Publ. Mat. **69** (2025), 445–471
DOI: 10.5565/PUBLMAT6922510

POINTWISE LOCALIZATION AND SHARP WEIGHTED BOUNDS FOR RUBIO DE FRANCIA SQUARE FUNCTIONS

Francesco Di Plinio, Mikel Flórez-Amatriain, Ioannis Parissis, and Luz Roncal

Abstract: Let $H_\omega f$ be the Fourier restriction of $f \in L^2(\mathbb{R})$ to an interval $\omega \subset \mathbb{R}$. If Ω is an arbitrary collection of pairwise disjoint intervals, the square function of $\{H_\omega f : \omega \in \Omega\}$ is termed the Rubio de Francia square function T^Ω_{RF}. This article proves a pointwise bound for T^Ω_{RF} by a sparse operator involving local L^2-averages. A pointwise bound for the smooth version of T^Ω_{RF} by a sparse square function is also proved. These pointwise localization principles lead to quantified $L^p(w)$, $p > 2$, and weak $L^p(w)$, $p \geq 2$, norm inequalities for T^Ω_{RF}. In particular, the obtained weak $L^p(w)$-norm bounds are new for $p \geq 2$ and sharp for $p > 2$. The proofs rely on sparse bounds for abstract balayages of Carleson sequences, local orthogonality, and very elementary time-frequency analysis techniques.

The paper also contains two results related to the outstanding conjecture that T^Ω_{RF} is bounded on $L^2(w)$ if and only if $w \in A_1$. The conjecture is verified for radially decreasing even A_1-weights, and in full generality for the Walsh group analogue of T^Ω_{RF}.

2020 Mathematics Subject Classification: Primary: 42B20. Secondary: 42B25.

Key words: Rubio de Francia square function, sparse domination, sparse square functions, exponential square integrability, sharp weighted bounds, time-frequency analysis, localization principles.

1. Introduction and main results

The L^p-norm, $1 < p < \infty$, equivalence between f and its Littlewood–Paley square function lies at the foundation of the modern treatment of singular integrals. The fact that this equivalence extends to weighted $L^p(w)$-norms for weights in the Muckenhoupt class testifies the localized nature of the Littlewood–Paley inequalities. In contrast to the lacunary Littlewood–Paley configuration, this article addresses the localization properties of square functions of both smooth and rough multipliers supported on frequency intervals forming an *arbitrary* pairwise disjoint, or finitely overlapping, collection; precise definitions are given below.

F. Di Plinio is partially supported by the FRA 2022 Program of University of Napoli Federico II, project ReSinAPAS - Regularity and Singularity in Analysis, PDEs, and Applied Sciences.
M. Flórez-Amatriain is partially supported by the projects CEX2021-001142-S, PID2021-122156NB-I00/AEI/10.13039/501100011033 funded by Agencia Estatal de Investigación and acronym "HAMIP", grants BERC 2022-2025 of the Basque Government and predoc Basque Government grant 2022 "Programa Predoctoral de Formación de Personal Investigador No Doctor".
I. Parissis is partially supported by grant PID2021-122156NB-I00 funded by MICIU/AEI/10.13039/501100011033 and FEDER, EU, grant IT1615-22 of the Basque Government and IKERBASQUE.
L. Roncal is partially supported by the projects CEX2021-001142-S, RYC2018-025477-I, and CNS2023-143893, funded by Agencia Estatal de Investigación, and PID2023-146646NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ESF+, grant BERC 2022-2025 of the Basque Government, and IKERBASQUE.

For intervals $\omega \subset \mathbb{R}$, define the class of multipliers adapted to ω as follows. Say $m \in \mathbb{M}_\omega$ if $m \in \mathcal{C}^D(\omega)$ for a fixed large integer D and

$$\operatorname{supp} m \subset \omega, \quad \sup_{\xi\in\omega} \sup_{0\le j\le D} \operatorname{dist}(\xi,\partial\omega)^j \|m^{(j)}\|_\infty \le 1.$$

To a collection of pairwise disjoint intervals Ω, and a choice $\{m_\omega \in \mathbb{M}_\omega : \omega\in\Omega\}$, associate the square function

$$T^\Omega f := \left(\sum_{\omega\in\Omega} |T_\omega f|^2\right)^{\frac12}, \quad T_\omega f(x) := \int_{\mathbb{R}} \widehat{f}(\xi) m_\omega(\xi) \mathrm{e}^{-i\xi x} \frac{\mathrm{d}\xi}{\sqrt{2\pi}}, \quad x\in\mathbb{R}.$$

The operator

$$H_\omega f(x) := \int_\omega \widehat{f}(\xi) \mathrm{e}^{-i\xi x} \frac{\mathrm{d}\xi}{\sqrt{2\pi}}, \quad x\in\mathbb{R}, \tag{1}$$

is an instance of T_ω corresponding to the choice $m_\omega = \mathbf{1}_\omega$. This specific case of T^Ω is the so-called *Rubio de Francia square function*, which is assigned the notation T^Ω_{RF}

$$T^\Omega_{\mathrm{RF}} f(x) := \left(\sum_{\omega\in\Omega} |H_\omega f(x)|^2\right)^{\frac12}, \quad x\in\mathbb{R}.$$

With more details and discussion to follow, one of the main results of this paper is the pointwise control of $T^\Omega_{\mathrm{RF}} f$ by a *sparse form*, see Subsection 1.1, in a sharp way, leading to new and in several cases best possible weighted norm inequalities for this operator.

Theorem A. *Let Ω be a collection of pairwise disjoint intervals and T^Ω_{RF} be as above. For every $f \in L^2(\mathbb{R})$ with compact support there exists a sparse collection $\mathcal{S}$ such that*

$$T^\Omega_{\mathrm{RF}} f \lesssim \sum_{Q\in\mathcal{S}} \langle f\rangle_{2,Q} \mathbf{1}_Q$$

and the L^2-average on the right hand side cannot be replaced by any L^p-average for any $p<2$. Furthermore there holds

$$\|T^\Omega_{\mathrm{RF}}\|_{L^2(w)\to L^{2,\infty}(w)} \lesssim [[w]_{A_1} [w]_{A_\infty} \log(\mathrm{e} + [w]_{A_\infty})]^{\frac12}$$

and for $2<p<\infty$

$$\|T^\Omega_{\mathrm{RF}}\|_{L^p(w)\to L^{p,\infty}(w)} \lesssim [w]^{\frac1p}_{A_{\frac p2}} [w]^{\frac{1}{p'}}_{A_\infty}.$$

The first estimate is best possible up to the logarithmic term while the second estimate is best possible.

We will get Theorem A as a consequence of more general corresponding results for square functions T^Ω defined in terms of more general multipliers in $\{\mathbb{M}_\omega\}_{\omega\in\Omega}$, as described above; see Theorem C, Corollary C.1, and Subsection 1.2.

A smooth, intrinsic counterpart of T^Ω is defined as follows. For each interval $\omega \subset \mathbb{R}$, let Φ_ω be the class of functions

$$\Phi_\omega := \left\{ \phi \in \mathcal{S}(R) : \operatorname{supp} \phi \subset \omega, \sup_{0 \le j \le D} \ell_\omega^j \|\phi^{(j)}\|_\infty \le 1 \right\}$$

for a positive integer D which we fix to be sufficiently large throughout the paper. Then the *intrinsic smooth Rubio de Francia square function* is the operator

$$G^\Omega f := \left(\sum_{\omega \in \Omega} f_\omega^2 \right)^{\frac{1}{2}}, \quad f_\omega(x) := \sup_{\phi \in \Phi_\omega} |f * \widehat{\phi}(x)|, \quad x \in \mathbb{R}.$$

Both definitions naturally extend to higher dimensions and/or parameters by considering collections of disjoint rectangles with respect to a fixed choice of a basis in $\mathbb{R}^n$ and defining the corresponding frequency projection operators. The two square functions G^Ω, T^Ω are related by vector-valued Littlewood–Paley inequalities, and their $L^p(\mathbb{R})$ behavior, and in fact their $L^p(w)$-boundedness for weights $w \in A_p$ as well, $1 < p < \infty$, are thus qualitatively equivalent.

