AALBORG UNIVERSITY

Some Remarks on Shrinkage Operators

by

Lasse Borup and Morten Nielsen

March 2003 R-2003-07

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

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

SOME REMARKS ON SHRINKAGE OPERATORS

LASSE BORUP[†] AND MORTEN NIELSEN

ABSTRACT. Wavelet shrinkage is known to be a useful practical method to obtain various types of nonlinear approximations to smooth functions in $L_p(\mathbb{R}^d)$. In this note we point out that it is the fact that wavelets form a so-called greedy basis in various $L_p(\mathbb{R}^d)$-spaces that make wavelet shrinkage work. We study thresholding and shrinkage operators for unconditional and greedy bases in a Banach space, and show that shrinkage operators "behave well" for such bases. Optimal convergence rates for a family of thresholding operators are obtained. Finally, we study shrinkage operators for certain types of redundant systems.

1. INTRODUCTION

Wavelet shrinkage is a method to obtain nonlinear approximants with "nice properties" to a given signal. The method has proven to be a succesfull tool in signal denoising and other areas such as non-parametric regression [6, 8]. Various shrinkage operators have been studied in recent years [17, 6, 7, 1]. The general idea is simple, we expand a given signal $f \in L_p(\mathbb{R}^d)$ in a sufficiently "nice" wavelet basis $\{\psi_k\}$: $f = \sum_k \langle f, \psi_k\rangle \psi_k$. An approximant to f is then obtained by applying the shrinkage rule $\delta(x,\lambda)$ to the wavelet coefficients of f,

$$f_\lambda = \sum_k \delta(\langle f, \psi_k\rangle, \lambda)\psi_k.$$

The shrinkage rule δ should be designed so that $f_\lambda \to f$ as $\lambda \to 0$ (the precise definition of a shrinkage rule will be given in Section 2). There are an infinite number of ways to choose the shrinkage rule δ to make this scheme work. Let us mention two well-known rules. The simplest shrinkage rule is hard thresholding, with $\delta(x,\lambda) = x\mathbf{1}_{(|x|>\lambda)}(x)$, and soft thresholding where $\delta(x,\lambda) = x\{1-\frac{\lambda}{|x|}\}\mathbf{1}_{(|x|>\lambda)}(x)$. For a smooth function f measured on either the Besov or Triebel-Lizorkin scale, one can also estimate the convergence rate of f_λ to f.

The smooth and hard thresholding rules have been investigated from a statistical point of view by Donoho and Johnstone in a series of papers (see [6, 5] and references therein). It has been shown that these operators when applied on wavelet coefficients give asymptotic near-minimax optimality properties.

One may well wonder why all those shrinkage schemes work well for wavelet bases in $L_p(\mathbb{R}^d)$? The purpose of this note is to identify the properties of wavelet

---

2000 *Mathematics Subject Classification.* Primary 41A17; Secondary 41A25.

Key words and phrases. wavelet bases, shrinkage operators, greedy bases, nonlinear approximation.

[†]This work is in part supported by the Danish Technical Science Foundation, Grant no. 9701481.

systems in $L_p(\mathbb{R}^d)$ that make shrinkage work well for such bases. With this information at hand, it is then possible to construct other bases that are also well adapted for shrinkage.

It turns out that the fact that wavelets are unconditional bases for $L_p(\mathbb{R}^d)$, $1 < p < \infty$, is important, but this is not all. The wavelets also form a so-called greedy basis for $L_p(\mathbb{R}^d)$, which is crucial for obtaining estimates for the convergence rate of shrinkage operators applied to smooth functions. In Section 2 we define unconditional bases, and we prove that shrinkage operators associated with such bases do converge. For unconditional bases, shrinkage always converges, but if convergence rates for "nice functions" are required, there is a need to impose additional conditions closely related to the concept of a greedy basis. Greedy type bases and convergence rates for shrinkage operators applied to "nice" functions are discussed in Section 3. In Section 3 we also study thresholding operators in more detail, and compare the approximation rate obtained from such operators with the rate from best nonlinear m-term approximation.

In Section 4 we generalize some of the results on shrinkage operators to atomic decompositions of Banach spaces. An atomic decomposition is basically a redundant system with an associated canonical expansion. Two examples of atomic decompositions where the estimates apply are given at the end of Section 4. The first example is tight wavelet frames in $L_p(\mathbb{R}^d)$ and the second example concerns Gabor systems in modulation spaces.

2. Shrinkage Operators for Unconditional bases

In this section we define and study shrinkage rules and their associated operators in a general Banach space. Before we can apply any type of shrinkage rule, we need to expand a given function in a fixed basis for the space, and then modify – or rather shrink – the size of the expansion coefficient from the expansion. This is not always a well defined operation if we have only a Schauder basis, so we need to work with unconditional bases. Let us recall the definition of an unconditional basis for a Banach space. Unconditional bases are the most general systems where shrinkage works.

Definition 2.1. *Let X be a separable Banach space. A family $\mathcal{B} = \{g_k\}_{k=1}^{\infty} \subset X$ is called a* ***Schauder basis*** *for X if*

- *$Span(\{g_k\}_{k=1}^{\infty})$ is dense in X,*
- *Every element $f \in X$ has a unique norm-convergent expansion*

$$f = \sum_{k=1}^{\infty} \alpha_k g_k.$$

A Schauder basis $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ for X is called an ***unconditional basis*** *if it satisfies*

- *There exists a constant K such that for any two sequences $\{\alpha_k\}_{k=1}^{\infty}$ and $\{\beta_k\}_{k=1}^{\infty}$ with $|\alpha_k| \le |\beta_k|$ for $k \in \mathbb{N}$, we have*

$$\Big\| \sum_{k=1}^{N} \alpha_k g_k \Big\|_X \le K \Big\| \sum_{k=1}^{N} \beta_k g_k \Big\|_X, \qquad N \in \mathbb{N}. \tag{2.1}$$

For a Schauder basis $\{g_k\}_{k=1}^{\infty} \subset X$ we let $\{g_k^\star\}_{k=1}^{\infty} \subset X^\star$ denote the corresponding dual system. Now we give the precise definition of a shrinkage rule.

