To appear in J. Math. Anal. Appl. (2009) doi:10.1016/j.jma.2009.04.21

MULTILINEAR EXTRAPOLATION AND APPLICATIONS TO THE BILINEAR HILBERT TRANSFORM

MARÍA JESÚS CARRO, LOUKAS GRAFAKOS, JOSÉ MARÍA MARTELL, AND FERNANDO SORIA

Abstract. We present two extrapolation methods for multi-sublinear operators that allow us to derive estimates for general functions from the corresponding estimates on characteristic functions. Of these methods, the first is applicable to general multi-sublinear operators while the second requires working with the so-called (ε,δ)-atomic operators. Among the applications, we discuss some new endpoint estimates for the bilinear Hilbert transform.

1. Introduction

For a variety of important operators in analysis, it is easier to derive a restricted type estimate, that is an estimate on characteristic functions of measurable sets, than to derive an estimate for general functions. It is therefore interesting to ask what kind of estimates can be obtained from a known restricted type estimate. This is, for example, the case for the Carleson operator [5]

$$Sf(x) = \sup_n |(D_n * f)(x)|,$$

where $f \in L^1(\mathbb{T})$ and D_n is the Dirichlet kernel on $\mathbb{T} = \{z \in \mathbb{C}; |z| = 1\}$, for which the following estimate is known (see [11])

$$\|S(\chi_E)\|_{L^{1,\infty}} \le C\, D(|E|)$$

with $D(t) = t(1 + \log^+ \frac{1}{t})$. Another example of this sort appears in the case of the bilinear Hilbert transform

$$H(f,g)(x) = \frac{1}{\pi} \lim_{\varepsilon\to 0} \int_{|t|\ge\varepsilon} f(x-t)g(x+t)\frac{dt}{t},$$

for which the following restricted type inequality has been proved in [3] (see also [4])

$$\big|\big\{x \in \mathbb{R} : |H(\chi_E,\chi_F)(x)| > \lambda\big\}\big| \le \frac{C}{\lambda^{2/3}}\left(1+\log^+\frac{1}{\lambda}\right)^{4/3}\Big(|E||F|\min(|E|,|F|)\Big)^{1/3},$$

Date: January 27, 2009. *Revised*: April 8, 2009.

2000 *Mathematics Subject Classification.* 47A30, 47A63, 42A99, 42B35.

Key words and phrases. Multilinear extrapolation, restricted weak type estimates, (ε,δ)-atomic operators, galb, bilinear Hilbert transform.

The first author is supported by MTM2007-60500 and by 2005SGR0556. The second author is supported by the National Science Foundation under grant DMS 0400387. The third and fourth authors are supported by MEC Grant MTM2007-60952 and by UAM-CM Grant CCG06-UAM/ESP-0286. The third author is also supported by CSIC PIE 200850I015.

for all $\lambda > 0$. Writing the last expression in terms of the decreasing rearrangement of $H(\chi_E, \chi_F)$, one can easily see that

$$\|H(\chi_E, \chi_F)\|_X \leq C D(|E|, |F|)\,, \tag{1.1}$$

where $D(s,t) = \big(st\min(s,t)\big)^{1/2}\Big(1+\log^+ \frac{1}{st\min(s,t)}\Big)^2$ and X is the weak type weighted Lorentz space $\Lambda^{2/3,\infty}(w)$ defined by

$$\|f\|_{\Lambda^{2/3,\infty}(w)} = \sup_{t>0} W(t)^{3/2} f^*(t),$$

with $W(t) = \int_0^t w(s)ds$ and, in this case, $W(t) = t(1+\log^+ 1/t)^{-4/3}$. In (1.1), the variables can be separated since for any $\alpha, \beta \in [0,1]$ with $\alpha+\beta = 1$, we have $D(s,t) \leq D_1(s)\, D_2(t)$ where

$$D_1(s) = s^{\frac{1+\alpha}{2}}\left(1+\log^+ \frac{1}{s}\right)^2, \qquad D_2(t) = t^{\frac{1+\beta}{2}}\left(1+\log^+ \frac{1}{t}\right)^2. \tag{1.2}$$

The preceding two examples provide the main motivation to investigate the boundedness properties of linear or multilinear operators for which restricted estimates are known. In the linear or sublinear case we assume that T satisfies

$$\|T(\chi_E)\|_X \leq C D(|E|) \tag{1.3}$$

for any measurable set E, $|E| < \infty$, where D is increasing with $D(0) = 0$ and X is a general quasi-Banach lattice space. Analogously, in the bilinear or bi-sublinear case, T may satisfy an estimate of the form

$$\|T(\chi_{E_1}, \chi_{E_2})\|_X \leq C D(|E_1|, |E_2|), \tag{1.4}$$

where D is a function which is increasing in both variables with $D(0,\cdot) = D(\cdot,0) = 0$. The analysis of m-linear or m-sublinear operators for $m \geq 3$ presents no significant differences and thus for simplicity in our exposition we may focus on the case $m = 2$.

In order to introduce the different approaches that we study in the present paper, we give an overview of the existing results in the linear (or sublinear) case. This study is motivated by the need to understand the a.e. convergence of Fourier series and thus the boundedness of the Carleson operator on spaces near L^1.

When $X = L^{1,\infty}$ and D is a concave function it is shown in [14] that if T satisfies (1.3) then T maps B^*_D (see [14] for the precise definition of this space) to $L^{1,\infty}$. The proof of this extrapolation result is based on decomposing each function f into simple functions to which the initial hypotheses are applied. Here it is worthwhile to point out that as X is a quasi-Banach space, additional issues appear since one needs to control the quasi-norm of a linear combination of functions. In this particular case one has that if $\{g_j\}_j$ is a sequence of functions with $\|g_j\|_{L^{1,\infty}} \leq 1$ then for any $\{c_j\}_j \subset \mathbb{R}$

$$\Big\|\sum_j c_j\, g_j\Big\|_{L^{1,\infty}} \lesssim \big\|\{c_j\}_j\big\|_{\ell(\log \ell)}. \tag{1.5}$$

These ingredients appear in the adaptation of this scheme to the m-linear setting. The method introduced in [14] applied to the Carleson operator S with $D(t) = t\,(1+\log^+ \frac{1}{t})$, gives that the Fourier series of each function in $L\,(\log L)\,(\log\log L)$

converges a.e.; this follows from the corresponding result in B_D^* since the latter space contains the former, see [14] for the precise details.

Concerning the a.e. convergence of Fourier series and the boundedness of the Carleson operator, a closer space to L^1 was obtained by Antonov in [1], namely, $L\,(\log L)\,(\log\log\log L)$. The ideas in [1] have been exploited in [13] (see also [9] for related results) to obtain that Antonov's result is a particular example of a general extrapolation result: the method developed in [14] can be improved when applied to maximal operators $T_*f(x) = \sup_j |K_j * f(x)|$ where $K_j \in L^1$ (similar results are obtained for variable kernels, see [13], [9]). A further extension of these techniques is introduced in [6] and [7] where a more general class of operators, called (ε, δ)-atomic (see the definition below), is considered. It is shown in [6] that if an (ε, δ)-atomic operator T satisfies (1.3), which is an estimate for characteristic functions, then the same estimate holds for any function $f \in L^1$ with $\|f\|_\infty \le 1$:

$$\|Tf\|_X \le CD(\|f\|_1). \tag{1.6}$$

This means that taking (1.6) as the initial assumption (note that this with $f = \chi_E$ is (1.3)), in place of decomposing f into simple functions as in [14], one can use more general bounded functions. This was used in [6] to give another proof of Antonov's result: The Carleson operator is (ε, δ)-atomic and (1.6) holds with $X = L^{1,\infty}$ and $D(t) = t\,(1 + \log^+ \frac{1}{t})$. The key idea in [1] relies on decomposing each function f according to the level sets $\{d_{k-1} < |f| \le d_k\}$ with $d_k = 2^{2^k}$. Again, to deal with linear combinations in X one uses (1.5). This allows one to obtain an estimate from $L(\log L)\,(\log\log\log L)$ to $L^{1,\infty}$. Let us observe that having taken the more "natural" sequence $d_k = 2^k$ would have led us to the smaller space $L(\log L)\,(\log\log L)$.

Motivated by the aforementioned results, in the present paper we extend the two approaches outlined above to the case of m-linear or m-sublinear operators. We first extend the approach in [14]: for general operators satisfying (1.4) we decompose the given functions into simple functions; the need to control the quasi-norm of linear combinations of simple functions requires a substitute for (1.5). Note that in this case it is natural to consider target spaces X that are quasi-Banach spaces below L^1 (this is the case for the bilinear Hilbert transform). We use the concept introduced by Turpin [15] of the Galb(X) of a quasi-Banach space X defined as follows

$$\mathrm{Galb}(X) = \Big\{(c_n)_n;\ \sum_n c_n f_n \in X, \quad \text{whenever } \|f_n\|_X \le 1\Big\},$$

endowed with the norm $\|c\|_{\mathrm{Galb}(X)} = \sup_{\|f_n\|_X \le 1} \big\| \sum_n c_n f_n \big\|_X$. This Galb space was studied in [7] for the case of the weighted Lorentz spaces $X = \Lambda^q(w)$, for $0 < q < \infty$, and also for the weak spaces $X = \Lambda^{q,\infty}(w)$.

Next, we introduce (ε, δ)-atomic operators in the multi-variable setting. For these, estimates for characteristic functions of the form (1.4) can be extended to L^1 functions bounded by 1 (as in (1.6)). Thus we take as initial assumption the more general estimate

$$\|T(f_1, \dots, f_m)\|_X \le CD(\|f_1\|_1, \dots, \|f_m\|_1),$$

for all functions $(f_1, \dots, f_m) \in L^1 \times \cdots \times L^1$ with $\|f_j\|_\infty \le 1$ for $j = 1, \dots, m$. By decomposing general functions not only into sums of simple functions but also into combinations of bounded functions, better results can be obtained following the ideas in [6] and [7]. Our main motivation in the study of this problem is to obtain estimates for the bilinear Hilbert transform. For this purpose, it is natural to consider quasi-Banach spaces with $\mathrm{Galb}(X) = \ell^q$ with $0 < q < 1$ and functions D_1, D_2 like in (1.2).

We denote by $L^0(\mathbb{R}^n)$ the class of Lebesgue measurable functions that are finite a.e. and by $g^*(t) = \inf\{s : \mu_g(s) \le t\}$ the decreasing rearrangement of $g \in L^0$, where $\mu_g(y) = |\{x \in \mathbb{R} : \ |g(x)| > y\}|$ is the distribution function of g with respect to the Lebesgue measure (we refer the reader to [2] for further information about distribution functions and decreasing rearrangements). For a measurable set E, χ_E denotes its characteristic function and $|E|$ its Lebesgue measure. For simplicity of presentation, we say that an operator T is *sublinear* if $|T(\lambda f)| = |\lambda|\,|Tf|$ and

$$\Big|T\Big(\sum_{n\in\mathbb{N}} f_n\Big)\Big| \le \sum_{n\in\mathbb{N}} |Tf_n|$$

for all functions f, f_n and $\lambda \in \mathbb{R}$. If we only have that $|T(f_1 + f_2)| \le |Tf_1| + |Tf_2|$, then to obtain our conclusions we need to assume an additional boundedness condition on our operator T such as

$$T : L^1 + L^\infty \longrightarrow L^0\,,$$

or assume some density property of the spaces in question.

For m-linear or m-sublinear operators we state and prove our results in the case $m = 2$, since the case with more variables only presents trivial notational changes.

Given a function $D(s,t)$, increasing in each variable with $D(0,\cdot) = D(\cdot,0) = 0$, we write $dD = dD(s,t)$ for the measure in $[0,\infty)^2$ defined by

$$dD\big([0,a)\times[0,b)\big) = \iint_{[0,a)\times[0,b)} dD(s,t) = D(a,b).$$

Note that if D is smooth then $dD(t,s) = \partial_t\partial_s D(t,s)\,dt\,ds$.