The well-known result by Rubio de Francia [**36**] tells us that the operators T^Ω_{RF}, G^Ω are bounded on $L^p(\mathbb{R})$ for $p \ge 2$; see [**24, 26**] for the higher parametric case. Rubio de Francia's reliance on local orthogonality in [**36**] is embodied by the main step of his proof, namely the sharp function pointwise inequality

$$[G^\Omega f]^\# \le C \sqrt{\mathrm{M}(|f|^2)}. \tag{2}$$

1.1. Pointwise sparse domination of T^Ω and G^Ω. Estimate (2) also yields $L^p(w)$-norm bounds for weights w in appropriate Muckenhoupt classes. With the dual intent of strengthening (2) and of precisely quantifying these weighted estimates, we establish pointwise domination principles for both T^Ω and G^Ω, respectively involving the case $p = 2$ of the *sparse operators*

$$T_{p,\mathcal{S}} f := \sum_{Q \in \mathcal{S}} \langle f \rangle_{p,Q} \mathbf{1}_Q, \quad G_{p,\mathcal{S}} f := \left(\sum_{Q \in \mathcal{S}} \langle f \rangle_{p,Q}^2 \mathbf{1}_Q \right)^{\frac{1}{2}}, \quad 0 < p < \infty, \tag{3}$$

associated to a *sparse* collection $\mathcal{S}$ of intervals on the real line. The notations and definitions appearing in (3) and in what follows are standard, and are recalled at the end of the introduction.

Theorem B. *Let Ω be a collection of pairwise disjoint intervals. For each $f \in L^2(\mathbb{R})$ with compact support there exists a sparse collection $\mathcal{S}$ such that*

$$G^\Omega f \lesssim G_{2,\mathcal{S}} f$$

pointwise almost everywhere. The implicit constant in the above inequality is absolute.

Theorem C. *Let Ω be a collection of pairwise disjoint intervals. For each $f \in L^2(\mathbb{R})$ with compact support there exists a sparse collection $\mathcal{S}$ such that*

$$T^\Omega f \lesssim T_{2,\mathcal{S}} f$$

pointwise almost everywhere. The implicit constant in the above inequality is absolute.

Pointwise domination of Hölder-continuous Calderón–Zygmund operators by the sparse operator $T_{1,\mathcal{S}}$ is the keystone of Lerner's simple re-proof [**28**] of Hytönen's A_2 theorem [**19**]. Since then, $T_{p,\mathcal{S}}$ have become ubiquitous in singular integral theory, to the point that an exhaustive list of references is well beyond the purview of this article. On the other hand, the sparse square functions $G_{1,\mathcal{S}}$, $G_{p,\mathcal{S}}$ have previously appeared in the context of weighted norm inequalities for square functions of Littlewood–Paley and Marcinkiewicz type; see e.g. [**4, 14, 29**] and references therein. Thus, the specific relevance of the sparse domination principles of Theorems B and C, beyond the strengthening of (2), is explained by the next proposition involving weights and A_p-weight constants, whose standard definitions are also recalled at the end of the introduction.

Proposition 1.1. *The estimates below hold with implicit constants possibly depending only on the exponents p, q appearing therein and in particular independent of the sparse collection $\mathcal{S}$.*

(i) $\|G_{2,\mathcal{S}}\|_{L^2(w)\to L^{2,\infty}(w)} \lesssim [[w]_{A_1} \log(\mathrm{e}+[w]_{A_\infty})]^{\frac12}$.

(ii) $\|T_{2,\mathcal{S}}\|_{L^2(w)\to L^{2,\infty}(w)} \lesssim [[w]_{A_1}[w]_{A_\infty} \log(\mathrm{e}+[w]_{A_\infty})]^{\frac12}$.

(iii) $\|G_{2,\mathcal{S}}\|_{L^p(w)\to L^{p,\infty}(w)} \lesssim [w]_{A_{\frac{p}{2}}}^{\frac1p}[w]_{A_\infty}^{\frac12-\frac1p}, \quad 2<p<\infty.$

(iv) $\|T_{2,\mathcal{S}}\|_{L^p(w)\to L^{p,\infty}(w)} \lesssim [w]_{A_{\frac{p}{2}}}^{\frac1p}[w]_{A_\infty}^{\frac{1}{p'}}, \quad 2<p<\infty.$

(v) $\|G_{2,\mathcal{S}}\|_{L^p(w)} \lesssim \min\{[w]_{A_{\frac{p}{2}}}^{\max\{\frac{1}{p-2},\frac12\}}, [w]_{A_q}^{\frac12}\}, \quad 2\le 2q<p<\infty.$

(vi) $\|T_{2,\mathcal{S}}\|_{L^p(w)} \lesssim \min\{[w]_{A_{\frac{p}{2}}}^{\max\{\frac{1}{p-2},1\}}, [w]_{A_q}\}, \quad 2\le 2q<p<\infty.$

An application of Proposition 1.1 immediately entails two corollaries of our main results.

Corollary B.1. *Estimates (i), (iii), and (v) of Proposition 1.1 hold for the intrinsic smooth square function G^Ω in place of $G_{2,\mathcal{S}}$.*

Corollary C.1. *Estimates (ii), (iv), and (vi) of Proposition 1.1 hold for T^Ω in place of $T_{2,\mathcal{S}}$.*

Proof of Proposition 1.1: Points (i), (iii), and the leftmost estimates in (v) and (vi) are essentially special cases of previously known results. For (i), (iii), and the leftmost estimate in (v), rely on the observation that

$$\|G_{2,\mathcal{S}}\|_{L^p(w)\to L^{p,\infty}(w)} = \|T_{1,\mathcal{S}}\|^{\frac12}_{L^{\frac{p}{2}}(w)\to L^{\frac{p}{2},\infty}(w)}, \quad \|G_{2,\mathcal{S}}\|_{L^p(w)} = \|T_{1,\mathcal{S}}\|^{\frac12}_{L^{\frac{p}{2}}(w)}$$

together with the sharp bound for the appropriate weighted norm of $T_{1,\mathcal{S}}$. The weak-type $L^q(w)$-bound for $T_{1,\mathcal{S}}$ was sharply quantified in [**32**, Theorem 1.2] for $q>1$ and in [**17**, Theorem 1.4] for $q=1$, whence (iii) and (i) respectively; the latter estimate for Calderón–Zygmund operators for $q=1$ is contained in [**31**]. The strong-type $L^q(w)$-bound for $T_{1,\mathcal{S}}$ is classical; see e.g. [**3, 9, 17, 22, 30, 34**]. Finally, the leftmost estimate in (vi) is from [**3**, Proposition 6.4].

The bounds (ii), (iv), and the rightmost estimates in (v) and (vi) seemingly do not appear in past literature. Estimates (ii) and (iv) are obtained by combining (i) and (iii), respectively, with Corollary F.1 below, cf. Section 2. This corollary is a sparse operator version of the exponential square good-λ of Chang, Wilson, and Wolff [**5**]. The rightmost estimate in (v) is obtained by interpolating the weak-type estimates in (iii) for $\frac{p}{2}\in(q,\infty)$. Likewise, the rightmost estimate in (vi) is obtained by interpolating the weak-type estimates of (iv) in the same open range of exponents. □

1.2. On the sharpness of Corollaries B.1 and C.1. As customary in the literature, the term *sharpness* of a weighted estimate in the Muckenhoupt class A_q, say, refers below to whether the functional dependence of the estimate on the weight characteristic $[w]_{A_q}$ is best possible.

With this language, estimate (i) is sharp up to the logarithmic term. It is conceivable that the appearance of such correction is related to whether $L^2(w)$-bounds for G^Ω hold true for all $w \in A_1$, a question that remains open at the time of writing. For G^Ω, the leftmost estimate in (v) is sharp for $p \geq 4$, while estimate (iii) is sharp for all $2 < p < \infty$. Analogously, it is expected that the presence of the logarithmic correction in (ii) is necessary if $L^2(w)$ fails for T^Ω. At the time of writing, we can only show that (ii) is sharp up to the logarithmic term. The leftmost estimate in (vi) is sharp for $p \geq 3$ and estimate (iv) is sharp for all $p > 2$. The rightmost estimates in (v) and (vi) are sharp.

The above claims are verified as follows. The claimed sharpness for strong-type $L^p(w)$-estimates ensues by combining the main results of [**33**] with the fact that the unweighted L^p-bounds for G^Ω are $O(p^{\frac{1}{2}})$, and the unweighted L^p-bounds for T^Ω are $O(p)$ as $p \to +\infty$. Similarly, in order to verify the sharpness of weak $L^p(w)$-estimates, interpolate any two such estimates for p in the open range $(2, \infty)$ with $w \in A_1$ and use [**33**] again.

