Industrial Mathematics Institute

2006:07

Greedy algorithms with prescribed coefficients

V. Temlyakov

IMI
Preprint Series

Department of Mathematics
University of South Carolina

GREEDY ALGORITHMS WITH PRESCRIBED COEFFICIENTS[1]

V. Temlyakov

University of South Carolina

Abstract. We study greedy algorithms in a Banach space from the point of view of convergence and rate of convergence. We concentrate on studying algorithms that provide expansions into a series. We call such expansions *greedy expansions*. It was pointed out in our previous paper that there is a great flexibility in choosing coefficients of greedy expansions. In that paper this flexibility was used for constructing a greedy expansion that converges in any uniformly smooth Banach space. In this paper we push the flexibility in choosing the coefficients of greedy expansions to the extreme. We make these coefficients independent of an element $f \in X$. Surprisingly, for a properly chosen sequence of coefficients we obtain results similar to the previous results on greedy expansions when the coefficients were determined by an element f.

1. Introduction

We continue the investigation of greedy approximation in Banach spaces. In this article we concentrate on studying convergence and rate of convergence of greedy expansions with coefficients assigned in advance. A new phenomenon of convergence of a greedy expansion of $f \in X$ with coefficients chosen a priori independently of f has been discovered in this paper. We remind notations that are standard in the theory of greedy approximations.

Let X be a Banach space with norm $\|\cdot\|$. We say that a set of elements (functions) $\mathcal{D}$ from X is a dictionary (symmetric dictionary) if each $g \in \mathcal{D}$ has norm bounded by one ($\|g\| \le 1$),

$$g \in \mathcal{D} \quad \text{implies} \quad -g \in \mathcal{D},$$

and $\overline{\operatorname{span}}\mathcal{D} = X$. We denote the closure (in X) of the convex hull of $\mathcal{D}$ by $A_1(\mathcal{D})$. We introduce a new norm, associated with a dictionary $\mathcal{D}$, in the dual space X^* by the formula

$$\|F\|_{\mathcal{D}} := \sup_{g \in \mathcal{D}} F(g), \quad F \in X^*.$$

We will study in this paper greedy algorithms with regard to $\mathcal{D}$. For a nonzero element $f \in X$ we denote by F_f a norming (peak) functional for f:

$$\|F_f\| = 1, \qquad F_f(f) = \|f\|.$$

[1] *Math Subject Classifications.* primary: 41A65; secondary: 41A25, 41A46, 46B20.
Keywords and Phrases. greedy algorithms, Banach spaces, greedy expansion, convergence.

The existence of such a functional is guaranteed by Hahn-Banach theorem.

We begin with the Pure Greedy Algorithm (PGA), introduced in [FS], and its generalization the Weak Greedy Algorithm (WGA) defined in a Hilbert space. We give a definition of the WGA and note that the PGA is the WGA with the weakness sequence $\tau = \{1\}$. Let a sequence $\tau = \{t_k\}_{k=1}^{\infty}$, $0 \le t_k \le 1$, be given. Following [T2] we define the Weak Greedy Algorithm.

Weak Greedy Algorithm (WGA(τ)). *We define $f_0 := f$. Then for each $m \ge 1$, we inductively define:*

1). $\varphi_m \in \mathcal{D}$ is any satisfying

$$\langle f_{m-1}, \varphi_m \rangle \ge t_m \sup_{g \in \mathcal{D}} \langle f_{m-1}, g \rangle;$$

2).

$$f_m := f_{m-1} - \langle f_{m-1}, \varphi_m \rangle \varphi_m;$$

3).

$$G_m(f, \mathcal{D}) := \sum_{j=1}^{m} \langle f_{j-1}, \varphi_j \rangle \varphi_j.$$

The greedy step (the first step) of the PGA can be interpreted in two ways. First, we look at the mth step for an element $\varphi_m \in \mathcal{D}$ and a number λ_m satisfying

$$\|f_{m-1} - \lambda_m \varphi_m\|_H = \inf_{g \in \mathcal{D}, \lambda} \|f_{m-1} - \lambda g\|_H. \tag{1.1}$$

Second, we look for an element $\varphi_m \in \mathcal{D}$ such that

$$\langle f_{m-1}, \varphi_m \rangle = \sup_{g \in \mathcal{D}} \langle f_{m-1}, g \rangle. \tag{1.2}$$

In a Hilbert space both versions (1.1) and (1.2) result in the same PGA. In a general Banach space the corresponding versions of (1.1) and (1.2) lead to different greedy algorithms. The Banach space version of (1.1) is straightforward: instead of the Hilbert norm $\|\cdot\|_H$ in (1.1) we use the Banach norm $\|\cdot\|_X$. This results in the following greedy algorithm.

X-Greedy Algorithm (XGA). *We define $f_0 := f$, $G_0 := 0$. Then, for each $m \ge 1$, we inductively define*

1). $\varphi_m \in \mathcal{D}$, $\lambda_m \in \mathbb{R}$ are such that (we assume existence)

$$\|f_{m-1} - \lambda_m \varphi_m\|_X = \inf_{g \in \mathcal{D}, \lambda} \|f_{m-1} - \lambda g\|_X. \tag{1.3}$$

2). Denote

$$f_m := f_{m-1} - \lambda_m \varphi_m, \qquad G_m := G_{m-1} + \lambda_m \varphi_m.$$

The second version of the PGA in a Banach space is based on the concept of a norming (peak) functional. We note that in a Hilbert space a norming functional F_f acts as follows

$$F_f(g) = \langle f/\|f\|, g\rangle.$$

Therefore, (1.2) can be rewritten in terms of the norming functional $F_{f_{m-1}}$ as

$$F_{f_{m-1}}(\varphi_m) = \sup_{g\in\mathcal{D}} F_{f_{m-1}}(g). \tag{1.4}$$

This observation leads to the class of dual greedy algorithms. We define the Weak Dual Greedy Algorithm with weakness τ (WDGA(τ)) that is a generalization of the Weak Greedy Algorithm.