2. Decompositions into simple functions and estimates on Lorentz spaces

Given an increasing function D such that $D(0) = 0$ and $0 < q < \infty$, the Lorentz space $\Lambda^q(dD)$ is given by

$$\|f\|_{\Lambda^q(dD)} = \Big(\int_0^\infty f^*(t)^q\, dD(t)\Big)^{\frac{1}{q}} \approx \Big(\int_0^\infty \lambda^q\, D\big(\mu_f(\lambda)\big)\,\frac{d\lambda}{\lambda}\Big)^{\frac{1}{q}}.$$

It is known that this space is quasi-Banach if and only if the function D satisfies the Δ_2-condition; that is $D(2\,t) \le C D(t)$ for some constant $C > 0$ and for every $t > 0$, see [8].

2.1. Sublinear case.

We start with the sublinear case which already contains many of the ideas that will be used in the m-linear setting. We have the following result.

Theorem 2.1. *Let T be a sublinear operator, let X be a quasi-Banach lattice space and let D be an increasing function such that $D(0) = 0$. Assume that, for any measurable set E with $|E| < \infty$, we have*

$$\|T(\chi_E)\|_X \leq C\, D(|E|). \tag{2.1}$$

Then, the following are valid:

(*a*) *If* $\mathrm{Galb}(X) = \ell^1$, *then*

$$T : \Lambda^1(dD) \longrightarrow X.$$

(*b*) *If* $\mathrm{Galb}(X) = \ell^p$ *with* $0 < p < 1$, *then*

$$T : \Lambda^p(dD^p) \longrightarrow X.$$

(*c*) *If* $\mathrm{Galb}(X) = \ell\,(\log \ell)^\alpha$ *with* $\alpha > 0$, *then*

$$T : \Lambda^*_\alpha(dD) \longrightarrow X,$$

where $\Lambda^*_\alpha(dD)$ *is the subspace of* $\Lambda^1(dD)$ *defined by the functional*

$$\begin{aligned}\|f\|_{\Lambda^*_\alpha(dD)} &= \int_0^\infty \lambda\, D\big(\mu_f(\lambda)\big) \left(1 + \log^+ \frac{\|f\|_{\Lambda^1(dD)}}{\lambda\, D\big(\mu_f(\lambda)\big)}\right)^\alpha \frac{d\lambda}{\lambda} \\ &= \|f\|_{\Lambda^1(dD)} \int_0^\infty \varphi_\alpha\left(\frac{\lambda\, D\big(\mu_f(\lambda)\big)}{\|f\|_{\Lambda^1(dD)}}\right) \frac{d\lambda}{\lambda}\end{aligned}$$

with $\varphi_\alpha(t) = t\,(1 + \log^+ 1/t)^\alpha$.

We do not prove Theorem 2.1 here. In Theorem 2.6 below we obtain similar results for bi-sublinear operators and the arguments in that proof can be easily adapted in the proof of Theorem 2.1.

Remark 2.2. We notice that in (*a*) and (*b*) we do not "lose" information since we may recover the initial assumption by applying the obtained estimate to characteristic functions since $\|\chi_E\|_{\Lambda^p(dD^p)} = D(|E|)$. More precisely, given X such that $\mathrm{Galb}(X) = \ell^p$ with $0 < p \leq 1$ then

$$\|T(\chi_E)\|_X \lesssim D(|E|), \quad |E| < \infty \qquad \Longleftrightarrow \qquad T : \Lambda^p(dD^p) \longrightarrow X.$$

The same occurs in (*c*) since

$$\|\chi_E\|_{\Lambda^*_\alpha(dD)} = D(|E|) \int_0^1 \varphi_\alpha\left(\frac{\lambda\, D(|E|)}{D(|E|)}\right) \frac{d\lambda}{\lambda} \approx D(|E|),$$

and therefore

$$\|T(\chi_E)\|_X \lesssim D(|E|), \quad |E| < \infty \qquad \Longleftrightarrow \qquad T : \Lambda^*_\alpha(dD) \longrightarrow X.$$

Remark 2.3. Let us observe that Theorem 2.1 part (*a*) (with D concave) is optimal in the sense that one cannot expect a space bigger than $\Lambda^1(dD)$ valid for every operator T satisfying (2.1): we take $X = \Lambda^1(dD)$ which is a Banach space, $T = Id$ and we observe that $\|T(\chi_E)\|_X = D(|E|)$.

Example 2.4. Suppose that X is any Banach space and hence $\mathrm{Galb}(X) = \ell^1$. If $D(t) = t^{1/q}$ then $\Lambda(dD) = L^{q,1}$ and we have that

$$\|T(\chi_E)\|_X \lesssim |E|^{1/q}, \quad |E| < \infty \qquad \Longleftrightarrow \qquad T : L^{q,1} \longrightarrow X.$$

This is the best estimate one can obtain from the restricted type assumption on T. Naturally, this conclusion may not be optimal if T itself maps the bigger space L^q (when $q \geq 1$) into X.

The method of Theorem 2.1 does not use any specific property of the operator T. For instance, if we know that T is a supremum of a sequence of linear operators (as in Moon's theorem [12]) or even more generally, that T is atomic (see the corresponding definition in the next section), then we are able to obtain a better conclusion. Let us examine a few more examples using the previous method.

Example 2.5. In these examples we set $D(t) = t$.

- Let $X = L^{q,\infty}$ with $q > 1$, hence X is a Banach space. Then, we have for any sublinear operator T,

$$\|T(\chi_E)\|_{L^{q,\infty}} \lesssim |E|, \quad |E| < \infty \qquad \Longleftrightarrow \qquad T : L^1 \longrightarrow L^{q,\infty}.$$

  We note that this equivalence can be obtained directly by working with simple functions.

- Let $X = L^{q,\infty}$ with $0 < q < 1$, hence $\mathrm{Galb}(X) = \ell^q$. In this case $\Lambda^q(dD^q) = L^{1,q}$ and, for any sublinear operator T,

$$\|T(\chi_E)\|_{L^{q,\infty}} \lesssim |E|, \quad |E| < \infty \qquad \Longleftrightarrow \qquad T : L^{1,q} \longrightarrow L^{q,\infty}.$$

  However, Moon's theorem [12] says that under certain conditions on T, one obtains that T maps L^1 into $L^{q,\infty}$ which is a stronger conclusion since $L^{1,q} \subsetneqq L^1$ for $0 < q < 1$.

- Let $X = L^{1,\infty}$, hence $\mathrm{Galb}(X) = \ell \log \ell$ and $\Lambda(dD) = L^1$. In this case we have $\Lambda_1^*(dD) = B_{\varphi_0}^*$ (see [14]) and thus

$$\|T(\chi_E)\|_{L^{1,\infty}} \lesssim |E|, \quad |E| < \infty \qquad \Longleftrightarrow \qquad T : \Lambda_1^*(dD) \longrightarrow L^{1,\infty}.$$

  Yet another comparison with Moon's theorem yields that, under some conditions on T, it is bounded from L^1 into $L^{1,\infty}$ which is a stronger conclusion since

$$\|f\|_{L^1} = \int_0^\infty \lambda\, \mu_f(\lambda)\, \frac{d\lambda}{\lambda} \leq \int_0^\infty \lambda\, \mu_f(\lambda) \left(1 + \log^+ \frac{\|f\|_{L^1}}{\lambda\, \mu_f(\lambda)}\right) \frac{d\lambda}{\lambda} = \|f\|_{\Lambda_1^*(dD)},$$

  and hence, $\Lambda_1^*(dD) \subset L^1$. To see that this inclusion is proper we take

$$f(x) = \frac{1}{x \log x\, (\log\log x)^2}\, \chi_{[e^e,\infty)}(x).$$

  We have that $f \in L^1$ but one can easily see that $\|f\|_{\Lambda_1^*(dD)} = \infty$. Thus, $\Lambda_1^*(dD)$ is a proper subspace of L^1.

2.2. **Bi-sublinear case.** Before discussing a bi-sublinear extension of Theorem 2.1 we introduce some notation. Given a function of two variables D such that it is increasing in each variable and $D(0,\cdot) = D(\cdot,0) = 0$, let $\vec{\Lambda}^p(dD^p)$ be the function space given by

$$\begin{aligned}\|(f_1,f_2)\|_{\vec{\Lambda}^p(dD^p)} &= \int_0^\infty\int_0^\infty f_1^*(s_1)^p\, f_2^*(s_2)^p\, dD^p(s_1,s_2)\\ &\approx \int_0^\infty\int_0^\infty s_1^p\, s_2^p\, D\big(\mu_{f_1}(s_1),\mu_{f_2}(s_2)\big)^p\, \frac{ds_1\, ds_2}{s_1\, s_2}.\end{aligned}$$

Notice that if $D(s_1,s_2) = D_1(s_1)\,D_2(s_2)$ then $dD^p(s_1,s_2) = dD_1^p(s_1)\,dD_2^p(s_2)$ and $\vec{\Lambda}^p(dD^p) = \Lambda^p(dD_1^p)\times\Lambda^p(dD_2^p)$ since $\|(f_1,f_2)\|_{\vec{\Lambda}^p(dD^p)} = \|f_1\|_{\Lambda^p(dD_1^p)}\,\|f_2\|_{\Lambda^p(dD_2^p)}$.

Analogously, we introduce the function space $\vec{\Lambda}^*_\alpha(dD)$ given by the functional:

$$\begin{aligned}&\|(f_1,f_2)\|_{\vec{\Lambda}^*_\alpha(dD)} =\\ &= \int_0^\infty\int_0^\infty D\big(\mu_{f_1}(s_1),\mu_{f_2}(s_2)\big)\left[1+\log^+\frac{\|(f_1,f_2)\|_{\vec{\Lambda}^1(dD)}}{s_1\, s_2\, D\big(\mu_{f_1}(s_1),\mu_{f_2}(s_2)\big)}\right]^\alpha ds_1 ds_2\\ &= \|(f_1,f_2)\|_{\vec{\Lambda}^1(dD)}\int_0^\infty\int_0^\infty \varphi_\alpha\left(\frac{s_1\, s_2\, D\big(\mu_{f_1}(s_1),\mu_{f_2}(s_2)\big)}{\|(f_1,f_2)\|_{\vec{\Lambda}^1(dD)}}\right)\frac{ds_1\, ds_2}{s_1\, s_2},\end{aligned}$$

with $\varphi_\alpha(t) = t\,(1+\log^+ 1/t)^\alpha$. In this case, if $D(s_1,s_2) = D_1(s_1)\,D_2(s_2)$ we have that $\Lambda^*_\alpha(dD_1)\times\Lambda^*_\alpha(dD_2) \hookrightarrow \vec{\Lambda}^*_\alpha(dD)$ since $\|(f_1,f_2)\|_{\vec{\Lambda}^*_\alpha(dD)} \le \|f_1\|_{\Lambda^*_\alpha(dD_1)}\,\|f_2\|_{\Lambda^*_\alpha(dD_2)}$.

We now state a bi-sublinear extension of Theorem 2.1:

Theorem 2.6. *Let T be a bi-sublinear operator and let X be a quasi-Banach space. Let D be a two-variable function increasing in each variable with $D(0,\cdot) = D(\cdot,0) = 0$. Assume that for all measurable sets E_1, E_2 with $|E_1|$, $|E_2| < \infty$, we have*

$$\big\|T(\chi_{E_1},\chi_{E_2})\big\|_X \lesssim D(|E_1|,|E_2|). \tag{2.2}$$

Then the following are valid:

(a) *If* $\mathrm{Galb}(X) = \ell^1$, *then* $T : \vec{\Lambda}^1(dD) \longrightarrow X$.

(b) *If* $\mathrm{Galb}(X) = \ell^p$ *with* $0 < p < 1$, *then* $T : \vec{\Lambda}^p(dD^p) \longrightarrow X$.

(c) *If* $\mathrm{Galb}(X) = \ell\,(\log\ell)^\alpha$ *with* $\alpha > 0$, *then* $T : \vec{\Lambda}^*_\alpha(dD) \longrightarrow X$.

As an immediate consequence of this result and the discussion above, in the particular case $D(s,t) = D_1(s)\,D_2(t)$, we obtain the following result.

Corollary 2.7. *Let T be a bi-sublinear operator and let X be a quasi-Banach space. Let D_1, D_2 be increasing functions that vanish at the origin. Assume that for all measurable sets E_1, E_2 with $|E_1|$, $|E_2| < \infty$ we have*

$$\|T(\chi_{E_1},\chi_{E_2})\|_X \lesssim D_1(|E_1|)\,D_2(|E_2|). \tag{2.3}$$

(a) *If* $\mathrm{Galb}(X) = \ell^1$, *then* $T : \Lambda^1(dD_1)\times\Lambda^1(dD_2) \longrightarrow X$.

