Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVIII (2018), 1443-1458

Positive sparse domination of variational Carleson operators

FRANCESCO DI PLINIO, YEN Q. DO AND GENNADY N. URALTSEV

Abstract. Due to its nonlocal nature, the r-variation norm Carleson operator C_r does not yield to the sparse domination techniques of Lerner [15, 17], Di Plinio and Lerner [6], Lacey [14]. We overcome this difficulty and prove that the dual form of C_r can be dominated by a positive sparse form involving L^p averages. Our result strengthens the L^p-estimates by Oberlin *et al*. [18]. As a corollary, we obtain quantitative weighted norm inequalities improving the results in [8] by Do and Lacey. Our proof relies on the localized outer L^p-embeddings of Di Plinio and Ou [7] and Uraltsev [19].

Mathematics Subject Classification (2010): 42B20 (primary); 42B25 (secondary).

1. Introduction and main results

The technique of controlling Calderón-Zygmund singular integrals, which are *a priori* non-local, by localized positive sparse operators has recently emerged as a leading trend in Euclidean Harmonic Analysis. We briefly review the advancements which are most relevant for the present article and postpone further references to the body of the introduction. The original domination in norm result of [17] for Calderón-Zygmund operators has since been upgraded to a *pointwise* positive sparse domination by Conde and Rey [2] and Lerner and Nazarov [16], and later by Lacey [14] by means of an inspiring stopping time argument forgoing local mean oscillation. Lacey's approach was further clarified in [15], resulting in the following principle: if T is a sub-linear operator of weak-type (p, p), and in addition the maximal operator

$$f \mapsto \sup_{Q\subset\mathbb{R} \text{ interval}} \left\| T(f\mathbf{1}_{\mathbb{R}\setminus 3Q})\right\|_{L^\infty(Q)} \mathbf{1}_Q, \tag{1.1}$$

F. Di Plinio was partially supported by the National Science Foundation under the grant NSF-DMS-1500449 and NSF-DMS-1650810. Y. Do was partially supported by the National Science Foundation under the grant NSF-DMS-1521293.

Received December 21, 2016; accepted April 5, 2017.
Published online July 2018.

embodying the non-locality of T, is of weak-type (s,s), for some $1 \leq p \leq s < \infty$, then T is pointwise dominated by a positive sparse operator involving L^s averages of f.

The principle (1.1) extends to certain modulated singular integrals. Of interest for us is the maximal partial Fourier transform

$$Cf(x) = \sup_N \left| \int_{\infty}^{N} \widehat{f}(\xi)\, \mathrm{e}^{ix\xi} \mathrm{d}\xi \right|,$$

also known as Carleson's operator on the real line. The crux of the matter is that (1.1) follows for $T = C$ from its representation as a maximally modulated Hilbert transform, a fact already exploited in the classical weighted norm inequalities for C by Hunt and Young [13], and in the more recent work [12]. Together with sharp forms of the Carleson-Hunt theorem near the endpoint $p = 1$ [5] this allows, as observed by the first author and Lerner in [6], the domination of C by sparse operators, and thus leads to sharp weighted norm inequalities for C.

In this article we consider the r-variation norm Carleson operator, which is defined for Schwartz functions on the real line as

$$C_r f(x) = \sup_{N \in \mathbb{N}} \sup_{\xi_0 < \cdots < \xi_N} \left(\sum_{j=1}^{N} \left| \int_{\xi_{j-1}}^{\xi_j} \widehat{f}(\xi)\, \mathrm{e}^{ix\xi} \mathrm{d}\xi \right|^r \right)^{1/r}.$$

The importance of C_r is revealed by the transference principle, presented in [18, Appendix B], which shows how r-variational convergence of the Fourier series of $f \in L^p(\mathbb{T}; w)$ for a weight w on the torus $\mathbb{T}$ follows from $L^p(\mathbb{R}; w)$-estimates for the sub-linear operator C_r. Values of interest for r are $2 < r < \infty$. Indeed the main result of [18] is that in this range, C_r maps into L^p whenever $p > r'$, while no L^p-estimates hold for variation exponents $r \leq 2$. Unlike the Carleson operator, its variation norm counterpart C_r does not have an explicit kernel form and thus fails to yield to Hunt-Young type techniques. The same essential difficulty is met in the search for L^q-bounds for the nonlocal maximal function (1.1) when $T = C_r$. Therefore, the approach via (1.1) does not seem to be applicable to C_r. In the series [8,9], the second author and Lacey circumvented this issue through a direct proof of A_p-weighted inequalities for C_r and its Walsh analogue, based on weighted phase plane analysis.

The main result of the present article is that a sparse domination principle for C_r holds in spite of the difficulties described above. More precisely, we sharply dominate the dual form to the r-variational Carleson operator C_r by a single positive sparse form involving L^p averages, leading to an effortless strengthening of the weighted theory of [8]. Our argument abandons (1.1) in favor of a stopping time construction, relying on the localized Carleson embeddings for suitably modified wave packet transforms of [7] by the first author and Yumeng Ou, and [19] by the third author. In particular, our technique requires no *a priori* weak-type information on the operator T. A similar approach was employed by Culiuc, Ou and the first author in [4] in the proof of a sparse domination principle for the family of modulation invariant multi-linear multipliers whose paradigm is the bilinear Hilbert transforms. Interestingly, unlike [4], our construction of the sparse collection in Section 4 seems to be the first in literature which does not make any use of dyadic grids.

We believe that intrinsic sparse domination can prove useful in the study of other classes of multi-linear operators lying way beyond the scope of Calderón-Zygmund theory, such as the iterated Fourier integrals of [10] and the sub-dyadic multipliers of [1].

To formulate our main theorem, we recall the notation

$$\langle f\rangle_{I,p} := \left(\frac{1}{|I|}\int |f|^p \,\mathrm{d}x\right)^{\frac{1}{p}}, \qquad 1 \le p < \infty,$$

where $I \subset \mathbb{R}$ is any interval, and the notion of a sparse collection of intervals. We say that the countable collection of intervals $I \in \mathcal{S}$ is η-sparse for some $0 < \eta \le 1$ if there exists a choice of measurable sets $\{E_I : I \in \mathbf{S}\}$ such that

$$E_I \subset I, \qquad |E_I| \ge \eta|I|, \qquad E_I \cap E_J = \varnothing \quad \forall I, J \in \mathcal{S},\ I \ne J.$$

Theorem 1.1. *Let $2 < r < \infty$ and $p > r'$. Given $f, g \in \mathcal{C}_0^\infty(\mathbb{R})$ there exists a sparse collection $\mathcal{S} = \mathcal{S}(f, g, p)$ and an absolute constant $K = K(p)$ such that*

$$|\langle C_r f, g\rangle| \le K(p)\sum_{I\in\mathcal{S}} |I|\langle f\rangle_{I,p}\langle g\rangle_{I,1}. \tag{1.2}$$