1.3. Past literature on weighted and sparse bounds for T^Ω and G^Ω. In [**36**, Theorem 6.1], Rubio de Francia proved that G^Ω, and hence T^Ω_{RF}, are bounded on $L^p(w)$ for $2 < p < \infty$ and $w \in A_{\frac{p}{2}}$. The $L^2(w)$-boundedness for $w \in A_1$ of T^Ω_{RF} and G^Ω, conjectured in [**36**, Section 6, p. 10], see also [**15**, Section 8.2, pp. 186–187], remains an open question at the time of writing. This conjecture is corroborated by the fact that it holds for the particular case of congruent intervals [**35**, Theorem A], as well as the partial result that T^Ω_{RF}, G^Ω are $L^2(w)$-bounded for $w(x) = |x|^{-\alpha} \in A_1$, $0 < \alpha < 1$. The latter was proved by Rubio de Francia in [**37**], and a different argument was later given by Carbery in [**39**, pp. 81–93]. Weighted weak-type estimates at the endpoint $p = 2$ were found in [**25**, Theorem B(ii)], yielding the weak variant of Rubio de Francia's conjecture.

Quantitative weighted strong (for $2 < p < \infty$) and weak (at $p = 2$) estimates for T^Ω_{RF} were recently obtained in [**18**, Corollaries 1.5 and 1.6] as a consequence of a sparse form domination [**3, 10**] of the bilinear form for the vector-valued version of the Rubio de Francia square function T^Ω_{RF}, cf. [**18**, Theorem 1.3]. In comparison with the arguments of the present paper, the sparse domination proof of [**18**] relied on a combination of the stopping forms techniques of [**10**] with deeper time-frequency tools, such as vector-valued tree estimates and size decompositions [**2**], circumventing the usual passing through the smooth operator G^Ω. The pointwise sparse bound of Theorem C is formally stronger than the vector-valued sparse estimate of [**18**]. Furthermore, forgoing the vector-valued formalism leads to a simpler argument devoid of vector-valued time-frequency analysis.

In [**18**], the quantification of the behavior of T^Ω_{RF} on $L^p(w)$ is sharp for $3 \leq p < \infty$. On the other hand, the quantitative weighted weak-type estimate at the endpoint $p = 2$ was of order $[w]^{\frac{1}{2}}_{A_1}[w]^{\frac{1}{2}}_{A_\infty} \log(\mathrm{e} + [w]_{A_\infty})$. In the present paper, the weak-type $(2, 2)$ bound of Proposition 1.1(ii) improves by a $[\log(\mathrm{e} + [w]_{A_\infty})]^{\frac{1}{2}}$ term in comparison to [**18**, Corollary 1.6], while the weak (p, p) bound, $2 < p < \infty$, is sharp.

1.4. The strong $L^2(w)$ inequality for the Walsh model. The Rubio de Francia square function T^Ω has an immediate Walsh group analogue. For direct comparison with the trigonometric case, the same notation is kept for corresponding operations, to the extent possible. In stark contrast with the former, we have a proof of $L^2(w)$-boundedness for the Walsh–Rubio de Francia square function. A precise statement is in Theorem D below. Albeit Theorems B and C continue to hold in the Walsh setting, here a sharp endpoint is available, and weighted extrapolation of the $L^2(w)$ results yields better quantified weighted $L^p(w)$-bounds for the Walsh–Rubio de Francia square function than those following from the corresponding sparse domination.

Here follow the definitions relevant to Theorem D. Let $\omega = [k, m)$ be an interval with $k, m \in \mathbb{N}$. Define the Walsh projection operator by

$$H_\omega f(x) := \sum_{n=0}^{\infty} \mathbf{1}_\omega(n)\langle f, W_n\rangle W_n(x), \quad x \in \mathbb{T},$$

where $\{W_n : n \in \mathbb{N}\}$ are the characters of the Walsh group on $\mathbb{T} = [0, 1)$; see (39). For a collection $\omega \in \Omega$ of pairwise disjoint intervals in $\mathbb{N}$ the Walsh–Rubio de Francia square function is

$$T^\Omega f(x) := \left(\sum_{\omega\in\Omega} |H_\omega f(x)|^2\right)^{\frac{1}{2}}, \quad x \in \mathbb{T}.$$

Due to the dyadic nature of the Walsh setting, it suffices to assume dyadic A_p conditions on the weight. The corresponding dyadic constant will be denoted by $A_{p,\mathcal{D}}$.

Theorem D. *Let $w \in A_1$. Then,*

$$\|T^\Omega f\|_{L^2(w)} \lesssim [w]_{A_{1,\mathcal{D}}}^{1/2} [w]_{A_\infty}^{1/2} \|f\|_{L^2(w)}.$$

Furthermore, the sharp bound

$$\|T^\Omega f\|_{L^p(w)} \lesssim [w]_{A_{\frac{p}{2},\mathcal{D}}} \|f\|_{L^2(w)}, \quad 2 < p < \infty,$$

holds with implicit constants depending only on p.

1.5. The strong $L^2(w)$ inequality for radially decreasing A_1-weights. Our final result extends the class of weights for which the $L^2(w)$-boundedness holds to even and radially decreasing A_1-weights in the form of the following theorem, giving new insight on the open question of the $L^2(w)$-boundedness for $w \in A_1$ of the Rubio de Francia square function.

Theorem E. *Let w be an even and radially decreasing A_1-weight on the real line. There holds*

$$\|G^\Omega\|_{L^2(w)} \lesssim [w]_{A_1} \|f\|_{L^2(w)}, \quad \|T^\Omega\|_{L^2(w)} \lesssim [w]_{A_\infty}^{\frac{1}{2}} [w]_{A_1} \|f\|_{L^2(w)}.$$

The proof of Theorem E combines local orthogonality with a stopping time argument and is presented in Section 6. Our argument actually yields the conclusions of Theorem E under the more general, albeit more technical, assumption (43). The latter is in general a strengthening of the A_1 condition, but is equivalent to A_1 for even, radially decreasing weights.

1.6. Notation and generalities. We shall write $X \lesssim Y$ to indicate that $X \le CY$ with a positive constant C independent of significant quantities and we denote $X \simeq Y$ when simultaneously $X \lesssim Y$ and $Y \lesssim X$.

The Fourier transform obeys the normalization

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) \mathrm{e}^{-ix\xi}\, \mathrm{d}x, \quad \xi \in \mathbb{R}.$$

Throughout the article, for $I \subset \mathbb{R}$ being any interval, denote

$$\chi_I(x) := \left[1 + \left(\frac{|x - c_I|}{\ell_I}\right)^2\right]^{-1}, \quad x \in \mathbb{R},$$

with c_I and ℓ_I being respectively the center and length of I. For positive localized averages and for their tailed counterpart, write

$$\langle f\rangle_{p,I} := |I|^{-\frac{1}{p}}\|f\mathbf{1}_I\|_p, \quad \langle f\rangle_{p,I,\dagger} := |I|^{-\frac{1}{p}}\|f\chi_I^9\|_p, \quad 0 < p < \infty.$$

When $p = 1$, the subscript is omitted, simply writing $\langle f\rangle_I$ and $\langle f\rangle_{I,\dagger}$ instead. The chosen 18th order decay is not a relevant feature.

Sparse collections. A collection of intervals $\mathcal{S}$ is called *η-sparse* if for every $I \in \mathcal{S}$ there exists a subset $E_I \subseteq I$ such that

$$|E_I| \geq \eta|I|$$

and the collection of sets $\{E_I : I \in \mathcal{S}\}$ is pairwise disjoint. In this article, the exact value of η may vary at each occurrence, although there is an absolute constant $\eta_0 > 0$ which bounds from below each occurrence of η. In accordance, η is omitted when referring to *η-sparse* collections.

Dyadic grids. The standard system of shifted dyadic grids on $\mathbb{R}$, see e.g. [**30**], is

$$\mathcal{D}^j = \left\{2^{-n}\left[k + \frac{(-1)^n j}{3}, k + 1 + \frac{(-1)^n j}{3}\right) : k, n \in \mathbb{Z}\right\}, \quad j = 0, 1, 2.$$

The superscript j in $\mathcal{D}^j$ is omitted whenever fixed and clear from the context. If I is an interval, write $\mathcal{D}(I) = \{J \in \mathcal{D} : J \subseteq I\}$. For $k \geq 0$, $j \in \mathbb{Z}$, and $Q \in \mathcal{D}$, denote by $Q^{(k)} \in \mathcal{D}$ the k-th dyadic parent of Q and define $Q^{(k,j)} := Q^{(k)} + j\ell_{Q^{(k)}}$, which also belongs to $\mathcal{D}$. To each $Q \in \mathcal{D}$, associate an instance of the decomposition

$$\mathcal{D} = \left(\bigcup_{|j|\leq 1} \mathcal{D}(Q^{(0,j)})\right) \cup \left(\bigcup_{\substack{k\geq 1\\ |j|\leq 1}} \{Q^{(k,j)}\}\right) \cup \left(\bigcup_{\substack{k\geq 0\\ 2\leq|j|\leq 3}} \mathcal{D}(Q^{(k,j)})\right). \tag{4}$$

