AALBORG UNIVERSITY

Bi-framelet systems with few vanishing moments characterize Besov spaces

by

Lasse Borup, Rémi Gribonval, and Morten Nielsen

November 2003 R-2003-18

DEPARTMENT OF MATHEMATICAL SCIENCES
AALBORG UNIVERSITY
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 96 35 80 80 ▪ Telefax: +45 98 15 81 29
URL: www.math.auc.dk/research/reports/reports.htm

ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

BI-FRAMELET SYSTEMS WITH FEW VANISHING MOMENTS CHARACTERIZE BESOV SPACES

LASSE BORUP, RÉMI GRIBONVAL[‡], AND MORTEN NIELSEN[†]

ABSTRACT. We study the approximation properties of wavelet bi-frame systems in $L_p(\mathbb{R}^d)$. For wavelet bi-frame systems the approximation spaces associated with best m-term approximation are completely characterized for a certain range of smoothness parameters limited by the number of vanishing moments of the functions in the dual frame. The approximation spaces turn out to be essentially Besov spaces, just as in the classical orthonormal wavelet case. We also prove that for smooth functions, the canonical expansion in the wavelet bi-frame system is sparse and one can reach the optimal rate of approximation by simply thresholding the canonical expansion. For twice oversampled MRA based wavelet frames, a characterization of the associated approximation space is proved without any restrictions given by the number of vanishing moments, but now the canonical expansion is replaced with another linear expansion.

1. INTRODUCTION

Given a finite collection of functions $\Psi = \{\psi^1, \psi^2, \dots, \psi^L\} \subset L_2(\mathbb{R}^d)$ we use the notation $X(\Psi)$ to denote the corresponding "wavelet" system,

$$X(\Psi) := \left\{2^{jd/2}\psi^\ell(2^j \cdot -k) \,|\, j \in \mathbb{Z}, k \in \mathbb{Z}^d, \ell = 1, 2, \dots, L\right\}.$$

A **wavelet bi-frame for** $L_2(\mathbb{R}^d)$ consists of two sequences of wavelets

$$\Psi = \{\psi^1, \psi^2, \dots, \psi^L\} \subset L_2(\mathbb{R}^d) \quad \text{and} \quad \tilde{\Psi} = \{\tilde{\psi}^1, \tilde{\psi}^2, \dots, \tilde{\psi}^L\} \subset L_2(\mathbb{R}^d)$$

for which the systems $X(\Psi)$ and $X(\tilde{\Psi})$ are Bessel systems, and satisfy the perfect reconstruction formula

$$f = \sum_{\ell=1}^{L} \sum_{j\in\mathbb{Z}, k\in\mathbb{Z}^d} \langle f, \tilde{\psi}^\ell_{j,k}\rangle \psi^\ell_{j,k}, \qquad \forall f \in L_2(\mathbb{R}^d), \tag{1.1}$$

where

$$\psi_{j,k} := 2^{jd/2}\psi(2^j \cdot -k), \qquad j \in \mathbb{Z},\ k \in \mathbb{Z}^d.$$

This definition implies that both $X(\Psi)$ and $X(\tilde{\Psi})$ are frames for $L_2(\mathbb{R}^d)$ and in fact the roles of Ψ and $\tilde{\Psi}$ are interchangeable in (1.1). The special case with $\Psi = \tilde{\Psi}$ corresponds to a so-called tight wavelet frame.

2000 *Mathematics Subject Classification.* Primary 41A17; Secondary 41A15, 41A46.

Key words and phrases. Wavelet bi-frames, framelets, Besov spaces, nonlinear approximation, Jackson inequality, Bernstein inequality, redundant dictionary.

Research supported by the Danish Technical Science Foundation, Grant no. 9701481.

[‡]Partially supported by the European Union's Human Potential Programme, under contract HPRN-CT-2002-00285 (HASSIP).

[†]Partially supported by the Danish SNF-PDE network.

The most common method to construct wavelet bi-frames relies on so-called extension principles. The resulting bi-frames are based on a multiresolution analysis, and the generators are often called bi-framelets. The construction of multiresolution-based wavelet frames have been studied extensively, see *e.g.* [6, 8, 22, 23, 24].

In this paper we study the nonlinear approximation properties of the systems $X(\Psi)$ and $X(\tilde{\Psi})$ when the approximation error is measured in $L_p(\mathbb{R}^d)$. That is, we consider the (nonlinear) set

$$\Sigma_m(X(\Psi)) := \Big\{ \sum_{i\in\Lambda} c_i g_i \;\Big|\; c_i \in \mathbb{C}, g_i \in X(\Psi), \operatorname{card}\Lambda \le m \Big\}$$

of all possible m-term expansions with elements from the system $X(\Psi)$. The error of the best m-term approximation to an element $f \in L_p(\mathbb{R}^d)$ is then $\sigma_m(f, X(\Psi))_p := \inf_{f_m\in\Sigma_m(X(\Psi))} \|f - f_m\|_{L_p(\mathbb{R}^d)}$. A first problem (Problem 1) is to characterize the class $\mathcal{A}^\alpha$ of functions $f \in L_p(\mathbb{R}^d)$ for which $\sigma_m(f, X(\Psi))_p = O(m^{-\alpha})$. A related problem (Problem 2) is to characterize the class $\mathcal{K}_\tau$ of functions f with expansions in $X(\Psi)$ that are sparse in the sense that the ℓ_τ norm of the expansion coefficients is finite.

In the special case where $X(\Psi)$ is indeed a (bi)orthogonal wavelet basis in $L^2(\mathbb{R}^d)$, it is well known that when $\alpha = 1/\tau - 1/p$, the two (families of) classes $\mathcal{A}^\alpha$ and $\mathcal{K}_\tau$ essentially coincide and are indeed Besov spaces $B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))$. The decay, smoothness, and number of vanishing moments of the wavelet determine the range of values of τ for which the characterization is true. We refer the reader to, *e.g.* [25, 2] for a definition and some properties of the Besov spaces. When $X(\Psi)$ is a redundant frame, since the perfect reconstruction formula (1.1) only provides one among an infinite number of possible expansions of f in the system $X(\Psi)$, one must a priori consider separately two families of "sparseness classes" depending whether sparseness is measured in terms of synthesis or analysis coefficients.

Some results on approximation with tight wavelet frames were obtained by the authors in [3], but the results in the present paper are more general even for the tight wavelet case. For oversampled tight wavelet frames based on splines, the approximation spaces associated with best m-term approximation in $L_p(\mathbb{R}^d)$ were completely characterized by two of the present authors in [15].

We address Problem 1 in Section 5 and 6, where it is proved that $\sigma_m(f, X(\Psi))_p = O(m^{-\alpha})$ if and only if f belongs to (essentially) a Besov space. Put another way, the wavelet bi-frame system completely characterize Besov spaces through the quantities $\sigma_m(f, X(\Psi))_p$. This characterization holds for smoothness parameters α in a certain range limited by the number of vanishing moments of the dual frame $X(\tilde{\Psi})$. For a univariate orthonormal wavelet system, smoothness and decay automatically imply a sufficient number of vanishing moments, see e.g. [16, Theorem 3.4], but this is no longer true for wavelet bi-frame systems.We use the standard approach to handle the characterization problem, deriving so-called Bernstein and Jackson estimates for the system $X(\Psi)$.

A nice corollary of the estimates in Section 5 and 6 is that smooth functions (in the Besov sense) have very sparse canonical expansions (1.1) in the bi-frame system. This addresses Problem 2 and, again, this is true for functions with smoothness α in a certain range given by the number of vanishing moments of the analysis system $X(\tilde{\Psi})$.

The set of approximation spaces we are able to characterize in Section 6 is limited by the number of vanishing moments of the functions in the dual frame $\tilde{\Psi}$. In fact, we show in Section 3 that this limit cannot be improved in the general case. In Section 7, we consider approximation with an *oversampled* version of the wavelet bi-frame dictionary $X(\Psi)$ in dimension $d = 1$. For such oversampled systems we prove that there is a Jackson inequality independent of the number of vanishing moments of the functions in Ψ. This leads to a complete characterization of the approximation spaces for oversampled wavelet bi-frame systems.

The structure of the paper is as follows. In Section 2 we review the most common method to construct wavelet bi-frames using the so-called extension principles. Section 3 is devoted to proving boundedness for certain matrix operators that will be needed to obtain Jackson inequalities for nonlinear approximation with wavelet bi-frame systems.

In Section 4 we give some elements of (nonlinear) approximation theory: we define approximation spaces for a general dictionary and state the corresponding Jackson and Bernstein inequalities. In Section 5 we prove a general Bernstein inequality for wavelet bi-frame systems with compact support.

The matrix lemmas from Section 3 are used in Section 6 where we consider Jackson inequalities for best m-term approximation with bi-frame systems in $L_p(\mathbb{R}^d)$, and we discuss the fairly general case where a complete characterization of the approximation spaces –associated with best m-term approximation in $L_p(\mathbb{R}^d)$ with bi-frame systems– in terms of (essentially) Besov spaces is possible.

In the final section of the paper, Section 7, we prove a characterization of the approximation spaces for oversampled wavelet bi-frame systems, again in terms of (essentially) Besov spaces.

Appendix A contains the analysis of the stability properties of wavelet bi-frame expansions in $L_p(\mathbb{R}^d)$. We give a complete characterization of the $L_p(\mathbb{R}^d)$-norm, $1 < p < \infty$, in terms of analysis coefficients associated with the frame, and prove that the bi-frame expansion gives an atomic decomposition for $L_p(\mathbb{R}^d)$. The characterization has the same form as the classical characterization of the $L_p(\mathbb{R}^d)$-norm by wavelet coefficients, see *e.g.* [20]. The main use of the results in this appendix is as a tool to derive a Jackson inequality for wavelet bi-frame systems.

2. MRA based wavelet bi-frame systems

In this section we will briefly describe how to construct MRA-based wavelet bi-frames -called **bi-framelets**- through so-called extension principles. The extension principles to construct bi-frames were introduced independently in [5] and [8]. We refer the reader to either [5] or [8] for a more detailed discussion of MRA based bi-frames. Below we use the notation of [8].

Let $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_L)$ and $\tilde{\boldsymbol{\tau}} = (\tilde{\tau}_0, \tilde{\tau}_1, \ldots, \tilde{\tau}_L)$ be two sequences of $2\pi\mathbb{Z}^d$-periodic essentially bounded functions. Assume that τ_0 and $\tilde{\tau}_0$ both generate refinable functions

$$\hat{\phi}(2\xi) = \tau_0(\xi)\hat{\phi}(\xi) \qquad \text{and} \qquad \hat{\tilde{\phi}}(2\xi) = \tilde{\tau}_0(\xi)\hat{\tilde{\phi}}(\xi),$$

satisfying

$$\lim_{\xi\to 0}\hat{\phi}(\xi)=1 \quad \text{and} \quad \lim_{\xi\to 0}\hat{\tilde{\phi}}(\xi)=1,$$

with

$$\operatorname{ess\,sup}_\xi \sum_{k\in\mathbb{Z}^d}|\hat{\phi}(\xi-k)|^2<\infty \quad \text{and} \quad \operatorname{ess\,sup}_\xi \sum_{k\in\mathbb{Z}^d}|\hat{\tilde{\phi}}(\xi-k)|^2<\infty,$$

where $\hat{\phi}(\xi)$ is the Fourier transform of the function $\phi(x)$. We associate the wavelets to $\boldsymbol{\tau}$ and $\tilde{\boldsymbol{\tau}}$ as follows

$$\hat{\psi}^\ell(2\xi)=\tau_\ell(\xi)\hat{\phi}(\xi), \qquad \hat{\tilde{\psi}}^\ell(2\xi)=\tilde{\tau}_\ell(\xi)\hat{\tilde{\phi}}(\xi).$$

The spectrum $\sigma(\phi)$ associated to ϕ is defined up to a null-set as

$$\sigma(\phi):=\{\omega\in[-\pi,\pi]^d : \hat{\phi}(\omega+2\pi k)\neq 0, \text{ for some } k\in\mathbb{Z}^d\}.$$

The spectrum $\sigma(\tilde{\phi})$ associated to $\tilde{\phi}$ is defined likewise. Assuming that the systems $X(\Psi)$ and $X(\tilde{\Psi})$ are both Bessel, we define the **mixed fundamental function of the parent vectors $\boldsymbol{\tau}$ and $\tilde{\boldsymbol{\tau}}$** by

$$\Theta(\xi):=\sum_{j=0}^{\infty}\sum_{\ell=1}^{L}\tau_\ell(2^j\xi)\overline{\tilde{\tau}_\ell(2^j\xi)}\prod_{m=0}^{j-1}\tau_0(2^m\xi)\overline{\tilde{\tau}_0(2^m\xi)}.$$

The following theorem proved in [8] is the main tool to create bi-framelet systems, the theorem is called the **Mixed Oblique Extension Principle**.

