_1 := \{\boldsymbol{\nu} \in \mathbb{Z}^2 : \nu_1 \le 0,\ \nu_2 \le -b_1\nu_1\}, \quad \mathcal{X}_2 := \{\boldsymbol{\nu} \in \mathbb{Z}^2 : \nu_2 \le 0,\ \nu_1 \le -b_2\nu_2\},$$
$$\mathcal{Y}_1 := \{\boldsymbol{\nu} \in \mathbb{Z}^2 : \nu_1 > 0,\ \nu_2 \ge -b_2^{-1}\nu_1\}, \quad \mathcal{Y}_2 := \{\boldsymbol{\nu} \in \mathbb{Z}^2 : \nu_2 > 0,\ \nu_1 \ge -b_1^{-1}\nu_2\}.$$

Evidently, $\mathbb{Z}^2 = \mathcal{X}_1 \cup \mathcal{X}_2 \cup \mathcal{Y}_1 \cup \mathcal{Y}_2$. From (8.11) it follows that

$$\begin{aligned}\sum_{\boldsymbol{\nu}\in\mathcal{X}_1} \|\varphi_{\boldsymbol{\nu}} * f\|_\infty &\le c\|f\|_{\dot{F}^0_{p2}} \sum_{\nu_1\le 0} 2^{\nu_1 n_1/p} \sum_{\nu_2 \le -b_1\nu_1} 2^{\nu_2 n_2/p}\\ &\le c\|f\|_{\dot{F}^0_{p2}} \sum_{\nu_1\le 0} 2^{\nu_1(n_1 - b_1 n_2)/p} = c\|f\|_{\dot{F}^0_{p2}} \sum_{\nu_1\le 0} 2^{\nu_1 n_1/2p} \le c\|f\|_{\dot{F}^0_{p2}}\end{aligned}$$

and similarly $\sum_{\boldsymbol{\nu}\in\mathcal{X}_2} \|\varphi_{\boldsymbol{\nu}} * f\|_\infty \le c\|f\|_{\dot{F}^0_{p2}}$. Therefore, $\sum_{\boldsymbol{\nu}\in\mathcal{X}_1\cup\mathcal{X}_2} \varphi_{\boldsymbol{\nu}} * f$ converges in $\mathcal{S}'$.

We will next prove that $\sum_{\boldsymbol{\nu}\in\mathcal{Y}_1\cup\mathcal{Y}_2} \varphi_{\boldsymbol{\nu}} * f$ converges in $\mathcal{S}'$ by only using that $f \in \mathcal{S}'$. Since the Fourier transform is an isomorphism of $\mathcal{S}'$, it suffices to show that $\sum_{\boldsymbol{\nu}\in\mathcal{Y}_1\cup\mathcal{Y}_2} \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}$ converges in $\mathcal{S}'$. From $\widehat{f} \in \mathcal{S}'$ it follows that there exist $\ell, m \ge 1$ such that for any $\phi \in \mathcal{S}$,

$$|\langle \widehat{f}, \phi\rangle| \le c \max_{|\alpha|\le\ell, |\beta|\le m} \|\phi\|_{\alpha,\beta}, \quad \|\phi\|_{\alpha,\beta} := \sup_{\boldsymbol{\xi}\in\mathbb{R}^{n_1+n_2}} |\boldsymbol{\xi}^\alpha||\partial^\beta\phi(\boldsymbol{\xi})|.$$

Hence,

$$|\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| = |\langle \widehat{f}, \widehat{\varphi_{\boldsymbol{\nu}}}\phi\rangle| \le c \max_{|\alpha|\le\ell, |\beta|\le m} \|\widehat{\varphi_{\boldsymbol{\nu}}}\phi\|_{\alpha,\beta}.$$

