Nonlinear Approximation Schemes Associated With Nonseparable Wavelet Bi-Frames

Martin Ehler*

January 31, 2008

Abstract

In the present paper, we study nonlinear approximation properties of multivariate wavelet bi-frames. For a certain range of parameters, the approximation classes associated with best N-term approximation are determined to be Besov spaces and thresholding the wavelet bi-frame expansion realizes the approximation rate. Our findings extend results about dyadic wavelets to more general scalings. Finally, we verify that the required linear independence assumption is satisfied for particular families of nondyadic wavelet bi-frames in arbitrary dimensions.

AMS subject classification: 41A17, 42C40, 41A50

Key words: Wavelet bi-frame, Besov spaces, nonlinear approximation, Jackson inequality, Bernstein inequality, isotropic scaling

1 Introduction

Almost any kind of application requires at least to a certain extent the analysis of data. Depending on the specific application, the collection of data is usually called a measurement, a signal, or an image. In a mathematical framework, all of these objects are represented as functions. In order to analyze them, they are decomposed into simple building blocks. Such methods are not only used in mathematics, but also in physics, electrical engeneering, and medical imaging.

The building blocks also provide a series expansion, which reconstructs the original function. In computational algorithms, the series has to be replaced by a finite sum. Hence, we must approximate from N terms. There arise two fundamental problems, which have to be solved. First, let the approximation class essentially collect all functions, whose best choice of N terms yields a specific rate of approximation. It is important to express the approximation class in terms of classical function spaces since the class serves as a benchmark in order to evaluate different selections of N terms. Second, in practical algorithms, we require a realization of the best N-term approximation, i.e., we must look for a simple rule of the selection of N particular terms such that they provide the same rate of approximation as the best N-term approximation.

In wavelet theory, one approximates functions from dilates and shifts. For dyadic orthonormal wavelets, at least up to a certain rate, the approximation class equals a Besov space and thresholding the coefficients of the series expansion realizes the best N-term approximation, cf. [6, 12]. The results require certain smoothness and vanishing moments of the wavelets as well as a linear

*Department of Mathematics and Computer Science, Philipps-Universität Marburg, Hans-Meerwein Str., 35032 Marburg/Germany, Email: ehler@mathematik.uni-marburg.de

independence condition on the underlying refinable function. In [26], Lindemann generalized the dyadic results regarding biorthogonal wavelet bases with isotropic dilation. Borup, Gribonval, and Nielsen address wavelet bi-frames in [1]. However, their results are restricted to dyadic dilation. The present work is dedicated to an extension to more general scalings.

In order to characterize the approximation class, one has to establish so-called matching Jackson and Bernstein inequalities. They imply that the approximation space equals a so-called interpolation space. Fortunately, interpolation is well-studied, and, in many particular situations, these classes can be identified with classical smoothness spaces, which yields the final characterization of the approximation class.

In order to derive the Jackson inequality, we require the characterization of Besov spaces by wavelet bi-frames, i.e., the Besov semi-norm must be equivalently expressed in terms of a sequence norm of wavelet bi-frame coefficients. The characterization of Besov spaces by dyadic orthonormal wavelet bases was derived by DeVore, Jawerth, and Popov in the early nineties, cf. [13]. Lindemann extended the characterization to pairs of biorthogonal wavelet bases with general isotropic scalings, see [26]. Recently, Borup, Gribonval, and Nielsen characterized Besov spaces by dyadic wavelet bi-frames, cf. [1]. To point out the difficulties of the extension to wavelet bi-frames with general isotropic scalings, we shall explain the main idea of the dyadic bi-frame approach. Initially, one chooses a dyadic orthonormal basis characterizing the Besov space. Recall that the characterization requires a sufficient order of smoothness, and one can choose a tensor product of Meyer wavelets or of sufficiently smooth Daubechies wavelets, cf. [8]. Then one applies a certain localization technique, i.e., the bi-frame is localized to the dyadic orthonormal wavelet basis such that the orthonormal characterization carries over to the wavelet bi-frame. Hence, the orthonormal basis plays the role of a reference system. In order to address general isotropic scalings, there arise two problems. First, for many isotropic dilation matrices, it is not clear whether there exist smooth compactly supported orthogonal wavelets. Hence, we need another reference system. Since, for most of the known dilation matrices, there exist smooth compactly supported biorthogonal wavelets, see for instance [11, 22], they constitute promising substitutes for the orthogonal wavelet basis. Second, we have to extend the localization technique from dyadic to isotropic dilation as well as from orthogonal to biorthogonal reference systems.

The Jackson inequality results from the Besov space characterization. In order to address the Bernstein inequality, we restrict us to so-called idempotent scalings, see Subsection 2.2. As in the dyadic setting, the underlying refinable function must have linearly independent integer shifts on $(0,1)^d$. Finally, the approximation classes of wavelet bi-frames with idempotent scalings are interpolation spaces. Since, for certain parameters, the arising interpolation spaces coincide with Besov spaces, we solve the first fundamental problem mentioned above. Facing the second problem, we derive that the best N-term approximation rate can be realized by thresholding the wavelet bi-frame expansion. Contrary to [1], we can allow for arbitrary thresholding rules.

It should be mentioned that the limitation to idempotent scalings is not too restrive since most wavelet bi-frames in the literature are included. In the remainder of the work, we verify that the nondyadic families of optimal wavelet bi-frames in arbitary dimensions with arbitrarily high smoothness and vanishing moments in [15, 16] satisfy the assumptions of the Jackson and Bernstein inequalities.

The present paper is organized as follows: In Section 2, we recall the basic elements of best N-term approximation in Banach space as well as wavelet bi-frames with respect to general scalings. In Section 3, we verify that the Besov semi-norm is equivalent to a sequence norm of wavelet bi-frame coefficients. The Jackson and Bernstein inequalities are established in Section 4, and Section 5 is dedicated to the realization of the best N-term approximation rate by applying arbitrary thresholding rules. Finally, in Section 6, we verify that particular families of nondyadic wavelet bi-frames in [15, 16] satisfy the assumptions of the Jackson and Bernstein estimates.

2 General Setting

2.1 Best N-Term Approximation

Let X be a Banach space. A countable collection $\mathcal{D} \subset X$ is called a *dictionary* if its elements are normalized in the sense of $\|g\|_X \sim 1$, for all $g \in \mathcal{D}$. Then let $\Sigma_N(\mathcal{D})$ be the collection of all linear combinations of at most N elements of $\mathcal{D}$. For any given $f \in X$,

$$\sigma_N(f, \mathcal{D})_X := \operatorname{dist}(f, \Sigma_N(\mathcal{D}))_X$$

is called the *error of best N-term approximation*. In order to approximate elements in X from $\Sigma_N(\mathcal{D})$, it is important to determine those $f \in X$ providing the *approximation rate* α, i.e.,

$$\sigma_N(f, \mathcal{D})_X \lesssim N^{-\alpha}, \quad \text{for all } N \in \mathbb{N},$$

where the constant may depend on f. This question leads to the following definition. For $0 < s < \infty$, $0 < q \le \infty$, the *approximation class* $\mathcal{A}^s_q(X, \mathcal{D})$ is the collection of all $f \in X$ such that

$$|f|_{\mathcal{A}^s_q(X,\mathcal{D})} := \begin{cases} \left(\sum_{N=1}^{\infty} (N^s \sigma_N(f, \mathcal{D})_X)^q \frac{1}{N}\right)^{\frac{1}{q}}, & \text{for } 0 < q < \infty, \\ \sup_{N \ge 1} (N^s \sigma_N(f, \mathcal{D})_X), & \text{for } q = \infty, \end{cases}$$

is finite. If we choose $q = \infty$, then the space $\mathcal{A}^s_\infty(X, \mathcal{D})$ precisely consists of all f in X having approximation rate s. For $0 < q < \infty$, membership in $\mathcal{A}^s_q(X, \mathcal{D})$ means a slightly stronger condition, see Chapter 7 in [14].

In order to determine the approximation class, the real method of interpolation is a valuable tool. The following so-called Jackson and Bernstein estimates provide the connection between approximation and interpolation, cf. [12] and Chapter 7 in [14]. Given $0 < s < \infty$ and a Banach space X, let Y be continuously embedded in X. If the *Jackson inequality*

$$\sigma_N(f, \mathcal{D})_X \lesssim N^{-s} |f|_Y, \quad \text{for all } f \in Y,\ N \in \mathbb{N},$$

holds, then the real interpolation space $[X, Y]_{\frac{\alpha}{s}, q}$ is contained in $\mathcal{A}^\alpha_q(X, \mathcal{D})$, for all $0 < \alpha < s$ and $0 < q \le \infty$. If the *Bernstein inequality*

$$|f|_Y \lesssim N^s \|f\|_X, \quad \text{for all } f \in \Sigma_N(\mathcal{D}),\ N \in \mathbb{N},$$

holds, then $\mathcal{A}^\alpha_q(X, \mathcal{D}) \subset [X, Y]_{\frac{\alpha}{s}, q}$, for $0 < \alpha < s$ and $0 < q \le \infty$.

In order to determine the approximation class, one has to establishing matching Jackson and Bernstein estimates. Then the approximation class equals an interpolation space, and in a next step, one has to describe the interpolation class by classical function spaces.

