Sharp Jackson and Bernstein Inequalities for N-term Approximation in Sequence Spaces with Applications

GUSTAVO GARRIGÓS & EUGENIO HERNÁNDEZ

ABSTRACT. We study N-term approximation for general families of sequence spaces, establishing sharp versions of Jackson and Bernstein inequalities. The sequence spaces used are adapted to provide characterizations of Triebel-Lizorkin and Besov spaces by means of wavelet-like systems using general dilation matrices, and thus they include spaces of anisotropic smoothness. As an application, we characterize the N-term approximation spaces when the error is measured in the first of the spaces mentioned above.

1. INTRODUCTION

In recent years, several methods involving non-linear approximation have been developed in the context of image compression and the numerical resolution of partial differential equations. Particularly successful are the approximation methods in terms of wavelet bases [6, 7, 10, 27, 32]. One of the advantages of wavelets is the simple characterization they provide for many classical function spaces. The norms in these spaces, say $\|\cdot\|_{\mathcal{F}}$, can be expressed as weighted sums of the wavelet coefficients, which for the purposes of compression considerably simplifies both, the numerical algorithms and the theoretical machinery behind them.

A particular instance is the characterization of the so-called *N-term approximation spaces* associated with a given space $\mathcal{F}$. These, usually denoted by $A_q^{\gamma}(\mathcal{F})$, consist of functions $f \in \mathcal{F}$ for which the error of approximation with just N basis coefficients:

$$\sigma_N(f)_{\mathcal{F}} := \inf\{\|f - S\|_{\mathcal{F}} : S \text{ has at most } N \text{ non-null basis coefficients}\},$$

has a rate of decay quantified by:

$$\Big[\sum_{N=1}^{\infty} (N^{\gamma}\sigma_N(f))^q \frac{1}{N}\Big]^{1/q} < \infty, \quad \text{for fixed } \gamma > 0 \text{ and } 0 < q \le \infty.$$

A well-known procedure in approximation theory places the crucial point in proving two *Jackson and Bernstein inequalities* of the form

$$\sigma_N(f)_{\mathcal{F}} \le CN^{-\varepsilon}\|f\|_{\mathcal{B}} \quad \text{and} \quad \|S\|_{\mathcal{B}} \le CN^{\varepsilon}\|S\|_{\mathcal{F}},$$

for a suitable function space $\mathcal{B}$ and some $\varepsilon > 0$ (see, e.g., [10] or [7]). The use of wavelets considerably simplifies this task, reducing matters to prove the corresponding inequalities in respective *sequence spaces* $\mathfrak{f}$ and $\mathfrak{b}$.

In this paper we shall establish sharp versions of such Jackson and Bernstein inequalities for quite general families of sequence spaces. The purpose is to provide a wide variety of choices for the norms measuring the error of approximation, as well as introducing a general setting of sequence spaces which can be applied to wavelet-like systems not necessarily of dyadic type.

We recall that Jackson inequalities related with N-term wavelet approximation were first studied in [9], with errors measured in L^p-norms, $1 < p < \infty$. Later on, it was proposed by other authors to measure errors with Besov norms $B^s_{p,p}$, or Sobolev norms H^s_p (see, e.g., [7, Chapter 4] or [20]). In these last cases one would expect that, once the size of the error is fixed, a best approximation S will capture better the smoothness of the original function f, while the number of coefficients which have to be used must be quantified. In this sense, Jackson and Bernstein inequalities reflect the precise interplay between quality and size of the approximation for a given pair of smoothness spaces $\mathcal{F}$ and $\mathcal{B}$.

For the results in this paper we shall use the sequence spaces $\mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^{\alpha}_{\tau,q}$ introduced by Frazier and Jawerth [14, 15], with a further generalization which allows anisotropic situations [13]. We shall prove sharp versions of the Jackson and Bernstein inequalities measured in the first of these norms, and introduce a transference principle leading to applications to N-term approximation in classical smoothness spaces.

We remark that our transference principle is general enough to be applied to many wavelet-like systems, including those whose dilations are given by powers of integer matrices [3]. As we shall see, the use of such wavelet bases leads directly to *anisotropic smoothness spaces*. Wavelet characterizations for these spaces, as well as various applications, have been recently presented in [16, 19, 24, 25], although a complete study involving non-linear approximation still remains to be done.

Finally, we mention that only recently we became aware of the work of Kyriazis [23], where some of these Jackson and Bernstein inequalities have been obtained for the special case of *isotropic homogeneous* Besov and Triebel-Lizorkin spaces. We also mention a very general setting for non-linear approximation in Banach spaces introduced by Temlyakov et al. [21, 22, 32], from which some of our results will be derived.

2. N-term Approximation in Sequence Spaces

We shall deal with a general setting of N-term approximation which can be described as follows. Consider a vector space $\mathfrak{f}$ consisting of sequences of complex numbers $\mathbf{s} = \{s_I\}_{I\in\mathcal{I}}$ defined over a fixed (countable) set of indices $\mathcal{I}$. For each $I \in \mathcal{I}$, we denote by $\mathbf{e}_I$ the sequence with entry 1 at I and 0 otherwise. Throughout this paper we assume that $\mathfrak{f}$ is a *quasi-Banach space*, endowed with a quasi-norm $\|\cdot\|_{\mathfrak{f}}$ satisfying the following properties:

(a) The set of all finite linear combinations of $\mathbf{e}_I$'s is contained in $\mathfrak{f}$.
(b) Monotonicity: if $\mathbf{t} \in \mathfrak{f}$ and $|s_I| \le |t_I|$, $\forall\, I \in \mathcal{I}$, then $\mathbf{s} \in \mathfrak{f}$ and $\|\{s_I\}\|_{\mathfrak{f}} \le \|\{t_I\}\|_{\mathfrak{f}}$;
(c) If $\mathbf{s} \in \mathfrak{f}$, then $\lim_{n\to\infty} \|s_{I_n}\mathbf{e}_{I_n}\|_{\mathfrak{f}} = 0$, for some enumeration $\mathcal{I} = \{I_1, I_2, \dots\}$.

It is easy to see that (b) implies $\|\{s_I\}\|_{\mathfrak{f}} = \|\{|s_I|\}\|_{\mathfrak{f}}$ for all $\mathbf{s} \in \mathfrak{f}$. Particular examples will be spaces $\mathfrak{f}$ for which $\{\mathbf{e}_I\}$ is an unconditional basis of $\mathfrak{f}$, and also some other situations which we present in Section 3 below.

Let Σ_N denote the set of all sequences with at most N non-null entries. Given $\mathbf{s} \in \mathfrak{f}$, we define the *$N$-term error of approximation to* $\mathbf{s}$ *in* $\mathfrak{f}$ by:

$$\sigma_N(\mathbf{s})_{\mathfrak{f}} := \inf_{\mathbf{t}\in\Sigma_N} \|\mathbf{s}-\mathbf{t}\|_{\mathfrak{f}}, \quad N = 0, 1, 2, \dots. \tag{2.1}$$

In applications, one is interested in an effective algorithm to determine sequences $\mathbf{t}$ which minimize (2.1). Such sequences, when they exist, are called *best N-term approximations to* $\mathbf{s}$ *in* $\mathfrak{f}$. Given a fixed constant $c \ge 1$ we shall say that $\mathbf{t} \in \Sigma_N$ is a *near best N-term approximation to* $\mathbf{s}$ (*with constant c*) if

$$\frac{1}{c}\|\mathbf{s}-\mathbf{t}\|_{\mathfrak{f}} \le \sigma_N(\mathbf{s})_{\mathfrak{f}}. \tag{2.2}$$

Observe that we always have $\sigma_N(\mathbf{s})_{\mathfrak{f}} \le \|\mathbf{s}-\mathbf{t}\|_{\mathfrak{f}}$, $\mathbf{t} \in \Sigma_N$, so the best approximation corresponds to the constant $c = 1$.

In practice, the computational cost of best approximations may be too high, so one is led to use simpler algorithms which provide near best approximations. One such case are the so-called *greedy algorithms* [32]. A greedy algorithm of step N is defined as a correspondence

$$\mathbf{s} = \{s_I\} \in \mathfrak{f} \longmapsto G_N(\mathbf{s}) := \sum_{\ell=1}^{N} s_{I_\ell}\mathbf{e}_{I_\ell} \in \Sigma_N,$$

for some enumeration of the index set $\mathcal{I} = \{I_\ell\}_{\ell=1}^{\infty}$ so that

$$\|s_{I_1}\mathbf{e}_{I_1}\|_{\mathfrak{f}} \geq \|s_{I_2}\mathbf{e}_{I_2}\|_{\mathfrak{f}} \geq \|s_{I_3}\mathbf{e}_{I_3}\|_{\mathfrak{f}} \geq \cdots . \tag{2.3}$$

Observe that such non-decreasing rearrangements exist by assumption (c) above. The following result, due to Konyagin and Temlyakov [22], characterizes the sequence spaces $\mathfrak{f}$ for which greedy algorithms give near optimal best approximations. For completeness, we provide an elementary proof in our special setting. Below we denote by

$$\tilde{\mathbf{1}}_\Gamma = \tilde{\mathbf{1}}_{\Gamma,\mathfrak{f}} = \sum_{I\in\Gamma} \frac{\mathbf{e}_I}{\|\mathbf{e}_I\|_{\mathfrak{f}}},$$

the (normalized) indicator sequence of a finite set of indices $\Gamma \subset \mathcal{I}$.

Theorem 2.1. *Let $\mathfrak{f}$ be a sequence space satisfying* (a)–(c) *above. Then, the following are equivalent:*

(i) *There exists $c \geq 1$ so that, for all $\mathbf{s} \in \mathfrak{f}$ and $N = 0, 1, 2, \ldots$,*

$$\frac{1}{c}\|\mathbf{s} - G_N(\mathbf{s})\|_{\mathfrak{f}} \leq \sigma_N(\mathbf{s})_{\mathfrak{f}}; \tag{2.4}$$

(ii) *There exists $c \geq 1$ so that, for all finite sets $\Gamma, \Gamma' \subset \mathcal{I}$ with* $\operatorname{Card}\Gamma = \operatorname{Card}\Gamma'$,

$$\frac{1}{c}\|\tilde{\mathbf{1}}_{\Gamma'}\|_{\mathfrak{f}} \leq \|\tilde{\mathbf{1}}_{\Gamma}\|_{\mathfrak{f}} \leq c\|\tilde{\mathbf{1}}_{\Gamma'}\|_{\mathfrak{f}}. \tag{2.5}$$

Proof. We prove first that (i) implies (ii). Take two finite sets $\Gamma, \Gamma' \subset \mathcal{I}$ with cardinality N, and let $\mathbf{s} = \tilde{\mathbf{1}}_{\Gamma\cup\Gamma'}$. If we assume first that $\Gamma \cap \Gamma' = \emptyset$, then (i) gives

$$\frac{1}{c}\|\tilde{\mathbf{1}}_{\Gamma'}\|_{\mathfrak{f}} = \frac{1}{c}\|\mathbf{s} - \tilde{\mathbf{1}}_{\Gamma}\|_{\mathfrak{f}} \leq \sigma_N(\mathbf{s})_{\mathfrak{f}} \leq \|\mathbf{s} - \tilde{\mathbf{1}}_{\Gamma'}\|_{\mathfrak{f}} = \|\tilde{\mathbf{1}}_{\Gamma}\|_{\mathfrak{f}}.$$

In general, if $\Gamma \cap \Gamma' \neq \emptyset$, then the quasi-triangle inequality and the previous case give us

$$\begin{aligned}\|\tilde{\mathbf{1}}_{\Gamma'}\|_{\mathfrak{f}} &= \|\tilde{\mathbf{1}}_{\Gamma'\setminus\Gamma} + \tilde{\mathbf{1}}_{\Gamma'\cap\Gamma}\|_{\mathfrak{f}} \leq C(\|\tilde{\mathbf{1}}_{\Gamma'\setminus\Gamma}\|_{\mathfrak{f}} + \|\tilde{\mathbf{1}}_{\Gamma'\cap\Gamma}\|_{\mathfrak{f}}) \\ &\leq C(c\|\tilde{\mathbf{1}}_{\Gamma\setminus\Gamma'}\|_{\mathfrak{f}} + \|\tilde{\mathbf{1}}_{\Gamma'\cap\Gamma}\|_{\mathfrak{f}}) \leq (c+1)C\|\tilde{\mathbf{1}}_{\Gamma}\|_{\mathfrak{f}},\end{aligned}$$

since $\operatorname{Card}\Gamma \setminus \Gamma' = \operatorname{Card}\Gamma' \setminus \Gamma$, and in the last step we use monotonicity.