Clearly, $\widehat{\varphi_{\boldsymbol{\nu}}}(\boldsymbol{\xi}) = \widehat{\varphi^1}(2^{-\nu_1}\xi_1)\widehat{\varphi^2}(2^{-\nu_2}\xi_2)$ and for any $\gamma = (\gamma_1, \gamma_2)$,

$$\partial^\gamma \widehat{\varphi_{\boldsymbol{\nu}}}(\boldsymbol{\xi}) = 2^{-\nu_1|\gamma_1| - \nu_2|\gamma_2|}\big(\partial^{\gamma_1}\widehat{\varphi^1}\big)(2^{-\nu_1}\xi_1)\big(\partial^{\gamma_2}\widehat{\varphi^2}\big)(2^{-\nu_2}\xi_2).$$

From (2.5), $\widehat{\varphi^i}$ is supported on $\{\xi_i : 2^{-1} \le |\xi_i| \le 2\}$, $i = 1, 2$, and hence $\widehat{\varphi_{\boldsymbol{\nu}}}$ is supported on the rectangle $R_{\boldsymbol{\nu}} := \{\boldsymbol{\xi} : 2^{\nu_1-1} \le |\xi_1| \le 2^{\nu_1+1},\ 2^{\nu_2-1} \le |\xi_2| \le 2^{\nu_2+1}\}$. From all of the above, it follows that

$$|\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| \le c2^{(-\nu_1 m)_+}2^{(-\nu_2 m)_+} \sup_{\boldsymbol{\xi}\in R_{\boldsymbol{\nu}}} (1 + |\boldsymbol{\xi}|)^\ell \max_{|\beta|\le m} |\partial^\beta\phi(\boldsymbol{\xi})|. \tag{8.12}$$

Here $y_+ := \max\{y, 0\}$. Choose $r_1 \ge (b_2^{-1} + 1)m$, $r_2 \ge (b_1^{-1} + 1)m$, $r_1, r_2 \in \mathbb{N}$. Clearly, (8.12) implies

$$|\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| \le c2^{-\nu_1 r_1 - \nu_2 m} \max_{|\alpha|\le\ell+r_1, |\beta|\le m} \|\phi\|_{\alpha,\beta}, \quad \nu_1 \ge 0, \nu_2 \le 0, \tag{8.13}$$

$$|\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| \le c2^{-\nu_1 m - \nu_2 r_2} \max_{|\alpha|\le\ell+r_2, |\beta|\le m} \|\phi\|_{\alpha,\beta}, \quad \nu_1 \le 0, \nu_2 \ge 0, \tag{8.14}$$

$$|\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| \le c2^{-\nu_1 r_1 - \nu_2 r_2} \max_{|\alpha|\le\ell+r_1+r_2, |\beta|\le m} \|\phi\|_{\alpha,\beta}, \quad \nu_1 \ge 0, \nu_2 \ge 0. \tag{8.15}$$

From (8.13) and (8.15) it follows that

$$\begin{aligned}\sum_{\boldsymbol{\nu}\in\mathcal{Y}_1} |\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| &\le c \max_{|\alpha|\le \ell+r_1+r_2, |\beta|\le m} \|\phi\|_{\alpha,\beta} \sum_{\nu_1>0} 2^{-\nu_1 r_1} \sum_{\nu_2\ge -b_2^{-1}\nu_1} 2^{-\nu_2 m}\\ &\le c \max_{|\alpha|\le \ell+r_1+r_2, |\beta|\le m} \|\phi\|_{\alpha,\beta} \sum_{\nu_1>0} 2^{-\nu_1(r_1-b_2^{-1}m)}\\ &\le c \max_{|\alpha|\le \ell+r_1+r_2, |\beta|\le m} \|\phi\|_{\alpha,\beta}.\end{aligned}$$

Similarly, $\sum_{\boldsymbol{\nu}\in\mathcal{Y}_2} |\langle \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}, \phi\rangle| \le c \max_{|\alpha|\le \ell+r_1+r_2, |\beta|\le m} \|\phi\|_{\alpha,\beta}$. These estimates yield that $\sum_{\boldsymbol{\nu}\in\mathcal{Y}_1\cup\mathcal{Y}_2} \widehat{\varphi_{\boldsymbol{\nu}}}\widehat{f}$ converges in $\mathcal{S}'$. Therefore, $\sum_{\boldsymbol{\nu}\in\mathbb{Z}^2} \varphi_{\boldsymbol{\nu}} * f$ convergence in $\mathcal{S}'$.