Definition 2.2. *The function* $\delta\colon \mathbb{C}\times\mathbb{R}^+ \mapsto \mathbb{C}$ *is called a* ***shrinkage rule*** *if there exists a constant* $A<\infty$ *such that*

$$|x-\delta(x,\lambda)| \le A\min\{|x|,\lambda\}.$$

Remark 2.3. Hard and soft thresholding easily fit into this general scheme, but there are also interesting examples of shrinkage rules that are not thresholding operators. One example is the n-degree garotte rule given by $\delta(x,\lambda) = x^{2n+1}/(\lambda^{2n} + x^{2n})$ for $x\in\mathbb{C}$. For the sake of generality, we do allow *degenerated* shrinkage rules, such as $\delta(x,\lambda) := x$, even though no shrinkage of $|x|$ takes place.

Given a quasi-normalized unconditional basis $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ for X, i.e., an unconditional basis for which there are constants, $0<c,C<\infty$, such that $c \le \|g_k\|_X \le C$ for $k\in\mathbb{N}$. We define the *shrinkage operator* $T_\lambda^\delta\colon X\mapsto X$, associated with the shrinkage rule δ, by

$$T_\lambda^\delta f = \sum_{k=1}^{\infty} \delta(g_k^\star(f),\lambda)g_k.$$

Notice that T_λ^δ is well-defined and bounded on X since $\mathcal{B}$ is unconditional. The idea is to use $T_\lambda^\delta f$ as an approximation to f for small λ. Let us first verify that this is possible.

Proposition 2.4. *Let* $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ *be an unconditional basis for the Banach space* X. *Let* T_λ^δ *be a shrinkage operator associated with* δ. *Then* $\|f - T_\lambda^\delta f\|_X \to 0$ *as* $\lambda\to 0$ *for every* $f\in X$.

Proof. Let $f\in X$. Given $\varepsilon>0$ we choose M such that $\left\|\sum_{k>M} g_k^\star(f)g_k\right\|_X \le \varepsilon$. Notice that $\left\|\sum_{k=1}^{M}\big(g_k^\star(f)-\delta(g_k^\star(f),\lambda)\big)g_k\right\|_X \to 0$ as $\lambda\to 0$ since $\delta(g_k^\star(f),\lambda)\to g_k^\star(f)$ as $\lambda\to 0$ for every $k\in\mathbb{N}$. Hence, we have

$$\begin{aligned}
\|f - T_\lambda^\delta f\|_X &= \left\| \sum_{k=1}^{\infty} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)g_k \right\|_X \\
&\le \left\| \sum_{k=1}^{M} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)g_k \right\|_X + \left\| \sum_{k>M} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)g_k \right\|_X \\
&\le \left\| \sum_{k=1}^{M} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)g_k \right\|_X + AK\left\| \sum_{k>M} g_k^\star(f)g_k \right\|_X \\
&\le \left\| \sum_{k=1}^{M} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)g_k \right\|_X + AK\varepsilon,
\end{aligned}$$

where we have used the uniform estimate $|g_k^\star(f) - \delta(g_k^\star(f),\lambda)| \le A|g_k^\star(f)|$ and property (2.1). Thus, for sufficiently small λ we have $\|f - T_\lambda^\delta f\|_X \le 2AK\varepsilon$. □

3. Greedy-Type Systems and Shrinkage Operators

At this point we know that shrinkage operators converge for a general unconditional basis in a Banach space. However, what is missing is estimates for the convergence *rate* when we use the operator to approximate the functions in X. For a general function in X, the convergence rate of the shrinkage operator can be arbitrarily slow as can be seen from simple Hilbert space examples. Thus, we need to focus on a class of "nice" functions in X. For wavelet type bases in $L_p(\mathbb{R}^d)$, "nice" usually means that the function should be contained in an appropriate Besov or Triebel-Lizorkin space. However, in the abstract setup, we cannot be so specific and below we introduce generalized smoothness or sparsity spaces.

Since we are going to study more general systems of functions than Schauder bases in Section 4, we will give the sparsity classes for a general dictionary. A dictionary $\mathcal{D}$ in X is a countable collection of quasi-normed elements that span a dense subset of X.

Recall the Lorentz space $\ell_{p,q}$, $0 < p < \infty$, $0 < q \leq \infty$, the set of sequences $\{a_m\}$ satisfying $\|\{a_m\}\|_{\ell_{p,q}} < \infty$, where

$$\|\{a_m\}\|_{\ell_{p,q}} = \begin{cases} \left(\sum_{j=0}^{\infty} \left(2^{j/p} a^*_{2^j}\right)^q\right)^{1/q}, & 0 < q < \infty, \\ \sup_{j\geq 0} 2^{j/p} a^*_{2^j}, & q = \infty, \end{cases}$$

with $\{a^*_j\}$ a decreasing rearrangement of $\{a_m\}$.

Definition 3.1 (Sparsity Class)**.** *Given a dictionary $\mathcal{D} = \{g_k\}$. For $\tau \in (0,\infty)$, $q \in (0,\infty]$, and $M < \infty$, we let $\mathcal{K}_{\tau,q}(X,\mathcal{D},M)$ denote the set*

$$\operatorname{clos}_X \left\{ f \in X \,\middle|\, \exists \Lambda \subset \mathbb{N}, \operatorname{card}\Lambda < \infty, f = \sum_{k\in\Lambda} c_k g_k, \|\{c_k\}\|_{\ell_{\tau,q}} \leq M \right\}.$$

Then we define

$$\mathcal{K}_{\tau,q}(X,\mathcal{D}) := \bigcup_{M>0} \mathcal{K}_{\tau,q}(X,\mathcal{D},M),$$

with $|f|_{\mathcal{K}_{\tau,q}(X,\mathcal{D})} = \inf\{M \colon f \in \mathcal{K}_{\tau,q}(X,\mathcal{D},M)\}$.

Since $\mathcal{B}$ is an unconditional basis, and therefore also a Schauder basis, we have the following more transparent characterization of $\mathcal{K}_{\tau,q}(X,\mathcal{B})$.