2.2 Wavelet Bi-Frames

Given a countable index set $\mathcal{K}$, a collection $\{f_\kappa : \kappa \in \mathcal{K}\}$ in a Hilbert space $\mathcal{H}$ is called a *frame* for $\mathcal{H}$ if there exist two positive constants A, B such that

$$A\|f\|^2_{\mathcal{H}} \le \|(\langle f, f_\kappa \rangle)_{\kappa \in \mathcal{K}}\|^2_{\ell_2} \le B\|f\|^2_{\mathcal{H}}, \quad \text{for all } f \in \mathcal{H}.$$

The collection $\{f_\kappa : \kappa \in \mathcal{K}\}$ is a frame for $\mathcal{H}$ iff its *synthesis operator*

$$F : \ell_2(\mathcal{K}) \to \mathcal{H}, \quad (c_\kappa)_{\kappa \in \mathcal{K}} \mapsto \sum_{\kappa \in \mathcal{K}} c_\kappa f_\kappa,$$

is well-defined and onto, see Section 5.5 in [5]. Hence, each $f \in \mathcal{H}$ has a series expansion in the frame. In order to derive the coefficients of such an expansion, one considers the *frame operator* $S := FF^*$. It is positive and boundedly invertible, and the system $\{S^{-1}f_\kappa : \kappa \in \mathcal{K}\}$ is called the *canonical dual frame*. It is a frame, and it provides the expansion

$$f = \sum_{\kappa\in\mathcal{K}} \langle f, S^{-1}f_\kappa \rangle f_\kappa, \quad \text{for all } f \in \mathcal{H}, \tag{1}$$

cf. Chapter 5 in [5].

In order to address wavelet frames, we shall clarify our concept of dilation. Throughout this paper, let M denote a *dilation matrix*, i.e., an integer matrix, whose eigenvalues are greater than one in modulus. For $\psi : \mathbb{R}^d \to \mathbb{C}$, let

$$\psi_{j,k}(x) := m^{\frac{j}{2}}\psi(M^j x - k), \quad \text{for } j \in \mathbb{Z}, k \in \mathbb{Z}^d,$$

where $m := |\det(M)|$ throughout. Given a finite number of $L_2(\mathbb{R}^d)$-functions $\psi^{(1)}, \dots, \psi^{(n)}$, the collection

$$X(\{\psi^{(1)}, \dots, \psi^{(n)}\}) := \{\psi^{(\mu)}_{j,k} : j \in \mathbb{Z}, k \in \mathbb{Z}^d, \mu = 1, \dots, n\} \tag{2}$$

is called a *wavelet frame* if it constitutes a frame for $L_2(\mathbb{R}^d)$. Unfortunately, its canonical dual frame may not have the wavelet structure as well. Nevertheless, the canonical dual in (1) can possibly be replaced by an alternative dual wavelet frame. This motivates the following definition. Two frames $\{f_\kappa : \kappa \in \mathcal{K}\}$ and $\{\widetilde{f}_\kappa : \kappa \in \mathcal{K}\}$ for $\mathcal{H}$ are called a *bi-frame* if the expansion

$$f = \sum_{\kappa\in\mathcal{K}} \langle f, \widetilde{f}_\kappa \rangle f_\kappa$$

holds for every $f \in \mathcal{H}$. We speak of a *wavelet bi-frame* if two systems as in (2) constitute a bi-frame for $L_2(\mathbb{R}^d)$.

Compactly supported wavelets are generally derived from two compactly supported *refinable functions* φ and $\widetilde{\varphi}$, i.e., there exist finitely supported sequence $(a_k)_{k\in\mathbb{Z}^d}$ and $(b_k)_{k\in\mathbb{Z}^d}$ such that

$$\varphi(x) = \sum_{k\in\mathbb{Z}} a_k \varphi(Mx - k), \quad \widetilde{\varphi}(x) = \sum_{k\in\mathbb{Z}} b_k \widetilde{\varphi}(Mx - k). \tag{3}$$

Next, one must find finitely supported sequences $(a^{(\mu)}_k)_{k\in\mathbb{Z}^d}$ and $(b^{(\mu)}_k)_{k\in\mathbb{Z}^d}$, $\mu = 1, \dots, n$, such that

$$\psi^{(\mu)}(x) = \sum_{k\in\mathbb{Z}^d} a^{(\mu)}_k \varphi(Mx - k), \quad \widetilde{\psi}^{(\mu)}(x) = \sum_{k\in\mathbb{Z}^d} b^{(\mu)}_k \widetilde{\varphi}(Mx - k), \tag{4}$$

generate a wavelet bi-frame, see [8, 9] for details.

A wavelet frame induced by a refinable function requires at least $m - 1$ wavelets. Hence, we need $2^d - 1$ wavelets for dyadic scalings $M = 2\mathcal{I}_d$. In order to reduce the number of wavelets, one considers nondyadic scalings. For instance, for $d = 2, 3$, let

$$M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad M = \begin{pmatrix} 0 & 2 & 1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \tag{5}$$

and, for $d > 3$, let

$$M = \begin{pmatrix} 0 & 2 & 1 & \cdots\cdots & 1 \\ \vdots & \ddots & 1 & 0 \ \cdots & 0 \\ \vdots & & \ddots & \ddots \quad \ddots & \vdots \\ 0 & \cdots\cdots & 0 & 1 & 0 \\ -1 & \cdots\cdots\cdots & & -1 & 0 \\ 1 & \cdots\cdots\cdots\cdots & & & 1 \end{pmatrix}. \tag{6}$$

These are dilation matrices with $m = 2$ in arbitrary dimensions. Next, we introduce two subclasses of scalings. A dilation matrix is called *isotropic* if it can be diagonalized and all eigenvalues have the same modulus. This class is mainly addressed in Section 3. A dilation matrix is called *idempotent* if there are $l, h \in \mathbb{N}$ such that $M^l = h\mathcal{I}_d$. Idempotent dilation matrices are of main interest in Subsection 4.2. One easily verifies that their minimum polynomial has pairwise distinct zeros. Hence, they can be diagonalized, see a standard textbook on linear algebra. Therefore, each idempotent dilation matrix is isotropic. Note that M in (5) and (6) is idempotent with $M^d = 2\mathcal{I}_d$.

3 Characterization of Besov Spaces by Wavelet Bi-Frames

Biorthogonal wavelet bases characterize Besov spaces, see [13] for dyadic scaling and [3, 17, 26] for the extension to isotropic dilation matrices. In this section, we establish the equivalence between the Besov semi-norm and a sequence norm of wavelet bi-frame coefficients. We extend the dyadic results in [1] to the more general class of isotropic scalings. In order to derive the equivalence, we apply the concept of localization: in a series of papers, Gröchenig considers localized frames, i.e., frames, whose Gramian matrices have certain decay outside the diagonal, see [20, 21]. In some sense, we follow these ideas. We also address Gramian type matrices, and we estimate the decay of their entries outside the diagonal. However, we apply localization to two different frames, i.e., we consider their mixed Gramian matrices. Finally, we establish that the mixed Gramian matrix of bi-frame wavelets and biorthogonal wavelets constitutes a bounded operator on certain sequence spaces. Then by applying some results about wavelet bi-frame expansions in $L_p(\mathbb{R}^d)$, the biorthogonal characterization carries over to the bi-frame.

3.1 A Characterization by Biorthogonal Wavelet Bases

Since we will study approximation in $L_p(\mathbb{R}^d)$, $1 < p < \infty$, let

$$\psi_{j,k}^{(\mu),p}(x) := m^{\frac{j}{p}} \psi^{(\mu)}(M^j x - k), \quad \text{for } j \in \mathbb{Z},\ k \in \mathbb{Z}^d,\ x \in \mathbb{R}^d,$$

denote the $L_p(\mathbb{R}^d)$-normalization of $\psi_{j,k}^{(\mu)}$, and let us use the short-hand notation $\psi_\lambda^p = \psi_{\mu,j,k}^p := \psi_{j,k}^{(\mu),p}$,where $\lambda = (\mu, j, k)$ and $\Lambda := \{1, \ldots, m-1\} \times \mathbb{Z} \times \mathbb{Z}^d$. We do so for the dual wavelets as well. Moreover, we say a function ψ has *s vanishing moments* if

$$\int_{\mathbb{R}^d} x^\alpha \psi(x) dx = 0, \quad \text{for all } |\alpha| < s.$$

For fixed $1 < p < \infty$, we write B^s for the Besov space $B_\tau^s(L_\tau(\mathbb{R}^d))$, where $\frac{1}{\tau} = \frac{s}{d} + \frac{1}{p}$. These spaces arise in the context of nonlinear approximation as described in Section 4, see also [12]. Let us recall the characterization of Besov spaces by biorthogonal wavelets. Given an isotropic dilation matrix M, let $1 < p < \infty$, $\frac{1}{p} + \frac{1}{p'} = 1$, and suppose that $X(\{\psi^{(1)}, \ldots, \psi^{(m-1)}\})$, $X(\{\widetilde{\psi}^{(1)}, \ldots, \widetilde{\psi}^{(m-1)}\})$ are a pair of compactly supported biorthogonal wavelet bases, whose underlying refinable functions φ and $\widetilde{\varphi}$ are contained in $L_p(\mathbb{R}^d)$ and $L_{p'}(\mathbb{R}^d)$, respectively. Let φ be also contained in $W^s(L_\infty(\mathbb{R}^d))$, $s \in \mathbb{N}$. Then, for all $0 < \alpha < s$ and $f \in B^\alpha$, the series expansion

$$f = \sum_{\lambda \in \Lambda} \langle f, \widetilde{\psi}_\lambda^{p'} \rangle \psi_\lambda^p \tag{7}$$

holds in $L_p(\mathbb{R}^d)$, and

$$|f|_{B^\alpha} \sim \big\| \big(\langle f, \widetilde{\psi}_\lambda^{p'} \rangle \big)_{\lambda \in \Lambda} \big\|_{\ell_\tau}, \quad \text{for } \tfrac{1}{\tau} = \tfrac{\alpha}{d} + \tfrac{1}{p}, \tag{8}$$

see [3, 17, 26] for details.

3.2 Localization Technique

Given two frames $\{f_\kappa : \kappa \in \mathcal{K}\}$, $\{g_{\kappa'} : \kappa' \in \mathcal{K}'\}$ for $\mathcal{H}$, their synthesis operators F and G, respectively, are bounded. Thus, G^*F is a bounded operator on $\ell_2(\mathcal{K})$. It coincides with the mixed Gramian matrix operator

$$(c_\kappa)_{\kappa\in\mathcal{K}} \mapsto \Big(\sum_{\kappa\in\mathcal{K}} \langle f_\kappa, g_{\kappa'}\rangle\, c_\kappa\Big)_{\kappa'\in\mathcal{K}'}.$$

The following theorem shows that, for wavelet systems, the mixed Gramian is bounded on a large scale of ℓ_τ-spaces. It is our main result of this section, and it extends the dyadic results in [1] to general isotropic scalings. Note that we do not assume strong differentiability as they do in [1]. We only require weak differentiability.