Weak Dual Greedy Algorithm (WDGA(τ)). *Let $\tau := \{t_m\}_{m=1}^\infty$, $t_m \in [0,1]$, be a weakness sequence. We define $f_0 := f$. Then, for each $m \geq 1$, we inductively define*

1). $\varphi_m \in \mathcal{D}$ is any satisfying

$$F_{f_{m-1}}(\varphi_m) \geq t_m \|F_{f_{m-1}}\|_{\mathcal{D}}. \tag{1.5}$$

2). Define a_m as

$$\|f_{m-1} - a_m\varphi_m\| = \min_{a\in\mathbb{R}} \|f_{m-1} - a\varphi_m\|.$$

3). Denote

$$f_m := f_{m-1} - a_m\varphi_m.$$

Let us make a remark that justifies the idea of the dual greedy algorithms in terms of real analysis. We consider here approximation in uniformly smooth Banach spaces. For a Banach space X we define the modulus of smoothness

$$\rho(u) := \sup_{\|x\|=\|y\|=1} \left(\frac{1}{2}(\|x+uy\| + \|x-uy\|) - 1\right).$$

The uniformly smooth Banach space is the one with the property

$$\lim_{u\to 0} \rho(u)/u = 0.$$

It is easy to see that for any Banach space X its modulus of smoothness $\rho(u)$ is an even convex function satisfying the inequalities

$$\max(0, u-1) \leq \rho(u) \leq u, \quad u \in (0,\infty).$$

The following well known proposition is a simple corollary of the definition of a uniformly smooth Banach space.

Proposition 1.1. *Let X be a uniformly smooth Banach space. Then for any $x \neq 0$ and y we have*

$$F_x(y) = \left(\frac{d}{du}\|x+uy\|\right)(0) = \lim_{u\to 0}(\|x+uy\| - \|x\|)/u. \tag{1.6}$$

Proposition 1.1 shows that in the WDGA we are looking for an element $\varphi_m \in \mathcal{D}$ that provides a big derivative of the quantity $\|f_{m-1}+ug\|$. Thus, we have two classes of greedy algorithms in Banach spaces. The first one is based on a greedy step of the form (1.3). We call this class the class of X-*greedy algorithms*. The second one is based on a greedy step of the form (1.5). We call this class the class of *dual greedy algorithms*. A very important feature of the dual greedy algorithms is that they can be modified into a weak form. The term "weak" in the definition of the WDGA means that at the greedy step (1.5) we do not shoot for the optimal element of the dictionary which realizes the corresponding sup but are satisfied with weaker property than being optimal. The obvious reason for this is that we do not know in general that the optimal one exists. Another, practical reason is that the weaker the assumption the easier to satisfy it and, therefore, easier to realize in practice.

From the definition of a dictionary it follows that any element $f \in X$ can be approximated arbitrarily well by finite linear combinations of the dictionary elements. The primary goal of this paper is to study representations of an element $f \in X$ by a series

$$f \sim \sum_{j=1}^{\infty} c_j(f) g_j(f), \quad g_j(f) \in \mathcal{D}, \quad c_j(f) > 0, \quad j = 1, 2, \dots. \tag{1.7}$$

In building a representation (1.7) we should construct two sequences: $\{g_j(f)\}_{j=1}^{\infty}$ and $\{c_j(f)\}_{j=1}^{\infty}$. In this paper the construction of $\{g_j(f)\}_{j=1}^{\infty}$ will be based on ideas used in greedy-type nonlinear approximation (greedy-type algorithms). This justifies the use of the term *greedy expansion* for (1.7) considered in the paper. We will consider two different classes of greedy expansions that correspond to the above two classes of greedy algorithms. The first class is the class of X-*greedy expansions* obtained by X-greedy algorithms and the second class is the class of *dual greedy expansions* obtained by dual greedy algorithms. The X-Greedy Algorithm provides a X-greedy expansion and the WDGA provides a dual greedy expansion.

We begin our discussion with known results on the X-greedy expansions. There is no results on X-Greedy Algorithms in general uniformly smooth Banach spaces. The reader can find some open problems on the X-Greedy Algorithms in a survey [T5, p.73, p.94]. We discuss here the results on the PGA. The first steps in the theory of greedy approximation were devoted to the study of convergence of the expansion (1.7) and were done in a Hilbert space. P.J. Huber [H] proved convergence of the PGA in the weak topology and conjectured that the PGA converges in the strong sense (in the norm of H). L. Jones [J1] proved this conjecture. It is a fundamental result in the theory of greedy approximation that guarantees convergence of (1.7), obtained by the PGA, for any $f \in H$ and any dictionary $\mathcal{D}$. Convergence of the WGA(τ) is also well studied. A criterion on the weakness sequence

τ for convergence of the WGA(τ) for each dictionary $\mathcal{D}$ and any element $f \in H$ has been found in [T4].

We proceed to the results on rate of approximation by the PGA. It was proved in [DT] that for a general dictionary $\mathcal{D}$ the Pure Greedy Algorithm provides the estimate

$$\|f_m\| \le m^{-1/6}, \quad f \in A_1(\mathcal{D}).$$

The above estimate was improved a little in [KT] to

$$\|f_m\| \le 4m^{-11/62}, \quad f \in A_1(\mathcal{D}).$$

The following problem (see [T5, p.65, Open Problem 3.1]) is a central theoretical problem in greedy approximation in Hilbert spaces.

Open problem. Find the order of decay of the sequence

$$\gamma(m) := \sup_{f \in A_1(\mathcal{D}),\mathcal{D}} \|f_m\|.$$

Recently, the known upper bounds in approximation by the Pure Greedy Algorithm have been improved in [S]. Sil'nichenko proved the estimate

$$\gamma(m) \le Cm^{-\frac{s}{2(2+s)}}$$

where s is a solution from $[1, 1.5]$ of the equation

$$(1+x)^{\frac{1}{2+x}} \left(\frac{2+x}{1+x}\right) - \frac{1+x}{x} = 0.$$

Numerical calculations of s (see [S]) give

$$\frac{s}{2(2+s)} = 0.182\dots > 11/62.$$

The technique used in [S] is a further development of a method from [KT].

There is also some progress in the lower estimates. The estimate

$$\gamma(m) \ge Cm^{-0.27},$$

with a positive constant C, has been proved in [LiT]. For previous lower estimates see [T5, p.59].