Equality (4) will be used in connection with tail estimates. It can be easily obtained as a consequence of the dyadic covering

$$5J^{(1)} \setminus 5J = J^{(1,-2)} \cup J^{(1,2)} \cup J^{(0,3\sigma)},$$

holding for each $J \in \mathcal{D}$, with $\sigma = 1$ if J is a left child of $J^{(1)}$, and $\sigma = -1$ otherwise. Indeed, let $Q \in \mathcal{D}$ and $\mathcal{J}$ be the maximal elements of $\mathcal{D}$ contained in $\mathbb{R} \setminus 5Q$. The elements of $\mathcal{J}$ partition $\mathbb{R} \setminus 5Q$ and one has the disjoint union

$$\mathcal{J} = \bigcup_{k=0}^{\infty} \mathcal{J}_k, \quad \mathcal{J}_k := \{J \in \mathcal{D} : J \cap 5Q^{(k,0)} = \varnothing,\ J \subset 5Q^{(k+1,0)}\}.$$

Notice that $\mathcal{J}_k$ is a partition of $5Q^{(k+1,0)} \setminus 5Q^{(k,0)}$. The maximality of $\mathcal{J}$ and the initial observation forces the equality

$$\mathcal{J}_k = \{Q^{(k+1,-2)}, Q^{(k+1,2)}, Q^{(k,3\sigma)}\}$$

for some $\sigma \in \{-1, 1\}$. Therefore, for a generic $I \in \mathcal{D}$, there are the following possibilities.

1. $I \subset 5Q$. Then $I \in \mathcal{D}(Q^{(0,j)})$ for some $|j| \leq 2$.
2. $I \cap 5Q \neq \varnothing$, $I \not\subset 5Q$. Then $I = Q^{(k,j)}$ for some $|j| \leq 2$ and $k \geq 1$.
3. $I \subset \mathbb{R} \setminus 5Q$. Then $I \subset J$ for some $J \in \mathcal{J}_k$ and $k \geq 0$, whence $I \subset Q^{(k,j)}$ for some $k \geq 0$ and $2 \leq |j| \leq 3$.

From here, equality (4) is deduced. If $Q \in \mathcal{D}$ is a dyadic cube, it is convenient to introduce the non-dyadic dilates of Q

$$\widetilde{Q^{(k)}} := \bigcup_{|j|\leq 2} Q^{(k;j)}, \quad k \geq 0, \quad \mathcal{R}(Q) := \{\widetilde{Q^{(k)}} : k \geq 0\}. \tag{5}$$

Note that $5Q^{(k)} = \widetilde{Q^{(k)}}$ and that $\mathcal{R}(Q)$ is a sparse collection, two facts used on several occasions below.

A general principle is that the operators associated to a sparse collection $\mathcal{S}$ may be estimated pointwise by a finite sum of operators associated to sparse collections coming from dyadic grids. More precisely, the three-grid lemma ([**30**, Theorem 3.1]) may be easily used to deduce that for each sparse collection $\mathcal{S}$ there exist sparse collections $\mathcal{S}^j \subset \mathcal{D}^j$, $j = 0, 1, 2$, such that, cf. (3),

$$T_{p,\mathcal{S}} f \lesssim \sum_{j=0,1,2} T_{p,\mathcal{S}^j} f, \quad G_{p,\mathcal{S}} f \lesssim \sum_{j=0,1,2} G_{p,\mathcal{S}^j} f,$$

pointwise, with implicit constants depending on p only. Any quasi-Banach function space operator norm estimate for operators (3) may thus be reduced to the case where $\mathcal{S}$ is a subset of a dyadic grid $\mathcal{D}$.

Weight characteristics. A *weight* w on $\mathbb{R}$ is a positive, locally integrable function. For $1 \leq p \leq \infty$, the A_p characteristic of w is defined by

$$[w]_{A_p} := \begin{cases} \sup_I \langle w \rangle_{1,I} (\inf_I w)^{-1}, & p = 1, \\ \sup_I \langle w \rangle_{1,I} \langle w^{-1} \rangle_{\frac{1}{p-1},I}, & 1 < p < \infty, \\ \sup_I \langle \mathrm{M}(w\mathbf{1}_I) \rangle_{1,I} \langle w \rangle_{1,I}^{-1}, & p = \infty, \end{cases}$$

where the suprema are being taken over all intervals $I \subset \mathbb{R}$ and M is the Hardy–Littlewood maximal function. Note that our definition of A_∞ coincides with that of Wilson, see e.g. [**17, 23, 40**], and that

$$[w]_{A_\infty} \lesssim [w]_{A_p} \leq [w]_{A_q}, \quad 1 \leq q < p < \infty,$$

with absolute implicit constant; see [**23**]. The formal definition of the dyadic A_p characteristic $[w]_{A_p,\mathcal{D}}$ is the same as the usual A_p constant, with the supremum therein being replaced by the supremum over all intervals in $\mathcal{D}^0(\mathbb{T})$, where $\mathbb{T} = [0, 1)$.

Structure of the paper. Section 2 introduces the sparse operators (3) as special cases of balayages of Carleson sequences and contains two relevant results: a weighted exponential good-λ inequality for balayages and a pointwise domination of balayages by a sparse operator. Sections 3 and 4 are devoted to the proofs of Theorem B and Theorem C, respectively. Theorem D is shown in Section 5, and Theorem E is proved in Section 6.

Acknowledgments. The development of this article began during F. Di Plinio's visit to the Basque Center of Applied Mathematics (BCAM) of Bilbao, Spain. The author gratefully acknowledges BCAM's warm hospitality. The authors would also like to thank Dariusz Kosz for many helpful suggestions on the first manuscript of this paper.

Our gratitude is also extended to the anonymous referees for an expert reading and for providing several helpful and detailed comments.

2. Balayages of Carleson sequences

Let $\mathcal{D}$ be a fixed dyadic grid and $\mathsf{a} = \{a_Q : Q \in \mathcal{D}\}$ be any sequence of complex numbers. If $\mathcal{E} \subset \mathcal{D}$, the $\mathcal{E}$-*balayage* of a is defined by

$$A_{\mathcal{E}}[\mathsf{a}] := \sum_{Q \in \mathcal{E}} |a_Q| \mathbf{1}_Q. \tag{6}$$

Remark 2.1. Sparse operators are special cases of (6). Indeed, if $0 < p < \infty$ and $f \in L^p_{\mathrm{loc}}(\mathbb{R})$,

$$T_{p,\mathcal{S}} f = A_{\mathcal{S}}[\{\langle f \rangle_{p,Q} : Q \in \mathcal{D}\}], \quad G_{p,\mathcal{S}} f = \sqrt{A_{\mathcal{S}}[\{\langle f \rangle_{p,Q}^2 : Q \in \mathcal{D}\}]}.$$

2.1. An exponential good-$\boldsymbol{\lambda}$ inequality for sparse balayages. The next theorem is an exponential good-λ inequality for balayages supported on sparse collections. Its corollary has been used in the deduction of estimates (ii), (iv) of Proposition 1.1 respectively from (i), (iii) of the same proposition. For ease of notation, given a complex sequence $\mathsf{a} = \{a_Q : Q \in \mathcal{D}\}$, indicate by $\mathsf{a}^2 := \{a_Q^2 : Q \in \mathcal{D}\}$.

Theorem F. *Let $w \in A_\infty$. There exist absolute constants $C, \delta > 0$ such that the following holds. Let $\mathcal{S} \subset \mathcal{D}$ be a sparse collection and $\mathsf{a} = \{a_Q : Q \in \mathcal{D}\}$ be any sequence. Then for all $\lambda, \gamma > 0$,*

$$w(\{A_{\mathcal{S}}[\mathsf{a}] > 2\lambda, \sqrt{A_{\mathcal{S}}[\mathsf{a}^2]} \leq \gamma\lambda\}) \leq C \exp\left(-\frac{\delta\gamma^2}{[w]_{A_\infty}}\right) w(\{A_{\mathcal{S}}[\mathsf{a}] > \lambda\}).$$

Corollary F.1. *Let $0 < q, s < \infty, 0 < r, t \leq \infty$. Then*

$$\sup \|T_{2,\mathcal{S}} : L^{q,r}(w) \to L^{s,t}(w)\| \lesssim [w]_{A_\infty}^{\frac{1}{2}} \sup \|G_{2,\mathcal{S}} : L^{q,r}(w) \to L^{s,t}(w)\|$$

with the supremum taken over all not necessarily dyadic sparse collections $\mathcal{S}$, and implied constant depending on q, r, s, t only.