(b) *If* $\mathrm{Galb}(X) = \ell^p$ *with* $0 < p < 1$, *then* $T : \Lambda^p(dD_1^p)\times\Lambda^p(dD_2^p) \longrightarrow X$.

(*c*) *If* $\mathrm{Galb}(X) = \ell\,(\log\ell)^\alpha$ *with* $\alpha > 0$, *then* $T : \Lambda^*_\alpha(dD_1) \times \Lambda^*_\alpha(dD_2) \longrightarrow X$.

We point out that this corollary is also a consequence of the corresponding one-variable result: we freeze one variable in T and apply Theorem 2.1 to the resulting sublinear operator, then we freeze the other variable and apply again Theorem 2.1.

Proof of Theorem 2.6. Assume without loss of generality that $f,\ g \ge 0$. By [14, Lemma 4],

$$f(x) = \sum_{k\in\mathbb{Z}} \sum_{j\ge 1} 2^k\,\chi_{E_{k,j}}(x) \quad \text{a.e.}; \qquad g(x) = \sum_{k'\in\mathbb{Z}} \sum_{j'\ge 1} 2^{k'}\,\chi_{F_{k',j'}}(x) \quad \text{a.e.},$$

where

$$|E_{k,j}| \le |\{x : f(x) > 2^{k+j}\}| = \mu_f(2^{k+j}),$$

and

$$|F_{k',j'}| \le |\{x : g(x) > 2^{k'+j'}\}| = \mu_g(2^{k'+j'}).$$

Thus, using (2.2) we have

$$\begin{aligned}\|T(f,g)\|_X &\le \Big\| \sum_{j,j'k,k'} 2^k\,2^{k'}\,\big|T(\chi_{E_{k,j}}, \chi_{F_{k',j'}})\big| \Big\|_X \\ &\le \big\| \big\{2^k\,2^{k'}\,D\big(\mu_f(2^{j+k}), \mu_g(2^{j'+k'})\big)\big\}_{j,j',k,k'} \big\|_{\mathrm{Galb}(X)}.\end{aligned}$$

We start with (*a*) and (*b*) in which case $\mathrm{Galb}(X) = \ell^p$ with $0 < p \le 1$. Then,

$$\begin{aligned}\|T(f,g)\|_X^p &\le \sum_{k,k'\in\mathbb{Z}} \sum_{j,j'\ge 1} 2^{k\,p}\,2^{k'\,p}\,D\big(\mu_f(2^{j+k}), \mu_g(2^{j'+k'})\big)^p \\ &\lesssim \sum_{j,j'\ge 1} \int_0^\infty\!\!\int_0^\infty s^p\,t^p\,D\big(\mu_f(s\,2^j), \mu_g(t\,2^{j'})\big)^p\,\frac{ds}{s}\,\frac{dt}{t} \\ &= \sum_{j,j'\ge 1} 2^{-j\,p}\,2^{-j'\,p} \int_0^\infty\!\!\int_0^\infty s^p\,t^p\,D\big(\mu_f(s), \mu_g(t)\big)^p\,\frac{ds}{s}\,\frac{dt}{t} \\ &\lesssim \int_0^\infty\!\!\int_0^\infty s^p\,t^p\,D\big(\mu_f(s), \mu_g(t)\big)^p\,\frac{ds}{s}\,\frac{dt}{t} \approx \|(f,g)\|^p_{\vec{\Lambda}^p(dD^p)}.\end{aligned}$$

Let us establish (*c*) in which case $\mathrm{Galb}(X) = \ell\,(\log\ell)^\alpha$ with $\alpha > 0$. Following [14], we write $\varphi_\alpha(t) = t\,(1 + \log^+ 1/t)^\alpha$ and we observe that given a non-trivial sequence of non-negative numbers $a = \{a_k\}_k$ we have

$$\|a\|_{\ell\,(\log\ell)^\alpha} \le \sum_k a_k \left(1 + \log\frac{\|a\|_{\ell^1}}{a_k}\right)^\alpha = \mathcal{N}_\alpha(a).$$

We write $F_j = 2^{-j}\,f$, $G_{j'} = 2^{-j'}\,g$ and

$$\beta_{k,k',j,j'} = 2^k\,2^{k'}\,D(\mu_f(2^{j+k}), \mu_g(2^{j'+k'})) = 2^k\;2^{k'}\,D(\mu_{F_j}(2^k), \mu_{G_{j'}}(2^{k'})).$$

As in [14, p. 239] (there, the computations are done with $\alpha = 1$ but the argument adapts trivially to an arbitrary $\alpha > 0$) we obtain

$$\|T(f,g)\|_X \lesssim \mathcal{N}_\alpha(\{\beta_{k,k',j,j'}\}_{k,k',j,j'}) \lesssim \mathcal{N}_\alpha\big(\big\{\mathcal{N}_\alpha(\{\beta_{k,k',j,j'}\}_{k,k'})\big\}_{j,j'}\big).$$

Then, for any $j, j' \geq 1$,

$$\sum_{k,k'} \beta_{k,k',j,j'} = \sum_{k,k'} 2^k\, 2^{k'}\, D(\mu_{F_j}(2^k), \mu_{G_{j'}}(2^{k'})) \lesssim \int_0^\infty\!\!\int_0^\infty s\,t\, D\big(\mu_{F_j}(s), \mu_{G_{j'}}(t)\big)\, \frac{ds\,dt}{s\,t}$$
$$\approx \|(F_j, G_{j'})\|_{\vec{\Lambda}^1(dD)} = 2^{-j}\, 2^{-j'} \|(f,g)\|_{\vec{\Lambda}^1(dD)}$$

and

$$\mathcal{N}_\alpha(\{\beta_{k,k',j,j'}\}_{k,k'}) \lesssim \sum_{k,k'} \beta_{k,k',j,j'} \left(1 + \log \frac{2^{-j}\, 2^{-j'}\, \|(f,g)\|_{\vec{\Lambda}^1(dD)}}{\beta_{k,k',j,j'}}\right)^\alpha$$
$$\lesssim \int_0^\infty\!\!\int_0^\infty s\,t\, D\big(\mu_{F_j}(s), \mu_{G_{j'}}(t)\big) \left(1 + \log^+ \frac{2^{-j}\, 2^{-j'}\, \|(f,g)\|_{\vec{\Lambda}^1(dD)}}{s\,t\, D\big(\mu_{F_j}(s), \mu_{G_{j'}}(t)\big)}\right)^\alpha \frac{ds\,dt}{s\,t}$$
$$= 2^{-j}\, 2^{-j'}\, \|(f,g)\|_{\vec{\Lambda}^*_\alpha(dD)}.$$

Therefore,

$$\|T(f,g)\|_X \lesssim \mathcal{N}_\alpha\big(\{2^{-j}\, 2^{-j'}\, \|(f,g)\|_{\vec{\Lambda}^*_\alpha(dD)}\}_{j,j'}\big) = \|(f,g)\|_{\vec{\Lambda}^*_\alpha(dD)}\, \mathcal{N}_\alpha\big(\{2^{-j}\, 2^{-j'}\}_{j,j'}\big)$$
$$= C\, \|(f,g)\|_{\vec{\Lambda}^*_\alpha(dD)}.$$

□

Remark 2.8. As already observed in Remark 2.2, in Theorem 2.6 (and thus in Corollary 2.7) we do not "lose" information and we recover the initial assumption by applying the obtained estimate to characteristic functions since $\|(\chi_E, \chi_F)\|_{\vec{\Lambda}^p(dD^p)} = D(|E|, |F|)$ and

$$\|(\chi_E, \chi_F)\|_{\vec{\Lambda}^*_\alpha(dD)} = D(|E|, |F)\int_0^1\!\!\int_0^1 \varphi_\alpha(s\,t)\, \frac{ds\,dt}{s\,t} \approx D(|E|, |F|).$$

Therefore, given X such that $\mathrm{Galb}(X) = \ell^p$ with $0 < p \leq 1$,

$$\|T(\chi_E, \chi_F)\|_X \lesssim D(|E|, |F|), \quad |E|, |F| < \infty \qquad \Longleftrightarrow \qquad T : \vec{\Lambda}^p(dD^p) \longrightarrow X$$

and, given X such that $\mathrm{Galb}(X) = \ell\,(\log \ell)^\alpha$,

$$\|T(\chi_E, \chi_F)\|_X \lesssim D(|E|, |F|), \quad |E|, |F| < \infty \qquad \Longleftrightarrow \qquad T : \vec{\Lambda}^*_\alpha(dD) \longrightarrow X.$$

3. Atomic operators

3.1. **Atomic and one-variable case.** Let us recall first some definitions and results from [6] and [7]. We work in $\mathbb{R}^n$, and Q represents a cube with sides parallel to the coordinate axes. The results can be extended in the natural way to $\mathbb{T}^N$ (identifying $\mathbb{T}^N$ with $[0,1)^N$). In [6], the following definitions were introduced:

Definition 3.1. *Given $\delta > 0$, a function $a \in L^1(\mathbb{R}^n)$ is called a δ-atom if it satisfies the following properties:*

(a) $\displaystyle\int_{\mathbb{R}^n} a(x)\, dx = 0.$

(b) *There exists a cube Q such that $|Q| \leq \delta$ and* $\operatorname{supp} a \subset Q$.

Definition 3.2.

(*a*) *A sublinear operator T is (ε,δ)-atomic if, for every $\varepsilon>0$, there exists $\delta>0$ such that, for every δ-atom a,*

$$\|Ta\|_{L^1+L^\infty}\le\varepsilon\|a\|_1.$$

(*b*) *A sublinear operator T is (ε,δ)-atomic approximable if there exists a sequence $(T_j)_j$ of (ε,δ)-atomic operators such that, for every measurable set E, $|T_j(\chi_E)|\le|T(\chi_E)|$ and, for every $f\in L^1$ such that $\|f\|_\infty\le 1$, and every $t>0$,*

$$(Tf)^*(t)\le\liminf_j(T_jf)^*(t)\cdot$$

In particular, any maximal operator of the form $\sup_j|K_j*f|$, where $K_j\in L^{p_j}$ for some $1\le p_j<\infty$, is (ε,δ)-atomic approximable (see [6] for more examples of this kind of operators). Also, it was proved in [7] that operators bounded on L^p with $0<p<1$ are not (ε,δ)-atomic approximable.

Theorem 3.3 ([7])**.** *Let T be a sublinear, (ε,δ)-atomic approximable operator. Let X be a quasi-Banach rearrangement invariant function space. Assume that, for every measurable set E,*

$$\|T(\chi_E)\|_X\le D(|E|),\tag{3.1}$$

for some positive function D. Then, for every function $f\in L^1$ with $\|f\|_\infty\le1$ we have

$$\|Tf\|_X\le D(\|f\|_1).\tag{3.2}$$

As a consequence of this result, we can improve Theorem 2.1 when $D(t)=t$.

Corollary 3.4 ([7])**.** *Let T be a sublinear, (ε,δ)-atomic approximable operator and let X be a quasi-Banach r.i. space. Assume that for any measurable set E with $|E|<\infty$ we have*

$$\|T(\chi_E)\|_X\le C\,|E|.\tag{3.3}$$

Then,

$$\|Tf\|_X\le C\,\|f\|_1,\qquad f\in L^1\cap L^\infty,$$

and thus $T:L^1\longrightarrow X$.

This result extends Moon's theorem since it includes a wider class of operators and holds for any quasi-Banach space X.

