

THE WCGA IN $L^p(\log L)^\alpha$ SPACES

G. GARRIGÓS

ABSTRACT. We present some new results concerning Lebesgue-type inequalities for the Weak Chebyshev Greedy Algorithm (WCGA) in uniformly smooth Banach spaces $\mathbb{X}$. First, we generalize Temlyakov's theorem [14] to cover situations in which the modulus of smoothness and the **A3** parameter are not necessarily power functions. Secondly, we apply this new theorem to the Zygmund spaces $\mathbb{X} = L^p(\log L)^\alpha$, with $1 < p < \infty$ and $\alpha \in \mathbb{R}$, and show that, when the Haar system is used, then optimal recovery of N -sparse signals occurs when the number of iterations is $\phi(N) = O(N^{\max\{1, 2/p'\}} (\log N)^{|\alpha|p'})$. Moreover, this quantity is sharp when $p \leq 2$. Finally, an expression for $\phi(N)$ in the case of the trigonometric system is also given.

1. INTRODUCTION

In this paper we consider several theoretical aspects regarding N -term approximation in a Banach space $(\mathbb{X}, \|\cdot\|)$, over a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

A fundamental question in this topic is, given a dictionary $\mathcal{D} = \{\varphi_i\}_{i \in \mathcal{I}}$ in $\mathbb{X}$, and the corresponding set of N -sparse vectors

$$\Sigma_N = \Sigma_N(\mathcal{D}) := \left\{ \sum_{j=1}^N c_j \varphi_{i_j} : c_j \in \mathbb{K}, \varphi_{i_j} \in \mathcal{D} \right\},$$

then find constructive procedures (algorithms) $\mathcal{A}_N : \mathbb{X} \rightarrow \Sigma_N$, where for all $f \in \mathbb{X}$ the quantity $\|f - \mathcal{A}_N(f)\|$ is as close as possible to the *best error of N -term approximation*, defined by

$$\sigma_N(f, \mathcal{D}) := \text{dist}(f, \Sigma_N) = \inf \left\{ \|f - g\| : g \in \Sigma_N(\mathcal{D}) \right\}.$$

Once an algorithm $\mathcal{A}_N$ is fixed, one can quantify the above statement by considering the associated *Lebesgue-type inequality*, which amounts to find the smallest value of $\phi(N)$ so that

$$(1.1) \quad \|f - \mathcal{A}_{\phi(N)}(f)\| \leq C \sigma_N(f), \quad \forall f \in \mathbb{X},$$

with C a fixed universal constant (if it exists). Observe, in particular, that (1.1) guarantees optimal recovery of all N -sparse signals after $\phi(N)$ iterations, that is

$$\mathcal{A}_{\phi(N)}(f) = f, \quad \forall f \in \Sigma_N(\mathcal{D}).$$

2010 *Mathematics Subject Classification.* 41A46, 41A25, 41A65, 46B15, 46B20, 46E30.

Key words and phrases. Non-linear approximation, greedy algorithm, uniformly smooth Banach space, Orlicz space, Haar system, trigonometric system.

Ideally, one would like to find algorithms $\mathcal{A}_N$ so that (1.1) holds with $\phi(N) = N$ (and $C = 1$). But this is hardly possible in many situations (a notable exception being when $\mathcal{D}$ is an orthonormal basis in a Hilbert space). For instance, in the classical case when $\mathcal{D}$ is the trigonometric system in $L^p(\mathbb{T})$, $p \neq 2$, it is still a relevant open question to find one such (constructive) algorithm.

In this paper we shall be interested in the *Weak Chebyshev Greedy Algorithm* (WCGA), which was introduced by Temlyakov in [12] as a generalization to Banach spaces of the celebrated *Orthogonal Matching Pursuit* (OMP) from Hilbert spaces. We refer to [13, 15, 16], and references therein, for background on this topic.

Lebesgue-type inequalities for the WCGA were proved in [7, 14]; see also [16, Chapter 8] for a historical overview. One the features of WCGA is that it has good approximation properties for the trigonometric system in L^p . Indeed, it was shown in [14, (4.3)] that, if $p > 2$, then Lebesgue inequalities hold with only $\phi(N) = O(N \log N)$ iterations. This seems to be the best known result with a constructive algorithm in that setting. Likewise, for the univariate Haar system in L^p , if $1 < p \leq 2$, then it suffices with $\phi(N) = O(N)$ iterations; see [14, (4.7)].

The above results are special cases a deep theorem proved by Temlyakov in [14, Theorem 2.8], which we describe in detail below. In that theorem, the number of iterations $\phi(N)$ is estimated in terms of some intrinsic properties of the pair $(\mathbb{X}, \mathcal{D})$, namely, the power type of the modulus of smoothness of $(\mathbb{X}, \|\cdot\|)$, and the power function associated with the so-called property **A3** of $\mathcal{D}$; see (1.20) below.

Our main result in this paper, Theorem 1.16, will be a generalization of Temlyakov's theorem, which allows to cover situations in which the modulus of smoothness and the **A3** parameters are not necessarily power functions. This is actually needed in some special cases, such as when $\mathbb{X} = L^p(\log L)^\alpha$, for which additional log factors appear naturally. Our next results, Theorems 1.21 and 5.22, will be applications of Theorem 1.16 to this setting, for two special dictionaries, the Haar and the trigonometric system.

We next give a more detailed description of these results.

1.1. Statements of results. We assume that $(\mathbb{X}, \|\cdot\|)$ is a uniformly smooth Banach space, meaning that its modulus of smoothness

$$(1.2) \quad \rho_{\mathbb{X}}(t) := \sup_{\|f\|=\|g\|=1} \frac{1}{2} \left(\|f + tg\| + \|f - tg\| - 2\|f\| \right), \quad t \in \mathbb{R}.$$

satisfies $\rho_{\mathbb{X}}(t) = o(t)$. Given $f \in \mathbb{X}$ with $f \neq 0$, we let $F_f \in \mathbb{X}^*$ be the associated norming functional, that is, the (unique) element in $\mathbb{X}^*$ such that

$$(1.3) \quad \|F_f\|_{\mathbb{X}^*} = 1, \quad \text{and} \quad F_f(f) = \|f\|.$$

Uniqueness follows from the smoothness of the norm $\|\cdot\|$.

We say that $\mathcal{D} = \{\varphi_i\}_{i \in \mathcal{I}}$ is a *dictionary* in $\mathbb{X}$, if it consists of non-null vectors whose closed linear span is $\mathbb{X}$, that is

$$[\varphi_i]_{i \in \mathcal{I}} = \mathbb{X}.$$

We do not assume the dictionary elements to be normalized, although as a consequence of later properties $\mathcal{D}$ will be *semi-normalized*, that is

$$\mathbf{c}_0 \leq \|\varphi_i\| \leq \mathbf{c}_1, \quad \forall \varphi_i \in \mathcal{D},$$

for some constants $\mathbf{c}_1 \geq \mathbf{c}_0 > 0$; see §2.1 below.

DEFINITION 1.4. Weak Chebyshev Greedy Algorithm (WCGA). *Given a fixed $\tau \in (0, 1]$, a τ -WCGA associated with $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ is any collection of mappings*

$$\mathcal{G}_N : \mathbb{X} \rightarrow \Sigma_N(\mathcal{D}), \quad N = 1, 2, \dots$$

with the following properties:

Given $f \in \mathbb{X} \setminus \{0\}$, we let $f_0 := f$ and define inductively vectors $\varphi_{i_1}, \dots, \varphi_{i_n}$ in $\mathcal{D}$ and $f_1, \dots, f_n \in \mathbb{X}$ by the following procedure: at step $n + 1$ we pick any $\varphi_{i_{n+1}} \in \mathcal{D}$ such that

$$(1.5) \quad |F_{f_n}(\varphi_{i_{n+1}})| \geq \tau \sup_{\varphi \in \mathcal{D}} |F_{f_n}(\varphi)|,$$

and let $\mathcal{G}_{n+1}(f)$ be any element in $[\varphi_{i_1}, \dots, \varphi_{i_{n+1}}]$ such that

$$(1.6) \quad \|f - \mathcal{G}_{n+1}(f)\| = \text{dist}(f, [\varphi_{i_1}, \dots, \varphi_{i_{n+1}}]).$$

Then we set $f_{n+1} = f - \mathcal{G}_{n+1}(f)$, and iterate the process (indefinitely, or until the remainder $f_{n+1} = 0$).

If at some stage we have $f_n = 0$, then we just let $\mathcal{G}_{n+k}(f) = \mathcal{G}_n(f) = f$ for all $k \geq 1$.

REMARK 1.7. Note that such algorithms can always be constructed when $\tau < 1$, and for some dictionaries also when $\tau = 1$ (namely, when the sup in (1.5) is attained within $\mathcal{D}$).

We next define the three key properties that are needed to prove Lebesgue-type inequalities for WCGA. The first one is a generalization of a property given in [2, Definition 1.13].

DEFINITION 1.8. *Let $Q(t)$ be a positive increasing function for $t \in (0, \infty)$, with $Q(0) = 0$. We say that $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ satisfies **property D(Q)** if*

$$(1.9) \quad \text{dist}(f, [\varphi]) \leq \|f\| \left(1 - Q(|F_f(\varphi)|)\right), \quad \forall \varphi \in \mathcal{D}, f \in \mathbb{X} \setminus \{0\}.$$

The next definition coincides with property A2 from [7, 14].

DEFINITION 1.10. *Let $N < D$ be positive integers and $k_N > 0$. We say that $\Sigma_N(\mathcal{D}) \in \mathbf{A2}(k_N, D)$ if*

$$(1.11) \quad \left\| \sum_{j \in A} a_j \varphi_j \right\| \leq k_N \left\| \sum_{j \in B} a_j \varphi_j \right\|, \quad \forall a_j \in \mathbb{K}, \forall A \subset B : |A| \leq N, |B| < D$$

If the above holds for all $D < \infty$, we just write $\Sigma_N(\mathcal{D}) \in \mathbf{A2}(k_N)$.

Our third definition is a slight generalization of property A3 from [14].

DEFINITION 1.12. *Let $N < D$ be positive integers and let $\{H(k)\}_{k=1}^\infty$ be an increasing sequence of positive numbers. We say $\Sigma_N(\mathcal{D}) \in \mathbf{A3}(H, D)$ if*

$$(1.13) \quad \sum_{j \in A} |a_j| \leq H(|A|) \left\| \sum_{j \in B} a_j \varphi_j \right\|, \quad \forall a_j \in \mathbb{K}, \forall A \subset B : |A| \leq N, |B| < D.$$

If the above holds for all $D < \infty$, we just write $\Sigma_N(\mathcal{D}) \in \mathbf{A3}(H)$.

Finally, we recall that a positive sequence $\{G(k)\}_{k=1}^\infty$ is called *1-quasi-convex* if

$$(1.14) \quad \frac{G(k)}{k} \leq \frac{G(k+1)}{k+1}, \quad \forall k \in \mathbb{N}.$$

As an example, if $G(t)$ is a positive convex function in $(0, \infty)$ with $G(0^+) = 0$, then $\{G(k)\}_{k=1}^\infty$ is 1-quasi-convex. This is the case, for instance, for the functions

$$(1.15) \quad G(t) = t^p (\log(c+t))^\alpha,$$

if $p = 1$ and $\alpha \geq 0$, or if $p > 1$ and $\alpha \in \mathbb{R}$ (for a sufficiently large $c \geq e$).

The precise statement of our main result is now the following.

THEOREM 1.16. *Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space, $\mathcal{D}$ a dictionary, $\tau \in (0, 1]$ and $\mathcal{G}_n : \mathbb{X} \rightarrow \Sigma_n$ a τ -WCGA. Let $D > N \geq 1$ be fixed. Let $k_N > 0$ and let $Q(t)$, $H(n)$ be positive and increasing functions such that the following properties hold*

- (i) $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ satisfies $D(Q)$
- (ii) Σ_N satisfies property $\mathbf{A2}(k_N, D)$.
- (iii) Σ_N satisfies property $\mathbf{A3}(H, D)$.