Proof of Theorem F and Corollary F.1: First, in view of Remark 2.1, Corollary F.1 follows from the theorem by standard good-λ method. To prove the theorem, by monotone convergence, it suffices to prove the claim for finite sparse collections $\mathcal{S}$ as long as the estimate obtained is uniform in $\#\mathcal{S}$. Denote $\mathcal{S}(Q) = \{Z \in \mathcal{S} : Z \subseteq Q\}$ for each $Q \in \mathcal{D}$. Also denote by F_λ and E_λ the sets appearing respectively in the left and right hand side of the conclusion of the theorem. Under our qualitative assumptions the set E_λ is a finite union of intervals of $\mathcal{D}$, whence $E_\lambda = \cup\{R : R \in \mathcal{R}\}$ and $\mathcal{R}$ is the collection of those elements of $\mathcal{D}$ contained in E_λ and maximal with respect to inclusion. Pairwise disjointness of the collection $\mathcal{R}$ thus reduces our claim to proving

$$w(F_\lambda \cap R) \leq C \exp\left(-\frac{\delta\gamma^2}{[w]_{A_\infty}}\right) w(R), \quad R \in \mathcal{R}. \tag{7}$$

If $x \in F_\lambda \cap R$, then

$$\begin{aligned} 2\lambda < A_{\mathcal{S}}[\mathsf{a}](x) &= A_{\mathcal{S}(R)}[\mathsf{a}](x) + \sum_{\substack{Z \in \mathcal{S} \\ Z \supseteq R^{(1)}}} |a_Z| \leq A_{\mathcal{S}(R)}[\mathsf{a}](x) + \lambda \\ &\leq \sqrt{A_{\mathcal{S}(R)}[\mathsf{a}^2](x)} \Bigg(\sum_{Z \in \mathcal{S}(R)} 1_Z\Bigg)^{\frac{1}{2}} + \lambda \leq \gamma\lambda \Bigg(\sum_{Z \in \mathcal{S}(R)} 1_Z\Bigg)^{\frac{1}{2}} + \lambda. \end{aligned}$$

For the second inequality on the first line we have used that $R^{(1)} \not\subset E_\lambda$ due to maximality of R. Therefore

$$F_\lambda \cap R \subset \left\{ x \in R : \sum_{Z \in \mathcal{S}(R)} 1_Z > \gamma^{-2} \right\}$$

and (7) follows from the weighted John–Nirenberg inequality. The proof of the theorem is thus complete. □

2.2. Subordinated Carleson sequences. Let $f \in L^1(\mathbb{R}^d)$ be a fixed function. Say that the sequence $\mathsf{a} = \{a_I : I \in \mathcal{D}\}$ is a *Carleson sequence subordinated to* f if

$$\frac{1}{|I|} \sum_{J \in \mathcal{D}(I)} |J||a_J| \le C\langle f\rangle_{I,\dagger} \tag{8}$$

uniformly over all $I \in \mathcal{D}$. The least constant C such that (8) holds is denoted by $\|\mathsf{a}\|_{\mathcal{D}}$ and termed *the Carleson norm* of a. The next proposition shows that balayages of Carleson sequences subordinated to f are dominated by 1-average sparse operators applied to f.

Proposition 2.2. *There exists an absolute constant C such that the following holds. For each $f \in L^1(\mathbb{R})$ with compact support there exists a sparse collection $\mathcal{S}$ of intervals with the property that for all Carleson sequences a subordinated to f there holds*

$$A_{\mathcal{D}}[\mathsf{a}] \le C\|\mathsf{a}\|_{\mathcal{D}} T_{1,\mathcal{S}} f, \tag{9}$$

pointwise almost everywhere.

The remainder of this section is devoted to the proof of Proposition 2.2. Fix a compactly supported function f, and choose $Q \in \mathcal{D}$ with $\operatorname{supp} f \subset (1+3^{-1})Q$. Further, fix $\mathsf{a} = \{a_I : I \in \mathcal{D}\}$ subordinated to f, and without loss of generality assume $a_I \ge 0$, and normalize $\|\mathsf{a}\|_{\mathcal{D}} = 1$. As a is fixed, for a generic collection $\mathcal{E}$, the notation $A_{\mathcal{E}}$ is used in the proof in place of $A_{\mathcal{E}}[\mathsf{a}]$.

For the proof of (9), note that (4) readily yields the splitting

$$A_{\mathcal{D}} \le \sum_{|j| \le 1} A_{\mathcal{D}(Q^{(0,j)})} + \sum_{\substack{k \ge 1 \\ |j| \le 1}} |a_{Q^{(k,j)}}| \mathbf{1}_{Q^{(k,j)}} + \sum_{\substack{k \ge 0 \\ 2 \le |j| \le 3}} A_{\mathcal{D}(Q^{(k,j)})}. \tag{10}$$

The proof is articulated into two constructions. The main term in (10) is the $|j| \le 1$ summation while the last two summands entail error terms. We first deal with those.

2.3. Tails. Here we control the latter two sums on the right hand side of (10). We note preliminarily that

$$\begin{aligned} I \in \mathcal{D},\ I^{(\ell,0)} \in \{Q^{(k,\pm 2)}, Q^{(k,\pm 3)}\} &\Longrightarrow \operatorname{dist}(\operatorname{supp} f, I) \gtrsim 2^{-\ell}\ell_I \\ &\Longrightarrow \langle f\rangle_{I,\dagger} \lesssim 2^{-6\ell}\langle f\rangle_{\widetilde{Q^{(k)}}}. \end{aligned} \tag{11}$$

To obtain the last implication we have used that $\|\chi_I^9 \mathbf{1}_{\operatorname{supp} f}\|_\infty \lesssim 2^{-9\ell}$, $\operatorname{supp} f \subset \widetilde{Q^{(k)}}$, and $|I| \gtrsim 2^{-\ell}|\widetilde{Q^{(k)}}|$. The second summand in (10) is estimated by the Carleson condition for a single scale, as follows:

$$\sum_{\substack{k \ge 1 \\ |j| \le 1}} |a_{Q^{(k,j)}}| \mathbf{1}_{Q^{(k,j)}} \le \sum_{\substack{k \ge 1 \\ |j| \le 1}} \langle f\rangle_{Q^{(k,j)},\dagger} \mathbf{1}_{Q^{(k,j)}} \lesssim T_{1,\mathcal{R}(Q)} f, \tag{12}$$

using the notation of Subsection 1.6, cf. (5) in particular. For the third summand in (10) we also proceed via a single scale. Indeed, applying the Carleson condition at the second step, and following with (11), we have for $\tau \in \{\pm 2, \pm 3\}$ that

$$A_{\mathcal{D}(Q^{(k,\tau)})} = \sum_{\ell \geq 0} \sum_{\substack{I \in \mathcal{D} \\ I^{(\ell,0)} = Q^{(k,\tau)}}} |a_I| \mathbf{1}_I \leq \sum_{\ell \geq 0} \sum_{\substack{I \in \mathcal{D} \\ I^{(\ell,0)} = Q^{(k,\tau)}}} \langle f \rangle_{I,\dagger} \mathbf{1}_I$$

$$\lesssim \sum_{\ell \geq 0} 2^{-6\ell} \sum_{\substack{I \in \mathcal{D} \\ I^{(\ell,0)} = Q^{(k,\tau)}}} \langle f \rangle_{\widetilde{Q^{(k)}}} \mathbf{1}_I \lesssim \langle f \rangle_{\widetilde{Q^{(k)}}} \mathbf{1}_{\widetilde{Q^{(k)}}}.$$

The last inequality shows that the third summation in (10) is also controlled by the sparse operator $T_{1,\mathcal{R}(Q)}$ as on the rightmost side of (12).

2.4. Main term. The main term in (10) will be controlled via the intermediate estimate

$$A_{\mathcal{D}(I)} \lesssim T_{1,\mathcal{Q}(I),\dagger} f, \quad T_{1,\mathcal{Q},\dagger} f := \sum_{I \in \mathcal{Q}} \langle f \rangle_{1,I,\dagger} \mathbf{1}_I \tag{13}$$

for each $I \in \{Q^{(0,j)}, |j| \leq 1\}$, where $\mathcal{Q}(I)$ is a suitably constructed sparse collection. Then,

$$T_{1,\mathcal{Q}(I),\dagger} f \lesssim \sum_{k \geq 0} 2^{-8k} \sum_{I \in \mathcal{Q}(I)} \langle f \rangle_{2,2^k I} \mathbf{1}_I$$

so that two applications of [**8**, Theorem A], cf. [**8**, proof of Corollary A.1], upgrade (13) to

$$A_{\mathcal{D}(I)} \lesssim T_{1,\mathcal{Q}'(I)} f, \quad I \in \{Q^{(0,j)}, |j| \leq 1\}$$

with a possibly different sparse collection $\mathcal{Q}'(I)$. Combining these bounds with the estimates of Subsection 2.3 completes the proof of (9), and in turn of Proposition 2.2.