Theorem 3.1. *Let M be isotropic and $s, s' \in \mathbb{N}$. For $\mu \in E := \{1,\dots,n\}$ and $\mu' \in E' := \{1,\dots,n'\}$, let compactly supported functions $f^{(\mu)} \in W^s(L_\infty(\mathbb{R}^d))$ and $g^{(\mu')} \in W^{s'}(L_\infty(\mathbb{R}^d))$ have s' and s vanishing moments, respectively. Given $1 \le p < \infty$ and $1 = \frac{1}{p} + \frac{1}{p'}$, we consider the matrix operator*

$$T : (c_\lambda)_{\lambda\in\Lambda} \mapsto \Big(\sum_{\lambda\in\Lambda} \langle f^p_\lambda, g^{p'}_{\lambda'}\rangle c_\lambda\Big)_{\lambda'\in\Lambda'},$$

where $\Lambda = E \times \mathbb{Z} \times \mathbb{Z}^d$ and $\Lambda' = E' \times \mathbb{Z} \times \mathbb{Z}^d$. Then $T : \ell_\tau(\Lambda) \to \ell_\tau(\Lambda')$ is bounded for any τ in the range

$$p\big(\tfrac{s'}{d} + 1\big) > \tau > \begin{cases} \left(\frac{s}{d} + \frac{1}{p}\right)^{-1}, & \text{for } \frac{s}{d} + \frac{1}{p} \ge 1,\\ p\left(1 - \frac{s}{d}\right), & \text{for } \frac{s}{d} + \frac{1}{p} \le 1.\end{cases} \tag{9}$$

Later, we only consider $p \ge \tau$. Therefore, the exact upper bound of τ in Theorem 3.1 is of minor interest. The lower bound is critical, and it will yield a restriction. Unfortunately, it cannot be improved in general, see [1] for a counterexample.

The proof of Theorem 3.1 keeps us busy for the remainder of the present subsection. One of the two fundamental ingredients is the following lemma. It extends the dyadic Lemma 8.10 in [27] also allowing for isotropic dilation matrices.

Lemma 3.2. *Let M be isotropic, and let $d < \delta$. For $j \in \mathbb{Z}$, consider the matrix operator T_j given by*

$$(d_k)_{k\in\mathbb{Z}^d} \mapsto \begin{cases} \Big(\sum_{k\in\mathbb{Z}^d} \left(1 + \|k - M^{-j}k'\|\right)^{-\delta} d_k\Big)_{k'\in\mathbb{Z}^d}, & \text{for } j > 0,\\ \Big(\sum_{k\in\mathbb{Z}^d} \left(1 + \|M^j k - k'\|\right)^{-\delta} d_k\Big)_{k'\in\mathbb{Z}^d}, & \text{for } j \le 0.\end{cases}$$

Then, T_j is bounded on $\ell_\tau(\mathbb{Z}^d)$, for any $1 \le \tau \le \infty$, and its operator norm satisfies

$$\|T_j\|_{\ell_\tau\to\ell_\tau} \lesssim \begin{cases} m^{\frac{j}{\tau}}, & j > 0,\\ m^{-\frac{j}{\tau'}}, & j \le 0, \text{ where } \frac{1}{\tau} + \frac{1}{\tau'} = 1.\end{cases}$$

Proof. First, we address $j \le 0$, and we consider $\tau = 1$ and $\tau = \infty$. For $1 < \tau < \infty$, we apply the Riesz-Thorin Interpolation Theorem. Let us choose $\tau = 1$. In order to derive

$$\|T_j\|_{\ell_1\to\ell_1} \lesssim 1, \tag{10}$$

we split $M^j k$ into the sum $l + r$ with $\|r\|_\infty < 1$, where $\|r\|_\infty$ denotes the maximum norm on $\mathbb{R}^d$. This yields

$$\begin{aligned}\sum_{k'\in\mathbb{Z}^d} (1 + \|M^j k - k'\|)^{-\delta} &\lesssim \sum_{k'\in\mathbb{Z}^d} (1 + \|l + r - k'\|_\infty)^{-\delta}\\ &= \sum_{k'\in\mathbb{Z}^d} (1 + \|r - k'\|_\infty)^{-\delta}.\end{aligned}$$

Applying the reverse triangle inequality $|\|r\|_\infty - \|k'\|_\infty| \leq \|r - k'\|_\infty$ provides

$$\sum_{k' \in \mathbb{Z}^d} (1 + \|M^j k - k'\|)^{-\delta} \lesssim \sum_{k' \in \mathbb{Z}^d} \big(1 + \big|\|r\|_\infty - \|k'\|_\infty\big|\big)^{-\delta}.$$

Since $\|r\|_\infty < 1$ and $d < \delta$, we obtain

$$\sum_{k' \in \mathbb{Z}^d} (1 + \|M^j k - k'\|)^{-\delta} \lesssim \sum_{k' \in \mathbb{Z}^d \setminus \{0\}} (\|k'\|_\infty)^{-\delta} + 1 \lesssim 1.$$

For $(d_k)_{k \in \mathbb{Z}^d} \in \ell_1(\mathbb{Z}^d)$, this yields

$$\begin{aligned}
\left\|T_j\left((d_k)_{k \in \mathbb{Z}^d}\right)\right\|_{\ell_1} &= \sum_{k' \in \mathbb{Z}^d} \Big| \sum_{k \in \mathbb{Z}^d} (1 + \|M^j k - k'\|)^{-\delta} d_k \Big| \\
&\leq \sum_{k \in \mathbb{Z}^d} \sum_{k' \in \mathbb{Z}^d} (1 + \|M^j k - k'\|)^{-\delta} |d_k| \\
&\lesssim \left\|(d_k)_{k \in \mathbb{Z}^d}\right\|_{\ell_1}.
\end{aligned}$$

Thus, (10) holds.

Now, let us address $\tau = \infty$. In the following, we verify

$$\|T_j\|_{\ell_\infty \to \ell_\infty} \lesssim m^{-j}. \tag{11}$$

This requires the introduction of a special norm: for isotropic dilation matrices M, there exists a norm $\|\cdot\|_M$ on $\mathbb{R}^d$ such that

$$\|Mx\|_M = \rho \|x\|_M, \quad \text{for all } x \in \mathbb{R}^d, \tag{12}$$

where ρ is the modulus of the eigenvalues of M, cf. [24]. Since all norms on $\mathbb{R}^d$ are equivalent, this leads to

$$\begin{aligned}
\sum_{k \in \mathbb{Z}^d} m^j (1 + \|M^j k - k'\|)^{-\delta} &\lesssim \sum_{k \in \mathbb{Z}^d} m^j (1 + \|M^j k - k'\|_M)^{-\delta} \\
&= \sum_{k \in \mathbb{Z}^d} m^j (1 + \|M^j (k - M^{-j} k')\|_M)^{-\delta}.
\end{aligned}$$

Due to $j \leq 0$, we have $M^{-j} k' \in \mathbb{Z}^d$. This provides with $m^j = \rho^{jd}$

$$\begin{aligned}
\sum_{k \in \mathbb{Z}^d} m^j (1 + \|M^j k - k'\|)^{-\delta} &= \sum_{k \in \mathbb{Z}^d} m^j (1 + \|M^j k\|_M)^{-\delta} \\
&\lesssim \sum_{k \in \mathbb{Z}^d} \rho^{jd} (1 + \|\rho^j k\|)^{-\delta}.
\end{aligned}$$

Since the last term is a Riemann sum of the integrable function $x \mapsto (1 + \|x\|)^{-\delta}$, we obtain

$$\sum_{k \in \mathbb{Z}^d} m^j (1 + \|M^j k - k'\|)^{-\delta} \lesssim 1.$$

For $(d_k)_{k\in\mathbb{Z}^d} \in \ell_\infty(\mathbb{Z}^d)$, the Cauchy-Schwartz inequality and the last estimate imply

$$\begin{aligned}\left\|T_j\left((d_k)_{k\in\mathbb{Z}^d}\right)\right\|_{\ell_\infty\to\ell_\infty} &= \sup_{k'\in\mathbb{Z}^d}\Big|\sum_{k\in\mathbb{Z}^d}(1+\|M^jk-k'\|)^{-\delta}d_k\Big| \\ &\le m^{-j}\sup_{k'\in\mathbb{Z}^d}\sum_{k\in\mathbb{Z}^d} m^j(1+\|M^jk-k'\|)^{-\delta}|d_k| \\ &= m^{-j}\sup_{k'\in\mathbb{Z}^d}\left\|\left(m^j(1+\|M^jk-k'\|)^{-\delta}d_k\right)_{k\in\mathbb{Z}^d}\right\|_{\ell_1} \\ &\le m^{-j}\sup_{k'\in\mathbb{Z}^d}\left\|\left(m^j(1+\|M^jk-k'\|)^{-\delta}\right)_{k\in\mathbb{Z}^d}\right\|_{\ell_1}\left\|(d_k)_{k\in\mathbb{Z}^d}\right\|_{\ell_\infty} \\ &\lesssim m^{-j}\left\|(d_k)_{k\in\mathbb{Z}^d}\right\|_{\ell_\infty}.\end{aligned}$$

Thus, (11) holds.