Theorem 2.1 (Mixed OEP)**.** *Let $\boldsymbol{\tau}$ and $\tilde{\boldsymbol{\tau}}$ be the combined mask of the systems $X(\Psi)$ and $X(\tilde{\Psi})$, respectively. Assume that the systems $X(\Psi)$ and $X(\tilde{\Psi})$ are Bessel systems. Suppose there exists a 2π-periodic function Θ satisfying*

a) *Θ is essentially bounded, continuous at the origin, and $\Theta(0)=1$.*
b) *If $\xi\in\sigma(\phi)\cap\sigma(\tilde{\phi})$ and $\nu\in\{0,\pi\}^d$ such that $\xi+\nu\in\sigma(\phi)\cap\sigma(\tilde{\phi})$, then*

$$\Theta(2\xi)\tau_0(\xi)\overline{\tilde{\tau}(\xi+\nu)}+\sum_{\ell=1}^{L}\tau_\ell(\xi)\overline{\tilde{\tau}_\ell(\xi+\nu)}=\begin{cases}\Theta(\xi), & \text{if } \nu=0\\ 0, & \text{otherwise.}\end{cases}$$

Then $X(\Psi)$, $X(\tilde{\Psi})$ is a bi-framelet system.

Remark 2.2. In many (most) interesting cases the spectra $\sigma(\phi)$ and $\sigma(\tilde{\phi})$ are both equal to $[-\pi,\pi]^d$. For example, if the integer translates of the scaling functions ϕ and $\tilde{\phi}$ are Riesz sequences, this is the case.

When $X(\Psi)=X(\tilde{\Psi})$, Theorem 2.1 gives the so-called Oblique Extension principle, see [8]. If, in addition, $\Theta\equiv 1$, Theorem 2.1 reduces to the Unitary Extension Principle, see [23, 24].

The reader can consult [5] and [8] for many explicit examples on how to construct framelet systems using the different extension principles.

3. Some matrix lemmas

This section contains mainly technical matrix lemmas that will be used to derive Jackson inequalities for wavelet bi-frame systems in Section 6. Given two functions ψ and η with a prescribed number of vanishing moments and degree of smoothness, we study the range of $\tau \in (0,\infty)$ for which the infinite matrix $[\langle \psi^{p'}_{j',k'}, \eta^p_{j,k}\rangle]_{j\in\mathbb{Z},k\in\mathbb{Z}^d}$ defines a bounded operator on ℓ_τ, where $p\in(1,\infty)$, $1 = 1/p + 1/p'$ and

$$\psi^p_{j,k} := 2^{jd(1/p-1/2)}\psi_{j,k}. \tag{3.1}$$

Notice that $\|\psi^p_{j,k}\|_{L_p(\mathbb{R}^d)} = \|\psi\|_{L_p(\mathbb{R}^d)}$ for all j and k. We are especially interested in the case where ψ is a function from a bi-frame system with some (possibly few) vanishing moments and η is a "nice" wavelet.

In order to prove the main result of the section (Proposition 3.5) we need the following technical results.

Lemma 3.1. *Given $p\in[1,\infty)$, $\gamma > d$ and positive constants c_1, c_2, let*

$$b_m = b_m(c_1,c_2) := \begin{cases} 2^{-mdc_1} & \text{for } m \ge 0,\\ 2^{mdc_2} & \text{for } m < 0.\end{cases}$$

Denoting $|x|$ the Euclidean norm of $x\in\mathbb{R}^d$, for $k,k'\in\mathbb{Z}^d$ let

$$a_{m;k,k'} = a_{m;k,k'}(\gamma) := \begin{cases} 2^{-md/p}(1+|k-2^{-m}k'|)^{-\gamma} & \text{for } m\ge 0,\\ 2^{md/p'}(1+|2^m k - k'|)^{-\gamma} & \text{for } m<0,\end{cases}$$

where $1/p' = 1-1/p$. Suppose $p\,(1-c_1) < \tau < p\,(1+c_2)$ and $\tau\ge 1$. Then

$$\sum_{m\in\mathbb{Z}}\sum_{j'\in\mathbb{Z}}\sum_{k'\in\mathbb{Z}^d} b_m\Big(\sum_{k\in\mathbb{Z}^d} a_{m;k,k'}|c_{j'-m,k}|\Big)^\tau \le C\|\{c_{j,k}\}\|^\tau_{\ell_\tau}.$$

Proof. Lemma 8.10 in [21] implies for any $\{d_k\}_k\in\ell_\tau$, $1\le\tau<\infty$,

$$\sum_{k'\in\mathbb{Z}^d}\Big(\sum_{k\in\mathbb{Z}^d} a_{m;k,k'}|d_k|\Big)^\tau \le C2^{md(1-\tau/p)}\sum_{k\in\mathbb{Z}^d}|d_k|^\tau, \qquad \text{for } m\in\mathbb{Z}.$$

The use of Lemma 8.10 in [21] is direct for $m<0$ and a duality argument is used to get the estimate for $m\ge 0$. This estimate yields

$$\begin{aligned}\sum_{m\in\mathbb{Z}} b_m \sum_{j'\in\mathbb{Z}}\sum_{k'\in\mathbb{Z}^d}\Big(\sum_{k\in\mathbb{Z}^d} a_{m;k,k'}|c_{j'-m,k}|\Big)^\tau &\\ \le C\sum_{m\in\mathbb{Z}} b_m 2^{md(1-\tau/p)}\sum_{j'\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^d}|c_{j'-m,k}|^\tau &\\ = C\Big(\sum_{m\in\mathbb{Z}} b_m(\delta_1,\delta_2)\Big)\|\{c_{j,k}\}\|^\tau_{\ell_\tau}, &\end{aligned}$$

where $\delta_1 := c_1 - (1-\tau/p)$ and $\delta_2 := c_2 + (1-\tau/p)$. Now, the lemma follows since $(1-c_1) < \tau/p < (1+c_2)$ implies $\delta_1 > 0$ and $\delta_2 > 0$. $\square$

With this lemma at hand we can now characterize a range of τ for which the matrix $[b_m a_{j'-j;k,k'}]_{j,j'\in\mathbb{Z},k,k'\in\mathbb{Z}^d}$ defines a bounded operator on ℓ_τ.

Lemma 3.2. *Given $p \in [1,\infty)$, $\gamma > d$ and positive constants c_1, c_2, let b_m and $a_{m;k,k'}$ be as in Lemma 3.1. Then*

$$\sum_{j'\in\mathbb{Z},k'\in\mathbb{Z}^d}\Big(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} b_{j'-j}a_{j'-j;k,k'}|c_{j,k}|\Big)^\tau \le C\|\{c_{j,k}\}\|_{\ell_\tau}^\tau$$

for any τ in the range

$$\Lambda(c_1) < \tau < p(1+c_2),$$

where

$$\Lambda(x) = \Lambda(x,p,\gamma/d) := \begin{cases} p(1-x) & \text{for } x \le 1-1/p, \\ (x+1/p)^{-1} & \text{for } 1-1/p < x \le \gamma/d - 1/p, \\ d/\gamma & \text{for } \gamma/d - 1/p < x. \end{cases} \tag{3.2}$$

Proof. For $\tau \ge 1$, Hölders inequality (for the sum over j), with $1 = 1/\tau + 1/\tau'$ yields

$$\begin{aligned}\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} b_{j'-j}a_{j'-j;k,k'}|c_{j,k}| &= \sum_{j\in\mathbb{Z}} b_{j'-j}^{1/\tau'+1/\tau}\Big(\sum_{k\in\mathbb{Z}^d} a_{j'-j;k,k'}|c_{j,k}|\Big) \\ &\le \Big(\sum_{j\in\mathbb{Z}} b_{j'-j}\Big)^{1/\tau'}\Big(\sum_{j\in\mathbb{Z}} b_{j'-j}\Big(\sum_{k\in\mathbb{Z}^d} a_{j'-j;k,k'}|c_{j,k}|\Big)^\tau\Big)^{1/\tau} \\ &\le C\Big(\sum_{m\in\mathbb{Z}} b_m\Big(\sum_{k\in\mathbb{Z}^d} a_{m;k,k'}|c_{j'-m,k}|\Big)^\tau\Big)^{1/\tau}.\end{aligned}$$

The result then follows using Lemma 3.1, provided that $\max(1, p(1-c_1)) < \tau < p(1+c_2)$. For $\tau < 1$ we have

$$\begin{aligned}\sum_{j'\in\mathbb{Z},k'\in\mathbb{Z}^d}\Big(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} b_{j'-j}a_{j'-j;k,k'}|c_{j,k}|\Big)^\tau &\le \sum_{j'\in\mathbb{Z},k'\in\mathbb{Z}^d}\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} b_{j'-j}^\tau a_{j'-j;k,k'}^\tau |c_{j,k}|^\tau \\ &= \sum_{m\in\mathbb{Z}}\sum_{j'\in\mathbb{Z}}\sum_{k'\in\mathbb{Z}^d} b_m^\tau\Big(\sum_{k\in\mathbb{Z}^d} a_{m;k,k'}^\tau|c_{j,k}|^\tau\Big) \\ &= \sum_{m\in\mathbb{Z}}\sum_{j'\in\mathbb{Z}}\sum_{k'\in\mathbb{Z}^d} \tilde{b}_m\Big(\sum_{k\in\mathbb{Z}^d} \tilde{a}_{m;k,k'}|c_{j,k}|^\tau\Big).\end{aligned}$$

In the last line we have used the fact that

$$b_m^\tau a_{m;k,k'}^\tau = b_m(c_1,c_2)^\tau a_{m;k,k'}^\tau(\gamma) = b_m(\tilde{c}_1,\tilde{c}_2)a_{m;k,k'}(\tilde{\gamma}) =: \tilde{b}_m\tilde{a}_{m;k,k'}$$

with $\tilde{c}_1 = c_1\tau - 1/p(1-\tau)$, $\tilde{c}_2 = c_2\tau - 1/p'(1-\tau)$ and $\tilde{\gamma} = \gamma\tau$. Notice that we have the same structure as in Lemma 3.1. Thus, using Lemma 3.1 with $\{d_{j,k}\} = \{|c_{j,k}|^\tau\} \in \ell^1$ we obtain that the result holds provided that $p(1-\tilde{c_1}) < 1 < p(1+\tilde{c}_2)$ and $\tilde{\gamma} > d$, which is equivalent to

$$\tau > \max(d/\gamma, (c_1+1/p)^{-1}). \tag{3.3}$$

Now combining the results for $\tau < 1$ and $\tau \ge 1$ we see that if $c_1 + 1/p \le 1$, the set of $\tau < 1$ that satisfy (3.3) is empty, and as $\max(1, p(1-c_1)) = p(1-c_1)$ we get the first part of the result. If $c_1 + 1/p > 1$, we get the rest. □

Definition 3.3. For $N \in \mathbb{N}$ and $\gamma > 0$ we let $D_\gamma^N(\mathbb{R}^d)$ be the set of all functions f defined on $\mathbb{R}^d$ with N derivatives and decay γ, *i.e.*, for which there exists a constant $c < \infty$ such that

$$|\partial^\alpha f(x)| \le c(1+|x|)^{-\gamma} \qquad \text{for } x \in \mathbb{R}^d,\ \alpha \in \mathbb{N}^d,\ |\alpha| \le N, \tag{3.4}$$

where $|\alpha|$ is the usual length of a multi-index. Likewise, we let $M_\gamma^N(\mathbb{R}^d)$ denote the set of all functions f with N vanishing moments and decay, *i.e.*, for which

$$\int x^\alpha f(x)dx = 0 \qquad \text{for } \alpha \in \mathbb{N}^d,\ |\alpha| < N,$$

and

$$|f(x)| \le C(1+|x|)^{-d-N-\gamma} \qquad \text{for } x \in \mathbb{R}^d. \tag{3.5}$$

We need the following lemma in order to prove Proposition 3.5.