Let $\lambda_1 > 1$. Assume further that the sequence

$$(1.17) \quad G(n) = \left[Q\left(\frac{c(\tau)}{H(n)}\right) \right]^{-1}, \quad \text{with } c(\tau) = \frac{\tau}{2} \left(1 - \frac{1}{\sqrt{\lambda_1}}\right).$$

is 1-quasi-convex. If we let

$$(1.18) \quad \phi(N) := 8 \ln \left[\frac{8(1 + \lambda_1)k_N}{\sqrt{\lambda_1} - 1} \right] G(2N),$$

then it holds

$$(1.19) \quad \left\| x - \mathcal{G}_{\phi(N)}(x) \right\| \leq \lambda_1 \|x - \Phi\|, \quad \forall x \in \mathbb{X}, \Phi \in \Sigma_N,$$

provided that $N + \phi(N) < D$.

We now make some comments about this theorem.

a) The result obtained by Temlyakov in [14, Theorem 2.8] corresponds to the case when λ_1 is a (possibly large) universal constant, and

$$H(N) = V_N N^r \quad \text{and} \quad Q(t) = c t^{q'},$$

where $q > 1$ is the power type of the modulus of smoothness, ie $\rho_{\mathbb{X}}(t) = O(t^q)$. In that case, the required number of iterations becomes

$$(1.20) \quad \phi(N) = C_1 (V_N/\tau)^{q'} \log(1 + k_N) N^{rq'},$$

for some $C_1 > 0$, provided that $rq' \geq 1$. Our contribution gives an additional explicit form for the constants when the parameter λ_1 approaches 1.

b) As we show in Proposition 2.4 below, if $\mathcal{D}$ is normalized, then condition $\mathbf{D}(Q)$ always holds with

$$Q(t) = 2\delta_{\mathbb{X}^*}(t/2),$$

where $\delta_{\mathbb{X}^*}(t)$ is the *modulus of convexity* of the dual space $\mathbb{X}^*$. This is also a new result. In many practical cases the asymptotic behavior of $\delta_{\mathbb{X}^*}(t)$ is well-known, so one can use property $\mathbf{D}(Q)$ with no need to compute $\rho_{\mathbb{X}}(t)$.

c) As was discussed in [2, Remark 2.10], in some special cases it is possible to prove that $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ satisfies property $\mathbf{D}(Q)$ with a function $Q(t)$ which is considerably better than $\delta_{\mathbb{X}^*}(t)$ (for t near 0). For instance, if $\mathbb{X} = \ell^p$ and $\mathcal{D}$ is the canonical basis, then one can take $Q(t) = c_p t^{p'}$, which gives better results than $\delta_{\mathbb{X}^*}(t) = O(t^{\max\{p', 2\}})$ when $p > 2$. Other examples (with power type) were given in [2, Proposition 4.12 and Lemma 5.7].

d) The assumption that $G(n)$ in (1.17) is 1-quasi-convex is only made for convenience. Alternatively, one could replace $G(n)$ by any convex majorant (hence, 1-quasi-convex). In practice, quasi-convexity is easily verified after substituting the functions $Q(t)$ and $H(n)$ into (1.17); see the example in (1.15).

e) As in [14], the conclusion (1.19) in the previous theorem also holds when the assumptions **A2** and **A3** are required *only* on the individual sparse element $\Phi = \sum_{j \in T} x_j \varphi_j$, with $|T| \leq N$ (and not necessarily in all $\Phi \in \Sigma_N$). Namely, in this case the requirement would be that (1.11) and (1.13) must hold for all $A \subset T$ and all scalars $a_j \in \mathbb{K}$ such that $a_j = x_j$, $j \in A$.

Our second result is an application of Theorem 1.16 to the case when $\mathbb{X} = L^p(\log L)^\alpha$; see §4 below for the precise definition. We stress that, when $1 < p \leq 2$, the number of iterations which are derived from the above theorem, namely

$$\phi(N) = O\left(N (\log(e + N))^{p'|\alpha|}\right),$$

is actually (asymptotically) optimal for all $\alpha \in \mathbb{R}$.

THEOREM 1.21. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$, and let $\mathbb{X} = L^p(\log L)^\alpha$ be as in §4. Let $\mathcal{D}$ be the (normalized) Haar basis in $\mathbb{X}$. Then*

a) there exists a constant $C > 1$ such that the WCGA satisfies

$$(1.22) \quad \left\| f - \mathcal{G}_{\phi(N)}(f) \right\| \leq 2 \sigma_N(f), \quad \forall f \in \mathbb{X}, N \in \mathbb{N},$$

where

$$(1.23) \quad \phi(N) = \begin{cases} C N^{\frac{2}{p'}} (\log(e + N))^{2\alpha_+} & \text{when } p > 2 \\ C N (\log(e + N))^{p'|\alpha|} & \text{when } 1 < p \leq 2. \end{cases}$$

b) if for some sequence $\psi(N)$ the WCGA satisfies

$$(1.24) \quad \left\| f - \mathcal{G}_{\psi(N)}(f) \right\| \leq C \sigma_N(f), \quad \forall f \in \mathbb{X}, N \in \mathbb{N},$$

then necessarily $\psi(N) \geq c' N (\log(e + N))^{| \alpha | p'}$, for some $c' > 0$.

REMARK 1.25. We remark that, when $p > 2$, it is an open question already for L^p spaces (case $\alpha = 0$) whether $\phi(N) \approx N^{2/p'}$ iterations are necessary to ensure (1.22); see [16, Open Problem 8.3, p. 448].

Finally, in §5 we give a similar application in the case that $\mathbb{X} = L^p(\log L)^\alpha$ and $\mathcal{D} = \{e^{inx}\}_{n \in \mathbb{Z}}$ is the trigonometric system. See Theorem 5.22 below for details.

2. PRELIMINARIES

2.1. About seminormalization of $\mathcal{D}$. We claim that the two properties $\mathbf{D}(Q)$ and $\mathbf{A3}(H, D)$ imply that the dictionary $\mathcal{D}$ must be semi-normalized. Indeed, if Σ_1 satisfies $\mathbf{A3}(H, D)$ then

$$(2.1) \quad \|\varphi\| \geq 1/H(1), \quad \forall \varphi \in \mathcal{D}.$$

On the other hand, $\mathbf{D}(Q)$ implies that $Q(|F_f(\varphi)|) \leq 1$ for all $\varphi \in \mathcal{D}$ and $f \in \mathbb{X} \setminus \{0\}$. Setting $f = \varphi$ and using $F_\varphi(\varphi) = \|\varphi\|$, this gives

$$(2.2) \quad \|\varphi\| \leq Q^{-1}(1), \quad \forall \varphi \in \mathcal{D}.$$

Conversely, suppose that $\mathcal{D} = \{\varphi_j\}$ is a dictionary satisfying any of the properties $\mathbf{D}(Q)$, $\mathbf{A2}(k_N, D)$ or $\mathbf{A3}(H, D)$, and let $\tilde{\varphi}_j = \lambda_j \varphi_j$ for scalars λ_j such that

$$0 < \mathbf{c}_0 \leq |\lambda_j| \leq \mathbf{c}_1, \quad \forall j.$$

It is then easily seen that the new dictionary $\tilde{\mathcal{D}} = \{\tilde{\varphi}_j\}$ satisfies the corresponding properties with new parameters, namely

$$\mathbf{D}(Q(\cdot/\mathbf{c}_1)), \quad \mathbf{A2}(k_N, D) \quad \text{or} \quad \mathbf{A3}(H/\mathbf{c}_0, D)$$

We also remark that if $\mathcal{G}_N$ is τ -WCGA for $\mathcal{D}$, then it is also a $(\tau\mathbf{c}_0/\mathbf{c}_1)$ -WCGA for $\tilde{\mathcal{D}}$.

2.2. About condition $\mathbf{D}(Q)$. We give a practical criterion which ensures that condition $\mathbf{D}(Q)$ holds. Let $(\mathbb{X}, \|\cdot\|)$ be a Banach space with *modulus of smoothness*

$$\rho_{\mathbb{X}}(t) = \sup_{\|x\|=\|y\|=1} \frac{1}{2} \left(\|x + ty\| + \|x - ty\| - 2\|x\| \right), \quad t \in \mathbb{R}.$$

We denote by $\delta(s) = \delta_{\mathbb{X}}(s)$ its *modulus of convexity*, that is

$$\delta_{\mathbb{X}}(s) = \inf_{\substack{\|x\|=\|y\|=1 \\ \|x-y\|=s}} \left(\frac{\|x\| + \|y\|}{2} - \left\| \frac{x+y}{2} \right\| \right), \quad s \in [0, 2].$$

Next, we consider the following related function, introduced by Figiel [3],

$$(2.3) \quad \tilde{\delta}_{\mathbb{X}^*}(s) := \sup_{t \geq 0} \left(\frac{1}{2}st - \rho_{\mathbb{X}}(t) \right), \quad s \geq 0.$$

Assume for simplicity that $\mathbb{X}$ is uniformly smooth, that is $\rho_{\mathbb{X}}(t) = o(t)$ when $t \rightarrow 0$ (so in particular, $\mathbb{X}$ is reflexive). Then, it is easily seen that $Q(s) = \tilde{\delta}_{\mathbb{X}^*}(s)$ is a convex increasing function with $Q(0) = 0$. Moreover, it is shown in [3, Proposition 1] (see also [6, Proposition 1.e.6]) that $\tilde{\delta}_{\mathbb{X}^*}(s)$ is “*equivalent*” to $\delta_{\mathbb{X}^*}(s)$ (for small s), in the sense that

$$\delta_{\mathbb{X}^*}(s/2) \leq \tilde{\delta}_{\mathbb{X}^*}(s) \leq \delta_{\mathbb{X}^*}(s), \quad s \in [0, 2].$$

Also, $\tilde{\delta}_{\mathbb{X}^*}(s)$ is the greatest convex minorant of $\delta_{\mathbb{X}^*}(s)$. In particular, $\tilde{\delta}_{\mathbb{X}^*} = \delta_{\mathbb{X}^*}$ when the later is a convex function. In many examples of Banach spaces $\mathbb{X}$, the behavior of the function $\delta_{\mathbb{X}^*}(s)$ is well-known (sometimes quite explicitly). For instance, if $\mathbb{X} = L^p$, $1 < p < \infty$, then

$$\delta_{L^{p'}}(s) = c_q s^q + o(s^q), \quad \text{with } q = \max\{2, p'\};$$

see [6, p.63]. Our main result in this section is the following.

PROPOSITION 2.4. *If $(\mathbb{X}, \|\cdot\|)$ is uniformly smooth, then every normalized dictionary $\mathcal{D}$ in $\mathbb{X}$ satisfies property $\mathbf{D}(Q)$ with $Q(s) = 2\tilde{\delta}_{\mathbb{X}^*}(s)$.*

Proof. It suffices to prove (1.9) for $f = x \in \mathbb{X}$ with $\|x\| = 1$. Let $F_x \in \mathbb{X}^*$ be the norming functional of $\mathbb{X}$, and given $\varphi \in \mathcal{D}$, let $\nu = \overline{\text{sign}} F_x(\varphi)$. Then, for every $t \geq 0$, using [2, Proposition 2.1], we have

$$\text{dist}(x, [\varphi]) \leq \|x - \nu t \varphi\| \leq \|x\| - t |F_x(\varphi)| + 2\rho_{\mathbb{X}}(t).$$

Taking the infimum over all $t \geq 0$ we obtain

$$\text{dist}(x, [\varphi]) \leq 1 - 2 \sup_{t \geq 0} \left(\frac{1}{2} t |F_x(\varphi)| - \rho_{\mathbb{X}}(t) \right) = 1 - 2\tilde{\delta}_{\mathbb{X}^*}(|F_x(\varphi)|).$$

□

REMARK 2.5. If the dictionary is not normalized, but we assume that $0 < \|\varphi\| \leq \mathbf{c}_1$, for all $\varphi \in \mathcal{D}$, then the previous result gives

$$\text{dist}(x, [\varphi]) \leq 1 - 2\tilde{\delta}_{\mathbb{X}^*}(|F_x(\varphi/\|\varphi\|)|) \leq 1 - 2\tilde{\delta}_{\mathbb{X}^*}(|F_x(\varphi)|/\mathbf{c}_1).$$

So property D(Q) holds with $Q(s) = 2\tilde{\delta}_{\mathbb{X}^*}(s/\mathbf{c}_1)$, which is also a function equivalent to $\delta_{\mathbb{X}^*}(s)$.

