

## APPROXIMATION WITH WAVE PACKETS GENERATED BY A REFINABLE FUNCTION

LASSE BORUP AND MORTEN NIELSEN

(Communicated by David R. Larson)

ABSTRACT. We consider best  $m$ -term approximation in  $L_p(\mathbb{R}^d)$  with wave packets generated by a single refinable function. The main examples of wave packets are orthonormal wavelets, or more generally wavelet frames based on a multiresolution analysis (so-called framelets). The approximation classes associated with best  $m$ -term approximation in  $L_p(\mathbb{R}^d)$  for a large class of wave packets are completely characterized in terms of Besov spaces.

As an application of the main result, we show that for  $m$ -term approximation in  $L_p(\mathbb{R}^d)$  with elements from an oversampled version of a framelet system with compactly supported generators, the associated approximation classes turn out to be (essentially) Besov spaces.

### 1. INTRODUCTION

The standard method used to generate “atoms” such as wavelets or wavelet frames is to begin with a refinable function  $\phi \in L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$  and then construct the generators  $\Psi = \{\psi^i\}_{i \in F}$  of the wavelet or wavelet frame system by applying a filter to  $\phi$ , i.e.,

$$(1.1) \quad \hat{\psi}^i(2\xi) = m^i(\xi)\hat{\phi}(\xi).$$

Usually  $m^i$  is a so-called high-pass filter, but this is not important for the results considered in this paper. We call a function  $\psi^i$  of the type given by (1.1) a wave packet generated by the refinable function  $\phi$ .

The main purpose of this paper is to study approximation of smooth functions with  $m$ -term approximants formed by dilating and translating one or more wave packets of the form given by (1.1). More precisely, given a finite collection of wave packets  $\Psi$  and  $K \geq 0$ , we consider the dictionary

$$(1.2) \quad X_K(\Psi) := \{\psi(2^j \cdot -k/2^K) : j \in \mathbb{Z}, k \in \mathbb{Z}^d, \psi \in \Psi\}.$$

We notice that  $K$  is an oversampling ratio, and for regular wavelet systems one usually considers  $K = 0$ . We write  $X(\Psi) := X_0(\Psi)$ .

---

Received by the editors July 15, 2003 and, in revised form, April 14, 2004.

2000 *Mathematics Subject Classification*. Primary 41A46; Secondary 41A17, 42C40.

*Key words and phrases*. Refinable functions, nonlinear approximation, framelet systems, Jackson inequality, Bernstein inequality, Besov spaces.

This work was supported in part by the Danish Technical Science Foundation, Grant no. 9701481.

The associated (nonlinear) space of all possible  $m$ -term expansions by elements from  $X_K(\Psi)$  is given by

$$\Sigma_m(X_K(\Psi)) := \left\{ S : S = \sum_{j=1}^m a_j g_j, \text{ with } a_j \in \mathbb{C}, g_j \in X_K(\Psi) \right\}.$$

We measure the approximation error in the Triebel-Lizorkin space  $\dot{F}_{p,q}^\beta(\mathbb{R}^d)$ ; see [14] for the definition. Suppose  $\Psi \subset \dot{F}_{p,q}^\beta(\mathbb{R}^d)$ . Then the error of the best  $m$ -term approximation of  $f \in \dot{F}_{p,q}^\beta(\mathbb{R}^d)$  from  $X_K(\Psi)$  is given by

$$\sigma_m(f, X_K(\Psi))_{\dot{F}_{p,q}^\beta} := \inf_{S \in \Sigma_m(X_K(\Psi))} \|f - S\|_{\dot{F}_{p,q}^\beta}.$$

We let  $\mathcal{A}_t^\gamma(\dot{F}_{p,q}^\beta(\mathbb{R}^d), X_K(\Psi))$ ,  $\gamma > 0$ ,  $0 < t \leq \infty$ , denote the approximation class of all functions  $f$  such that

$$(1.3) \quad |f|_{\mathcal{A}_t^\gamma(\dot{F}_{p,q}^\beta, X_K(\Psi))} := \left( \sum_{m=1}^{\infty} (m^\gamma \sigma_m(f, X_K(\Psi))_{\dot{F}_{p,q}^\beta})^t \frac{1}{m} \right)^{1/t} < \infty,$$

with the usual modification when  $t = \infty$ . Roughly speaking, the approximation class contains functions whose  $m$ -term approximation error  $\sigma_m(f, X_K(\Psi))_{\dot{F}_{p,q}^\beta}$  decays like  $O(m^{-\gamma})$ ;  $t$  is just a fine tuning parameter. Now the fundamental question is whether it is possible to completely characterize  $\mathcal{A}_t^\gamma(\dot{F}_{p,q}^\beta(\mathbb{R}^d), X_K(\Psi))$  in terms of known smoothness spaces.