Lemma 3.2. *Let $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ be a quasi-normalized Schauder basis for X. For $\tau \in (0,\infty)$ and $q \in (0,\infty]$,*

$$\mathcal{K}_{\tau,q}(X,\mathcal{B}) = \left\{ f \in X \,\middle|\, \{g^\star_k(f)\}_{k\in\mathbb{N}} \in \ell_{\tau,q} \right\}, \tag{3.1}$$

and $|f|_{\mathcal{K}_{\tau,q}(X,\mathcal{B})} = \|\{g^\star_k(f)\}_k\|_{\ell_{\tau,q}}$.

Proof. Denote the right hand side of (3.1) by $\tilde{\mathcal{K}}_{\tau,q}(X,\mathcal{B})$. Suppose $f \in \tilde{\mathcal{K}}_{\tau,q}(X,\mathcal{B})$, with $\|\{g^\star_k(f)\}_k\|_{\ell_{\tau,q}} \leq M < \infty$, and let $f_n = \sum_{k=1}^n g^\star_k(f) g_k$ for $n \in \mathbb{N}$. Then, clearly $f_n \in \mathcal{K}_{\tau,q}(X,\mathcal{B},M)$ for all $n \in \mathbb{N}$. Furthermore, since $\mathcal{B}$ is a Schauder basis for X, we have $f_n \to f$, thus $f \in \mathcal{K}_{\tau,q}(X,\mathcal{B})$. Conversely, suppose $f \in \mathcal{K}_{\tau,q}(X,\mathcal{B})$, i.e., there exists a sequence $\{f_n\}_{n\in\mathbb{N}} \subset X$, such that $f_n \to f$ in X and $\|\{g_k^\star(f_n)\}_{k\in\mathbb{N}}\|_{\tau,q} \leq M + 1/n$. Since $g_k^\star$ is continuous for each $k \in \mathbb{N}$, we have $g_k^\star(f) = \lim_{n\to\infty} g_k^\star(f_n)$. Now, an applications of Fatou's lemma gives that $\|\{g_k^\star(f)\}_{k\in\mathbb{N}}\|_{\tau,q} \leq M < \infty$ and thus $f \in \tilde{\mathcal{K}}_{\tau,q}(X,\mathcal{B})$. $\square$

It can be verified that $\mathcal{K}_{\tau,q}(X,\mathcal{B})$ is (essentially) a Besov space when the greedy basis is a wavelet basis (see e.g. [4]).

For a quasi-normalized unconditional basis $\mathcal{B} = \{g_k\}_{k=1}^\infty$ for X, we define

$$\varphi(n,\mathcal{B}) := \sup_{\Lambda\subset\mathbb{N}:\mathrm{card}(\Lambda)\leq n} \Big\| \sum_{k\in\Lambda} g_k \Big\|_X.$$

The quantity $\varphi(n,\mathcal{B})$ can be used to obtain the following estimate that will be crucial for the main result, Proposition 3.5 below.

Lemma 3.3. *Let $\mathcal{B} = \{g_k\}_{k=1}^\infty$ be an unconditional basis for X. Suppose that for some $p \in (1,\infty)$ there is a constant $C < \infty$ such that*

$$\varphi(n,\mathcal{B}) \leq Cn^{1/p}, \qquad \forall n \in \mathbb{N}.$$

Then

$$\|\cdot\|_X \leq C' \|\{g_k^\star(\cdot)\}\|_{\ell_{p,1}(\mathbb{N})}.$$

Proof. Since $\mathcal{B}$ is an unconditional basis we have for all finite sets $\Lambda \subset \mathbb{N}$,

$$\Big\| \sum_{k\in\Lambda} c_k g_k \Big\| \leq \sup_{k\in\Lambda} |c_k| \Big\| \sum_{\ell\in\Lambda} g_\ell \Big\| \leq C \sup_{k\in\Lambda} |c_k| \,\mathrm{card}(\Lambda)^{1/p}.$$

Let $f \in X$ and denote by $\Lambda_j = \{k \in \mathbb{N} \colon |g_k^\star(f)| \geq 2^{-j}\}$. As $\mathcal{B}$ is unconditional we can write

$$\begin{aligned}
\|f\|_X = \Big\| \sum_{j=-\infty}^{\infty} \sum_{k\in\Lambda_j\setminus\Lambda_{j-1}} g_k^\star(f) g_k \Big\|_X &\leq \sum_{j=-\infty}^{\infty} \Big\| \sum_{k\in\Lambda_j\setminus\Lambda_{j-1}} g_k^\star(f) g_k \Big\|_X \\
&\leq \sum_{j=-\infty}^{\infty} C 2^{-(j-1)} \mathrm{card}(\Lambda_j\setminus\Lambda_{j-1})^{1/p} \leq \tilde{C} \sum_{j=-\infty}^{\infty} 2^{-j} \mathrm{card}(\Lambda_j)^{1/p} \\
&\leq C' \|\{c_k(f)\}\|_{\ell_{p,1}}.
\end{aligned}$$

$\square$

Remark 3.4. Lemma 3.3 also holds true for the so-called quasi-greedy bases, a larger class of bases than the unconditional ones, see [10, Lemma 4.2]. However, shrinkage operators are not well-defined on quasi-greedy bases, in general, so they will not be considered here.

We can now state the main result of this section, giving us convergence rates for shrinkage operators when applied to nice functions.

Proposition 3.5. *Let $\mathcal{B} = \{g_k\}_{k=1}^\infty$ be a quasi-normalized unconditional basis for the Banach space X, and let T_λ^δ be a shrinkage operator associated with δ. Suppose $\varphi(n,\mathcal{B}) \leq Cn^{1/p}$ for some $p > 1$. Then for $\tau < p$ there exists a constant C'*

such that

$$\|f - T_\lambda^\delta f\|_X \le C' \|f\|_{\mathcal{K}_{\tau,\infty}(\mathcal{B})}^{\tau/p} \lambda^{1-\tau/p}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B}), \tag{3.2}$$

with C' independent of f and λ.