The results on the PGA discussed above as results on the X-greedy expansions can also be considered as results on the dual greedy expansions. We proceed to further discussion of results on the dual greedy expansions. For the WGA(t) with $\tau = \{t\}$, $t \in (0, 1]$ the following error estimate has been proved in [T2]

$$\|f_m\| \le (1 + mt^2)^{-t/(4+2t)}, \quad f \in A_1(\mathcal{D}).$$

This estimate implies the inequality

$$\|f_m\| \le Cm^{-t/6}, \quad f \in A_1(\mathcal{D}),$$

with the exponent $t/6$ approaching 0 linearly in t. It was proved in [LiT] that the corresponding exponent cannot decrease to 0 at a slower rate than linearly. These results provide understanding of the dependence of rate of approximation on the weakness parameter t.

The WDGA provides a dual greedy expansion in a Banach space. We do not have a general convergence result for the WDGA. The reader can find some open problems on the behavior of the WDGA in [T5, p.73]. There is a convergence result due to Ganichev and Kalton [GK] for uniformly smooth Banach spaces satisfying an extra condition *property* Γ (see Section 2). In Section 2 we give other proof of this convergence result.

The construction of $\{g_j(f)\}_{j=1}^{\infty}$ is, clearly, the most important and difficult part in building a representation (1.7). It was pointed out in [T6] that we have a great flexibility in choosing the coefficients $c_j(f)$ of the expansion (1.7). In [T6] this flexibility was used for constructing a dual greedy expansion that converges in any uniformly smooth Banach space. We discuss these results in more detail. The construction from [T6] depends on two numerical parameters $t \in (0,1]$ (the weakness parameter) and $b \in (0,1)$ (the tuning parameter of the approximation method). The construction also depends on a majorant μ of the modulus of smoothness of the Banach space X.

Dual Greedy Algorithm with parameters (t,b,μ) **(DGA**(t,b,μ)**).** *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u)$ and let $\mu(u)$ be a continuous majorant of $\rho(u)$: $\rho(u) \le \mu(u)$, $u \in [0,\infty)$. For parameters $t \in (0,1]$, $b \in (0,1]$ we define sequences $\{f_m\}_{m=0}^{\infty}$, $\{\varphi_m\}_{m=1}^{\infty}$, $\{c_m\}_{m=1}^{\infty}$ inductively. Let $f_0 := f$. If for $m \ge 1$ $f_{m-1} = 0$ then we set $f_j = 0$ for $j \ge m$ and stop. If $f_{m-1} \ne 0$ then we conduct the following three steps:*

1). take any $\varphi_m \in \mathcal{D}$ such that

$$F_{f_{m-1}}(\varphi_m) \ge t\|F_{f_{m-1}}\|_{\mathcal{D}};$$

2). choose $c_m > 0$ from the equation

$$\|f_{m-1}\|\mu(c_m/\|f_{m-1}\|) = \frac{tb}{2}c_m\|F_{f_{m-1}}\|_{\mathcal{D}};$$

3). define

$$f_m := f_{m-1} - c_m\varphi_m.$$

In [T6] we proved the following convergence result.

Theorem 1.1. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u)$ and let $\mu(u)$ be a continuous majorant of $\rho(u)$ with the property $\mu(u)/u \downarrow 0$ as $u \to +0$. Then for any $t \in (0,1]$ and $b \in (0,1)$ the DGA(t,b,μ) converges for each dictionary $\mathcal{D}$ and all $f \in X$.*

We proceed to results on Banach spaces with power type modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. It is well known (see [LT]) that power type modulus of smoothness of nontrivial Banach spaces are limited to the case $q \in [1,2]$. The following result from [T6] gives the rate of convergence.

Theorem 1.2. *Assume X has modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. Denote $\mu(u) := \gamma u^q$. Then for any dictionary $\mathcal{D}$ and any $f \in A_1(\mathcal{D})$ the rate of convergence of the DGA(t,b,μ) is given by*

$$\|f_m\| \le C(t,b,\gamma,q) m^{-\frac{t(1-b)}{p(1+t(1-b))}}, \quad p := \frac{q}{q-1}.$$

In this paper we push the flexibility in choosing the coefficients $c_j(f)$ in the expansion (1.7) to the extreme. We make these coefficients independent of an element $f \in X$. Surprisingly, for properly chosen coefficients we obtain results for the corresponding dual greedy expansion similar to the above Theorems 1.1 and 1.2. Even more surprisingly, we obtain similar results for the corresponding X-greedy expansions. We proceed to the formulation of these results. Let $\mathcal{C} := \{c_m\}_{m=1}^\infty$ be a fixed sequence of positive numbers. We restrict ourselves to positive numbers because of the symmetry of the dictionary $\mathcal{D}$.

X-Greedy Algorithm with coefficients $\mathcal{C}$ (XGA($\mathcal{C}$)). *We define $f_0 := f$, $G_0 := 0$. Then, for each $m \ge 1$, we inductively define*

1). $\varphi_m \in \mathcal{D}$ is such that (we assume existence)

$$\|f_{m-1} - c_m\varphi_m\|_X = \inf_{g\in\mathcal{D}} \|f_{m-1} - c_m g\|_X.$$

2). Denote

$$f_m := f_{m-1} - c_m\varphi_m, \qquad G_m := G_{m-1} + c_m\varphi_m.$$

Dual Greedy Algorithm with weakness τ and coefficients $\mathcal{C}$ (DGA(τ,$\mathcal{C}$)). *Let $\tau := \{t_m\}_{m=1}^\infty$, $t_m \in [0,1]$, be a weakness sequence. We define $f_0 := f$, $G_0 := 0$. Then, for each $m \ge 1$, we inductively define*

1). $\varphi_m \in \mathcal{D}$ is any satisfying

$$F_{f_{m-1}}(\varphi_m) \ge t_m \|F_{f_{m-1}}\|_{\mathcal{D}}.$$

2). Define

$$f_m := f_{m-1} - c_m\varphi_m, \qquad G_m := G_{m-1} + c_m\varphi_m.$$

In this paper we consider the case $\tau = \{t\}$, $t \in (0,1]$. We will write t instead of τ in the notation. The first result on convergence properties of the DGA$(t,\mathcal{C})$ has been obtained in [T6]. We formulate it here.