We turn to the converse: (ii) implies (i). If Δ is any index subset of $\mathcal{I}$ and $\mathbf{s} \in \mathfrak{f}$, we shall denote the truncation by Δ as

$$\mathbf{s}\chi_\Delta = \{s_I\chi_\Delta\}_{I\in\mathcal{I}} = \begin{cases} s_I & \text{if } I \in \Delta, \\ 0 & \text{otherwise.}\end{cases}$$

We first observe that monotonicity yields the following simplification

$$\sigma_N(\mathbf{s})_{\mathfrak{f}} := \inf_{\mathbf{t}\in\Sigma_N} \|\mathbf{s}-\mathbf{t}\|_{\mathfrak{f}} = \inf_{\mathrm{Card}\,\Gamma=N} \|\{s_I\chi_{\Gamma^c}\}\|_{\mathfrak{f}}. \tag{2.6}$$

Indeed, for any $\mathbf{t}\in\Sigma_N$, if $\Gamma_{\mathbf{t}} = \mathrm{Supp}\,\mathbf{t}$, then

$$\begin{aligned}\|\mathbf{s}-\mathbf{t}\|_{\mathfrak{f}} &= \|\{(s_I-t_I)\chi_{\Gamma_{\mathbf{t}}}\} + \{s_I\chi_{\Gamma_{\mathbf{t}}^c}\}\|_{\mathfrak{f}}\\ &\geq \|\{s_I\chi_{\Gamma_{\mathbf{t}}^c}\}\|_{\mathfrak{f}} \geq \|\{s_I\chi_{\Gamma^c}\}\|_{\mathfrak{f}},\end{aligned}$$

where Γ is any set containing $\Gamma_{\mathbf{t}}$ with cardinality exactly N. From here (2.6) follows immediately. Now, if we call $\Gamma_0 = \mathrm{Supp}\,G_N(\mathbf{s})$, (i) reduces to show

$$\|\{s_I\chi_{\Gamma^c}\}\|_{\mathfrak{f}} \geq \frac{1}{c'}\|\{s_I\chi_{\Gamma_0^c}\}\|_{\mathfrak{f}}, \qquad \forall\,\Gamma : \mathrm{Card}\,\Gamma = N. \tag{2.7}$$

Starting from the right-hand side and using the quasi-triangular inequality,

$$\|\{s_I\chi_{\Gamma_0^c}\}\|_{\mathfrak{f}} \leq C(\|\{s_I\chi_{\Gamma_0^c\cap\Gamma}\}\|_{\mathfrak{f}} + \|\{s_I\chi_{\Gamma_0^c\cap\Gamma^c}\}\|_{\mathfrak{f}}).$$

Clearly, by monotonicity, the second summand is smaller than $\|\{s_I\chi_{\Gamma^c}\}\|_{\mathfrak{f}}$, so we just need to estimate the first term. Now, a new use of monotonicity, together with (ii) and the definition of the set Γ_0 yields

$$\begin{aligned}\|\{s_I\chi_{\Gamma_0^c\cap\Gamma}\}\|_{\mathfrak{f}} &\leq \max_{I\in\Gamma_0^c\cap\Gamma} |s_I|\,\|\mathbf{e}_I\|_{\mathfrak{f}}\,\|\tilde{\mathbf{1}}_{\Gamma_0^c\cap\Gamma}\|_{\mathfrak{f}}\\ &\leq c \min_{I\in\Gamma_0\cap\Gamma^c} |s_I|\,\|\mathbf{e}_I\|_{\mathfrak{f}}\,\|\tilde{\mathbf{1}}_{\Gamma_0\cap\Gamma^c}\|_{\mathfrak{f}}\\ &\leq c\|\{|s_I|\chi_{\Gamma_0\cap\Gamma^c}\}\|_{\mathfrak{f}}\\ &\leq c\|\{s_I\chi_{\Gamma^c}\}\|_{\mathfrak{f}}.\end{aligned}$$

Thus, $\|\{s_I\chi_{\Gamma_0^c}\}\|_{\mathfrak{f}} \leq (c+1)C\|\{s_I\chi_{\Gamma^c}\}\|_{\mathfrak{f}}$, establishing the theorem. □

Remark 2.2. When $\{\mathbf{e}_I\}$ is a Schauder basis of $\mathfrak{f}$ (hence unconditional, by (b)), the previous theorem states that $\{\mathbf{e}_I\}$ is a *greedy basis* (i.e., (2.4) holds) if and only if $\{\mathbf{e}_I\}$ is *democratic* (i.e., (2.5) holds). This result (and terminology) was first introduced in a more general context of Banach spaces in [22]. Our proof follows the lines of [20].

Remark 2.3. Some examples concerning condition (2.5) are the following. For any $0<p<\infty$, the space $\ell_p(\mathbb{Z})$ satisfies $\|\tilde{\mathbf{1}}_\Gamma\|_{\ell_p} = (\mathrm{Card}\,\Gamma)^{1/p}$, for all $\Gamma\subset\mathbb{Z}$. On the other hand, this condition fails for mixed normed spaces $\ell_q(\ell_p)$, with norm given by

$$\|\{s_{j,k}\}\|_{\ell_q(\ell_p)} = \Big[\sum_{j\in\mathbb{Z}}\Big(\sum_{k\in\mathbb{Z}}|s_{j,k}|^p\Big)^{q/p}\Big]^{1/q}, \quad \text{and } p\neq q.$$

Indeed, in this case $\Gamma = \{(1,0),\dots,(N,0)\}$ and $\Gamma' = \{(0,1),\dots,(0,N)\}$ lead respectively to $\|\tilde{\mathbf{1}}_\Gamma\|_{\ell_q(\ell_p)} = N^{1/q}$ and $\|\tilde{\mathbf{1}}_{\Gamma'}\|_{\ell_q(\ell_p)} = N^{1/p}$. For a discussion on non-greedy algorithms for N-term approximation in such situations we refer to [20, Section 11.6]

Using the terminology in [20], for $0 < p < \infty$ we say that $\mathfrak{f}$ is a *p-space* if

$$\frac{1}{c}(\operatorname{Card}\Gamma)^{1/p} \le \|\tilde{\mathbf{1}}_\Gamma\|_{\mathfrak{f}} \le c(\operatorname{Card}\Gamma)^{1/p}, \quad \text{for all finite } \Gamma \subset \mathcal{I}. \tag{2.8}$$

This condition is also referred to in the literature as the *p-Temlyakov property* (see, e.g. [21]). A natural family of spaces satisfying (2.8) are the *discrete Lorentz spaces* $\ell_{p,r}(\mathcal{I})$, $0 < p < \infty$. When $0 < r < \infty$, these consist of sequences $\mathbf{s} = \{s_I\}$ for which $\lim s_I = 0$ and

$$\|\mathbf{s}\|_{\ell_{p,r}} := \left[\sum_{k=1}^{\infty} (k^{1/p}|s_{I_k}|)^r \frac{1}{k}\right]^{1/r} < \infty,$$

for some enumeration of the index set $\mathcal{I} = \{I_k\}_{k=1}^\infty$ so that $|s_{I_1}| \ge |s_{I_2}| \ge \cdots$. When $r = \infty$, $\ell_{p,\infty}$ is the *discrete weak-ℓ_p space* consisting of sequences so that

$$\|\mathbf{s}\|_{\ell_{p,\infty}} := \sup_{\lambda > 0} \lambda(\operatorname{Card}\{I : |s_I| \ge \lambda\})^{1/p} = \sup_{k\in\mathbb{N}} k^{1/p}|s_{I_k}| < \infty.$$

In this paper, given a sequence space $\mathfrak{f}$ we define

$$\ell_{p,r}(\mathfrak{f}) := \left\{\mathbf{s} \in \mathbb{C}^{\mathcal{I}} : \{\|s_I\mathbf{e}_I\|_{\mathfrak{f}}\} \in \ell_{p,r}(\mathcal{I})\right\}, \quad \text{and}$$
$$\|\mathbf{s}\|_{\ell_{p,r}(\mathfrak{f})} := \left[\sum_{k=1}^{\infty} (k^{1/p}\|s_{I_k}\mathbf{e}_{I_k}\|_{\mathfrak{f}})^r \frac{1}{k}\right]^{1/r},$$

where $\|s_{I_k}\mathbf{e}_{I_k}\|_{\mathfrak{f}}$ are ordered decreasingly as in (2.3). These are quasi-Banach spaces which are isomorphic to $\ell_{p,r}(\mathcal{I})$. Since

$$\|\tilde{\mathbf{1}}_\Gamma\|_{\ell_{p,r}(\mathfrak{f})} = \Big(\sum_{k \le \operatorname{Card}\Gamma} k^{r/p-1}\Big)^{1/r} \sim (\operatorname{Card}\Gamma)^{1/p},$$

it follows that $\ell_{p,r}(\mathfrak{f})$ is also a p-space for any sequence space $\mathfrak{f}$ and any indices $0 < r \le \infty$ and $0 < p < \infty$.

The following result shows that p-spaces are precisely those $\mathfrak{f}$'s which can be embedded in between two such Lorentz spaces (compare with Theorem 3.11 in [30, Chapter 5], or 11.19 in [20]). We shall assume that $\mathfrak{f} \hookrightarrow \mathbb{C}^{\mathcal{I}}$, that is,

$$\lim_{n\to\infty} \mathbf{s}^{(n)} = \mathbf{s}, \text{ in } \mathfrak{f} \Rightarrow \lim_{n\to\infty} s_I^{(n)} = s_I, \quad \forall\, I \in \mathcal{I}.$$

Theorem 2.4. *Let $0 < p < \infty$, and $\mathfrak{f} \hookrightarrow \mathbb{C}^{\mathcal{I}}$ be a sequence space satisfying* (a)-(b) *above. Then, $\mathfrak{f}$ is a p-space if and only if for some $r > 0$ we have*

$$\ell_{p,r}(\mathfrak{f}) \hookrightarrow \mathfrak{f} \hookrightarrow \ell_{p,\infty}(\mathfrak{f}). \tag{2.9}$$

Moreover, if $\mathfrak{f}$ satisfies the ρ-triangular inequality $\|\mathbf{s}+\mathbf{t}\|_{\mathfrak{f}}^{\rho} \leq \|\mathbf{s}\|_{\mathfrak{f}}^{\rho} + \|\mathbf{t}\|_{\mathfrak{f}}^{\rho}$, then one can take $r = \rho$.