A corollary of Theorem 1.1 is that C_r extends to a bounded sub-linear operator on $L^q(\mathbb{R})$ whenever $q > r'$. As a matter of fact, let us fix $q \in (r', \infty]$, and choose $p \in (r', q)$. Denoting by

$$\mathrm{M}_p f(x) = \sup_{I\ni x}\langle f\rangle_{I,p}$$

the p-th Hardy-Littlewood maximal function, the estimate of Theorem 1.1 and the fact that $\mathcal{S}$ is sparse yields

$$\begin{aligned}|\langle C_r f, g\rangle| &\lesssim \sum_{I\in\mathcal{S}} |E_I|\langle f\rangle_{I,p}\langle g\rangle_{I,1} \le \langle \mathrm{M}_p f, \mathrm{M}_1 g\rangle \\ &\lesssim \|\mathrm{M}_p f\|_q \|\mathrm{M}_1 g\|_{q'} \lesssim \|f\|_q\|g\|_{q'}.\end{aligned}$$

Bounds on L^q for C_r were first proved in [18], where it is also shown that the restriction $q > r'$ is necessary, whence no sparse domination of the type occurring in Theorem 1.1 will hold for $p < r'$. We can thus claim that Theorem 1.1 is sharp, short of the endpoint $p = r'$. In fact, sparse domination as in (1.2) also entails $C_r : L^p(\mathbb{R}) \to L^{p,\infty}(\mathbb{R})$. Such an estimate is currently unknown for $p = r'$.

However, Theorem 1.1 yields much more precise information than mere L^q-boundedness. In particular, we obtain precisely quantified weighted norm inequalities for C_r. Recall the definition of the A_t constant of a locally integrable nonnegative function w as

$$[w]_{A_t} := \begin{cases} \sup_{I\subset\mathbb{R}} \langle w\rangle_{I,1}\left\langle w^{\frac{1}{1-t}}\right\rangle_{I,1}^{t-1} & 1 < t < \infty \\ \inf\left\{A : \mathrm{M}w(x) \le Aw(x) \text{ for a.e. } x\right\} & t = 1.\end{cases}$$

Theorem 1.2. *Let* $2 < r < \infty$ *and* $q > r'$ *be fixed. Then*

(i) *there exists* $K : [1, \frac{q}{r'}) \to (0, \infty)$ *nondecreasing such that*

$$\|C_r\|_{L^q(\mathbb{R};w)\to L^q(\mathbb{R};w)} \le K(t)[w]_{A_t}^{\max\left\{1,\frac{t}{q(t-1)}\right\}};$$

(ii) *there exists a positive increasing function* $\mathcal{Q}$ *such that for* $t = \frac{q}{r'}$

$$\|C_r\|_{L^q(\mathbb{R};w)\to L^q(\mathbb{R};w)} \le \mathcal{Q}\left([w]_{A_t}\right). \tag{1.3}$$

We omit the standard deduction of Theorem 1.2 from Theorem 1.1, which follows along lines analogous to the proofs of [4, Theorem 3] and [16, Theorem 17.1]. Estimate (i) of Theorem 1.2 yields in particular that

$$w \in A_t \implies \|C_r\|_{L^q(\mathbb{R};w)\to L^q(\mathbb{R};w)} < \infty \qquad \forall r > \max\left\{2, \frac{q}{q-t}\right\},$$

an improvement over [8, Theorem 1.2], where $L^q(\mathbb{R}; w)$ boundedness is only shown for variation exponents $r > \max\left\{2t, \frac{qt}{q-t}\right\}$ when $w \in A_t$. Fixing r instead, part (ii) of Theorem 1.2 is sharp in the sense that $t = \frac{q}{r'}$ is the largest exponent such that an estimate of the type of (1.3) is allowed to hold. Indeed, if (1.3) were true for any $q = q_0 \in (r', \infty)$ and some $t = \frac{q_0}{s}$ with $s < r'$, a version of the Rubio de Francia extrapolation theorem (see for instance [3, Theorem 3.9]) would yield that C_r maps L^q into itself for all $q \in (s, \infty)$, contradicting the already mentioned counterexample from [18].

We turn to further comments on the proof and on the structure of the paper. In the upcoming Section 2 we reduce the bilinear form estimate (1.2) to an analogous statement for a bilinear form involving integrals over the upper-three space of symmetry parameters for the Carleson operator of a wave packet transforms of f and a variational-truncated wave packet transform of g. The natural framework for L^p-boundedness of such forms, the L^p-theory of outer measures, has been developed by the second author and Thiele in [11]. In Section 3, we recall the basics of this theory as well as the localized Carleson embeddings of [7] and [19]. These will come to fruition in Section 4, where we give the proof of Theorem 1.1. A significant challenge in the course of the proof is the treatment of the nonlocal (tail) components, which are handled via novel *ad-hoc* embedding theorems incorporating the fast decay of the wave packet coefficients away from the support of the input functions.

ACKNOWLEDGEMENTS. This work was initiated and continued during G. Uraltsev's visit to the Brown University and University of Virginia Mathematics Departments, whose hospitality is gratefully acknowledged. The authors would like to thank Amalia Culiuc, Michael Lacey, Ben Krause and Yumeng Ou for useful conversations about sparse domination principles.

2. Reduction to wave packet transforms

In this section we reduce the inequality (1.2) to an analogous statement involving wave packet transforms. Throughout this section, the variation exponent $r \in (2, \infty)$ is fixed, and we take $f, g \in \mathcal{C}_0^\infty(\mathbb{R})$. First of all we linearize the variation norm appearing in C_r. Begin by observing that the map

$$(x, \xi) \mapsto \int_{-\infty}^{\xi} \widehat{f}(\zeta)\, e^{ix\zeta} d\zeta$$

is uniformly continuous. By duality and standard considerations

$$C_r f(x) = \sup_N \sup_{\Xi \subset \mathbb{R}, \#\Xi \le N} \sup_{\|\{a_j\}\|_{\ell^{r'}} \le 1} \sum_{j=1}^{N} a_j \int_{\xi_{j-1}}^{\xi_j} \widehat{f}(\zeta)\, e^{ix\zeta} d\zeta.$$

Therefore, (1.2) will be a consequence of the estimate

$$\begin{aligned} \Lambda_{\vec{\xi},\vec{a}}(f, g) &:= \int_{\mathbb{R}} g(x) \left(\sum_{j=1}^{N} a_j(x) \int_{\xi_{j-1}(x)}^{\xi_j(x)} \widehat{f}(\zeta)\, e^{ix\zeta}\, d\zeta \right) dx \\ &\le K(p) \sum_{I \in \mathcal{S}} |I| \langle f \rangle_{I,p} \langle g \rangle_{I,1}, \end{aligned} \tag{2.1}$$

with right-hand side independent of $N \in \mathbb{N}$, $\Xi \subset \mathbb{R}$, $\#\Xi \le N$, and of the measurable Ξ^{N+1}-valued function $\vec{\xi}(x) = \{\xi_j(x)\}$ with $\xi_0(x) < \cdots < \xi_N(x)$, and $\mathbb{C}^{N+1}$-valued $\vec{a}(x) = \{a_j(x)\}$ with $\|\vec{a}(x)\|_{\ell^{r'}} = 1$.