Theorem 1.3. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u)$. Assume $\mathcal{C} = \{c_j\}_{j=1}^\infty$ is such that*

$$\sum_{j=1}^\infty c_j = \infty, \quad \text{and, for any} \quad y > 0, \quad \sum_{j=1}^\infty \rho(yc_j) < \infty.$$

Then for the DGA(t,$\mathcal{C}$) we have

$$\liminf_{m\to\infty} \|f_m\| = 0.$$

In Section 2 we prove an analogue of Theorem 1.3 for the XGA($\mathcal{C}$). In Section 3 we improve upon the convergence in Theorem 1.3 in the case of uniformly smooth Banach spaces with power type modulus of smoothness. Under an extra assumption on $\mathcal{C}$ we replace $\liminf$ by $\lim$.

Theorem 1.4. *Let $\mathcal{C} \in \ell_q \setminus \ell_1$ be a monotone sequence. Then the DGA(t,$\mathcal{C}$) and the XGA($\mathcal{C}$) converge for each dictionary and all $f \in X$ in any uniformly smooth Banach space X with modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$.*

In Section 4 we address the question of rate of approximation for $f \in A_1(\mathcal{D})$. We prove the following theorem.

Theorem 1.5. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. We set $s := (1+1/q)/2$ and $\mathcal{C}_s := \{k^{-s}\}_{k=1}^\infty$. Then the DGA(t,$\mathcal{C}_s$) and XGA($\mathcal{C}_s$) (for this algorithm $t=1$) converge for $f \in A_1(\mathcal{D})$ with the following rate: for any $r \in (0, t(1-s))$*

$$\|f_m\| \le C(r,t,q,\gamma) m^{-r}.$$

In the case $t = 1$ Theorem 1.5 provides rate of convergence m^{-r} for $f \in A_1(\mathcal{D})$ with r arbitrarily close to $(1-1/q)/2$. Theorem 1.2 provides similar rate of convergence. It would be interesting to understand if the rate $m^{-(1-1/q)/2}$ is the best that can be achieved in greedy expansions (for each $\mathcal{D}$, any $f \in A_1(\mathcal{D})$, and any X with $\rho(u) \le \gamma u^q$, $q \in (1,2]$). We note that there are greedy approximation methods that provide error bound of the order $m^{1/q-1}$ for $f \in A_1(\mathcal{D})$ (see survey [T5] and [T8] for recent results). However, these approximation methods do not provide an expansion. We discuss one of such approximation methods in Section 4.

2. Convergence

We begin with a simple well known lemma.

Lemma 2.1. *For any elements $\psi \ne 0$, φ, and a number c one has*

$$\|\psi - c\varphi\| \le \|\psi\|(1 + 2\rho(c\|\varphi\|/\|\psi\|)) - cF_\psi(\varphi). \tag{2.1}$$

Proof. From the definition of modulus of smoothness we have

$$\|\psi - c\varphi\| + \|\psi + c\varphi\| \le 2\|\psi\|(1 + \rho(c\|\varphi\|/\|\psi\|)). \tag{2.2}$$

Next,

$$\|\psi + c\varphi\| \ge F_\psi(\psi + c\varphi) = \|\psi\| + cF_\psi(\varphi). \tag{2.3}$$

Combining (2.2) and (2.3) we obtain (2.1). □

Lemma 2.2. *Let f, f^ϵ, $\epsilon \geq 0$, $A(\epsilon) > 0$, be such that*

$$\|f - f^\epsilon\| \leq \epsilon, \quad f^\epsilon / A(\epsilon) \in A_1(\mathcal{D}).$$

Then for

$$f_k := f - \sum_{j=1}^{k} c_j \varphi_j, \quad \varphi_j \in \mathcal{D}, \quad j = 1, \dots, k,$$

we have

$$\|F_{f_k}\|_{\mathcal{D}} \geq (\|f_k\| - \epsilon)/(A(\epsilon) + A_k), \quad A_k := \sum_{j=1}^{k} |c_j|.$$

Proof. Denote

$$G_k := \sum_{j=1}^{k} c_j \varphi_j.$$

Then

$$\|f_k\| = F_{f_k}(f_k) = F_{f_k}(f - f^\epsilon + f^\epsilon - G_k) \leq \epsilon + F_{f_k}(f^\epsilon - G_k) \leq \epsilon + \|F_{f_k}\|_{\mathcal{D}}(A(\epsilon) + A_k). \quad \square$$

We give a new proof of the following theorem from [GK] on convergence of the Weak Dual Greedy Algorithm (WDGA) defined in the Introduction. We will prove the convergence result under an extra assumption on a Banach space X.

Definition 2.1 (Property Γ). *A uniformly smooth Banach space has property Γ if there is a constant $\beta > 0$ such that for any $x, y \in X$, satisfying $F_x(y) = 0$, we have*

$$\|x + y\| \geq \|x\| + \beta F_{x+y}(y).$$

Theorem 2.1. *Let X be a uniformly smooth Banach space with property Γ. Then the WDGA(τ) with $\tau = \{t\}$, $t \in (0, 1]$, converges for each dictionary and all $f \in X$.*

Proof. Let $\{f_m\}_{m=0}^{\infty}$ be a sequence generated by the WDGA(t). Then

$$f_{m-1} = f_m + a_m \varphi_m, \quad F_{f_m}(\varphi_m) = 0. \tag{2.4}$$

We use property Γ with $x := f_m$ and $y := a_m \varphi_m$ and obtain

$$\|f_{m-1}\| \geq \|f_m\| + \beta a_m F_{f_{m-1}}(\varphi_m). \tag{2.5}$$

This inequality and monotonicity of the sequence $\{\|f_m\|\}$ imply that

$$\sum_{m=1}^{\infty} a_m F_{f_{m-1}}(\varphi_m) < \infty \quad \Rightarrow \quad \sum_{m=1}^{\infty} a_m \|F_{f_{m-1}}\|_{\mathcal{D}} < \infty. \tag{2.6}$$

We consider separately two cases

$$(I) \quad \sum_{m=1}^{\infty} a_m = \infty, \qquad (II) \quad \sum_{m=1}^{\infty} a_m < \infty.$$

Consider the case (I). By property Γ we get (see (2.5))

$$\|f_m\| \le \|f_{m-1}\| - \beta a_m F_{f_{m-1}}(\varphi_m) \le \|f_{m-1}\| - t\beta a_m \|F_{f_{m-1}}\|_{\mathcal{D}}. \tag{2.7}$$