We now show that the limit f_0 in (6.16) is independent of the specific selection of φ. Assume that $\tilde{\varphi} = \tilde{\varphi}^1 \otimes \tilde{\varphi}^2$, where $\tilde{\varphi}^1, \tilde{\varphi}^2$ is another pair of functions satisfying (2.3), (2.5) with $n = n_1, n_2$, respectively. Let $\tilde{f}_0 := \sum_{\boldsymbol{\nu}\in\mathbb{Z}^2} \tilde{\varphi}_{\boldsymbol{\nu}} * f$. From above it follows that this series convergences in $\mathcal{S}'$. Then for any $\phi \in \mathcal{S}$, we have

$$\langle f_0 - \tilde{f}_0, \phi\rangle = \lim_{N\to\infty} \sum_{\nu_1\ge -N} \sum_{\nu_2\ge -N} \big\langle (\varphi_{\boldsymbol{\nu}} - \tilde{\varphi}_{\boldsymbol{\nu}}) * f, \phi\big\rangle,$$

where the series converges absolutely. Further, by (2.3) and (2.5) it follows that

$$\operatorname{supp}\Big(\sum_{\nu_1\ge -N} \sum_{\nu_2\ge -N} \big((\varphi_{\boldsymbol{\nu}} - \tilde{\varphi}_{\boldsymbol{\nu}}) * f\big)^{\wedge}\Big) \subset \{\boldsymbol{\xi} : |\boldsymbol{\xi}| \le 2^{-N}\}.$$

Let $\omega \in \mathcal{S}(\mathbb{R}^{n_1}\times\mathbb{R}^{n_2})$ be such that $\widehat{\omega}(\boldsymbol{\xi}) = 1$ if $|\boldsymbol{\xi}| \le 1$ and $\widehat{\omega}(\boldsymbol{\xi}) = 0$ if $|\boldsymbol{\xi}| \ge 2$. Then

$$\begin{aligned}\sum_{\nu_1\ge -N} \sum_{\nu_2\ge -N} (\varphi_{\boldsymbol{\nu}} - \tilde{\varphi}_{\boldsymbol{\nu}}) * f &= \omega_{-N} * \sum_{\nu_1\ge -N} \sum_{\nu_2\ge -N} (\varphi_{\boldsymbol{\nu}} - \tilde{\varphi}_{\boldsymbol{\nu}}) * f\\ &= \sum_{\nu_1=-N}^{-N+1} \sum_{\nu_2=-N}^{-N+1} \big(\omega_{-N} * \varphi_{\boldsymbol{\nu}} * f - \omega_{-N} * \tilde{\varphi}_{\boldsymbol{\nu}} * f\big),\end{aligned}$$

and hence

$$\begin{aligned}|\langle f_0 - \tilde{f}_0, \phi\rangle| &= \Big| \lim_{N\to\infty} \sum_{\nu_1=-N}^{-N+1} \sum_{\nu_2=-N}^{-N+1} \big\langle \omega_{-N} * (\varphi_{\boldsymbol{\nu}} * f - \tilde{\varphi}_{\boldsymbol{\nu}} * f), \phi\big\rangle\Big|\\ &\le \lim_{N\to\infty} \sum_{\nu_1=-N}^{-N+1} \sum_{\nu_2=-N}^{-N+1} \|\omega_{-N}\|_{L^1}\big(\|\varphi_{\boldsymbol{\nu}} * f\|_{L^\infty} + \|\tilde{\varphi}_{\boldsymbol{\nu}} * f\|_{L^\infty}\big)\|\phi\|_{L^1} = 0.\end{aligned}$$

Here for the last equality we used (8.11) and the fact that (8.11) holds with $\varphi_{\boldsymbol{\nu}}$ replaced by $\tilde{\varphi}_{\boldsymbol{\nu}}$. This completes the proof of part (a) of Proposition 6.5. The proofs of parts (b) and (c) are carried out in a similar manner. We omit them. □

8.4 Proof of Proposition 7.7

For simplicity we assume that $n_1 = n_2 = 1$.