The proof of (13) is a simple John–Nirenberg type iteration argument: details are as follows. For each $R \in \mathcal{D}(I)$, define the collection

$$\mathcal{S}(R) := \text{maximal elements of } \left\{ Z \in \mathcal{D}(R) : \sum_{\substack{W \in \mathcal{D}(R) \\ Z \subset W}} a_W > 4 \langle f \rangle_{R,\dagger} \right\}.$$

As $\mathcal{S}(R)$ is a pairwise disjoint collection, an application of the Carleson condition in the last step yields the packing estimate

$$\sum_{Z \in \mathcal{S}(R)} |Z| \leq \frac{1}{4 \langle f \rangle_{R,\dagger}} \sum_{Z \in \mathcal{S}(R)} \int_Z A_{\mathcal{D}(R)} \leq \frac{\|A_{\mathcal{D}(R)}\|_1}{4 \langle f \rangle_{R,\dagger}} \leq \frac{|R|}{4}$$

while, setting $\mathcal{D}^\star(R) := \mathcal{D}(R) \setminus \bigcup_{Z \in \mathcal{S}(R)} \mathcal{D}(Z)$,

$$A_{\mathcal{D}(R)} \leq A_{\mathcal{D}^\star(R)} + \sum_{Z \in \mathcal{S}(R)} A_{\mathcal{D}(Z)} \leq 4 \langle f \rangle_{R,\dagger} + \sum_{Z \in \mathcal{S}(R)} A_{\mathcal{D}(Z)}. \tag{14}$$

Setting $\mathcal{Q}_0 \coloneqq \{I\}$, inductively define

$$\mathcal{Q}_{k+1} \coloneqq \bigcup_{R\in\mathcal{Q}_k} \mathcal{S}(R), \quad k=0,1,\dots, \quad \mathcal{Q}(I) \coloneqq \bigcup_{k\geq 0} \mathcal{Q}_k,$$

and observe that the previously obtained packing estimate ensures $\mathcal{Q}(I)$ is a sparse collection. Finally, iterating (14),

$$A_{\mathcal{D}(I)} \leq \sum_{R\in\mathcal{Q}} A_{\mathcal{D}^\star(R)} \leq 4T_{1,\mathcal{Q}(I),\dagger}f,$$

which is the claimed (13).

3. Proof of Theorem B

The proof of Theorem B relies upon a suitable discretization of G^Ω into a wave packet coefficient square function. It is not difficult to show that the square sum of the wave packet coefficients of f localized on a single spatial interval is a Carleson sequence subordinated to $|f|^2$, so that the claim of the theorem readily follows from Proposition 2.2.

We turn to the details. For $j=0,1,2$ define the corresponding *j-th tile universe* $\mathbb{S}^j \subset \mathcal{D}^0\times\mathcal{D}^j$ as the set of those $I\times\omega \in \mathcal{D}^0\times\mathcal{D}^j$ with $\ell_I\ell_\omega = 1$. The superscript j is omitted whenever fixed and clear from the context. As customary, the notation $s = I_s\times\omega_s$ is employed for $s\in\mathbb{S}$. If $\mathbb{P}\subset\mathbb{S}$, we write

$$\mathbb{P}(I) = \{s\in\mathbb{P} : I_s = I\}, \quad \mathbb{P}_{\subseteq}(I) = \{s\in\mathbb{P} : I_s\subseteq I\} \tag{15}$$

for each interval $I\subset\mathbb{R}$. For our purposes, we are especially interested in subcollections of tiles whose frequency intervals are pairwise disjoint, the precise definition being as follows. If $\Omega\subset\mathcal{D}$ is a collection of pairwise disjoint intervals, write

$$\mathbb{P}^\Omega \coloneqq \{s\in\mathbb{P} : \omega_s = \omega \text{ for some } \omega\in\Omega\}. \tag{16}$$

With these notations, let $\mathbb{P}^\Omega(I) \coloneqq \{s\in\mathbb{P}^\Omega : I_s = I\}$ for each interval $I\subset\mathbb{R}$. Fix a large integer D. To each tile $s = I_s\times\omega_s$, recalling that c_{ω_s} denotes the center of ω_s, we associate the L^1-normalized wavelet class Ψ_s consisting of those $\phi\in\mathcal{C}^\infty(\mathbb{R})$ with

$$\operatorname{supp}\widehat{\phi}\subset\omega_s, \quad \sup_{0\leq j\leq D} |I_s|^{1+j}\|\chi_{I_s}^{-D}(\exp(ic_{\omega_s}\cdot)\phi)^{(j)}\|_\infty \leq 1. \tag{17}$$

The intrinsic wave packet coefficient of $f\in L^2(\mathbb{R})$ is then defined by the maximal quantity

$$s(f) \coloneqq \sup_{\phi\in\Psi_s} |\langle f,\phi\rangle|, \quad s\in\mathbb{S}. \tag{18}$$

The coefficients (18) may be used to construct a smooth, approximately localized analogue of the L^2-norm of f on the torus $I\in\mathcal{D}$. Namely, if Ω is a collection of pairwise disjoint dyadic intervals, set

$$[f]_{\mathbb{S}^\Omega(I)} \coloneqq \Bigg(\sum_{s\in\mathbb{S}^\Omega(I)} s(f)^2\Bigg)^{\frac12}, \quad I\in\mathcal{D}.$$

The next lemma shows that whenever $\Omega\subset\mathcal{D}$ is a pairwise disjoint collection, and $f\in L^2(\mathbb{R})$, the sequence $\{[f]^2_{\mathbb{S}^\Omega(I)} : I\in\mathcal{D}\}$ is a Carleson sequence subordinated to the function $|f|^2$.

Lemma 3.1. *There holds* $\sum_{J\in\mathcal{D}(I)} |J|[f]^2_{\mathbb{S}^\Omega(J)} \lesssim |I|\langle|f|^2\rangle_{I,\dagger}$ *uniformly over* $I\in\mathcal{D}$.

Proof: Fix $I \in \mathcal{D}$. It suffices to show that

$$\sum_{J\in\mathcal{D}(I)} \sum_{s\in\mathbb{S}^\Omega(J)} |J||\langle f, \phi_s\rangle|^2 \lesssim |I|\langle |f|^2\rangle_{I,\dagger} \tag{19}$$

for an arbitrary choice of $\phi_s \in \Psi_s$, for each $J \in \mathcal{D}(I)$ and $s \in \mathbb{S}^\Omega(J)$. Set $\varphi_s := \chi_I^{-9}\phi_s$. Due to localization and to the pairwise disjoint nature of the collection Ω, the almost orthogonality estimate

$$|\langle \varphi_s, \varphi_{s'}\rangle| \begin{cases} = 0, & \omega_s \neq \omega_{s'}, \\ \lesssim |I_s|^{-1}\operatorname{dist}(I_s, I_{s'})^{-100}, & \omega_s = \omega_{s'}, \end{cases}$$

holds for all $s, s' \in \mathbb{S}^\Omega$ with $I_s, I_{s'} \in \mathcal{D}(I)$. A standard TT^* type argument, see for example [**1**, §4.3], yields the almost orthogonality bound

$$\sum_{J\in\mathcal{D}(I)} \sum_{s\in\mathbb{S}^\Omega(J)} |J||\langle g, \varphi_s\rangle|^2 \lesssim \|g\|_2^2 \tag{20}$$

and (19) follows by applying (20) to $g = f\chi_I^9$ and relying on the definition of $\langle\cdot\rangle_{I,\dagger}$. □

A combination of Proposition 2.2 and Lemma 3.1 immediately yields a sparse domination result for the intrinsic wave packet square function

$$W^\Omega f := \left(\sum_{s\in\mathbb{S}^\Omega} s(f)^2 \mathbf{1}_{I_s}\right)^{\frac{1}{2}} = \sqrt{A_{\mathcal{D}}[\{[f]^2_{\mathbb{S}^\Omega(I)} : I \in \mathcal{D}\}]}. \tag{21}$$

Proposition 3.2. *Let $f \in L^2(\mathbb{R})$ be a compactly supported function and $\Omega \subset \mathcal{D}$ be a pairwise disjoint collection. Then there exists a sparse collection $\mathcal{S}$ with the property that*

$$W^\Omega f \lesssim G_{2,\mathcal{S}} f$$

pointwise almost everywhere. The implicit constant in the above inequality is absolute.