Proof. Let $f\in L^1\cap L^\infty$ and write $\tilde f=f/\|f\|_\infty$ such that $\|\tilde f\|_\infty\le1$. Thus, (3.3) and Theorem 3.3 imply

$$\|Tf\|_X=\|f\|_\infty\,\|T\tilde f\|_X\le C\,\|f\|_\infty\,\|\tilde f\|_1=C\,\|f\|_1.$$

To complete the proof let us give the density argument. Let $f\in L^1$. As $L^1\cap L^\infty$ is dense in L^1, there exists a sequence $\{f_k\}_k\subset L^1\cap L^\infty$ such that $f_k\longrightarrow f$ in L^1 as $k\to0$. As T is sublinear we have

$$\big\||Tf_j|-|Tf_k|\big\|_X\le\|T(f_j-f_k)\|_X\le C\,\|f_j-f_k\|_{L^1}.$$

Thus $\{Tf_k\}_k$ is a Cauchy sequence on X and hence is convergent in X. This allows us to define Tf and to conclude that T maps L^1 into X. □

We compare Theorem 2.1 with Corollary 3.4. Note that Theorem 2.1 part (a) and Corollary 3.4 give the same estimate since $\Lambda(dD) = L^1$ (in the latter we assume that T is (ε,δ)-atomic approximable). In Theorem 2.1 part (b) (resp. (c)) we obtain that T maps $\Lambda^p(dD^p) = L^{1,p}$ (resp. $\Lambda^*(dD)$) into X. These are improved in Corollary 3.4 since $L^{1,p} \subsetneq L^1$ for $0 < p < 1$ and $\Lambda_1^*(dD) \subsetneq L^1$. Therefore, the atomicity assumption allows one to obtain better estimates.

This reflects that, in principle, the estimates in Theorem 2.1 can be improved when the operator T is (ε,δ)-atomic approximable. We notice that, once we know that (3.2) holds, we can take this as our initial assumption. This contains in particular the restricted type estimate (3.1) and as we start with a more general estimate more decompositions of the functions are allowed. We follow this approach in Section 3.3 where we consider the functions $D(t) = t^q$, $D(t) = t^q\,(1+\log^+ 1/t)^\alpha$, etc. Eventually we apply these results to the bilinear Hilbert Transform.

Remark 3.5. When X is a Banach space and D is concave then (3.1) implies (3.2), whether or not T is atomic. To see this, let $f \in L^1$ with $\|f\|_\infty \le 1$. As X is a Banach space $\mathrm{Galb}(X) = \ell^1$. This fact, Theorem 2.1 part (a) and the concavity of D yield

$$\|Tf\|_X \le C\,\|f\|_{\Lambda^1(dD)} = C\int_0^1 D\big(\mu_f(\lambda)\big)\,d\lambda \le C\,D\Big(\int_0^1 \mu_f(\lambda)\,d\lambda\Big) = C\,D\big(\|f\|_1\big).$$

This means that the fact that a given operator is atomic only matters when X is a quasi-Banach space.

3.2. **Atomic and multi-variable case.**

Definition 3.6. *Given $\delta > 0$, a pair of functions $(a_1, a_2) \in L^1(\mathbb{R}^n) \times L^1(\mathbb{R}^n)$ is called a δ-atom if it satisfies the following properties:*

(a) $\displaystyle\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} a_1(x_1)\,a_2(x_2)\,dx_1\,dx_2 = 0.$

(b) *There exist cubes Q_1, Q_2 with $|Q_1|$, $|Q_2| \le \delta$ such that* $\operatorname{supp} a_1 \subset Q_1$, $\operatorname{supp} a_2 \subset Q_2$.

Definition 3.7.

(a) *A bi-sublinear operator T is (ε,δ)-atomic if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every δ-atom (a_1, a_2),*

$$\|T(a_1,a_2)\|_{L^1+L^\infty} \le \varepsilon \|a_1\|_1\,\|a_2\|_1.$$

(b) *A bi-sublinear operator T is (ε,δ)-atomic approximable if there exists a sequence $(T_n)_n$ of (ε,δ)-atomic operators such that, for all measurable sets E_1, E_2*

$$|T_n(\chi_{E_1},\chi_{E_2})| \le |T(\chi_{E_1},\chi_{E_2})|$$

and, for all $(f_1,f_2) \in L^1 \times L^1$ such that $\|f_1\|_\infty$, $\|f_2\|_\infty \le 1$, and every $t > 0$,

$$(T(f_1,f_2))^*(t) \le \liminf_n (T_n(f_1,f_2))^*(t).$$

(c) *A multi-bilinear operator is "iterative" (ε,δ)-atomic (approximable), if for every $f_0 \in L^1$ with $\|f_0\|_\infty \le 1$, the sublinear operators $T_1 g = T(g, f_0)$ and $T_2 g = T(f_0, g)$ are (ε,δ)-atomic (approximable).*

Example 3.8. Consider

$$T(f_1, f_2)(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(x, y_1, y_2)\, f_1(y_1)\, f_2(y_2)\, dy_1\, dy_2.$$

Assume that it is well defined for $(f_1, f_2) \in L^1 \times L^1$ and the kernel K satisfies that

$$\lim_{(y_1,y_2)\to(x_1,x_2)} \|K(\cdot, y_1, y_2) - K(\cdot, x_1, x_2)\|_{L^1+L^\infty} = 0, \tag{3.4}$$

uniformly in $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$, then T is (ε,δ)-atomic. To see this, observe that if (a_1, a_2) is a δ-atom, then

$$\begin{aligned}
\|T(a_1, a_2)\|_{L^1+L^\infty} &= \Big\| \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K(\cdot, y_1, y_2)\, a_1(y_1)\, a_2(y_2)\, dy_1\, dy_2 \Big\|_{L^1+L^\infty} \\
&= \Big\| \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (K(\cdot, y_1, y_2) - K(\cdot, x_{Q_1}, x_{Q_2}))\, a_1(y_1)\, a_2(y_2)\, dy_1\, dy_2 \Big\|_{L^1+L^\infty} \\
&\le \int_{Q_1}\int_{Q_2} \|K(\cdot, y_1, y_2) - K(\cdot, x_{Q_1}, x_{Q_2})\|_{L^1+L^\infty} |a_1(y_1)|\, |a_2(y_2)| dy_1 dy_2,
\end{aligned}$$

with x_{Q_j} being the center of the cube Q_j where a_j is supported. Therefore, given ε, we can choose δ in such a way that the above quantity is bounded by $\varepsilon\|a_1\|_1\, \|a_2\|_1$.

In particular we have the following examples:

(A) For functions f_1, f_2 on $\mathbb{R}^n$ define their tensor on $\mathbb{R}^{2n}$ by $(f_1 \otimes f_2)(x, y) = f_1(x) f_2(y)$ for $x, y \in \mathbb{R}^n$. If $K \in L^p(\mathbb{R}^n \times \mathbb{R}^n)$ for some $1 \le p < \infty$ and $T(f_1, f_2)(x) = (K * (f_1 \otimes f_2))(x, x)$, $x \in \mathbb{R}^n$, then (3.4) holds since

$$\lim_{(y_1,y_2)\to(x_1,x_2)} \|K(\cdot - (y_1, y_2)) - K(\cdot - (x_1, x_2))\|_{L^1+L^\infty} \le \lim_{(y_1,y_2)\to(x_1,x_2)} \|\dots\|_{L^p} = 0.$$

(B) Consider a family of kernels $\{K_j\}_j$ satisfying (3.4) for each $j \in \mathbb{N}$. Let

$$T_m(f_1, f_2)(x) = \sup_{1\le j\le m} \Big| \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_j(x, y_1, y_2)\, f_1(y_1)\, f_2(y_2)\, dy_1\, dy_2 \Big|,$$

where $m \in \mathbb{N}$, then T_m is (ε,δ)-atomic. Consequently,

$$T_*(f_1, f_2)(x) = \sup_{j\in\mathbb{N}} \Big| \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_j(x, y_1, y_2)\, f_1(y_1)\, f_2(y_2)\, dy_1\, dy_2 \Big|,$$

is (ε,δ)-atomic approximable. In general, $T_*(f_1, f_2)(x) = \sup_n |T_n(f_1, f_2)(x)|$, where T_n is (ε,δ)-atomic, is (ε,δ)-atomic approximable.

We state our main result concerning bi-sublinear atomic operators:

Theorem 3.9. *Let T be a bi-sublinear operator that is (ε,δ)-atomic approximable or iterative (ε,δ)-atomic approximable.*

(i) *Assume that for all measurable sets E_1, E_2,*

$$\big(T(\chi_{E_1}, \chi_{E_2})\big)^*(t) \le h(t; |E_1|, |E_2|), \tag{3.5}$$

where, for all s_1,$s_2 > 0$, $h(t; s_1, s_2)$ *is continuous as a function of* $t > 0$. *Then, for all* f_1, $f_2 \in L^1$ *such that* $\|f_1\|_\infty$, $\|f_2\|_\infty \leq 1$, *we have that*

$$(T(f_1, f_2))^*(t) \leq h(t; \|f_1\|_1, \|f_2\|_1). \tag{3.6}$$

(*ii*) *Let* X *be a quasi-Banach r.i. space and assume that, for all measurable sets* E_1, E_2,

$$\|T(\chi_{E_1}, \chi_{E_2})\|_X \leq D(|E_1|, |E_2|), \tag{3.7}$$

where D *is increasing in each variable and* $D(0, \cdot) = D(\cdot, 0) = 0$. *Then, for all* f_1, $f_2 \in L^1$ *such that* $\|f_1\|_\infty$, $\|f_2\|_\infty \leq 1$, *we have*

$$\|T(f_1, f_2)\|_X \lesssim D(\|f_1\|_1, \|f_2\|_1). \tag{3.8}$$

Proof. When T is iterative (ε, δ)-atomic approximable, the desired estimates follow by applying two times the sublinear case (see [6] and [7]): each time we freeze one of the variables. In the other case, we use the ideas in [6] and [7] with the appropriate changes.

First of all, let us assume that T is (ε, δ)-atomic. Let $\{(a_1^i, a_2^k)\}_{i,k}$ be a collection of δ-atoms. For every $s > 0$ we have

$$\begin{aligned}
\left(\sum_{i,k} |T(a_1^i, a_2^k)|\right)^*(s) &\leq \frac{1}{s}\int_0^s \left(\sum_{i,k} |T(a_1^i, a_2^k)|\right)^*(t)\, dt \\
&\leq \frac{1}{s}\sum_{i,k}\int_0^s (T(a_1^i, a_2^k))^*(t)\, dt \leq \max(s^{-1}, 1)\sum_{i,k}\|T(a_1^i, a_2^k)\|_{L^1+L^\infty} \\
&\leq \max(s^{-1}, 1)\, \varepsilon \sum_i \|a_1^i\|_1 \sum_k \|a_2^k\|_1.
\end{aligned} \tag{3.9}$$

Let $(f_1, f_2) \in L^1 \times L^1$ be a pair of positive functions such that $\|f_j\|_\infty \leq 1$. Given $\varepsilon > 0$, let δ be the number associated to ε by the property that T is (ε, δ)-atomic. Let $\mathcal{F}_\delta$ be any collection of pairwise disjoint cubes $\{Q^i\}_i$ such that $\cup_i Q^i = \mathbb{R}^n$, and $|Q^i| = \delta$ for every i. Given $Q^i \in \mathcal{F}_\delta$ and $j = 1$ or 2, let $f_j^i = f_j \chi_{Q^i}$. Then,

$$\int_{\mathbb{R}^n} f_j^i(x)dx \leq |Q^i|,$$

and hence, we can take a set $\tilde{Q}_j^i \subset Q^i$ (this set can be empty) such that

$$|\tilde{Q}_j^i| = \int_{\mathbb{R}^n} f_j^i(x)\, dx = \int_{Q^i} f_j(x)\, dx.$$

We define $g_j^i = f_j^i - \chi_{\tilde{Q}_j^i}$, which clearly has vanishing integral, and satisfies that

$$\|g_j^i\|_1 \leq \int_{Q^i} f_j(x)\, dx + |\tilde{Q}_j^i| = 2\int_{Q^i} f_j(x)\, dx.$$

Therefore,

$$\sum_i \|g_j^i\|_1 \leq 2\, \|f_j\|_1, \qquad \sum_i \|\chi_{\tilde{Q}_j^i}\|_1 = \sum_i |\tilde{Q}_j^i| = \|f_j\|_1. \tag{3.10}$$

Note that $f_j = \sum_i f_j^i = \sum_i g_j^i + \chi_{E_j} = A_j + \chi_{E_j}$, where $E_j = \cup_i \tilde{Q}_j^i$. Then, by sublinearity

$$\begin{aligned}|T(f_1, f_2)| &\le |T(f_1, A_2)| + |T(f_1, \chi_{E_2})| \\ &\le |T(A_1, A_2)| + |T(\chi_{E_1}, A_2)| + |T(A_1, \chi_{E_2})| + |T(\chi_{E_1}, \chi_{E_2})| \\ &\le \sum_{i,k} |T(g_1^i, g_2^k)| + \sum_{i,k} |T(\chi_{\tilde{Q}_1^i}, g_2^k)| + \sum_{i,k} |T(g_1^i, \chi_{\tilde{Q}_2^k})| + |T(\chi_{E_1}, \chi_{E_2})|\end{aligned}$$

and therefore,

$$\begin{aligned}(T(f_1, f_2))^*(t) \le \Big(\sum_{i,k} |T(g_1^i, g_2^k)|\Big)^*(\alpha_1 t) + \Big(\sum_{i,k} |T(\chi_{\tilde{Q}_1^i}, g_2^k)|\Big)^*(\alpha_2 t) \\ + \Big(\sum_{i,k} |T(g_1^i, \chi_{\tilde{Q}_2^k})|\Big)^*(\alpha_3 t) + (T(\chi_{E_1}, \chi_{E_2}))^*(\alpha_4 t), \qquad (3.11)\end{aligned}$$

for all $\alpha_j > 0$, $1 \le j \le 4$, with $\sum_{j=1}^4 \alpha_j = 1$. Let us point out that (g_1^i, g_2^k), $(\chi_{\tilde{Q}_1^i}, g_2^k)$ and $(g_1^i, \chi_{\tilde{Q}_2^k})$ are δ-atoms.