By applying the Riesz-Thorin Interpolation Theorem to (10) and (11), we obtain, for all $1 \le \tau \le \infty$,

$$\|T_j\|_{\ell_\tau\to\ell_\tau} \lesssim m^{-\frac{j}{\tau'}}, \text{ where } \tfrac{1}{\tau}+\tfrac{1}{\tau'}=1.$$

In order to address $j > 0$, we observe that, for $1 \le \tau < \infty$, the operator $T_{-j} : \ell_{\tau'} \to \ell_{\tau'}$ is the dual matrix operator of $T_j : \ell_\tau \to \ell_\tau$. Thus,

$$\|T_j\|_{\ell_\tau\to\ell_\tau} = \|T_{-j}\|_{\ell_{\tau'}\to\ell_{\tau'}} \lesssim m^{\frac{j}{\tau}}.$$

Since $T_j : \ell_\infty \to \ell_\infty$ is the dual of $T_{-j} : \ell_1 \to \ell_1$, this inequality still holds for $\tau = \infty$, which concludes the proof. □

By following the lines of the proof in [1], Lemma 3.2 implies the next Proposition.

Proposition 3.3. *Let M be isotropic, and let $1 \le p < \infty$, $\delta > d$, and $s, s' \in \mathbb{N}$. Then the matrix operator*

$$(c_{j,k})_{j,k} \mapsto \left(\sum_{\substack{k\in\mathbb{Z}^d,\\ j\le j'}} \frac{m^{(j-j')\left(\frac{s}{d}+\frac{1}{p}\right)}c_{j,k}}{(1+\|k-M^{j-j'}k'\|)^\delta} + \sum_{\substack{k\in\mathbb{Z}^d,\\ j>j'}} \frac{m^{(j'-j)\left(\frac{s'}{d}+\frac{1}{p'}\right)}c_{j,k}}{(1+\|k'-M^{j'-j}k\|)^\delta}\right)_{j',k'}$$

is bounded on $\ell_\tau(\mathbb{Z}\times\mathbb{Z}^d)$, for

$$p\left(\tfrac{s'}{d}+1\right) > \tau > \begin{cases} \frac{d}{\delta}, & \text{for } \frac{s}{d}+\frac{1}{p} \ge \frac{\delta}{d}, \\ \left(\frac{s}{d}+\frac{1}{p}\right)^{-1}, & \text{for } 1 < \frac{s}{d}+\frac{1}{p} \le \frac{\delta}{d}, \\ p(1-\frac{s}{d}), & \text{for } \frac{s}{d}+\frac{1}{p} \le 1. \end{cases}$$

The second fundamental ingredient for the proof of Theorem 3.1 is the following version of the Bramble-Hilbert Lemma, see [10]. Let Π_{s-1} denote the space of all polynomials of degree up to $s-1$:

Theorem 3.4. *Given $\Omega \subset \mathbb{R}^d$ convex, $s \in \mathbb{N}$, and $1 \le p \le \infty$, let $f \in W^s(L_p(\Omega))$. Then there exists a polynomial $q \in \Pi_{s-1}$ such that*

$$|f-q|_{W^l(L_p(\Omega))} \lesssim \operatorname{diam}(\Omega)^{s-l}|f|_{W^s(L_p(\Omega))}, \quad l = 0,\dots,s,$$

where

$$|f|_{W^s(L_p)} := \sum_{|\beta|=s}\|\partial^\beta f\|_{L_p}$$

denotes the Sobolev semi-norm of order s.

The following proposition results by combining Proposition 3.3 with Theorem 3.4.

Proposition 3.5. *Let M be isotropic, $s, s' \in \mathbb{N}$, and suppose that compactly supported functions $f \in W^s(L_\infty(\mathbb{R}^d))$ and $g \in W^{s'}(L_\infty(\mathbb{R}^d))$ have s' and s vanishing moments, respectively. Given $1 \le p < \infty$ and $1 = \frac{1}{p} + \frac{1}{p'}$, we consider the matrix operator*

$$T : (c_{j,k})_{j,k} \mapsto \Big(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} \left\langle f^p_{j,k}, g^{p'}_{j',k'} \right\rangle c_{j,k} \Big)_{j',k'}.$$

Then T is bounded on $\ell_\tau(\mathbb{Z}\times\mathbb{Z}^d)$ for any τ as in (9).

Proof. Fix $\delta > d$ sufficiently large. First, we address $j' \ge j$. Let $R > 0$ such that

$$\operatorname{supp}(g) \subset G := \{x \in \mathbb{R}^d : \|x\|_M \le R\},$$

where $\|\cdot\|_M$ denotes the norm in (12). Then G is convex and $M^{j-j'}G \subset G$. According to the vanishing moments and the Hölder inequality, we obtain

$$\begin{aligned}
\left|\left\langle f^p_{j,k}, g^{p'}_{j',k'} \right\rangle\right| &= m^{\frac{j}{p}} m^{\frac{j'}{p'}} \int_{\mathbb{R}^d} f(M^{j-j'}x + M^{j-j'}k' - k)\overline{g(x)} m^{-j'} dx \\
&= m^{(j-j')\frac{1}{p}} \inf_{q\in\Pi_{s-1}} \int_G \Big(f(M^{j-j'}x + M^{j-j'}k' - k) - q(x) \Big) \overline{g(x)} dx \\
&\le m^{(j-j')\frac{1}{p}} \inf_{q\in\Pi_{s-1}} \left\| f(M^{j-j'}\cdot + M^{j-j'}k' - k) - q(\cdot) \right\|_{L_\infty(G)} \|g\|_{L_1(G)}.
\end{aligned}$$

The space Π_{s-1} is affine invariant, i.e., $q \in \Pi_{s-1}$ yields $q(A\cdot + t) \in \Pi_{s-1}$, for all $A \in \mathbb{R}^{d\times d}$ and $t \in \mathbb{R}^d$. Thus, Theorem 3.4 with $l = 0$ implies

$$\begin{aligned}
\left|\left\langle f^p_{j,k}, g^{p'}_{j',k'} \right\rangle\right| &\lesssim m^{(j-j')\frac{1}{p}} \inf_{q\in\Pi_{s-1}} \|f - q\|_{L_\infty(M^{j-j'}G + M^{j-j'}k' - k)} \\
&\lesssim m^{(j-j')\frac{1}{p}} \operatorname{diam}(M^{j-j'}G)^s |f|_{W^s(L_\infty(M^{j-j'}G + M^{j-j'}k' - k))} \\
&\lesssim m^{(j-j')\frac{1}{p}} m^{(j-j')\frac{s}{d}} |f|_{W^s(L_\infty(G + M^{j-j'}k' - k))}.
\end{aligned}$$

Since f is compactly supported, there exists $r > 0$ such that, for all $v \in \mathbb{R}^d$ with $\|v\| \ge r$, the intersection $(G + v) \cap \operatorname{supp}(f)$ is empty. Hence, the Sobolev semi-norm can be estimated by

$$|f|_{W^s(L_\infty(G + M^{j-j'}k' - k))} \le \begin{cases} |f|_{W^s(L_\infty(\mathbb{R}^d))}, & \text{for } \|M^{j-j'}k' - k\| < r, \\ 0, & \text{for } \|M^{j-j'}k' - k\| \ge r. \end{cases}$$

This provides the final inequalities

$$\begin{aligned}
\left|\left\langle f^p_{j,k}, g^{p'}_{j',k'} \right\rangle\right| &\lesssim m^{(j-j')(\frac{s}{d}+\frac{1}{p})} |f|_{W^s(L_\infty(\mathbb{R}^d))} \left(\frac{1+r}{1 + \|M^{j-j'}k' - k\|} \right)^\delta \\
&\lesssim \frac{m^{(j-j')(\frac{s}{d}+\frac{1}{p})}}{(1 + \|M^{j-j'}k' - k\|)^\delta}.
\end{aligned}$$

Next, we address $j > j'$. Following the lines above with interchanged roles of f and g, we obtain

$$\left|\left\langle f^p_{j,k}, g^{p'}_{j',k'} \right\rangle\right| \lesssim \frac{m^{(j'-j)(\frac{s'}{d}+\frac{1}{p'})}}{(1 + \|M^{j'-j}k - k'\|)^\delta}.$$

By applying Proposition 3.3, the operator T is bounded on ℓ_τ. □

Proposition 3.5 addresses single f and g. In order to consider a finite number of functions as in Theorem 3.1, one applies norm coherences between $\ell_\tau(E \times \mathbb{Z} \times \mathbb{Z}^d)$ and $\ell_\tau(\mathbb{Z} \times \mathbb{Z}^d)$. We omit the detailed elaboration.

3.3 Hilbertian Dictionaries

Given a sufficiently smooth pair of compactly supported biorthogonal wavelet bases, then for $f \in B^\alpha$, the series expansion $\sum_{\lambda\in\Lambda} \langle f, \widetilde{\psi}_\lambda^{p'} \rangle \psi_\lambda^p$ converges towards f in $L_p(\mathbb{R}^d)$. This subsection provides some fundamentals, in order to generalize this statement regarding wavelet bi-frames. We extend the dyadic results in [1] to isotropic scalings. Given a wavelet system $\{\psi_\lambda : \lambda \in \Lambda\}$, we derive a classical decay condition on the sequence $(c_\lambda)_{\lambda\in\Lambda}$ such that

$$\sum_{\lambda\in\Lambda} c_\lambda \psi_\lambda^p \tag{13}$$

converges in $L_p(\mathbb{R}^d)$. In order to obtain a sufficient variety of decay conditions, we recall the following family of sequence spaces. For $0 < p < \infty$, $0 < q \le \infty$ and a countable index set $\mathcal{K}$, the *Lorentz space* $\ell_{p,q}(\mathcal{K})$ is the collection of bounded sequences $(c_\kappa)_{\kappa\in\mathcal{K}}$ satisfying $\|(c_\kappa)_{\kappa\in\mathcal{K}}\|_{\ell_{p,q}} < \infty$, where

$$\|(c_\kappa)_{\kappa\in\mathcal{K}}\|_{\ell_{p,q}} := \begin{cases} \left(\sum_{j=1}^\infty (j^{\frac{1}{p}} c_j^*)^q \frac{1}{j}\right)^{\frac{1}{q}}, & \text{for } 0 < q < \infty, \\ \sup_{j\ge 1} (j^{\frac{1}{p}} c_j^*), & \text{for } q = \infty, \end{cases}$$

while $(c_j^*)_{j\in\mathbb{N}}$ denotes a decreasing rearrangement of $(|c_\kappa|)_{\kappa\in\mathcal{K}}$.