The next step is to uniformly dominate the form $\Lambda_{\vec{\xi},\vec{a}}(f, g)$ by an outer form involving wave packet transforms of f and g; in the terminology of [11], *embedding* maps into the upper 3-space

$$(u, t, \eta) \in \mathbb{X} = \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}.$$

The parameters $\xi, \vec{a}$ will enter the definition of the embedding map for g. We introduce the wave packets

$$\psi_{t,\eta}(x) := t^{-1} e^{i\eta z}\, \psi\left(\frac{x}{t}\right), \qquad \eta \in \mathbb{R},\ t \in (0, \infty),$$

where ψ is a real valued, even Schwartz function with frequency support of width b containing the origin. The wave packet transform of f is thus defined, as in [11], by

$$F(f)(u, t, \eta) = \left| f * \psi_{t,\eta}(u) \right|, \qquad (u, t, \eta) \in \mathbb{X}. \tag{2.2}$$

For our fixed choice of $\vec{\xi}, \vec{a}$ we introduce the modified wave packet transform of g that is dual to (2.2) for the sake of bounding the left hand side of (2.1). Following [19, Equation (1.14)], it is given by

$$A(g)(u,t,\eta) := \sup_{\Psi} \left| \int_{\mathbb{R}} g(x) \sum_{j=1}^{N} a_j(x) \Psi_{t,\eta}^{\xi_j(x),\xi_{j+1}(x)}(x-u)\, dx \right|, \quad (u,t,\eta) \in \mathbb{X}, \tag{2.3}$$

with supremum being taken over all choices of *truncated wave packets* $\Psi_{t,\eta}^{\xi_-,\xi_+}$, that for each $t, \eta \in \mathbb{R}_+ \times \mathbb{R}$ are functions in $\mathcal{S}(\mathbb{R})$ parameterized by $\xi_-, \xi_+ \in \Xi$. We summarize the basic defining properties of the truncated wave packets in Remark 2.3 below, and we refer to [19] for a precise definition.

The duality of the embeddings (2.2) and (2.3) is a consequence of the following wave packet domination Lemma. We refer to [19] for the proof.

Lemma 2.1 (Wave packet domination). *Let $f, g, \Xi, \vec{\xi}, \vec{a}$ be as above and consider the bilinear form defined on the wave packets transforms, given by*

$$\mathsf{B}_{\vec{\xi},\vec{a}}(f,g) := \int_{\mathbb{X}} F(f)(u,t,\eta)A(g)(u,t,\eta)\,\mathrm{d}u\mathrm{d}t\mathrm{d}\eta. \tag{2.4}$$

Then

$$\Lambda_{\vec{\xi},\vec{a}}(f,g) \lesssim \mathsf{B}_{\vec{\xi},\vec{a}}\big(F(f), A(g)\big)$$

with uniform implied constant.

Using the above Lemma, we see that inequality (2.1) and thus Theorem 1.1 will follow from the bounds of the next proposition.

Proposition 2.2. *Let $p > r'$ be fixed. For all compactly supported $f, g \in L^\infty(\mathbb{R})$ there exist a sparse collection $\mathcal{S} = \mathcal{S}(f, g, p)$ and an absolute constant $K = K(p)$ such that*

$$\sup_N \sup_{\#\Xi \le N} \sup_{\vec{\xi},\vec{a}} \mathsf{B}_{\vec{\xi},\vec{a}}(f,g) \le K(p) \sum_{I\in\mathcal{S}} |I| \langle f\rangle_{I,p}\langle g\rangle_{I,1}, \tag{2.5}$$

where $\vec{\xi}, \vec{a}$ range over $\Xi^{N+1}, \mathbb{C}^{N+1}$-valued functions as above.

We now make a brief digression to justify definitions (2.2) and (2.3) of the wave packet transforms and the result of Lemma 2.1. Consider the term

$$\int_{\xi_{j-1}(x)}^{\xi_j(x)} \widehat{f}(\eta)e^{ix\eta}\mathrm{d}\eta$$

appearing in (2.1) and let us think for a moment of $\xi_{j-1}(x) = \xi_-$ and $\xi_j(x) = \xi_+$ as frozen. Then the following representation holds for the multiplier $1_{(\xi_-,\xi_+)}(\zeta)$:

$$1_{(\xi_-,\xi_+)}(\zeta) = \int_{\mathbb{R}_+\times\mathbb{R}} \widehat{\Psi}_{t,\eta}^{\xi_-,\xi_+}(\zeta)\,\mathrm{d}t\mathrm{d}\eta, \tag{2.6}$$

where $\Psi_{t,\eta}^{\xi_-,\xi_+}$ are truncated wave packets. Choosing a $\phi \in \mathcal{S}(\mathbb{R})$ such that $\widehat{\phi}_{t,\eta}(\zeta) = 1$ whenever $\widehat{\Psi}_{t,\eta}^{\xi_-,\xi_+}(\zeta) \neq 0$ for any $\xi_- < \xi_+ \in \mathbb{R}$, we obtain the pointwise identity

$$\int_{\xi_{j-1}(x)}^{\xi_j(x)} \widehat{f}(\zeta)e^{ix\zeta}\mathrm{d}\zeta = \int_{\mathbb{X}} f * \phi_{t,\eta}(u)\Psi_{t,\eta}^{\xi_{j-1}(x),\xi_j(x)}(x-u)\mathrm{d}u\mathrm{d}t\mathrm{d}\eta.$$

The results of Lemma 2.1 follow by Fubini and the triangle inequality.

We briefly illustrate identity (2.6), for a more careful discussion we refer to [19, Section 3]. Start by choosing $\psi \in \mathcal{S}(\mathbb{R})$ with $\widehat{\psi}$ non-negative and supported on a ball of radius $b/2$, and let $\chi \in \mathcal{C}_0^\infty(\mathbb{R})$ be a non-negative bump function supported on $[d-\epsilon, d+\epsilon]$ with $d > b$ and $\epsilon \ll b$. Suppose formally that $\xi_+ = +\infty$ so that, up to a suitable normalization of χ, a Littlewood-Paley type decomposition centered at ξ_- of the multiplier $1_{(\xi_-,+\infty)}$ gives

$$1_{(\xi_-,+\infty)}(\zeta) = \int_{\mathbb{R}_+\times\mathbb{R}} \widehat{\psi}(t(\zeta-\eta))\chi(t(\eta-\xi_-))\mathrm{d}t\mathrm{d}\eta.$$

A similar expression holds if $\xi_- = -\infty$ and $\xi_+ \in \mathbb{R}$. We choose truncated wave packets so that

$$\begin{aligned}
\Psi_{t,\eta}^{\xi_-,\xi_+}(x) &:= \psi_{t,\eta}(x)\chi(t(\eta-\xi_-)) && t(\eta-\xi_-) \ll t(\xi_+-\eta)\\
\Psi_{t,\eta}^{\xi_-,\xi_+}(x) &:= \psi_{t,\eta}(x)\chi(t(\xi_+-\eta)) && t(\eta-\xi_-) \gg t(\xi_+-\eta)\\
\Psi_{t,\eta}^{\xi_-,\xi_+}(x) &:= 0 && \eta \notin (\xi_-,\xi_+).
\end{aligned}$$

Finally if $t(\eta-\xi_-) \approx t(\xi_+-\eta)$ then $\Psi_{t,\eta}^{\xi_-,\xi_+}$ is chosen to appropriately model the transition between the above regimes and justifies identity (2.6).