Proof. The sufficiency of (2.9) follows easily from the fact that $\ell_{p,r}(\mathfrak{f})$ are p-spaces for all $0 < r \leq \infty$ and $0 < p < \infty$. For the necessity, let us first show the right inclusion in (2.9). Given $\mathbf{s} = \{s_I\} \in \mathfrak{f}$ and $\lambda > 0$, define the set of indices $\Gamma_\lambda = \{I : \|s_I \mathbf{e}_I\|_{\mathfrak{f}} \geq \lambda\}$. Then, the p-space assumption on $\mathfrak{f}$ and the monotonicity imply that this set must be finite and, moreover,

$$\lambda(\operatorname{Card}\Gamma_\lambda)^{1/p} \leq c\Big\|\sum_{I\in\Gamma_\lambda} s_I \mathbf{e}_I\Big\|_{\mathfrak{f}} \leq c\|\mathbf{s}\|_{\mathfrak{f}}.$$

Taking the supremum in $\lambda > 0$ we obtain the desired inclusion. For the left hand inclusion in (2.9), let $\mathbf{s} = \{s_I\} \in \ell_{p,\rho}(\mathfrak{f})$ and define $\mathbf{s}^{(j)} = \sum_{1\leq k<2^j} s_{I_k}\mathbf{e}_{I_k}$, with the coefficients $\|s_{I_k}\mathbf{e}_{I_k}\|_{\mathfrak{f}}$ rearranged decreasingly as in (2.3). We claim that $\{\mathbf{s}^{(j)}\}_{j=1}^{\infty}$ converges to $\mathbf{s}$ in $\mathfrak{f}$. Indeed, by the ρ-triangular inequality, the monotonicity and the p-space property,

$$\begin{aligned} \|\mathbf{s}^{(j+m)} - \mathbf{s}^{(j)}\|_{\mathfrak{f}}^{\rho} &\leq \sum_{\ell=j}^{j+m} \Big\|\sum_{2^\ell\leq k<2^{\ell+1}} s_{I_k}\mathbf{e}_{I_k}\Big\|_{\mathfrak{f}}^{\rho} \\ &\leq c\sum_{\ell=j}^{j+m} \|s_{I_{2^\ell}}\mathbf{e}_{I_{2^\ell}}\|_{\mathfrak{f}}^{\rho}(2^\ell)^{\rho/p} \\ &\leq c\sum_{\ell=0}^{\infty}(2^{\ell/p}\|s_{I_{2^\ell}}\mathbf{e}_{I_{2^\ell}}\|_{\mathfrak{f}})^{\rho}. \end{aligned} \tag{2.10}$$

Since the last series is finite (and bounded by $\|\mathbf{s}\|_{\ell_{p,\rho}(\mathfrak{f})}^{\rho}$), it follows that $\{\mathbf{s}^{(j)}\}_{j=1}^{\infty}$ is a Cauchy sequence in $\mathfrak{f}$, which by the assumption $\mathfrak{f} \hookrightarrow \mathbb{C}^{\mathcal{I}}$ must converge to $\mathbf{s}$. Considering $j = 0$ in the above inequalities and passing to the limit, we obtain the desired estimate

$$\|\mathbf{s}\|_{\mathfrak{f}} = \lim_{m\to\infty}\|\mathbf{s}^{(m)}\|_{\mathfrak{f}} \leq c'\|\mathbf{s}\|_{\ell_{p,\rho}(\mathfrak{f})}. \qquad \square$$

Corollary 2.5. *Let $\mathfrak{f} \hookrightarrow \mathbb{C}^{\mathcal{I}}$ be a p-space satisfying the ρ-triangular inequality for some $0 < \rho \leq 1$. If either $0 < \tau < p$ and $0 < r \leq \infty$, or $\tau = p$ and $0 < r \leq \rho$, then*

$$\ell_{\tau,r}(\mathfrak{f}) = \{\mathbf{s} \in \mathfrak{f} : \{s_I\|\mathbf{e}_I\|_{\mathfrak{f}}\} \in \ell_{\tau,r}(\mathcal{I})\}. \tag{2.11}$$

Proof. This is a consequence of the left inclusion in (2.9) and the trivial embedding $\ell_{\tau,r} \hookrightarrow \ell_{p,\rho}$. ◻

3. General Sequence Spaces Associated with Smoothness

In this paper we shall be interested in the sequence spaces introduced by Frazier and Jawerth in [14]: $\mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^\alpha_{\tau,q}$. To allow a greater generality, and for later applications to anisotropic smoothness spaces in $\mathbb{R}^d$, we adopt the following definitions (see also [13, 24]).

We fix a real $d\times d$-matrix M with $|\det M| > 1$. We define the "M-adic" blocks

$$I_{j,\mathbf{k}} := M^{-j}([0,1)^d + \mathbf{k}), \quad j \in \mathbb{Z},\ \mathbf{k} \in \mathbb{Z}^d, \tag{3.1}$$

and a real number $\lambda > 1$ by the relation $|\det M| = \lambda^d$. We shall regard the set $\mathcal{I}$ of all M-adic blocks as an index set, which we shall identify, when necessary, with $\mathbb{Z} \times \mathbb{Z}^d$.

Roughly speaking, and keeping in mind the functional setting, $I_{j,\mathbf{k}}$ gives the space-localization of a wavelet $\psi_{I_{j,\mathbf{k}}}(x) = \psi_{j,\mathbf{k}}(x) = \lambda^{jd/2}\psi(M^j x - \mathbf{k})$. Observe that, when $M = 2I$, one obtains the usual family of dyadic cubes. Other choices of M, however, give rise to tilings of $\mathbb{R}^d$ by parallelograms, which in general will not have the nesting property of dyadic intervals at different scales.

Having fixed one such matrix M, we are ready to define the associated sequence spaces of Besov and Triebel-Lizorkin type. We shall denote below $\chi_I^{(p)} = \chi_I/\|\chi_I\|_p = |I|^{-1/p}\chi_I$.

Definition 3.1.

1. Given $s \in \mathbb{R}$, $0 < p < \infty$ and $0 < r \le \infty$, we let $\mathfrak{f}^s_{p,r}$ be the space of sequences $\mathbf{s} = \{s_I\}_{I\in\mathcal{I}}$ such that

$$\|\mathbf{s}\|_{\mathfrak{f}^s_{p,r}} = \Big\|\Big[\sum_{I\in\mathcal{I}}(|I|^{-s/d+1/p-1/2}\,|s_I|\chi_I^{(p)}(\cdot))^r\Big]^{1/r}\Big\|_{L^p(\mathbb{R}^d)} < \infty. \tag{3.2}$$

2. Given $\alpha \in \mathbb{R}$, $0 < \tau, q \le \infty$, we let $\mathfrak{b}^\alpha_{\tau,q}$ be the space of sequences $\mathbf{s} = \{s_I\}_{I\in\mathcal{I}}$ such that

$$\|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}} = \Big[\sum_{j\in\mathbb{Z}}\Big(\sum_{|I|=\lambda^{-jd}}(|I|^{-\alpha/d+1/\tau-1/2}\,|s_I|)^\tau\Big)^{q/\tau}\Big]^{1/q} < \infty. \tag{3.3}$$

Of course, one takes the obvious modifications when $r, \tau, q = \infty$. We avoid the case $p = \infty$, which requires a different definition. It is clear that these are quasi-Banach spaces satisfying properties (a) and (b) in Section 2. Property (c) also holds for the whole family $\mathfrak{f}^s_{p,r}$, as a consequence of the right inclusion in (2.9) and the p-space condition in our next proposition. Finally we observe that $\{\mathbf{e}_I\}$ is an unconditional basis for these spaces only when $r, \tau, q < \infty$.

Proposition 3.2. *Let $s \in \mathbb{R}$, $0 < p < \infty$ and $0 < r \le \infty$. Then $\mathfrak{f}^s_{p,r}$ is a p-space. That is, there exists $c > 0$ so that*

$$\frac{1}{c}(\operatorname{Card}\Gamma)^{1/p} \le \|\tilde{\mathbf{1}}_\Gamma\|_{\mathfrak{f}^s_{p,r}} \le c(\operatorname{Card}\Gamma)^{1/p}, \quad \text{for all finite } \Gamma \subset \mathcal{I}.$$

Proof. The proof follows a similar argument as in [20, Theorem 11.2]. Since $\|\mathbf{s}\|_{\mathfrak{f}^s_{p,\infty}} \le \|\mathbf{s}\|_{\mathfrak{f}^s_{p,r}}$, for all $r > 0$, the left-hand estimate will follow from

$$\frac{1}{c}(\operatorname{Card}\Gamma)^{1/p} \le \|\sup_{I\in\Gamma} \chi_I^{(p)}(x)\|_{L^p}. \tag{3.4}$$

Now, calling $F(x) = \sup_{I\in\Gamma}\chi_I^{(p)}(x)$, which is a finite valued function, we have

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \lambda^{jd}|\{F(x)^p \ge \lambda^{jd}\}| &= \sum_{j\in\mathbb{Z}} \lambda^{jd} \sum_{\ell\ge 0} |\{F(x)^p = \lambda^{(j+\ell)d}\}| \\ &= \sum_{\ell\ge 0}\sum_{j\in\mathbb{Z}} \lambda^{-\ell d}\int_{\{F(x)^p=\lambda^{(j+\ell)d}\}} \lambda^{(j+\ell)d}\,dx \\ &= \Big(\sum_{\ell\ge 0}\lambda^{-\ell d}\Big)\int_{\mathbb{R}^d} F(x)^p\,dx.\end{aligned}$$

If we denote $\Gamma_j = \{I \in \Gamma : |I| = \lambda^{-jd}\}$, then the disjoint union $\bigcup_{I\in\Gamma_j} I \subset \{F(x)^p \ge \lambda^{jd}\}$, from which we conclude

$$\int_{\mathbb{R}^d} F(x)^p\,dx \ge \frac{1}{c}\sum_{j\in\mathbb{Z}}\lambda^{jd}\sum_{I\in\Gamma_j}|I| = \frac{1}{c}\operatorname{Card}\Gamma,$$

with the constant $c = \lambda^d/(\lambda^d - 1)$. Conversely, we need to estimate the L^p-norm of $[\sum_{I\in\Gamma}(\chi_I^{(p)}(x))^r]^{1/r}$. Now, for each $x \in \bigcup_{I\in\Gamma} I$, let I_x be the smallest interval in Γ containing x. Then we have the pointwise estimate

$$\begin{aligned}\sum_{I\in\Gamma}(\chi_I^{(p)}(x))^r &= \sum_{j\in\mathbb{Z}}\sum_{I\in\Gamma_j}|I|^{-r/p}\chi_I(x) \\ &\le c'|I_x|^{-r/p}\chi_{I_x}(x) \\ &\le c'\Big(\sum_{I\in\Gamma}|I|^{-1}\chi_I(x)\Big)^{r/p}.\end{aligned}$$

Thus, raising to the power $1/r$ and taking the L^p-norm, we conclude

$$\Big\|\Big[\sum_{I\in\Gamma}(\chi_I^{(p)})^r\Big]^{1/r}\Big\|_{L^p} \le c''\Big(\int_{\mathbb{R}^d}\sum_{I\in\Gamma}|I|^{-1}\chi_I(x)\,dx\Big)^{1/p}$$
$$= c''(\operatorname{Card}\Gamma)^{1/p}. \qquad \square$$

4. JACKSON INEQUALITIES IN $\mathfrak{f}^s_{p,r}$-NORMS

In this section we study the validity of Jackson-type inequalities of the form

$$\sigma_N(\mathbf{s})_{\mathfrak{f}^s_{p,r}} \le C(N+1)^{-\varepsilon}\|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}}, \quad N = 0,1,2,\dots, \tag{4.1}$$

where ε is a positive number. Observe that the case $N = 0$ just corresponds to the embedding $\mathfrak{b}^\alpha_{\tau,q} \hookrightarrow \mathfrak{f}^s_{p,r}$. Our first result gives a necessary relation which must be satisfied by the "smoothness and integrability" indices s, α, p, τ.