Proof. For $\lambda > 0$ we have $T_\lambda^\delta f = \sum_{k=1}^\infty \delta(g_k^\star(f),\lambda) g_k$ so

$$f - T_\lambda^\delta f = \sum_{k=1}^\infty \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)\big) g_k.$$

Define $\Lambda_j = \{k \in \mathbb{N}\colon 2^{-j} < |g_k^\star(f) - \delta(g_k^\star(f),\lambda)| \le 2^{-j+1}\}$, $j \in \mathbb{Z}$. We have the uniform estimate $|g_k^\star(f) - \delta(g_k^\star(f),\lambda)| \le A|g_k^\star(f)|$ so it follows that

$$\operatorname{card}\Lambda_j \le \operatorname{card}\{\ell : |g_\ell^\star(f)| > A^{-1}2^{-j}\} \le A^\tau \|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}}^\tau 2^{j\tau}.$$

We also have the estimate $|g_k^\star(f) - \delta(g_k^\star(f),\lambda)| \le A\lambda$ which implies that $\Lambda_j = \emptyset$ whenever $A\lambda \le 2^{-j}$. Using the estimate from Lemma 3.3, we have

$$\begin{aligned}
\|f - T_\lambda^\delta f\|_X &= \Big\| \sum_{k=1}^\infty \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)\big) g_k \Big\|_X \\
&\le \sum_{j\in\mathbb{Z}} \Big\| \sum_{k\in\Lambda_j} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)\big) g_k \Big\|_X \\
&= \sum_{j > -\log_2(A\lambda)} \Big\| \sum_{k\in\Lambda_j} \big(g_k^\star(f) - \delta(g_k^\star(f),\lambda)\big) g_k \Big\|_X \\
&\le KD \sum_{j > -\log_2(A\lambda)} 2^{-j+1} \varphi\big(A^\tau \|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}}^\tau 2^{j\tau}, \mathcal{B}\big) \\
&\le 2KDCA^{\tau/p} \|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}}^{\tau/p} \sum_{j > -\log_2(A\lambda)} 2^{-j(1-\tau/p)} \\
&\le C' \|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}}^{\tau/p} \lambda^{1-\tau/p},
\end{aligned}$$

where $C' = KDCA(2^{-\tau/p} - 2^{-1})^{-1}$. $\square$

We have the following Corollary.

Corollary 3.6. *Let $\mathcal{B} = \{g_k\}_{k=1}^\infty$ be a quasi-normalized unconditional basis for the Banach space X, and let T_λ^δ be a shrinkage operator associated with δ. Then for $0 < \tau < 1$ there exists a constant C' such that*

$$\|f - T_\lambda^\delta f\|_X \le C' \|f\|_{\mathcal{K}_{\tau,\infty}(\mathcal{B})}^{\tau/p} \lambda^{1-\tau}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B}),$$

with C' independent of f and λ.

Moreover, if X is uniformly convex, there exists $P > 1$ such that for $0 < \tau < P$,

$$\|f - T_\lambda^\delta f\|_X \le C'' \|f\|_{\mathcal{K}_{\tau,\infty}(\mathcal{B})}^{\tau/p} \lambda^{1-\tau/P}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B}),$$

with C'' a constant independent of f and λ.

Proof. By the triangle inequality, we have $\varphi(n,\mathcal{B}) \le Cn$ since the elements of $\mathcal{B}$ are quasi-normalized. If, in addition, X is uniformly convex it is known [12] that there exists $P > 1$ such that $\varphi(n,\mathcal{B}) \le Cn^{1/P}$. □

3.1. **Greedy bases.** The main challenge in order to be able to apply Proposition 3.5 is to have a good estimate for $\varphi(n,\mathcal{B})$. Next we consider a type of basis $\mathcal{B}$, where we have more information on the function $\varphi(n,\mathcal{B})$, the so-called greedy bases. Greedy bases are unconditional bases that also satisfy the democracy property. To be more specific, we have

Definition 3.7. *Let X be a separable Banach space. The family $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ is called* ***democratic*** *if there exists a constant $0 < D < \infty$ such that*

$$\Big\| \sum_{k\in\Gamma} g_k \Big\|_X \le D \Big\| \sum_{k\in\Lambda} g_k \Big\|_X, \tag{3.3}$$

whenever $\Lambda,\Gamma \subset \mathbb{N}$ with $\operatorname{card}(\Lambda) = \operatorname{card}(\Gamma)$. A family of elements $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ in X is called a ***greedy basis*** *if it is an unconditional basis for X that is also democratic.*

Example 3.8. *Given $p \in (1,\infty)$. Two Schauder bases $\{g_k\}$, $\{\tilde{g}_k\}$ for $L_p(\mathbb{R}^d)$ are said to be L_p-equivalent if*

$$\Big\| \sum_k c_k g_k \Big\|_{L_p(\mathbb{R}^d)} \asymp \Big\| \sum_k c_k \tilde{g}_k \Big\|_{L_p(\mathbb{R}^d)},$$

with equivalence independent of the coefficients c_k. In [18] *Temlyakov proved that the L_p-normalized Haar system $\{h_k^p\}$ and any system equivalent to the Haar system satisfies*

$$\Big\| \sum_{k\in\Lambda} h_k^p \Big\|_{L_p} \asymp (\operatorname{card}\Lambda)^{1/p},$$

for any finite index set Λ. Thus, these systems are greedy bases for $L_p(\mathbb{R}^d)$. An example of such a system is an L_p-normalized wavelet system.

For a greedy basis $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ in $L_p(\mathbb{R}^d)$ we have the following complete asymptotic estimate for $\varphi(n,\mathcal{B})$.

Proposition 3.9. *Let $\mathcal{B}$ be a greedy basis in $L_p(\mathbb{R}^d)$ for some $p \in (1,\infty)$. Then $\varphi(n,\mathcal{B}) \asymp n^{1/p}$.*

Proof. We notice that $\mathcal{B} = \{g_k\}_k$ is also a quasi-normalized unconditional basis for $L_p(\mathbb{R}^d)$. It follows from [13] that there is a subsequence $\{g_{k_j}\}$ which is equivalent to the unit vector basis in ℓ_p, i.e. the subsequence $\{g_{k_j}\}$ can be mapped by an isomorphism to the unit vector basis $\{e_k\}_{k\in\mathbb{N}}$ for ℓ_p. Since $\mathcal{B}$ is democratic, it follows that $\varphi(n,\mathcal{B}) \asymp \big\| \sum_{j=1}^n g_{k_j} \big\|_{L_p(\mathbb{R}^d)} \asymp n^{1/p}$. □

We have the following immediate Corollary, that applies to, e.g., L_p-normalized wavelet systems.