Some results on characterizing  $\mathcal{A}_t^\gamma(L_p(\mathbb{R}^d), X_K(\Psi))$  are known when  $\Psi$  is a sufficiently nice wavelet system. For nice bi-orthogonal wavelet systems in  $L_2(\mathbb{R}^d)$  with parameters  $\beta = 0$ ,  $q = 2$ , and  $1 < p < \infty$  in (1.3) (we recall that  $\dot{F}_{p,2}^0(\mathbb{R}^d) \approx L_p(\mathbb{R}^d)$ ), it was proved by DeVore, Jawerth, and Popov in [4] that the approximation classes are essentially Besov spaces. A wavelet frame is a wavelet system that constitutes a frame for  $L_2(\mathbb{R}^d)$ , and it is thus a more general concept than a bi-orthogonal wavelet basis. It was proved in [8] that for tight wavelet frames based on a spline multiresolution for  $L_2(\mathbb{R})$ , the approximation classes are essentially Besov spaces provided the system is oversampled by a factor of two.

The main contribution of the present paper is to show that given one or more compactly supported wave packets of the form (1.1), based on a compactly supported refinable function, it is possible to find  $K_0 \in \mathbb{N}$  such that for  $K \geq K_0$ , the approximation class  $\mathcal{A}_t^\gamma(L_p(\mathbb{R}^d), X_K(\Psi))$  is essentially a Besov space. The approximation classes are exactly the same as those for orthonormal wavelet systems. The precise statement of the result is given in Section 3; see Theorem 3.3. For  $d \geq 2$  there is a technical restriction on the filters  $m^i$  for Theorem 3.3 to apply, but in the univariate case  $d = 1$  the existence of such a  $K_0$  does not depend at all on the particular form of the filters  $m^i$  from (1.1). In the general case, the system  $X(\Psi)$  might not be a frame for  $L_2(\mathbb{R}^d)$ ; in fact,  $X(\Psi)$  might not even be dense in  $L_2(\mathbb{R}^d)$ . However, Theorem 3.3 shows that there is an oversampled version of  $X(\Psi)$  that is suited for nonlinear approximation in  $L_p(\mathbb{R}^d)$ .

The idea of oversampling wavelet type systems to improve the stability in various function spaces was used by Frazier and Jawerth in [6, 7] for their  $\phi$ -transform. Nonlinear approximation with general (oversampled) wave packets has been considered by Kyriazis and Petrushev in [12] and by Petrushev in [13]. The main difference between the results in [12] and the present paper is that we obtain a complete characterization of  $\mathcal{A}_t^\gamma(L_p(\mathbb{R}^d), X_K(\Psi))$  for a large class of generators  $\Psi$

based on refinable functions. In [12] so-called Jackson estimates were obtained for approximation with general (oversampled) wave packets.

The structure of the paper is as follows. In Section 2 we prove a Jackson inequality for best  $m$ -term approximation in  $\dot{F}_{p,q}^\beta(\mathbb{R}^d)$  with elements from the dictionary  $X_K(\Psi)$  for sufficiently large values of  $K$ . The idea is to use a finite number of functions from  $X_K(\Psi)$  to build a stable wavelet basis for the various Triebel-Lizorkin spaces using the  $\phi$ -transform machinery by Frazier and Jawerth [7]. Then the properties of this new wavelet basis are used to derive a Jackson estimate for the system  $X_K(\Psi)$ . To obtain the complete characterization of the approximation class, we establish a Bernstein inequality for  $m$ -term expansions from  $X_K(\Psi)$ . This is done in Section 3. Put together, the Jackson and Bernstein inequalities lead to the main result, Theorem 3.3. In the last part of Section 3 similar results are considered for  $X_K(\phi)$ . Finally, we consider an application of Theorem 3.3 to framelet systems, i.e., wavelet frames based on an MRA. In Section 4 we give a characterization of the approximation classes with  $m$ -term approximation in  $L_p(\mathbb{R}^d)$ .

## 2. JACKSON INEQUALITIES FOR GENERAL WAVE PACKETS

In this section we establish a Jackson inequality for best  $m$ -term approximation with elements from the dictionary  $X_K(\Psi)$  for sufficiently large values of  $K$ . The proof consists of two steps. First we show that for large  $K$  it is possible to build a stable wavelet basis for  $\dot{F}_{p,q}^\beta(\mathbb{R}^d)$  using a finite number of functions from  $X_K(\Psi)$ . Then we set up a mapping from the Meyer wavelet system to the new wavelet system. This mapping allows us to “translate” the approximation results for the Meyer basis to the new wavelet system, and thereby obtain a Jackson estimate for  $X_K(\Psi)$ .

In order to obtain a stable wavelet basis, we will use the  $\phi$ -transform machinery of Frazier and Jawerth. In [7] Frazier and Jawerth give sufficient conditions for a function  $u$  to generate an atomic decomposition of the Triebel-Lizorkin and Besov spaces. Two of the conditions are

$$(2.1) \quad |\hat{u}(\xi)| \geq c > 0 \quad \text{if } 2^{-1} \leq |\xi| \leq 2,$$

and

$$(2.2) \quad \int x^\gamma u(x) dx = 0 \quad \text{if } |\gamma| \leq N,$$

for sufficiently large  $N \in \mathbb{N}$ . However, many interesting functions do not satisfy the conditions (2.1) and (2.2). For example, whenever  $\int u \neq 0$ , then (2.2) clearly fails, but in the paper [12] it was shown how to build appropriate wave packets from such functions with the required number of vanishing moments. The wave packets then form atomic decompositions of the Triebel-Lizorkin and Besov spaces. Another type of function often encountered in applications is one that is very smooth, but with a comparatively small number of vanishing moments. Often generators of tight wavelet frames will have these characteristics. We can apply the results in [7] directly to such functions, but we will only get atomic decompositions valid for a more restricted range of smoothness parameters than one would expect from the smoothness of the generator. Inspired by [12], we show in Lemma 2.1 below how to create wave packets from a function which does not satisfy the conditions (2.1) and (2.2).