Proposition 4.1. *If Jackson's inequality* (4.1) *holds for any* $N \ge 0$*, then we must have* $s/d - 1/p = \alpha/d - 1/\tau$.

Proof. The proof follows from a simple homogeneity argument. We denote by D the dilation operator mapping sequences $\mathbf{s} = \{s_I\}$ into $\mathbf{t} = \{t_I = s_{(MI)}\}$. Observe that this just produces a shift in the indices $\{s_{j,\mathbf{k}}\} \mapsto \{t_{j,\mathbf{k}} = s_{j-1,\mathbf{k}}\}$. It is immediate to verify that

$$\begin{aligned}\|D\mathbf{s}\|_{\mathfrak{f}^s_{p,r}} &= \lambda^{d(s/d-1/p+1/2)}\|\mathbf{s}\|_{\mathfrak{f}^s_{p,r}}, \\ \|D\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}} &= \lambda^{d(\alpha/d-1/\tau+1/2)}\|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}}.\end{aligned} \tag{4.2}$$

Thus, since D preserves Σ_N, if Jackson's inequality (4.1) holds, we deduce

$$\begin{aligned}\lambda^{d\ell(s/d-1/p+1/2)}\sigma_N(\mathbf{s})_{\mathfrak{f}^s_{p,r}} &= \sigma_N(D^\ell\mathbf{s})_{\mathfrak{f}^s_{p,r}} \le C(N+1)^{-\varepsilon}\|D^\ell\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}} \\ &= C(N+1)^{-\varepsilon}\lambda^{d\ell(\alpha/d-1/\tau+1/2)}\|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}},\end{aligned}$$

for any $\ell \in \mathbb{Z}$. This gives

$$0 < \frac{(N+1)^{\varepsilon}\sigma_N(\mathbf{s})_{\mathfrak{f}^s_{p,r}}}{C\|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}}} \le \lambda^{d\ell((\alpha-s)/d-1/\tau+1/p)},$$

at least for finitely supported (non-null) $\mathbf{s}$. Now, if $\delta := (\alpha-s)/d-1/\tau+1/p \ne 0$, we may let $\ell \to \pm\infty$ (depending on the sign of δ), obtaining a contradiction. $\square$

Next we establish a lower bound for the error decay in a given pair of smoothness spaces $\mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^\alpha_{\tau,q}$.

Proposition 4.2. *If Jackson's inequality* (4.1) *holds for some* $\varepsilon > 0$*, all* $N \geq 0$ *and* $s/d - 1/p = \alpha/d - 1/\tau$*, then necessarily*

$$\max\{\tau, q\} < p \quad \text{and} \quad \varepsilon \leq \min\left\{\frac{1}{\tau} - \frac{1}{p}, \frac{1}{q} - \frac{1}{p}\right\}.$$

Proof. The result follows by testing with two simple examples of the form $\mathbf{s} = \tilde{\mathbf{1}}_\Gamma$, for suitable choices of Γ. We first observe that, under the condition $1/\tau = 1/p + (\alpha - s)/d$, we have $\|\mathbf{e}_I\|_{\mathfrak{f}^s_{p,r}} = \|\mathbf{e}_I\|_{\mathfrak{b}^\alpha_{\tau,q}}$, for all $I \in \mathcal{I}$, so we will just write $\tilde{\mathbf{1}}_\Gamma = \tilde{\mathbf{1}}_{\Gamma, \mathfrak{f}^s_{p,r}} = \tilde{\mathbf{1}}_{\Gamma, \mathfrak{b}^\alpha_{\tau,q}}$. Now, using the notation $\mathbf{s} = \{s_{j,\mathbf{k}}\}$, we can choose $\Gamma = \{(1,0), \ldots, (2N, 0)\}$, so that a simple calculation, Theorem 2.1 and Proposition 3.2 give

$$\frac{1}{c} N^{1/p} \leq \sigma_N(\tilde{\mathbf{1}}_\Gamma)_{\mathfrak{f}^s_{p,r}} \leq C N^{-\varepsilon} \|\tilde{\mathbf{1}}_\Gamma\|_{\mathfrak{b}^\alpha_{\tau,q}} = C N^{-\varepsilon} (2N)^{1/q}.$$

Since $\varepsilon > 0$, it follows that necessarily $q < p$ and $\varepsilon \leq 1/q - 1/p$. One proves similarly the restrictions on τ by choosing $\Gamma = \{(0,1), \ldots, (0, 2N)\}$. □

Finally, we are ready to prove the main result of this section, which establishes Jackson's inequality for all the possible choices of indices. This extends a result in [9], where such inequality was shown for the pair $\mathfrak{f}^0_{p,2}$, $\mathfrak{b}^\alpha_{\tau,\tau}$. Our approach to the proof follows the lines of [11].

Theorem 4.3. *Let* $s, \alpha \in \mathbb{R}$, $0 < p, \tau, q < \infty$ *and* $0 < r \leq \infty$ *be so that*

$$\max\{\tau, q\} < p \quad \text{and} \quad \frac{\alpha}{d} - \frac{1}{\tau} = \frac{s}{d} - \frac{1}{p}. \tag{4.3}$$

Then, for all $\mathbf{s} \in \mathfrak{b}^\alpha_{\tau,q}$ *we have*

$$\sigma_N(\mathbf{s})_{\mathfrak{f}^s_{p,r}} \leq C(N+1)^{-(1/(\tau \vee q) - 1/p)} \|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,q}}, \quad N = 0, 1, 2, \ldots.$$

Proof. It suffices to show the theorem for $\tau = q$. Indeed, under such assumption the cases $\tau > q$ would follow from

$$\sigma_N(\mathbf{s})_{\mathfrak{f}^s_{p,r}} \leq C(N+1)^{-(1/\tau - 1/p)} \|\mathbf{s}\|_{\mathfrak{b}^\alpha_{\tau,\tau}},$$

and the inclusion $\mathfrak{b}^\alpha_{\tau,q} \hookrightarrow \mathfrak{b}^\alpha_{\tau,\tau}$. On the other hand, if $\tau < q$ and we choose $\beta \in \mathbb{R}$ so that $\beta/d - 1/q = \alpha/d - 1/\tau = s/d - 1/p$, then we will have

$$\sigma_N(\mathbf{s})_{\mathfrak{f}^s_{p,r}} \leq C(N+1)^{-(1/q - 1/p)} \|\mathbf{s}\|_{\mathfrak{b}^\beta_{q,q}},$$

which, together with the inclusion $\mathfrak{b}^\alpha_{\tau,q} \hookrightarrow \mathfrak{b}^\beta_{q,q}$, gives the desired result.

We will deduce the theorem when $\tau = q$ from the following more general result.

Proposition 4.4. *Let $\mathfrak{f} \hookrightarrow \mathbb{C}^{\mathcal{I}}$ be a p-space (as in Section 2). Then, for all $0 < \tau < p$, we have the embedding $\ell_{\tau,\infty}(\mathfrak{f}) \hookrightarrow \mathfrak{f}$ and, moreover,*

$$\sigma_N(\mathbf{s})_{\mathfrak{f}} \le CN^{-(1/\tau - 1/p)}\|\mathbf{s}\|_{\ell_{\tau,\infty}(\mathfrak{f})}, \quad \forall\, \mathbf{s} \in \ell_{\tau,\infty}(\mathfrak{f}),\ N = 1, 2, \dots. \tag{4.4}$$

Proof. The first assertion follows from the trivial embedding of Lorentz spaces $\ell_{\tau,\infty}(\mathfrak{f}) \hookrightarrow \ell_{p,\rho}(\mathfrak{f})$, whenever $\tau < p$, and then an application of the left hand embedding in (2.9).

In order to show (4.4) we may assume $\|\mathbf{s}\|_{\ell_{\tau,\infty}(\mathfrak{f})} = 1$, while by monotonicity of σ_N, it suffices to consider $N = 2^j$, $j = 0, 1, \dots$. For each such j define $\mathbf{s}^{(j)} = \sum_{1\le k<2^j} s_{I_k}\mathbf{e}_{I_k}$ as in the proof of Theorem 2.4. Then, recalling that $\mathbf{s}^{(j)}$ converges to $\mathbf{s}$ in $\mathfrak{f}$ and using the estimate in (2.10), we obtain

$$\begin{aligned}
\|\mathbf{s} - \mathbf{s}^{(j)}\|_{\mathfrak{f}} &\le c\Big[\sum_{\ell=j}^{\infty} (2^{\ell/p}\|s_{I_{2^\ell}}\mathbf{e}_{I_{2^\ell}}\|_{\mathfrak{f}})^{\rho}\Big]^{1/\rho} \\
&\le c\|\mathbf{s}\|_{\ell_{\tau,\infty}(\mathfrak{f})}\Big[\sum_{\ell=j}^{\infty} 2^{\rho(\ell/p - \ell/\tau)}\Big]^{1/\rho} \\
&= c' 2^{-(j/\tau - j/p)},
\end{aligned}$$

since $\tau < p$. Finally, we use $\sigma_N(\mathbf{s})_{\mathfrak{f}} \le \|\mathbf{s} - \mathbf{s}^{(j)}\|_{\mathfrak{f}}$, since $\mathbf{s}^{(j)} \in \Sigma_{2^j}$. □

To complete the proof of Theorem 4.3, it only remains to show that $\mathfrak{b}^{\alpha}_{\tau,\tau} = \ell_{\tau,\tau}(\mathfrak{f}^s_{p,r})$. Assuming this identity, the result follows immediately from (4.4) and the trivial embedding $\ell_{\tau,\tau}(\mathfrak{f}^s_{p,r}) \hookrightarrow \ell_{\tau,\infty}(\mathfrak{f}^s_{p,r})$. Finally, the above identity is a simple consequence of (4.3), which implies $\|\mathbf{e}_I\|_{\mathfrak{f}^s_{p,r}} = \|\mathbf{e}_I\|_{\mathfrak{b}^{\alpha}_{\tau,q}}$. □

Remark 4.5. Proposition 4.4 is sharp, in the sense that for a p-space $\mathfrak{f}$ the inequalities $\sigma_N(\mathbf{s})_{\mathfrak{f}} \lesssim N^{-\varepsilon}\|\mathbf{s}\|_{\ell_{\tau,r}(\mathfrak{f})}$, $N = 1, 2, \dots$, can only hold if $\tau < p$ and $\varepsilon \le 1/\tau - 1/p$. The same example as in Proposition 4.2 shows this fact.

Remark 4.6. An interesting particular case of Theorem 4.3 corresponds to $N = 0$. Our result proves the non-trivial embeddings:

$$\mathfrak{b}^{\alpha}_{\tau,q} \hookrightarrow \mathfrak{f}^s_{p,r},$$

when

$$\frac{\alpha}{d} - \frac{1}{\tau} = \frac{s}{d} - \frac{1}{p}, \quad \max\{\tau, q\} < p, \quad 0 < r \le \infty.$$

These embeddings continue to hold in some cases when $\max\{\tau, q\} = p$ (see [15]), although as our examples above show, such cases will not admit a Jackson type inequality.

5. Bernstein Type Inequalities

In this section we turn to Bernstein-type inequalities of the form:

$$\|\mathbf{s}\|_{\mathfrak{b}^{\alpha}_{\tau,q}} \leq CN^{\varepsilon}\|\mathbf{s}\|_{\mathfrak{f}^{s}_{p,r}}, \quad \mathbf{s} \in \Sigma_N, \ N = 1, 2, \ldots, \tag{5.1}$$

where ε is a real number. The first observation is similar to Proposition 4.1, and gives the same relation between the smoothness and integrability indices s, α, p, τ. The proof is completely analogous, and left to the reader.