Remark 2.3. In general we call a function $\Psi_{t,\eta}^{\xi_-,\xi_+} \in \mathcal{S}(\mathbb{R})$ parameterized by $\xi_- < \xi_+ \in \mathbb{R}$ a truncated wave-packet adapted to $t, \eta \in \mathbb{R}_+ \times \mathbb{R}$ if

$$e^{-i\eta t z} t\Psi_{t,\eta}^{\xi_-,\xi_+}(tx), \quad t^{-1}\partial_{\xi_-}\Big(e^{-i\eta t z} t\Psi_{t,\eta}^{\xi_-,\xi_+}(tx)\Big), \quad t^{-1}\partial_{\xi_+}\Big(e^{-i\eta t z} t\Psi_{t,\eta}^{\xi_-,\xi_+}(tx)\Big)$$

are uniformly bounded in $\mathcal{S}(\mathbb{R})$ as functions of x. Furthermore we require that $\widehat{\Psi}_{t,\eta}^{\xi_-,\xi_+}$ be supported on $(\eta - t^{-1}b,\ \eta + t^{-1}b)$ for some $b > 0$. Finally, for some constants $d, d', d'' > 0$ and $\epsilon > 0$ it must hold that

$$\begin{aligned}
&\Psi_{t,\eta}^{\xi_-,\xi_+} \neq 0 && \text{only if } \begin{cases} t(\eta-\xi_-) \in (d-\epsilon, d+\epsilon) \\ t(\xi_+-\eta) > d' > 0 \end{cases}\\
&\partial_{\xi_+}\Psi_{t,\eta}^{\xi_-,\xi_+} = 0 && \text{if } t(\xi_+-\eta) > d'' > d' > 0.
\end{aligned} \tag{2.7}$$

3. Localized outer-L^p embeddings

We now turn to the description of the analytic tools which we rely on in the proof of estimate (1.2). We work in the framework of outer measure spaces [11], see also [4,7]. In particular, we define a distinguished collection of subsets of the upper 3-space $\mathbb{X}$ which we refer to as *tents* above the time-frequency loci (I,ξ) where I is an interval of center $c(I)$ and length $|I|$, and $\xi \in \mathbb{R}$:

$$\begin{aligned}
&\mathsf{T}(I,\xi) := \mathsf{T}^\ell(I,\xi) \cup \mathsf{T}^o(I,\xi);\\
&\mathsf{T}^o(I,\xi) := \big\{(u,t,\eta) : t\eta - t\xi \in \Theta^o, t < |I|, |u-c(I)| < |I|-t\big\};\\
&\mathsf{T}^\ell(I,\xi) := \big\{(u,t,\eta) : t\eta - t\xi \in \Theta \setminus \Theta^o, t < |I|, |u-c(I)| < |I|-t\big\},
\end{aligned}$$

where $\Theta^o = [\beta^-, \beta^+]$, $\Theta = [\alpha^-, \alpha^+]$ are two geometric parameter intervals such that $0 \in \Theta^o \subset \Theta$. The specific values of the parameters do not matter. What is important is that given the geometric parameters of the wave packets appearing in (2.2) and (2.3) there exists a choice of parameters of the tents such that the statements of the subsequent discussion hold. For example it must hold that $(-b, b) \subset \Theta^o$ where b is the parameter that governs the frequency support of $\psi_{t,\eta}$ and $\Psi_{t,\eta}^{\xi_-,\xi_+}$. For a complete discussion see [19, Section 2]. As usual, we denote by μ the outer measure generated by countable coverings by tents $\mathsf{T}(I, \xi)$, $I \subset \mathbb{R}$, $\xi \in \mathbb{R}$ via the pre-measure $\mathsf{T}(I, \xi) \mapsto |I|$.

Let s be a size [11], *i.e.*, a family of quasi-norms indexed by tents T, defined on Borel functions $F : \mathbb{X} \to \mathbb{C}$. The corresponding outer-L^p space on $(\mathbb{X}, \mu)$ is defined by the quasi-norm

$$\|F\|_{L^p(\mathsf{s})} := \left(p \int_0^\infty \lambda^{p-1} \mu(\mathsf{s}(F) > \lambda)\, \mathrm{d}\lambda \right)^{\frac{1}{p}}, \qquad 0 < p < \infty,$$

$$\mu(\mathsf{s}(F) > \lambda) := \inf\left\{ \mu(E) : E \subset \mathbb{X},\ \sup_{\mathsf{T}} \mathsf{s}\left(F\mathbf{1}_{\mathbb{X}\setminus E}\right)(\mathsf{T}) \le \lambda \right\},$$

where the supremum on the right is taken over all tents $\mathsf{T} = \mathsf{T}(I, \xi)$. We will work with outer L^p spaces based on the sizes

$$\mathsf{s}^e(F)(\mathsf{T}) := \left(\frac{1}{|I|} \int_{\mathsf{T}^\ell} |F(u,t,\eta)|^2\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta \right)^{\frac{1}{2}} + \sup_{(u,t,\eta)\in\mathsf{T}} |F(u,t,\eta)|,$$

$$\mathsf{s}^m(A)(\mathsf{T}) := \left(\frac{1}{|I|} \int_{\mathsf{T}} |A(u,t,\eta)|^2\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta \right)^{\frac{1}{2}} + \frac{1}{|I|} \int_{\mathsf{T}^o} |A(u,t,\eta)|\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta,$$

which are related to the two embeddings (2.2) and (2.3) respectively. The dual relation of the sizes $\mathsf{s}^e, \mathsf{s}^m$ is given by the fact that for any two Borel functions $F, A : \mathbb{X} \to \mathbb{C}$ there holds

$$\int_{\mathsf{T}} |F(u,t,\eta)A(u,t,\eta)|\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta \le 2\mathsf{s}^\ell(F)(\mathsf{T})\mathsf{s}^o(A)(\mathsf{T}).$$

The abstract outer Hölder inequality [11, Proposition 3.4] and Radon-Nikodym type bounds [11, Proposition 3.6] yield

$$\int_{\mathsf{T}} |F(u,t,\eta)A(u,t,\eta)|\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta \lesssim \|F\|_{L^\sigma(\mathsf{s}^\ell)} \|A\|_{L^\tau(\mathsf{s}^o)} \tag{3.1}$$

whenever $1 \le \sigma, \tau \le \infty$ are Hölder dual exponents *i.e.*, $\frac{1}{\sigma} + \frac{1}{\tau} = 1$.