Let us introduce the following notation. Let  $D$  denote the set of dyadic cubes  $Q = Q_{\nu k} = 2^{-\nu}([0, 1]^d + k)$ ,  $\nu \in \mathbb{Z}$ ,  $k \in \mathbb{Z}^d$ . Following the notation in [7] we will let  $\{\phi^Q\}_{Q \in D}$  denote a sequence of functions indexed by the dyadic cubes while  $\phi_Q(x) := 2^{-\nu d/2} \phi(2^\nu x - k)$ . We let  $\dot{f}_{p,q}^\beta(D)$  and  $\dot{b}_{p,q}^\beta(D)$ , denote the discrete Triebel-Lizorkin and Besov spaces, respectively; see [7] for the definition.

**Lemma 2.1.** *Given  $s \geq 0$ , let  $\phi \in C^s(\mathbb{R}^d)$  be a compactly supported function with  $\hat{\phi}(0) \neq 0$  and let  $m$  be a trigonometric polynomial satisfying*

$$(2.3) \quad m(\xi) \neq 0 \quad \text{for } 0 < |\xi| < r$$

for some  $r > 0$ . Suppose  $\psi$  is a wave packet satisfying  $\hat{\psi}(2\xi) = m(\xi)\hat{\phi}(\xi)$ . Then for  $N \in \mathbb{N}_0$  there exists  $J \in \mathbb{N}$  and a finite set of coefficients  $\{c_k\}_{k=1}^n$ ,  $n \leq 2N + 1$ , such that the function  $u = \sum_{k=1}^n c_k \psi(2^J(\cdot - k))$  satisfies (2.1) and (2.2).

Moreover, suppose  $0 < q \leq \infty$  and  $0 < p < \infty$  satisfy  $\min(p, q) \geq d/(N + d + s)$ . Then there is a  $K_0 \in \mathbb{Z}$  such that if  $K \geq K_0$  and  $g := u(2^{-K}\cdot)$ , there exists a family of functions  $\{\tilde{g}^Q\}_{Q \in D}$  such that

$$f = \sum_{Q \in D} \langle f, \tilde{g}^Q \rangle g_Q,$$

for all  $f \in \dot{B}_{p,q}^s(\mathbb{R}^d)$ , with  $\|\langle f, \tilde{g}^Q \rangle\|_{\dot{b}_{p,q}^s} \leq c \|f\|_{\dot{B}_{p,q}^s}$ , and for any sequence  $s = \{s_Q\}_{Q \in D}$ , we have

$$\left\| \sum_{Q \in D} s_Q g_Q \right\|_{\dot{B}_{p,q}^s} \leq c \|s\|_{\dot{b}_{p,q}^s}.$$

*Proof.* Since  $\phi$  is continuous there exists an  $r' > 0$  such that  $\hat{\phi}(\xi) \neq 0$  for  $|\xi| < r'$ . If we choose  $J \in \mathbb{N}$  such that  $2^J > 2/\min(r, r')$ , then  $\tilde{u} = \psi(2^J\cdot)$  satisfies (2.1).

In order to obtain  $N$  vanishing moments, we simply apply a high-pass filter with symbol  $p(\xi)$  which has a zero of order  $N$  at  $\xi = 0$  and no zeros on  $K := \{\xi \in \mathbb{R}^d : 1/2 \leq |\xi| \leq 2\}$ . For example, let  $u = (\sum_{j=1}^d (\Delta_{e_j} + \Delta_{-e_j}))^N \tilde{u}$ , where  $\Delta_e$  is the difference operator in the direction  $e \in \mathbb{R}^d$ ,  $\Delta_e f(x) = f(x+e) - f(x)$ , and  $e_j$  is a unit vector in the  $j$ -th direction. Then it is easy to see that  $u$  satisfies (2.2). Moreover, since  $\mathcal{F}(\Delta_{e_j} + \Delta_{-e_j})f(\xi) = 2(\cos(\xi \cdot e_j) - 1)\hat{f}$  and  $|\sum_{j=1}^d \cos(\xi \cdot e_j) - 1| \geq 1 - \cos(1/2)$  on  $K$ ,  $u$  satisfies (2.1) too.

Since  $\phi$  has compact support and  $m$  is a trigonometric polynomial,  $u$  is compactly supported. Thus, the second part of the lemma follows from [7, Section 4].  $\square$

*Remark 2.2.* For  $d = 1$ , (2.3) is satisfied by any non-trivial trigonometric polynomial  $m$ , since its zeroes are isolated. For  $d > 1$  the condition is more restrictive. In particular, (2.3) is not satisfied for separable wavelet systems. However, in this case a slightly different condition suffices; see Remark 4.3.