We first prove (i). Using (3.11), (3.9), (3.10), and (3.5) we have

$$\begin{aligned}(T(f_1, f_2))^*(t) &\le \Big(\sum_{j=1}^3 4\max\Big(\frac{1}{\alpha_j\, t}, 1\Big)\Big)\, \varepsilon\, \|f_1\|_1\, \|f_2\|_1 + (T(\chi_{E_1}, \chi_{E_2}))^*(\alpha_4 t) \\ &\le \Big(\sum_{j=1}^3 4\max\Big(\frac{1}{\alpha_j\, t}, 1\Big)\Big)\, \varepsilon\, \|f_1\|_1\, \|f_2\|_1 + h(\alpha_4\, t; |E_1|, |E_2|),\end{aligned}$$

and, since $|E_j| = \|f_j\|_1$ by (3.10), we obtain

$$(T(f_1, f_2))^*(t) \le \Big(\sum_{j=1}^3 4\max\Big(\frac{1}{\alpha_j t}, 1\Big)\Big)\, \varepsilon\, \|f_1\|_1\, \|f_2\|_1 + h(\alpha_4 t; \|f_1\|_1, \|f_2\|_1).$$

Letting first $\varepsilon \to 0$ and then $\alpha_4 \to 1$, we obtain the desired estimate for T.

Next we obtain (ii). We take $\alpha_1 = \alpha_2 = \alpha_3 = 1/(3\,N^2)$ and $\alpha_4 = 1 - 1/N^2$ with $N \ge 2$. Then, for $t \in (1/N, N)$ we have that $0 \le t - 1/N \le (1 - 1/N^2)\, t = \alpha_4\, t$ and

$$R_N(t) = (T(\chi_{E_1}, \chi_{E_2}))^*(\alpha_4 t)\, \chi_{(1/N,N)}(t) \le (T(\chi_{E_1}, \chi_{E_2}))^*(t - 1/N)\, \chi_{(1/N,N)}(t).$$

This yields that $R_N^*(t) \le (T(\chi_{E_1}, \chi_{E_2}))^*(t)$ for every $t > 0$. Let $\overline{X}$ be the quasi-Banach r.i. space given by the Luxemburg representation theorem such that $\|h\|_X = \|h^*\|_{\overline{X}}$. Then, using (3.7) and that $|E_j| = \|f_j\|_1$ we have

$$\|R_N\|_{\overline{X}} \le \|(T(\chi_{E_1}, \chi_{E_2}))^*\|_{\overline{X}} = \|T(\chi_{E_1}, \chi_{E_2})\|_X \le D(|E_1|, |E_2|) = D(\|f_1\|_1, \|f_2\|_1).$$

On the other hand by (3.11), (3.9) and (3.10) we obtain for every $t \in (1/N, N)$

$$\begin{aligned}(T(f_1, f_2))^*(t) \le \Big(\sum_{i,k} \dots\Big)^*(1/(3\,N^3)) + \Big(\sum_{i,k} \dots\Big)^*(1/(3\,N^3)) \\ + \Big(\sum_{i,k} \dots\Big)^*(1/(3\,N^3)) + R_N(t)\end{aligned}$$

$$\leq 24\, N^3\, \varepsilon\, \|f_1\|_1\, \|f_2\|_1 + R_N(t).$$

Therefore,

$$\begin{aligned}\|(T(f_1,f_2))^*\, \chi_{(1/N,N)}\|_{\overline{X}} &\lesssim N^3\, \varepsilon\, \|f_1\|_1\, \|f_2\|_1\, \|\chi_{(1/N,N)}\|_{\overline{X}} + \|R_N\|_{\overline{X}} \\ &\leq N^3\, \varepsilon\, \|f_1\|_1\, \|f_2\|_1\, \|\chi_{(1/N,N)}\|_{\overline{X}} + D(\|f_1\|_1, \|f_2\|_1).\end{aligned}$$

Letting first $\varepsilon \to 0$ and then $N \to \infty$, we deduce the desired estimate as a consequence of the Fatou property for $\overline{X}$ ($h_N \uparrow h$ a.e. implies $\|h_N\|_{\overline{X}} \uparrow \|h\|_{\overline{X}}$).

To finish, we consider the case when T is (ε,δ)-atomic approximable. Let $(T_n)_n$ be the corresponding sequence of (ε,δ)-atomic operators given in Definition 3.7.

To obtain (i) we observe that

$$(T_n(\chi_{E_1},\chi_{E_2}))^*(t) \leq (T(\chi_{E_1},\chi_{E_2}))^*(t) \leq h(t, |E_1|, |E_2|),$$

and hence $(T_n(f_1,f_2))^*(t) \leq h(t, \|f_1\|_1, \|f_2\|_1)$, for all pairs of positive functions (f_1,f_2) such that $\|f_j\|_\infty \leq 1$. Using $(T(f_1,f_2))^*(t) \leq \liminf_n (T_n(f_1,f_2))^*(t)$, the desired estimate for T follows at once.

To derive (ii) we notice that

$$\|T_n(\chi_{E_1},\chi_{E_2})\|_X \leq \|T(\chi_{E_1},\chi_{E_2})\|_X \leq D(|E_1|, |E_2|),$$

and we deduce that T_n satisfies (3.8). Thus we conclude the same estimate for T using that $(T(f_1,f_2))^*(t) \leq \liminf_n (T_n(f_1,f_2))^*(t)$ and the Fatou property. □

As a consequence of Theorem 3.9, we can improve Theorem 2.6 (and also Corollary 2.7) when $D(t,s) = t\, s$.

Corollary 3.10. *Let T be a bi-sublinear operator that is (ε,δ)-atomic approximable or iterative (ε,δ)-atomic approximable. Let X be quasi-Banach r.i. space. Assume that for all measurable sets E_1, E_2 with $|E_1|, |E_2| < \infty$ we have*

$$\|T(\chi_{E_1},\chi_{E_2})\|_X \leq C\, |E_1|\, |E_2|. \tag{3.12}$$

Then,

$$\|T(f_1,f_2)\|_X \leq C\, \|f_1\|_1\, \|f_2\|_1, \qquad f_1, f_2 \in L^1 \cap L^\infty,$$

and thus $T : L^1 \times L^1 \longrightarrow X$.

This gives a bilinear (and thus multilinear) version of Moon's theorem improving the result in [10], where only the case $X = L^{q,\infty}$ with $q > 0$ was considered.

3.3. **Decompositions into level sets and estimates on Orlicz spaces.** If T is an operator as in Theorem 3.3, then (3.1) implies (3.2). At this point, we make this latter condition our starting assumption. That is, from now on we will be working with sublinear operators T for which (3.2) holds. Whether this condition follows from the assumption that T is atomic or not plays no role in the arguments below. Let us emphasize that in (3.1) we only allow characteristic functions while in (3.2) a wider a class of functions is considered ($f \in L^1$ with $\|f\|_\infty \leq 1$). Notice that in Theorem 2.1, functions are decomposed as linear combinations of characteristic functions. Starting with (3.2) we can use more general decompositions: characteristic functions can be replaced by L^1-functions bounded by 1. In the following argument we will use decompositions based on the level sets of the functions. As mentioned before in Remark 3.5, only the case where X is quasi-Banach matters, since being atomic or not makes a difference.

Next we explain the general scheme that we are going to follow:

Step 0. We start from

$$\|Tf\|_X \leq C\, D(\|f\|_1), \qquad \|f\|_\infty \leq 1. \tag{3.13}$$

Step 1. Given $f \geq 0$ and an increasing sequence of non-negative numbers $\{d_k\}$ such that $d_k \longrightarrow 0$ as $k \to -\infty$ and $d_k \longrightarrow +\infty$ as $k \to +\infty$, we write

$$f = \sum_k d_k\, \tilde{f}_k, \qquad \tilde{f}_k = \frac{1}{d_k}\, f\, \chi_{\{d_{k-1} < f \leq d_k\}}.$$

Let us observe that in some cases we will take $d_k = 0$ for $k \leq -1$ and $d_0 = 1$. Thus, the summation runs for $k \geq 0$ and $\tilde{f}_0 = f\, \chi_{\{f \leq 1\}}$.

Step 2. We use (3.13) (as $\|\tilde{f}_k\|_\infty \leq 1$) and the definition of the Galb:

$$\|Tf\|_X \leq \Big\| \sum_k d_k\, |T\tilde{f}_k| \Big\|_X \lesssim \big\| \{d_k\, D(\|\tilde{f}_k\|_1)\}_k \big\|_{\mathrm{Galb}(X)} = \big\| \{A_k\}_k \big\|_{\mathrm{Galb}(X)}.$$

Step 3. We pick a non-negative function φ that it is essentially constant in the intervals $[d_{k-1}, d_k]$ and write $c_k = \varphi(d_k)$. Then, setting $a_k = \int_{d_{k-1} < f \leq d_k} f\, \varphi(f)$ we have

$$A_k = d_k\, D\Big(\frac{1}{d_k} \int_{d_{k-1} < f \leq d_k} f\Big) \approx d_k\, D\Big(\frac{1}{d_k\, c_k} \int_{d_{k-1} < f \leq d_k} f\, \varphi(f)\Big) = d_k\, D\Big(\frac{a_k}{d_k\, c_k}\Big).$$

Step 4. We show that for every non-negative sequence $\{a_k\}_k \in \ell^1$ with $\|\{a_k\}_k\|_{\ell^1} = 1$, we have

$$\Big\| \Big\{ d_k\, D\Big(\frac{a_k}{d_k\, c_k}\Big) \Big\}_k \Big\|_{\mathrm{Galb}(X)} \lesssim 1. \tag{3.14}$$

Step 5. If we are able to check all the steps in this procedure, then we will get

$$\|Tf\|_X \lesssim 1$$

for all f such that

$$1 = \sum_k a_k = \sum_k \int_{d_{k-1} < f \leq d_k} f\, \varphi(f) = \int_{\mathbb{R}^n} f\, \varphi(f).$$

Therefore, T maps L^ψ into X where L^ψ is the Orlicz-type space defined by the function $\psi(t) = t\, \varphi(t)$.

Looking at all these steps, the strategy consists in finding an appropriate function φ as in Step 3 such that the estimate in Step 4 holds. The choice of φ should depend on the sequence $\{d_k\}_k$ (as φ has to be essentially constant in the intervals defined by the sequence) and also on the function D and $\mathrm{Galb}(X)$. Motivated by the restricted estimates for the bilinear Hilbert transform we consider quasi-Banach spaces X with $\mathrm{Galb}(X)=\ell^q$ for $0<q<1$. We start with functions $D(t)=t\,(1+\log^+1/t)^\alpha$. For every $k\geq 1$ we write $\varepsilon_k=k\,(1+\log k)^{1+\varepsilon}$ with $\varepsilon>0$.