2.3. About condition A3(H). In practice, it is quite common that $(\mathbb{X}, \mathcal{D})$ satisfies properties A2 or A3 with depth $D = \infty$. Our first observation is that this implies that $\mathcal{D}$ has a biorthogonal dual system.

LEMMA 2.6. *Let $\mathcal{D} = \{\varphi_j\}_{j=1}^{\infty}$ be a dictionary in $\mathbb{X}$. Assume that one of the following properties hold*

- (i) *There exists $k_1 > 0$ such that $\Sigma_1(\mathcal{D}) \in \mathbf{A2}(k_1; D)$, for all $D < \infty$*
- (ii) *There exists $H(1) > 0$ such that $\Sigma_1(\mathcal{D}) \in \mathbf{A3}(H; D)$, for all $D < \infty$.*

Then, there exists $\{\varphi_j^\}_{j=1}^{\infty}$ in $\mathbb{X}^*$ such that $\{\varphi_j, \varphi_j^*\}_{j=1}^{\infty}$ is a biorthogonal system, ie*

$$\varphi_j^*(\varphi_i) = 0, \text{ if } j \neq i, \quad \text{and} \quad \varphi_j^*(\varphi_j) = 1.$$

Proof. This is a consequence of [11, Theorem 6.1, page 54]. Indeed, if (i) holds then

$$\left\| \sum_{j=1}^n a_j \varphi_j \right\| \leq \sum_{j=1}^n \|a_j \varphi_j\| \leq n k_1 \left\| \sum_{i=1}^{n+m} a_i \varphi_i \right\|,$$

which implies biorthogonality by [11, Theorem 6.1, “8° $\Rightarrow$ 2°”]. Similarly, if (ii) holds then

$$\sum_{j=1}^n \frac{|a_j|}{2^j H(1)} \leq \sum_{j=1}^n 2^{-j} \left\| \sum_{i=1}^n a_i \varphi_i \right\| \leq \left\| \sum_{i=1}^n a_i \varphi_i \right\|,$$

which implies biorthogonality by [11, Theorem 6.1, “4° $\Rightarrow$ 2°”]. □

So under this situation, the dictionary $\mathcal{D}$ generates a dual system $\mathcal{D}^*$. Then, a variation of [2, Lemma 2.17] gives the following.

LEMMA 2.7. *Let $\mathcal{D} = \{\varphi_j\}_{j=1}^{\infty}$ be a dictionary, with dual system $\mathcal{D}^* = \{\varphi_j^*\}_{j \geq 1}$. Then, $\Sigma_N \in \mathbf{A3}(H, D)$, for all $N < D < \infty$, if we choose*

$$(2.8) \quad H(n) = \sup_{|A| \leq n, |\varepsilon_j| = 1} \left\| \sum_{j \in A} \varepsilon_j \varphi_j^* \right\|_{\mathbb{X}^*}.$$

Proof. Take sets $A \subset B$, with $|A| \leq N$, and scalars $a_j \in \mathbb{K}$. Let $\varepsilon_j = \overline{\text{sign}} a_j$, and denote

$$\mathbb{1}_{\varepsilon A}^* := \sum_{j \in A} \varepsilon_j \varphi_j^* \in \mathbb{X}^*.$$

Then

$$\sum_{n \in A} |a_n| = \mathbb{1}_{\varepsilon A}^* \left(\sum_{n \in B} a_n \varphi_n \right) \leq \|\mathbb{1}_{\varepsilon A}^*\|_{\mathbb{X}^*} \left\| \sum_{n \in B} a_n \varphi_n \right\| \leq H(N) \left\| \sum_{n \in B} a_n \varphi_n \right\|.$$

□

In practice, the sequence $H(n)$ in (2.8) is equivalent to the fundamental function of $\mathcal{D}^*$ in $\mathbb{X}^*$, which in many examples has an explicit expression.

2.4. Quasi-convex sequences. Given a positive sequence $w = \{w(j)\}_{j=1}^\infty$ we define its associated *summing sequence* $\tilde{w}$ by

$$(2.9) \quad \tilde{w}(n) := \sum_{j=1}^n \frac{w(j)}{j}, \quad n \geq 1.$$

LEMMA 2.10. *If $w = \{w(j)\}_{j=1}^\infty$ is non-decreasing then for all $N \in \mathbb{N}$*

$$\sum_{j: 1 \leq 2^j \leq N} w(2^j) < 2\tilde{w}(N).$$

Proof. Let $\Delta_j = \{n \in \mathbb{N} : 2^j \leq n < 2^{j+1}\}$, which has cardinality 2^j , $j = 0, 1, \dots$. Then, if J is the largest integer with $2^J \leq N$ we have

$$\sum_{j=0}^J w(2^j) \leq \sum_{j=0}^J \frac{\sum_{n \in \Delta_j} w(n)}{2^j} < 2 \sum_{j=0}^J \sum_{n \in \Delta_j} \frac{w(n)}{n} = 2\tilde{w}(2^J) \leq 2\tilde{w}(N).$$

□

LEMMA 2.11. *If $w = \{w(j)\}_{j=1}^\infty$ is 1-quasi-convex then*

- a) $\tilde{w}(N) \leq w(N)$ for all $N \in \mathbb{N}$
- b) $\tilde{w}$ is superadditive, that is,

$$\tilde{w}(M + N) \geq \tilde{w}(M) + \tilde{w}(N), \quad \forall M, N \in \mathbb{N}.$$

Proof. The assertion a) follows from the definition of 1-quasi-convex, since

$$\tilde{w}(N) = \sum_{n=1}^N \frac{w(n)}{n} \leq \sum_{n=1}^N \frac{w(N)}{N} = w(N).$$

The assertion b) follows similarly from

$$\begin{aligned}
\tilde{w}(M+N) &= \sum_{n=1}^N \frac{w(n)}{n} + \sum_{n=N+1}^{N+M} \frac{w(n)}{n} \\
&= \tilde{w}(N) + \sum_{j=1}^M \frac{w(j+N)}{j+N} \\
&\geq \tilde{w}(N) + \sum_{j=1}^M \frac{w(j)}{j} = \tilde{w}(N) + \tilde{w}(M).
\end{aligned}$$

□

3. THE PROOF OF THEOREM 1.16

In this section we give the proof of Theorem 1.16. We shall follow the main steps in the original proof of Temlyakov, see [14, Theorem 2.8] or [16, Theorem 8.7.18], adapted to the new properties $D(Q)$ and $A3(H, D)$. For completeness, we give self-contained arguments of all the steps, although the main changes will mostly appear in steps 1 and 4.

3.1. Step 1. The iteration theorem. The following result is a generalization of [2, Theorem 3.1], so we follow the notation presented there. Namely, if $f \in \mathbb{X} \setminus \{0\}$, then we write

$$f_n := f - \mathcal{G}_n(f) \quad \text{and} \quad \Gamma_n := \text{supp } \mathcal{G}_n(f),$$

for the remainder and the supporting set of the n -th WCGA applied to f ; see Definition 1.4. Also, if $\Phi = \sum_{j \in T} a_j \varphi_j \in \Sigma_N$ and $A \subset T$, then we denote

$$\Phi_A := \sum_{j \in T \cap A} a_j \varphi_j, \quad \text{and} \quad T_n := T \setminus \Gamma_n.$$

We shall also make frequent use of [2, Lemma 2.12], which asserts that

$$(3.1) \quad F_{f_n}(g) = 0, \quad \forall g \in [\varphi_j]_{j \in \Gamma_n}.$$

THEOREM 3.2. *Let $D > N \geq 1$. Assume that*

- (i) $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ satisfies $D(Q)$
- (ii) Σ_N satisfies property $A3(H, D)$

Then, for every $f \in \mathbb{X} \setminus \{0\}$, $\Phi = \sum_{j \in T} a_j \varphi_j \in \Sigma_N$, and $\lambda > 1$, and for all integers $m, M \geq 0$ such that $N + m + M < D$ the following holds

$$(3.3) \quad \|f_{m+M}\| \leq e^{-M/G(|A|)} \|f_m\| + \lambda \left(\|f - \Phi\| + \|\Phi_B\| \right),$$

for all sets $A \subset T_k$ (with $A \neq \emptyset$), $B = T_k \setminus A$ and all $k \in [0, m)$, and where

$$(3.4) \quad G(n) = \frac{1}{Q(c(\tau)/H(n))}, \quad \text{with} \quad c(\tau) = \frac{\tau}{2} \left(1 - \frac{1}{\lambda} \right).$$

Proof. Given a fixed $n \in [m, m + M)$, condition $\mathbf{D}(Q)$ implies

$$(3.5) \quad \|f_{n+1}\| \leq \text{dist}(f_n, [\varphi_{i_{n+1}}]) \leq \|f_n\| \left(1 - Q(|F_{f_n}(\varphi_{i_{n+1}})|)\right).$$

By definition of the WCGA and (3.1), for each (non-empty) $A \subset T$ we have

$$\tau |F_{f_n}(\Phi_A)| = \tau |F_{f_n}(\Phi_{A \cap T_n})| \leq \left(\sum_{A \cap T_n} |a_i| \right) |F_{f_n}(\varphi_{i_{n+1}})|.$$

Now, the assumption $\Sigma_N \in \mathbf{A3}(H, D)$ implies that

$$\sum_{A \cap T_n} |a_i| \leq H(|A|) \|\Phi - \mathcal{G}_n(f)\| \leq H(|A|) (\|\Phi - f\| + \|f_n\|).$$

In order to apply $\mathbf{A3}$ we have used that $|A \cap T_n| \leq |T| \leq N$ and $|T \cup \Gamma_n| \leq N + n \leq N + m + M < D$. Thus, inserting these estimates into (3.5) we obtain

$$\|f_{n+1}\| \leq \|f_n\| \left(1 - Q\left(\frac{\tau |F_{f_n}(\Phi_A)|}{H(|A|)(\|\Phi - f\| + \|f_n\|)}\right)\right),$$

which is valid for all sets $A \subset T$.