Fix $N \in 2\mathbb{N}$ and denote by $\mathcal{Y}_j$, $j = 1, \dots, N$, the set of all dyadic rectangles $\mathbf{I} = I_1 \times I_2 \subset [0,1]^2$ such that $|I_1| = 2^{-j}$ and $|I_2| = 2^{-N+j}$; hence $|\mathbf{I}| = 2^{-N}$. Set $\mathcal{Y} := \cup_{j=1}^N \mathcal{Y}_j$. Clearly, $\#\mathcal{Y}_j = 2^N$ and hence $\#\mathcal{Y} = \#\big(\cup_{j=0}^N \mathcal{Y}_j\big) = N2^N$.

For any $\mathbf{I} \in \mathcal{Y}$, we write $\psi_{\mathbf{I}}(\mathbf{x}) := \psi_{I_1}(x_1)\psi_{I_2}(x_2)$. Consider the function

$$F(\mathbf{x}) := \sum_{\mathbf{I} \in \mathcal{Y}} \psi_{\mathbf{I}}(\mathbf{x}), \quad \cup\{\mathbf{I} \colon \mathbf{I} \in \mathcal{Y}\} \subset [0,1]^2. \tag{8.16}$$

By the wavelet characterization of $\dot{B}^{\mathbf{s}}_{\tau\tau}$ in (4.1) along with (6.6), it follows that

$$\begin{aligned} \|F\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} &\sim \Big(\sum_{\mathbf{I} \in \mathcal{Y}} \|\psi_{\mathbf{I}}\|_p^\tau\Big)^{1/\tau} \sim \Big(\sum_{\mathbf{I} \in \mathcal{Y}} |\mathbf{I}|^{(1/p-1/2)\tau}\Big)^{1/\tau} \\ &\sim 2^{-N(1/p-1/2)}(\#\mathcal{Y})^{1/\tau} \sim 2^{N(1/2-1/p)}(N2^N)^{1/\tau}. \end{aligned} \tag{8.17}$$

(a) We next prove estimate (7.11) (the sharpness of (7.7)). Assume $2 < p < \infty$. Let $m := \#\mathcal{Y}/2 = N2^N/2$. We claim that

$$\sigma_m(F) \ge c^\star m^{-\alpha} (\ln m)^{1/2-1/p} \|F\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}, \tag{8.18}$$

where $c^\star > 0$ is a constant independent of m. Indeed, clearly

$$\begin{aligned} \sigma_m(F)_p &= \inf_{g \in \Sigma_m} \|F - g\|_p \ge c \inf_{g \in \Sigma_m} \|F - g\|_{\dot{F}^0_{p2}} \\ &\ge c \Big\| \Big(\sum_{\mathbf{I} \in \mathcal{Y} \setminus \mathcal{X}_m} [|\mathbf{I}|^{-1/2} \mathbb{1}_{\mathbf{I}}(\cdot)]^2 \Big)^{1/2} \Big\|_{L^p} = c2^{N/2} \Big\| \Big(\sum_{\mathbf{I} \in \mathcal{Y} \setminus \mathcal{X}_m} \mathbb{1}_{\mathbf{I}}(\cdot) \Big)^{1/2} \Big\|_{L^p} \end{aligned} \tag{8.19}$$

for some set $\mathcal{X}_m \subset \mathcal{Y}$ with $\#\mathcal{X}_m = m$.