Example 3.11. Let $D(t)=t\,(1+\log^+1/t)^\alpha$ and $\mathrm{Galb}(X)=\ell^q$ with $0<q<1$. We take $d_k=2^k$ for $k\geq 1$, $d_0=1$ and $d_k=0$ for $k\leq -1$. We pick

$$\varphi(t)=(1+\log^+t)^\alpha\left((1+\log^+t)\,(1+\log^+\log^+t)^{1+\varepsilon}\right)^{\frac{1-q}{q}}$$

with $\varepsilon>0$. Then, $c_k\approx k^\alpha\left(k\,(1+\log k)^{1+\epsilon}\right)^{\frac{1-q}{q}}=k^\alpha\,\varepsilon_k^{\frac{1-q}{q}}$. We have to estimate

$$\Big\|d_k\,D\Big(\frac{a_k}{d_k\,c_k}\Big)\Big\|^q_{\mathrm{Galb}(X)}\approx D(a_0)^q+\sum_{k\geq1}2^{k\,q}\,D\Big(\frac{a_k}{2^k\,k^\alpha\,\varepsilon_k^{(1-q)/q}}\Big)^q\leq 1+\Sigma_I+\Sigma_{II},$$

where Σ_I, Σ_{II} are the corresponding sums where the indices run over the following sets

$$I=\{k\geq1:a_k\leq\varepsilon_k^{-1}\},\qquad II=\{k\geq1:a_k>\varepsilon_k^{-1}\}.$$

Then,

$$\begin{aligned}\Sigma_I&\leq\sum_{k\geq1}2^{k\,q}\,D\Big(\frac{1}{2^k\,k^\alpha\,\varepsilon_k^{1/q}}\Big)^q\leq\sum_{k\geq1}\Big(2^k\,\frac{1}{2^k\,k^\alpha\varepsilon_k^{1/q}}\,(1+\log^+2^k\,k^\alpha\varepsilon_k^{1/q})^\alpha\Big)^q\\&\lesssim\sum_{k\geq1}\frac{1}{\varepsilon_k}\lesssim1.\end{aligned}$$

Also,

$$\begin{aligned}\Sigma_{II}&\leq\Big\|\Big\{\frac{a_k}{k^\alpha\,\varepsilon_k^{(1-q)/q}}\,\Big(1+\log^+\frac{2^k\,k^\alpha\,\varepsilon_k^{(1-q)/q}}{a_k}\Big)^\alpha\Big\}_{k\in II}\Big\|^q_{\ell^q}\\&\leq\Big\|\Big\{\frac{a_k}{k^\alpha\,\varepsilon_k^{(1-q)/q}}\,\Big(1+\log^+2^k\,k^\alpha\,\varepsilon_k^{1/q}\Big)^\alpha\Big\}_{k\in II}\Big\|^q_{\ell^q}\\&\lesssim\Big\|\Big\{a_k\,\varepsilon_k^{-(1-q)/q}\Big\}_{k\in II}\Big\|^q_{\ell^q}\lesssim\sum_{k\geq1}a_k=1.\end{aligned}$$

Thus we have shown (3.14) and therefore for every $\varepsilon>0$ we obtain

$$T:L\,(\log L)^{\alpha+\frac{1-q}{q}}\,(\log\log L)^{(1+\varepsilon)\,\frac{1-q}{q}}\longrightarrow X.$$

Let us observe that taking a little bigger sequence in ℓ^1, that is, $\varepsilon_k=k^{1+\varepsilon}$ we can replace the first space by $L\,(\log L)^{\alpha+\frac{1-q}{q}\,(1+\varepsilon)}$.

Example 3.12. We proceed as in the previous example but now we choose a different sequence d_k. Let $D(t)=t\,(1+\log^+1/t)^\alpha$ and $\mathrm{Galb}(X)=\ell^q$ with $0<q<1$. We take $d_k=2^{2^k}$ for $k\geq1$, $d_0=1$ and $d_k=0$ for $k\leq-1$. We pick

$$\varphi(t)=(1+\log^+t)^\alpha\left((1+\log^+\log^+t)\,(1+\log^+\log^+\log^+t)^{1+\varepsilon}\right)^{\frac{1-q}{q}}$$

with $\beta>0$. Then, $c_k\approx2^{k\,\alpha}\,\varepsilon_k^{\frac{1-q}{q}}$. The sets I and II are the same and we estimate Σ_I and Σ_{II}:

$$\Sigma_I\leq\Big\|\Big\{2^{2^k}\,D\Big(\frac{a_k}{2^{2^k}\,2^{k\,\alpha}\,\varepsilon_k^{(1-q)/q}}\Big)\Big\}_{k\in I}\Big\|^q_{\ell^q}\leq\Big\|\Big\{2^{2^k}\,D\Big(\frac{1}{2^{2^k}\,2^{k\,\alpha}\,\varepsilon_k^{1/q}}\Big)\Big\}_{k\geq1}\Big\|^q_{\ell^q}$$

$$\lesssim \sum_{k\geq 1} \frac{1}{\varepsilon_k} \lesssim 1,$$

and

$$\Sigma_{II} \leq \Big\| \Big\{ \frac{a_k}{2^{k\,\alpha}\, \varepsilon_k^{(1-q)/q}} \Big(1 + \log^+ \frac{2^{2^k}\, 2^{k\,\alpha}\, \varepsilon_k^{(1-q)/q}}{a_k} \Big)^{\alpha} \Big\}_{k\in II} \Big\|_{\ell^q}^q \lesssim \Big\| \Big\{ a_k\, \varepsilon_k^{-\frac{1-q}{q}} \Big\}_{k\in II} \Big\|_{\ell^q}^q$$

$$\lesssim \sum_{k\geq 1} a_k = 1.$$

Thus we have shown (3.14) and therefore, for every $\varepsilon > 0$, we obtain

$$T : L\,(\log L)^{\alpha}\,(\log\log L)^{\frac{1-q}{q}}\,(\log\log\log L)^{(1+\varepsilon)\,\frac{1-q}{q}} \longrightarrow X.$$

As before, the space of origin can be replaced by $L\,(\log L)^{\alpha}\,(\log\log L)^{\frac{1-q}{q}\,(1+\varepsilon)}$ or even more by $L\,(\log L)^{\alpha+\varepsilon}$. Note that these improve what was obtained in the previous example, thus the sequence 2^{2^k} gives better estimates than 2^k.

Let us observe that taking $d_k = 2^{2^{2^k}}$ then $(1+\log^+ t)^{\alpha}$ is not essentially constant on the interval $[d_{k-1}, d_k]$. In some sense, as we have started with a restricted weak type associated with the space $L(\log L)^{\alpha}$ we should take sequences d_k for which the function $(1+\log^+ t)^{\alpha}$ is essentially constant on the intervals $[d_{k-1}, d_k]$.

In the following two examples we want to illustrate how this method behaves with respect to different logarithms. We give the final results leaving the details to the interested reader.

Example 3.13. Let $D(t) = t\,(1+\log^+ 1/t)^{\alpha}\,(1+\log^+\log^+ 1/t)^{\beta}$ and $\mathrm{Galb}(X) = \ell^q$ with $0 < q < 1$. We take $d_k = 2^k$ for $k \geq 1$, $d_0 = 1$ and $d_k = 0$ for $k \leq -1$. The ideas used before lead to the space $L\,(\log L)^{\alpha+\frac{1-q}{q}}\,(\log\log L)^{\beta+(1+\varepsilon)\,\frac{1-q}{q}}$. A better result is proved by choosing the sequence $d_k = 2^{2^k}$ for $k \geq 1$, $d_0 = 1$ and $d_k = 0$ for $k \leq -1$ in which case one gets $L\,(\log L)^{\alpha}\,(\log\log L)^{\beta+\frac{1-q}{q}}\,(\log\log\log L)^{(1+\varepsilon)\,\frac{1-q}{q}}$. Let us emphasize that as before we cannot take $d_k = 2^{2^{2^k}}$ since $(1+\log^+ t)^{\alpha}$ is not essentially constant on the interval $[d_{k-1}, d_k]$.

Example 3.14. Let $D(t) = t\,(1+\log^+\log^+ 1/t)^{\alpha}$ and $\mathrm{Galb}(X) = \ell^q$ with $0 < q < 1$. The previous ideas indicate that one should find sequences for which the function $(1+\log^+\log^+ t)^{\alpha}$ is essentially constant in the interval $[d_{k-1}, d_k]$. In this way, we take $d_k = 2^{2^{2^k}}$ for $k \geq 1$ (notice that we cannot work with $d_k = 2^{2^{2^{2^k}}}$), and then the space obtained by this method is

$$L\,(\log\log L)^{\alpha}\,(\log\log\log L)^{\frac{1-q}{q}}\,(\log\log\log\log L)^{(1+\varepsilon)\,\frac{1-q}{q}}.$$

Note that if we had taken the sequences $d_k = 2^k$, $d_k = 2^{2^k}$, we would have obtained the smaller spaces, respectively

$$L\,(\log L)^{\frac{1-q}{q}}\,(\log\log L)^{\alpha+(1+\varepsilon)\,\frac{1-q}{q}}, \qquad L\,(\log\log L)^{\alpha+\frac{1-q}{q}}\,(\log\log\log L)^{(1+\varepsilon)\,\frac{1-q}{q}}.$$

As in the case on the bilinear Hilbert transform, we also have functions of the form $D(t) = t^{1/p}\,(1+\log^+ 1/t)^{\alpha}$ with $1 < p < \infty$, and we investigate what spaces one obtains via this method. Note that this function is associated with an Orlicz space near L^p (indeed $L^p\,(\log L)^{\alpha\,p}$), thus we look for sequences d_k for which t^{p-1} is essentially constant in the intervals $[d_{k-1}, d_k]$, that is, $d_k \approx 2^k$. Notice that in the previous examples the Orlicz functions satisfy $\psi(t) = t$ for $t \le 1$. Thus, it was not necessary to decompose the function $f\,\chi_{\{f\le 1\}}$ into level sets. Here, as the Orlicz function is going to be near t^p, we investigate whether decomposing this function leads or not to a better estimate.

We start with the simpler case $\alpha = 0$.

Example 3.15. Let $D(t) = t^{1/p}$ with $1 < p < \infty$ and $\mathrm{Galb}(X) = \ell^q$ with $0 < q < 1$. We first take $d_k = 2^k$ for $k \ge 1$, $d_0 = 1$ and $d_k = 0$ for $k \ge -1$. We take $\varphi(t) = 1$ for $t \ge 1$ and $\varphi(t) = t^{p-1}\,(1+\log^+ t)^{p/q-1}\,(1+\log^+\log^+ t)^{(p/q-1)\,(1+\varepsilon)}$. Then $c_k \approx 2^{k\,(p-1)}\,\varepsilon_k^{p/q-1}$. We have to estimate

$$\Big\| d_k\, D\Big(\frac{a_k}{d_k\, c_k}\Big)\Big\|_{\mathrm{Galb}(X)}^q \approx D(a_0)^q + \sum_{k\ge1} 2^{k\,q}\, D\Big(\frac{a_k}{2^{k\,p}\varepsilon_k^{p/q-1}}\Big)^q \le 1 + \sum_{k\ge1} \frac{a_k^{q/p}}{\varepsilon_k^{1-q/p}}$$

and we split the sum in the right-hand side as $\Sigma_I + \Sigma_{II}$ (with the same definition of I and II). Note that we trivially have that $\Sigma_I \le \sum_{k\ge1} \varepsilon_k^{-1} \lesssim 1$. On the other hand, since $0 < q < 1 < p$, we have that $\Sigma_{II} \le \sum_{k\ge1} a_k = 1$. Then we obtain that T maps L^ψ into X where $\psi(t) = t$ for $t \le 1$ and

$$\psi(t) = t^p\,(1+\log^+ t)^{p/q-1}\,(1+\log^+\log^+ t)^{(p/q-1)\,(1+\varepsilon)}$$

for $t \ge 1$.