Proposition 5.1. *If Bernstein's inequality* (5.1) *holds for any $N \geq 1$, then we must have $\alpha/d - 1/\tau = s/d - 1/p$.*

In a similar way, one finds a lower bound for the rate of growth of N^{ε}.

Proposition 5.2. *If Bernstein's inequality* (5.1) *holds for some $\varepsilon \in \mathbb{R}$, all $N \geq 1$ and $\alpha/d - 1/\tau = s/d - 1/p$, then necessarily*

$$\varepsilon \geq \max\left\{\frac{1}{\tau} - \frac{1}{p}, \frac{1}{q} - \frac{1}{p}\right\}.$$

Proof. The proof is similar to the one given in Proposition 4.2. One tests with $S = \tilde{\mathbf{1}}_{\Gamma} \in \Sigma_N$, for the choices $\Gamma = \{(1,0), \ldots, (N,0)\}$ and $\Gamma = \{(0,1), \ldots, (0,N)\}$. We leave details to the reader. □

Remark 5.3. It is clear that (5.1) can never hold with $\varepsilon < 0$ (test, e.g, with a fixed $\mathbf{s} = \mathbf{e}_I$ and let $N \to \infty$). It is interesting, however, to observe that the case $\varepsilon = 0$ corresponds precisely to the collection of embeddings: $\mathfrak{f}^{s}_{p,r} \hookrightarrow \mathfrak{b}^{\alpha}_{\tau,q}$. These are known to hold exactly when

$$p \leq \min\{\tau, q\} \quad \text{and} \quad \frac{s}{d} - \frac{1}{p} = \frac{\alpha}{d} - \frac{1}{\tau},$$

with no restriction in $0 < r \leq \infty$ if $p < \tau$, and just for $0 < r \leq q$ if $p = \tau$ (see, e.g., [15, Theorem 3.8]). For the remaining cases, the Bernstein inequality (5.1) gives an upper bound for the norm of the inclusion operator $\Sigma_N \to \mathfrak{b}^{\alpha}_{\tau,q}$, when Σ_N is seen as a subspace of $\mathfrak{f}^{s}_{p,r}$. Our next theorem shows that the lower bounds in Proposition 5.2 are actually the best possible in the range $p > \min\{\tau, q\}$.

Theorem 5.4. *Let s, $\alpha \in \mathbb{R}$, $0 < p$, τ, $q < \infty$, and $0 < r \leq \infty$ be so that*

$$\min\{\tau, q\} < p \quad \text{and} \quad \frac{\alpha}{d} - \frac{1}{\tau} = \frac{s}{d} - \frac{1}{p}. \tag{5.2}$$

Then, there exists $C > 0$ so that, for all $N = 1, 2, \ldots,$

$$\|\mathbf{s}\|_{\mathfrak{b}^{\alpha}_{\tau,q}} \leq CN^{1/(\tau\wedge q)-1/p}\|\mathbf{s}\|_{\mathfrak{f}^{s}_{p,r}}, \quad \forall\, \mathbf{s} \in \Sigma_N. \tag{5.3}$$

Proof. As before, it will suffice to consider the case $q = \tau$. Indeed, if $q > \tau$ one can use $\|\mathbf{s}\|_{\mathfrak{b}^{\alpha}_{\tau,\tau}} \le CN^{1/\tau-1/p}\|\mathbf{s}\|_{\mathfrak{f}^s_{p,r}}$ and the trivial embedding $\mathfrak{b}^{\alpha}_{\tau,\tau} \hookrightarrow \mathfrak{b}^{\alpha}_{\tau,q}$. If on the other hand $q < \tau$, we define β so that $\beta/d-1/q = \alpha/d-1/\tau = s/d-1/p$. Then we will have

$$\|\mathbf{s}\|_{\mathfrak{b}^{\beta}_{q,q}} \le CN^{1/q-1/p}\|\mathbf{s}\|_{\mathfrak{f}^s_{p,r}},$$

and we obtain the desired inequality from the embedding $\mathfrak{b}^{\beta}_{q,q} \hookrightarrow \mathfrak{b}^{\alpha}_{\tau,q}$.

As for Jackson inequalities, the case $\tau = q$ will be a straightforward consequence of the equality $\mathfrak{b}^{\alpha}_{\tau,\tau} = \ell_{\tau,\tau}(\mathfrak{f}^s_{p,r})$ and the following general result. ◻

Proposition 5.5. *Let $\mathfrak{f}$ be a p-space as in Section 2. Then, for all $0 < \tau < p$ and all $0 < r \le \infty$ there exists $C = C(p,\tau,r) > 0$ such that*

$$\|\mathbf{s}\|_{\ell_{\tau,r}(\mathfrak{f})} \le CN^{1/\tau-1/p}\|\mathbf{s}\|_{\mathfrak{f}}, \quad \forall\, \mathbf{s} \in \Sigma_N,\ N = 1,2,\dots. \tag{5.4}$$

Proof. Let $\mathbf{s} = \sum_{k=1}^{N} s_{I_k}\mathbf{e}_{I_k} \in \Sigma_N$, where we can assume $\|s_{I_1}\mathbf{e}_{I_1}\|_{\mathfrak{f}} \ge \cdots \ge \|s_{I_N}\mathbf{e}_{I_N}\|_{\mathfrak{f}}$. Then, using the p-space property and monotonicity we have

$$\begin{aligned}
\|\mathbf{s}\|_{\ell_{\tau,r}(\mathfrak{f})} &= \Big[\sum_{k=1}^{N} (k^{1/\tau}\|s_{I_k}\mathbf{e}_{I_k}\|_{\mathfrak{f}})^r \frac{1}{k}\Big]^{1/r} \\
&\le c\Big[\sum_{k=1}^{N} \Big(k^{1/\tau-1/p}\Big\|\sum_{\ell=1}^{k} s_{I_\ell}\mathbf{e}_{I_\ell}\Big\|_{\mathfrak{f}}\Big)^r \frac{1}{k}\Big]^{1/r} \\
&\le c\|\mathbf{s}\|_{\mathfrak{f}}\Big[\sum_{k=1}^{N} k^{r(1/\tau-1/p)}\frac{1}{k}\Big]^{1/r} \\
&= C\|\mathbf{s}\|_{\mathfrak{f}}N^{1/\tau-1/p}.
\end{aligned}$$
◻

Remark 5.6. With the same proof as before, in the limiting case $\tau = p$ we have

$$\|\mathbf{s}\|_{\ell_{p,r}(\mathfrak{f})} \le C(\log N)^{1/r}\|\mathbf{s}\|_{\mathfrak{f}}, \quad \forall\, \mathbf{s} \in \Sigma_N,\ N = 2,3,\dots. \tag{5.5}$$

Particularizing to the spaces $\mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^{\alpha}_{\tau,q}$, this produces an additional Bernstein-type inequality in the limiting situation when $p = \tau = \min\{\tau,q\}$ (and thus $s = \alpha$):

$$\|\mathbf{s}\|_{\mathfrak{b}^{\alpha}_{p,q}} \le C(\log N)^{1/p}\|\mathbf{s}\|_{\mathfrak{f}^{\alpha}_{p,r}}, \quad \forall\, \mathbf{s} \in \Sigma_N,\ N = 2,3,\dots.$$

Of course, when $0 < r \le q$ we can remove the $\log N$, due to the embedding $\mathfrak{f}^{\alpha}_{p,r} \hookrightarrow \mathfrak{b}^{\alpha}_{p,q}$ (see Remark 5.3). For $q < r \le \infty$, we do not know whether the power in the $\log N$ can be improved.

6. Applications of Jackson and Bernstein Inequalities

6.1. Approximation and interpolation of sequence spaces. We can apply the general setting introduced by DeVore and Popov [12] to characterize approximation and interpolation spaces from Jackson and Bernstein inequalities. Recall that for a quasi-normed space X, a collection of subsets $\{\Sigma_N\}_{N=0}^{\infty}$ is an *approximating family* whenever:

1. $\Sigma_0 = \{0\} \subset \Sigma_1 \subset \Sigma_2 \subset \cdots \subset X$;
2. There exists $c \geq 1$ so that $\Sigma_N \pm \Sigma_M \subset \Sigma_{c(N+M)}$, $\forall\, N, M \geq 0$.

Given $\gamma > 0$ and $0 < q \leq \infty$ we define the *approximation space of order* (γ, q), $A_q^\gamma(X)$, as the space of vectors $x \in X$ for which

$$\|x\|_{A_q^\gamma(X)} := \|x\|_X + \left[\sum_{N=1}^{\infty} (N^\gamma \sigma_N(x))^q \frac{1}{N}\right]^{1/q} < \infty,$$

where as usual $\sigma_N(x) = \sigma_N(x)_X = \inf_{z\in\Sigma_N} \|x - z\|_X$.

Similarly, if $Y \hookrightarrow X$ are quasi-normed spaces, we define for $0 < \theta < 1$ and $0 < q \leq \infty$ the (*real*) *interpolation space* $(X, Y)_{\theta,q}$ as the space of vectors $x \in X$ for which

$$\|x\|_{(X,Y)_{\theta,q}} := \left[\int_0^\infty (t^{-\theta} K(t;x))^q \frac{dt}{t}\right]^{1/q} < \infty,$$

where $K(t;x) = K(t;x)_{X,Y} = \inf_{y\in Y}\{\|x - y\|_X + t\|y\|_Y\}$.

If we set $X = \mathfrak{f}$ be a sequence space with the approximating family Σ_N in Section 2, then a direct application of the formalism of DeVore-Popov produces the following result[1].

Theorem 6.1. *Let* $\mathfrak{f} \hookrightarrow \mathbb{C}^{\mathbb{1}}$ *be a* p*-space for some* $0 < p < \infty$*. Then, for all* $\gamma > 0$ *and* $0 < q \leq \infty$ *we have*

$$A_q^\gamma(\mathfrak{f}) = \ell_{\tau,q}(\mathfrak{f}), \quad \text{whenever } \frac{1}{\tau} = \gamma + \frac{1}{p}. \tag{6.1}$$

Proof. In Propositions 4.4 and 5.5 (with $r = \infty$) we have established the Jackson and Bernstein inequalities for the spaces $\mathfrak{f}$ and $\ell_{\tau,\infty}(\mathfrak{f})$ and every fixed $\tau < p$. Then, quoting Theorems 3.1 and 4.2 from [12] we obtain the following identities for all $0 < q \leq \infty$:

$$A_q^\beta(\mathfrak{f}) = (\mathfrak{f}, \ell_{\tau,\infty}(\mathfrak{f}))_{\beta/\gamma,q}, \quad \text{whenever } 0 < \beta < \gamma := \frac{1}{\tau} - \frac{1}{p}, \tag{6.2}$$

$$A_q^\gamma(\mathfrak{f}) = (\ell_{\tau_0,\infty}(\mathfrak{f}), \ell_{\tau_1,\infty}(\mathfrak{f}))_{(\gamma-\gamma_0)/(\gamma_1-\gamma_0),q}, \tag{6.3}$$

$$\text{when } 0 < \gamma_0 < \gamma < \gamma_1 \text{ and } \frac{1}{\tau_i} := \gamma_i + \frac{1}{p}.$$

[1] This result, in a slightly different context, has been independently obtained in [17], [21].

Observe that each space $\ell_{\tau_i,\infty}(\mathfrak{f})$ is isometrically isomorphic to the Lorentz space $\ell_{\tau_i,\infty}(\mathcal{I})$. Thus, using the interpolation properties of Lorentz spaces (see e.g. [1]) we see that the right hand side of (6.3) equals $\ell_{\tau,q}(\mathfrak{f})$, with $1/\tau = \gamma + 1/p$. □

Particularizing to the p-spaces $\mathfrak{f} = \mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^s_{p,p}$, we obtain the following corollary.