Lemma 3.4. *Consider $N_1, N_2 \in \mathbb{N}$ and $\gamma_1, \gamma_2 > 0$. Suppose $\eta \in D_{\gamma_1}^{N_1}(\mathbb{R}^d) \cap M_{\gamma_2}^{N_2}(\mathbb{R}^d)$ and $\psi \in D_{\gamma_2}^{N_2}(\mathbb{R}^d) \cap M_{\gamma_1}^{N_1}(\mathbb{R}^d)$. Then, for $j' \ge j$, we have*

$$|(\eta_{j,k} \star \psi_{j',k'})(x)| \le \frac{C 2^{-(j'-j)(N_1+d/2)}}{(1+|2^j x - 2^{j-j'}k' - k|)^{\gamma_1}} \tag{3.6}$$

and for $j' \le j$

$$|(\eta_{j,k} \star \psi_{j',k'})(x)| \le \frac{C 2^{-(j-j')(N_2+d/2)}}{(1+|2^{j'} x - k' - 2^{j'-j}k|)^{\gamma_2}} \tag{3.7}$$

where $\star$ denotes the convolution product.

The result is well known (see *e.g.* Lemma 3.3 in [13]) but will be shown here for the sake of completeness.

Proof. We will simply prove Eq. (3.6), since interchanging the role of ψ and η will then provide Eq. (3.7). We notice that $(\eta_{j,k} \star \psi_{j',k'})(x) = (\eta \star \psi_{j'-j,0})(2^j x - 2^{j-j'}k' - k)$. Thus, it suffices to prove Eq. (3.6) for $j = k = k' = 0$ and $m := j' - j \ge 0$. Since ψ has N_1 vanishing moments, a Taylor expansion of η gives

$$|\eta \star \psi_{m,0}(x)| \le \left(\int_{|y-x|\le|x|/2} + \int_{|y-x|>|x|/2} \right) |\psi_{m,0}(x-y)||x-y|^{N_1} E(x,y)\, dy = A + B,$$

where $E(x,y) := \sup_{|\beta|=N_1} \sup_{\varepsilon\in(0,1)} |\partial^\beta \eta(x+\varepsilon(y-x))|/\beta!$ (with $\beta!$ the standard factorial for a multi-index $\beta \in \mathbb{N}^d$). By (3.4) we have $|E(x,y)| \le c(1+|x|)^{-\gamma_1}$ when $|y-x| \le |x|/2$. This estimate together with (3.5) applied on ψ, yield

$$\begin{aligned} A &\le C \int |\psi_{m,0}(x-y)||x-y|^{N_1}\, dy \cdot (1+|x|)^{-\gamma_1} \\ &\le C\ 2^{md/2} \int_{|y-x|\le|x|/2} (1+2^m|y|)^{-(N_1+d+\gamma_1)} |y|^{N_1}\, dy \cdot (1+|x|)^{-\gamma_1} \\ &\le C\ 2^{md/2}\ 2^{-m(N_1+d)} (1+|x|)^{-\gamma_1}. \end{aligned}$$

Moreover, since E is bounded, we have

$$\begin{aligned}
B &\le C \int_{|y-x|>|x|/2} |\psi_{m,0}(x-y)||x-y|^{N_1}\,dy \\
&\le C\, 2^{md/2} \int_{|y-x|>|x|/2} (1+2^m|x-y|)^{-(N_1+d+\gamma_1)}|x-y|^{N_1}\,dy \\
&\le C\, 2^{md/2}\, 2^{-m(N_1+d)}(1+|x|)^{-\gamma_1}.
\end{aligned}$$

This proves (3.6). □

We can now state the main result of this section.

Proposition 3.5. *Suppose $\eta \in D_\gamma^{N_1}(\mathbb{R}^d) \cap M_\gamma^{N_2}(\mathbb{R}^d)$ and $\psi \in D_\gamma^{N_2}(\mathbb{R}^d) \cap M_\gamma^{N_1}(\mathbb{R}^d)$ for some $N_1, N_2 \in \mathbb{N}$ and $\gamma > d$. Consider the matrix operator $\mathbf{T}$ given by*

$$(\mathbf{T}(c_{j,k}))_{j',k'} := \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} c_{j,k}\langle \eta^p_{j,k}, \psi^{p'}_{j',k'}\rangle \tag{3.8}$$

for $p \in [1,\infty)$ and $1 = 1/p + 1/p'$. Then $\mathbf{T}$ is bounded on $\ell_\tau(\mathbb{Z}\times\mathbb{Z}^d)$ for any τ in the range

$$\Lambda\left(\frac{N_1}{d}\right) < \tau < p\left(1+\frac{N_2}{d}\right), \tag{3.9}$$

with $\Lambda(x) = \Lambda(x,p,d/\gamma)$ given by (3.2).

Proof. By Lemma 3.4,

$$|\langle \eta_{j,k}, \psi_{j',k'}\rangle| \le \begin{cases} \dfrac{C2^{(j-j')(d/2+N_1)}}{(1+|k-2^{j-j'}k'|)^\gamma} & \text{for } j' \ge j, \\[2ex] \dfrac{C2^{(j'-j)(d/2+N_2)}}{(1+|k'-2^{j'-j}k|)^\gamma} & \text{for } j' \le j. \end{cases}$$

Recalling the definition of $\psi^p_{j,k}$ and $\eta^p_{j,k}$ in Eq. (3.1), this gives for $j' \ge j$,

$$\begin{aligned}
|\langle \eta^p_{j,k}, \psi^{p'}_{j',k'}\rangle| &\le C2^{jd(1/p-1/2)}2^{j'd(1/p'-1/2)}\frac{2^{(j-j')(d/2+N_1)}}{(1+|k-2^{j-j'}k'|)^\gamma} \\
&= C2^{j(d/p-d/2+d/2)}2^{-j'(-d/p'+d/2+d/2)}\frac{2^{(j-j')N_1}}{(1+|k-2^{j-j'}k'|)^\gamma} \\
&= C\frac{2^{(j-j')(d/p+N_1)}}{(1+|k-2^{j-j'}k'|)^\gamma}.
\end{aligned}$$

A similar argument shows that for $j > j'$,

$$|\langle \eta^p_{j,k}, \psi^{p'}_{j',k'}\rangle| \le C\frac{2^{(j'-j)(d/p'+N_2)}}{(1+|k'-2^{j'-j}k|)^\gamma}.$$

We can thus write

$$|\langle \eta^p_{j,k}, \psi^{p'}_{j',k'}\rangle| \le b_{j'-j}(N_1/d, N_2/d)\; a_{j'-j;k,k'}(\gamma) \tag{3.10}$$

The result now follows using Lemma 3.2. □

It is the lower bound in (3.9), $\Lambda(N_1/d) < \tau$, that will be most important to us in the subsequent sections. The following example shows that this bound cannot be improved in the general case.

Proposition 3.6. *Suppose $\psi \in L_2(\mathbb{R})$ has fast decay and exactly $N_1 \in \mathbb{N}$ vanishing moments, i.e., $\psi \in M_\gamma^{N_1}(\mathbb{R})$ but $\psi \notin M_\gamma^{N_1+1}(\mathbb{R})$ for any $\gamma > 1$, and suppose $\eta \in C^\infty(\mathbb{R})$ is compactly supported. Consider the matrix operator $\mathbf{T}$ given by (3.8) in Proposition 3.5 for some $p \in [1,\infty)$. Then $\mathbf{T}$ is not bounded on $\ell_\tau(\mathbb{Z}\times\mathbb{Z})$ for any $\tau \le (N_1+1/p)^{-1}$.*

Proof. It suffices to prove that $\sum_{j,k} |\langle \eta, \psi_{j,k}^{p'}\rangle|^\tau = \infty$ for $\tau \le 1/(N_1+1/p)$. Since ψ has exactly N_1 vanishing moments and fast decay, there exists a function θ with fast decay, such that $\psi = (-1)^{N_1}\theta^{(N_1)}$ and $\int_{\mathbb{R}} \theta(x)\,dx \neq 0$, see e.g. [19, Theorem 6.2]. Observe that

$$\begin{aligned}\langle \eta, \psi_{j,k}\rangle &= \int_{\mathbb{R}} \eta(x) 2^{j/2}\psi(2^j x - k)\,dx \\ &= 2^{-jN_1}(-1)^{N_1}\int_{\mathbb{R}} \eta(x)2^{j/2}\theta^{(N_1)}(2^j x - k)\,dx \\ &= 2^{-jN_1}\int_{\mathbb{R}} \eta^{(N_1)}(x)2^{j/2}\theta(2^j x - k)\,dx \\ &= 2^{-j(N_1+1/2)}\big(\eta^{(N_1)} \star w_{2^{-j}}\big)(2^{-j}k),\end{aligned}$$

where $w_\varepsilon(x) := \varepsilon^{-1}\theta(-x/\varepsilon)$. Since $\eta \in C^\infty(\mathbb{R})$ has compact support, $\lim_{\varepsilon\to 0}\big(\eta^{(N_1)} \star w_\varepsilon\big)(x) = a\cdot\eta^{(N_1)}(x)$ uniformly, where $a = \int_{\mathbb{R}}\theta(x)\,dx$, see e.g. [12]. Moreover, if we define the Riemann sums $s_\varepsilon := \varepsilon\sum_{k\in\mathbb{Z}} |\eta^{(N_1)}(\varepsilon k)|^\tau$, $\tau \in (0,\infty)$, we have $\lim_{\varepsilon\to 0} s_\varepsilon = \|\eta^{(N_1)}\|_\tau^\tau$. Combining these two properties, there exists $\varepsilon_0 > 0$ such that

$$\varepsilon\sum_{k\in\mathbb{Z}} |\big(\eta^{(N_1)}\star w_\varepsilon\big)(\varepsilon k)|^\tau \ge c\|\eta^{(N_1)}\|_\tau^\tau \qquad \text{for all } \varepsilon \in (0,\varepsilon_0).$$

Thus, since $\psi_{j,k}^{p'} = 2^{j(1/p'-1/2)}\psi_{j,k}$, there exists $j_0 > 0$ such that for all $j > j_0$,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} |\langle \eta, \psi_{j,k}^{p'}\rangle|^\tau &= 2^{-j(N_1+1-1/p')\tau}\sum_{k\in\mathbb{Z}} |\big(\eta^{(N_1)}\star w_{2^{-j}}\big)(2^{-j}k)|^\tau \\ &\ge c2^{-j(N_1+1/p)\tau}2^j\|\eta^{(N_1)}\|_\tau^\tau.\end{aligned}$$

Now, in order to get $\sum_{j,k}|\langle \eta, \psi_{j,k}^{p'}\rangle|^\tau < \infty$ we need to have $-(N_1+1/p)\tau + 1 < 0$, that is to say $\tau > 1/(N_1+1/p)$. □

4. Approximation spaces

The remaining sections contain the core of the paper. We begin by introducing some notions of approximation theory that will be used throughout the rest of the paper. In Section 5 we prove a Bernstein inequality that provides "one half" of the characterization of m-term approximation with "nice" wavelet bi-frame systems. In Section 6 we will consider Jackson inequalities that match the Bernstein inequalities in order to get complete characterizations.

A dictionary $\mathcal{D} = \{g_k\}_{k\in\mathbb{N}}$ in $L_p(\mathbb{R}^d)$ is a countable collection of quasi-normalized elements from $L_p(\mathbb{R}^d)$. For $\mathcal{D}$ we consider the collection of all possible m-term expansions with elements from $\mathcal{D}$:

$$\Sigma_m(\mathcal{D}) := \Big\{ \sum_{i\in\Lambda} c_i g_i \,\Big|\, c_i \in \mathbb{C}, \operatorname{card}\Lambda \le m \Big\}.$$

The error of the best m-term approximation to an element $f \in L_p(\mathbb{R}^d)$ is then

$$\sigma_m(f, \mathcal{D})_p := \inf_{f_m\in\Sigma_m(\mathcal{D})} \|f - f_m\|_{L_p(\mathbb{R}^d)}.$$

Definition 4.1 (Approximation spaces)**.** The approximation space $\mathcal{A}_q^\alpha(L_p(\mathbb{R}^d), \mathcal{D})$ is defined by

$$|f|_{\mathcal{A}_q^\alpha(L_p(\mathbb{R}^d),\mathcal{D})} := \Big(\sum_{m=1}^{\infty} \big(m^\alpha \sigma_m(f, \mathcal{D})_p\big)^q \frac{1}{m} \Big)^{1/q} < \infty,$$

and (quasi)normed by $\|f\|_{\mathcal{A}_q^\alpha(L_p(\mathbb{R}^d),\mathcal{D})} = \|f\|_p + |f|_{\mathcal{A}_q^\alpha(L_p(\mathbb{R}^d),\mathcal{D})}$, for $0 < q, \alpha < \infty$. When $q = \infty$, the ℓ_q norm is replaced by the sup-norm.