The nature of the wave packet transforms $f \mapsto F(f)$, $g \mapsto A(g)$ defined by (2.2), (2.3) is heavily exploited in the stopping-type outer L^p-embedding theorems below. We state the embedding theorems after some necessary definitions. It is convenient to use the notation

$$\mathsf{T}(I) := \{(u,t,\eta) : t < |I|,\ |u - c(I)| < |I| - t,\ \eta \in \mathbb{R}\}$$

for the set of the upper 3-space associated to the usual spatial tent over I. Given an open set $E \subset \mathbb{R}$ we associate to it the subset of $\mathsf{T}(E) \subset \mathbb{X}$ given by

$$\mathsf{T}(E) = \bigcup_{I \subset E} \mathsf{T}(I), \tag{3.2}$$

where the union is taken over all intervals $I \subset E$.

The first stopping embedding theorem, a reformulation of a result first obtained in [7], deals with the wave packet transform $f \mapsto F(f)$ of (2.2).

Proposition 3.1. *Let* $1 < p < 2$, $\sigma \in (p', \infty)$, *then there exists* $K > 0$ *such that the following holds. For all* $f \in L^p_{\mathrm{loc}}(\mathbb{R})$, *all intervals* Q, *and all* $c \in (0,1)$ *there exists an open set* $U_{f,p,Q}$ *satisfying*

$$\left|U_{f,p,Q}\right| \le c|Q|,$$

such that

$$\left\| F(f\mathbf{1}_{3Q})\mathbf{1}_{\mathsf{T}(Q)\setminus\mathsf{T}(U_{f,p,Q})} \right\|_{L^\sigma(\mathsf{S}^e)} \le K|Q|^{\frac{1}{\sigma}} \langle f\rangle_{3Q,p}. \tag{3.3}$$

The embedding theorem we use to treat the variationally truncated wave packet transform $g \mapsto A(g)$ of (2.3) stems from the main result of [19].

Proposition 3.2. *Let* $\tau \in (r', \infty)$, *then there exists* $K > 0$ *such that the following holds. For all* $g \in L^1_{\mathrm{loc}}(\mathbb{R})$, *all intervals* Q *and all* $c \in (0,1)$ *there exists an open set* $V_{g,1,Q}$ *satisfying*

$$\left|V_{g,1,Q}\right| \le c|Q|,$$

such that

$$\left\| A(g\mathbf{1}_{3Q})\mathbf{1}_{\mathsf{T}(Q)\setminus\mathsf{T}(V_{g,1,Q})} \right\|_{L^\tau(\mathsf{S}^m)} \le K|Q|^{\frac{1}{\tau}} \langle g\rangle_{3Q,1}. \tag{3.4}$$

We stress that the constant K in Proposition 3.2 does not depend on the parameters $\vec{a}, \vec{\xi}, \Xi, N$ appearing in the definition (2.3) of the map A.

The two propositions above appear in [19] in a somewhat different form that uses the notion of iterated outer measure spaces introduced therein. We derive the statement of Propositions 3.2 by using the weak boundedness on $L^1(\mathbb{R})$ of the map (2.3) of [19, Theorem 1.3]. In particular that result, applied to the function $g\mathbf{1}_{3Q}$ for $\lambda = cK\langle g\rangle_{3Q,1}$, yields a collection of disjoint open intervals $\mathcal{I}$ and

$$V_{g,1,Q} := \bigcup_{I \in \mathcal{I}} I, \qquad \left|V_{g,1,Q}\right| \le \frac{C|Q|}{K},$$

so that (3.4) holds as required. We conclude by choosing $K \ge C/c$. A similar procedure can be used to obtain Proposition 3.1 from [19, Theorem. 1.2].

In effect, we have shown that the formulation of the boundedness properties of the embedding maps (2.2) and (2.3) as expressed in Propositions 3.1 and 3.2 are equivalent to the iterated outer measure formulation of [19]. Furthermore the use of iterated outer measure L^p-norms allowed us to bootstrap the above results to $L^p_{\mathrm{loc}}(\mathbb{R})$ generality from an *a priori* type statement, as illustrated in [19, Section 2.1].

4. Proof of Proposition 2.2

Throughout this proof, the exponent $p \in (r', \infty)$ is fixed and all the implicit constants are allowed to depend on r, p without explicit mention. Since the linearization parameters play no explicit role in the upcoming arguments we omit them from the notation: assume them fixed and simply write $\mathsf{B}(f, g)$ for the form $\mathsf{B}_{\vec{\xi},\vec{a}}(f, g)$ defined in (2.4). Given any interval Q, we introduce the localized version

$$\mathsf{B}_Q(f, g) := \int_{\mathsf{T}(Q)} F(f)(u, t, \eta) A(g)(u, t, \eta)\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta. \tag{4.1}$$

4.1. The principal iteration

The main step of the proof of Proposition 2.2 is contained in the following lemma, which we will apply iteratively.

Lemma 4.1. *There exists a positive constant K such that the following holds. Let $f, g \in L^\infty(\mathbb{R})$ and compactly supported, and $Q \subset \mathbb{R}$ be any interval. There exists a countable collection of disjoint open intervals $\mathcal{I}_Q$ such that*

$$\bigcup_{I \in \mathcal{I}_Q} I \subset Q, \qquad \sum_{I \in \mathcal{I}_Q} |I| \le 2^{-12}|Q|, \tag{4.2}$$

and such that

$$\mathsf{B}_Q(f\mathbf{1}_{3Q}, g\mathbf{1}_{3Q}) \le K|Q|\langle f\rangle_{3Q,p}\langle g\rangle_{3Q,1} + \sum_{I \in \mathcal{I}_Q} \mathsf{B}_I(f\mathbf{1}_{3I}, g\mathbf{1}_{3I}). \tag{4.3}$$

The proof of the lemma consists of several steps, which we now address. Notice that there is no loss in generality with assuming that f, g are supported on $3Q$: we do so for mere notational convenience.