Naturally, convergence problems as in (13) also arise in more abstract settings. In order to point out the key ingredients of its solution, we study the problem in a general framework. Following [1, 19], a dictionary $\mathcal{D} = \{g_\kappa : \kappa \in \mathcal{K}\}$ in a Banach space X is called $\ell_{p,q}(\mathcal{K})$-*hilbertian* if the synthesis-type operator

$$F : \ell_{p,q}(\mathcal{K}) \to X, \quad (c_\kappa)_{\kappa\in\mathcal{K}} \mapsto \sum_{\kappa\in\mathcal{K}} c_\kappa g_\kappa$$

is well-defined and bounded. For $q = 1$, hilbertian dictionaries are characterized in the following Proposition.

Proposition 3.6. *Let $\mathcal{D} = \{g_\kappa : \kappa \in \mathcal{K}\}$ be a dictionary in a Banach space X and $1 \le p < \infty$. Then the following properties are equivalent:*

(i) *$\mathcal{D}$ is $\ell_{p,1}(\mathcal{K})$-hilbertian.*

(ii) *For all index sets $\mathcal{K}_N \subset \mathcal{K}$ of cardinality N and every choice of signs*

$$\Big\| \sum_{\kappa\in\mathcal{K}_N} \pm g_\kappa \Big\|_X \lesssim N^{\frac{1}{p}}.$$

(iii) *For all index sets $\mathcal{K}_N \subset \mathcal{K}$ of cardinality N and every sequence $(d_\kappa)_{\kappa\in\mathcal{K}_N} \in \ell(\mathcal{K}_N)$*

$$\Big\| \sum_{\kappa\in\mathcal{K}_N} d_\kappa g_\kappa \Big\|_X \lesssim N^{\frac{1}{p}} \max_{\kappa\in\mathcal{K}_N} |d_\kappa|. \tag{14}$$

The equivalence between (i) and (ii) has already been derived in [19]. We extend the result to condition (iii).

Proof. Obviously, (iii) implies (ii). Let us show that (i) implies (iii). Given $(d_\kappa)_{\kappa\in\mathcal{K}_N} \in \ell(\mathcal{K}_N)$, its zero extension $(c_\kappa)_{\kappa\in\mathcal{K}}$ is contained in $\ell_{p,1}(\mathcal{K})$. Applying (i) yields

$$\begin{aligned} \Big\| \sum_{\kappa\in\mathcal{K}_N} d_\kappa g_\kappa \Big\|_X &\lesssim \|(c_\kappa)_{\kappa\in\mathcal{K}}\|_{\ell_{p,1}} = \sum_{j=1}^\infty j^{\frac{1}{p}-1} c_j^* \\ &\le \max_{\kappa\in\mathcal{K}_N} |c_\kappa| \sum_{j=1}^N j^{\frac{1}{p}-1} = \max_{\kappa\in\mathcal{K}_N} |d_\kappa|\, N^{\frac{1}{p}} \frac{1}{N} \sum_{j=1}^N \big(\tfrac{j}{N}\big)^{\frac{1}{p}-1}. \end{aligned}$$

A Riemann sum argument provides

$$\frac{1}{N}\sum_{j=1}^{N}\left(\tfrac{j}{N}\right)^{\frac{1}{p}-1} \le \int_0^1 x^{\frac{1}{p}-1}dx = p.$$

This concludes the proof. □

According to the results in [1], compactly supported dyadic wavelet systems, properly normalized in $L_p(\mathbb{R}^d)$, are $\ell_{p,1}$-hilbertian. We can extend this result to isotropic scalings:

Corollary 3.7. *Given M isotropic, let $\psi^{(\mu)}$, $\mu = 1,\dots,n$, be compactly supported functions in $L_\infty(\mathbb{R}^d)$ and $1 \le p < \infty$. Then, with $\Lambda = \{1,\dots,n\}\times\mathbb{Z}\times\mathbb{Z}^d$, the L_p-normalized wavelet system $\{\psi_\lambda^p : \lambda\in\Lambda\}$ is an $\ell_{p,1}(\Lambda)$-hilbertian dictionary in $L_p(\mathbb{R}^d)$.*

Proof. The following estimate is a standard component in nonlinear approximation theory for dyadic dilation, cf. [6, 12],

$$\Big\| \sum_{\lambda\in\Lambda_N} d_\lambda\psi_\lambda\Big\|_{L_p} \lesssim N^{\frac{1}{p}} \max_{\lambda\in\Lambda_N}\|d_\lambda\psi_\lambda\|_{L_p}. \tag{15}$$

See [26], for this estimate with respect to wavelet bases with isotropic scaling. An analysis of its proof yields that the bases assumption is not necessary, and (15) holds in our situation. Actually, (15) is just a rephrasing of (14) involving the $L_p(\mathbb{R}^d)$-normalization, and applying Proposition 3.6 concludes the proof. □

According to Corollary 3.7, the series in (13) converges in $L_p(\mathbb{R}^d)$ if $(c_\lambda)_{\lambda\in\Lambda}$ is contained in $\ell_{p,1}(\Lambda)$. In order to consider wavelet bi-frame expansions

$$f = \sum_{\lambda\in\Lambda}\langle f, \widetilde{\psi}_\lambda^{p'}\rangle\psi_\lambda^p \tag{16}$$

in $L_p(\mathbb{R}^d)$, there still remain two problems. First, we have to verify that the coefficient sequence $\big(\langle f,\widetilde{\psi}_\lambda^{p'}\rangle\big)_{\lambda\in\Lambda}$ is contained in $\ell_{p,1}(\Lambda)$. Then the right-hand side of (16) converges in $L_p(\mathbb{R}^d)$. Second, we have to verify that the series converges towards f. Both problems are addressed in the following Subsection.

3.4 A Characterization by Wavelet Bi-Frames

In this subsection, we finally derive the characterization of Besov spaces by wavelet bi-frames with general isotropic scalings. The following theorem extends dyadic results in [1].

Theorem 3.8. *Given $1 < p < \infty$, $\frac{1}{p}+\frac{1}{p'} = 1$, let $X(\{\psi^{(1)},\dots,\psi^{(n)}\})$, $X(\{\widetilde{\psi}^{(1)},\dots,\widetilde{\psi}^{(n)}\})$ be a compactly supported wavelet bi-frame. In addition, suppose that $X(\{\eta^{(1)},\dots,\eta^{(m-1)}\})$, $X(\{\widetilde{\eta}^{(1)},\dots,\widetilde{\eta}^{(m-1)}\})$ is a pair of compactly supported biorthogonal wavelet bases. Given $s,s'\in\mathbb{N}$, then let, for $\mu = 1,\dots,n$ and $\nu = 1,\dots,m-1$, $\psi^{(\mu)},\eta^{(\nu)}\in W^s(L_\infty(\mathbb{R}^d))$ and $\widetilde{\psi}^{(\mu)},\widetilde{\eta}^{(\nu)}\in W^{s'}(L_\infty(\mathbb{R}^d))$ have s' and s vanishing moments, respectively. If the pair of biorthogonal wavelet bases characterizes B^α in the sense of (7) and (8), then we have, for α in the range*

$$0 < \alpha < \begin{cases} s, & \text{for } \frac{s}{d}+\frac{1}{p}\ge 1,\\ \frac{s}{p\left(1-\frac{s}{d}\right)}, & \text{for } \frac{s}{d}+\frac{1}{p}\le 1,\end{cases} \tag{17}$$

and for all $f \in B^\alpha$, that

$$f = \sum_{\lambda \in \Lambda} \langle f, \widetilde{\psi}_\lambda^{p'} \rangle \psi_\lambda^p \tag{18}$$

holds in $L_p(\mathbb{R}^d)$ and

$$|f|_{B^\alpha} \sim \big\| \big(\langle f, \widetilde{\psi}_\lambda^{p'} \rangle \big)_{\lambda \in \Lambda} \big\|_{\ell_\tau(\Lambda)}, \quad \textit{for } \tfrac{1}{\tau} = \tfrac{\alpha}{d} + \tfrac{1}{p}.$$

Remark 3.9. Theorem 3.8 requires the existence of a pair of biorthogonal reference wavelet bases, which already characterize the Besov space. In the dyadic setting of [1], this assumption is not explicitly mentioned since one can simply choose tensor products of sufficiently smooth orthonormal Daubechies wavelets or the Meyer wavelet, see [8]. As far as we know, it is still an open problem, whether, for each isotropic dilation matrix, one can find families of arbitrarily smooth compactly supported pairs of biorthogonal wavelet bases. Hence, we had to formulate the existence of a reference system as an assumption in Theorem 3.8. We should point out, that, for many nondyadic scalings, these families exist, cf. [11, 22], and the characterization is applicable. Since one allows for arbitrarily large support sizes, we expect that the overwhelming majority of isotropic dilation matrices has such biorthogonal reference wavelets.

For preparation, we need the following lemma. We omit the simple proof.