As a consequence of Lemma 2.1 we have the following Jackson inequality.

**Proposition 2.3.** *Given  $s \geq 0$ , let  $\phi \in C^s(\mathbb{R}^d)$  be a compactly supported refinable function and  $m$  a trigonometric polynomial satisfying (2.3). Let  $\psi$  be the associated wave packet. Suppose  $0 < p < \infty$ ,  $0 < t \leq \infty$ ,  $0 \leq \beta < \gamma < s$  and  $\psi \in \dot{F}_{p,t}^\beta(\mathbb{R}^d)$ , and define  $1/\tau := (\gamma - \beta)/d + 1/p$ . Then there exists a finite constant  $C$  such that*

$$\sigma_m(f, X_K(\psi))_{\dot{F}_{p,t}^\beta} \leq C m^{-(\gamma-\beta)/d} \|f\|_{\dot{B}_{\tau,\tau}^\gamma}, \quad \forall m \in \mathbb{N}, f \in \dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d).$$

*Proof.* By Lemma 2.1, we can construct a function  $g \in \Sigma_n(X_K(\psi))$  for some finite  $n$ , and a sequence  $\{\tilde{g}^Q\}_{Q \in D}$ , such that

$$f = \sum_{Q \in D} \langle f, \tilde{g}^Q \rangle g_Q,$$

for every  $f \in \dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d)$  with  $\|\langle f, \tilde{g}^Q \rangle\|_{i_{\tau,\tau}^\gamma} \leq C\|f\|_{\dot{B}_{\tau,\tau}^\gamma}$ . Fix  $f \in \dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d)$ , and let  $\eta$  be one of the  $2^d - 1$  orthonormal Meyer wavelets on  $\mathbb{R}^d$ . We notice that the function  $h := \sum_{Q \in D} \langle f, \tilde{g}^Q \rangle \eta_Q$  belongs to  $\dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d)$  with  $\|h\|_{\dot{B}_{\tau,\tau}^\gamma} \leq C\|f\|_{\dot{B}_{\tau,\tau}^\gamma}$ . From [11], we have

$$(2.4) \quad \sigma_m(h, \{\eta_Q\}_Q)_{\dot{F}_{p,t}^\beta} \leq Cm^{-(\gamma-\beta)/d} \|h\|_{\dot{B}_{\tau,\tau}^\gamma}, \quad m \in \mathbb{N}.$$

Let  $h_m \in \Sigma_m(\{\eta_Q\}_Q)$  be a sequence that realizes (2.4) up to the relaxed constant  $2C$ . We want to map  $h_m$  to an element of  $f_m \in \Sigma_m(\{g_Q\}_Q)$ . To accomplish this, we consider the operator  $T$  with kernel

$$K(x, y) := \sum_{Q \in D} g_Q(x) \eta_Q(y).$$

The matrix of this operator in the Meyer wavelet basis is  $M = [\langle g_P, \eta_Q \rangle]_{P,Q \in D}$ . It is easy to verify that  $M \in \mathbf{ad}_{p,t}^\beta$ , where  $\mathbf{ad}_{p,t}^\beta$  is the algebra of almost diagonal matrices; see [7, Sec. 9]. We notice that  $f_m := Th_m \in \Sigma_m(\{g_Q\}_Q)$ , and moreover,  $f - f_m = T(h - h_m)$ . However, the matrix representation of  $T$  shows that  $T$  is bounded on  $\dot{F}_{p,t}^\beta(\mathbb{R}^d)$ , so

$$\begin{aligned} \|f - f_m\|_{\dot{F}_{p,t}^\beta} &\leq C_1 \|h - h_m\|_{\dot{F}_{p,t}^\beta} \\ &\leq 2C_1 Cm^{-(\gamma-\beta)/d} \|h\|_{\dot{B}_{\tau,\tau}^\gamma} \\ &\leq C_2 m^{-(\gamma-\beta)/d} \|f\|_{\dot{B}_{\tau,\tau}^\gamma}. \end{aligned}$$

Hence, we have the wanted Jackson inequality for the dictionary  $\{g_Q\}_{Q \in D}$ , and thus for  $X_K(\psi)$ .  $\square$

*Remark 2.4.* If  $\psi \in \Psi$  for some finite set  $\Psi \subset \dot{F}_{p,t}^\beta(\mathbb{R}^d)$ , we immediately obtain a Jackson inequality for the dictionary  $X_K(\Psi)$ .

### 3. APPROXIMATION WITH WAVE PACKETS

In this section we derive a fairly general Bernstein inequality for the system  $X_K(\Psi)$ . Then using the Jackson inequality for such systems, valid for large  $K$ , we give a complete characterization of  $\mathcal{A}_q^{\alpha/d}(L_p(\mathbb{R}^d), X_K(\Psi))$  for  $1 < p < \infty$ . The Bernstein inequality for  $X_K(\Psi)$  is given by the following proposition. We refer the reader to [9] for the definition of locally linearly independent sets.