Let $\epsilon > 0$, $A(\epsilon) > 0$, and f^ϵ be such that

$$\|f - f^\epsilon\| \le \epsilon, \quad f^\epsilon / A(\epsilon) \in A_1(\mathcal{D}).$$

By Lemma 2.2

$$\|F_{f_{m-1}}\|_{\mathcal{D}} \ge (\|f_{m-1}\| - \epsilon)/(A(\epsilon) + A_{m-1}), \quad A_k := \sum_{j=1}^{k} a_j. \tag{2.8}$$

We complete the proof in case (I) by the contradiction argument. Assume $\lim_{m\to\infty} \|f_m\| = \alpha > 0$. Set $\epsilon := \alpha/2$. Then (2.7) and (2.8) imply

$$\|f_m\| \le \|f_{m-1}\| \left(1 - \frac{t\beta a_m}{2(A(\epsilon) + A_{m-1})}\right).$$

The assumption (I) implies

$$\sum_{m=1}^{\infty} \frac{a_m}{A(\epsilon) + A_{m-1}} = \infty \quad \Rightarrow \quad \|f_m\| \to 0.$$

In the second case (II) we also argue by contradiction. The argument in this case is the same as in [GK]. Assume

$$\lim_{m\to\infty} \|f_m\| = \alpha > 0.$$

Then by (II) we have $f_m \to f_\infty \ne 0$ as $m \to \infty$. By uniform smoothness of X we get

$$\lim_{m\to\infty} \|F_{f_m} - F_{f_\infty}\| = 0, \quad \lim_{m\to\infty} \|F_{f_m} - F_{f_{m-1}}\| = 0. \tag{2.9}$$

In particular, (2.4) and (2.9) imply that

$$\lim_{m\to\infty} F_{f_{m-1}}(\varphi_m) = 0 \quad \Rightarrow \quad \lim_{m\to\infty} \|F_{f_m}\|_{\mathcal{D}} = 0. \tag{2.10}$$

We have $F_{f_\infty} \ne 0$ and therefore there is a $g \in \mathcal{D}$ such that $F_{f_\infty}(g) > 0$. However, by (2.9) and (2.10)

$$F_{f_\infty}(g) = \lim_{m\to\infty} F_{f_m}(g) \le \lim_{m\to\infty} \|F_{f_m}\|_{\mathcal{D}} = 0.$$

The obtained contradiction completes the proof. $\square$

We now proceed to the convergence results in general uniformly smooth Banach spaces.

Theorem 2.2. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u)$. Assume that the coefficients sequence $\mathcal{C}$ satisfies the conditions*

$$\sum_{k=1}^{\infty} c_k = \infty, \tag{2.11}$$

$$\sum_{k=1}^{\infty} \rho(\theta c_k) < \infty \quad \textit{for any} \quad \theta > 0. \tag{2.12}$$

Then for the DGA($t,\mathcal{C}$) and for the XGA($\mathcal{C}$) we have for each dictionary and any $f \in X$

$$\liminf_{m\to\infty} \|f_m\| = 0.$$

Proof. In the case of the DGA($t,\mathcal{C}$) Theorem 2.2 has been proved in [T6]. We give here other proof that works for both algorithms from Theorem 2.2. Let f_{m-1} be a residual after $m-1$ steps of either the DGA($t,\mathcal{C}$) or the XGA($\mathcal{C}$). Let φ_m be such that

$$F_{f_{m-1}}(\varphi_m) \geq t\|F_{f_{m-1}}\|_{\mathcal{D}}. \tag{2.13}$$

Then

$$\inf_{g\in\mathcal{D}} \|f_{m-1} - c_m g\| \leq \|f_{m-1} - c_m\varphi_m\|.$$

Thus, in both cases (DGA($t,\mathcal{C}$) and XGA($\mathcal{C}$)) it is sufficient to estimate $\|f_{m-1} - c_m\varphi_m\|$ with φ_m satisfying (2.13). By Lemma 2.1 we get

$$\|f_{m-1} - c_m\varphi_m\| \leq \|f_{m-1}\|(1 + 2\rho(c_m/\|f_{m-1}\|)) - tc_m\|F_{f_{m-1}}\|_{\mathcal{D}}.$$

Applying Lemma 2.2 we obtain

$$\|f_{m-1} - c_m\varphi_m\| \leq \|f_{m-1}\|(1 + 2\rho(c_m/\|f_{m-1}\|)) - \frac{tc_m(\|f_{m-1}\| - \epsilon)}{A(\epsilon) + A_{m-1}}. \tag{2.14}$$

We complete the proof by the contradiction argument. Assume that $\liminf_{m\to\infty}\|f_m\| > 0$. Then there exist $\alpha > 0$ and N such that $\|f_m\| \geq \alpha$ for $m \geq N$. We set $\epsilon = \alpha/2$ and obtain from (2.14) the following inequality for both algorithms

$$\|f_m\| \leq \|f_{m-1}\| \left(1 - \frac{tc_m}{2(A(\epsilon) + A_{m-1})} + 2\rho(c_m/\alpha)\right), \quad m > N. \tag{2.15}$$

We now use our assumption on the $\mathcal{C}$: (2.11) implies

$$\sum_{m=1}^{\infty} \frac{tc_m}{A(\epsilon) + A_{m-1}} = \infty, \tag{2.16}$$

and (2.12) gives

$$\sum_{m=1}^{\infty} \rho(c_m/\alpha) < \infty. \tag{2.17}$$

The relations (2.15)–(2.17) imply that $\|f_m\| \to 0$ as $m \to \infty$. We got a contradiction that completes the proof. □

3. Convergence in the case of power type modulus of smoothness

In this section we prove some convergence results in the uniformly smooth Banach spaces with $\rho(u) \le \gamma u^q$, $q \in (1, 2]$. Assume that the coefficients sequence $\mathcal{C}$ satisfies the following conditions

$$\mathcal{C} \in \ell_q \setminus \ell_1, \tag{3.1}$$

$$\lim_{k\to\infty} c_k^{q-1} \sum_{j=1}^{k} c_j = 0. \tag{3.2}$$

The condition (3.1) corresponds to the conditions (2.11) and (2.12). The condition (3.2) is an extra condition. However, in some cases (3.2) is automatically satisfied. We give one result to that effect.