In what follows, we use the notation $V \hookrightarrow W$ to indicate V is continuously embedded in W for two (quasi)normed spaces V and W, i.e., $V \subset W$ and there is a constant $C < \infty$ such that $\|\cdot\|_W \le C\|\cdot\|_V$.

It is well known that the main tool in the characterization of $\mathcal{A}_q^\alpha(L_p(\mathbb{R}^d), \mathcal{D})$ comes from the link between approximation theory and interpolation theory (see *e.g.* [11, Theorem 9.1, Chapter 7]). Let $Y_p(\mathbb{R}^d)$ be a (quasi)Banach space continuously embedded in $L_p(\mathbb{R}^d)$ with semi-(quasi)norm $|\cdot|_{Y_p}$. Given $\alpha > 0$, the Jackson inequality

$$\sigma_m(f, \mathcal{D})_p \le Cm^{-\alpha}|f|_{Y_p(\mathbb{R}^d)}, \qquad \forall f \in Y_p(\mathbb{R}^d),\ \forall m \in \mathbb{N} \tag{4.1}$$

and the Bernstein inequality

$$|S|_{Y_p(\mathbb{R}^d)} \le C' m^\alpha \|S\|_p, \qquad \forall S \in \Sigma_m(\mathcal{D}) \tag{4.2}$$

(with some constants C and C' independent of f, S and m) imply, respectively, the continuous embedding

$$\big(L_p(\mathbb{R}^d), Y_p(\mathbb{R}^d)\big)_{\beta/\alpha,q} \hookrightarrow \mathcal{A}_q^\beta(L_p(\mathbb{R}^d), \mathcal{D})$$

and the converse embedding

$$\big(L_p(\mathbb{R}^d), Y_p(\mathbb{R}^d)\big)_{\beta/\alpha,q} \hookleftarrow \mathcal{A}_q^\beta(L_p(\mathbb{R}^d), \mathcal{D})$$

for all $0 < \beta < \alpha$ and $q \in (0, \infty]$. Here $(X, Y)_{\theta,q}$ denotes the real interpolation space between the Banach spaces X and Y. We refer the reader to [1] for the definition of the real interpolation method.

Thus, to get a complete characterization of the approximation space $\mathcal{A}_q^\alpha(L_p(\mathbb{R}^d), X(\Psi))$, we need to prove associated Jackson and Bernstein inequalities. In Section 5, a Bernstein inequality for an MRA based bi-frame system is proved. In Section 6, a Jackson inequality is proved for a general bi-frame system under mild assumptions on the smoothness, decay and number of vanishing moments of its generators. The proof relies on the matrix lemmas proved in Section 3 and provides, at the same time, a characterization of Besov spaces in terms of the ℓ_τ norm of the dual (properly re-normalized) frame coefficients that appear in Eq. (1.1). The set of approximation spaces we are able to characterize in Section 6 is limited by the number of vanishing moments of the functions in the dual frame $\tilde{\Psi}$. In Section 7, Jackson and Bernstein inequalities for a twice oversampled MRA based bi-frame system are considered. For such a system the Jackson inequalities are, in fact, independent of the number of vanishing moments of the functions in $\tilde{\Psi}$. This leads to a complete characterization of the approximation spaces for oversampled framelet systems without any assumptions on the number of vanishing moments.

5. Bernstein estimates for the bi-frame system

Let us begin by proving a Bernstein inequality for bi-framelet systems, since such an inequality will show us "how big" $\mathcal{A}_q^\beta(L_p(\mathbb{R}^d), X(\Psi))$ can possibly be. In the following we denote by $W^s(L_\infty(\mathbb{R}^d))$ the Sobolev space consisting of functions with s distributional derivatives in $L_\infty(\mathbb{R}^d)$. Given a function $\phi \in L_\infty(\mathbb{R}^d)$, let

$$\Gamma = \{k \in \mathbb{Z}^d \colon |\{x \in (0,1)^d \colon \phi(x-k) \neq 0\}| > 0\}.$$

We say that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ is a **locally linearly independent set** if the set $\{\phi(\cdot - k)\}_{k\in\Gamma}$ is linearly independent. For MRA based bi-frame systems we have the following Bernstein inequality.

Proposition 5.1. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a bi-framelet system and assume that $X(\Psi)$ is based on a compactly supported refinable function ϕ where:*

(1) *$\phi \in W^s(L_\infty(\mathbb{R}^d))$ with $s \geq 0$;*
(2) *$\{\phi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ is a locally linearly independent set (this condition is void if $d = 1$);*
(3) *The functions $\tau_\ell(\xi)$, $1 \leq \ell \leq L$ are trigonometric polynomials (see Section 2).*

Then the Bernstein inequality

$$|S|_{B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))} \leq Cm^\alpha \|S\|_{L_p(\mathbb{R}^d)}, \qquad \forall S \in \Sigma_m(X(\Psi)), \qquad \forall m \geq 1 \tag{5.1}$$

holds true for each $0 < \alpha < s/d$, $0 < p \leq \infty$, with $1/\tau := \alpha + 1/p$ and $C = C(\alpha, p)$.

Proof. In the case $d = 1$, if the integer shifts of the function ϕ are not already linearly independent, we can always find a perfect generator $\tilde{\phi}$ for the shift invariant space $S_0 := \operatorname{span}\{\phi(\cdot - k) \colon k \in \mathbb{Z}\}$, *i.e.*, $\tilde{\phi}$ is a compactly supported refinable function with linearly independent shifts that generates S_0, see [18, Theorem 1]. In particular, there exists a finite sequence $\{a_k\}_k$ such that $\phi(x) = \sum_k a_k \tilde{\phi}(x-k)$. In the arguments below we may use $\tilde{\phi}$ in place of ϕ.

By the result of Jia [17], for each $0 < \alpha < s/d$, the Bernstein inequality

$$|S|_{B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))} \leq Cm^\alpha \|S\|_{L_p(\mathbb{R}^d)}, \qquad \forall S \in \Sigma_m(X(\phi)),$$

$1/\tau := \alpha + 1/p$, $0 < p \leq \infty$, holds true for the system

$$X(\phi) := \{\phi(2^j x - k)\}_{j\in\mathbb{Z}, k\in\mathbb{Z}^d}.$$

Now, since $X(\Psi)$ is based on ϕ we have finite masks $\{b_k^\ell\}_k$ such that

$$\psi^\ell(x) = \sum_{k\in\mathbb{Z}^d} b_k^\ell \phi(2x - k).$$

Thus, for $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^d$, we have

$$\psi^\ell(2^j x - i) = \sum_{k\in\mathbb{Z}^d} b_k^\ell \phi\big(2^{j+1}x - 2i - k\big) \tag{5.2}$$

That is to say $\psi_{j,i}^\ell \in \Sigma_K(X(\phi))$ for some uniform constant K depending only the length of the finite masks used above. Take any $S \in \Sigma_m(X(\Psi))$, then $S \in \Sigma_{Km}(X(\phi))$. Using the Bernstein inequality for $X(\phi)$ we obtain the wanted inequality,

$$\begin{aligned} |S|_{B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))} &\le C(Km)^\alpha \|S\|_{L_p(\mathbb{R}^d)} \\ &\le \tilde{C} m^\alpha \|S\|_{L_p(\mathbb{R}^d)}, \qquad \forall S \in \Sigma_m(X(\Psi)). \end{aligned}$$

□

Proposition 5.1 shows that, at best, the approximation space $\mathcal{A}_q^\beta(L_p(\mathbb{R}^d), X(\Psi))$ will be (essentially) a Besov space. Indeed, as mentioned in Section 4, the Bernstein inequality (5.1) implies the continuous embedding

$$\mathcal{A}_q^\beta(L_p(\mathbb{R}^d), X(\Psi)) \hookrightarrow \big(L_p(\mathbb{R}^d), B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))\big)_{\beta/\alpha,q} \tag{5.3}$$

for $0 < \beta < \alpha$ and $q \in (0, \infty]$. It is well known (see, *e.g.*, [9]) that the right hand side of (5.3) equals the Besov space $B_q^{\beta d}(L_q(\mathbb{R}^d))$ when $1/q = \beta + 1/p$.

Thus, in order to be able to completely characterize $\mathcal{A}_q^\beta(L_p(\mathbb{R}^d), X(\Psi))$, we need to obtain a matching Jackson estimate for smooth functions in $B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))$. We address this problem in the following section.

6. Jackson estimates for the bi-frame system

In this section, we derive a Jackson inequality that matches the Bernstein inequality obtained in the previous section, and we obtain a complete characterization of m-term approximation with "nice" wavelet bi-frame systems. The basic tool is the matrix lemmas given in Section 3. The main result of this section is the following theorem.

Theorem 6.1. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a wavelet bi-frame system and assume that $X(\Psi)$ is based on a compactly supported refinable function ϕ where:*

1. *$\phi \in W^s(L_\infty(\mathbb{R}^d))$ with $s \ge 0$;*
2. *$\{\phi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ is a locally linearly independent set (this condition is void if $d = 1$);*
3. *The functions $\tau_\ell(\xi)$ are trigonometric polynomials (see Section 2);*
4. *$\tilde{\Psi} \subset C^\beta(\mathbb{R}^d) \cap M_\gamma^{N_1}(\mathbb{R}^d)$ for some $\beta > 0$, $N_1 \in \mathbb{N}$ and $\gamma > d$.*

Let $p \in (1, \infty)$ and $\tau := (\alpha + 1/p)^{-1}$ where we assume

$$0 < \alpha < \min\left\{\frac{s}{d}, \frac{1}{\Lambda\left(\frac{N_1}{d}\right)} - \frac{1}{p}\right\}, \tag{6.1}$$

with $\Lambda(x) = \Lambda(x, p, d/\gamma)$ given by (3.2). Then, for each $0 < \beta < \alpha$, $q \in (0, \infty]$, we have the characterization

$$\mathcal{A}_q^\beta(L_p(\mathbb{R}^d), X(\Psi)) = \big(L_p(\mathbb{R}^d), B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))\big)_{\beta/\alpha,q}. \tag{6.2}$$

We refer the reader to Definition 3.3 for the class $M_\gamma^N(\mathbb{R}^d)$. The above theorem is obtained simply by combining Proposition 5.1 above with Proposition 6.2 below.

Proposition 6.2. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a bi-frame. Suppose $\tilde{\Psi} \subset M_\gamma^{N_1}(\mathbb{R}^d)$ for some $N_1 \in \mathbb{N}$ and $\gamma > d$ and suppose there exist $\beta, \varepsilon > 0$ such that for all $\psi \in \Psi \cup \tilde{\Psi}$, $\psi \in C^\beta(\mathbb{R}^d)$ and $|\psi(x)| \leq C(1+|x|)^{-d-\varepsilon}$. Then, we have the Jackson inequality*

$$\sigma_m(f, X(\Psi))_p \leq Cm^{-\alpha}\|f\|_{B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))}$$

for $p \in (1,\infty)$, $\Lambda\left(\frac{N_1}{d}\right) < \tau < p$, and $\alpha = 1/\tau - 1/p$, with $\Lambda(x) = \Lambda(x, p, d/\gamma)$ given by (3.2).

We will provide a detailed proof of Proposition 6.2 later in this section, but let us first give an outline of the proof. First, using the characterization of Besov spaces in terms of wavelet coefficients, we know that the wavelet coefficients of any $f \in B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d)$ are in ℓ_τ. Then, using the matrix lemmas from Section 3, we show that the wavelet-frame coefficients in Eq. (1.1), once properly normalized, are also in ℓ_τ for τ admissible. Thanks to the $L_p(\mathbb{R}^d)$-stability of bi-frame expansions –which we prove in Appendix A– we show that this implies $f \in \mathcal{K}_{\tau,\tau}\big(L_p(\mathbb{R}^d), X(\Psi)\big)$, where the latter class is a "sparseness class" of functions in $L_p(\mathbb{R}^d)$ with well-defined ℓ_τ-summable frame expansions in $X(\Psi)$. We can then conclude and get the desired Jackson inequality by relying on a more general "abstract" Jackson inequality with $Y_p(\mathbb{R}^d) = \mathcal{K}_{\tau,\tau}\big(L_p(\mathbb{R}^d), X(\Psi)\big)$ proved in [14, Theorem 6].