Corollary 6.2. *Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < r \le \infty$. Then, for all $\gamma > 0$*

$$A_\tau^{\gamma/d}(\mathfrak{f}^s_{p,r}) = A_\tau^{\gamma/d}(\mathfrak{b}^s_{p,p}) = \mathfrak{b}^{s+\gamma}_{\tau,\tau}, \quad \text{if } \frac{1}{\tau} = \frac{\gamma}{d} + \frac{1}{p}. \tag{6.4}$$

Proof. We know that $\ell_{\tau,\tau}(\mathfrak{f}^s_{p,r}) = \ell_{\tau,\tau}(\mathfrak{b}^s_{p,p}) = \mathfrak{b}^\alpha_{\tau,\tau}$ provided $\alpha/d - 1/\tau = s/d - 1/p$. Setting $\alpha = s + \gamma$, and using (6.1) (with $\tau = q$) we obtain the desired identity. □

Remark 6.3. When $p = \infty$, the proofs given for Propositions 4.4, 5.5 and Theorem 6.1 are also valid for the space $\mathfrak{b}^s_{\infty,\infty}$ (which is an "∞-space"), and therefore it holds

$$A_\tau^{1/\tau}(\mathfrak{b}^s_{\infty,\infty}) = \mathfrak{b}^{s+d/\tau}_{\tau,\tau}, \quad \text{if } 0 < \tau < \infty,\ s \in \mathbb{R}$$

(see also [20, Section 11.5]). In this case, however, condition (c) in Section 2 fails, meaning that greedy algorithms will not always be well defined.

Finally, we state a non-trivial interpolation property between $\mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^\alpha_{\tau,\tau}$.

Corollary 6.4. *Let $s \in \mathbb{R}$, $0 < p < \infty$, $0 < r \le \infty$. Then, for all $\gamma > 0$ and $0 < \theta < 1$ we have*

$$(\mathfrak{f}^s_{p,r}, \mathfrak{b}^{s+\gamma}_{\tau,\tau})_{\theta,\tau_\theta} = \mathfrak{b}^{s+\theta\gamma}_{\tau_\theta,\tau_\theta}, \quad \text{where} \quad \frac{1}{\tau} = \frac{\gamma}{d} + \frac{1}{p} \text{ and } \frac{1}{\tau_\theta} = \frac{\theta\gamma}{d} + \frac{1}{p}.$$

Proof. From Theorems 4.3 and 5.4 (with $\tau = q$), and the general theory of DeVore-Popov, we obtain for all $0 < q \le \infty$

$$A_q^{\beta/d}(\mathfrak{f}^s_{p,r}) = (\mathfrak{f}^s_{p,r}, \mathfrak{b}^{s+\gamma}_{\tau,\tau})_{\beta/\gamma,q}, \quad \text{whenever} \quad 0 < \beta < \gamma \quad \text{and } \frac{\gamma}{d} := \frac{1}{\tau} - \frac{1}{p}.$$

Using (6.4) we obtain the desired result. □

6.2. An abstract transference framework. In this section we present a general procedure which allows to transfer results from sequence spaces into distribution or function spaces. This is essentially the φ-transform formalism of Frazier and Jawerth [14], which nowadays can be applied in many different contexts.

The general setting can be described as follows. We consider a TVS $\mathcal{F}$ and its dual $\mathcal{F}'$, endowed with the ω^*-topology. Typically, $\mathcal{F}$ is a space of test functions containing a family of "analyzing wavelets" $\{\tilde\psi_I\}_{I\in\mathcal{I}}$, while its dual $\mathcal{F}'$ is a distribution space containing the "synthesizing wavelets" $\{\psi_I\}_{I\in\mathcal{I}}$. For example, one may have $\mathcal{F} = S_L(\mathbb{R}^d)$, the class of Schwartz functions in $\mathbb{R}^d$ with L vanishing moments, in which case we can identify $\mathcal{F}'$ with the space $S'/\mathcal{P}_L$ of distributions modulo polynomials of degree $< L$. We shall also consider below the case $\mathcal{F} = L^p$ and $\mathcal{F}' = L^{p'}$ for $1 < p < \infty$. Of course, many other situations are possible, depending on the regularity and vanishing moments of the wavelet systems. We also recall from Section 3 that, if $I = I_{j,\mathbf{k}}$ as in (3.1), then ψ_I is typically generated from a single ψ by $\psi_{I_{j,\mathbf{k}}}(x) = \lambda^{jd/2}\psi(M^j x - \mathbf{k})$, although our general setting admits as well other possibilities. Finally we observe that the same reasonings can be carried out for wavelet systems with more than one generator $\{\psi_I^1, \ldots, \psi_I^m\}_{I\in\mathcal{I}}$, after the obvious replacement of $\mathcal{I}$ by $\mathcal{I}\times\cdots\times\mathcal{I}$.

Let $\mathbf{f} = \mathbb{C}^{\mathcal{I}}$ denote the TVS of all sequences indexed by $\mathcal{I}$, and $\mathbf{f}_c$ those which are compactly supported. Observe that, with the product topology in $\mathbb{C}^{\mathcal{I}}$, the subspace $\mathbf{f}_c$ is dense, and the system of vectors $\{\mathbf{e}_I\}_{I\in\mathcal{I}}$ is an unconditional basis of $\mathbf{f}$.

Given a pair of systems $\tilde{\Psi} = \{\tilde{\psi}_I\}_{I\in\mathcal{I}} \subset \mathcal{F}$ and $\Psi = \{\psi_I\}_{I\in\mathcal{I}} \subset \mathcal{F}'$, we define the *analysis and synthesis operators* by

$$\begin{aligned} A_{\tilde{\Psi}} : \mathcal{F}' &\longrightarrow \mathbf{f} \\ u &\longmapsto \{\langle u, \tilde{\psi}_I\rangle\}_{I\in\mathcal{I}} \end{aligned}$$

and

$$\begin{aligned} S_{\Psi} : \mathbf{f}_c &\longrightarrow \mathcal{F}' \\ \{s_I\}_{I\in\mathcal{I}} &\longmapsto \sum_{I\in\mathcal{I}} s_I\psi_I. \end{aligned}$$

We say that $(\tilde{\Psi}, \Psi)$ is an *analysis-synthesis pair for* $\mathcal{F}'$ when S_Ψ can be continuously extended to the range of $A_{\tilde{\Psi}}$, and moreover, the identity $S_\Psi \circ A_{\tilde{\Psi}} = I_{\mathcal{F}'}$ holds. Equivalently, when every $u \in \mathcal{F}'$ can be written as $u = \sum_I \langle u, \tilde{\psi}_I\rangle\psi_I$, with unconditional convergence in the ω^*-topology of $\mathcal{F}'$.

We say that a quasi-normed space $F \hookrightarrow \mathcal{F}'$ is a *retract* of a quasi-normed sequence space $\mathfrak{f} \hookrightarrow \mathbf{f}$ via $(A_{\tilde{\Psi}}, S_{\tilde{\Psi}})$ whenever we can define in a bounded way the mappings

$$A_{\tilde{\Psi}} : F \longrightarrow \mathfrak{f} \quad \text{and} \quad S_\Psi : \mathfrak{f} \longrightarrow F, \tag{6.5}$$

and the identity $S_\Psi \circ A_{\tilde{\Psi}} = I_F$ holds. In this case we have the identification

$$F = \{u \in \mathcal{F}' : \{\langle u, \tilde{\psi}_I\rangle\}_I \in \mathfrak{f}\}, \tag{6.6}$$

and the equivalence of norms

$$\frac{1}{c}\|\{\langle u, \tilde{\psi}_I\rangle\}_I\|_{\mathfrak{f}} \le \|u\|_F \le c\|\{\langle u, \tilde{\psi}_I\rangle\}_I\|_{\mathfrak{f}}. \tag{6.7}$$

A well-known application of this setting gives interpolation and inclusion results for distribution spaces from the corresponding ones in sequence spaces (see, e.g., [1, p. 150]).

Proposition 6.5. *Let* F_0, $F_1 \hookrightarrow \mathcal{F}'$ *be two quasi-normed spaces which are, respectively, retracts of two sequence spaces* $\mathfrak{f}_0$ *and* $\mathfrak{f}_1$ *via* $(A_{\tilde{\Psi}}, S_\Psi)$. *Then* $\mathfrak{f}_1 \hookrightarrow \mathfrak{f}_0$ *implies* $F_1 \hookrightarrow F_0$. *Moreover, for every* $0 < \theta < 1$ *and* $0 < q \le \infty$ *the interpolation space* $(F_0, F_1)_{\theta,q}$ *is a retract of* $(\mathfrak{f}_0, \mathfrak{f}_1)_{\theta,q}$.

In the next proposition we summarize the transference results concerning approximation spaces. The proof is completely elementary and left to the reader.

Proposition 6.6. *Let* F, $B \hookrightarrow \mathcal{F}'$ *be two quasi-normed spaces which are, respectively, retracts of two sequence spaces* $\mathfrak{f}$ *and* $\mathfrak{b}$ *via* $(A_{\tilde{\Psi}}, S_\Psi)$. *Then,*

(1) *If* $\{S_N\}_{N=0}^\infty \subset F$ *is an approximating family for* F, *then so is* $\Sigma_N := A_{\tilde{\Psi}}(S_N)$ *for* $\mathfrak{f}$.

(2) *For all* $\gamma > 0$, $0 < q \le \infty$ *the approximation space* $A_q^\gamma(F; S_N)$ *is a retract of* $A_q^\gamma(\mathfrak{f}; \Sigma_N)$.

(3) *For every* $\varepsilon > 0$, *Jackson's inequality is preserved in the sense:*

$$\begin{aligned} &\sigma_N(\mathbf{s}, \Sigma_N)_{\mathfrak{f}} \lesssim N^{-\varepsilon}\|\mathbf{s}\|_{\mathfrak{b}}, \quad \forall\, \mathbf{s} \in \mathfrak{b} \\ \Rightarrow\ &\sigma_N(u, S_N)_F \lesssim N^{-\varepsilon}\|u\|_B, \quad \forall\, u \in B. \end{aligned}$$

(4) *For every* $\varepsilon > 0$, *Bernstein's inequality is preserved in the sense:*

$$\begin{aligned} &\|\mathbf{s}\|_{\mathfrak{b}} \lesssim N^{\varepsilon}\|\mathbf{s}\|_{\mathfrak{f}}, \quad \forall\, \mathbf{s} \in \Sigma_N \\ \Rightarrow\ &\|u\|_B \lesssim N^{\varepsilon}\|u\|_F, \quad \forall\, u \in S_N. \end{aligned}$$

Finally, we point out that, in order to transfer results about N-term approximation (from $\mathfrak{f}$ to F), we need the stronger condition that $(\tilde{\Psi}, \Psi)$ are *biorthogonal systems*, that is

$$\langle \tilde{\psi}_J, \psi_I \rangle = \delta_{I,J}, \quad \forall\, I, J \in \mathcal{I}. \tag{6.8}$$

This ensures that the operators $A_{\tilde{\Psi}}$ and S_Ψ in (6.5) are actually *isomorphisms* between F and $\mathfrak{f}$, and therefore that the set S_N of linear combinations of at most N ψ_I's, is mapped *exactly* onto the set Σ_N of sequences with at most N non-null entries.

The biorthogonality condition (6.8) is a natural requirement in many applications, but is also a bit too restrictive since it excludes for instance the use of Frazier-Jawerth ψ-functions, or more generally, frame systems in F. For the moment we do not know of any general result about N-term approximation involving systems which are not biorthogonal.