**Proposition 3.1.** *Given  $s \geq 0$ , let  $\phi \in W^s(L_\infty(\mathbb{R}^d))$  be a compactly supported refinable function with associated finite mask and let  $\Psi = \{\psi^i\}_{i \in F}$  be a finite sequence of wave packets generated by (1.1) using trigonometric polynomials  $m^i$ . If  $d > 1$  we suppose  $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}^d}$  is a locally linearly independent set (this condition is void if  $d = 1$ ). Then for each  $0 < \gamma < s$  and  $K \geq 0$ , the Bernstein inequality*

$$(3.1) \quad |S|_{\dot{B}_{\tau,\tau}^\gamma} \leq Cm^{\gamma/d} \|S\|_{L_p}, \quad \forall S \in \Sigma_m(X_K(\Psi)),$$

$1/\tau := \gamma/d + 1/p$ ,  $0 < p \leq \infty$ , holds true.

*Proof.* In the case  $d = 1$ , if the integer shifts of the function  $\phi$  are not already linearly independent, we can always find a perfect generator  $\tilde{\phi}$  for the shift-invariant space  $S_0 := \text{span}\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ , i.e.,  $\tilde{\phi}$  is a compactly supported refinable function with linearly independent shifts that generates  $S_0$ ; see [10]. In the arguments below we may use  $\tilde{\phi}$  in place of  $\phi$ .

By the result of Jia [9], for each  $0 < \gamma < s$ , the Bernstein inequality

$$|S|_{\dot{B}^{\gamma,\tau}} \leq Cm^{\gamma/d} \|S\|_{L_p}, \quad \forall S \in \Sigma_m(X(\phi)),$$

$1/\tau := \gamma/d + 1/p$ ,  $0 < p \leq \infty$ , holds true.

For  $\psi \in \Psi$  there is a finite mask  $\{b_\ell\}_\ell$  such that

$$\psi(x) = \sum_{\ell \in \mathbb{Z}^d} b_\ell \phi(2x - \ell),$$

and by assumption there is another finite mask  $\{a_\ell\}_\ell$  such that

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

Thus, for  $K \geq 0$ ,  $j \in \mathbb{Z}$  and  $k \in \mathbb{Z}^d$ , we have

$$\begin{aligned} &\psi(2^j x - k/2^K) \\ &= \sum_{\ell_1 \in \mathbb{Z}^d} b_{\ell_1} \phi(2^{j+1} x - k/2^{K-1} - \ell_1) \\ &= \sum_{\ell_1, \ell_2 \in \mathbb{Z}^d} b_{\ell_1} a_{\ell_2} \phi(2^{j+2} x - k/2^{K-2} - 2\ell_1 - \ell_2) \\ &\quad \vdots \\ &= \sum_{\ell_1, \ell_2, \dots, \ell_K \in \mathbb{Z}^d} b_{\ell_1} a_{\ell_2} \cdots a_{\ell_K} \phi(2^{j+K} x - k - 2^{K-1} \ell_1 - 2^{K-2} \ell_2 - \cdots - \ell_K). \end{aligned}$$

That is to say,  $\psi(2^j x - k/2^K) \in \Sigma_L(X(\phi))$  for some uniform  $L$  depending only on  $K$  and the length of the finite masks used above. Take any  $S \in \Sigma_m(X_K(\Psi))$ . Then  $S \in \Sigma_{Mm}(X(\phi))$  for some uniform constant  $M$  (depending on  $K$  and the length of the finite masks). Using the Bernstein inequality for  $X(\phi)$ , we obtain the wanted inequality,

$$\begin{aligned} |S|_{\dot{B}^{\gamma,\tau}} &\leq C(Mm)^{\gamma/d} \|S\|_{L_p} \\ &\leq \tilde{C}m^{\gamma/d} \|S\|_{L_p}, \quad \forall S \in \Sigma_m(X_K(\Psi)). \end{aligned}$$

□

*Remark 3.2.* Notice that the inequality (3.1) holds for all  $S \in \Sigma_m(X_K(\psi))$ , where  $\psi$  is any single function from  $\Psi$ .

We can now apply the Jackson and Bernstein inequalities obtained so far to get the main result of this paper: that we can characterize the approximation class by an interpolation space; see [1] for the definition of the real method of interpolation.

**Theorem 3.3.** *Given  $s \geq 0$ , let  $\phi \in W^s(L_\infty(\mathbb{R}^d))$  be a compactly supported refinable function with associated finite mask and let  $\Psi = \{\psi^i\}_{i \in F}$  be a finite sequence of wave packets with associated trigonometric polynomials, where at least one, say*

$m^{i_0}$ , satisfies (2.3). If  $d > 1$  we suppose  $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}^d}$  is a locally linearly independent set (condition is void if  $d = 1$ ). Then there exists  $K_0 \in \mathbb{N}_0$  such that for  $K \geq K_0$ ,

$$\mathcal{A}_q^{\alpha/d}(L_p(\mathbb{R}^d), X_K(\Psi)) = \mathcal{A}_q^{\alpha/d}(L_p(\mathbb{R}^d), X_K(\psi^{i_0})) = (L_p(\mathbb{R}^d), B_{\tau,\tau}^\gamma(\mathbb{R}^d))_{\alpha/\gamma,q},$$

for  $1 < p < \infty$ ,  $0 < \alpha < \gamma < s$ , and  $1/\tau := \gamma/d + 1/p$ .