Fix now an integer $k \in [0, m]$ and a set $A \subset T_k$, and let $B = T_k \setminus A$. Since $\Gamma_k \subset \Gamma_n$ we can use (3.1) to obtain

$$\begin{aligned} |F_{f_n}(\Phi_A)| &= |F_{f_n}(\Phi_A + \Phi_{\Gamma_k \cap T} - \mathcal{G}_n(f))| = |F_{f_n}(\Phi_T - \Phi_B - f + f_n)| \\ &\geq \|f_n\| - \|f - \Phi\| - \|\Phi_B\|. \end{aligned}$$

So, we conclude that

$$\|f_{n+1}\| \leq \|f_n\| \left(1 - Q\left(\frac{\tau (\|f_n\| - \|f - \Phi\| - \|\Phi_B\|)_+}{H(|A|)(\|\Phi - f\| + \|f_n\|)}\right)\right).$$

Using in the denominator that $\|\Phi - f\| \leq \|f_n\|$ (when the numerator is not zero), this further simplifies into

$$\begin{aligned} \|f_{n+1}\| &\leq \|f_n\| \left(1 - Q\left(\frac{\tau (\|f_n\| - \|f - \Phi\| - \|\Phi_B\|)_+}{2H(|A|)\|f_n\|}\right)\right) \\ &= \|f_n\| \left(1 - Q\left(\frac{\tau (1 - u)_+}{2H(|A|)}\right)\right), \end{aligned}$$

where we have let $u = (\|f - \Phi\| + \|\Phi_B\|)/\|f_n\|$. Now, call

$$\beta = Q\left(\frac{(1 - \frac{1}{\lambda})\tau}{2H(|A|)}\right) = \frac{1}{G(|A|)},$$

and observe that, if $u \leq 1/\lambda$ then we have

$$(3.6) \quad \|f_{n+1}\| \leq (1 - \beta) \|f_n\|.$$

On the other hand, if $u \geq 1/\lambda$, by definition of u we have

$$\|f_n\| \leq \lambda (\|f - \Phi\| + \|\Phi_B\|),$$

and therefore,

$$\begin{aligned}
 \|f_{n+1}\| &\leq \|f_n\| = (1 - \beta) \|f_n\| + \beta \|f_n\| \\
 (3.7) \qquad &\leq (1 - \beta) \|f_n\| + \beta \lambda (\|f - \Phi\| + \|\Phi_B\|).
 \end{aligned}$$

So, combining (3.6) and (3.7), and calling $v = \|f - \Phi\| + \|\Phi_B\|$ we obtain

$$\|f_{n+1}\| - \lambda v \leq (1 - \beta) (\|f_n\| - \lambda v).$$

Since $\|f_{n+1}\| \leq \|f_n\|$ this implies

$$(\|f_{n+1}\| - \lambda v)_+ \leq (1 - \beta) (\|f_n\| - \lambda v)_+.$$

We can now iterate for all $n \in [m, m + M)$ to obtain

$$\|f_{m+M}\| - \lambda v \leq (1 - \beta)^M (\|f_m\| - \lambda v)_+ \leq (1 - \beta)^M \|f_m\|.$$

Finally, using the value of v and $1 - \beta \leq e^{-\beta}$ we obtain

$$\|f_{m+M}\| \leq e^{-M\beta} \|f_m\| + \lambda (\|f - \Phi\| + \|\Phi_B\|).$$

This corresponds exactly to (3.3). □

3.2. Step 2: Selection of sets A_j . In the next step, we shall follow [16, pp. 435–437], and iteratively apply Theorem 3.2, with a suitably chosen selection of sets A_j , in order to obtain the following result. We have adapted the proof to include the new conditions $D(Q)$ and $A3(H, D)$, and have made more precise the value of the constants.

THEOREM 3.8. *Let $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ and $1 \leq N < D$ be such that*

- (i) $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ satisfies $D(Q)$
- (ii) Σ_N satisfies property $A2(k_N, D)$.
- (iii) Σ_N satisfies property $A3(H, D)$.

Given $\lambda > 1$ and $\delta > 0$, there exists $\beta_0 = \beta_0(\lambda, \delta, k_N) > 0$ such that, if $\beta \geq \beta_0$ and $\Phi = \sum_{j \in T} a_j \varphi_j \in \Sigma_N \setminus \{0\}$, then there exist positive integers $L, m_L \in \mathbb{N}$ such that

$$(3.9) \qquad 2^{L-2} < |T| \quad \text{and} \quad m_L \leq \beta \sum_{j=1}^L G(2^{j-1}),$$

(with $G(n)$ defined in (3.4)), and so that for all $x \in \mathbb{X} \setminus \{0\}$ it holds

$$\text{either } \|x_{m_L}\| \leq (1 + \delta) \lambda \|x - \Phi\|, \quad \text{or} \quad |T \cap \Gamma_{m_L}| > 2^{L-2},$$

provided that $N + m_L < D$. Moreover, we can set

$$\beta_0 = 2 \ln \left(\frac{8k_N(1 + (1 + \delta)\lambda)}{\delta} \right).$$

Proof. Let $n \geq 0$ be such that $2^{n-1} < |T| \leq 2^n$. Now, for each $j = 1, 2, \dots, n+1$, choose $A_j \subset T$ such that

$$\|\Phi - \Phi_{A_j}\| = \min_{\substack{A \subset T \\ |A| \leq 2^{j-1}}} \|\Phi - \Phi_A\|.$$

Then, define $B_j = T \setminus A_j$. Picking the sets A_j with smallest cardinality we may assume that $|A_1| \leq |A_2| \leq \dots \leq |A_{n+1}|$ (although these sets may not be nested). Observe that, as special cases we can let

$$(A_{n+1}, B_{n+1}) = (T, \emptyset) \quad \text{and} \quad (A_0, B_0) = (\emptyset, T).$$

This construction implies the following

$$(3.10) \quad \textbf{Key Fact:} \quad \text{If } T' \subset T \text{ and } \|\Phi - \Phi_{T'}\| < \|\Phi_{B_j}\| \text{ then } |T'| > 2^{j-1}.$$

This will be a crucial argument later to conclude the proof of the theorem.

Let $\beta > 0$ be a large number to be determined later, and define

$$\eta = e^{-\beta/2} \quad \text{and} \quad b = \frac{1}{2\eta}.$$

For the moment assume that β is large enough so that $\eta < 1/2$. With that choice of β we have $b > 1$. Now, pick the first positive integer $L = L(b, \Phi) \in \mathbb{N}$ such that

$$(3.11) \quad \|\Phi_{B_{j-1}}\| < b \|\Phi_{B_j}\|, \quad j = 1, 2, \dots, L-1, \quad \text{and} \quad \|\Phi_{B_{L-1}}\| \geq b \|\Phi_{B_L}\|.$$

Note that we could have $L = 1$ if the first condition never holds, i.e. whenever $\|\Phi\| = \|\Phi_{B_0}\| \geq b \|\Phi_{B_1}\|$. At the other extreme, we always have

$$\|\Phi_{B_n}\| \geq b \|\Phi_{B_{n+1}}\| = 0,$$

which implies that $1 \leq L \leq n+1$. Thus,

$$2^{L-2} \leq 2^{n-1} < |T|,$$

which is the first assertion in (3.9). Observe also that $A_L \neq \emptyset$, since otherwise we would have $A_j = \emptyset$, for all $j \leq L$, and hence $\Phi_{B_L} = \Phi_{B_{L-1}} = \Phi$, which would contradict the right hand side of (3.11).

We now apply iteratively Theorem 3.2. Consider the numbers $m_0 = 0$ and

$$m_j = m_{j-1} + \lfloor \beta G(|A_j|) \rfloor, \quad j = 1, \dots, L.$$

Actually, to avoid trivial cases, we should restrict to $j = j_0, \dots, L$, where j_0 is the first integer such that $|A_{j_0}| \neq 0$ (and let $m_j = 0$ for $j < j_0$). We also assume that β is large enough so that

$$(3.12) \quad \beta G(1) \geq 1.$$

Observe that

$$m_L = \sum_{j=j_0}^L \lfloor \beta G(|A_j|) \rfloor \leq \beta \sum_{j=1}^L G(2^{j-1}),$$

since G is increasing. This is the second inequality in (3.9).

Now, for each $j = j_0, \dots, L$ we apply Theorem 3.2 with $k = 0$, $m = m_{j-1}$, $M = \lfloor \beta G(|A_j|) \rfloor$ and $A = A_j$ to obtain

$$\begin{aligned} \|x_{m_j}\| &\leq e^{-\frac{\lfloor \beta G(|A_j|) \rfloor}{G(|A_j|)}} \|x_{m_{j-1}}\| + \lambda \left(\|\Phi - x\| + \|\Phi_{B_j}\| \right) \\ &\leq \eta \|x_{m_{j-1}}\| + \lambda \left(\|\Phi - x\| + \|\Phi_{B_j}\| \right), \end{aligned}$$

using in the last line that $\lfloor a \rfloor \geq a/2$ if $a \geq 1$. Observe that the above inequalities hold trivially for $1 \leq j < j_0$ (if there any such j) since

$$\|x_{m_j}\| = \|x_0\| = \|x\| \leq \|x - \Phi\| + \|\Phi\| = \|x - \Phi\| + \|\Phi_{B_j}\|,$$

and in this case $B_j = T$. Therefore, we can iterate the inequalities to obtain

$$\|x_{m_L}\| \leq \eta^L \|x_{m_0}\| + \lambda \sum_{j=1}^L \eta^{L-j} \left(\|\Phi - x\| + \|\Phi_{B_j}\| \right).$$

For the first summand we can also use

$$(3.13) \quad \|x_{m_0}\| = \|x\| \leq \|x - \Phi\| + \|\Phi\| = \|x - \Phi\| + \|\Phi_{B_0}\|.$$

We now use the crucial assumption (3.11), that is,

$$\|\Phi_{B_j}\| \leq b^{L-1-j} \|\Phi_{B_{L-1}}\|, \quad j = 0, \dots, L-1, \quad \text{and} \quad \|\Phi_{B_L}\| \leq b^{-1} \|\Phi_{B_{L-1}}\|,$$

which inserted in the above expression gives

$$\begin{aligned} \|x_{m_L}\| &\leq \lambda \sum_{j=0}^L \eta^{L-j} \|\Phi - x\| + \lambda b^{-1} \sum_{j=0}^L (\eta b)^{L-j} \|\Phi_{B_{L-1}}\| \\ &\leq \frac{\lambda}{1-\eta} \|\Phi - x\| + \frac{\lambda b^{-1}}{1-\eta b} \|\Phi_{B_{L-1}}\| \\ (3.14) \quad &= \frac{\lambda}{1-\eta} \|\Phi - x\| + 4\lambda \eta \|\Phi_{B_{L-1}}\|, \end{aligned}$$

using in the last step the choice of $b = 1/(2\eta)$.

On the other hand, we can use that Σ_N satisfies property **A2**(k_N, D) to obtain the following estimate

$$\begin{aligned} \|\Phi - \Phi_{T \cap \Gamma_{m_L}}\| &\leq k_N \|\Phi - \mathcal{G}_{m_L}(x)\| \leq k_N (\|\Phi - x\| + \|x_{m_L}\|) \\ (3.15) \quad &\leq k_N \left[\left(1 + \frac{\lambda}{1-\eta}\right) \|\Phi - x\| + 4\lambda \eta \|\Phi_{B_{L-1}}\| \right]. \end{aligned}$$

At this point we wish to use the Key Fact in (3.10). So we distinguish two cases.