Let us now give the definition of the sparseness classes $\mathcal{K}_{\tau,\tau}\big(L_p(\mathbb{R}^d), X(\Psi)\big)$. Then, we will have all the tools in our hands to prove Proposition 6.2.

6.1. **Sparseness Classes.** It will be proved in the appendix that the bi-frame system $X(\Psi)$, once normalized in $L_p(\mathbb{R}^d)$, has the so-called $\ell_{p,1}$-hilbertian property (see Corollary A.4). Applying Proposition 3 in [14], it follows that we can define, for $p \in (1,\infty)$, $\tau < p$ and $q \in (0,\infty]$:

$$\mathcal{K}_{\tau,q}\big(L_p(\mathbb{R}^d), X(\Psi)\big) = \left\{ f \in L_p(\mathbb{R}^d) \,\middle|\, \exists \{c_{j,k}^\ell\}_{j,k,\ell} \in \ell_{\tau,q},\ f = \sum_{\ell=1}^{L} \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} c_{j,k}^\ell \psi_{j,k}^{\ell,p} \right\},$$

and $|f|_{\mathcal{K}_{\tau,q}(L_p(\mathbb{R}^d),X(\Psi))}$ the smallest Lorentz norm $\|\{c_{j,k}^\ell\}_{j,k,\ell}\|_{\ell_{\tau,q}}$ such that $f = \sum_{j,k,\ell} c_{j,k}^\ell \psi_{j,k}^{\ell,p}$ (see, *e.g.*, (A.3) for the definitions of the Lorentz norm on $\ell_{p,q}$).

6.2. **Proof of a Jackson inequality for the bi-frame system.** We are now ready to give the proof of Proposition 6.2.

Proof of Proposition 6.2. Take $f \in B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))$. The expansion of f in the Meyer wavelet system $\{\eta^i\}_{i=1}^{2^d-1}$ is given by

$$f = \sum_{i=1}^{2^d-1} \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} c_{j,k}^i \eta_{j,k}^{i,p}, \qquad 1/\tau - 1/p = \alpha,$$

with coefficients $c^i_{j,k}$ satisfying $\|\{c^i_{j,k}\}\|_{\ell_\tau} \asymp \|f\|_{B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))}$, see [10]. Now we calculate the bi-frame coefficients of f:

$$\langle f, \tilde{\psi}^{\ell,p'}_{j',k'}\rangle = \sum_{i=1}^{2^d-1}\Big(\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} c^i_{j,k}\big\langle \eta^{i,p}_{j,k}, \tilde{\psi}^{\ell,p'}_{j',k'}\big\rangle\Big).$$

Since $\eta^i \in D^N_\gamma(\mathbb{R}^d)\cap M^N_\gamma(\mathbb{R}^d)$ for any N and γ, we can apply Proposition 3.5 for each ℓ and i to get that $\|\{\langle f, \tilde{\psi}^{\ell,p'}_{j',k'}\rangle\}\|_{\ell_\tau} \leq C\|f\|_{B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))}$. According to Corollary A.4, the system $X(\Psi)$ is $\ell_{p,1}$-hilbertian so the canonical frame expansion (1.1), which can be rewritten as

$$f = \sum_{\ell=1}^{L}\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} \langle f, \tilde{\psi}^{\ell,p'}_{j,k}\rangle \psi^{\ell,p}_{j,k}, \tag{6.3}$$

is unconditionally convergent in $L_p(\mathbb{R}^d)$. It follows that $f \in \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi))$ and $|f|_{\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d),X(\Psi))} \leq C\|f\|_{B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))}$. Eventually, using again the fact that the system $X(\Psi)$ is $\ell_{p,1}$-hilbertian, we obtain by [14, Theorem 6]

$$\begin{aligned}\sigma_m(f, X(\Psi))_p &\leq Cm^{-\alpha}|f|_{\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d),X(\Psi))}\\ &\leq Cm^{-\alpha}\|\{\langle f, \tilde{\psi}^{\ell,p'}_{j',k'}\rangle\}\|_{\ell_\tau}\\ &\leq Cm^{-\alpha}\|f\|_{B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))}. \qquad (6.4)\end{aligned}$$

□

Remark 6.3. Notice that the sparse representation of $f \in B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))$ we use to prove that $f \in \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi))$ is exactly the canonical bi-frame expansion (properly normalized in $L_p(\mathbb{R}^d)$). Thus, to realize the rate of approximation given by the Jackson estimate, we simply *threshold* the coefficients of the expansion (6.3).

6.3. **Characterization of the sparseness classes.** One of the interesting byproducts of the proof of Proposition 6.2 is that, under the hypotheses of Theorem 6.1, we also get a characterization of Besov spaces and approximation spaces in terms of the sparseness classes. Indeed, one can deduce from the proof of Proposition 6.2 that

$$B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d)) = \Big\{f \in L_p(\mathbb{R}^d), \{\langle f, \psi^{\ell,p'}_{j,k}\rangle\}_{j,k,\ell} \in \ell^\tau\Big\} = \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi)) \tag{6.5}$$

for $\alpha = 1/\tau - 1/p$ and τ admissible (see the statement of Proposition 6.2). The embedding of $B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))$ into the middle class in Eq. (6.5) is derived explicitly in the proof of Proposition 6.2. The embedding of the middle class into $\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi))$ is trivial, and the converse embedding, $\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi)) \hookrightarrow B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))$ is a consequence of the Jackson estimate $\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi)) \hookrightarrow \mathcal{A}^\alpha_\tau(L_p(\mathbb{R}^d), X(\Psi))$ –which follows from [14, Theorem 6]– and the Bernstein "embedding" (5.3) for the system $X(\Psi)$.

7. On the role of Vanishing Moments

In this section, we discuss the role of the number N_1 of vanishing moments of the bi-framelet system with respect to its ability to characterize Besov spaces through sparseness spaces and/or approximation spaces. We say that a collection of wavelets Ψ has vanishing moments of order N if each $\psi \in \Psi$ has at least N vanishing moments (*i.e.*, $\psi \in M^N_\gamma(\mathbb{R}^d)$

for some $\gamma > 0$, see Eq. (3.5)) , and at least one of the functions in Ψ has exactly N vanishing moments, *i.e.*, $\psi \notin M_\gamma^{N'}$ for any $N' > N$. We denote by $\mathrm{VM}(\Psi)$ the order of vanishing moments of Ψ.

At the beginning of this section, we discuss the fact that when $\mathrm{VM}(\tilde{\Psi})$ is "small", the range of smoothness parameters α for which Besov spaces are characterized by bi-framelet expansions/approximations (see Theorem 6.1 and Eq. (6.5)) can be rather limited. Following an idea introduced in [15] we propose means to overcome this limitation: since the system $X(\Psi)$ is redundant, one can choose another representation of any f than its canonical frame expansion; by replacing the canonical frame representation with a "better" one, it is sometimes possible to recover the characterization of Besov spaces even when $\mathrm{VM}(\tilde{\Psi})$ is small. However, the price we have to pay for getting better expansions is that we can no longer expand the functions in the canonical system $X(\Psi)$, but we have to consider expansions in the twice oversampled system, which will be defined soon.

7.1. **Limitations of the canonical frame expansion.** Let us first motivate the need for the results of this section by discussing the role of the number $N_1 := \mathrm{VM}(\tilde{\Psi})$ of vanishing moments of the dual bi-frame, and some limitations of Theorem 6.1 when we use framelet systems with few vanishing moments.

As stated in Theorem 6.1 the number of vanishing moments puts a limit to the range of approximation spaces for the bi-framelets that can be characterized even if we assume the wavelets are smooth and compactly supported. Notice that by Eq.(3.2),

$$\frac{1}{\Lambda(x)} - \frac{1}{p} = \begin{cases} (1/x-1)^{-1}\cdot 1/p & \text{if } x \le 1-1/p, \\ x & \text{if } 1-1/p < x \le \gamma/d - 1/p, \\ \gamma/d - 1/p & \text{if } \gamma/d - 1/p < x. \end{cases}$$

Thus, if $N_1/d > 1-1/p$, the set of admissible α in Theorem 6.1 is given by

$$0 < \alpha < \min\left\{\frac{s}{d}, \frac{N_1}{d}, \frac{\gamma}{d} - \frac{1}{p}\right\}. \tag{7.1}$$

If $N_1/d \le 1-1/p$, the set is even more restricted. This shows the importance of the three qualities of a wavelet: smoothness, vanishing moments and decay. A "good" bi-frame system, $X(\Psi)$, $X(\tilde{\Psi})$, in this context is when the functions Ψ are compactly supported and smooth, and the functions $\tilde{\Psi}$ have fast decay and a large number of vanishing moments.

On Figure 1 we display, as a function of N_1/d (and with s, d, p and γ fixed), the range of admissible values of α.

$\mathrm{VM}(\psi)$ *for univariate wavelets.* For a univariate orthonormal wavelet $\psi \in L_2(\mathbb{R}) \cap C^r(\mathbb{R})$ with decay $|\psi(x)| \le C(1+|x|)^{-r-\varepsilon}$ for some $\varepsilon > 0$, it is well known that ψ has (at least) r vanishing moments, see e.g. [16, Theorem 2.3.4]. So for orthonormal wavelets, smoothness and decay enforce vanishing moments, and we get the characterization

$$\mathcal{A}_q^\beta(L_p(\mathbb{R}^d), X(\psi)) = (L_p(\mathbb{R}), B_\tau^\alpha(L_\tau(\mathbb{R})))_{\beta/\alpha,q} \tag{7.2}$$

for all $1 < p < \infty$, $0 < q \le \infty$, $0 < \beta < \alpha < r$, and $\tau = (\alpha + 1/p)^{-1}$.

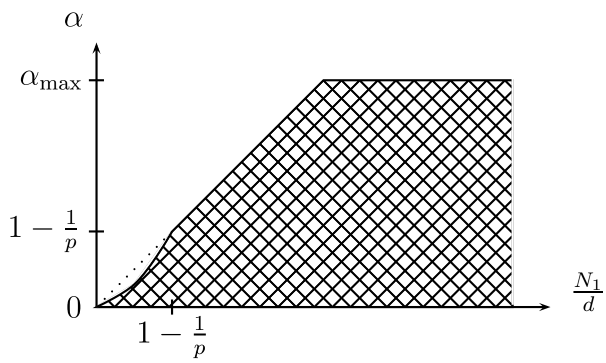

FIGURE 1. A crosshatch of the range of admissible values of α in Theorem 6.1 (given by (6.1)) as a function of N_1/d, $\alpha_{\max} = \min\{s/d, \gamma/d - 1/p\}$.

VM(Ψ) *for univariate tight spline framelets.* The relation between smoothness and vanishing moments satisfied by orthonormal wavelets, is far from true for wavelet frames. The most well-known example is the family of spline-based tight wavelet frames built through the unitary extension principle: the smoothness s, which corresponds to the degree of the splines, can be arbitrarily high; however at least one of the framelets will have only one vanishing moment, see [8], hence $N_1 = \mathrm{VM}(\Psi) = \mathrm{VM}(\tilde{\Psi}) = 1$. In this case, since the wavelets have compact support, we can take $\gamma \gg N_1$ and check that the range of admissible α is exactly $0 < \alpha < N_1 = 1$ while the smoothness s of Ψ might be arbitrarily large.

Wavelet bi-frames with few vanishing moments in $L_p(\mathbb{R}^d)$. Suppose the decay γ and smoothness s are large compared to the number N_1 of vanishing moments of $X(\tilde{\Psi})$:

$$\begin{aligned} s &> N_1 \\ \gamma &> d + N_1 \end{aligned}$$

and that the latter is not too small, *i.e.*, $N_1 \geq d$. Then, again, the range of admissible α is exactly $0 < \alpha < N_1/d$.

At first sight, it seems that the possibly limited number of vanishing moments of bi-framelet systems is an obstacle that prevents them from having good approximation properties, in the sense that the characterization (7.2) is for a limited range of α. The rest of this section is devoted to showing that this restriction can be avoided by using an oversampled version of the framelet system.