Corollary 3.10. *Let $\mathcal{B} = \{g_k\}_{k=1}^{\infty}$ be a greedy basis for $L_p(\mathbb{R}^d)$, with $1 < p < \infty$. Let T_λ^δ be a shrinkage operator associated with δ. Then for $\tau < p$ there exists a constant C' such that*

$$\|f - T_\lambda^\delta f\|_X \leq C' \|f\|_{\mathcal{K}_{\tau,\infty}(\mathcal{B})}^{\tau/p} \lambda^{1-\tau/p}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B}),$$

with C' independent of f and λ.

3.2. **Inverse estimates.** The inequality (3.2) in Proposition 3.5 gives an estimate on how fast the approximant $T_\lambda^\delta f$ converges to $f \in X$ as $\lambda \to 0$. Thus in some sense (3.2) can be seen as a Jackson inequality. Since we do not require that the shrinkage operator sets any coefficients to zero, we cannot hope for a corresponding "standard" Bernstein inequality. However, if we impose some extra restriction on the shrinkage rule, we can show an alternative dual inequality to (3.2).

Definition 3.11. *We call the shrinkage rule $\delta\colon \mathbb{C} \times \mathbb{R}^+ \mapsto \mathbb{C}$ a **strict shrinkage rule** if it is nondecreasing in x and there exists a constant $\varepsilon > 0$ such that*

$$|\delta(x,\lambda)| \leq |x| \min\{1, |x|^\varepsilon \lambda^{-\varepsilon}\}, \qquad \text{for all } x \in \mathbb{C}, \quad \lambda \in \mathbb{R}^+.$$

Lemma 3.12. *Given $1 < \tau < p < \infty$, and $0 < q \leq \infty$. Suppose δ is a strict shrinkage rule with $\varepsilon = p/\tau - 1$. Then*

$$\|\{\delta(g_k^\star(f),\lambda)\}\|_{\ell_{\tau,q}} \leq \lambda^{-p/\tau+1} \|\{g_k^\star(f)\}\|_{\ell_{p,pq/\tau}}^{p/\tau}. \tag{3.4}$$

In particular, if $n^{1/p} \leq C\varphi(n,\mathcal{B})$, we have

$$\|T_\delta^\lambda f\|_{\mathcal{K}_{\tau,\infty}(X,\mathcal{B})} \leq C\lambda^{-p/\tau+1} \|f\|_X^{p/\tau}, \tag{3.5}$$

for all $f \in X$.

Proof. Denote by $\{c_k^*\}_{k=1}^\infty$ the decreasing rearrangement of $\{c_k\}_{k=1}^\infty$. We only prove (3.4) for $q < \infty$, the proof for $q = \infty$ is similar. We have

$$\begin{aligned}
\|\{\delta(g_k^\star(f),\lambda)\}\|_{\ell_{\tau,q}} &= \Big(\sum_{j=0}^{\infty} \big(2^{j/\tau}(\delta(g_{2^j}^\star(f),\lambda))^*\big)^q\Big)^{1/q} \\
&\leq \lambda^{-\varepsilon} \sum_{j=0}^{\infty} \big(2^{j/\tau}((g_{2^j}^\star(f))^*)^{1+\varepsilon}\big)^q\Big)^{1/q} \\
&= \lambda^{-\varepsilon} \sum_{j=0}^{\infty} \big(2^{j/p}(g_{2^j}^\star(f))^*\big)^{pq/\tau}\Big)^{1/q} \\
&= \lambda^{-\varepsilon} \|\{g_k^\star(f)\}\|_{\ell_{p,pq/\tau}}^{p/\tau}.
\end{aligned}$$

According to [10, Lemma 4.1], $n^{1/p} \leq C\varphi(n,\mathcal{B})$ implies that $\|\{g_k^\star(f)\}\|_{\ell_{p,\infty}} \leq C\|f\|_X$. Thus, (3.5) follows if we let $q = \infty$ in (3.4). $\square$

3.3. **Thresholding operators.** We consider a special class of shrinkage operators where more detailed direct and inverse estimates can be proved. The shrinkage rule δ is called a **thresholding rule** if there exists a constant $\alpha > 0$ such that for any x, with $|x| \le \alpha\lambda$, we have $\delta(x,\lambda) = 0$. For a thresholding rule δ, the function $T_\lambda^\delta f$ has a finite expansion in the basis $\mathcal{B} = \{g_k\}_{k=1}^\infty$. This follows from the fact that $g_k^\star(f) \to 0$ as $k \to \infty$, so there exists M_λ such that $k \ge M_\lambda \Rightarrow |g_k^\star(f)| < \alpha\lambda$. We can thus define a function $K\colon X \times \mathbb{R}^+ \mapsto \mathbb{N}$ by

$$K(f,\lambda) := \operatorname{card}\{k \in \mathbb{N}\colon \delta(g_k^\star(f),\lambda) \neq 0\}.$$

For such operators we have the following Jackson estimate.

Proposition 3.13. *Let $\mathcal{B} = \{g_k\}_{k=1}^\infty$ be an unconditional basis for the Banach space X, and let T_λ^δ be a thresholding operator associated with δ. Suppose that $\varphi(n,\mathcal{B}) \le Cn^{1/p}$ for some $p \in (1,\infty)$. Then for $\tau < p$, there exists a constant C' such that*

$$\|f - T_\lambda^\delta f\|_X \le C'\|f\|_{\mathcal{K}_{\tau,\infty}(\mathcal{B})} K(f,\lambda)^{-\alpha}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B}), \quad \alpha = \frac{1}{\tau} - \frac{1}{p},$$

with C' independent of f and λ.