Case 1: $\|\Phi - x\| < A \|\Phi_{B_{L-1}}\|$, for some $A > 0$ to be determined. Then

$$\|\Phi - \Phi_{T \cap \Gamma_{m_L}}\| < k_N \left[\left(1 + \frac{\lambda}{1-\eta}\right) A + 4\lambda\eta \right] \|\Phi_{B_{L-1}}\|.$$

In this case, we wish to select A and η (and hence β) so that

$$(3.16) \quad k_N \left[\left(1 + \frac{\lambda}{1-\eta}\right) A + 4\lambda\eta \right] \leq 1,$$

which by the Key Fact would imply that

$$|T \cap \Gamma_{m_L}| > 2^{L-2}.$$

Case 2: $\|\Phi_{B_{L-1}}\| \leq A^{-1} \|\Phi - x\|$. In this case, using (3.14) we have

$$\|x_{m_L}\| \leq \lambda \left(\frac{1}{1-\eta} + 4\eta A^{-1} \right) \|\Phi - x\|.$$

So, we wish to select A and η (hence β) such that

$$(3.17) \quad \frac{1}{1-\eta} + 4\eta A^{-1} \leq 1 + \delta.$$

Overall, we have reduced the theorem to find numbers A and η so that (3.16) and (3.17) hold. Writing $A = \eta B$, this amounts to find B and η so that

$$\frac{1}{1-\eta} + \frac{4}{B} \leq 1 + \delta \quad \text{and} \quad k_N \eta \left[\left(1 + \frac{\lambda}{1-\eta}\right) B + 4\lambda \right] \leq 1.$$

This is clearly possible if B is chosen sufficiently large and η sufficiently small. In order to make an explicit choice, we let $B = 8/\delta$, so we need to select η so that

$$\frac{1}{1-\eta} \leq 1 + \delta/2 \quad \text{and} \quad k_N \eta \left[\left(1 + \frac{\lambda}{1-\eta}\right) 8\delta^{-1} + 4\lambda \right] \leq 1$$

If we impose the first condition, the second one will hold provided

$$k_N \eta \left[\left(1 + (1 + \delta/2)\lambda\right) 8\delta^{-1} + 4\lambda \right] = 8k_N \eta \left[\lambda + \frac{1+\lambda}{\delta} \right] \leq 1.$$

That is, we can choose

$$\eta \leq \min \left\{ (8k_N [\lambda + \frac{1+\lambda}{\delta}])^{-1}, \frac{\delta}{2 + \delta} \right\} = (8k_N [\lambda + \frac{1+\lambda}{\delta}])^{-1},$$

with the last equality following easily from $k_N \geq 1$ and $\lambda \geq 1$. So, simplifying a bit we can choose

$$\eta = \frac{\delta}{8k_N(1 + (1 + \delta)\lambda)},$$

and using that $\eta = e^{-\beta/2}$, we find the expression

$$\beta = 2 \ln(1/\eta) = 2 \ln \left[\frac{8k_N(1 + (1 + \delta)\lambda)}{\delta} \right].$$

We finally observe that (3.12) is also satisfied, as in fact we have $G(1) \geq 1$. This is a simple consequence of

$$\frac{1}{G(1)} = Q\left(\frac{(1 - \frac{1}{\lambda})\tau}{2H(1)}\right) \leq Q\left(\frac{1}{H(1)}\right) \leq Q(\|\varphi\|) = Q(|F_\varphi(\varphi)|) \leq 1,$$

with the second inequality due to (2.1) (for any $\varphi \in \mathcal{D}$), and the last one due to $D(Q)$. $\square$

REMARK 3.18. In order to ensure that $\|x_{m_L}\| \leq 2\|x - \Phi\|$ we must choose λ and δ so that $(1 + \delta)\lambda = 2$. For instance, $\lambda = \sqrt{2}$ and $\delta = \sqrt{2} - 1$ will give the value

$$\beta = 2 \ln \left[\frac{24k_N}{\sqrt{2} - 1} \right] = 2 \ln (24(1 + \sqrt{2})k_N).$$

3.3. Step 3. The next step is a slight generalization of Theorem 3.8, which corresponds to the special case $k = 0$.

THEOREM 3.19. *Let $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ and $1 \leq N < D$ be such that*

- (i) $(\mathbb{X}, \|\cdot\|, \mathcal{D})$ *satisfies* $D(Q)$
- (ii) Σ_N *satisfies property* $A2(k_N, D)$.
- (iii) Σ_N *satisfies property* $A3(H, D)$.

Given $\lambda > 1$ and $\delta > 0$, let $\beta_0 = \beta_0(\lambda, \delta, k_N) > 0$ be as in Theorem 3.8, and let $\beta \geq \beta_0$. If $x \in \mathbb{X}$ and $\Phi = \sum_{j \in T} a_j \varphi_j \in \Sigma_N$ are not null, and if $k \in \mathbb{N}_0$ is such that

$$T_k := T \setminus \Gamma_k \neq \emptyset, \quad (\text{where } \Gamma_k = \text{supp } \mathcal{G}_k(x)),$$

then there exist integers $L \in \mathbb{N}$ and $m_L \geq k + 1$ such that

$$(3.20) \quad 2^{L-2} < |T_k| \quad \text{and} \quad m_L - k \leq \beta \sum_{j=1}^L G(2^{j-1}),$$

and so that

$$\text{either } \|x_{m_L}\| \leq (1 + \delta)\lambda \|x - \Phi\|, \quad \text{or } |T_k \cap \Gamma_{m_L}| > 2^{L-2},$$

provided that $N + m_L < D$.

Proof. Apply the construction in the first part of Theorem 3.8 to the vector Φ_{T_k} (instead of Φ). So for η and b fixed as above, this gives an integer $L \in \mathbb{N}$ such that $2^{L-2} < |T_k|$ and sets $A_j \subset T_k$ and $B_j = T_k \setminus A_j$ such that the inequalities in (3.11) hold.

At this point we let $m_0 = k$ and consider

$$m_j = m_{j-1} + \lfloor \beta G(|A_j|) \rfloor, \quad j = j_0, \dots, L,$$

where j_0 is the first integer such that $|A_{j_0}| \neq 0$. Otherwise we let $m_j = m_0 = k$ when $1 \leq j < j_0$. As before, this choice (and the size of the sets A_j) gives the second assertion in (3.20).

Now, if $j_0 \leq j \leq L$ we apply Theorem 3.2 with $m = m_{j-1}$, $M = \lfloor \beta G(|A_j|) \rfloor$ and $A = A_j$ to obtain

$$\begin{aligned} \|x_{m_j}\| &\leq e^{-\frac{\lfloor \beta G(|A_j|) \rfloor}{G(|A_j|)}} \|x_{m_{j-1}}\| + \lambda \left(\|\Phi - x\| + \|\Phi_{B_j}\| \right) \\ &\leq \eta \|x_{m_{j-1}}\| + \lambda \left(\|\Phi - x\| + \|\Phi_{B_j}\| \right). \end{aligned}$$

When $0 \leq j < j_0$, we have instead

$$\begin{aligned} \|x_{m_j}\| = \|x_k\| = \text{dist}(x, [\varphi_i]_{i \in \Gamma_k}) &\leq \|x - \Phi_{T \cap \Gamma_k}\| = \|x - \Phi + \Phi_{T_k}\| \\ &\leq \|x - \Phi\| + \|\Phi_{B_j}\|, \end{aligned}$$

since $A_j = \emptyset$ and hence $B_j = T_k$. Thus, we can proceed exactly as we did in (3.14) to obtain the same conclusion, namely

$$(3.21) \quad \|x_{m_L}\| \leq \frac{\lambda}{1-\eta} \|\Phi - x\| + 4\lambda\eta \|\Phi_{B_{L-1}}\|.$$

On the other hand, using property **A2**(k_N, D) we obtain

$$\begin{aligned} \|\Phi_{T_k} - \Phi_{T_k \cap \Gamma_{m_L}}\| &\leq k_N \|\Phi_{T_k} - \mathcal{G}_{m_L}(x) + \Phi_{T \cap \Gamma_k}\| = k_N \|\Phi - x + x_{m_L}\| \\ &\leq k_N (\|\Phi - x\| + \|x_{m_L}\|) \\ (3.22) \quad &\leq k_N \left[(1 + \frac{\lambda}{1-\eta}) \|\Phi - x\| + 4\lambda\eta \|\Phi_{B_{L-1}}\| \right]. \end{aligned}$$

which is the analogous inequality to (3.15) in the previous theorem.

At this point one considers the same two cases as in the lines following (3.15). Namely

Case 1: $\|\Phi - x\| < A \|\Phi_{B_{L-1}}\|$, with the same $A > 0$ as in Theorem 3.8. This implies

$$\|\Phi_{T_k} - \Phi_{T_k \cap \Gamma_{m_L}}\| < k_N \left[(1 + \frac{\lambda}{1-\eta}) A + 4\lambda\eta \right] \|\Phi_{B_{L-1}}\| \leq \|\Phi_{B_{L-1}}\|,$$

so by the construction of the sets (A_j, B_j) and the Key Fact one obtains

$$|T_k \cap \Gamma_{m_L}| > 2^{L-2}.$$

Case 2: $\|\Phi_{B_{L-1}}\| \leq A^{-1} \|\Phi - x\|$. In this case, the same reasoning as in Theorem 3.8 gives

$$\|x_{m_L}\| \leq \lambda \left(\frac{1}{1-\eta} + 4\eta A^{-1} \right) \|\Phi - x\| \leq (1 + \delta) \lambda \|\Phi - x\|.$$

This completes the proof of Theorem 3.19. $\square$

3.4. Step 4: Conclusion of the proof of Theorem 1.16. This part of the proof requires substantial modifications compared to [14, 16], so we present it in detail.

Write $\lambda_1 = (1 + \delta)\lambda$, say with

$$\lambda = \sqrt{\lambda_1} > 1 \quad \text{and} \quad \delta = \sqrt{\lambda_1} - 1 > 0.$$

The iterative process discussed in the previous subsections produces a positive constant $\beta = 2 \ln \left[\frac{8k_N(1+\lambda_1)}{\sqrt{\lambda_1}-1} \right]$, and the following sequences of numbers

- there exist positive integers L_1 and m_{L_1} such that

$$(3.23) \quad m_{L_1} \leq \beta \sum_{j=1}^{L_1} G(2^{j-1}) \quad \text{and} \quad 2^{L_1-2} \leq |T|,$$

with the property that

$$\text{either } \|x_{m_{L_1}}\| \leq \lambda_1 \|x - \Phi\| \quad \text{or} \quad |T \cap \Gamma_{m_{L_1}}| \geq 2^{L_1-2}.$$

In the first case one stops; if not one iterates and applies Theorem 3.19 with $k = m_{L_1}$, which implies

- there exist positive integers L_2 and $m_{L_2} > m_{L_1}$ such that

$$(3.24) \quad m_{L_2} - m_{L_1} \leq \beta \sum_{j=1}^{L_2} G(2^{j-1}) \quad \text{and} \quad 2^{L_2-2} \leq |T_{m_{L_1}}|,$$

with the property that

$$\text{either } \|x_{m_{L_2}}\| \leq \lambda_1 \|x - \Phi\| \quad \text{or} \quad |T_{m_{L_1}} \cap \Gamma_{m_{L_2}}| \geq 2^{L_2-2}.$$

Again, in the first case one stops; if not one applies iteratively Theorem 3.19, with values of $k = m_{L_i}$, $i = 2, \dots, s-1$, until some step s , where can one ensure that

- there are positive integers L_s and $m_{L_s} > m_{L_{s-1}}$ such that

$$(3.25) \quad m_{L_s} - m_{L_{s-1}} \leq \beta \sum_{j=1}^{L_s} G(2^{j-1}) \quad \text{and} \quad 2^{L_s-2} \leq |T_{m_{L_{s-1}}}|,$$

where

$$(3.26) \quad \begin{cases} \text{either } \|x_{m_{L_s}}\| \leq \lambda_1 \|x - \Phi\|, \\ \text{or } |T_{m_{L_{s-1}}} \cap \Gamma_{m_{L_s}}| \geq 2^{L_s-2} \quad \text{and} \quad m_{L_s} \geq 2\beta \tilde{G}(2N). \end{cases}$$

Here $G(n)$ denotes the sequence defined in (3.4), and the notation $\tilde{G}(n)$ stands for the associated summing sequence as in (2.9).

In the first case of (3.26) one stops; if not, we shall show that the greedy algorithm actually covers the whole set T , that is

$$(3.27) \quad |T \cap \Gamma_{m_{L_s}}| \geq N.$$

This would imply that $x_{m_{L_s}} = 0$, and so we would also stop.