Proposition 3.1. *Any monotone sequence $\mathcal{C}$ satisfying (3.1) satisfies (3.2).*

Proof. A monotone sequence $\mathcal{C} \in \ell_q$ of positive numbers is a nonincreasing sequence. Let $\epsilon > 0$ be given. Find N such that $\sum_{j>N} c_j^q < \epsilon$. Then for any $k > N$ we have

$$c_k^{q-1} \sum_{j=1}^{k} c_j = c_k^{q-1} \sum_{j=1}^{N} c_j + c_k^{q-1} \sum_{j=N+1}^{k} c_j \tag{3.3}$$

$$\le c_k^{q-1} \sum_{j=1}^{N} c_j + \sum_{j=N+1}^{k} c_j^q \le c_k^{q-1} \sum_{j=1}^{N} c_j + \epsilon.$$

Taking into account that $c_k \to 0$ as $k \to \infty$, we complete the proof. □

Theorem 3.1. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1, 2]$. Assume that the coefficients sequence $\mathcal{C}$ satisfies (3.1) and (3.2). Then for each dictionary $\mathcal{D}$ and any $f \in X$ we have for the DGA$(t,\mathcal{C})$ and the XGA$(\mathcal{C})$*

$$\lim_{m\to\infty} \|f_m\| = 0.$$

Proof. We begin with the inequality (2.14) proved in Section 2 for φ_m satisfying (2.13)

$$\|f_{m-1} - c_m\varphi_m\| \le \|f_{m-1}\|(1 + 2\rho(c_m/\|f_{m-1}\|)) - \frac{tc_m(\|f_{m-1}\| - \epsilon)}{A(\epsilon) + A_{m-1}}. \tag{3.4}$$

We set $a_k := \|f_{k-1}\| - \epsilon$ and note that

$$a_{k+1} \le a_k + c_k. \tag{3.5}$$

Using the fact that the function $\rho(u)/u$ is monotone increasing on $[0,\infty)$ we obtain from (3.4) for $a_k > 0$

$$a_{k+1} \le a_k\left(1 - \frac{tc_k}{A(\epsilon)+A_{k-1}} + 2\frac{\|f_{k-1}\|}{a_k}\rho\left(\frac{c_k}{\|f_{k-1}\|}\right)\right) \tag{3.6}$$

$$\le a_k\left(1 - \frac{tc_k}{A(\epsilon)+A_{k-1}} + 2\rho\left(\frac{c_k}{a_k}\right)\right) \le a_k\left(1 - \frac{tc_k}{A(\epsilon)+A_{k-1}} + 2\gamma\left(\frac{c_k}{a_k}\right)^q\right).$$

Denote

$$c'_k := \frac{tc_k}{A(\epsilon)+A_{k-1}}.$$

We choose a sequence $\{b_k\}$ such that

$$2\gamma(c_k/b_k)^q = c'_k, \quad b_k = (2\gamma)^{1/q}c_k(c'_k)^{-1/q}.$$

Representing

$$b_k = (2\gamma)^{1/q}t^{-1/q}c_k^{1-1/q}(A(\epsilon)+A_{k-1})^{1/q}$$

we see from the assumption (3.2) that $b_k \to 0$ as $k \to \infty$.

The inequality (3.6) guarantees that for k such that $a_k \ge b_k$ we have $a_{k+1} \le a_k$.

Let

$$U := \{k : \quad a_k \ge b_k\}.$$

If the set U is finite then we get

$$\limsup_{k\to\infty} a_k \le \lim_{k\to\infty} b_k = 0.$$

This implies

$$\limsup_{m\to\infty}\|f_m\| \le \epsilon.$$

Consider the case when U is infinite. We note that Theorem 2.2 implies that there is a subsequence $\{k_j\}$ such that $a_{k_j} \le 0$, $j = 1, 2, \dots$. This means that

$$U = \cup_{j=1}^{\infty}[l_j, n_j]$$

with the property $n_{j-1} < l_j - 1$. For $k \notin U$ we have

$$a_k < b_k. \tag{3.7}$$

For $k \in [l_j, n_j]$ we have by (3.5) and the monotonicity property of a_k, when $k \in [l_j, n_j]$, that

$$a_k \le a_{l_j} \le a_{l_j-1} + c_{l_j-1} \le b_{l_j-1} + c_{l_j-1}. \tag{3.8}$$

By (3.7) and (3.8) we obtain

$$\limsup_{k\to\infty} a_k \le 0 \quad \Rightarrow \quad \limsup_{m\to\infty}\|f_m\| \le \epsilon.$$

Taking into account that $\epsilon > 0$ is arbitrary we complete the proof. □

Corollary 3.1. *Let $\mathcal{C} \in \ell_q \setminus \ell_1$ be a monotone sequence. Then the DGA($t,\mathcal{C}$) and the XGA($\mathcal{C}$) converge for each dictionary and all $f \in X$ in any uniformly smooth Banach space X with modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$.*

Proof. It follows from Theorem 3.1 and Proposition 3.1. □

4. Rate of approximation

In this section we consider the DGA($t,\mathcal{C}$) and the XGA($\mathcal{C}$) with a specific sequence $\mathcal{C}$. For a special $\mathcal{C}$ we prove the rate of convergence results in the uniformly smooth Banach spaces with modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$.

Theorem 4.1. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. We set $s := (1+1/q)/2$ and $\mathcal{C}_s := \{k^{-s}\}_{k=1}^{\infty}$. Then the DGA($t,\mathcal{C}_s$) and XGA($\mathcal{C}_s$) (for this algorithm $t=1$) converge for $f \in A_1(\mathcal{D})$ with the following rate: for any $r \in (0, t(1-s))$*

$$\|f_m\| \le C(r,t,q,\gamma)m^{-r}.$$

Proof. We begin with the inequality (2.14) proved in Section 2. Using the assumption $f \in A_1(\mathcal{D})$, we write (2.14) with $\epsilon = 0$ and $A(\epsilon) = 1$

$$\|f_m\| \le \|f_{m-1}\| \left(1 - \frac{tc_m}{1+A_{m-1}} + 2\gamma(c_m/\|f_{m-1}\|)^q\right). \tag{4.1}$$

We have

$$A_{m-1} = \sum_{k=1}^{m-1} k^{-s} \le 1 + \int_1^m x^{-s}dx = 1 + (1-s)^{-1}(m^{1-s} - 1).$$

Taking into account that $1 - s < 1/2$ we get

$$1 + A_{m-1} \le (1-s)^{-1}m^{1-s}.$$

Therefore,

$$\frac{tc_m}{1+A_{m-1}} \ge \frac{t(1-s)}{m}. \tag{4.2}$$

We will need the following technical lemma. This lemma is a more general version of Lemma 2.1 from [T1] (see also Remark 5.1 in [T7]).