Next, we take $d_k = 2^k$ for every $k \in \mathbb{Z}$. Consider the function

$$\widetilde{\varphi}(t) = t^{p-1}\,(1+|\log t|)^{p/q-1}\,(1+\log^+|\log t|)^{(p/q-1)\,(1+\varepsilon)}$$

and then $c_k \approx 2^{k\,(p-1)}\,\varepsilon_{|k|}^{(p/q-1)}$ for $k \ne 0$, and $c_0 = 1$. Since $\widetilde{\varphi}(t) = \varphi(t)$ for $t \ge 1$, we only have to estimate the terms $k \le 0$. Proceeding as before (now we compare a_k with $1/\varepsilon_{|k|}$) we conclude that

$$\sum_{k\le0} d_k^q\, D\Big(\frac{a_k}{d_k\, c_k}\Big)^q \lesssim 1 + \sum_{k\le-1} \frac{a_k^{q/p}}{\varepsilon_{|k|}^{1-q/p}} \lesssim 1.$$

Then we obtain that T maps $L^{\widetilde{\psi}}$ into X where

$$\widetilde{\psi}(t) = t^p\,(1+|\log t|)^{p/q-1}\,(1+\log^+|\log t|)^{(p/q-1)\,(1+\varepsilon)}.$$

Let us observe that $\widetilde{\psi}(t) \le \psi(t) = t$ for $t \le 1$ (as $p > 1$) and also that $\widetilde{\psi}(t) = \psi(t)$ for $t \ge 1$. Thus, $L^\psi \subset L^{\widetilde{\psi}}$. Therefore, decomposing $f\,\chi_{\{f\le1\}}$ leads to a better estimate.

Example 3.16. Let $D(t) = t^{1/p}\,(1+\log^+ 1/t)^\alpha$ with $1 < p < \infty$ and $\mathrm{Galb}(X) = \ell^q$ with $0 < q < 1$. We take $d_k = 2^k$ for every $k \in \mathbb{Z}$. Consider the function $\widetilde{\varphi}(t) = t^{p-1}\,(1+\log^+ t)^\alpha\,(1+|\log t|)^{p/q-1}\,(1+\log^+|\log t|)^{(p/q-1)\,(1+\varepsilon)}$ and then $c_k \approx 2^{k\,(p-1)}\,\max(1,k)^\alpha\,\varepsilon_{|k|}^{(p/q-1)}$ for $k \ne 0$, and $c_0 = 1$. The same ideas allow us to show that T is bounded from L^ψ into X where

$$\psi(t) = t^p\,(1+\log^+ t)^\alpha\,(1+|\log t|)^{p/q-1}\,(1+\log^+|\log t|)^{(p/q-1)\,(1+\varepsilon)}.$$

Remark 3.17. We would like to emphasize that the method used in this section, of picking the sequence $d_k = a^k$, $a > 1$, cannot improve Theorem 2.1. Indeed, going back to Step 2, we need to estimate $\|\{A_k\}_k\|_{\mathrm{Galb}(X)}$. Notice that if a is big enough,

$$\begin{aligned}\|\{A_k\}_k\|_{\mathrm{Galb}(X)} =& \big\|\{a^k\, D(\|\tilde f_k\|_1)\}_k\big\|_{\mathrm{Galb}(X)} \approx \big\|\{a^k\, D\big(|\{|f| \approx a^k\}|\big)\}_k\big\|_{\mathrm{Galb}(X)} \\ &\approx \big\|\{a^k\, D\big(\mu_f(a^k)\big)\}_k\big\|_{\mathrm{Galb}(X)}.\end{aligned}$$

We observe that the last quantity is a discretized version of the norms appearing in Theorem 2.1.

Therefore, the approach developed in this section becomes meaningful when d_k growths faster (when $d_k = 2^{2^k}, 2^{2^k}, \dots$ the previous quantities are no longer comparable since $d_k \not\approx d_{k+1}$). This shows that the spaces obtained in Examples 3.15, 3.16 are worse than the ones that follow from Theorem 2.1. We will use this when working with the bilinear Hilbert transform.

3.3.1. *The multi-variable case.* As observed before, having some extra information about the operator leads us, in some cases, to better estimates. Thus we will study different cases for which the previous arguments in the linear case can be also exploited. For simplicity we first consider the case where D can be broken up into two functions. We start with a bi-sublinear operator satisfying

$$\|T(f,g)\|_X \le C\, D_1(\|f\|_1)\, D_2(\|g\|_1), \qquad \|f\|_\infty \le 1, \quad \|g\|_\infty \le 1.$$

This occurs when T is (ε,δ)-atomic approximable or when it is iterative (ε,δ)-atomic approximable. Once we have the last inequality we will not use these properties anymore. For these operators we can freeze one of the variables and work with the other one. Thus, there is no difference with the 1-sublinear case considered before.

In the case general case where D is not split we have to work with two sequences at the same time, one for each variable. We take $\{d_k\}_k$ and $c_k = \varphi_1(d_k)$ with φ_1 essentially constant in the intervals $[d_{k-1}, d_k]$. This sequence is related to the function f. For the function g we take δ_j and $\eta_j = \varphi_2(\delta_j)$ with φ_2 essentially constant in the intervals $[\delta_{j-1}, \delta_j]$. We define A_k, a_k and B_j, b_j as in Step 3 (A_k is for f, d_k, c_k; B_j is for g, δ_j, η_j). Everything reduces to show the following analog of Step 4: for all non-negative sequences $\{a_k\}_k$, $\{b_j\}_j \in \ell^1$ with $\|\{a_k\}_k\|_{\ell^1} = \|\{b_j\}_j\|_{\ell^1} = 1$ we have

$$\Big\|\Big\{d_k\,\delta_j\, D\Big(\frac{a_k}{d_k\, c_k}, \frac{b_j}{\delta_j\, \eta_j}\Big)\Big\}_{k,j}\Big\|_{\mathrm{Galb}(X)} \lesssim 1. \tag{3.15}$$

If we are able to show this, we obtain that T maps $L^{\psi_1} \times L^{\psi_2}$ into X, where $\psi_1(t) = t\,\varphi_1(t)$ and $\psi_2(t) = t\,\varphi_2(t)$.

4. The Bilinear Hilbert Transform

We start with the basic estimate proved in [3] (see also [4]):

$$\sup_t \Phi(t)\, H\big(\chi_{E_1}, \chi_{E_2}\big)^*(t) \lesssim D(|E_1|, |E_2|), \tag{4.1}$$

where

$$\Phi(t) = t^{3/2}\,(1+\log^+ t)^{-2}, \qquad D(s,t) = \big(s\,t\min(s,t)\big)^{1/2}\Big(1 + \log^+ \frac{1}{s\,t\min(s,t)}\Big)^2.$$

Notice that for any $\alpha, \beta \in [0,1]$ with $\alpha + \beta = 1$, we have $D(s,t) \leq D_1(s)\, D_2(t)$ where

$$D_1(s) = s^{\frac{1+\alpha}{2}} \left(1 + \log^+ \frac{1}{s}\right)^2, \qquad D_2(t) = t^{\frac{1+\beta}{2}} \left(1 + \log^+ \frac{1}{t}\right)^2.$$

We define X to be the space given by the quasi-norm

$$\|f\|_X = \sup_{t>0} \Phi(t)\, f^*(t).$$

It is known (see [7]) that $\mathrm{Galb}(X) = \ell^{\frac{2}{3}}$.

4.1. **Estimates on Lorentz spaces.** Applying Corollary 2.7 we obtain

$$\|H(f,g)\|_X \lesssim \|f\|_{\Lambda^{2/3}(dD_1^{2/3})}\, \|g\|_{\Lambda^{2/3}(dD_2^{2/3})}.$$

It remains to identify these Λ-Lorentz spaces. We have:

$$\begin{aligned}\|f\|^{2/3}_{\Lambda^{2/3}(dD_1^{2/3})} &= \int_0^\infty f^*(t)^{2/3}\, dD_1^{2/3}(t) \approx \int_0^\infty \left(f^*(t)\, t^{\frac{1+\alpha}{2}} \left(1 + \log^+ \frac{1}{t}\right)^2\right)^{\frac{2}{3}} \frac{dt}{t} \\ &= \|f\|^{\frac{2}{3}}_{L^{\frac{2}{(1+\alpha)},\frac{2}{3}}(\log L)^{\frac{4}{3}}}.\end{aligned}$$

Considering the extreme cases $\alpha = 1$ and $\beta = 0$, or vice versa, we obtain

$$\begin{aligned} H &: L^{1,\frac{2}{3}}(\log L)^{\frac{4}{3}} \times L^{2,\frac{2}{3}}(\log L)^{\frac{4}{3}} \longrightarrow X, \\ H &: L^{2,\frac{2}{3}}(\log L)^{\frac{4}{3}} \times L^{1,\frac{2}{3}}(\log L)^{\frac{4}{3}} \longrightarrow X. \end{aligned}$$

We see below that in the extreme case $\alpha = 1$ and $\beta = 0$ the previous estimate can be "improved" exploiting the fact that the bilinear Hilbert transform is atomic.

We can also use Theorem 2.6 with the original function D and then

$$\begin{aligned}\|H(f,g)\|_X &\lesssim \int_0^\infty\!\!\int_0^\infty s^{2/3}\, t^{2/3}\, D\big(\mu_f(s), \mu_g(t)\big)^{2/3} \frac{ds\, dt}{s\, t} \\ &\lesssim \int_0^\infty\!\!\int_0^\infty f^*(s)^{2/3}\, g^*(t)^{2/3}\, dD^{2/3}(s,t),\end{aligned}$$

with

$$D(s,t) = \big(s\, t \min(s,t)\big)^{1/2} \left(1 + \log^+ \frac{1}{s\, t \min(s,t)}\right)^2.$$

As observed in Remark 2.8, here we do not lose any information in the following sense:

$$\|H(\chi_E, \chi_F)\|_X \lesssim D(|E|,|F|), \quad |E|,|F| < \infty \qquad \Longleftrightarrow \qquad H : \vec{\Lambda}^{2/3}(dD^{2/3}) \longrightarrow X.$$

4.2. **Atomicity and estimates on Orlicz spaces.** We show that the following truncations of H

$$H_N(f,g) = \int_{1/N<|t|<N} f(x-t)\, g(x+t)\, \frac{dt}{t} = \int_{\mathbb{R}} f(x-t)\, g(x+t)\, k_N(t)\, dt$$

are iterative (ε,δ)-atomic. Let $g \in L^1$ be such that $\|g\|_\infty \leq 1$ and consider the 1-linear operator $T_N f$ defined by $T_N f(x) = H_N(f,g)$. We obtain that T_N is (ε,δ)-atomic (the other case in which f is frozen can be obtained in the same manner).

We write T_N in the following way,

$$T_N f(x) = \int_{\mathbb{R}} f(t)\, g(2\,x-t)\, k_N(x-t)\, dt = \int_{\mathbb{R}} f(t)\, K_N(x,t)\, dt.$$

Given $\varepsilon > 0$, let a be a δ-atom with $\delta > 0$ to be chosen. Then $\operatorname{supp} a \subset I_0$ for some interval I_0 with $|I_0| \le \delta$. Let t_0 be the center of I_0.