Indeed, by Lemma 3.1, $\{[f]^2_{\mathbb{S}^\Omega(I)} : I \in \mathcal{D}\}$ is a Carleson sequence subordinated to the function $|f|^2$. Thus Proposition 3.2 is obtained via an application of Proposition 2.2, followed by the observation that $\sqrt{T_{1,\mathcal{S}}(|f|^2)} = G_{2,\mathcal{S}} f$.

3.1. Sparse estimates for smooth square functions: proof of Theorem B. The relation of G^Ω with the wave packet square function W^Ω defined above is given by the pointwise estimate

$$G^\Omega f \lesssim \sup_{1\le k\le 9} W^{\Omega^{k,\star}} f, \tag{22}$$

where each $\Omega^{k,\star}$, $1 \le k \le 9$, is a collection of pairwise disjoint intervals contained in one of the three grids $\mathcal{D}^j$, $j = 0, 1, 2$. To obtain this pointwise bound, associate to each $\omega \in \Omega$ an index $j \in \{0, 1, 2\}$ and a smoothing interval $\omega^\star \in \mathcal{D}^j$, that is the unique interval of $\mathcal{D}^j$ with $\omega \subset \omega^\star$ and $3\ell_\omega \le \ell_{\omega^\star} < 6\ell_\omega$. As the intervals of Ω are pairwise disjoint, Ω can be split into collections Ω^k, $1 \le k \le 9$, with the property that $\Omega^{k,\star} := \{\omega^\star : \omega \in \Omega^k\} \subset \mathcal{D}^j$ for some j and is a pairwise disjoint collection. A standard discretization procedure, see for example [**1**, Lemma 5.9], then entails

$$G^{\Omega^k} f \lesssim W^{\Omega^{k,\star}} f$$

and (22) follows. Finally we may combine Proposition 3.2 with (22) to conclude Theorem B, using also that the union of nine sparse collections is still a sparse collection; see e.g. [**30**].

4. Proof of Theorem C

The proof of Theorem C, finalized at the end of this section, rests on a well-known Littlewood–Paley type reduction to a model time-frequency square function appearing on the left hand side of (24), which we now introduce.

4.1. Time-frequency square function. Fix a dyadic grid $\mathcal{D}$. Given an interval ω we let $\boldsymbol{\omega} \subset \mathcal{D}$ be a collection of dyadic intervals with the following properties.

(i) The collection $\boldsymbol{\omega}$ is pairwise disjoint.

(ii) For each $k \in \mathbb{Z}$ there exists at most one $\alpha \in \boldsymbol{\omega}$ with $\ell_\alpha = 2^k$.

(iii) Each $\alpha \in \boldsymbol{\omega}$ satisfies $7^3\alpha \subset \omega$ and $7^4\alpha \not\subset \omega$.

Observe that by (iii) the collection $\boldsymbol{\omega}$ is a subcollection of a dyadic Whitney covering of ω but not necessarily the whole Whitney cover. As a result properties (i) and (ii) can always be achieved by splitting ω into finitely many subcollections. Let $I \subset \mathbb{R}$ be any interval, possibly unbounded. Recalling the definitions (15), (16), (17), let $\Phi_{\mathbb{S}} \coloneqq \{\varphi_s \in \Psi_s,\, \vartheta_s \in |I_s|\Psi_s : s \in \mathbb{S}\}$ be a choice of wave packets. We say that

$$P^{\Phi_{\mathbb{S}}}_{I,\omega,\boldsymbol{\omega}} f \coloneqq \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(I)} \langle f, \varphi_s\rangle \vartheta_s, \quad \omega \in \Omega,$$

is a *time-frequency projection* of f on the time-frequency region $I \times \omega$. Note the L^1, L^∞ normalizations of φ_s, ϑ_s respectively. We will drop the subindex $\mathbb{S}$ from $\Phi_{\mathbb{S}}$ for the rest of the section and, whenever the choices of Φ and $\boldsymbol{\omega}$ are fixed and clear from the context, we will simplify the notation by writing $P_{I,\omega}$, suppressing the dependence on $\Phi_{\mathbb{S}}$ and $\boldsymbol{\omega}$. It is easy to check that

$$\widetilde{P_{I,\omega}} f \coloneqq P_{I,\omega}(\chi_I^{-9} f)$$

is a standard Calderón–Zygmund operator, whence the estimates

$$\|P_{I,\omega} f\|_p \lesssim pp' |I|^{\frac{1}{p}} \langle f\rangle_{p,I,\dagger}, \quad 1 < p < \infty, \tag{23}$$

hold uniformly over all bounded intervals I, which we will only use for $p = 2$. Furthermore, due to the frequency localization of the Ψ_s classes for $s \in \mathbb{S}^{\boldsymbol{\omega}}$, the equality

$$P_{I,\omega} f = P_{I,\omega} H_\omega f$$

holds for the frequency projection H_ω defined in (1). The next theorem is a sparse domination principle for the square function $\|P_{\mathbb{R},\omega}\|_{\ell^2(\omega\in\Omega)}$ under the pairwise disjointness assumption of the corresponding collection $\omega \in \Omega$.

Proposition 4.1. *Let Φ be a choice of wave packets, $f \in L^2(\mathbb{R})$ be a compactly supported function, and Ω be a qualitatively finite, pairwise disjoint collection of intervals. Then there exists a sparse collection $\mathcal{S}$ depending on Φ, f, Ω only with the property that*

$$\|P^{\Phi}_{\mathbb{R},\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim T_{2,\mathcal{S}} f \tag{24}$$

pointwise almost everywhere, with implicit absolute numerical constant.

The proof of the proposition is given in Subsection 4.3. It relies on two lemmas which we state now. The first deals with domination of tails.

Lemma 4.2. *Let $J \in \mathcal{D}$ and $\mathrm{M}_2 f := (\mathrm{M}|f|^2)^{\frac{1}{2}}$. Let Φ be a choice of wave packets. The following pointwise bounds hold.*

(i) $\displaystyle \Big\| \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}}(J)} \langle f, \varphi_s \rangle \vartheta_s \Big\|_{\ell^2(\omega \in \Omega)} \lesssim \chi_J^9 \langle f \rangle_{2,J,\dagger}$.

(ii) *Suppose* $\operatorname{dist}(x, J) \gtrsim \ell_J$. *Then* $\|P_{J,\omega} f(x)\|_{\ell^2(\omega\in\Omega)} \lesssim \chi_J^6 \mathrm{M}_2 f(x)$.

(iii) *If* $\operatorname{supp} f \subset 2J^{(0,\pm 2)}$, *then* $\|P_{J,\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim \chi_J^9 \langle f \rangle_{2,7J}$.

Proof: The first estimate follows immediately from the two controls

$$\sqrt{\sum_{\omega\in\Omega} \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}}(J)} |\langle f, \varphi_s \rangle|^2} \lesssim \langle f \rangle_{2,J,\dagger}, \qquad \sup_{s \in \mathbb{S}^{\boldsymbol{\omega}}(J)} |\vartheta_s| \lesssim \chi_J^9,$$

the second meant pointwise. To obtain the bound in (ii),

$$\begin{aligned} \|P_{J,\omega} f(x)\|_{\ell^2(\omega\in\Omega)} &\le \sum_{\ell\ge 0} \sum_{\substack{I \in \mathcal{D}(J)\\ I^{(\ell,0)}=J}} \Big\| \sum_{s\in\mathbb{S}^{\boldsymbol{\omega}}(I)} \langle f, \varphi_s\rangle \vartheta_s(x) \Big\|_{\ell^2(\omega\in\Omega)} \\ &\lesssim \sum_{\ell\ge 0} \sum_{\substack{I \in \mathcal{D}(J)\\ I^{(\ell,0)}=J}} \chi_I^9(x) \langle f \rangle_{2,I,\dagger} \\ &\lesssim \mathrm{M}_2 f(x) \sum_{\ell\ge 0} \sum_{\substack{I \in \mathcal{D}(J)\\ I^{(\ell,0)}=J}} \chi_I^8(x) \lesssim \chi_J^6 \mathrm{M}_2 f(x), \end{aligned} \tag{25}$$

having applied (i) with $J = I$ for each I such that $I^{(\ell,0)} = J$. We have employed the easily verified inequalities

$$\chi_I(x) \langle f \rangle_{2,I,\dagger} \lesssim \mathrm{M}_2 f(x), \qquad \sum_{\substack{I \in \mathcal{D}(J)\\ I^{(\ell,0)}=J}} \chi_I^8(x) \lesssim 2^{-\ell} \chi_J^6(x),$$

valid for $I \subset J$, $\operatorname{dist}(x, J) \gtrsim \ell_J$. To obtain the bound of (iii), start again from the right hand side of the first line of (25), and apply (i) for each I in the summation, so that

$$\|P_{J,\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim \sum_{\ell\ge 0} \sum_{\substack{I \in \mathcal{D}(J)\\ I^{(\ell,0)}=J}} \chi_I^9 \langle f \rangle_{2,I,\dagger} \lesssim \chi_J^9 \langle f \rangle_{2,7J} \tag{26}$$

as claimed. We have used that for each I as above, $\operatorname{dist}(\operatorname{supp} f, I) \gtrsim 2^\ell \ell_I$. Together with $\operatorname{supp} f \subset 2J^{(0,\pm 2)} \subset 7J$, it follows that $\langle f \rangle_{2,I,\dagger} \lesssim 2^{-8\ell} \langle f \rangle_{2,7J}$, whence the last inequality in (26). This completes the proof of the lemma. □

The second lemma encapsulates the main iteration of the proof of Proposition 4.1.