Let us prove (3.27). Here we shall use the assumption that the sequence $G(n)$ in (3.4) is increasing and 1-quasi-convex. Observe that

$$(3.28) \quad \begin{aligned} |T \cap \Gamma_{m_{L_s}}| &= |T \cap \Gamma_{m_{L_1}}| + |T_{m_{L_1}} \cap \Gamma_{m_{L_2}}| + \dots + |T_{m_{L_{s-1}}} \cap \Gamma_{m_{L_s}}| \\ &\geq 2^{L_1-2} + 2^{L_2-2} + \dots + 2^{L_s-2}. \end{aligned}$$

Now, by Lemma 2.10 and the inductive assumptions, see (3.24), for each $i = 1, \dots, s$, we have

$$(3.29) \quad 2\beta \tilde{G}(2^{L_i-1}) > \beta \sum_{j=0}^{L_i-1} G(2^j) = \beta \sum_{j=1}^{L_i} G(2^{j-1}) \geq m_{L_i} - m_{L_{i-1}},$$

with the notation $m_{L_0} = 0$. Thus, applying the (non-decreasing) function $2\beta\tilde{G}(2\cdot)$ to both sides of (3.28) and using part b) of Lemma 2.11 we obtain

$$\begin{aligned} 2\beta\tilde{G}(2|T \cap \Gamma_{m_{L_s}}|) &\geq 2\beta\tilde{G}(2^{L_1-1} + 2^{L_2-1} + \dots + 2^{L_s-1}) \\ &\geq 2\beta\left(\tilde{G}(2^{L_1-1}) + \tilde{G}(2^{L_2-1}) + \dots + \tilde{G}(2^{L_s-1})\right) \\ &> m_{L_s} \geq 2\beta\tilde{G}(2N), \end{aligned}$$

using in the last line (3.29) and the second assertion in (3.26). Since $\tilde{G}$ is increasing this implies

$$|T \cap \Gamma_{m_{L_s}}| > N,$$

which proves (3.27).

Thus, the process will indeed end after m_{L_s} iterations. We now estimate this number using the remaining conditions in (3.25). Since the last inequality in (3.26) occurs for the first time at step s , we must have

$$m_{L_{s-1}} < 2\beta\tilde{G}(2N).$$

Thus,

$$\begin{aligned} m_{L_s} &= m_{L_{s-1}} + (m_{L_s} - m_{L_{s-1}}) \\ &\leq 2\beta\tilde{G}(2N) + (m_{L_s} - m_{L_{s-1}}) \\ &\leq 2\beta\tilde{G}(2N) + \beta \sum_{j=1}^{L_s} G(2^{j-1}) \\ &\leq 2\beta\tilde{G}(2N) + 2\beta\tilde{G}(2^{L_s-1}) \leq 4\beta\tilde{G}(2N). \end{aligned}$$

Therefore, using also part a) of Lemma 2.11, we see that (1.19) will be true with

$$\phi(N) = m_{L_s} \leq 4\beta G(2N) = 8 \ln \left[\frac{8k_N(1 + \lambda_1)}{\sqrt{\lambda_1} - 1} \right] G(2N),$$

as asserted in (1.18). □

4. AN APPLICATION: WCGA IN $L^p(\log L)^\alpha$ SPACES

4.1. Property $\mathcal{D}(Q)$ in $L^p(\log L)^\alpha$. In this section we shall apply Theorem 1.16 in the case when

$$\mathbb{X} = L^p(\log L)^\alpha, \quad 1 < p < \infty, \alpha \in \mathbb{R}.$$

Following [1, Definition IV.6.11], this is the set of all measurable $f : \mathbb{R}^d \rightarrow \mathbb{C}$ such that

$$\left\| |f(x)| \log(2 + |f(x)|) \right\|_{L^p(\mathbb{R}^d)}^\alpha < \infty.$$

These classes satisfy the elementary inclusions

$$L^p(\log L)^{|\alpha|} \subset L^p \subset L^p(\log L)^{-|\alpha|}.$$

We shall regard $\mathbb{X}$ as an Orlicz space L^Φ associated with the function

$$(4.1) \quad \Phi(t) = \int_0^t s^{p-1} (\log(c+s))^{\alpha p}, \quad t \geq 0,$$

which for a sufficiently large $c > 1$ is a (smooth) Young function*. The corresponding (Luxemburg) norm is then defined by

$$\|f\|_{L^\Phi} = \inf \left\{ \lambda : \int_{\mathbb{R}^d} \Phi(|f(x)|/\lambda) dx \leq 1 \right\}.$$

Let Ψ be the complementary function[†] of $\Phi(t)$. Then it is known that $(L^\Phi)^* = L^\Psi$ (isometrically, when the latter space is endowed with the Orlicz norm); see [1, Corollary IV.8.15]. In these examples it is not difficult to check that

$$(4.2) \quad \Phi(t) \approx t^p (\log(e+t))^{\alpha p}, \quad t \geq 0,$$

and

$$(4.3) \quad \Psi(t) \approx t^{p'} (\log(e+t))^{-\alpha p'}, \quad t \geq 0;$$

see e.g. [5, Theorem I.7.2].

We recall how the norming functional F_f of a (normalized) element $f \in L^\Phi$ is defined; see [5, Theorem 18.5]. Let $P(z)$, $z \in \mathbb{C}$, be an extension of $P(t) = \Phi'(t)$, $t \geq 0$, such that $\text{sign}(zP(z)) = 1$, $\forall z \neq 0$. Then F_f is explicitly given by

$$F_f(g) = \frac{\int_{\mathbb{R}^d} P(f(x)) g(x) dx}{\int_{\mathbb{R}^d} P(f(x)) f(x) dx}.$$

In our case of interest we will have $P(z) = |z|^{p-2} \bar{z} (\log(c+|z|))^{\alpha p}$, $z \in \mathbb{C}$, and hence

$$(4.4) \quad F_f(g) = \frac{1}{A(f)} \int_{\mathbb{R}^d} |f(x)|^{p-2} \overline{f(x)} g(x) (\log(c+|f(x)|))^{\alpha p} dx,$$

with $A(f) = \int_{\mathbb{R}^d} |f|^p (\log(c+|f|))^{\alpha p} dx$.

The moduli of smoothness and convexity for Orlicz spaces L^Φ have been studied in [8, 3]. According to [8, Theorem 1], there exists a Young function $\bar{\Phi}$, equivalent to Φ , such that

$$(4.5) \quad \rho_{L^{\bar{\Phi}}}(t) \lesssim \sup_{u \in [s,1], v > 0} \frac{t^2 \Phi(uv)}{u^2 \Phi(v)} \quad \text{and} \quad \delta_{(L^{\bar{\Phi}})^*}(s) \gtrsim \inf_{u \in [s,1], v > 0} \frac{s^2 \Psi(uv)}{u^2 \Psi(v)}$$

*A Young function is a convex, non-decreasing function $\Phi : [0, \infty) \rightarrow [0, \infty]$ such that $\Phi(0^+) = 0$ and $\lim_{t \rightarrow +\infty} \Phi(t) = +\infty$. See [10, §I.3] or [1, §IV.8] for properties and various equivalent definitions.

[†]The complementary function of a Young function $\Phi(t)$ is defined by $\Psi(s) = \sup_{t \geq 0} (ts - \Phi(t))$. See [10, §I.3] or [1, §IV.8] for properties and other equivalent definitions.

under suitable doubling conditions in $\Phi(t)$ and $\Psi(t)$ (which are always held in the cases considered in (4.2) and (4.3)). Moreover, our specific examples satisfy the regularity conditions stated in [3, Proposition 19], so one may actually take $\bar{\Phi} = \Phi$.

Therefore, inserting into (4.5) the expressions for Φ and Ψ from (4.2) and (4.3), and performing some straightforward computations, one obtains the following result. Here we use the standard notation $\alpha = \alpha_+ - \alpha_-$, where

$$\alpha_+ = \max\{\alpha, 0\} \quad \text{and} \quad \alpha_- = \max\{-\alpha, 0\}.$$

PROPOSITION 4.6. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$, and let $\mathbb{X} = L^p(\log L)^\alpha$ as above. Then, for the Luxemburg norm associated with $\Phi(t)$ in $\mathbb{X}$ it holds*

- if $p > 2$ then

$$\rho_{\mathbb{X}}(t) \lesssim t^2 \quad \text{and} \quad \delta_{\mathbb{X}^*}(s) \gtrsim s^2,$$

- if $1 < p \leq 2$ then

$$\rho_{\mathbb{X}}(t) \lesssim t^p (\log(e + \frac{1}{t}))^{p\alpha_-} \quad \text{and} \quad \delta_{\mathbb{X}^*}(s) \gtrsim \frac{s^{p'}}{(\log(e + \frac{1}{s}))^{p'\alpha_-}}.$$

In view of the discussion in §2.2, we then obtain the following.

COROLLARY 4.7. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$, and let $\mathbb{X} = L^p(\log L)^\alpha$ as above. Then, for every normalized dictionary $\mathcal{D}$ in $\mathbb{X}$, property $D(Q)$ holds with*

$$(4.8) \quad Q(s) := \begin{cases} \frac{c_0 s^{p'}}{(\log(e + \frac{1}{s}))^{p'\alpha_-}} & \text{if } 1 < p \leq 2 \\ c_0 s^2 & \text{if } p > 2, \end{cases}$$

for a suitably small constant $c_0 > 0$.

4.2. The Haar system in $L^p(\log L)^\alpha$. Next we consider the dictionary $\mathcal{D} = \{\psi_j\}_{j=1}^\infty$ in $\mathbb{X} = L^p(\log L)^\alpha$ given by the normalized Haar basis in $\mathbb{R}^d$ (or any sufficiently smooth wavelet basis). These bases are unconditional, so in this case property $A2(k_N, D)$ will hold with $k_N = O(1)$ (and any $D < \infty$). It remains to verify property $A3(H, D)$. As mentioned in Lemma 2.7, this property holds with

$$(4.9) \quad H(N) = \sup_{|A| \leq N, |\varepsilon_j|=1} \left\| \sum_{j \in A} \varepsilon_j \psi_j^* \right\|_{\mathbb{X}^*}$$

(also for all $D < \infty$). Here $\{\psi_j^*\}$ is the dual dictionary, which is again the Haar basis, this time normalized in $\mathbb{X}^*$. Since the basis is unconditional, the parameter $H(N)$ is equivalent to the *upper democracy function* of the dual space $\mathbb{X}^*$, that is

$$H(N) \approx h_{\mathbb{X}^*}(N) := \sup_{|A| \leq N} \left\| \sum_{j \in A} \psi_j^* \right\|_{\mathbb{X}^*}.$$

Democracy functions, for the Orlicz classes L^Φ , were studied in [4], where it was proved that, if the Boyd indices of Φ are non trivial, then

$$h_{L^\Phi}(N) \approx \sup_{s>0} \frac{\varphi(Ns)}{\varphi(s)},$$

where $\varphi(t) := 1/\Phi^{-1}(1/t)$ is the fundamental function of L^Φ .

In our case of interest, where $\Phi(t)$ satisfies (4.2), we have

$$\varphi(t) \approx t^{1/p} \left(\log(e + \frac{1}{t}) \right)^\alpha \quad \text{and} \quad h_{L^\Phi}(N) \approx N^{1/p} \left(\log(e + N) \right)^{\alpha-};$$

see [4, Proposition 3.4]. Thus, for the dual space $\mathbb{X}^* = L^\Psi$ with Ψ as in (4.3) we have

$$h_{L^\Psi}(N) \approx N^{1/p'} \left(\log(e + N) \right)^{\alpha+}.$$

Overall we conclude that property **A3**(H) holds with

$$(4.10) \quad H(N) \approx N^{1/p'} \left(\log(e + N) \right)^{\alpha+}.$$

Thus, combining (4.8) and (4.10), we see that, for $p > 2$ we have

$$G(N) = \frac{1}{Q(\frac{c(\tau)}{H(N)})} \approx H(N)^2 \approx N^{\frac{2}{p'}} \left(\log(e + N) \right)^{2\alpha+}$$

while for $1 < p \leq 2$ we have

$$G(N) = \frac{1}{Q(\frac{c(\tau)}{H(N)})} \approx H(N)^{p'} \left(\log(e + H(N)) \right)^{p'\alpha-} \approx N \left(\log(e + N) \right)^{p'|\alpha|}.$$

4.3. Proof of Theorem 1.21.a. Collecting the values of the parameters $G(N)$ obtained in the previous subsection, and inserting them into Theorem 1.16, we deduce the first assertions (1.22) and (1.23) in Theorem 1.21.