Lemma 4.1. *Let a sequence $\{a_n\}_{n=1}^{\infty}$ have the following property. For given positive numbers $\alpha < \beta \le 1$, $A > a_1$ we have for all $n \ge 2$*

$$a_n \le a_{n-1} + A(n-1)^{-\alpha}; \tag{4.3}$$

if for some $\nu \ge 2$ we have

$$a_\nu \ge A\nu^{-\alpha}$$

then

$$a_{\nu+1} \le a_\nu(1 - \beta/\nu). \tag{4.4}$$

Then there exists a constant $C(\alpha,\beta)$ such that for all $n=1,2,\dots$ we have

$$a_n \le C(\alpha,\beta)An^{-\alpha}.$$

We will apply this lemma with $a_n := \|f_{n-1}\|$, $\alpha := r$, $\beta := (r+t(1-s))/2$ and A specified later. Let us check the conditions (4.3) and (4.4) of Lemma 4.1. By the inequality

$$\|f_m\| \le \|f_{m-1}\| + c_m = \|f_{m-1}\| + m^{-s}$$

the condition (4.3) holds for $A \ge 1$. Assume that $a_m \ge Am^{-r}$. Then

$$(c_m/a_m)^q \le A^{-q}m^{-(s-r)q} \le A^{-q}m^{-1}. \tag{4.5}$$

Setting $A := (2\gamma)^{1/q}(t(1-s)-\beta)^{-1/q}$ we obtain from (4.1), (4.2), and (4.5)

$$a_{m+1} \le a_m(1-\beta/m)$$

provided $a_m \ge Am^{-r}$. Thus (4.4) holds. Applying Lemma 4.1 we get

$$a_m \le C(r,t,q,\gamma)m^{-r}. \quad \square$$

We suggest to study the following asymptotic characteristics of greedy expansions. We begin with a pair of a Banach space X and a dictionary $\mathcal{D}$

$$v_m(X,\mathcal{D}) := \inf_{\mathcal{C}} \sup_{f\in A_1(\mathcal{D})} \|f_m\|,$$

where $\{f_m\}$ is a sequence of residuals of the XGA($\mathcal{C}$) (in this case we write $v_m^X(X,\mathcal{D})$) or of the DGA($t,\mathcal{C}$) (in this case we write $v_m^t(X,\mathcal{D})$). Next, for a Banach space X

$$v_m(X) := \inf_{\mathcal{C}} \sup_{\mathcal{D}} \sup_{f\in A_1(\mathcal{D})} \|f_m\|$$

for both the XGA($\mathcal{C}$) and the DGA($t,\mathcal{C}$). Finally, for a collection of Banach spaces with modulus of smoothness $\rho(u)\le\gamma u^q$, $q\in(0,1]$

$$v_m(\gamma,q) := \inf_{\mathcal{C}} \sup_{X:\rho(u)\le\gamma u^q} \sup_{\mathcal{D}} \sup_{f\in A_1(\mathcal{D})} \|f_m\|.$$

Theorem 4.1 gives an upper estimate for the $v_m(\gamma,q)$. It would be interesting to find the behavior of these characteristics and also to find optimal coefficients $\mathcal{C}$.

Our main interest in this paper is the study of greedy expansions with coefficients assigned in advance. We now discuss a different setting that is close in spirit to the main setting of the paper. This setting concerns a convex approximation of elements from $A_1(\mathcal{D})$. Generalizing a setting for a Hilbert space (see [J2], [B]) the authors of [DGDS] considered the following setting in a Banach space. We remind some definitions from [DGDS]. An incremental sequence is any sequence $a_1, a_2, \dots$ of X so that $a_1 \in \mathcal{D}$ and for each $n \geq 1$ there are some $g_n \in \mathcal{D}$ and $\lambda_n \in [0,1]$ so that

$$a_n = (1-\lambda_n)a_{n-1} + \lambda_n g_n, \quad (a_0 = 0).$$

We say that an incremental sequence $a_1, a_2, \dots$ is ϵ-greedy (with respect to f) if ($a_0 = 0$)

$$\|f - a_n\| < \inf_{\lambda\in[0,1]; g\in\mathcal{D}} \|f - ((1-\lambda)a_{n-1} + \lambda g)\| + \epsilon_n, \quad n = 1, 2, \dots.$$

It is pointed out in [DGDS] that the sequence $\{\lambda_n\}$ can be selected beforehand. They say that an incremental sequence $a_1, a_2, \dots$ is ϵ-greedy (with respect to f) with *convexity schedule* $\lambda_1, \lambda_2, \dots$ if ($a_0 = 0$)

$$\|f - a_n\| < \inf_{g\in\mathcal{D}} \|f - ((1-\lambda_n)a_{n-1} + \lambda_n g)\| + \epsilon_n, \quad n = 1, 2, \dots.$$

This is a X-greedy type algorithm. It is proved in [DGDS] that if X is a uniformly smooth Banach space with $\rho(u) \leq \gamma u^q$, $q \in (1, 2]$, then one can choose the convexity schedule $\lambda_n := 1/(n+1)$ and $\epsilon_n := \epsilon n^{-q}$ and get the rate of convergence

$$\|f - a_n\| \leq C(q, \gamma, \epsilon) n^{1/q-1}, \quad f \in A_1(\mathcal{D}).$$

We discuss an analogue of the above setting in the case of dual greedy algorithm. The following dual greedy algorithm was considered in [T3].