Proof. We notice that

$$\begin{aligned} K(f,\lambda) &= \operatorname{card}\{k \in \mathbb{N}\colon \delta(g_k^\star(f),\lambda) \neq 0\} \\ &\le \operatorname{card}\{k\colon |g_k^\star(f)| > \alpha\lambda\} \le \|\{g_k^\star(f)\}\|_{\ell_{\tau,\infty}}^\tau (\alpha\lambda)^{-\tau}. \end{aligned}$$

Hence,

$$K(f,\lambda)^{1/p-1/\tau} \ge \|\{g_k^\star(f)\}\|_{\ell_{\tau,\infty}}^{\tau/p-1} (\alpha\lambda)^{1-\tau/p}.$$

We now use Proposition 3.5 to obtain the wanted estimate

$$\begin{aligned} \|f - T_\lambda^\delta f\|_X &\le C'\|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}}^{\tau/p} \lambda^{1-\tau/p} \\ &= C'\|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}} \|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}}^{\tau/p-1} \lambda^{1-\tau/p} \\ &\le C''\|\{g_k^\star(f)\}_k\|_{\ell_{\tau,\infty}} K(f,\lambda)^{1/p-1/\tau}. \end{aligned}$$

□

One interesting question is whether the convergence rate provided by Proposition 3.13 is "good" or it can be improved. In the present case, where we have a thresholding operator, we can actually compare the approximant $T_\lambda^\delta f$ to the corresponding best m-term approximation. The good news is that every thresholding operator achieve the rate of the best m-term approximation. To see this, let us introduce some notation. Let $\mathcal{B} = \{g_k\}_{k\in\mathbb{N}}$ be a greedy basis in X. The error associated to the *best m-term* approximation to $f \in X$ from $\mathcal{B}$ is given by

$$\sigma_m(f,\mathcal{B})_X = \inf_{\Lambda \subset \mathbb{N}\colon \operatorname{card}(\Lambda) \le m, \{c_k\}_{k\in\Lambda} \in \mathbb{C}^m} \Big\|f - \sum_{k\in\Lambda} c_k g_k\Big\|_X. \tag{3.6}$$

The class of elements in X that can essentially be approximated at rate $O(m^{-\alpha})$ by m-term expansions from $\mathcal{B}$ is given by the following approximation spaces. Given $\alpha > 0$ and $0 < q \le \infty$ we define

(3.7)

$$\mathcal{A}_q^\alpha(\mathcal{B}) := \Big\{ f \in X, |f|_{\mathcal{A}_q^\alpha(\mathcal{B})} := \|f\|_X + \Big(\sum_{m\in\mathbb{N}} \big(m^\alpha \sigma_m(f,\mathcal{B})_X\big)^q 1/m \Big)^{1/q} < \infty \Big\}.$$

with the usual change to sup when $q = \infty$. Suppose that $\varphi(n,\mathcal{B}) \asymp n^{1/p}$, then it was proved in [10, Cor. 6.1] that

$$\mathcal{A}_q^\alpha(\mathcal{B}) = \mathcal{K}_{\tau,q}(X,\mathcal{B}),$$

for $\alpha = 1/\tau - 1/p$. Thus, in cases where $\varphi(n,\mathcal{B}) \asymp n^{1/p}$ [e.g. in $L_p(\mathbb{R}^d)$, $1 < p < \infty$], and $q = \infty$, Corollary 3.13 shows that thresholding operators actually reach the same rate as best nonlinear m-term approximation. For hard thresholding, i.e., $\delta(x,\lambda) = x\mathbf{1}_{\{|x|>\lambda\}}(x)$, this is no surprise since the definition of a greedy basis is actually equivalent to having a constant $0 < C < \infty$ such that $\|f - T_\lambda^\delta f\|_X \le C\sigma_{K(f,\lambda)}(f,\mathcal{B})_X$, see [14].

3.4. **Inverse estimates for thresholding operators.** For greedy bases, we also have the following Bernstein embedding.

Proposition 3.14. *Let $\mathcal{B} = \{g_k\}_{k\in\mathbb{N}}$ be a greedy basis for X with $\varphi(n,\mathcal{B}) \asymp n^{1/p}$ for some $p \in (1,\infty)$, and let δ be a thresholding rule. Suppose that $f \in X$ satisfies*

$$\|f - T_\lambda^\delta f\|_X \le CK(f,\lambda)^{1/p-1/\tau}, \qquad \text{as } \lambda \to 0^+.$$

Then $f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B})$.

Proof. Since $\mathcal{B}$ is a greedy basis with $\varphi(n,\mathcal{B}) \asymp n^{1/p}$, it suffices to prove that $\sigma_n(f,\mathcal{B}) \le Cn^{1/p-1/\tau}$ for $n \in \mathbb{N}$, see [10, Cor. 6.1]. For $E \subset \mathbb{N}$, we define the projection $P_E(\sum_k c_k g_k) = \sum_{k\in E} c_k g_k$. We have,

$$\begin{aligned}
\sigma_{K(f,\lambda)}(f,\mathcal{B})_X &\le \|f - P_{\mathrm{supp}\{\delta(g_k^\star(f),\lambda)\}} f\|_X \\
&= \|(f - T_\lambda^\delta f) + (T_\lambda^\delta f - P_{\mathrm{supp}\{\delta(g_k^\star(f),\lambda)\}} f)\|_X \\
&\le CK(f,\lambda)^{1/p-1/\tau} + C'\|(f - T_\lambda^\delta f)\|_X \\
&\le C''K(f,\lambda)^{1/p-1/\tau},
\end{aligned}$$

where we have used that $\mathcal{B}$ is an unconditional basis to obtain the estimate

$$\|T_\lambda^\delta f - P_{\mathrm{supp}\{\delta(g_k^\star(f),\lambda)\}} f\|_X \le C'\|(f - T_\lambda^\delta f)\|_X.$$

This completes the proof since $\sigma_n(f,\mathcal{B})_X$ is decreasing in n. □

4. Shrinkage operators for redundant systems

Proposition 3.5 can be extended to certain redundant systems provided they have a very special structure. However, it is clear that for an arbitrary dictionary $\mathcal{D}$, we cannot even define a shrinkage operator due to the fact that a given element $f \in X$ can have many possible expansions in the dictionary. So which expansion should we "shrink"? Therefore, we will focus on redundant dictionaries with "canonical" expansions. We need the following definition.

Definition 4.1. *Let X be a Banach space and X_d a Banach sequence space indexed by $\mathbb{N}$. Let $\{g_k\} \subset X$, $\{g_k^\star\} \subset X^*$. Then $(\{g_k\},\{g_k^\star\})$ is an atomic decomposition of X with respect to X_d if*

(1) $\{g_k^\star(f)\} \in X_d$ *for all* $f \in X$.
(2) *there exist constants* $A, B > 0$ *such that*

$$A\|f\|_X \leq \|\{g_k^\star(f)\}\|_{X_d} \leq B\|f\|_X, \quad \forall f \in X.$$

(3) *The reconstruction operator* $R\colon X_d \mapsto X$, *defined by* $R(\{\alpha_k\}) = \sum_{k\in\mathbb{N}} \alpha_k g_k$, *is bounded with* $R(\{g_k^\star(f)\}) = f$ *for all* $f \in X$.