4.1.1. *Construction of $\mathcal{I}_Q$*

Referring to the notation of Section 3, set

$$\begin{aligned} E_{f,Q} &= U_{f,p,Q} \cup \left\{x \in \mathbb{R} : \mathrm{M}_p f(x) > c^{-1}\langle f\rangle_{3Q,p}\right\}, \\ E_{g,Q} &= V_{g,1,Q} \cup \left\{x \in \mathbb{R} : \mathrm{M}_1 g(x) > c^{-1}\langle g\rangle_{3Q,1}\right\}, \\ E_Q &= Q \cap \left(E_{f,Q} \cup E_{g,Q}\right). \end{aligned}$$

Write the open set E_Q as the union of a countable collection $I \in \mathcal{I}_Q$ of disjoint open intervals. Then (4.2) holds provided that c is chosen small enough. Also, necessarily $3I \cap E_Q^c \neq \varnothing$ if $I \in \mathcal{I}_Q$, so that

$$\inf_{x \in 3I} \mathrm{M}_1 f(x) \lesssim \langle f\rangle_{3Q,p}, \qquad \inf_{x \in 3I} \mathrm{M}_1 g(x) \lesssim \langle g\rangle_{3Q,1}. \tag{4.4}$$

For further use we note that, with reference to the notation of Propositions 3.1 and 3.2,

$$\mathsf{T}(Q) \setminus \mathsf{T}(E_Q) \subset \mathsf{T}(Q) \setminus \big(\mathsf{T}(U_{f,p,Q}) \cup \mathsf{T}(V_{g,1,Q})\big). \tag{4.5}$$

This completes the construction of $\mathcal{I}_Q$.

4.1.2. *Proof of* (4.3)

We begin by using (3.2) to partition the outer integral over $\mathsf{T}(Q)$ as

$$\mathsf{B}_Q(f,g) \le \int_{\mathsf{T}(Q)\setminus\mathsf{T}(E_Q)} F(f)A(g)\,\mathrm{d}u\mathrm{d}t\mathrm{d}\eta + \sum_{I\in\mathcal{I}_Q} \mathsf{B}_I(f,g). \tag{4.6}$$

Choosing $\tau \in (r', p)$, the dual exponent $\sigma = \tau' \in (p', \infty)$. By virtue of (4.5), we may apply the outer Hölder inequality (3.1) and the embeddings in Propositions 3.1 and 3.2 to control the first summand in (4.6) by an absolute constant times

$$\Big\| F(f)\mathbf{1}_{\mathsf{T}(Q)\setminus\mathsf{T}(U_{f,p,Q})}\Big\|_{L^\sigma(\mathsf{s}^\ell)} \Big\| A(g)\mathbf{1}_{\mathsf{T}(Q)\setminus\mathsf{T}(V_{g,1,Q})}\Big\|_{L^\tau(\mathsf{s}^o)} \lesssim |Q|\langle f\rangle_{3Q,p}\langle g\rangle_{3Q,1}.$$

We turn to the second summand in (4.6), which is less than or equal to

$$\sum_{I\in\mathcal{I}_Q} \mathsf{B}_I\big(f\mathbf{1}_{3I}, g\mathbf{1}_{3I}\big) + \sum_{\substack{(\mathsf{a},\mathsf{b})\in\{\mathsf{in},\mathsf{out}\}^2\\ (\mathsf{a},\mathsf{b})\neq(\mathsf{in},\mathsf{in})}} \sum_{I\in\mathcal{I}_Q} \mathsf{B}_I\big(f\mathbf{1}_{I^\mathsf{a}}, g\mathbf{1}_{I^\mathsf{b}}\big),$$

where $I^{\mathsf{in}} = 3I$, $I^{\mathsf{out}} = 3Q \setminus 3I$. The first term in the above expression appears on the right hand side of (4.3). We claim that

$$\sum_{I\in\mathcal{I}_Q} \mathsf{B}_I\big(f\mathbf{1}_{I^\mathsf{a}}, g\mathbf{1}_{I^\mathsf{b}}\big) \lesssim |Q|\langle f\rangle_{3Q,p}\langle g\rangle_{3Q,1}, \qquad (\mathsf{a},\mathsf{b}) \neq (\mathsf{in},\mathsf{in}), \tag{4.7}$$

thus leading to the required estimate for (4.3). Assume $\mathsf{a} = \mathsf{in}$, $\mathsf{b} = \mathsf{out}$ for the sake of definiteness, the other cases being identical. Fix $I \in \mathcal{I}_Q$. We will show that

$$\mathsf{B}_I\big(f\mathbf{1}_{I^{\mathsf{in}}}, g\mathbf{1}_{I^{\mathsf{out}}}\big) \lesssim |I|\langle f\rangle_{3Q,p}\langle g\rangle_{3Q,1}, \tag{4.8}$$

whence (4.7) follows by summing over $I \in \mathcal{I}_Q$ and taking advantage of (4.2).

4.1.3. *Proof of* (4.8)

We introduce the Carleson box over the interval P:

$$\mathsf{box}(P) = \Big\{(u,t,\eta) \in \mathbb{X} : u \in P, \tfrac{1}{2}|P| \le t < |P|\Big\}.$$

Fix $I \in \mathcal{I}_Q$. The root of our argument for (4.8) is the fact that $\operatorname{supp} g\mathbf{1}_{I^{\mathsf{out}}}$ lies outside $3I$. This leads to the exploitation of the following lemma, whose proof is given at the end of the paragraph.

Lemma 4.2. *Let P be any interval, $h \in L^p_{\mathrm{loc}}(\mathbb{R})$, and τ, σ as above. There holds*

$$\|A(h)\mathbf{1}_{\mathsf{box}(P)}\|_{L^\tau(\mathsf{s}^m)} \lesssim |P|^{\frac{1}{\tau}} \left(1 + \frac{\operatorname{dist}(P, \operatorname{supp} h)}{|P|}\right)^{-100} \inf_{x\in 3P} \mathrm{M}_1 h(x), \tag{4.9}$$

$$\|F(h)\mathbf{1}_{\mathsf{box}(P)}\|_{L^\sigma(\mathsf{s}^e)} \lesssim |P|^{\frac{1}{\sigma}} \left(1 + \frac{\operatorname{dist}(P, \operatorname{supp} h)}{|P|}\right)^{-100} \inf_{x\in 3P} \mathrm{M}_p h(x). \tag{4.10}$$

Now let $P \in \mathcal{P}_k(I)$ be the collection of dyadic subintervals of I with $|P| = 2^{-k}|I|$. If $P \in \mathcal{P}_k(I)$ it holds $\operatorname{dist}(P, I^{\mathrm{out}}) \geq |I| = 2^k|P|$. Moreover

$$\sum_{P\in\mathcal{P}_k(I)} |P| = |I|, \qquad \inf_{x\in 3P} \mathrm{M}_1 h(x) \lesssim 2^k \inf_{x\in 3I} \mathrm{M}_1 h(x),$$

for all locally integrable h. Since

$$\mathsf{T}(I) \subset \bigcup_{k=0}^{\infty} \bigcup_{P\in\mathcal{P}_k(I)} \mathsf{box}(P),$$

we obtain, using the outer Hölder inequality (3.1) to pass to the third line, the following chain of inequalities:

$$\begin{aligned}
\mathsf{B}_I(f\mathbf{1}_{I^{\mathrm{in}}}, g\mathbf{1}_{I^{\mathrm{out}}}) &\leq \sum_{k\geq 0} \sum_{P\in\mathcal{P}_k(I)} \int_{\mathsf{box}(P)} F(f\mathbf{1}_{I^{\mathrm{in}}}) A(g\mathbf{1}_{I^{\mathrm{out}}})\, \mathrm{d}u\mathrm{d}t\mathrm{d}\eta \\
&\leq \sum_{k\geq 0} \sum_{P\in\mathcal{P}_k(I)} \|F(f\mathbf{1}_{I^{\mathrm{in}}})\mathbf{1}_{\mathsf{box}(P)}\|_{L^\sigma(\mathsf{s}^e)} \|A(g\mathbf{1}_{I^{\mathrm{out}}})\mathbf{1}_{\mathsf{box}(P)}\|_{L^\tau(\mathsf{s}^m)} \\
&\lesssim \sum_{k\geq 0} \sum_{P\in\mathcal{P}_k(I)} |P| \left(\inf_{x\in 3P} \mathrm{M}_p f(x)\right) \left(2^{-99k} \inf_{x\in 3P} \mathrm{M}_p g(x)\right) \\
&\leq \sum_{k\geq 0} 2^{-98k} \sum_{P\in\mathcal{P}_k(I)} |P| \left(\inf_{x\in 3I} \mathrm{M}_p f(x)\right) \left(\inf_{x\in 3I} \mathrm{M}_1 g(x)\right) \\
&\lesssim |I| \left(\inf_{x\in 3I} \mathrm{M}_p f(x)\right) \left(\inf_{x\in 3I} \mathrm{M}_1 g(x)\right),
\end{aligned}$$

which, by virtue of (4.4), complies with (4.8).

Proof of Lemma 4.2. We show how estimate (4.9) follows from Proposition 3.2. Then, (4.10) is obtained from Proposition 3.1 in a similar manner. By quasi-sublinearity and monotonicity of the outer measure $L^\tau(s^m)$ norm we have that

$$\begin{aligned}
&\|A(h)\mathbf{1}_{\mathsf{box}(P)}\|_{L^\tau(s^m)} \\
&\leq C\|A\big(h\mathbf{1}_{9P}\big)\mathbf{1}_{\mathsf{box}(P)}\|_{L^\tau(s^m)} + \sum_{k=3}^{\infty} C^k \|A\big(h\mathbf{1}_{3^kP\setminus 3^{k-1}P}\big)\mathbf{1}_{\mathsf{box}(P)}\|_{L^\tau(s^m)}.
\end{aligned} \tag{4.11}$$

Applying the embedding bound (3.4) with $c = 3^{-2}$ and $Q = 3P$ provides us with $V_{h,1,3P}$ such that $\mathsf{box}(P) \subset \mathsf{T}(9P) \setminus \mathsf{T}(V_{h,1,3P})$, whence

$$\left\| A\big(h\mathbf{1}_{9P}\big)\mathbf{1}_{\mathsf{box}(P)}\right\|_{L^\tau(s^m)} \le CK|P|^{\frac{1}{\tau}}\langle h\rangle_{9P,1} \le CK|P|^{\frac{1}{\tau}} \inf_{x\in 3P} M_1 h(x).$$

Indeed, we chose c in such a way that $|V_{h,1,3P}| < 3^{-1}Q$, which guarantees that $\mathsf{T}(V_{h,1,9P})$ does not intersect $\mathsf{box}(P)$. We claim that, similarly, we have that for $k > 2$ and for an arbitrarily large $N \gg 1$ there holds

$$\begin{aligned}\left\| A\big(h\mathbf{1}_{3^kP\setminus 3^{k-1}P}\big)\mathbf{1}_{\mathsf{box}(P)}\right\|_{L^\tau(s^m)} &\le CK3^{-Nk}|P|^{\frac{1}{\tau}}\langle h\rangle_{3^kP,1}\\ &\le CK|P|^{\frac{1}{\tau}}3^{-Nk}\inf_{x\in 3P} M_1 h(x).\end{aligned}$$

Let

$$(u,t,\eta)\mapsto \Psi^{\xi_-,\xi_+}_{t,\eta}(\cdot - u)$$

be a choice of truncated wave packets which approximately achieves the supremum in

$$A\big(h\mathbf{1}_{3^kP\setminus 3^{k-1}P}\big)(u,t,\eta),$$

cf. (2.3). Then

$$\widetilde{\Psi}^{\xi_-,\xi_+}_{t,\eta}(\cdot - u) := \left(1 + \frac{|(x-u)-c(P)|}{|P|}\right)^{2N}\Psi^{\xi_-,\xi_+}_{t,\eta}(\cdot - u)$$

are adapted truncated wave packets as well since multiplying by a polynomial does not change the frequency support of $\Psi^{\xi_-,\xi_+}_{t,\eta}$ and so the conditions on being truncated wave packets is maintained. Let $\tilde{A}(h\mathbf{1}_{3^kP\setminus 3^{k-1}P})(u,t,\eta)$ be the embedding obtained by using the wave packets $\tilde{\Psi}^{\xi_-,\xi_+}_{t,\eta}(\cdot - u)$ instead of $\Psi^{\xi_-,\xi_+}_{t,\eta}(\cdot - u)$. Given that $(u,t,\eta)\in \mathsf{box}(P)$ we have that

$$\left|A(h\mathbf{1}_{3^kP\setminus 3^{k-1}P})(u,t,\eta)\right| \le C3^{-2Nk}\tilde{A}\big(h\mathbf{1}_{3^kP\setminus 3^{k-1}P}\big)(u,t,\eta).$$

However the bounds (3.4) also hold for $\tilde{A}$ with an additional multiplicative constant that depends at most on N. Applying these bounds with $P = 3^{k-1}Q$ and $c = 3^{-k}$ we have once again that

$$\left\|\tilde{A}\big(h\mathbf{1}_{3^kP\setminus 3^{k-1}P}\big)\mathbf{1}_{\mathsf{box}(P)}\right\|_{L^\tau(s^m)} \le CK|P|^{\frac{1}{\tau}}3^k\langle h\rangle_{3^kP,1}.$$

As long as N is chosen large enough with respect to $C > 1$ appearing in (4.11), the above inequality gives the required bound. The decay factor in term of $\operatorname{dist}(P, \operatorname{supp} h)$ follows from the fact that the the first k_0 terms in (4.11) vanish if $\operatorname{supp} h \cap 3^{k_0}P = \varnothing$. □