4.4. Proof of Theorem 1.21.b. We shall use the following result whose proof can be found in [4, Lemma 3.1]. For simplicity in the notation, we assume in this section that the underlying space $\mathbb{R}^d$ has dimension $d = 1$.

LEMMA 4.11. *Let $\mathbb{X} = L^\Phi(\mathbb{R})$ be an Orlicz space with non-trivial Boyd indices, and let $\mathcal{D} = \{h_I\}$ be the (normalized) Haar basis in $\mathbb{X}$. Then, if A is a finite collection of disjoint dyadic intervals with the same size s , then*

$$\left\| \sum_{I \in A} h_I \right\|_{L^\Phi} \approx \frac{\varphi(|A|s)}{\varphi(s)},$$

where $\varphi(t)$ is the fundamental function of $\mathbb{X}$.

We now show the lower bound for the function $\psi(N)$ stated in (1.24).

Proof of (1.24). Write $\mathcal{D} = \{h_I\}$ where h_I is the (normalized) Haar function supported in I , and I runs over all dyadic intervals in $\mathbb{R}$. Pick any two collections A and B , of pairwise disjoint dyadic intervals with cardinalities $|A| = N$ and $|B| = M$, such that

$$|I| = 1 \quad \text{if } I \in A, \quad \text{and} \quad |I| = 1/M \quad \text{if } I \in B.$$

For instance, we could take

$$A = \{[n, n+1) : n = 1, \dots, N\} \quad \text{and} \quad B = \{J \subset [0, 1) : |J| = 2^{-m}\},$$

with $M = 2^m$. For $b > 0$ to be determined, consider the function

$$(4.12) \quad f = f_1 + f_2 = \sum_{I \in A} h_I + b \sum_{I \in B} h_I.$$

Using Lemma 4.11 and $\varphi(t) \approx t^{1/p} (\log(e + \frac{1}{t}))^\alpha$, observe that

$$\|f_1\| \approx \varphi(N) \approx N^{1/p}, \quad \text{and} \quad \|f_2\| \approx \frac{b}{\varphi(1/M)} \approx \frac{b M^{\frac{1}{p}}}{(\log(e + M))^\alpha}.$$

Also, since $\|h_I\|_{L^\Phi} = 1$ we have

$$(4.13) \quad |h_I(x)| = \frac{1}{\varphi(|I|)} \mathbf{1}_I(x).$$

In particular, if $I \in B$ we have

$$(4.14) \quad |f(x)| = b |h_I(x)| = \frac{b}{\varphi(1/M)} \approx \|f_2\| \lesssim \|f\|, \quad x \in I,$$

and similarly, if $I \in A$ we have

$$(4.15) \quad |f(x)| = 1 \lesssim \|f_1\| \lesssim \|f\|, \quad x \in I.$$

Using the formula for the norming functional in (4.4) we see that

$$|F_f(h_I)| = \frac{1}{C(f)} \times \begin{cases} \int |h_I|^p \left(\log(c + |f(x)|/\|f\|) \right)^{\alpha p} dx, & I \in A, \\ b^{p-1} \int |h_I|^p \left(\log(c + |f(x)|/\|f\|) \right)^{\alpha p} dx, & I \in B, \\ 0, & I \notin A \cup B. \end{cases}$$

In view of (4.14) and (4.15), the logarithmic factors inside the integrals are approximately constant, so can be disregarded. Also, (4.13) implies

$$\int |h_I|^p dx = \frac{|I|}{\varphi(|I|)^p} \approx \frac{1}{(\log(e + |I|^{-1}))^{\alpha p}},$$

so we have

$$(4.16) \quad |F_f(h_I)| \approx \frac{1}{C(f)}, \quad I \in A, \quad \text{and} \quad |F_f(h_I)| \approx \frac{1}{C(f)} \frac{b^{p-1}}{(\log(e + M))^{\alpha p}}, \quad I \in B.$$

Thus, the above quantities are approximately the same provided we choose

$$(4.17) \quad b = c_1 (\log(e + M))^{\alpha p'}.$$

Therefore, if $c_1 > 0$ is chosen properly, the WCGA, $\mathcal{G}_n(f)$, can be formed either by selecting consecutive elements I from A (if $n \leq N$), or by selecting consecutive elements I from B (if $n \leq M/2$). To verify these assertions one should note that the equivalences in (4.16) remain also true* when f is replaced by the remainder $f - \mathcal{G}_n(f)$.

*In the latter case, we have restricted to $n \leq M/2$ to ensure that (4.14) continues to hold when f is replaced by $f - \mathcal{G}_n(f)$. Indeed, in such case one would use that $\|\sum_{I \in B'} h_I\| \approx 1/\varphi(1/M)$, when $B' \subset B$ with $|B'| \geq M/2$, by Lemma 4.11.

So suppose now that (1.24) holds. If $\alpha \geq 0$, we let $N = \psi(M)$, and in view of the previous comment we can select c_1 such that $\mathcal{G}_N(f) = f_1$. Then

$$\|f_2\| = \|f - \mathcal{G}_{\psi(M)}(f)\| \leq 2\sigma_M(f) \leq \|f_1\|,$$

which in view of (4.17) and (4.12) implies

$$c_1^p M (\log(e + M))^{\alpha p'} = \frac{b^p M}{(\log(e + M))^{\alpha p}} \approx \|f_2\|^p \lesssim \|f_1\|^p \approx N = \psi(M).$$

This proves the assertion in the Theorem when $\alpha \geq 0$.

If $\alpha \leq 0$, then we take $M = 2\psi(N)$, and select c_1 such that $\mathcal{G}_{M/2}(f) = b \sum_{I \in B'} h_I$, for some $B' \subset B$ with $|B'| = M/2$. Then

$$\|f_1\| \lesssim \|f_1 + b \sum_{I \in B \setminus B'} h_I\| = \|f - \mathcal{G}_{\psi(N)}(f)\| \leq 2\sigma_N(f) \leq \|f_2\|,$$

which this time implies

$$N \approx \|f_1\|^p \lesssim \|f_2\|^p \approx M (\log(e + M))^{\alpha p'}.$$

Solving for M this gives

$$\psi(N) = M/2 \gtrsim \frac{N}{(\log(e + N))^{\alpha p'}} = N (\log(e + N))^{|\alpha| p'}.$$

This establishes (1.24), and therefore completes the proof of Theorem 1.21. $\square$

5. WCGA FOR TRIGONOMETRIC SYSTEM IN $L^p(\log L)^\alpha$

In this section we give a second application of Theorem 1.16, this time to the trigonometric system in the torus $\mathbb{T} \equiv [-\pi, \pi)$, that is,

$$\mathcal{D} = \mathcal{T} := \{e^{inx}\}_{n \in \mathbb{Z}}.$$

So, from now on, all functions $f \in L^p(\log L)^\alpha$ are understood as defined in $\mathbb{T}$. Otherwise, we regard $L^p(\log L)^\alpha$ as an Orlicz space L^Φ in the same sense as in §4. Since [8] covers also this setting, the estimates for the moduli of convexity and smoothness in Proposition 4.6 remain true, and so does the estimate (4.8) for the function $Q(s)$ in Corollary 4.7.

We still have to compute the parameters k_N and $H(N)$. To do so, we shall make use of the following interpolation lemma.

LEMMA 5.1. *Consider the Young function $\bar{\Phi}(t) = t^p (\log(c + t))^\alpha$, for some $c \geq e$. Assume that*

$$(5.2) \quad 2 < p < \infty \text{ and } \alpha \in \mathbb{R}, \quad \text{or} \quad p = 2 \text{ and } \alpha \geq 0.$$

Then,

$$(5.3) \quad \|f\|_{L^{\bar{\Phi}}} \leq \|f\|_\infty^{1-\frac{2}{p}} \left(\log \left(c + \left[\frac{\|f\|_\infty}{\|f\|_2} \right]^{\frac{p}{2}} \right) \right)^\alpha \|f\|_2^{\frac{2}{p}}, \quad \forall f \in L^\infty(\mathbb{T}).$$

Proof. We may assume that $\|f\|_2 = 1$. Define the functions

$$(5.4) \quad a(t) = t^{\frac{2}{p}} \left(\log(c + t^{\frac{2}{p}}) \right)^{-\alpha} \quad \text{and} \quad b(t) = \frac{t}{a(t)} = t^{1-\frac{2}{p}} \left(\log(c + t^{\frac{2}{p}}) \right)^\alpha, \quad t > 0.$$

By the lattice property of the Luxemburg norm in $L^{\bar{\Phi}}$ we have

$$\|f\|_{L^{\bar{\Phi}}} = \|a(f)b(f)\|_{L^{\bar{\Phi}}} \leq \|b(f)\|_{L^\infty} \|a(f)\|_{L^{\bar{\Phi}}} = b(\|f\|_\infty) \|a(f)\|_{L^{\bar{\Phi}}},$$

using that $b(t)$ is increasing under the conditions in (5.2). So, it suffices to show that

$$\int_{\mathbb{T}} \bar{\Phi}(a(|f(x)|)) dx \leq 1,$$

as this will imply that $\|a(f)\|_{L^{\bar{\Phi}}} \leq 1$. Write

$$\int \bar{\Phi}(a(|f|)) dx = \int a(|f|)^p \left[\log(c + a(|f|)) \right]^{\alpha p} dx.$$

Observe that, regardless of the sign of $\alpha \in \mathbb{R}$, we always have

$$\left[\log(c + a(|f|)) \right]^{\alpha p} = \left[\log \left(c + \frac{|f|^{\frac{2}{p}}}{[\log(c + |f|^{2/p})]^\alpha} \right) \right]^{\alpha p} \leq \left[\log \left(c + |f|^{\frac{2}{p}} \right) \right]^{\alpha p}.$$

Thus,

$$\int \bar{\Phi}(a(|f|)) dx \leq \int a(|f|)^p \left[\log(c + |f|^{\frac{2}{p}}) \right]^{\alpha p} dx = \int |f|^2 = 1.$$

□

REMARK 5.5. Observe that, when the indices p and α satisfy (5.2), then it holds

$$(5.6) \quad L^p(\log L)^\alpha \hookrightarrow L^2(\mathbb{T}).$$

This is easily proved using that $t^2 \lesssim \Phi(t)$ for $t \geq 1$, since

$$\int_{\mathbb{T}} |f|^2 = \int_{\{|f|<1\}} |f|^2 + \int_{\{|f|\geq 1\}} |f|^2 \leq 1 + c' \int \Phi(|f|) dx < \infty.$$

Likewise, by duality, one proves that $L^2(\mathbb{T}) \hookrightarrow L^p(\log L)^\alpha$ when

$$(5.7) \quad 1 < p < 2 \text{ and } \alpha \in \mathbb{R}, \quad \text{or} \quad p = 2 \text{ and } \alpha \leq 0.$$

5.1. Property A3 for $\mathcal{T}$ in $L^p(\log L)^\alpha$.

LEMMA 5.8. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$. Then, for all $|\varepsilon_n| \leq 1$ and all $A \subset \mathbb{Z}$ with $|A| \leq N$ it holds*

$$(5.9) \quad \left\| \sum_{n \in A} \varepsilon_n e^{inx} \right\|_{L^p(\log L)^\alpha} \lesssim \max \left\{ N^{1/2}, N^{1-\frac{1}{p}} (\log(e+N))^\alpha \right\}.$$

Proof. When p and α satisfy (5.7), the right hand side of (5.9) is $\approx N^{1/2}$, so the assertion follows from the inclusion $L^2 \hookrightarrow L^p(\log L)^\alpha$. On the other hand, if p and α satisfy (5.2), then by Lemma 5.1 we have

$$(5.10) \quad \|f\|_{L^p(\log L)^\alpha} \lesssim b\left(\frac{\|f\|_\infty}{\|f\|_2}\right) \|f\|_2,$$

where $b(t) = t^{1-\frac{2}{p}} (\log(c+t^{\frac{2}{p}}))^\alpha$. Applying this to $f = \sum_{n \in A} \varepsilon_n e^{inx}$, and using that $b(t)$ is increasing and

$$\|f\|_\infty / \|f\|_2 \leq N / \sqrt{N} = \sqrt{N},$$

one easily obtains (5.9). $\square$

REMARK 5.11. The upper bounds in (5.9) cannot be improved, even when all signs $\varepsilon_n = 1$. Indeed, if one considers the Dirichlet kernel $D_N(x) = \sum_{|n| \leq N} e^{inx}$, then we have

$$(5.12) \quad \|D_N\|_{L^p(\log L)^\alpha} \approx N^{1-\frac{1}{p}} (\log(e+N))^\alpha;$$

see e.g. [9, Lemma 3.1]. On the other hand, if A is a lacunary set (say, $A = \{2^j\}_{j=1}^N$), then

$$\left\| \sum_{n \in A} e^{inx} \right\|_{L^p(\log L)^\alpha} \approx \sqrt{N}.$$

Indeed, this is easily obtained from a similar result for all the L^q spaces, $0 < q < \infty$, and the inclusions $L^{p+\varepsilon} \hookrightarrow L^p(\log L)^\alpha \hookrightarrow L^{p-\varepsilon}$.