7.2. **Beyond the canonical frame decomposition.** We have mentioned that to get characterizations of Besov spaces with systems with few vanishing moments, we will need to oversample them. Given a wavelet bi-frame $X(\Psi)$, $X(\tilde{\Psi})$ and $R \geq 1$ we let $X_R(\Psi)$ denote the oversampled system,

$$X_R(\Psi) := \left\{ 2^{jd/2}\psi^\ell(2^j \cdot -k/R) | j \in \mathbb{Z}, k \in \mathbb{Z}^d, \ell = 1, 2, \ldots, L \right\}.$$

Just as the non-oversampled system, the oversampled one $X_R(\Psi)$ is a frame in $L_2(\mathbb{R})$ and is $\ell_{p,1}$-hilbertian in $L_p(\mathbb{R}^d)$ after proper normalization. See, *e.g.*, [15, Theorem 4.11] for a proof in the case $R = 2$, which directly extends to arbitrary oversampling factors $R \in \mathbb{N}$. Denoting $\psi_I^\ell(\cdot) = 2^{jd/2}\psi^\ell(2^j \cdot -k/R), I \in D_R$, where D_R is the collection of all "oversampled dyadic intervals" $I = 2^{-j}([0,1]^d + k/R)$, we can thus define sparseness classes for $p \in (1,\infty)$, $\tau < p$ and $q \in (0,\infty]$ as follows

$$\mathcal{K}_{\tau,q}\big(L_p(\mathbb{R}^d), X_R(\Psi)\big) := \left\{ f \in L_p(\mathbb{R}^d) \;\middle|\; \exists \{c_I^\ell\}_{I,\ell} \in \ell_{\tau,q},\ f = \sum_{I \in D_R, \ell \in E} c_I^\ell \psi_I^{\ell,p} \right\},$$

and $|f|_{\mathcal{K}_{\tau,q}(L_p(\mathbb{R}^d),X(\Psi))}$ the smallest Lorentz norm $\|\{c_I^\ell\}_{I,\ell}\|_{\ell_{\tau,q}}$ such that $f = \sum_{I,\ell} c_I^\ell \psi_I^{\ell,p}$.

It turns out that, when $X(\Psi), X(\tilde{\Psi})$ is a "nice" wavelet bi-frame, the oversampled system $X_R(\Psi)$, again, gives rise to an approximation space no larger than a Besov space. Indeed, using refinability in (5.2) we can prove a corresponding Bernstein inequality for the oversampled system $X_R(\Psi)$, by exactly the same arguments as given in the proof of Proposition 5.1 (see [4] for details on how to use refinability to deal with the oversampling).

Proposition 7.1. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a bi-framelet system with $X(\Psi)$ based on a compactly supported refinable function ϕ where:*

1. *$\phi \in W^s(L_\infty(\mathbb{R}^d))$ with $s \geq 0$;*
2. *$\{\phi(\cdot - k)\}_{k \in \mathbb{Z}^d}$ is a locally linearly independent set (this condition is void if $d = 1$);*
3. *The functions $\tau_\ell(\xi)$, $1 \leq \ell \leq L$ are trigonometric polynomials (see Section 2).*

Then, the Bernstein inequality

$$|S|_{B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))} \leq C m^\alpha \|S\|_{L_p(\mathbb{R}^d)}, \qquad \forall S \in \Sigma_m(X_R(\Psi)), \qquad \forall m \geq 1$$

holds true for each $0 < \alpha < s/d$, $0 < p \leq \infty$, with $1/\tau := \alpha + 1/p$ and $C = C(\alpha, p)$.

Again, in order to get a complete characterization of Besov spaces in terms of the approximation spaces based on $X_R(\Psi)$, we need to prove a matching Jackson estimate. Just as in Section 6, we will use the fact that we can find a "nice" orthogonal wavelet system with a sparse expansion in $X_R(\Psi)$. However, the sparse expansion that we will consider is no longer expressed in terms of canonical frame coefficients of the orthogonal wavelet(s), but rather on adequately chosen sparse synthesis coefficients, as can be seen in the next lemma.

Lemma 7.2. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a wavelet bi-frame system. Let $\psi \in L_2(\mathbb{R}^d)$ for which there exists a sequence $\{d_k^\ell\}_{\ell \in E,\ k \in \mathbb{Z}^d}$ and $R \in \mathbb{N}$ such that*

$$\psi(x) = \sum_{\ell=1}^{L} \sum_{k \in \mathbb{Z}^d} d_k^\ell \psi^\ell(x - k/R).$$

Then, for $1 < p < \infty$, and $\tau < p$ such that $\{d_k^\ell\} \in \ell_1 \cap \ell_\tau$, we have

$$\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\psi)) \hookrightarrow \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X_R(\Psi)).$$

Proof. Let $f \in \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X(\psi))$. Then $f = \sum_{j,k} c_{j,k}\psi^p_{j,k}$, for some sequence $\{c_{j,k}\} \in \ell_\tau$. We rewrite this as

$$\begin{aligned} f &= \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d} c_{j,k}2^{j/p}\psi(2^j\cdot -k) \\ &= \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d}\sum_{\ell=1}^{L}\sum_{k'\in\mathbb{Z}^d} c_{j,k}d^\ell_{k'}2^{j/p}\psi^\ell(2^j\cdot -k-k'/R) \\ &= \sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d}\sum_{\ell=1}^{L}\sum_{n\in\mathbb{Z}^d} c_{j,k}d^\ell_{n-Rk}2^{j/p}\psi^\ell(2^j\cdot -n/R) \\ &= \sum_{j\in\mathbb{Z},n\in\mathbb{Z}^d}\sum_{\ell=1}^{L}\left(\sum_k c_{j,k}d^\ell_{n-Rk}\right)2^{j/p}\psi^\ell(2^j\cdot -n/R). \end{aligned}$$

It is easy to check using brute force for $0 < \tau \le 1$, and Young's inequality for $1 < \tau < p$, that

$$\sum_{j\in\mathbb{Z},n\in\mathbb{Z}^d}\sum_{\ell=1}^{L}\left|\sum_k c_{j,k}d^\ell_{n-Rk}\right|^\tau \le L\cdot\max\left(\|\{d^\ell_k\}\|_{\ell_1}, \|\{d^\ell_k\}\|^\tau_{\ell_\tau}\right)\cdot\sum_{j\in\mathbb{Z},k\in\mathbb{Z}^d}|c_{j,k}|^\tau,$$

and we conclude that indeed $f \in \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X_R(\Psi))$. □

The following corollary gives more details on how Lemma 7.2 will be used to prove the desired Jackson inequality.

Corollary 7.3. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a wavelet bi-frame system, $X(\{\psi^i\}_{i=1}^{2^d-1})$ a bi-orthogonal wavelet basis and $r > 0$ such that*

$$B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d)) = \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\{\psi^i\}_{i=1}^{2^d-1})),\ 0 < \alpha = 1/\tau - 1/p < r.$$

Assume that for $1 \le i \le 2^d - 1$ there exists sequences $\{d^{\ell,i}_k\}_{\ell\in E,k\in\mathbb{Z}^d} \in \ell_{1/(r+1)}$, such that

$$\psi^i(x) = \sum_{\ell=1}^{L}\sum_{k\in\mathbb{Z}^d} d^{\ell,i}_k\psi^\ell(x-k/R).$$

Then, for $1 < p < \infty$, and $0 < \alpha = 1/\tau - 1/p < r$, we have the Jackson inequality

$$\sigma_m(f, X_R(\Psi))_p \le Cm^{-\alpha}\|f\|_{B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))}.$$

Proof. First, notice that we have (see [14]), for $0 < \alpha = 1/\tau - 1/p < r$,

$$B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d)) = \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\{\psi^i\}_{i=1}^{2^d-1})) = \sum_{i=1}^{2^d-1}\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\psi^i)).$$

Then, since $\tau = (\alpha+1/p)^{-1} > (r+1/p)^{-1} \ge (r+1)^{-1}$ we have $\ell_{1/(r+1)} \subset \ell_1$. Hence, using the fact that $\{d^{\ell,i}_k\} \in \ell_{1/(r+1)} \cap \ell_1 \subset \ell_\tau \cap \ell_1$ we can apply Lemma 7.2 to get, for each i,

$$\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\psi^i)) \hookrightarrow \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X_R(\Psi)).$$

The conclusion follows just as in the proof of Proposition 6.2. □

We can now combine Corollary 7.3 and Proposition 7.1 to get the following complete characterization of the approximation spaces $\mathcal{A}^\beta_q(L_p(\mathbb{R}^d), X_2(\Psi))$. Notice the extended range for α as compared to Theorem 6.1.

Corollary 7.4. *Suppose $X(\Psi)$, $X(\tilde{\Psi})$ satisfy all the assumptions of Proposition 7.1 and Corollary 7.3 with parameters s and r, respectively. Then, for $0 < \alpha < \min\{s, r\}$, $0 < \beta < \alpha$, $q \in (0, \infty]$, we have the characterization*

$$\mathcal{A}^\beta_q(L_p(\mathbb{R}^d), X_2(\Psi)) = \big(L_p(\mathbb{R}^d), B^{d\alpha}_\tau(L_\tau(\mathbb{R}^d))\big)_{\beta/\alpha,q}. \tag{7.3}$$

Now the task is to show that the assumption of Lemma 7.2 is satisfied for interesting framelet systems. This will be done in the following subsection where we restrict to the univariate case $d = 1$.

7.3. **Constructing wavelets out of framelets in the univariate case.**

In our strategy to get a Jackson inequality for the (oversampled) framelet system $X_R(\Psi)$, the crucial issue is to identify some "nice" wavelet(s) that can be expanded sparsely in terms of oversampled bi-framelets. In [4] such a construction is proposed in the multivariate case for wavelet type systems that need not be frames, however there is no control on how large the oversampling factor R must be, and the proof is not constructive. In the *univariate* case for spline-based *tight* wavelet frames, it was shown in [15] how to get a finite expansion of a nice semi-orthogonal wavelet in the twice oversampled ($R = 2$) framelet system. Here, still in the *univariate* case, we propose a construction which is valid for more general wavelet bi-frames and only requires $R = 2$. It is an interesting but open question whether similar constructions are possible in the multivariate case.

The constructed wavelet will be the "standard" orthogonal wavelet associated with the MRA underlying the bi-framelet system.

Definition 7.5. Let ϕ be a univariate scaling function generated by the refinement filter $\tau_0(\xi)$, and let $P(\xi) := \sum_{k\in\mathbb{Z}} |\hat{\phi}(\xi-k)|^2$. The "standard" orthonormal wavelet ψ associated with the scaling function ϕ is defined by

$$\hat{\psi}(2\xi) = e^{-i\xi}\overline{\tau_0(\xi+\pi)}\frac{\hat{\phi}(\xi)}{\sqrt{P(\xi)}}. \tag{7.4}$$

Let us recall that the number of vanishing moments of the standard orthonormal wavelet associated to ϕ is given by

$$\mathrm{VM}(\psi) = \min\{N, |\tau_0(\xi+\pi)| = O(|\xi|^N) \text{ around } \xi = 0\},$$

see, *e.g.*, [20]. Moreover, as discussed previously, if ϕ is r-regular (see [20]) then so is ψ and we have $\mathrm{VM}(\psi) \geq r$.

Proposition 7.6. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be an MRA-based wavelet bi-frame system with combined mask $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_L)$ and $\tilde{\boldsymbol{\tau}}$, and let ϕ and ψ be respectively the scaling function generated by τ_0 and the associated standard orthonormal wavelet. Suppose that*

- *each filter τ_ℓ, $0 \leq \ell \leq L$, is a trigonometric polynomial;*
- *$\sum_{\ell=1}^L |\tau_\ell(\xi)|^2 > 0$ for $\xi \neq 0$;*
- *ϕ is an r-regular scaling function (not necessarily orthonormal);*

- $\mathrm{VM}(\Psi) \le \mathrm{VM}(\psi)$.

Then ψ can be expressed as a linear combination

$$\psi(\cdot) := \sum_{\ell=1}^{L} \sum_{k\in\mathbb{Z}} d_k^\ell \psi^\ell(\cdot - k/2), \qquad \Psi = \{\psi^1, \psi^2, \ldots, \psi^L\},$$

where $\{d_k^\ell\} \in \bigcap_{\tau>0} \ell_\tau$.

Remark 7.7. It is clear that such a construction will fail if we have $\mathrm{VM}(\psi) < \mathrm{VM}(\Psi)$. This follows from the fact that a sparse expansion of functions with at least $\mathrm{VM}(\Psi)$ vanishing moments will also have (at least) $\mathrm{VM}(\Psi)$ vanishing moments, while the orthonormal wavelet will have exactly $\mathrm{VM}(\psi) < \mathrm{VM}(\Psi)$ vanishing moments. However this is not a problem since, in such a case, $X(\Psi)$ has enough vanishing moments to get a Jackson inequality from the theory developed in Section 6.