Lemma 3.10. *Let $1 \leq q < p \leq \infty$ and let $f_n \in L_p(\mathbb{R}^d) \cap L_q(\mathbb{R}^d)$, $n \in \mathbb{N}$, converge to f in $L_p(\mathbb{R}^d)$ and to g in $L_q(\mathbb{R}^d)$. Then $f = g$ up to a set of measure zero.*

Proof of Theorem 3.8. Let $f \in B^\alpha$ and $\Lambda' = \{1, \ldots, m-1\} \times \mathbb{Z} \times \mathbb{Z}^d$, then

$$f = \sum_{\lambda' \in \Lambda'} \left\langle f, \widetilde{\eta}_{\lambda'}^{p'} \right\rangle \eta_{\lambda'}^p$$

holds in $L_p(\mathbb{R}^d)$ and

$$|f|_{B^\alpha} \sim \big\| \big(\langle f, \widetilde{\eta}_{\lambda'}^{p'} \rangle \big)_{\lambda' \in \Lambda'} \big\|_{\ell_\tau}. \tag{19}$$

For $\frac{s}{d} + \frac{1}{p} \geq 1$, we have $\frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p} < \frac{s}{d} + \frac{1}{p}$. Hence, $p > \tau > \left(\frac{s}{d} + \frac{1}{p}\right)^{-1}$, and τ is in the admissible range of Theorem 3.1. For $\frac{s}{d} + \frac{1}{p} \leq 1$, we have

$$\begin{aligned} \frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p} &< \frac{s}{p(d-s)} + \frac{1}{p} \\ &= \frac{d}{p(d-s)} = \frac{1}{p\left(1 - \frac{s}{d}\right)}. \end{aligned}$$

Thus, Theorem 3.1 can be applied in both cases. Then we obtain

$$\begin{aligned} \big\| \big(\langle f, \widetilde{\psi}_\lambda^{p'} \rangle \big)_{\lambda \in \Lambda} \big\|_{\ell_\tau} &= \Big\| \Big(\sum_{\lambda' \in \Lambda'} \langle f, \widetilde{\eta}_{\lambda'}^{p'} \rangle \langle \eta_{\lambda'}^p, \widetilde{\psi}_\lambda^{p'} \rangle \Big)_{\lambda \in \Lambda} \Big\|_{\ell_\tau} &(20) \\ &\lesssim \big\| \big(\langle f, \widetilde{\eta}_{\lambda'}^{p'} \rangle \big)_{\lambda' \in \Lambda'} \big\|_{\ell_\tau}. &(21) \end{aligned}$$

With (19), this implies

$$\big\| \big(\langle f, \widetilde{\psi}_\lambda^{p'} \rangle \big)_{\lambda \in \Lambda} \big\|_{\ell_\tau} \lesssim |f|_{B^\alpha}. \tag{22}$$

For the reverse estimate, wavelet bi-frame and biorthogonal wavelets change roles in the localization process. First, we establish (18). According to Corollary 3.7, the primal bi-frame wavelets $\left\{ \psi_\lambda^p : \lambda \in \Lambda \right\}$ are $\ell_{p,1}$-hilbertian. Hence, the synthesis-type operator

$$F : \ell_{p,1} \to L_p(\mathbb{R}^d), \quad (d_\lambda)_{\lambda \in \Lambda} \mapsto \sum_{\lambda \in \Lambda} d_\lambda \psi_\lambda^p$$

is well-defined and bounded. By applying (22), the analysis-type operator

$$\widetilde{F}^* : B^\alpha \to \ell_\tau, \quad f \mapsto \big(\langle f, \widetilde{\psi}^{p'}_{\lambda'}\rangle\big)_{\lambda'}$$

is bounded (the notation may only remind of the original analysis operator on Hilbert spaces. The present operator $\widetilde{F}^*$ is neither considered as any adjoint on Hilbert spaces nor any dual operator on Banach spaces). Due to $\ell_\tau(\Lambda) \hookrightarrow \ell_{p,1}(\Lambda)$, we can consider the bounded operator

$$F\widetilde{F}^* : B^\alpha \to L_p(\mathbb{R}^d)$$

more closely. Since B^α is contained in $L_p(\mathbb{R}^d)$, Lemma 3.10 and the bi-frame expansion in $L_2(\mathbb{R}^d)$ imply that $F\widetilde{F}^*$ is the identity on $B^\alpha \cap L_2(\mathbb{R}^d)$. According to the results of Chapter 1 in [29], the intersection $B^\alpha \cap L_2(\mathbb{R}^d)$ is dense in B^α. Hence, the continuity of $F\widetilde{F}^*$ finally yields that (18) holds in $L_p(\mathbb{R}^d)$.

By following (20), (21) with interchanged roles of $\widetilde{\psi}$, $\widetilde{\eta}$ as well as η replaced by ψ, we obtain the reverse estimate of (22). □

4 Determining the Approximation Classes

4.1 Jackson Inequality

The following theorem establishs a Jackson inequality for wavelet bi-frames with isotropic scalings.

Theorem 4.1. *Let M be isotropic and $1 < p < \infty$. Given a compactly supported wavelet bi-frame $X(\{\psi^{(1)}, \dots, \psi^{(n)}\})$, $X(\{\widetilde{\psi}^{(1)}, \dots, \widetilde{\psi}^{(n)}\})$, let the assumptions of Theorem* 3.8 *hold. If α is in the range of* (17)*, then*

$$\sigma_N(f, X(\{\psi^{(1)}, \dots, \psi^{(n)}\}))_{L_p} \lesssim N^{-\frac{\alpha}{d}} |f|_{B^\alpha}, \quad \textit{for all } f \in B^\alpha,\ N \in \mathbb{N}.$$

Proof. By Corollary 3.7, the system $\{\psi^p_\lambda : \lambda \in \Lambda\}$ is $\ell_{p,1}(\Lambda)$-hilbertian. Thus, the general Jackson inequality for hilbertian dictionaries in [19] is applicable, which yields

$$\sigma_N(f, X(\{\psi^{(1)}, \dots, \psi^{(n)}\}))_{L_p} \lesssim N^{-\frac{\alpha}{d}} \inf\Big\{ \|(c_\lambda)_{\lambda\in\Lambda}\|_{\ell_\tau} : f = \sum_{\lambda\in\Lambda} c_\lambda \psi^p_\lambda \Big\}, \tag{23}$$

where $\frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p}$. Given $f \in B^\alpha$ and according to Theorem 3.8, the expansion

$$f = \sum_{\lambda\in\Lambda} \langle f, \widetilde{\psi}^{p'}_\lambda\rangle \psi^p_\lambda$$

holds in $L_p(\mathbb{R}^d)$ and $\big(\langle f, \widetilde{\psi}^{p'}_\lambda\rangle\big)_{\lambda\in\Lambda} \in \ell_\tau(\Lambda)$, where $\frac{1}{p} + \frac{1}{p'} = 1$. Thus, the right-hand side of (23) is bounded by $N^{-\frac{\alpha}{d}} \big\|\big(\langle f, \widetilde{\psi}^{p'}_\lambda\rangle\big)_{\lambda\in\Lambda}\big\|_{\ell_\tau}$ Then the norm equivalence of Theorem 3.8 concludes the proof. □

4.2 Bernstein Inequality

In this subsection, we establish a Bernstein inequality for wavelet bi-frames with idempotent scaling. It requires an independence assumption as we shall introduce next. Given a nonempty open subset $A \subset \mathbb{R}^d$, we say a compactly supported distribution φ has *linearly independent integer shifts on A* if

$$\sum_{k\in\mathbb{Z}^d} c_k \varphi(\cdot - k) = 0 \quad \text{on } A,$$

implies $c_k\varphi(\cdot - k) = 0$ on A, for all $k \in \mathbb{Z}^d$. Our result is based on the generalization of the following dyadic Bernstein inequality from [23].

Theorem 4.2. *Given $M = 2\mathcal{I}_d$ and $1 < p < \infty$, let $\varphi \in W^s(L_\infty(\mathbb{R}^d))$, $s \in \mathbb{N}$, be a compactly supported refinable function with linearly independent integer shifts on $(0,1)^d$. Then, for each $0 < \alpha < s$,*

$$|f|_{B^\alpha} \lesssim N^{\frac{\alpha}{d}} \|f\|_{L_p(\mathbb{R}^d)}, \quad \text{for all } f \in \Sigma_N\left(X(\{\varphi\})\right).$$

By following the lines of the proof in [23], one verifies that Theorem 4.2 still holds for a dilation matrix $M = h\mathcal{I}_d$, where $h \in \mathbb{N}$. This observation is the key ingredient for the proof of the following corollary. It generalizes the dyadic result in [1] regarding idempotent dilation matrices M (recall that a dilation matrix M is called idempotent if there exist $l, h \in \mathbb{N}$ such that $M^l = h\mathcal{I}_d$).

Corollary 4.3. *Given an idempotent dilation matrix M and $1 < p < \infty$, let $\varphi \in W^s(L_\infty(\mathbb{R}^d))$ be a compactly supported refinable function with finitely supported mask and with linearly independent integer shifts on $(0,1)^d$. Moreover, let $\psi^{(1)}, \ldots, \psi^{(n)}$ be wavelets with finitely supported sequences $\big(a_k^{(\mu)}\big)_{k\in\mathbb{Z}^d}$ such that (4) holds. Then, for $0 < \alpha < s$,*

$$|f|_{B^\alpha} \lesssim N^{\frac{\alpha}{d}} \|f\|_{L_p(\mathbb{R}^d)}, \quad \text{for all } f \in \Sigma_N(X(\{\psi^{(1)}, \ldots, \psi^{(n)}\})).$$

Proof. According to (4), we have for each $\mu = 1, \ldots, n$,

$$\psi^{(\mu)}(M^j x - k') = \sum_{k\in\mathbb{Z}^d} a_k^{(\mu)} \varphi(M^{j+1}x - Mk' - k), \quad \text{for all } j \in \mathbb{Z},\ k' \in \mathbb{Z}^d.$$

Thus, there exists a constant C_1 such that $\psi_{j,k}^{(1)}, \ldots, \psi_{j,k}^{(n)} \in \Sigma_{C_1}(X(\{\varphi\}))$. This implies

$$\Sigma_N(X(\{\psi^{(1)}, \ldots, \psi^{(n)}\})) \subset \Sigma_{C_1 N}(X(\{\varphi\})). \tag{24}$$

Let $(a_k)_{k\in\mathbb{Z}^d}$ be the finitely supported mask of φ, and let l and h be contained in $\mathbb{N}$ such that $M^l = h\mathcal{I}_d$. In the sequel, we verify that there exists a uniform constant C_2 such that, for all $j' \in \mathbb{Z}$ and $k' \in \mathbb{Z}^d$

$$\varphi(M^{j'}x - k') \in \Sigma_{C_2}(\{\varphi(h^j x - k) : j \in \mathbb{Z},\ k \in \mathbb{Z}^d\}). \tag{25}$$

Note that we can find $u \in \mathbb{Z}$ and $r \in \mathbb{N}$, $r < l$ such that $j' + r = lu$. Then r-times applying the refinement equation (3) provides

$$\begin{aligned}\varphi(M^{j'}x - k') &= \sum_{k_1,\ldots,k_r} a_{k_1}\cdots a_{k_r} \varphi(M^r M^{j'}x - M^{r-1}k_1 - \ldots - Mk_{r-1} - k_r)\\ &= \sum_{k_1,\ldots,k_r} a_{k_1}\cdots a_{k_r} \varphi(M^{lu}x - M^{r-1}k_1 - \ldots - Mk_{r-1} - k_r).\end{aligned}$$

According to $M^l = h\mathcal{I}_d$, the last term is contained in

$$\Sigma_{C^r}(\{\varphi(h^j x - k) : j \in \mathbb{Z},\ k \in \mathbb{Z}^d\}),$$

where C denotes the number of nonzero entries of the mask $(a_k)_{k\in\mathbb{Z}^d}$. Since $r < l$, (25) holds with $C_2 = C^{l-1}$.