*Proof.* By Proposition 3.1, we have the Bernstein inequality

$$|S|_{\dot{B}_{\tau,\tau}^\gamma} \leq \tilde{C}_K m^{\gamma/d} \|S\|_{L_p}, \quad \forall S \in \Sigma_m(X_K(\Psi)), \quad \gamma < s.$$

Using Proposition 2.3, there is a  $K \geq 0$  such that the Jackson inequality

$$\sigma_m(f, X_K(\psi^{i_0}))_{L_p} \leq C m^{-\gamma/d} \|f\|_{\dot{B}_{\tau,\tau}^\gamma}, \quad \forall m \in \mathbb{N}, f \in \dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d),$$

holds for  $\gamma < s$ . Hence, the result follows by the well-known theorem of DeVore and Popov [5].  $\square$

*Remark 3.4.* For  $p \leq 1$ , we cannot get a Bernstein inequality valid for  $H_p(\mathbb{R}^d)$  using the techniques of this paper. This is due to the fact that the refinable function  $\phi$  is not contained in  $H_p(\mathbb{R}^d)$ . Indeed, whether  $\mathcal{A}_q^{\alpha/d}(H_p(\mathbb{R}^d), X_K(\psi))$  is even well-defined depends on the properties of the filters  $m^i$  from (1.1).

Notice that Theorem 3.3 applies to the wave packet  $\psi(x) := \phi(2x)$ , since the function  $m(\xi) \equiv 1$  satisfies (2.3). Thus, Theorem 3.3 also characterizes the approximation classes associated with the oversampled scaling system  $X_K(\phi)$ . However, since  $\phi$  is a refinable function we can actually do much better. We have the following result.

**Proposition 3.5.** *Consider an oversampled system  $X_K(\phi)$  for which the following Jackson inequality holds:*

$$\sigma_m(f, X_K(\phi))_{\dot{F}_{p,t}^\beta} \leq C m^{-(\gamma-\beta)/d} \|f\|_{\dot{B}_{\tau,\tau}^\gamma}, \quad \forall m \in \mathbb{N}, f \in \dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d),$$

where  $1/\tau := (\gamma - \beta)/d + 1/p$  and  $0 < t \leq \infty$ . Suppose  $\phi$  is refinable and compactly supported with an associate finite mask. Then the Jackson inequality (with the same parameters as above)

$$\sigma_m(f, X(\phi))_{\dot{F}_{p,t}^\beta} \leq C m^{-(\gamma-\beta)/d} \|f\|_{\dot{B}_{\tau,\tau}^\gamma}, \quad \forall m \in \mathbb{N}, f \in \dot{B}_{\tau,\tau}^\gamma(\mathbb{R}^d),$$

holds for the non-oversampled system  $X(\phi)$ .

*Proof.* We notice that it suffices to prove that  $\Sigma_m(X_K(\phi)) \subset \Sigma_{Lm}(X(\phi))$  for some uniform constant  $L$ , since then we would obtain the estimate

$$\sigma_{Lm}(f, X(\phi))_{\dot{F}_{p,t}^\beta} \leq \sigma_m(f, X_K(\phi))_{\dot{F}_{p,t}^\beta}.$$

Using the refinement relation

$$\phi(x) = \sum_{\ell \in \mathbb{Z}^d} a_\ell \phi(2x - \ell)$$

successively, we notice that

$$\begin{aligned} & \phi(2^j x - k/2^K) \\ &= \sum_{\ell_1 \in \mathbb{Z}^d} a_{\ell_1} \phi(2^{j+1} x - k/2^{K-1} - \ell_1) \\ &= \sum_{\ell_1, \ell_2 \in \mathbb{Z}^d} a_{\ell_1} a_{\ell_2} \phi(2^{j+2} x - k/2^{K-2} - 2\ell_1 - \ell_2) \\ & \quad \vdots \\ &= \sum_{\ell_1, \ell_2, \dots, \ell_K \in \mathbb{Z}^d} a_{\ell_1} a_{\ell_2} \cdots a_{\ell_K} \phi(2^{j+K} x - k - 2^{K-1} \ell_1 - 2^{K-2} \ell_2 - \cdots - \ell_K). \end{aligned}$$

It follows that  $\phi(2^j x - k/2^K) \in \Sigma_L(X(\phi))$ , where  $L$  only depends on the length of the mask  $\{a_k\}_k$  and on  $K$ . This proves the claim.  $\square$

We easily deduce the following result.

**Corollary 3.6.** *Let  $\phi \in W^s(L_\infty(\mathbb{R}^d))$ ,  $s > 0$ , be a compactly supported refinable function with associated finite mask. Then we have the Jackson inequality*

$$(3.2) \quad \sigma_m(f, X(\phi))_{\dot{F}_{p,t}^{\beta}} \leq C m^{-(\gamma-\beta)/d} \|f\|_{\dot{B}_{\tau,\tau}^{\gamma}}, \quad \forall m \in \mathbb{N}, f \in \dot{B}_{\tau,\tau}^{\gamma}(\mathbb{R}^d),$$

where  $1 < p < \infty$ ,  $0 < \gamma < s$ , and  $1/\tau := (\gamma - \beta)/d + 1/p$ .