Proof. We want to expand the standard orthonormal wavelet ψ in the twice oversampled framelet system. In the frequency domain the problem is to find "nice" 2π-periodic functions $Q_\ell(\xi)$ such that

$$\hat{\psi}(\xi) = \sum_{\ell=1}^{L} Q_\ell(\xi/2)\tau_\ell(\xi/2)\hat{\phi}(\xi/2).$$

We will look for Q_ℓ of the form $Q_\ell(\xi) = Q(\xi)\overline{\tau_\ell(\xi)}$. Using Eq. (7.4), we see that the problem will be solved if Q_ℓ has fast decaying Fourier coefficients and Q satisfies

$$Q(\xi)\sum_{\ell=1}^{L} |\tau_\ell(\xi)|^2 = \frac{e^{-i\xi}\overline{\tau_0(\xi+\pi)}}{\sqrt{P(\xi)}}.$$

Hence, we define for $\xi \neq 0$

$$Q_\ell(\xi) := \frac{\tau_\ell(\xi)\cdot\overline{\tau_0(\xi+\pi)}}{\sum_{\ell=1}^{L}|\tau_\ell(\xi)|^2}\cdot\frac{e^{-i\xi}}{\sqrt{P(\xi)}}. \tag{7.5}$$

Let us check that Q_ℓ can be readily extended at $\xi = 0$ and that the resulting extension has no pole on the unit circle. First, we have[1], for ξ close to zero, $\sum_{\ell=1}^{L}|\tau_\ell(\xi)|^2 \asymp |\xi|^{2\cdot\mathrm{VM}(\Psi)}$. Then, we use the fact that, for ξ close to zero,

$$|\tau_\ell(\xi)\overline{\tau_0(\xi+\pi)}| = O(|\xi|^{\mathrm{VM}(\Psi)+\mathrm{VM}(\psi)}) = O(|\xi|^{2\cdot\mathrm{VM}(\Psi)}).$$

We conclude by proving that the Fourier coefficients of Q_ℓ decay faster than any polynomial. Notice that $P(\xi)^{-1/2}$ is C^∞ (see, *e.g.*, [20]) so its Fourier coefficients decay faster than any polynomial. The factor

$$\frac{\tau_\ell(\xi)\cdot\overline{\tau_0(\xi+\pi)}}{\sum_{\ell=1}^{L}|\tau_\ell(\xi)|^2}$$

in (7.5) is a quotient of two trigonometric polynomials with no pole on the unit circle, so its Fourier coefficients decay exponentially which can be seen from its Laurent expansion. □

By combining Corollary 7.3 and Proposition 7.6 we get

[1] By $F \asymp G$ we mean that there exist two constants $0 < c \le C < \infty$ such that $cF \le G \le CF$.

Corollary 7.8. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be an MRA-based wavelet bi-frame system with combined mask $\boldsymbol{\tau} = (\tau_0, \tau_1, \dots, \tau_L)$ and $\tilde{\boldsymbol{\tau}}$, and let ϕ and ψ be respectively the scaling function generated by τ_0 and the associated standard orthonormal wavelet. Suppose that*

- *each filter τ_ℓ, $0 \le \ell \le L$, is a trigonometric polynomial;*
- $\sum_{\ell=1}^{L} |\tau_\ell(\xi)|^2 > 0$ *for* $\xi \neq 0$;
- *ϕ is an r-regular scaling function (not necessarily orthonormal);*
- $\mathrm{VM}(\Psi) \le \mathrm{VM}(\psi)$.

Then, for $1 < p < \infty$, $0 < \tau < p$,

$$\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X(\psi)) \hookrightarrow \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X_2(\Psi)).$$

In particular, for $1 < p < \infty$ and $0 < \alpha = 1/\tau - 1/p < r$, we have the Jackson inequality

$$\sigma_m(f, X_2(\Psi))_p \le C m^{-\alpha} \|f\|_{B_\tau^\alpha(L_\tau(\mathbb{R}))}.$$

Proof. Using Proposition 7.6, we get the sparse expansion coefficients which make it possible to apply Lemma 7.2 to obtain $\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X(\psi)) \hookrightarrow \mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X_2(\Psi))$. The second claim follows from Corollary 7.3 using the fact that $\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}), X(\psi)) = B_\tau^\alpha(L_\tau(\mathbb{R}))$, $\alpha \in (0, r)$, $\alpha = 1/\tau - 1/p$, see *e.g.* [10]. □

Remark 7.9. In particular, Corollary 7.8 applies to any tight framelet system Ψ with $\mathrm{VM}(\Psi) = 1$. For example, it applies to tight framelet systems constructed using the UEP based on the B-splines [8, 15].

8. Conclusion

In this paper we have studied approximation with wavelet bi-frame systems in $L_p(\mathbb{R}^d)$, $1 < p < \infty$, and we have characterized the associated approximation spaces $\mathcal{A}^\alpha$ and shown that they are essentially Besov spaces $B_\tau^{d\alpha}(L_\tau(\mathbb{R}^d))$, with $\alpha = 1/\tau - 1/p$. The characterization holds true for the smoothness parameter α in a certain range depending on the number of vanishing moments for the bi-frame system. It is also shown that for a function f in a Besov space with smoothness parameter in this range, the corresponding canonical $L_p(\mathbb{R}^d)$-normalized bi-frame expansion of f is sparse in the sense that the frame coefficients are contained in ℓ_τ. Moreover, the rate of best m-term approximation to f is obtained simply by thresholding the canonical expansion.

For twice oversampled univariate wavelet bi-frames, we give a complete characterization of the approximation spaces in terms of the Besov spaces $B_\tau^\alpha(L_\tau(\mathbb{R}))$. The characterization holds true even for systems with few vanishing moments, and there is no restriction on the smoothness parameter α except the natural requirement that α is less than the smoothness of the generators of the wavelet frame. To obtain a characterization for wavelet bi-frames with few vanishing moments, we prove that there exists a "nice" orthonormal wavelet with a highly sparse expansion in the framelet system. This fact is then used to show that smooth functions in $B_\tau^\alpha(L_\tau(\mathbb{R}))$ have sparse expansions in the twice oversampled wavelet bi-frame system with expansion coefficients in ℓ_τ.

There is one fundamental difference between the sparse expansions obtained for systems with many vanishing moments and for the twice oversampled wavelet bi-frames. When we have enough vanishing moments, we can use the canonical frame expansion and it will be sparse for smooth functions. For wavelet bi-frames with few vanishing moments we no longer use the canonical expansion, but show that a smooth function has another sparse expansion in the twice oversampled system, obtained through an expansion of the function in an orthonormal wavelet.

Appendix A. Stability of wavelet bi-frame systems in $L_p(\mathbb{R}^d)$

In this appendix we study stability of wavelet bi-frame systems in $L_p(\mathbb{R}^d)$. Theorem A.1 below will show that we can characterize the $L_p(\mathbb{R}^d)$-norms by the analysis coefficients associated with the bi-frame. Theorem A.3 will show that there is a stable way to reconstruct $L_p(\mathbb{R}^d)$-functions using the bi-frame expansion. The main application of the stability result in this paper is to prove the Jackson inequality for the bi-frame system (Proposition 6.2). The technique to obtain the various characterizations in this appendix is similar to the technique introduced by the authors in [3].

Below we let D denote the set of dyadic intervals $I = 2^{-j}([0,1]^d + k)$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^d$, $\psi_I := \psi_{j,k}$, and χ_I denotes the indicator function for I.

Theorem A.1. *Let $X(\Psi)$, $X(\tilde{\Psi})$ be a wavelet bi-frame system. Suppose for all $\psi \in \Psi \cup \tilde{\Psi}$ there are $\beta > 0$ and $\varepsilon > 0$ such that $\psi \in C^\beta(\mathbb{R}^d)$, and*

$$|\psi(x)| \le C(1+|x|)^{-d-\varepsilon}.$$

Then

(A.1)

$$\|f\|_p \asymp \Big\| \Big(\sum_{I\in D, \ell\in E} |\langle f, \psi_I^\ell\rangle|^2 |I|^{-1}\chi_I(x)\Big)^{1/2}\Big\|_p \asymp \Big\| \Big(\sum_{I\in D, \ell\in E} |\langle f, \tilde{\psi}_I^\ell\rangle|^2 |I|^{-1}\chi_I(x)\Big)^{1/2}\Big\|_p,$$

for $1 < p < \infty$, where $E := \{1, 2, \ldots, L\}$.

Proof. Let $\{\eta^s\}_{s=1}^{2^d-1}$ be the orthonormal Meyer wavelet(s) defined on $\mathbb{R}^d$. For each $\ell \in E$ we consider the integral kernel

$$K^\ell(x,y) := \sum_{I\in D} \eta_I^1(x)\overline{\psi_I^\ell(y)}.$$

Notice that the corresponding operator

$$T^\ell \colon f \mapsto \int_{\mathbb{R}^d} K^\ell(x,y) f(y)\, dy$$

is bounded on $L_2(\mathbb{R}^d)$ due to the fact that $\{\psi_I^\ell\}_{I\in D}$ is a subset of a frame. Also, standard estimates show that (see *e.g.* [7])

$$|K^\ell(x,y)| \le C|x-y|^{-d},$$

$$|K^\ell(x',y) - K^\ell(x,y)| \le C|x-x'|^\alpha |x-y|^{-d-\alpha},$$

and

$$|K^\ell(x,y') - K^\ell(x,y)| \le C|y-y'|^\alpha |x-y|^{-d-\alpha},$$

because of the smoothness and decay of ψ^ℓ. Thus T^ℓ is a Calderón-Zygmund operator and therefore bounded on $L_p(\mathbb{R}^d)$, $1 < p < \infty$. However $T^\ell f$ has a nice expansion in the orthonormal Meyer wavelet, so using the $L_p(\mathbb{R}^d)$-characterization of such expansions we get

$$\Big\|\Big(\sum_{I\in D}|\langle f,\psi_I^\ell\rangle|^2|I|^{-1}\chi_I(x)\Big)^{1/2}\Big\|_p \asymp \|T^\ell f\|_p \le C\|f\|_p.$$

Using this estimate for $\ell = 1, 2, \dots, L$, and the fact that $\ell_1 \hookrightarrow \ell_2$ we get

$$\begin{aligned}
&\Big\|\Big(\sum_{I\in D,\ell\in E}|\langle f,\psi_I^\ell\rangle|^2|I|^{-1}\chi_I(x)\Big)^{1/2}\Big\|_p \\
&\qquad = \Big\|\Big(\sum_{\ell\in E}\Big(\Big\{\sum_{I\in D}|\langle f,\psi_I^\ell\rangle|^2|I|^{-1}\chi_I(x)\Big\}^{1/2}\Big)^2\Big)^{1/2}\Big\|_p \\
&\qquad \le \Big\|\sum_{\ell\in E}\Big\{\sum_{I\in D}|\langle f,\psi_I^\ell\rangle|^2|I|^{-1}\chi_I(x)\Big\}^{1/2}\Big\|_p \\
&\qquad \le L\cdot C\|f\|_p.
\end{aligned}$$

Using a similar argument applied to the frame $X(\tilde{\Psi})$, we may conclude that

$$\Big\|\Big(\sum_{I\in D,\ell\in E}|\langle f,\tilde{\psi}_I^\ell\rangle|^2|I|^{-1}\chi_I(x)\Big)^{1/2}\Big\|_p \le L\cdot C\|f\|_p.$$

Now we turn to the converse estimate. Notice that since we have a bi-frame we have the identity

$$\langle f,g\rangle = \sum_{I\in D,\ell\in E}\langle f,\tilde{\psi}_I^\ell\rangle\overline{\langle g,\psi_I^\ell\rangle}, \quad f,g\in L_2(\mathbb{R}^d).$$

Write

$$Wf(x) = \{|I|^{-1/2}\langle f,\psi_I^\ell\rangle\chi_I(x)\}_{I,\ell}, \qquad \tilde{W}g(x) = \{|I|^{-1/2}\langle g,\tilde{\psi}_I^\ell\rangle\chi_I(x)\}_{I,\ell},$$

and notice that for $f\in L_2(\mathbb{R}^d)\cap L_p(\mathbb{R}^d)$ and $g\in L_2(\mathbb{R}^d)\cap L_{p'}(\mathbb{R}^d)$, with $p^{-1}+(p')^{-1}=1$,

$$\begin{aligned}
|\langle f,g\rangle| &= \Big|\int\langle Wf(x),\tilde{W}g(x)\rangle_{\ell_2}\,dx\Big| \\
&\le \Big|\int\|Wf(x)\|_{\ell_2}\|\tilde{W}g(x)\|_{\ell_2}\,dx\Big| \\
&\le \big\|\|Wf(x)\|_{\ell_2}\big\|_p\big\|\|\tilde{W}g(x)\|_{\ell_2}\big\|_{p'} \\
&\le C\begin{cases}\big\|\|Wf(x)\|_{\ell_2}\big\|_p\|g\|_{p'}, \\ \big\|\|\tilde{W}g(x)\|_{\ell_2}\big\|_{p'}\|f\|_p.\end{cases}
\end{aligned} \tag{A.2}$$

Taking the supremum of the estimate (A.2) for $\{g\in L_2(\mathbb{R}^d)\cap L_{p'}(\mathbb{R}^d) : \|g\|_{p'}\le 1\}$ we obtain

$$\|f\|_p \le \tilde{C}\big\|\|Wf(x)\|_{\ell_2}\big\|_p.$$

This proves the result for $f \in L_2(\mathbb{R}^d) \cap L_p(\mathbb{R}^d)$. To complete the proof for $f \in L_p(\mathbb{R}^d)$ we just notice that from the first part of the proof it follows that $f \mapsto \|Wf(x)\|_{\ell_2}$ is continuous on $L_p(\mathbb{R}^d)$.