We first show that there exists $\delta = \delta(\varepsilon, g, N)$ such that

$$\big\|K_N(\cdot, s) - K_N(\cdot, s_0)\big\|_{L^1} \le \varepsilon, \qquad \text{for all } s,\ s_0 \in \mathbb{R} \text{ with } |s-s_0| < \delta/2. \tag{4.2}$$

Using that $\|g\|_\infty \le 1$ it follows that

$$\begin{aligned} \big\|K_N(\cdot, s) - K_N(\cdot, s_0)\big\|_{L^1} &\le \int_{\mathbb{R}} |g(2\,x-s)\, k_N(x-s) - g(2\,x-s_0)\, k_N(x-s_0)|\, dx \\ &\le \int_{\mathbb{R}} |k_N(x-s)|\, |g(2\,x-s) - g(2\,x-s_0)|\, dx \\ &\qquad + \int_{\mathbb{R}} |g(2\,x-s_0)|\, |k_N(x-s) - k_N(x-s_0)|\, dx \\ &\le N \int_{\mathbb{R}} |g(x+(s-s_0)) - g(x)|\, dx + \int_{\mathbb{R}} |k_N(x+(s-s_0)) - k_N(x)|\, dx. \end{aligned}$$

Thus, since $g,\ k_N \in L^1(\mathbb{R})$, using properties of the translation operator in $L^1(\mathbb{R})$, there exists $\delta = \delta(\varepsilon, g, N)$ such that for every $|\Delta| < \delta$

$$N \int_{\mathbb{R}} |g(x+\Delta) - g(x)|\, dx + \int_{\mathbb{R}} |k_N(x+\Delta) - k_N(x)|\, dx \le N\, \frac{\varepsilon}{2\,N} + \frac{\varepsilon}{2} = \varepsilon.$$

Applying this with $\Delta = s - s_0$, we obtain (4.2). In this way, using that a has vanishing integral and (4.2), we conclude that

$$\begin{aligned} \|T_N a\|_{L^1+L^\infty} \le \|T_N a\|_1 &\le \Big\| \int_{\mathbb{R}} a(t)\, K_N(\cdot, t)\, dt \Big\|_1 \\ &= \Big\| \int_{\mathbb{R}} a(t)\, (K_N(\cdot, t) - K_N(\cdot, t_0))\, dt \Big\|_1 \\ &\le \int_{|t-t_0|<\delta/2} |a(t)|\, \big\|K_N(\cdot, t) - K_N(\cdot, t_0)\big\|_{L^1}\, dt \le \varepsilon \|a\|_1. \end{aligned}$$

Therefore, we have shown that T_N is (ε, δ)-atomic and H_N is iterative (ε, δ)-atomic. We observe that H_N satisfies (4.1) uniformly in N. Thus, by (ii) in Theorem 3.9 we conclude that for all $\|f\|_\infty \le 1$, $\|g\|_\infty \le 1$,

$$\|H_N(f,g)\|_X = \sup_t \Phi(t)\, H_N\big(f,g\big)^*(t) \lesssim D(\|f\|_1, \|g\|_1),$$

where the constants involved are uniform in N, and

$$\Phi(t) = t^{3/2}\, (1+\log^+ t)^{-2}, \qquad D(s,t) = \big(s\, t \min(s,t)\big)^{1/2} \Big(1 + \log^+ \frac{1}{s\, t \min(s,t)}\Big)^2.$$

Notice that for any $\alpha, \beta \in [0,1]$ with $\alpha + \beta = 1$, we have $D(s,t) \le D_1(s)\, D_2(t)$ where

$$D_1(s) = s^{\frac{1+\alpha}{2}} \Big(1 + \log^+ \frac{1}{s}\Big)^2, \qquad D_2(t) = t^{\frac{1+\beta}{2}} \Big(1 + \log^+ \frac{1}{t}\Big)^2.$$

As mentioned before $\mathrm{Galb}(X)=\ell^{\frac{2}{3}}$. We have already observed that the method presented in Section 3.3 is useful when $D_i(t)$ is of the form $t\,(1+\log^+ t)^2$. Hence, we fix $\alpha=1$, $\beta=0$ and then

$$D_1(s)=s\left(1+\log^+\frac{1}{s}\right)^2,\qquad D_2(t)=t^{\frac{1}{2}}\left(1+\log^+\frac{1}{t}\right)^2.$$

Working with the first variable and applying Example 3.12 with $q=2/3$ and $\alpha=2$, we deduce that the domain space for f is $L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}}\,(\log\log\log L)^{\frac{1}{2}+\varepsilon}$ for any $\varepsilon>0$. We can take smaller spaces such as $L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}+\varepsilon}$ or $L\,(\log L)^{2+\varepsilon}$ for any $\varepsilon>0$.

For the other variable, we use the non-atomic approach and obtain that the domain space is $L^{2,\frac{2}{3}}(\log L)^{4/3}$. Thus, using the symmetry of the problem, we have

$$H_N: L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}}\,(\log\log\log L)^{\frac{1}{2}+\varepsilon}\times L^{2,\frac{2}{3}}\,(\log L)^{\frac{4}{3}}\longrightarrow X,$$
$$H_N: L^{2,\frac{2}{3}}\,(\log L)^{\frac{4}{3}}\times L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}}\,(\log\log\log L)^{\frac{1}{2}+\varepsilon}\longrightarrow X.$$

From here one can interpolate by the complex method to conclude some other estimates. Notice that all these estimates are uniform in N.

Next we are going to show how to derive these estimates for H. By density (in the domain spaces), it suffices to consider Schwartz functions f, g. In that case we have $\lim_{N\to\infty}H_N(f,g)=H(f,g)$ a.e. and consequently $H\big(f,g\big)^*\le\liminf_{N\to\infty}H_N\big(f,g\big)^*$. Then, for any $0<t<\infty$ we have

$$\begin{aligned}\Phi(t)\,H\big(f,g\big)^*(t)&\le\liminf_{N\to\infty}\Phi(t)\,H_N\big(f,g\big)^*(t)\le\liminf_{N\to\infty}\sup_t\Phi(t)\,H_N\big(f,g\big)^*(t)\\&=\liminf_{N\to\infty}\|H_N(f,g)\|_X.\end{aligned}$$

Taking the supremum for $0<t<\infty$ we conclude that

$$\|H(f,g)\|_X\le\liminf_{N\to\infty}\|H_N(f,g)\|_X.$$

This, the uniform estimates obtained before for H_N and a standard density argument lead us to

$$H: L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}}\,(\log\log\log L)^{\frac{1}{2}+\varepsilon}\times L^{2,\frac{2}{3}}\,(\log L)^{\frac{4}{3}}\longrightarrow X,$$
$$H: L^{2,\frac{2}{3}}\,(\log L)^{\frac{4}{3}}\times L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}}\,(\log\log\log L)^{\frac{1}{2}+\varepsilon}\longrightarrow X.$$

We finish this section by comparing the different spaces that we have obtained using the two approaches. When $\alpha=1$ and $\beta=0$, the two methods have led us to the following spaces

$$X_1=L^{1,\frac{2}{3}}\,(\log L)^{\frac{4}{3}},\qquad X_2=L\,(\log L)^2\,(\log\log L)^{\frac{1}{2}}\,(\log\log\log L)^{\frac{1}{2}+\varepsilon}.$$

We see that X_1, X_2 are not comparable. Our first function is given by

$$h^*(t)=\frac{1}{t\,(1+\log^+1/t)^{7/2}}\,\chi_{(0,e^{-e^e})}(t).$$

Then,

$$\|h^*\|_{X_1}^{\frac{2}{3}}=\int_0^\infty\big(h^*(t)\,t\,\big(1+\log^+1/t\big)^2\big)^{\frac{2}{3}}\,\frac{dt}{t}=\int_0^{e^{-e^e}}\frac{1}{(1+\log 1/t)}\,\frac{dt}{t}=\infty.$$

On the other hand, we have that $\widetilde{X}_2 = L\,(\log L)^{9/4} \hookrightarrow X_2$ and therefore

$$\|h^*\|_{X_2} \le \|h^*\|_{\widetilde{X}_2} = \int_0^\infty h^*(t)\,t\,\big(1+\log^+ 1/t\big)^{9/4}\frac{dt}{t}$$
$$= \int_0^{e^{-e^e}} \frac{1}{(1+\log 1/t)^{5/4}}\,\frac{dt}{t} < \infty.$$

Next, we consider a second function

$$h^*(t) = \sum_{j=1}^\infty A_j\,\chi_{(a_{j+1},a_j)}(t), \qquad A_j = e^{e^{j^4}}\,j^{-3}, \qquad a_j = e^{-(e^{j^4}+2\,j^4)}.$$

Notice that A_j is increasing and a_j decreasing. We set $m(t) = t\,(1+\log^+ 1/t)^2$ and notice that

$$m(a_j) \approx e^{-(e^{j^4}+2\,j^4)}\,(e^{j^4}+2\,j^4)^2 \approx e^{-e^{j^4}} = A_j^{-1}\,j^{-3}$$

Then,

$$\|h^*\|_{X_1}^{\frac{2}{3}} = \sum_{j=1}^\infty A_j^{\frac{2}{3}} \int_{a_{j+1}}^{a_j} m(t)^{\frac{2}{3}}\,\frac{dt}{t} \lesssim \sum_{j=1}^\infty A_j^{\frac{2}{3}}\,m(a_j)^{\frac{2}{3}} \approx \sum_{j=1}^\infty \frac{1}{j^2} < \infty.$$

On the other hand, let us observe that $X_2 \hookrightarrow \widehat{X}_2 = L\,(\log L)^2\,(\log\log L)^{1/2}$. We write $\varphi(t) = m(t)\,(1+\log^+\log^+ 1/t)^{1/2}$. Observe that $a_j \ge e\,a_{j+1}$ and

$$\int_{a_{j+1}}^{a_j} \varphi(t)\,\frac{dt}{t} \ge \int_{a_j/e}^{a_j} \varphi(t)\,\frac{dt}{t} \ge \varphi(a_j/e) \approx m(a_j)\,(\log^+\log^+ a_j^{-1})^{1/2} \approx A_j^{-1}\,j^{-1}.$$

Therefore,

$$\|h^*\|_{X_2} \ge \|h^*\|_{\widehat{X}_2} = \sum_{j=1}^\infty A_j \int_{a_{j+1}}^{a_j} \varphi(t)\,\frac{dt}{t} \gtrsim \sum_{j=1}^\infty \frac{1}{j} = \infty.$$

These two examples show that the symmetric difference of X_1 and X_2 is nonempty and hence the two approaches developed in the present paper give independent estimates. Consequently, combining both methods, we obtain estimates for functions in the larger space $X_1 + X_2$, that is, the bilinear Hilbert transform H satisfies

$$H : (X_1 + X_2) \times L^{2,\frac{2}{3}}\,(\log L)^{\frac{4}{3}} \longrightarrow X, \qquad H : L^{2,\frac{2}{3}}\,(\log L)^{\frac{4}{3}} \times (X_1 + X_2) \longrightarrow X.$$

References

[1] N.Y. Antonov, *Convergence of Fourier series*, Proceedings of the XXth Workshop on Function Theory (Moscow, 1995), East J. Approx. **2** (1996), 187–196.

[2] C. Bennett and R.C. Sharpley, *Interpolation of Operators*, Pure and Appl. Math. **129**, Academic Press, 1988.

[3] D. Bilyk and L. Grafakos, *Distributional estimates for the bilinear Hilbert transform*, Journal of Geom. Anal. **16** (2006), no. 4, 563–584.

[4] D. Bilyk and L. Grafakos, *A new way of looking at distributional estimates; applications for the bilinear Hilbert transform*, Proceedings of the 7th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2004), Collect. Math. 2006, 141–169.

[5] L. Carleson, *Convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.

[6] M.J. Carro, *From restricted weak type to strong type estimates*, J. London Math. **70** (2004), 750–762.

[7] M.J. Carro, L. Colzani, and G. Sinnamon, *From restricted weak type to strong type estimates on quasi-Banach rearrangement invariant spaces*, Studia Math. **182** (2007), no. 1, 1–27.

[8] M.J. Carro and J. Soria, *Weighted Lorentz spaces and the Hardy operator*, J. Funct. Anal. **112** (1993), 480–494.

[9] L. Grafakos, J.M. Martell, and F. Soria, *Weighted norm inequalities for maximally modulated operators*, Math. Ann. **331** (2005), no. 2, 359–394.

[10] L. Grafakos and M. Mastyło, *Restricted weak type versus weak type,* Proc. Amer. Math. Soc. **133** (2005), 1075–1081.

[11] R. Hunt, *On the convergence of Fourier series*, Orthogonal Expansions and Their Continuous Analogues (Edwardsville, Ill., 1967), 235–255, D. T. Haimo (ed.), Southern Illinois Univ. Press, Carbondale IL, 1968.

[12] K. H. Moon, *On restricted weak type* $(1, 1)$, Proc. Amer. Math. Soc. **42** (1974), 148–152.

[13] P. Sjölin and F. Soria, *Remarks on a theorem by N.Y. Antonov*, Studia Math. **158** (2003), 79–97.

[14] F. Soria, *On an extrapolation theorem of Carleson-Sjölin with applications to a.e. convergence of Fourier series*, Studia Math. **94** (1989), 235–244.

[15] P. Turpin, *Convexités dans les espaces vectoriels topologiques généraux*, Dissertationes Math. **131** (1976).

María Jesús Carro, Departament de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, E-08071 Barcelona, Spain
E-mail address: carro@ub.edu

Loukas Grafakos, Department of Mathematics, University of Missouri, Columbia, MO 65211, USA
E-mail address: loukas@math.missouri.edu

José María Martell, Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM, Consejo Superior de Investigaciones Científicas, C/ Serrano 121, E-28006 Madrid, Spain
E-mail address: chema.martell@uam.es

Fernando Soria, Departamento de Matemáticas, C-XV, Universidad Autónoma de Madrid, E-28049 Madrid, Spain
E-mail address: fernando.soria@uam.es