Set $h(\mathbf{x}) := \sum_{\mathbf{I} \in \mathcal{Y} \setminus \mathcal{X}_m} \mathbb{1}_{\mathbf{I}}(\mathbf{x})$ and let $\Omega := \big\{\mathbf{x} \colon h(\mathbf{x}) > N/4\big\}$. We claim that $|\Omega| \ge 1/5$. Indeed, assume $|\Omega| < 1/5$. Then

$$N/2 = \int_{[0,1]^2} h(\mathbf{x}) d\mathbf{x} = \int_\Omega h(\mathbf{x}) d\mathbf{x} + \int_{[0,1]^2 \setminus \Omega} h(\mathbf{x}) d\mathbf{x} \le N/5 + N/4 < N/2,$$

where we used that $\|h\|_\infty \le N$. We got a contradiction that proves that $|\Omega| \ge 1/5$. This coupled with (8.19) leads to

$$\sigma_m(F)_p \ge c2^{N/2} \|h^{1/2}\|_{L^p} \ge c2^{N/2} \|h^{1/2}\|_{L^p(\Omega)} \ge c' 2^{N/2} N^{1/2}.$$

A little algebra shows that this and (8.17) yield (8.18). Estimate (7.11) follows at once from (8.18).

To prove (7.10) (assuming $0 < p \le 2$) is easier. One simply takes $F := \sum_{\mathbf{I}\in\mathcal{Y}} \psi_{\mathbf{I}}$, where $\mathcal{Y}$ consists of $2m$ disjoint rectangles of the same area. Then it is easy to verify that $\sigma_m(F) \ge cm^{-\alpha}\|F\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}}$, which implies (7.10).

(b) We now prove estimate (7.12). Let $0 < p \le 2$. We will use again the function F from (8.16). This time we choose $m := N2^N$ and let $g := F$. Observe that $g \in \Sigma_m$. We have

$$\|g\|_p \sim \|g\|_{\dot{F}^0_{p2}} \sim \Big\|\Big(\sum_{\mathbf{I}\in\mathcal{Y}}[|\mathbf{I}^{-1/2}\mathbb{1}_{\mathbf{I}}(\cdot)]^2\Big)^{1/2}\Big\|_{L^p} \sim 2^{N/2}\Big\|\Big(\sum_{\mathbf{I}\in\mathcal{Y}}\mathbb{1}_{\mathbf{I}}(\cdot)\Big)^{1/2}\Big\|_{L^p}$$
$$\sim 2^{N/2}\Big(\Big\|\sum_{\mathbf{I}\in\mathcal{Y}}\mathbb{1}_{\mathbf{I}}(\cdot)\Big\|_{L^{p/2}}\Big)^{1/2} \sim 2^{N/2}N^{1/2}.$$

On the other hand, from (8.17) we know that $\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \sim 2^{N(1/2-1/p)}(N2^N)^{1/\tau}$. Now again a straightforward calculation shows that

$$\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \sim m^{\alpha}(\ln m)^{1/p-1/2}\|g\|_p,$$

which readily implies (7.12).

To prove (7.13) we take $g := \sum_{\mathbf{I}\in\mathcal{Y}} \psi_{\mathbf{I}}$, where $\mathcal{Y}$ consists of m disjoint rectangles of the same area. Then it is easy to see that $\|g\|_{\dot{B}^{\mathbf{s}}_{\tau\tau}} \ge cm^{\alpha}\|g\|_p$, which yields (7.13). The proof is complete. □