From (25), we derive

$$\Sigma_N(X(\{\varphi\})) \subset \Sigma_{C_2 N}(\{\varphi(h^j x - k) : j \in \mathbb{Z},\ k \in \mathbb{Z}^d\}),$$

which provides with (24)

$$\Sigma_N(X(\{\psi^{(1)}, \ldots, \psi^{(n)}\})) \subset \Sigma_{C_2 C_1 N}(\{\varphi(h^j x - k) : j \in \mathbb{Z},\ k \in \mathbb{Z}^d\}).$$

Then applying Theorem 4.2 to $h\mathcal{I}_d$ yields

$$\begin{aligned}|f|_{B^{\alpha}_{\tau}(L_\tau(\mathbb{R}^d))} &\lesssim (C_2C_1N)^{\frac{\alpha}{d}}\|f\|_{L_p(\mathbb{R}^d)}\\ &\lesssim N^{\frac{\alpha}{d}}\|f\|_{L_p(\mathbb{R}^d)},\end{aligned}$$

for all $f \in \Sigma_N(X(\{\psi^{(1)},\dots,\psi^{(n)}\}))$. □

Remark 4.4. First, the restriction of the Bernstein inequality to idempotent dilation matrices is only a technical requirement. Most of the isotropic dilation matrices addressed in the literature are idempotent. Second, the arising constants are already far from being optimal in the Jackson inequality of Theorem 4.1, see also Remark 3.9. However, our proof of the Bernstein inequality yields to a certain extent an explosion since the constants linearly depend on the number of nonzero entries of the underlying masks and they even exponentially depend on the idempotence of the scaling. Nevertheless, we could derive the qualitative result, and we are convinced that the true constants are much better.

The Bernstein inequality in Corollary 4.3 requires that the shifts of the underlying refinable function are linearly independent on the unit cube. Jia conjectures in [23] that the assumption can be removed. However, there is no proof so far, and the application of the corollary requires the verification of this condition.

4.3 Approximation Classes as Besov Spaces

Let us collect the results of the previous subsections. If the assumptions of Theorem 4.1 and Corollary 4.3 are satisfied, we have established matching Jackson and Bernstein inequalities, which yields

$$\mathcal{A}^{\frac{\alpha}{d}}_{\tau}\big(L_p(\mathbb{R}^d), X(\{\psi^{(1)},\dots,\psi^{(n)}\})\big) = \Big[L_p(\mathbb{R}^d), B^s\Big]_{\frac{\alpha}{s},\tau}.$$

For $\frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p}$, the right-hand side equals the Besov space B^α, cf. [25], see also [7, 13] and the survey article [12]. Hence, the approximation class is essentially a Besov space. The following theorem is an explicit summary of our results:

Theorem 4.5. *Given an idempotent dilation matrix M and $1 < p < \infty$, let $X(\{\psi^{(1)},\dots,\psi^{(n)}\})$, $X(\{\widetilde{\psi}^{(1)},\dots,\widetilde{\psi}^{(n)}\})$ be a compactly supported wavelet bi-frame. Moreover, let their primal refinable function $\varphi \in W^s(L_\infty(\mathbb{R}^d))$, $s \in \mathbb{N}$, have linearly independent integer shifts on $(0,1)^d$. Suppose that the assumptions of Theorem 3.8 hold. Then for α in the range of (17), we have*

$$\mathcal{A}^{\frac{\alpha}{d}}_{\tau}\big(L_p(\mathbb{R}^d), X(\{\psi^{(1)},\dots,\psi^{(n)}\})\big) = B^\alpha, \quad \text{where } \frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p}.$$

In order to apply Theorem 4.5 to the Checkerboard wavelet bi-frames in [15, 16], we have to verify that their underlying refinable functions have linearly independent integer shifts on the unit cube. We address this topic in the final Section 6. Before, we complete the theoretical framework, and we derive a realization of the best N-term approximation rate in the following Section 5.

5 N-term Approximation by Thresholding

Theorem 4.5 describes best N-term approximation. In order to implement practical algorithms, we still need a rule for the selection of N particular terms. In other words, we want to realize the best N-term approximation rate. For pairs of biorthogonal wavelet bases, one can simply select the N largest coefficients of the series expansion, cf. [12, 26]. This procedure also works for dyadic wavelet bi-frames by thresholding the bi-frame expansion, see [1] for details. In the following, we extend these results to wavelet bi-frames with general isotropic scalings. Moreover, we allow for more general thresholding operators. They are considered in [2] with respect to unconditional bases. An analysis of the proof yields that the results hold true for wavelet bi-frames as well. The critical ingredients are the following:

- $\{\psi^p_\lambda : \lambda \in \Lambda\}$ is $\ell_{p,1}$-hilbertian,
- $\|(\langle f, \widetilde{\psi}^{p'}_\lambda\rangle)_{\lambda\in\Lambda}\|_{\ell_\tau} \lesssim |f|_{B^\alpha}$, for all $f \in B^\alpha$ and $\frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p}$.

Given a wavelet bi-frame $X(\{\psi^{(1)}, \dots \psi^{(n)}\})$, $X(\{\widetilde{\psi}^{(1)}, \dots \widetilde{\psi}^{(n)}\}$ satisfying the assumptions of Theorem 4.1, these points are satisfied. Then following [2], we call a function $\varrho : \mathbb{C} \times \mathbb{R}_+ \to \mathbb{C}$ a *thresholding rule* if

$$|x - \varrho(x, \delta)| \lesssim \min(|x|, \delta)$$

and $|x| \lesssim \delta$ implies $\varrho(x, \delta) = 0$. It should be mentioned that hard- and soft-thresholding, see for instance [4, 30], as well as garotte-thresholding as described in [18] constitute thresholding rules. Under the notation and the assumptions of Theorem 3.8, let ϱ be a thresholding rule. Then

$$T_\varrho : B^\alpha \times \mathbb{R}_+ \to L_p(\mathbb{R}^d), \quad (f, \delta) \mapsto \sum_{\lambda\in\Lambda} \varrho\big(\langle f, \widetilde{\psi}^{p'}_\lambda\rangle, \delta\big)\psi^p_\lambda \tag{26}$$

is called the associated *thresholding operator*. Since $\big(\langle f, \widetilde{\psi}^{p'}_\lambda\rangle\big)_{\lambda\in\Lambda}$ is contained in $\ell_\tau(\Lambda)$, the series (26) is actually a finite sum. Note that the operator is applied to the bi-frame coefficients, and one does not allow for thresholding an arbitrary expansion. By denoting

$$N_{f,\delta} := \operatorname{card}\big\{\lambda \in \Lambda : \varrho\big(\langle f, \widetilde{\psi}^{p'}_\lambda\rangle, \delta\big) \neq 0\big\},$$

we have for α in the range of (17) and for all $f \in B^\alpha$,

$$\|f - T_\varrho(f, \delta)\|_{L_p} \lesssim N_{f,\delta}^{-\frac{\alpha}{d}} |f|_{B^\alpha}.$$

Thus, the best N-term approximation rate as described in Theorem 4.5 can be realized by thresholding the wavelet bi-frame expansion. Note that this result does not require any linear independence. Hence, even if the assumptions of the Bernstein inequality are not satisfied and so the best N-term approximation is not completely described, thresholding still provides the same approximation rate as predicted by the Jackson inequality.

6 Checkerboard Wavelet Bi-Frames

Given M as in (5), (6), a family of wavelet bi-frames in arbitary dimensions with arbitrarily high smoothness and an arbitrarily high number of vanishing moments is derived from only 3 wavelets in [15]. They satisfy a variety of optimality conditions and we refer to them as *Checkerboard wavelet bi-frames* in the present paper since M generates the *checkerboard lattice*, i.e.,

$$M\mathbb{Z}^d = \Big\{(k_1, \dots, k_d)^\top \in \mathbb{Z}^d : \sum_{i=1}^d k_i \in 2\mathbb{Z}\Big\}.$$

The present section is dedicated to determining the best N-term approximation class of the Checkerboard wavelet bi-frames.

The scaling is idempotent, but before we can apply our results of the previous sections, we have to verify that the underlying primal refinable function of the wavelet bi-frame has linearly independent integer shifts on the unit cube.