Moreover, if for  $d > 1$ ,  $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}^d}$  is a locally linearly independent set (the condition is void if  $d = 1$ ), then we have for each  $K \geq 0$ ,

$$\mathcal{A}_q^{\alpha/d}(L_p(\mathbb{R}^d), X_K(\phi)) = (L_p(\mathbb{R}^d), B_{\tau,\tau}^{\gamma}(\mathbb{R}^d))_{\alpha/\gamma,q},$$

for  $1 < p < \infty$ ,  $0 < \alpha < \gamma < s$ , and  $1/\tau := \gamma/d + 1/p$ .

*Proof.* The Jackson inequality for  $X_K(\phi)$  (for large  $K$ ) follows from Proposition 2.3 (put  $m(\xi) \equiv 1$ ), and then the Jackson inequality for smaller  $K$  can be deduced from Proposition 3.5. The Bernstein inequality for  $X_K(\phi)$  follows from the same string of arguments as used in the proof of Proposition 3.1.  $\square$

*Remark 3.7.* The Jackson inequality (3.2) given by Corollary 3.6 generalizes [4, Theorem 2.1] to the case where the approximation errors are measured on the full scale of Triebel-Lizorkin spaces and not just in the  $L_p$ -spaces.

#### 4. APPROXIMATION WITH FRAMELET DICTIONARIES

We conclude this paper by considering one application of the results to wavelet frames for  $L_2(\mathbb{R}^d)$  associated with a multiresolution analysis (MRA). Such systems are also called **framelets**. We refer to [2, 3] for a more detailed discussion of the construction of framelets. For simplicity, we only consider tight wavelet frames here, but the results below apply to bi-frames as well (see [3]).

Let  $m = (m_0, m^1, \dots, m^L)$  be a vector of  $2\pi\mathbb{Z}^d$ -periodic measurable functions with  $m_0$  the mask of a refinable scaling function  $\phi$  of an MRA  $\{V_j\}_{j \in \mathbb{Z}}$ . We assume that  $\phi$  is uniformly bounded, has compact support, and that  $\{\phi(\cdot - k)\}_{k \in \mathbb{Z}^d}$  is a Riesz basis for the scaling space  $V_0$  of the MRA. We associate the ‘‘wavelets’’  $\Psi = \{\psi^\ell\}_{\ell \in E}$  to  $m$  by letting  $\widehat{\psi^\ell}(2\xi) = m^\ell(\xi)\widehat{\phi}(\xi)$ .

The following theorem is the fundamental tool for constructing framelets. Theorem 4.1 can be stated in a slightly more general way, see [3], but this will not be needed here.

**Theorem 4.1** (The Oblique Extension Principle (OEP) [3]). *Suppose there exists a  $2\pi\mathbb{Z}^d$ -periodic function  $\Theta$  that is non-negative, essentially bounded, and continuous at the origin with  $\Theta(0) = 1$ . If for every  $\xi \in [-\pi, \pi]^d$  and  $\nu \in \{0, \pi\}^d$ ,*

$$(4.1) \quad \Theta(2\xi)m_0(\xi)\overline{m_0(\xi + \nu)} + \sum_{\ell=1}^L m^\ell(\xi)\overline{m^\ell(\xi + \nu)} = \begin{cases} \Theta(\xi), & \nu = 0, \\ 0, & \text{otherwise,} \end{cases}$$

then the wavelet system  $\{2^{jd/2}\psi^\ell(2^j \cdot -k) \mid j \in \mathbb{Z}, k \in \mathbb{Z}^d, \ell \in E\}$  defined by  $m$  is a tight wavelet frame for  $L_2(\mathbb{R}^d)$ .

We can apply Theorem 3.3 directly to such systems  $\Psi$  for  $d = 1$ , and for  $d > 1$  we can use that  $\Psi$  satisfies (4.1). We state the result as a proposition.

**Proposition 4.2.** *Let  $\Psi$  be a framelet system in  $L_2(\mathbb{R}^d)$  associated with the vector  $m = (m_0, m^1, \dots, m^L)$  of trigonometric polynomials and the  $2\pi\mathbb{Z}^d$ -periodic function  $\Theta$  satisfying (4.1). Assume that the associated scaling function  $\phi \in W^s(L_\infty(\mathbb{R}^d))$  for some  $s > 0$ , and that the integer translates of  $\phi$  generate a Riesz basis. In case  $d > 1$ , assume in addition that there exists  $\varepsilon > 0$  such that  $\Theta(\xi) - \Theta(2\xi)|m_0(\xi)|^2 \neq 0$  for  $0 < |\xi| < \varepsilon$ . Then there exists  $K_0 \geq 1$  such that for  $K \geq K_0$ ,*

$$(4.2) \quad \mathcal{A}_q^{\alpha/d}(L_p(\mathbb{R}^d), X_K(\Psi)) = (L_p(\mathbb{R}^d), B_{\tau, \tau}^\gamma(\mathbb{R}^d))_{\alpha/\gamma, q},$$

for  $1 < p < \infty$ ,  $0 < \alpha < \gamma < s$ , and  $1/\tau := \gamma/d + 1/p$ .