References

1. Bownik, M., Li, B., Yang, D., Zhou, Y.: Weighted anisotropic product Hardy spaces and boundedness of sublinear operators. Math. Nachr. **283**(3), 392–442 (2010)
2. Chang, S.-Y.A., Fefferman, R.: Some recent developments in Fourier analysis and H^p-theory on product domains. Bull. Am. Math. Soc. (N.S.) **12**, 1–43 (1985)
3. Cohen, A., Daubechies, I., Feauveau, J.-C.: Biorthogonal bases of compactly supported wavelets. Commun. Pure Appl. Math. **45**, 485–560 (1992)
4. Cohen, A., DeVore, R., Hochmuth, R.: Restricted nonlinear approximation. Constr. Approx. **16**, 85–113 (2000)
5. Daubechies, I.: Ten Lectures on Wavelets, CBMS–NSF Regional Conference Series in Applied Mathematics, vol. 61. SIAM, Philadelphia (1992)
6. DeVore, R.: Nonlinear approximation. Acta Numer. **7**, 51–150 (1998)
7. DeVore, R., Konyagin, S., Temlyakov, V.: Hyperbolic wavelet approximation. Constr. Approx. **14**, 1–26 (1998)
8. DeVore, R., Lorentz, G.G.: Constructive Approximation. Springer, Berlin (1993)
9. Fefferman, R.: Singular integrals on product H^p spaces. Rev. Mat. Iberoam. **1**(2), 25–31 (1985)
10. Fefferman, R.: Harmonic analysis on product spaces. Ann. Math. (2) **126**(1), 109–130 (1987)
11. Frazier, M., Jawerth, B.: Decomposition of Besov Spaces. Indiana Univ. Math. J. **34**, 777–799 (1985)
12. Frazier, M., Jawerth, B.: The φ-transform and applications to distribution spaces. In: Function Spaces and Applications (Lund, 1986), pp. 223–246, Lecture Notes in Mathematics, vol. 1302, Springer, Berlin (1988)
13. Frazier, M., Jawerth, B.: A discrete transform and decompositions of distribution. J. Funct. Anal. **93**, 34–170 (1990)
14. Frazier, M., Jawerth, B., Weiss, G.: Littlewood–Paley theory and the study of function spaces, CBMS 79 (1991). AMS

15. Grafakos, L.: Classical and Modern Fourier Analysis. Pearson Education, Inc., Upper Saddle River (2004)
16. Gundy, R.F., Stein, E.M.: H^p theory for the poly-disc. Proc. Nat. Acad. Sci. **76**, 1026–1029 (1979)
17. Kyriazis, G., Petrushev, P.: New bases for Triebel–Lizorkin and Besov spaces. Trans. Am. Math. Soc. **354**(2), 749–776 (2002)
18. Li, B., Bownik, M., Yang, D., Zhou, Y.: Anisotropic singular integrals in product spaces. Sci. China Math. **53**, 3163–3178 (2010)
19. Li, B., Bownik, M., Yang, D.: Littlewood–Paley characterization and duality of weighted anisotropic product Hardy spaces. J. Funct. Anal. **266**(5), 2611–2661 (2014)
20. Meyer, Y.: Wavelets and Operators. Cambridge University Press, Cambridge (1992)
21. Narcowich, F.J., Petrushev, P., Ward, J.D.: Decomposition of Besov and Triebel–Lizorkin spaces on the sphere. J. Funct. Anal. **238**, 530–564 (2006)
22. Peetre, J.: New Thoughts on Besov Spaces, Duke University Mathematics Series. Duke University, Durham (1976)
23. Schmeisser, H.-J.: Recent developments in the theory of function spaces with dominating mixed smoothness. In: Rakosnik, J. (ed.) Proceedings of the Conference NAFSA-8 in Prague 2006, pp. 145–204. Institute of Mathematical Academy of Sciences, Prague (2007)
24. Schmeisser, H.-J., Triebel, H.: Topics in Fourier Analysis and Function Spaces. Wiley, New York (1987)
25. Stein, E.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton (1993)
26. Triebel, H.: Theory of Function Spaces, Monographs in Mathematics, vol. 78. Birkhäuser Verlag, Basel (1983)
27. Triebel, H.: Theory of Function Spaces II. Monographs in Mathematics. Birkhäuser, Basel (1992)
28. Triebel, H.: The Structure of Functions. Birkhäuser, Basel (2001)
29. Vybiral, J.: Function spaces with dominating mixed smoothness. Diss. Math. **436**, 1–73 (2006)
30. Wojtaszczyk, P.: A Mathematical Introduction to Wavelets. Cambridge University Press, Cambridge (1997)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.