Remark 4.2. If $\ell_{p,1} \hookrightarrow X_d$ for some $p \in (1,\infty)$, then it can be shown that for $1 \leq q \leq \infty$ and $0 < \tau < p$,

$$\mathcal{K}_{\tau,q}(X,\mathcal{D}) = \left\{ f \in X \,\middle|\, \{g_k^\star(f)\}_k \in \ell_{\tau,q} \right\},$$

with $|f|_{\mathcal{K}_{\tau,q}(X,\mathcal{D})} = \|\{g_k^\star(f)\}_k\|_{\ell_{\tau,q}}$ [9, Proposition 3].

Given an atomic decomposition $(\{g_k\},\{g_k^\star\})$ and a shrinkage rule δ we define

$$T_\lambda^\delta f = \sum_{k=1}^{\infty} \delta(g_k^\star(f),\lambda) g_k.$$

We have the following result.

Proposition 4.3. *Let $(\{g_k\},\{g_k^\star\})$ be an atomic decomposition of X with respect to X_d. Assume that $\ell_{p,1} \hookrightarrow X_d$ for some $p \in (1,\infty)$. Let T_λ^δ be a shrinkage operator associated with δ. Then for $\tau < p$ there exists a constant C such that*

$$\|f - T_\lambda^\delta f\|_X \leq C \|f\|_{\mathcal{K}_{\tau,\infty}(\mathcal{B})}^{\tau/p} \lambda^{1-\tau/p}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(X,\mathcal{B}),$$

with C independent of f and λ.

Proof. We notice that there exists a constant $0 < C' < \infty$ such that for any finite set $\Lambda \subset \mathbb{N}$,

$$\Big\| \sum_{k\in\Lambda} \delta(g_k^\star(f),\lambda) g_k \Big\|_X \leq C' \big(\mathrm{card}(\Lambda)\big)^{1/p} \max_{k\in\Lambda} |\delta(g_k^\star(f),\lambda)|,$$

which follows from the assumption that $\ell_{p,1} \hookrightarrow X_d$ and condition (2) in Definition 4.1. Now, the proof is completed using arguments similar to those in the proof of Proposition 3.5. □

One may wonder whether there are any *interesting* redundant dictionaries that satisfy all the requirements of Proposition 4.3. Below we consider two examples. Tight wavelet frames in $L_p(\mathbb{R}^d)$ and Gabor dictionaries in the modulation spaces $M_w^{p,p}(\mathbb{R}^d)$.

4.1. **Tight wavelet frames.** We consider tight wavelet frames in $L_2(\mathbb{R}^d)$. Let us introduce some notation. Let D denote the set of dyadic cubes $I = 2^{-j}([0,1]^d + k)$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^d$, and for a function $\psi \in L_2(\mathbb{R}^d)$, we let $\psi_I := 2^{jd/2}\psi(2^j \cdot - k)$.

Definition 4.4. *A tight wavelet frame for $L_2(\mathbb{R}^d)$ is a finite collection of functions $\Psi = \{\psi^\ell\}_{\ell \in E}$ in $L_2(\mathbb{R}^d)$, $E = \{1, 2, \ldots, L\}$, for which the system $X(\Psi) = \{\psi_I^\ell \colon I \in D, \ell \in E\}$ is a tight frame for $L_2(\mathbb{R}^d)$.*

The most common methods to construct tight wavelet frames are the extension principles of Ron and Shen. Tight wavelet frames build through this principle are based on a multiresolution analysis. For more information on extension principles see e.g. [3, 16, 15]. We need the following result proved in [2]:

Theorem 4.5. *Let $\Psi = \{\psi^\ell\}_{\ell \in E}$ be a tight wavelet frame with frame bound A. Suppose $\int \psi^\ell(x)\,dx = 0$, $\psi^\ell \in C^\beta(\mathbb{R}^d)$ for some $\beta > 0$ and $\ell \in E$, and assume that $|\psi^\ell(x)| \le C(1+|x|)^{-d-\varepsilon}$ for some $\varepsilon > 0$ and $\ell \in E$. Then the system $(\{A^{-1}\psi_I^\ell\}_{I,\ell}, \{\psi_I^\ell\}_{I,\ell})$ is an atomic decomposition of $L_p(\mathbb{R}^d)$, $1 < p < \infty$, and*

$$\|f\|_{L_p} \asymp \left\| \left(\sum_{I \in D, \ell \in E} |\langle f, \psi_I^\ell \rangle|^2 |I|^{-1} \chi_I(x) \right)^{1/2} \right\|_{L_p}. \tag{4.1}$$

Furthermore, there exists a constant $0 < C < \infty$ such that for any finite set $\Lambda \subset D \times E$

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

Here, $\psi_I^{\ell,p}$ denotes the function ψ_I^ℓ normalized in $L_p(\mathbb{R})$.

Hence, from Proposition 4.3, we have the following corollary.

Corollary 4.6. *Let $\Psi = \{\psi^\ell\}_{\ell \in E}$ be a tight wavelet frame that satisfies the hypothesis of Theorem 4.5. Define the associated shrinkage operator $T_\lambda^\delta \colon L_p(\mathbb{R}^d) \mapsto L_p(\mathbb{R}^d)$ by*

$$T_\lambda^\delta f = \sum_{I \in D, \ell \in E} \delta\big(\langle f, \psi_I^{\ell,p'} \rangle, \lambda\big) \psi_I^{\ell,p}, \qquad 1 = 1/p + 1/p'.$$

Then for $\tau < p$ there exists a constant C such that

$$\|f - T_\lambda^\delta f\|_{L_p} \le C \|f\|_{\mathcal{K}_{\tau,\infty}(L_p(\mathbb{R}^d), X(\Psi))}^{\tau/p} \lambda^{1-\tau/p}, \qquad \text{for all } f \in \mathcal{K}_{\tau,\infty}(L_p(\mathbb{R}^d), X(\Psi)).$$

Remark 4.7. It was proved in [2] that whenever Ψ consists of “nice” functions (i.e., the generators are sufficiently smooth with vanishing moments), the sparsity space $\mathcal{K}_{\tau,\tau}(L_p(\mathbb{R}^d), X(\Psi))$, for $1 < \tau < p < \infty$, can actually be identified with the Besov space $B_\tau^{\alpha d}(L_\tau(\mathbb{R}^d))$, where $\alpha = 1/\tau - 1/p$.