Using (A.2) once more, taking the supremum for $\{f \in L_2(\mathbb{R}^d) \cap L_p(\mathbb{R}^d) : \|f\|_p \leq 1\}$ we obtain $\|g\|_{p'} \leq \tilde{C} \big\| \|\tilde{W} g(x)\|_{\ell_2} \big\|_{p'}$. □

¿From Theorem A.1 we see that the following sequence space plays an important role.

Definition A.2. Let d_p denote the set of sequences $\{c_I^\ell\}_{I\in D,\ell\in E}$ for which

$$\|\{c_I^\ell\}\|_p := \left\| \Big(\sum_{I\in D,\ell\in E} |c_I^\ell|^2 |I|^{-1} \chi_I(x) \Big)^{1/2} \right\|_p < \infty.$$

In fact, let us show that there is a stable reconstruction operator defined on d_p for bi-frame systems.

Theorem A.3. *Let $X(\Psi)$ and $X(\tilde{\Psi})$ be a bi-frame for $L_2(\mathbb{R}^d)$. Suppose for all $\psi \in \Psi \cup \tilde{\Psi}$ there exist $\beta > 0$ and $\varepsilon > 0$ such that $\psi \in C^\beta(\mathbb{R}^d)$ and $|\psi(x)| \leq C(1+|x|)^{-d-\varepsilon}$. Then the maps $T\colon d_p \mapsto L_p(\mathbb{R}^d)$ and $\tilde{T}\colon d_p \mapsto L_p(\mathbb{R}^d)$ defined by*

$$T\{c_I^\ell\} = \sum_{I\in D,\ell\in E} c_I^\ell \psi_I^\ell, \qquad \tilde{T}\{c_I^\ell\} = \sum_{I\in D,\ell\in E} c_I^\ell \tilde{\psi}_I^\ell$$

are both bounded linear maps.

Proof. We consider the operator U^ℓ with kernel

$$\tilde{K}^\ell(x,y) := \sum_{I\in D} \psi_I^\ell(x) \overline{\eta_I^1(y)}.$$

By exactly the same arguments as given in the first part of the proof of Theorem A.1, it can be shown that U^ℓ is bounded on $L_p(\mathbb{R}^d)$. Take $\{c_I^\ell\}_{I\in D,\ell\in E} \in d_p$ and consider $f^\ell := \sum_{I\in D} c_I^\ell \eta_I^1$. This is a well-defined function in $L_p(\mathbb{R}^d)$ with

$$\|f^\ell\|_p \asymp \left\| \Big(\sum_{I\in D} |c_I^\ell|^2 |I|^{-1} \chi_I(x) \Big)^{1/2} \right\|_p,$$

where we used the characterization of $L_p(\mathbb{R}^d)$ using wavelets. Thus,

$$\begin{aligned}
\Big\| \sum_{I\in D,\ell\in E} c_I^\ell \psi_I^\ell \Big\|_p &\leq \sum_{\ell\in E} \Big\| \sum_{I\in D} c_I^\ell \psi_I^\ell \Big\|_p = \sum_{\ell\in E} \|U^\ell f^\ell\|_p \\
&\leq C \sum_{\ell\in E} \|f^\ell\|_p \leq \tilde{C} \sum_{\ell\in E} \left\| \Big(\sum_{I\in D} |c_I^\ell|^2 |I|^{-1} \chi_I(x) \Big)^{1/2} \right\|_p \\
&\leq L\tilde{C} \left\| \Big(\sum_{I\in D,\ell\in E} |c_I^\ell|^2 |I|^{-1} \chi_I(x) \Big)^{1/2} \right\|_p,
\end{aligned}$$

and it follows that $T\colon d_p \mapsto L_p(\mathbb{R}^d)$ is bounded. The claim for $\tilde{T}$ is proved using similar arguments with appropriately modified operator kernels. □

Recall the Lorentz space $\ell_{p,q}(\Lambda)$, $1 \le p < \infty$, $0 < q \le \infty$, for some countable set Λ, as the set of sequences $\{a_m\}_{m\in\Lambda}$ satisfying $\|\{a_m\}\|_{\ell_{p,q}} < \infty$, where

$$\|\{a_m\}\|_{\ell_{p,q}} = \begin{cases} \left(\sum_{j=0}^{\infty}\left(2^{j/\tau}a_{2^j}^*\right)^q\right)^{1/q}, & 0 < q < \infty, \\ \sup_{j\ge 0} 2^{j/\tau}a_{2^j}^*, & q = \infty, \end{cases} \tag{A.3}$$

with $\{a_j^*\}_{j=0}^{\infty}$ a decreasing rearrangement of $\{a_m\}_{m\in\Lambda}$.

It can be shown that there exist constants $c, C > 0$ such that

$$c\|\{c_I^\ell\}\|_{\ell_{p,\infty}(D\times E)} \le \left\| \left(\sum_{I\in D,\ell\in E} |c_I^\ell|^2 |I|^{-2/p}\chi_I(x) \right)^{1/2} \right\|_p \le C\|\{c_I^\ell\}\|_{\ell_{p,1}(D\times E)}, \tag{A.4}$$

for any $\{c_I^\ell\} \in \ell_{p,1}(D\times E)$, see *e.g.* [3].

¿From the above results we easily deduce the following important corollary. As before, we denote by $\psi_I^{\ell,p}$ the function ψ_I^ℓ normalized in $L_p(\mathbb{R}^d)$, *i.e.* $\|\psi_I^{\ell,p}\|_p \asymp |I|^{1/2-1/p}\|\psi_I^\ell\|_p$.

Corollary A.4. *Let $X(\Psi)$ and $X(\tilde\Psi)$ be a wavelet bi-frame for $L_2(\mathbb{R}^d)$. Suppose for all $\psi \in \Psi \cup \tilde\Psi$ there exist $\beta > 0$ and $\varepsilon > 0$ such that $\psi \in C^\beta(\mathbb{R}^d)$ and $|\psi(x)| \le C(1+|x|)^{-d-\varepsilon}$. Then $X(\Psi)$ and $X(\tilde\Psi)$ are* $\ell_{\mathbf{p},\mathbf{1}}$**-hilbertian** *systems in $L_p(\mathbb{R}^d)$, $1 < p < \infty$, that is to say we have*

$$\Big\| \sum_{I\in D,\ell\in E} c_I^\ell \psi_I^{\ell,p} \Big\|_p \le C_p \|\{c_I^\ell\}\|_{\ell_{p,1}(D\times E)},$$

$$\Big\| \sum_{I\in D,\ell\in E} c_I^\ell \tilde\psi_I^{\ell,p} \Big\|_p \le C_p \|\{c_I^\ell\}\|_{\ell_{p,1}(D\times E)},$$

for any sequence $\{c_I^\ell\} \in \ell_{p,1}(D\times E)$.

Proof. Follows from Theorem A.3 and equation (A.4). □

This property is used to prove a Jackson inequality in Section 6.

References

[1] C. Bennett and R. Sharpley. *Interpolation of operators.* Academic Press Inc., Boston, MA, 1988.

[2] J. Bergh and J. Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[3] L. Borup, R. Gribonval, and M. Nielsen. Tight wavelet frames in Lebesgue and Sobolev spaces. *Preprint*, 2003. (Preprint available from: http://www.math.auc.dk/ research/reports/R-2003-05.ps).

[4] L. Borup and M. Nielsen. Approximation with wave packets generated by a refinable function. Technical report, Aalborg University, 2003. (Available from: http://www.math.auc.dk/ ˜mnielsen/fpackets.pdf).

[5] C. K. Chui, W. He, and J. Stöckler. Compactly supported tight and sibling frames with maximum vanishing moments. *Appl. Comput. Harmon. Anal.*, 13(3):224–262, 2002.

[6] C. K. Chui, W. He, and J. Stöckler. Tight frames with maximum vanishing moments and minimum support. In *Approximation theory, X (St. Louis, MO, 2001)*, Innov. Appl. Math., pages 187–206. Vanderbilt Univ. Press, Nashville, TN, 2002.

[7] C. K. Chui and X. L. Shi. On L^p-boundedness of affine frame operators. *Indag. Math. (N.S.)*, 4(4):431–438, 1993.

[8] I. Daubechies, B. Han, A. Ron, and Z. Shen. Framelets: MRA-based constructions of wavelet frames. *Appl. Comput. Harmon. Anal.*, 14(1):1–46, 2003.

[9] R. A. DeVore. Nonlinear approximation. In *Acta numerica, 1998*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.

[10] R. A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992.

[11] R. A. DeVore and G. G. Lorentz. *Constructive approximation*. Springer-Verlag, Berlin, 1993.

[12] G. B. Folland. *Introduction to partial differential equations*. Princeton University Press, Princeton, NJ, second edition, 1995.

[13] M. Frazier and B. Jawerth. Decomposition of Besov spaces. *Indiana Univ. Math. J.*, 34(4):777–799, 1985.

[14] R. Gribonval and M. Nielsen. Nonlinear approximation with dictionaries. I. Direct estimates. *J. Fourier Anal. Appl.*, 2002. (To appear. Preprint available from: http://www.math.auc.dk/research/reports/R-02-2018.ps).

[15] R. Gribonval and M. Nielsen. On approximation with spline generated framelets. *Constr. Approx.*, Published online July 7, 2003. (Available from: http://www.springerlink.com/link.asp?id=nj0hqqvlc1nytg9g).

[16] E. Hernández and G. Weiss. *A first course on wavelets*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1996. With a foreword by Yves Meyer.

[17] R. Q. Jia. A Bernstein-type inequality associated with wavelet decomposition. *Constr. Approx.*, 9(2-3):299–318, 1993.

[18] R.-Q. Jia. Shift-invariant spaces on the real line. *Proc. Amer. Math. Soc.*, 125(3):785–793, 1997.

[19] S. Mallat. *A wavelet tour of signal processing*. Academic Press Inc., San Diego, CA, 1998.

[20] Y. Meyer. *Wavelets and operators*, volume 37 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1992.

[21] Y. Meyer and R. Coifman. *Wavelets*. Cambridge University Press, Cambridge, 1997. Calderón-Zygmund and multilinear operators, Translated from the 1990 and 1991 French originals by David Salinger.

[22] A. Petukhov. Explicit construction of framelets. *Appl. Comput. Harmon. Anal.*, 11(2):313–327, 2001.

[23] A. Ron and Z. Shen. Affine systems in $L_2(\mathbf{R}^d)$. II. Dual systems. *J. Fourier Anal. Appl.*, 3(5):617–637, 1997. Dedicated to the memory of Richard J. Duffin.

[24] A. Ron and Z. Shen. Affine systems in $L_2(\mathbf{R}^d)$: the analysis of the analysis operator. *J. Funct. Anal.*, 148(2):408–447, 1997.

[25] H. Triebel. *The structure of functions*. Birkhäuser Verlag, Basel, 2001.

Department of Mathematical Sciences, Aalborg University, Fr. Bajers Vej 7G, DK-9220 Aalborg East, Denmark

E-mail address: lasse@math.auc.dk

IRISA-INRIA, Campus de Beaulieu, 35042 Rennes cedex, France

E-mail address: remi.gribonval@inria.fr

Department of Mathematical Sciences, Aalborg University, Fr. Bajers Vej 7G, DK-9220 Aalborg East, Denmark

E-mail address: mnielsen@math.auc.dk