Advantageously, the underlying refinable function φ of the Checkerboard wavelet bi-frame is explicitly given by

$$\varphi(x) = \varphi_0 \otimes \cdots \otimes \varphi_0 \, (Dx)\,, \tag{27}$$

where φ_0 is a univariate dyadic refinable function and D is a square matrix with ones in the diagonal as well as above, and zeros elsewhere, cf. [15]. Moreover, φ_0 is *fundamental*, i.e., $\varphi_0(k) = \delta_{0,k}$, for all $k \in \mathbb{Z}$. Therefore, φ_0 has linearly independent integer shifts on $\mathbb{R}$. Since φ_0 is univariate, this yields that it even has *locally linearly independent integer shifts*, i.e., its integer shifts are linearly independent on each nonempty open subset in $\mathbb{R}^d$, cf. [28].

In the following, we will verify that local linear independence of integer shifts is invariant under tensor products of univariate refinable functions as well as under the action of D.

Proposition 6.1. *Let φ_0 be a univariate, continuous, dyadic refinable function with compact support. If its integer shifts are linearly independent on $\mathbb{R}$, then the tensor product $\varphi = \bigotimes_{i=1}^{d} \varphi_0$ has locally linearly independent integer shifts.*

The following proof of Proposition 6.1 is direct, see [17] for an alternative proof in terms of the mask of the refinable function.

Proof. Given a nonempty open subset A in $\mathbb{R}^d$, let x be an arbitrary point in A. Then there exists an open cube $U_x \subset A$, whose edges are parallel to the coordinate axis, and x is contained in U_x. Thus, we have open subsets U_{x_i} in $\mathbb{R}$, $i = 1, \ldots, d$, such that

$$U_x = U_{x_1} \times \cdots \times U_{x_d}.$$

According to [28], φ_0 has locally linearly independent integer shifts. Hence, for each $i = 1, \ldots, d$, the collection

$$B_i := \{\varphi_0(\cdot - k_i) : \operatorname{supp}(\varphi_0(\cdot - k_i)) \cap U_{x_i} \neq \emptyset,\ k_i \in \mathbb{Z}\}$$

is linearly independent. Therefore, the collection of tensor products

$$\begin{aligned} B_1 \otimes \cdots \otimes B_d &= \{\varphi(\cdot - k) : \operatorname{supp}(\varphi_0(\cdot - k_i)) \cap U_{x_i} \neq \emptyset,\ k_i \in \mathbb{Z},\ i = 1, \ldots, d\} \\ &= \left\{\varphi(\cdot - k) : \operatorname{supp}(\varphi(\cdot - k)) \cap U_x \neq \emptyset,\ k \in \mathbb{Z}^d\right\} \end{aligned}$$

is also linearly independent. Thus, φ has linearly independent integer shifts on U_x. Since $A = \bigcup_{x \in A} U_x$, the integer shifts of φ are linearly independent on A. □

Next, we address the action of D in (27).

Lemma 6.2. *Let $\varphi : \mathbb{R}^{d_1} \to \mathbb{C}$ have locally linearly independent integer shifts, and let $D \in \mathbb{Z}^{d_1 \times d_2}$ be an integer matrix of rank d_1. Then $\varphi(D\cdot) : \mathbb{R}^{d_2} \to \mathbb{C}$ has locally linearly independent integer shifts.*

Proof. Given some nonempty open subset A in $\mathbb{R}^{d_2}$, let

$$\sum_{k \in \mathbb{Z}^{d_2}} c_k \varphi(D(\cdot - k)) = 0, \quad \text{on } A.$$

This implies $\sum_{k \in \mathbb{Z}^{d_2}} c_k \varphi(\cdot - Dk) = 0$, on DA. A trivial zero extension yields

$$\sum_{k \in \mathbb{Z}^{d_2}} c_k \varphi(\cdot - Dk) + \sum_{k \in \mathbb{Z}^{d_1} \setminus D\mathbb{Z}^{d_2}} 0 \cdot \varphi(\cdot - k) = 0, \quad \text{on } DA.$$

Since $D : \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}$ is linear and onto, it constitutes an open mapping, i.e., DA is an open subset of $\mathbb{R}^{d_1}$. Hence, the local linear independence of φ provides $c_k \varphi(\cdot - Dk) = 0$, on DA, for all $k \in \mathbb{Z}^{d_2}$. Finally, this yields $c_k \varphi(D(\cdot - k)) = 0$, on A, for all $k \in \mathbb{Z}^{d_2}$, which concludes the proof. □

By applying Proposition 6.1 and Lemma 6.2, the underlying refinable function φ in (27) of the Checkerboard wavelet bi-frame has locally linearly independent integer shifts. Hence, Theorem 4.5 can be applied:

Example 6.3. Given $1 < p < \infty$ and an arbitrarily large number $0 < s \in \mathbb{N}$, there is a Checkerboard wavelet bi-frame with a sufficiently high order of smoothness and sufficiently many vanishing moments such that, for all $0 < \alpha < s$,

$$\mathcal{A}_\tau^{\frac{\alpha}{d}}\left(L_p(\mathbb{R}^d), X(\{\psi^{(1)}, \psi^{(2)}, \psi^{(3)}\})\right) = B^\alpha, \quad \text{where } \tfrac{1}{\tau} = \tfrac{\alpha}{d} + \tfrac{1}{p}.$$

Note that the above equality holds in arbitrary dimensions with only three wavelets. According to Section 5, the best N-term approximation rate can be realized by an arbitrary thresholding operator.

Remark 6.4. In [16], the number of wavelets of the Checkerboard wavelet bi-frames could even be reduced to 2 and the underlying refinable function is still given by (27). Therefore, Example 6.3 even holds with respect to this reduced counterpart with only two wavelets in arbitrary dimensions.

Acknowledgments

The author thanks Professor Stephan Dahlke for valuable comments on the presentation of the manuscript. He also thanks Professor Bin Han for simplifying the proof of Proposition 6.1.

References

[1] L. Borup, R. Gribonval, and M. Nielsen, *Bi-framelet systems with few vanishing moments characterize Besov spaces*, Appl. Comput. Harm. Anal. **17** (2004), no. 1, 3–28.

[2] L. Borup and M. Nielsen, *Some remarks on shrinkage operators*, preprint.

[3] C. Canuto and A. Tabacco, *Multilevel decompositions of functional spaces*, J. Fourier Anal. Appl. **3** (1997), no. 6, 715–742.

[4] A. Chambolle, R. A. DeVore, N. Y. Lee, and B. J. Lucier, *Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Trans. Image Process. **7** (1998), 319–335.

[5] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, 2003.

[6] A. Cohen, *Numerical Analysis of Wavelet Methods*, vol. 32, Studies in Mathematics and its Applications, Amsterdam, 2003.

[7] A. Cohen, R. A. DeVore, and R. Hochmuth, *Restricted nonlinear approximation*, Constr. Approx. **16** (2000), 85–113.

[8] I. Daubechies, *Ten Lectures on Wavelets*, vol. 9, SIAM, Philadelphia, 1992, CBMS-NSF Regional Conf. Ser. in Appl. Math. 61.

[9] I. Daubechies, B. Han, A. Ron, and Z. Shen, *Framelets: MRA-based constructions of wavelet frames*, Appl. Comput. Harm. Anal. **14** (2003), no. 1, 1–46.

[10] S. Dekel and D. Leviatan, *The Bramble-Hilbert lemma for convex domains*, SIAM J. Math. Anal. **35** (2004), 1203–1212.

[11] J. Derado, *Multivariate refinable interpolating functions*, Appl. Comput. Harm. Anal. **7** (1999), 165–183.

[12] R. A. DeVore, *Nonlinear approximation*, Acta Numerica (1998), 51–150.

[13] R. A. DeVore, B. Jawerth, and V. A. Popov, *Compression of wavelet decompositions*, American Journal of Math. **114** (1992), 737–785.

[14] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, 1993.

[15] M. Ehler, *Compactly supported multivariate pairs of dual wavelet frames obtained by convolution*, to appear in: Int. J. Wavelets Multiresolut. Inf. Process.

[16] ______, *On multivariate compactly supported bi-frames*, J. Fourier Anal. Appl. **13** (2007), no. 5, 511–532.

[17] ______, *The construction of nonseparable wavelet bi-frames and associated approximation schemes*, Ph.D. thesis, University of Marburg, 2007.

[18] H. Y. Gao, *Wavelet shrinkage denoising using the non-negative garotte*, J. Comput. Graph. Statist. **7** (1998), no. 4, 469–488.

[19] R. Gribonval and M. Nielsen, *Nonlinear approximation with dictionaries. I. Direct estimates*, J. Fourier Anal. Appl. **10** (2004), no. 1, 51–71.

[20] K. Gröchenig, *Localized frames are finite unions of riesz sequences*, Adv. Comp. Math. **18** (2003), 149–157.

[21] ______, *Localization of frames, banach frames, and the invertibility of the frame operator*, J. Fourier Anal. Appl. **10** (2004), no. 2, 105–132.

[22] H. Ji, S. D. Riemenschneider, and Z. Shen, *Multivariate compactly supported fundamental refinable functions, duals and biorthogonal wavelets*, Stud. Appl. Math. **102** (1999), 173–204.

[23] R. Q. Jia, *A bernstein-type inequality associated with wavelet decomposition*, Constr. Approx **9** (1993), 299–318.

[24] ______, *Approximation properties of multivariate wavelets*, Math. Comp. **67** (1998), 647–665.

[25] G. Kyriazis, *Non-linear approximation and interpolation spaces*, J. Approx. **113** (2001), 110–126.

[26] M. Lindemann, *Approximation properties of non-separable wavelet bases with isotropic scaling matrices and their relation to besov spaces*, Ph.D. thesis, University of Bremen, 2005.

[27] Y. Meyer and R. Coifman, *Wavelets. Calderon zygmund and multilinear operators*, Cambridge University Press, 1997.

[28] A. Ron, *Introduction to shift-invariant spaces I: Linear independence*, Multivariate Approximation and Applications (N. Dyn et al., ed.), 1999, pp. 112–151.

[29] T. Runst and W. Sickel, *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations*, Walter de Gruyter, Berlin, 1996.

[30] T. Tao, *On the almost everywhere convergence of wavelet summation methods*, Appl. Comput. Harm. Anal. **3** (1996), 384–387.