COROLLARY 5.13. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$. Let $\mathbb{X} = L^p(\log L)^\alpha$ and $\mathcal{D} = \{e^{inx}\}_{n \in \mathbb{Z}}$ in $\mathbb{T}$. Then, property A3(H) holds with*

$$H(N) \approx \max \left\{ N^{1/2}, N^{\frac{1}{p}} (\log(e+N))^{-\alpha} \right\}.$$

Proof. Apply Lemmas 2.7 and 5.8, and the duality relation $\mathbb{X}^* = L^{p'}(\log L)^{-\alpha}$. $\square$

5.2. Property A2 for $\mathcal{T}$ in $L^p(\log L)^\alpha$.

Given a finite set $A \subset \mathbb{Z}$, we denote

$$S_A(g) = \sum_{n \in A} \hat{g}(n) e^{inx},$$

where $\hat{g}(n)$, $n \in \mathbb{Z}$, are the Fourier coefficients of $g \in L^1(\mathbb{T})$. As noticed in [2, Lemma 2.15], property A2(k_N) holds trivially when we let

$$(5.14) \quad k_N = \sup_{|A| \leq N} \|S_A\|_{L^p(\log L)^\alpha \rightarrow L^p(\log L)^\alpha}.$$

In this section we compute this last expression.

LEMMA 5.15. *Let $2 < p < \infty$ and $\alpha \in \mathbb{R}$, or $p = 2$ and $\alpha \geq 0$. Then,*

$$(5.16) \quad \sup_{|A| \leq N} \|S_A\|_{L^p(\log L)^\alpha \rightarrow L^p(\log L)^\alpha} \lesssim N^{\frac{1}{2} - \frac{1}{p}} (\log(e + N))^\alpha.$$

Proof. Let $g \in L^p(\log L)^\alpha$. Using the inequality in (5.10) from the previous section, applied to $f = S_A(g)$ we see that

$$\|S_A(g)\|_{L^p(\log L)^\alpha} \lesssim b\left(\frac{\|S_A(g)\|_\infty}{\|S_A(g)\|_2}\right) \|S_A(g)\|_2.$$

Now,

$$\|S_A(g)\|_\infty \leq |A|^{1/2} \left(\sum_{n \in A} |\hat{g}(n)|^2 \right)^{1/2} \leq \sqrt{N} \|S_A(g)\|_2,$$

so using that $b(t)$ is increasing we obtain

$$\|S_A(g)\|_{L^p(\log L)^\alpha} \lesssim b(\sqrt{N}) \|S_A(g)\|_2.$$

On the other hand, the inclusion in (5.6) gives

$$\|S_A(g)\|_2 \leq \|g\|_2 \lesssim \|g\|_{L^p(\log L)^\alpha}.$$

Thus, we obtain

$$\|S_A\| \lesssim b(\sqrt{N}) \approx N^{\frac{1}{2} - \frac{1}{p}} (\log(e + N))^\alpha.$$

□

Since $S_A^* = S_A$, by duality one obtains the following complementary result.

LEMMA 5.17. *Let $1 < p < 2$ and $\alpha \in \mathbb{R}$, or $p = 2$ and $\alpha \leq 0$. Then,*

$$(5.18) \quad \sup_{|A| \leq N} \|S_A\|_{L^p(\log L)^\alpha \rightarrow L^p(\log L)^\alpha} \lesssim N^{\frac{1}{p} - \frac{1}{2}} (\log(e + N))^{-\alpha}.$$

REMARK 5.19. The estimate in (5.18) is best possible (and by duality, also (5.16)).

One can prove this by noticing that there exist choices of signs ± 1 such that

$$(5.20) \quad \left\| \sum_{|n| \leq N} \pm e^{inx} \right\|_{L^p(\log L)^\alpha} \gtrsim \sqrt{N}.$$

This last assertion can be easily obtained from a similar property of the L^q -spaces, and the inclusions at the end of Remark 5.11. From (5.20), there will be a set $A \subset [-N, N]$, either corresponding to the positive or the negative signs, so that

$$\left\| \sum_{n \in A} e^{inx} \right\|_{L^p(\log L)^\alpha} \gtrsim \frac{1}{2} \sqrt{N}.$$

Thus, omitting the subindices $L^p(\log L)^\alpha$ from the norms, we have

$$\|S_A\| \geq \|S_A(D_N)\| / \|D_N\| = \left\| \sum_{n \in A} e^{inx} \right\| / \|D_N\| \gtrsim N^{\frac{1}{p} - \frac{1}{2}} (\log(e + N))^{-\alpha},$$

using (5.12) in the last step.

COROLLARY 5.21. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$. Let $\mathbb{X} = L^p(\log L)^\alpha$ and $\mathcal{D} = \{e^{inx}\}_{n \in \mathbb{Z}}$ in $\mathbb{T}$. Then, property **A2**(k_N) holds with*

$$k_N \approx \begin{cases} N^{\frac{1}{2}-\frac{1}{p}} (\log(e+N))^\alpha & \text{if } p > 2, \text{ or } p = 2 \text{ and } \alpha \geq 0 \\ N^{\frac{1}{p}-\frac{1}{2}} (\log(e+N))^{-\alpha} & \text{if } 1 < p < 2, \text{ or } p = 2 \text{ and } \alpha \leq 0. \end{cases}$$

Proof. Apply Lemmas 5.15 and 5.17 to the expression in (5.14). $\square$

5.3. WCGA for $\mathcal{T}$ in $L^p(\log L)^\alpha$. Combining the estimates from the previous subsections, we obtain the following.

THEOREM 5.22. *Let $1 < p < \infty$ and $\alpha \in \mathbb{R}$. Let $\mathbb{X} = L^p(\log L)^\alpha$ and $\mathcal{D} = \{e^{inx}\}_{n \in \mathbb{Z}}$ in $\mathbb{T}$. Then, there exists a constant $C > 1$ such that the WCGA satisfies*

$$\left\| f - \mathcal{G}_{\phi(N)}(f) \right\|_{L^p(\log L)^\alpha} \leq 2 \sigma_N(f), \quad \forall f \in L^p(\log L)^\alpha, \quad N \geq 2,$$

where

$$(5.23) \quad \phi(N) = C \begin{cases} N \log N & \text{when } p > 2 \\ N \log \log N & \text{when } p = 2 \text{ and } \alpha > 0 \\ N & \text{when } p = 2 \text{ and } \alpha = 0 \\ N (\log N)^{4\alpha-} \log \log N & \text{when } p = 2 \text{ and } \alpha < 0 \\ N^{p'-1} \frac{(\log N)^{p'\alpha-}}{(\log N)^{p'\alpha}} \log N & \text{when } 1 < p < 2. \end{cases}$$

Proof. Combine Theorem 1.16, with the estimates for $Q(t)$, $H(N)$ and k_N in Corollaries 4.7, 5.13 and 5.21. $\square$

REMARK 5.24. The necessity of the log factors and the powers in the above expression of $\phi(N)$ is not known, even in the case $\alpha = 0$ (except, of course, if $\mathbb{X} = L^2$). See [16, Open Question 8.2].

ACKNOWLEDGMENTS

Research partially supported by grants MTM2017-83262-C2-2-P, PID2019-105599GB-I00 from *Ministerio de Ciencia e Innovación* (Spain), and grant 20906/PI/18 from *Fundación Séneca* (Región de Murcia, Spain).

The author wishes to thank E. Hernández and D. Kutzarova for useful comments at different stages of this work.

REFERENCES

- [1] C. BENETT AND R.C. SHARPLEY, *Interpolation of operators*. Academic Press, 1988.
- [2] S. DILWORTH, G. GARRIGÓS, E. HERNÁNDEZ, D. KUTZAROVA, V. TEMLYAKOV, Lebesgue-type inequalities in greedy approximation. *Jour. Funct. Anal.* **280** (5) (2021), 108885
- [3] T. FIGIEL, On the moduli of convexity and smoothness. *Studia Math* **56** (1976), 121–155.

- [4] G. GARRIGÓS, E. HERNÁNDEZ, J.M. MARTELL, Wavelets, Orlicz spaces, and greedy bases. *Appl. Comput. Harmon. Anal.* 24 (1) (2008), 70–93.
- [5] M. KRASNOSEL'SKII, J. RUTICKII. *Convex functions and Orlicz spaces*, Noordhoff Ltd., Groningen 1961.
- [6] J. LINDENSTRAUSS, L. TZAFRIRI, *Classical Banach spaces*, vol II, Springer-Verlag 1979.
- [7] E. LIVSHITZ, V. TEMLYAKOV, Sparse approximation and recovery by greedy algorithms. *IEEE Trans. Inform. Theory* 60 (7) (2014), 3989–4000.
- [8] R. MALEEV, S. TROYANSKI, On the moduli of convexity and smoothness in Orlicz spaces. *Studia Math.* 54 (2) (1975), 131–141.
- [9] A. PAWLEWICZ, M. WOJCIECHOWSKI, Marcinkiewicz sampling theorem for Orlicz spaces. *Positivity* 26 (3) (2022), Paper No. 56.
- [10] M.M. RAO AND Z.D. REN, *Theory of Orlicz spaces*. Monographs and Textbooks in Pure and Applied Mathematics, **146**, Marcel Dekker Inc., 1991.
- [11] I. SINGER. *Bases in Banach spaces I*. Springer-Verlag, 1970.
- [12] V.N. TEMLYAKOV, Greedy algorithms in Banach spaces. *Adv. Comput. Math.* **14** (3) (2001), 277–292.
- [13] V.N. TEMLYAKOV. *Greedy Approximation*. Cambridge University Press, Cambridge, 2011.
- [14] V.N. TEMLYAKOV, Sparse approximation and recovery by greedy algorithms in Banach spaces. *Forum Math, Sigma* **2** (12) (2014), 26 p.
- [15] V.N. TEMLYAKOV, Sparse Approximation with Bases, *Advanced Courses in Mathematics CRM Barcelona*, Birkhäuser, Springer Basel 2015.
- [16] V.N. TEMLYAKOV. *Multivariate Approximation*. Cambridge University Press, Cambridge, 2018.
- [17] T. ZHANG, Sparse recovery with orthogonal matching pursuit under RIP. *IEEE Trans Inform Theory* 57 (2011), 6215–6221.

GUSTAVO GARRIGÓS, DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DE MURCIA, 30100 MURCIA, SPAIN

Email address: `gustavo.garrigos@um.es`