6.3. Applications to distribution spaces: some examples. We turn to specific applications of our results to wavelet decompositions of classical distribution and function spaces.

Throughout this section we fix a matrix M which is *expansive* (i.e., all its eigenvalues have modulus larger than 1). As in Section 3 we let $\lambda := |\det M|^{1/d} > 1$ and $I = I_{j,\mathbf{k}} = M^{-j}([0,1)^d + \mathbf{k})$, and for a function η we shall denote

$$\eta_I(x) = \eta_{j,\mathbf{k}}(x) = \lambda^{jd/2}\eta(M^j x - \mathbf{k}), \quad j \in \mathbb{Z},\ \mathbf{k} \in \mathbb{Z}^d.$$

As a first assumption, suppose we are given two functions $\tilde{\psi}$, $\psi \in L^2(\mathbb{R}^d)$ for which the systems $\tilde{\Psi} = \{\tilde{\psi}_I\}$ and $\Psi = \{\psi_I\}$ form an *analysis-synthesis pair for* $L^2(\mathbb{R}^d)$. This happens, e.g., if $(\tilde{\Psi}, \Psi)$ is a pair of dual frames in $L^2(\mathbb{R}^d)$. For the study of distribution spaces one must require some extra regularity in $\tilde{\psi}$ and ψ, which is typically measured as follows:

1. *r-regularity*: the function $\eta(x) \in C^r(\mathbb{R}^d)$ and

$$\sup_{x \in \mathbb{R}^d} |(1+|x|)^n D^\alpha \eta(x)| < \infty, \quad \forall\, n \geq 1,\ 0 \leq |\alpha| \leq r;$$

2. *ℓ-vanishing moments:* for all $0 \leq |\alpha| < \ell$

$$\int_{\mathbb{R}^d} x^\alpha \eta(x)\, dx = 0.$$

We shall denote by $S^r_\ell(\mathbb{R}^d)$ the class of all r-regular functions with ℓ-vanishing moments.

Systems $(\tilde{\Psi}, \Psi)$ as above are usually called *wavelet systems*, and under certain conditions provide frames or unconditional bases for many classical distribution spaces. The first examples were initially introduced for the matrix $M = 2I$ [14, 26], and later on extended to more general expansive matrices. We refer to [3–5, 29, 31] for many such constructions whose main features we shall not discuss here. We recall one more time that these systems often have more than one generator: $\{\tilde{\psi}^1, \dots, \tilde{\psi}^m; \psi^1, \dots, \psi^m\}$, but for notational simplicity there is no loss if we just consider $m = 1$.

For the study of distribution spaces we shall fix r, $\ell \in \mathbb{N} \cup \{\infty\}$, and let $\mathcal{F} = S^r_\ell(\mathbb{R}^d)$. Observe that $r = \ell = \infty$ corresponds to Schwartz functions with infinite vanishing moments, and hence $\mathcal{F}'$ is the space of tempered distributions modulo polynomials. In general, $\mathcal{F}'$ consists of distributions of order r modulo polynomials of degree $< \ell$. We shall make the final assumption that $(\tilde{\Psi}, \Psi)$ *is an analysis-synthesis pair for* $\mathcal{F}'$, that is, the identity $u = \sum_I \langle u, \tilde{\psi}_I \rangle \psi_I$ also holds with ω^*-unconditional convergence in $\mathcal{F}'$. This is typically a consequence of the regularity in ψ, $\tilde{\psi}$, but we shall not discuss this matter here (see e.g. [24]).

Next, associated with the sequence spaces $\mathfrak{f}^s_{p,r}$ and $\mathfrak{b}^\alpha_{\tau,q}$ in Section 3 we can define in an abstract way corresponding distribution spaces by (6.6) and (6.7), that is

$$(6.9)\qquad \dot F^s_{p,r} = \dot F^s_{p,r}(\widetilde\Psi,\Psi,\mathcal F') := \{u\in\mathcal F' : \|u\|_{\dot F^s_{p,r}} := \|\langle u,\tilde\psi_I\rangle\|_{\mathfrak f^s_{p,r}} < \infty\}$$

$$(6.10)\qquad \dot B^\alpha_{\tau,q} = \dot B^\alpha_{\tau,q}(\widetilde\Psi,\Psi,\mathcal F') := \{u\in\mathcal F' : \|u\|_{\dot B^\alpha_{\tau,q}} := \|\langle u,\tilde\psi_I\rangle\|_{\mathfrak b^\alpha_{\tau,q}} < \infty\}.$$

It follows from the definition that $\dot F^s_{p,r}$ is a retract of $\mathfrak f^s_{p,r}$ and similarly for $\dot B^\alpha_{\tau,q}$ and $\mathfrak b^\alpha_{\tau,q}$. When $\mathcal F = S^r_\ell$ are large enough, these spaces are related with the *classical anisotropic homogeneous Triebel-Lizorkin and Besov spaces* (see [33], or [13] for a general definition in anisotropic cases). In fact, a main topic in wavelet theory is to show that such equivalences hold with minimal assumptions on r and ℓ. We refer to [26] for the first results when $M = 2I$, and to [2, 24] for recent research concerning more general matrices. The scope of our general framework is to cover all these situations independently of further refinements which concerning such characterizations may appear in the future.

Our main result below is a direct application of the theory in Sections 6.1, 6.2 to the distribution spaces in (6.9) and (6.10). Combined with the above mentioned characterization of classical spaces, it extends in particular results from [9, 20, 23].

Corollary 6.7. *Let $(\widetilde\Psi,\Psi)$ be a biorthogonal analysis-synthesis pair as above, and consider the approximating sets S_N, of linear combinations of at most N ψ_I's. Let $s\in\mathbb R$, $0<p<\infty$, $0<r\le\infty$, and $\dot F^s_{p,r}$ be defined as in (6.9). Then, for all $\gamma>0$ and $0<q\le\infty$,*

$$A_q^{\gamma/d}(\dot F^s_{p,r};S_N) = (\dot F^s_{p,r},\dot B^{s+\beta_1}_{\tau_1,\tau_1})_{\gamma/\beta_1,q}$$
$$= (\dot B^{s+\beta_0}_{\tau_0,\tau_0},\dot B^{s+\beta_1}_{\tau_1,\tau_1})_{(\gamma-\beta_0)/(\beta_1-\beta_0),q},$$

when $\beta_1>\gamma>\beta_0>0$ and $1/\tau_i = \beta_i/d + 1/p$, $i = 0, 1$. Moreover,

$$(6.11)\qquad \begin{aligned} A_q^{\gamma/d}(\dot F^s_{p,r};S_N) &= A_q^{\gamma/d}(\dot B^s_{p,p};S_N)\\ &= \dot B^{s+\gamma}_{q,q}, \end{aligned}\qquad \text{if } \frac1q = \frac\gamma d + \frac1p.$$

In applications, it is perhaps of more interest to consider *spaces of functions* (rather than equivalence classes of distributions). A typical case in the literature is the n-term approximation by means of wavelets in L^p-spaces [8, 9]. In our setting, this corresponds to taking $\mathcal F' = L^p(\mathbb R^d)$, for a fixed $1<p<\infty$. It is well known that under certain conditions in $\tilde\psi$, ψ, the systems $(\widetilde\Psi,\Psi)$ become a pair of unconditional bases in L^p, and moreover, $L^p = \dot F^0_{p,2}(\widetilde\Psi,\Psi,L^p)$ as defined in (6.9) (see e.g. [2, 28] for a proof involving systems with expansive matrices). If we denote $\mathcal B^\alpha_{\tau,q} := \dot B^\alpha_{\tau,q}(\widetilde\Psi,\Psi,L^p)$, then from the previous corollary we obtain the identity

$$A_q^{\gamma/d}(L^p(\mathbb{R}^d); S_N) = \mathcal{B}^\gamma_{q,q}, \quad \text{if} \quad \frac{1}{q} = \frac{\gamma}{d} + \frac{1}{p}. \tag{6.12}$$

It should be pointed out, however, that $\mathcal{B}^\gamma_{q,q}$ *is not equal to the classical Besov space* $B^\gamma_{q,q}(\mathbb{R}^d)$, as it has been stated by some authors in the literature. If for simplicity we consider the isotropic case, and take the usual definition:

$$B^\alpha_{\tau,q}(\mathbb{R}^d) = \Big\{ u \in L^\tau(\mathbb{R}^d) \\ : |u|_{B^\alpha_{\tau,q}(\mathbb{R}^d)} = \Big[\sum_{i=1}^d \int_0^\infty (t^{-\alpha} \|\Delta^{[\alpha]+1}_{t\mathbf{e}_i} u\|_\tau)^q \frac{dt}{t} \Big]^{1/q} < \infty \Big\}, \tag{6.13}$$

then for $\alpha > d(1/\tau - 1)_+$ and $1/p := 1/\tau - \alpha/d$ it is well known that $B^\alpha_{\tau,\tau}(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$ and, moreover, for sufficiently nice wavelets,

$$c_1 |u|_{B^\alpha_{\tau,\tau}(\mathbb{R}^d)} \leq \|\{\langle u, \tilde{\psi}_I \rangle\}_I\|_{\mathfrak{b}^\alpha_{\tau,\tau}} \leq c_2 |u|_{B^\alpha_{\tau,\tau}(\mathbb{R}^d)}$$

(see e.g. [9], or in the anisotropic case [16]). That is, $B^\alpha_{\tau,\tau}(\mathbb{R}^d) \hookrightarrow \mathcal{B}^\alpha_{\tau,\tau}$. However, it is easy to construct examples of functions $u \in \mathcal{B}^\alpha_{\tau,\tau}$ *which are not in* $L^\tau(\mathbb{R}^d)$, and thus cannot belong to $B^\alpha_{\tau,\tau}(\mathbb{R}^d)$. This is a typical phenomenon of "growth at infinity," which does not happen in the "compact" case $B^\alpha_{\tau,\tau}[0,1]^d$.

As a final application, which takes care of this phenomenon and is probably closer to the problems in signal compression, we replace $(\tilde{\Psi}, \Psi)$ by the *truncated systems* (at a fixed level $j_0 \in \mathbb{Z}$):

$$\tilde{\Psi}^+ = \{\tilde{\psi}_I : |I| \leq \lambda^{-j_0 d}\} \cup \{\tilde{\varphi}_I : |I| = \lambda^{-j_0 d}\},$$

and likewise for Ψ^+. In practice, $\tilde{\varphi}$ and φ are a pair of *scaling functions* from which one derives $\tilde{\psi}$ and ψ via usual multiresolution analysis methods [7, 18, 26]. Typically, φ, $\tilde{\varphi}$ are supposed to be r-regular but with no vanishing moments. To adapt our setting to such applications, we take as starting assumption that $(\tilde{\Psi}^+, \Psi^+)$ *is a biorthogonal analysis-synthesis pair for* $L^2(\mathbb{R}^d)$ (which happens, e.g., if these are dual Riesz bases in L^2). The new analysis operator $u \mapsto A_{\tilde{\Psi}^+}(u) = \{s_{j,\mathbf{k}}\}$ is defined by:

$$\begin{aligned} s_{j,\mathbf{k}} &= \langle u, \tilde{\psi}_{j,\mathbf{k}} \rangle \qquad \text{if } j \geq j_0, \\ s_{j_0-1,\mathbf{k}} &= \langle u, \tilde{\varphi}_{j_0,\mathbf{k}} \rangle, \\ s_{j,\mathbf{k}} &= 0 \qquad\qquad\quad \text{if } j < j_0 - 1, \end{aligned}$$

and similarly for S_Ψ. Thus, the corresponding sequence spaces $\mathfrak{f}^{s,+}_{p,r}$ and $\mathfrak{b}^{\alpha,+}_{\tau,q}$ are subspaces of the previous ones, consisting of sequences supported in the truncated set

$$\begin{aligned}\mathcal{I}^+ &= \{I \in \mathcal{I} : |I| \le \lambda^{-(j_0-1)d}\}\\ &= \{I_{j,\mathbf{k}} : j \ge j_0 - 1,\ \mathbf{k} \in \mathbb{Z}^d\}.\end{aligned}$$