4.2. The iteration argument

With Lemma 4.1 at hand, we proceed to the proof of Proposition 2.2. Fix $f, g \in L^\infty(\mathbb{R})$ with compact support. By an application of Fatou's lemma, it suffices to prove (2.5) with B_{Q_0} *an lieu* of B for an arbitrary interval Q_0 with $\operatorname{supp} f, \operatorname{supp} g \subset Q_0$. That is, it suffices to construct a sparse collection $\mathcal{S}$ such that

$$\mathsf{B}_{Q_0}(f, g) \leq C \sum_{I \in \mathcal{S}} |I| \langle f \rangle_{I,p} \langle g \rangle_{I,1}, \tag{4.12}$$

provided that the constant C does not depend on Q_0. We fix such a Q_0. Furthermore, as

$$\mathsf{B}_{Q_0}(f, g) = \sup_{\varepsilon > 0} \mathsf{B}_{Q_0,\varepsilon}(f, g),$$
$$\mathsf{B}_{Q,\varepsilon}(f, g) := \int_{\mathsf{T}(Q)} F(f)(u, t, \eta) A(g)(u, t, \eta) \mathbf{1}_{\{t>\varepsilon\}} \,\mathrm{d}u \mathrm{d}t \mathrm{d}\eta,$$

it suffices to prove (4.12) with $\mathsf{B}_{Q_0,\varepsilon}$ replacing B_{Q_0}, with constants uniform in $\varepsilon > 0$. We also notice that Lemma 4.1 holds uniformly, if one replaces all the instances of B_Q in (4.1) with $\mathsf{B}_{Q,\varepsilon}$. From now on we fix $\varepsilon > 0$ and drop it from the notation.

We now perform the following iterative procedure. Set $\mathcal{S}_0 = \{Q_0\}$. Suppose that the collection of open intervals $Q \in \mathcal{S}_n$ has been constructed, and define inductively

$$\mathcal{S}_{n+1} = \bigcup_{Q \in \mathcal{S}_n} \mathcal{I}_Q,$$

where $\mathcal{I}_Q$ is obtained as in the Lemma 4.1. It can be seen inductively that

$$Q \in \mathcal{S}_n \implies |Q| \leq 2^{-12n} |Q_0|.$$

We iterate this procedure as long as $n \leq N$, where N is taken so that $2^{-12N}|Q_0| < \varepsilon$ holds. At that point we stop the iteration and set

$$\mathcal{S}^\star = \bigcup_{n=0}^{N} \mathcal{S}_n.$$

Making use of estimate (4.3) along the iteration of Lemma 4.1 we readily obtain

$$\begin{aligned} \mathsf{B}_{Q_0}(f, g) &\lesssim \sum_{n=0}^{N-1} \sum_{Q \in \mathcal{S}_n} |Q| \langle f \rangle_{3Q,p} \langle g \rangle_{3Q,1} + \sum_{Q \in \mathcal{S}_N} \sum_{I \in \mathcal{I}_Q} \mathsf{B}_I(f \mathbf{1}_{3I}, g \mathbf{1}_{3I}) \\ &= \sum_{Q \in \mathcal{S}^\star} |Q| \langle f \rangle_{3Q,p} \langle g \rangle_{3Q,1}, \end{aligned}$$

as each term B_I, $I \in \mathcal{S}_N$ vanishes by the condition on N. Observe that the sets

$$X_Q := Q \setminus \left(\bigcup_{I \in \mathcal{S}^\star : I \subsetneq Q} I \right) = Q \setminus \left(\bigcup_{I \in \mathcal{I}_Q} I \right) \qquad Q \in \mathcal{S}^\star$$

are pairwise disjoint and, from (4.2), that $|Q \backslash X_Q| \geq (1 - 2^{-12})|Q|$ yields that $\mathcal{S}^\star$ is sparse, and so is $\mathcal{S} = \{3Q : Q \in \mathcal{S}^\star\}$. This completes the proof of Proposition 2.2.

References

[1] D. BELTRAN and J. BENNETT, *Subdyadic square functions and applications to weighted harmonic analysis*, Adv. Math. **307** (2017), 72–99.

[2] J. M. CONDE-ALONSO and G. REY, *A pointwise estimate for positive dyadic shifts and some applications*, Math. Ann. **365** (2016), 1111–1135.

[3] D. V. CRUZ-URIBE, J. M. MARTELL and C. PÉREZ, "Weights, Extrapolation and the Theory of Rubio de Francia", Operator Theory: Advances and Applications, Vol. 215, Birkhäuser/Springer Basel AG, Basel, 2011.

[4] A. CULIUC, F. DI PLINIO and Y. OU, *Domination of multilinear singular integrals by positive sparse forms*, J. London Math. Soc., to appear, preprint arXiv:1603.05317.

[5] F. DI PLINIO, *Weak-L^p bounds for the Carleson and Walsh-Carleson operators*, C. R. Math. Acad. Sci. Paris **352** (2014), 327–331.

[6] F. DI PLINIO and A. K. LERNER, *On weighted norm inequalities for the Carleson and Walsh-Carleson operator*, J. Lond. Math. Soc. (2) **90** (2014), 654–674.

[7] F. DI PLINIO and Y. OU, *A modulation invariant Carleson embedding theorem outside local L^2*, J. Anal. Math., to appear, preprint arXiv:1506.05827.

[8] Y. DO and M. LACEY, *Weighted bounds for variational Fourier series*, Studia Math. **211** (2012), 153–190.

[9] Y. DO and M. LACEY, *Weighted bounds for variational Walsh-Fourier series*, J. Fourier Anal. Appl. **18** (2012), 1318–1339.

[10] Y. DO, C. MUSCALU and C. THIELE, *Variational estimates for the bilinear iterated Fourier integral*, J. Funct. Anal. **272** (2017), 2176–2233.

[11] Y. DO and C. THIELE, *L^p theory for outer measures and two themes of Lennart Carleson united*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), 249–296.

[12] L. GRAFAKOS, JOSÉ MARÍA MARTELL and F. SORIA, *Weighted norm inequalities for maximally modulated singular integral operators*, Math. Ann. **331** (2005), 359–394.

[13] R. A. HUNT and W. SANG YOUNG, *A weighted norm inequality for Fourier series*, Bull. Amer. Math. Soc. **80** (1974), 274–277.

[14] M. T. LACEY, *An elementary proof of the A_2 bound*, Israel J. Math. **217** (2017), 181–195.

[15] A. LERNER, *On pointwise estimates involving sparse operators*, New York J. Math. **22** (2016), 341–349.

[16] A. LERNER and F. NAZAROV, *Intuitive dyadic calculus: the basics*, Expo. Math. (2018), to appear, preprint arXiv:1508.05639 (2015).

[17] A. K. LERNER, *A simple proof of the A_2 conjecture*, Int. Math. Res. Not. IMRN (2013), 3159–3170.

[18] R. OBERLIN, A. SEEGER, T. TAO, C. THIELE and J. WRIGHT, *A variation norm Carleson theorem*, J. Eur. Math. Soc. (JEMS) **14** (2012), 421–464.

[19] G. URALTSEV, *Variational Carleson embeddings into the upper* 3*-space*, preprint arXiv:1610.07657.

Department of Mathematics
The University of Virginia
Charlottesville, VA 22904-4137, USA
francesco.diplinio@virginia.edu
yen.do@virginia.edu

Mathematics Department
Universität Bonn,
Endenicher Allee 60
D - 53115 Bonn, Germany
gennady.uraltsev@math.uni-bonn.de