4.2. **Gabor Banach frames.** Let us consider another example of a redundant dictionary where shrinkage works. We consider the weighted modulation space $M_w^{p,p}(\mathbb{R}^d)$ with α-moderate weight w, for $1 \leq p < \infty$. We refer the reader to [11, Chapters 11-12] for the definition and basic properties of such spaces. For a given window function g, the Gabor dictionary is given by

$$\{g_{n,m} := g(\cdot - na)e^{2i\pi mb\cdot}\}_{m,n\in\mathbb{Z}^d},$$

with (a,b) fixed positive numbers. One can prove that for small parameters a and b, and sufficiently "nice" g,

- The synthesis operator $T\mathbf{c} = \sum_{m,n} c_{n,m}g_{n,m}$ is bounded from ℓ_p to $M_w^{p,p}(\mathbb{R}^d)$ [11, Theorem 12.2.4].
- There exists a dual system $\{\tilde{g}_{n,m}\}_{m,n}$ such that

$$f = \sum_{n,m} \langle f, \tilde{g}_{m,n}\rangle \frac{g_{m,n}}{\|g_{m,n}\|_{M_w^{p,p}(\mathbb{R}^d)}}, \qquad f \in M_w^{p,p}(\mathbb{R}^d).$$

A Gabor system $(\{g_{m,n}\}, \{\tilde{g}_{m,n}\})$ with these properties is called a Gabor Banach frame. It was noticed in [9] that for $\tau \leq p$ (and for the particular dual system $\{\tilde{g}_{m,n}\}$),

$$M_w^{\tau,\tau}(\mathbb{R}^d) = \mathcal{K}_{\tau,\tau}(M_w^{p,p}(\mathbb{R}^d), \mathcal{D}).$$

With these facts about Gabor Banach frames, we have the following Corollary, which concludes this note,

Corollary 4.8. *Let $\mathcal{D} = \{g_{m,n}\}_{m,n\in\mathbb{Z}^d}$ be a Gabor Banach frame for $M_w^{p,p}(\mathbb{R}^d)$ for some $1 \leq p < \infty$. Define the associated shrinkage operator $T_\lambda^\delta \colon M_w^{p,p}(\mathbb{R}^d) \mapsto M_w^{p,p}(\mathbb{R}^d)$ by*

$$T_\lambda^\delta f = \sum_{m,n\in\mathbb{Z}^d} \delta\big(\langle f, \tilde{g}_{m,n}\rangle, \lambda\big) g_{m,n},$$

with $\{\tilde{g}_{m,n}\}$ the dual system mentioned above. Then for $\tau < p$ there exists a constant C such that

$$\|f - T_\lambda^\delta f\|_{M_w^{p,p}(\mathbb{R}^d)} \leq C\|f\|_{M_w^{\tau,\tau}(\mathbb{R}^d)}\lambda^{1-\tau/p}, \qquad \textit{for all } f \in M_w^{\tau,\tau}(\mathbb{R}^d).$$

REFERENCES

[1] F. Abramovich, T. Sapatinas, and B. W. Silverman. Wavelet thresholding via a Bayesian approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 60(4):725–749, 1998.

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

[3] Ingrid Daubechies, Bin Han, Amos Ron, and Zuowei Shen. Framelets: MRA-based constructions of wavelet frames. *Preprint*, 2001.

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

[5] David L. Donoho and Iain M. Johnstone. Minimax estimation via wavelet shrinkage. *Ann. Statist.*, 26(3):879–921, 1998.

[6] David L. Donoho, Iain M. Johnstone, Gérard Kerkyacharian, and Dominique Picard. Wavelet shrinkage: asymptopia? *J. Roy. Statist. Soc. Ser. B*, 57(2):301–369, 1995. With discussion and a reply by the authors.

[7] Hong-Ye Gao. Wavelet shrinkage denoising using the non-negative garrote. *J. Comput. Graph. Statist.*, 7(4):469–488, 1998.

[8] Hong-Ye Gao and Andrew G. Bruce. WaveShrink with firm shrinkage. *Statist. Sinica*, 7(4):855–874, 1997.

[9] R. Gribonval and M Nielsen. Nonlinear approximation with dictionaries. I. direct estimates. *J. Fourier Anal. Appl.* (to appear).

[10] R. Gribonval and M. Nielsen. Some remarks on non-linear approximation with Schauder bases. *East J. Approx.*, 7(3):267–285, 2001.

[11] Karlheinz Gröchenig. *Foundations of time-frequency analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston Inc., Boston, MA, 2001.

[12] V. I. Gurariĭ and N. I. Gurariĭ. Bases in uniformly convex and uniformly smooth Banach spaces. *Izv. Akad. Nauk SSSR Ser. Mat.*, 35:210–215, 1971.

[13] M. I. Kadec and A. Pełczyński. Bases, lacunary sequences and complemented subspaces in the spaces L_p. *Studia Math.*, 21:161–176, 1961/1962.

[14] S. V. Konyagin and V. N. Temlyakov. A remark on greedy approximation in Banach spaces. *East J. Approx.*, 5(3):365–379, 1999.

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

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

[17] Terence Tao and Brani Vidakovic. Almost everywhere behavior of general wavelet shrinkage operators. *Appl. Comput. Harmon. Anal.*, 9(1):72–82, 2000.

[18] V. N. Temlyakov. The best m-term approximation and greedy algorithms. *Adv. Comput. Math.*, 8(3):249–265, 1998.

DEPARTMENT OF MATHEMATICAL SCIENCES, AALBORG UNIVERSITY,
FREDRIK BAJERS VEJ 7G, 9220 AALBORG EAST, DENMARK

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

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