*Proof.* For  $d = 1$ , the result follows from Theorem 3.3 and Remark 2.2. Now we consider the case  $d > 1$ . Using (4.1) with  $\nu = 0$ , we notice that the trigonometric polynomial  $M$  given by

$$M(\xi) = \sum_{i=1}^L \overline{m^i(\xi)}m^i(\xi) = \Theta(\xi) - \Theta(2\xi)|m_0(\xi)|^2$$

satisfies  $M(\xi) \neq 0$  for  $0 < |\xi| < \varepsilon$ . We define the function  $g$  by  $\hat{g}(2\xi) = M(\xi)\hat{\phi}(\xi) = \sum_{i=1}^L \overline{m^i(\xi)}\hat{\psi}^i(2\xi)$ . Notice that  $g \in \Sigma_L(X_1(\Psi))$  for some finite constant  $L$ . For  $1 < p < \infty$ ,  $0 < \alpha < \gamma < s$ , and  $1/\tau := \gamma/d + 1/p$ , the Jackson estimate from Proposition 2.3 holds for  $X_{\tilde{K}}(g)$  for large  $\tilde{K}$ . It follows that there is a Jackson estimate for  $X_{\tilde{K}+1}(\Psi)$  for large  $\tilde{K}$ . There is a Bernstein inequality for  $X_K(\Psi)$  for any  $K \geq 0$  by Proposition 3.1, so we may conclude using the same argument as in the proof of Theorem 3.3.  $\square$

*Remark 4.3.* A special case to which Proposition 4.2 applies is orthonormal tensor product wavelets in  $\mathbb{R}^d$ ,  $d \geq 2$ . The filters associated with a tensor product wavelet basis  $\Psi$  satisfy (4.1) with  $\Theta \equiv 1$ , and the low-pass filter for the underlying separable MRA is given by  $m_0(\xi_1)m_0(\xi_2) \cdots m_0(\xi_d)$ . The low-pass filter always satisfies  $1 - |m_0(\xi_1)m_0(\xi_2) \cdots m_0(\xi_d)|^2 \neq 0$  on some annulus  $0 < |\xi| < \varepsilon$  provided that the univariate orthonormal filter  $m_0$  is a trigonometric polynomial.

## REFERENCES

- [1] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223. MR0482275 (58:2349)
- [2] 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. MR1942743 (2004a:94011)
- [3] 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. MR1971300 (2004a:42046)
- [4] R. A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992. MR1175690 (94a:42045)
- [5] R. A. DeVore and V. A. Popov. Interpolation spaces and nonlinear approximation. In *Function spaces and applications (Lund, 1986)*, volume 1302 of *Lecture Notes in Math.*, pages 191–205. Springer, Berlin, 1988. MR0942269 (89d:41035)
- [6] M. Frazier and B. Jawerth. Decomposition of Besov spaces. *Indiana Univ. Math. J.*, 34(4):777–799, 1985. MR0808825 (87h:46083)
- [7] M. Frazier and B. Jawerth. A discrete transform and decompositions of distribution spaces. *J. Funct. Anal.*, 93(1):34–170, 1990. MR1070037 (92a:46042)
- [8] R. Gribonval and M. Nielsen. On approximation with spline generated framelets. *Constr. Approx.*, 20(2):207–232, 2004. MR2036641
- [9] R. Q. Jia. A Bernstein-type inequality associated with wavelet decomposition. *Constr. Approx.*, 9(2-3):299–318, 1993. MR1215774 (94h:41026)
- [10] R.-Q. Jia. Shift-invariant spaces on the real line. *Proc. Amer. Math. Soc.*, 125(3):785–793, 1997. MR1350950 (97e:41039)
- [11] G. Kyriazis. Non-linear approximation and interpolation spaces. *J. Approx. Theory*, 113(1):110–126, 2001. MR1866250 (2002h:41036)
- [12] G. Kyriazis and P. Petrushev. New bases for Triebel-Lizorkin and Besov spaces. *Trans. Amer. Math. Soc.*, 354(2):749–776 (electronic), 2002. MR1862566 (2002k:46082)
- [13] P. Petrushev. Bases consisting of rational functions of uniformly bounded degrees or more general functions. *J. Funct. Anal.*, 174(1):18–75, 2000. MR1761363 (2001k:46016)
- [14] H. Triebel. *The structure of functions*. Birkhäuser Verlag, Basel, 2001. MR1851996 (2002k:46087)

DEPARTMENT OF MATHEMATICAL SCIENCES, AALBORG UNIVERSITY, FREDRIK BAJERS VEJ 7G,  
 DK-9220 AALBORG EAST, DENMARK  
*E-mail address:* `lasse@math.auc.dk`

DEPARTMENT OF MATHEMATICAL SCIENCES, AALBORG UNIVERSITY, FREDRIK BAJERS VEJ 7G,  
 DK-9220 AALBORG EAST, DENMARK  
*E-mail address:* `mnielsen@math.auc.dk`