This new setting is also quite natural for the study of function spaces. As before we set $\mathcal{F}' = L^p(\mathbb{R}^d)$ for fixed $1 < p < \infty$. If for simplicity we assume that M is diagonal, then it is known that the abstract space $F^s_{p,r}(\tilde{\Psi}^+, \Psi^+, L^p)$ (defined as in (6.9)) coincides with the classical *inhomogeneous Triebel-Lizorkin space* $F^s_{p,r}(\mathbb{R}^d)$, with an equivalence of norms

$$\|u\|_{F^s_{p,r}(\mathbb{R}^d)} \sim \|u\|_{L^p(\mathbb{R}^d)} + \|A_{\tilde{\Psi}^+}(u)\|_{\mathfrak{f}^s_{p,r}} \sim \|A_{\tilde{\Psi}^+}(u)\|_{\mathfrak{f}^s_{p,r}},$$

at least for $s \ge 0$ and regular enough wavelet bases [24]. This is also the case for the space $B^\alpha_{\tau,\tau}(\tilde{\Psi}^+, \Psi^+, L^p)$, which when $\alpha > d(1/\tau - 1)_+$ and $1/p = 1/\tau - \alpha/d$ coincides with the *inhomogeneous Besov space* $B^\alpha_{\tau,\tau}(\mathbb{R}^d)$ with the norm equivalence

$$\|u\|_{B^\alpha_{\tau,\tau}(\mathbb{R}^d)} \sim \|u\|_{L^\tau(\mathbb{R}^d)} + \|A_{\tilde{\Psi}^+}(u)\|_{\mathfrak{b}^\alpha_{\tau,\tau}} \sim \|A_{\tilde{\Psi}}(u)\|_{\mathfrak{b}^\alpha_{\tau,\tau}}$$

(see [24], and [16] for a general L^p formulation).

Since Jackson and Bernstein inequalities remain valid for the subspaces $\mathfrak{f}^{s,+}_{p,r}$ and $\mathfrak{b}^{\alpha,+}_{\tau,q}$, one obtains an analog of Corollary 6.7 for the spaces $F^s_{p,r}(\tilde{\Psi}^+, \Psi^+, \mathcal{F}')$ and $B^{s+\beta}_{\tau,\tau}(\tilde{\Psi}^+, \Psi^+, \mathcal{F}')$. In particular, for sufficiently regular wavelet bases we obtain the characterization:

$$\begin{aligned}A^{\gamma/d}_q(F^s_{p,r}(\mathbb{R}^d); S^+_N) &= A^{\gamma/d}_q(B^s_{p,p}(\mathbb{R}^d); S^+_N)\\ &= B^{s+\gamma}_{q,q}(\mathbb{R}^d), \quad \text{if} \quad \frac{1}{q} = \frac{\gamma}{d} + \frac{1}{p},\end{aligned}$$

where the approximating family S^+_N is made up of all linear combinations of at least N functions in Ψ^+. When $s = 0$ and $r = 2$, the identity becomes $A^{\gamma/d}_q(L^p; S^+_N) = B^{s+\gamma}_{q,q}$, which corrects the phenomenon in (6.12) for approximation with $\{S_N\}$ in $L^p(\mathbb{R}^d)$.

Acknowledgements. Gustavo Garrigós was supported by grant BFM2001-0189 (MEC) and "Programa Ramón y Cajal," 2001 (MCyT). Eugenio Hernández was supported by grant BFM2001-0189 (MEC).

The authors wish to thank an anonymous referee for comments and references leading to a considerable improvement in the presentation of this paper.

REFERENCES

[1] JÖRAN BERGH and JÖRGEN LÖFSTRÖM, *Interpolation Spaces. An Introduction*, Springer-Verlag, Berlin, 1976. MR0482275 (58 #2349)

[2] MARCIN BOWNIK, *Anisotropic Hardy spaces and wavelets*, Ph.D. Thesis, Washington University, St. Louis, 1999, http://www.math.lsa.umich.edu/~marbow.

[3] ______, *The construction of r-regular wavelets for arbitrary dilations*, J. Fourier Anal. Appl. **7** (2001), 489–506. MR1845100 (2002i:42044)

[4] MARCIN BOWNIK and DARRIN SPEEGLE, *Meyer type wavelet bases in* $\mathbf{R}^2$, J. Approx. Theory **116** (2002), 49–75. MR1909012 (2003g:42056)

[5] ANDREA CALOGERO, *Wavelets on general lattices, associated with general expanding maps of* $\mathbf{R}^n$, Electron. Res. Announc. Amer. Math. Soc. **5** (1999), 1–10 (electronic). MR1667201 (99i:42042)

[6] ALBERT COHEN, WOLFGANG DAHMEN, and RONALD A. DEVORE, *Adaptive wavelet methods for elliptic operator equations: convergence rates*, Math. Comp. **70** (2001), 27–75 (electronic). MR1803124 (2002h:65201)

[7] ALBERT COHEN, *Wavelet methods in numerical analysis*, Handbook of Numerical Analysis, Vol. VII, Handb. Numer. Anal., VII, North-Holland, Amsterdam, 2000, pp. 417–711. MR1804747 (2002c:65252)

[8] ALBERT COHEN, RONALD A. DEVORE, and REINHARD HOCHMUTH, *Restricted nonlinear approximation*, Constr. Approx. **16** (2000), 85–113. MR1848843 (2002g:41019)

[9] RONALD A. DEVORE, BJÖRN JAWERTH, and VASIL A. POPOV, *Compression of wavelet decompositions*, Amer. J. Math. **114** (1992), 737–785. MR1175690 (94a:42045)

[10] RONALD A. DEVORE, *Nonlinear approximation*, Acta Numerica, 1998, Acta Numer., vol. 7, Cambridge Univ. Press, Cambridge, 1998, pp. 51–150. MR1689432 (2001a:41034)

[11] RONALD A. DEVORE, GUERGANA PETROVA, and VLADIMIR N. TEMLYAKOV, *Best basis selection for approximation in* L_p, Found. Comput. Math. **3** (2003), 161–185. MR1966298 (2004c:41033)

[12] RONALD A. DEVORE and VASIL A. POPOV, *Interpolation spaces and nonlinear approximation*, Function Spaces and Applications (Lund, 1986), Lecture Notes in Math., vol. 1302, Springer, Berlin, 1988, pp. 191–205. MR942269 (89d:41035)

[13] PETER DINTELMANN, *Classes of Fourier multipliers and Besov-Nikolskij spaces*, Math. Nachr. **173** (1995), 115–130. MR1336956 (96h:42007)

[14] MICHAEL FRAZIER and BJÖRN JAWERTH, *A discrete transform and decompositions of distribution spaces*, J. Funct. Anal. **93** (1990), 34–170. MR1070037 (92a:46042)

[15] ______, *The φ-transform and applications to distribution spaces*, Function Spaces and Applications (Lund, 1986), Lecture Notes in Math., vol. 1302, Springer, Berlin, 1988, pp. 223–246. MR942271 (89g:46064)

[16] GUSTAVO GARRIGÓS, REINHARD HOCHMUTH, and ANITA TABACCO, *Wavelet characterizations for anisotropic Besov spaces with* $0 < p < 1$, Proc. Edin. Math. Soc. **47** (2004), 573–595, http://www.uam.es/gustavo.garrigos.

[17] RÉMI GRIBONVAL and MORTEN NIELSEN, *Some remarks on non-linear approximation with Schauder bases*, East J. Approx. **7** (2001), 267–285. MR1861393 (2002i:46007)

[18] EUGENIO HERNÁNDEZ and GUIDO WEISS, *A First Course on Wavelets*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1996, ISBN 0-8493-8274-2. MR1408902 (97i:42015)

[19] REINHARD HOCHMUTH, *Wavelet characterizations for anisotropic Besov spaces*, Appl. Comput. Harmon. Anal. **12** (2002), 179–208. MR1884234 (2002m:46051)

[20] C. C. HSIAO, BJÖRN JAWERTH, BRADLEY J. LUCIER, and XIANG MING YU, *Near Optimal Compression of Orthonormal Wavelet Expansions*, Wavelets: Mathematics and Applications, Stud. Adv. Math., CRC, Boca Raton, FL, 1994, pp. 425–446. MR1247523 (94h:42060)

[21] GÉRARD KERKYACHARIAN and DOMINIQUE PICARD, *Entropy, universal coding, approximation, and bases properties*, Constr. Approx. **20** (2004), 1–37. MR2025412 (2004k:41052)

[22] SERGEI V. KONYAGIN and VLADIMIR N. TEMLYAKOV, *A remark on greedy approximation in Banach spaces*, East J. Approx. **5** (1999), 365–379. MR1716087 (2000j:46020)

[23] GEORGE KYRIAZIS, *Non-linear approximation and interpolation spaces*, J. Approx. Theory **113** (2001), 110–126. MR1866250 (2002h:41036)

[24] ______, *Multilevel characterizations of anisotropic function spaces*, SIAM J. Math. Anal. **36** (2004), 441–462.

[25] CHRISTOPHER LEISNER, *Nonlinear wavelet approximation in anisotropic Besov spaces*, Indiana Univ. Math. J. **52** (2003), 437–455. MR1976084 (2004c:42069)

[26] YVES MEYER, *Ondelettes et Opérateurs. I*, Actualités Mathématiques. [Current Mathematical Topics], Hermann, Paris, 1990, ISBN 2-7056-6125-0. MR1085487 (93i:42002) (French)

[27] ______, *Wavelets and functions with bounded variation from image processing to pure mathematics*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. (2000), 77–105. MR1845666 (2002h:42076)

[28] WALDEMAR POMPE, *Unconditional biorthogonal wavelet bases in* $L^p(\mathbf{R}^d)$, Colloq. Math. **92** (2002), 19–34. MR1899234 (2003c:42038)

[29] MAURA SALVATORI and PAOLO M. SOARDI, *Multivariate compactly supported biorthogonal spline wavelets*, Ann. Mat. Pura Appl. (4) **181** (2002), 161–179. MR1911934 (2003g:42064)

[30] ELIAS M. STEIN and GUIDO WEISS, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, N.J., 1971. MR0304972 (46 #4102)

[31] ROBERT S. STRICHARTZ, *Wavelets and self-affine tilings*, Constr. Approx. **9** (1993), 327–346. MR1215776 (94f:42039)

[32] VLADIMIR N. TEMLYAKOV, *Nonlinear methods of approximation*, Found. Comput. Math. **3** (2003), 33–107. MR1951502 (2003j:41029)

[33] HANS TRIEBEL, *Theory of function spaces* **78** (1983), 284. MR781540 (86j:46026)

Matemáticas
Universidad Autónoma de Madrid
28049, Madrid, Spain .
E-MAIL, Gustavo Garrigós: `gustavo.garrigos@uam.es`
E-MAIL, Eugenio Hernández: `eugenio.hernandez@uam.es`

KEY WORDS AND PHRASES: Non-linear approximation, greedy algorithm, anisotropic Besov and Triebel-Lizorkin spaces, Lorentz spaces, φ-transforms, wavelets.

2000 MATHEMATICS SUBJECT CLASSIFICATION: 41A17, 42C40.

Received: *October 21st, 2003; revised: May 19th, 2004.*