Weak Relaxed Greedy Algorithm (WRGA). *We define* $f_0^r := f_0^{r,\tau} := f$ *and* $G_0^r := G_0^{r,\tau} := 0$. *Then for each* $m \geq 1$ *we inductively define*

1). $\varphi_m^r := \varphi_m^{r,\tau} \in \mathcal{D}$ *is any satisfying*

$$F_{f_{m-1}^r}(\varphi_m^r - G_{m-1}^r) \geq t_m \sup_{g\in\mathcal{D}} F_{f_{m-1}^r}(g - G_{m-1}^r).$$

2). Find $0 \leq \lambda_m \leq 1$ *such that*

$$\|f - ((1-\lambda_m)G_{m-1}^r + \lambda_m \varphi_m^r)\| = \inf_{0\leq\lambda\leq1} \|f - ((1-\lambda)G_{m-1}^r + \lambda\varphi_m^r)\|$$

and define

$$G_m^r := G_m^{r,\tau} := (1-\lambda_m)G_{m-1}^r + \lambda_m \varphi_m^r.$$

3). Denote

$$f_m^r := f_m^{r,\tau} := f - G_m^r.$$

The following result from [T3] provides rate of approximation of the WRGA.

Theorem 4.2. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)\le\gamma u^q$, $1<q\le 2$. Then for a sequence $\tau:=\{t_k\}_{k=1}^\infty$, $t_k\le 1$, $k=1,2,\dots$, we have for any $f\in A_1(\mathcal{D})$ that*

$$\|f_m^{r,\tau}\|\le C_1(q,\gamma)(1+\sum_{k=1}^m t_k^p)^{-1/p},\quad p:=\frac{q}{q-1},$$

with a constant $C_1(q,\gamma)$ which may depend only on q and γ.

We consider the following variant of the WRGA with prescribed coefficients.

Weak Relaxed Greedy Algorithm with Equal Coefficients (WRGAEC). *Let $t\in(0,1]$. We define $f_0:=f$ and $G_0:=0$. Then for each $m\ge 1$ we inductively define*

1). $\varphi_m\in\mathcal{D}$ is any satisfying

$$F_{f_{m-1}}(\varphi_m-G_{m-1})\ge t\sup_{g\in\mathcal{D}}F_{f_{m-1}}(g-G_{m-1}).$$

2). Define

$$G_m:=(1-1/m)G_{m-1}+\varphi_m/m.$$

3). Denote

$$f_m:=f-G_m.$$

Theorem 4.3. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u)\le\gamma u^q$, $1<q\le 2$. Then for $t>1-1/q$ the WRGAEC converges for any $f\in A_1(\mathcal{D})$ with the rate*

$$\|f_m\|\le C(t,q,\gamma)m^{1/q-1}$$

with a constant $C(t,q,\gamma)$ which may depend only on t, q, and γ.

Proof. We use the following inequality proved in [T3, (3.2)] for $\lambda\in[0,1]$ with $\lambda=1/m$

$$\|f_{m-1}-\lambda(\varphi_m-G_{m-1})\|\le\|f_{m-1}\|(1-\lambda t+2\rho(2\lambda/\|f_{m-1}\|)). \tag{4.6}$$

We will use Lemma 4.1 in the same way as in the proof of Theorem 4.1. We specify $a_n:=\|f_{n-1}\|$, $\alpha:=1-1/q$, $\beta:=(t+1-1/q)/2$, and $A:=2(2\gamma)^{1/q}(t-\beta)^{-1/q}$. Then one can check that (4.6) implies (4.4). It is easy to check that (4.3) holds. Thus, application of Lemma 4.1 completes the proof. □

References

[B] Andrew R. Barron, *Universal approximation bounds for superposition of n sigmoidal functions*, IEEE Transactions on Information Theory **39** (1993), 930– 945.

[DT] R.A. DeVore and V.N. Temlyakov, *Some remarks on Greedy Algorithms*, Advances in Computational Mathematics **5** (1996), 173–187.

[DGDS] M. Donahue, L. Gurvits, C. Darken, E. Sontag, *Rate of convex approximation in non-Hilbert spaces*, Constr. Approx. **13** (1997), 187–220.

[FS] J.H. Friedman and W. Stuetzle, *Projection pursuit regression*, J. Amer. Statist. Assoc. **76** (1981), 817–823.

[GK] M. Ganichev and N.J. Kalton, *Convergence of the Weak Dual Greedy Algorithm in L_p-spaces*, J. Approx. Theory **124** (2003), 89–95.

[H] P.J. Huber, *Projection Pursuit*, The Annals of Statistics **13** (1985), 435–475.

[J1] L. Jones, *On a conjecture of Huber concerning the convergence of projection pursuit regression*, Annals of Stat. **15** (1987), 880–882.

[J2] L. Jones, *A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training*, The Annals of Statistics **20** (1992), 608–613.

[KT] S.V. Konyagin and V.N. Temlyakov, *Rate of convergence of Pure Greedy Algorithm*, East J. on Approx. **5** (1999), 493–499.

[LT] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces I*, Springer-Verlag, Berlin, 1977.

[LiT] E.D. Livshitz and V.N. Temlyakov, *Two lower estimates in greedy approximation*, Constr. Approximation **19** (2003), 509–523.

[S] A.V. Sil'nichenko, *On the rate of convergence of greedy algorithms*, manuscript (2004), 1–5.

[T1] V.N. Temlyakov, *Greedy Algorithms and m-term Approximation With Regard to Redundant Dictionaries*, J. Approx. Theory **98** (1999), 117–145.

[T2] V.N. Temlyakov, *Weak greedy algorithms*, Advances in Comput. Math. **12** (2000), 213–227.

[T3] V.N. Temlyakov, *Greedy algorithms in Banach spaces*, Advances in Comput. Math. **14** (2001), 277–292.

[T4] V.N. Temlyakov, *A criterion for convergence of Weak Greedy Algorithms*, Advances in Comput. Math. **17** (2002), 269–280.

[T5] V.N. Temlyakov, *Nonlinear Methods of Approximation*, Found. Comput. Math. **3** (2003), 33–107.

[T6] V.N. Temlyakov, *Greedy expansions in Banach spaces*, IMI Preprints Series **6** (2003), 1–21.

[T7] V.N. Temlyakov, *Greedy-Type Approximation in Banach Spaces and Applications*, Constr. Approx. **21** (2005), 257–292.

[T8] V.N. Temlyakov, *Relaxation in greedy approximation*, IMI-Preprints **03** (2006), 1–25.