Lemma 4.3. *Let $J \in \mathcal{D}$ and $f \in L^2(\mathbb{R})$. Let Φ be a choice of wave packets. Then there exists a sparse collection $\mathcal{Q} = \mathcal{Q}(\Phi, J, f, \Omega)$ with the property that, pointwise almost everywhere,*

$$\|1_J P^{\Phi}_{3J,\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim T_{2,\mathcal{Q}} f. \tag{27}$$

The proof of this lemma is more involved and thus occupies its own subsection.

4.2. Proof of Lemma 4.3. The collection Φ is fixed throughout this proof and thus omitted from the notation. Arguing as in Subsection 2.4, cf. (13), it suffices to prove the weaker result that for some sparse collection $\mathcal{Q}$

$$\|1_J P_{3J,\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim T_{2,\mathcal{Q},\dagger} f, \quad T_{2,\mathcal{Q},\dagger} f := \sum_{I\in\mathcal{Q}} \langle f\rangle_{2,I,\dagger} \mathbf{1}_I, \tag{28}$$

and later upgrade (28) to (27) via [**8**], with a possibly different sparse collection $\mathcal{Q}$.

The proof of (28) rests on an iterative inequality whose first step is a stopping construction. Fix again a large constant Θ to be determined. For each $I \in \mathcal{D}(J)$, define the stopping sets and collections

$$\begin{aligned} E_1(I) &:= \{x \in I : \|\mathrm{M} P_{3I,\omega} f\|_{\ell^2(\omega\in\Omega)} > \Theta\langle f\rangle_{2,I,\dagger}\},\\ E_2(I) &:= \{x \in I : \|\mathrm{M}[H_\omega(f\chi_I^9)]\|_{\ell^2(\omega\in\Omega)} > \Theta\langle f\rangle_{2,I,\dagger}\},\\ \mathcal{S}(I) &:= \{\text{maximal elements } Z \in \mathcal{D} : Z \subset E(I) := E_1(I) \cup E_2(I)\}. \end{aligned}$$

This time, the maximality condition ensures

$$\inf_{3Z} \|\mathrm{M} P_{3I,\omega} f\|_{\ell^2(\omega\in\Omega)} + \inf_{3Z} \|\mathrm{M}[H_\omega(f\chi_I^9)]\|_{\ell^2(\omega\in\Omega)} \lesssim \langle f\rangle_{2,I,\dagger}, \qquad Z \in \mathcal{S}(I), \tag{29}$$

$$\|\mathrm{M} P_{3I,\omega} f(x)\|_{\ell^2(\omega\in\Omega)} \le \Theta\langle f\rangle_{2,I,\dagger}, \quad x \in I \setminus E(I). \tag{30}$$

Now set $\mathcal{Q}_0 := \{J\}$. Proceed inductively, defining

$$\mathcal{Q}_{k+1} := \bigcup_{I\in\mathcal{Q}_k} \mathcal{S}(I), \quad k = 0, 1, \dots, \quad \mathcal{Q} := \bigcup_{k\ge 0} \mathcal{Q}_k.$$

Arguing in the same way as [**6**, proof of equation (2.22)], see also [**7**, Section 4], the fact that $\mathcal{Q}$ is a sparse collection is easily verified once the estimates $|E_j(I)| < 2^{-16}|I|$, $j = 1, 2$, are proved. In the case of $E_1(I)$, provided Θ is large enough, this follows from Chebyshev and

$$\frac{1}{|I|}\int \|\mathrm{M} P_{3I,\omega} f\|^2_{\ell^2(\omega\in\Omega)} \lesssim \frac{1}{|I|}\sum_{\omega\in\Omega} \|\widetilde{P_{3I,\omega}} H_\omega(f\chi_I^9)\|_2^2 \lesssim \frac{1}{|I|}\sum_{\omega\in\Omega} \|H_\omega(f\chi_I^9)\|_2^2 \le \langle f\rangle^2_{2,I,\dagger}$$

having used the maximal theorem in the first step, (23) for the second inequality, and orthogonality of the projections H_ω in the last. A shorter computation leads to the same estimate for $E_2(I)$. The next lemma is the main device that controls the oscillation. The maximal frequency truncation idea dates back to the single tree estimate in Lacey and Thiele's seminal paper on the Carleson operator [**27**]. The proof is given at the end of this subsection.

Lemma 4.4. *Let $Z \in \mathcal{S}(I)$. There holds*

$$\sup_Z |P_{3Z,\omega} f - P_{3I,\omega} f| \lesssim \inf_{3Z} \mathrm{M} P_{3I,\omega} f + \inf_{3Z} \mathrm{M}[H_\omega(f\chi_I^9)].$$

Now for each $I \in \mathcal{D}(J)$, with the stopping collection $\mathcal{S}(I)$ at hand, pick $x \in I$. Then either $x \in I \setminus E(I)$, in which case

$$\|P_{3I,\omega} f(x)\|_{\ell^2(\omega\in\Omega)} \le C\langle f\rangle_{2,I,\dagger},$$

by virtue of (30), or $x \in Z$ for some $Z \in \mathcal{S}(I)$, in which case

$$\|P_{3I,\omega} f(x)\|_{\ell^2(\omega\in\Omega)} \le \|P_{3Z,\omega} f(x)\|_{\ell^2(\omega\in\Omega)} + C\langle f\rangle_{2,I,\dagger}$$

via an application of Lemma 4.4 and (29). It follows that

$$\mathbf{1}_I \|P_{3I,\omega} f\|_{\ell^2(\omega\in\Omega)} \le C\langle f\rangle_{2,I,\dagger} \mathbf{1}_I + \sum_{Z\in\mathcal{Q}} \mathbf{1}_Z \|P_{3Z,\omega} f\|_{\ell^2(\omega\in\Omega)}. \tag{31}$$

Starting from $I = J$, iterate (31) to obtain

$$\mathbf{1}_J \|P_{3J,\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim \sum_{I\in\mathcal{Q}} \langle f\rangle_{2,I,\dagger} \mathbf{1}_I,$$

completing the proof of (28), and in turn of Lemma 4.3.

Proof of Lemma 4.4: Note that

$$\begin{aligned} |P_{3Z,\omega} f - P_{3I,\omega} f| &= \left| \sum_{s\in\mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(3I)\setminus\mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(3Z)} \langle f,\varphi_s\rangle\vartheta_s \right| \\ &\leq \left| \sum_{\substack{s\in\mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(3I)\\ \ell_{I_s}>\ell_Z}} \langle f,\varphi_s\rangle\vartheta_s \right| + \left| \sum_{\substack{|j|\geq 2\\ Z^{(0,j)}\subset 3I}} P_{Z^{(0,j)},\omega} f \right|. \end{aligned} \tag{32}$$

Let us deal with the tail term in (32). The separation between Z and the small scales contained in $Z^{(0,j)}$ for some $j \geq 2$ allows for the standard Calderón–Zygmund tail estimate

$$\begin{aligned} \mathbf{1}_Z \left| \sum_{\substack{|j|\geq 2\\ Z^{(0,j)}\subset 3I}} P_{Z^{(0,j)},\omega} f \right| &= \mathbf{1}_Z \left| \sum_{\substack{|j|\geq 2\\ Z^{(0,j)}\subset 3I}} \sum_{s\in\mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(Z^{(0,j)})} \langle H_\omega(f\chi_I^9),\varphi_s\rangle\vartheta_s \right| \\ &\lesssim \inf_{3Z} \mathrm{M}[H_\omega(f\chi_I^9)]. \end{aligned} \tag{33}$$

The proof is essentially a repetition of the one for Lemma 4.2(ii) using L^1-averages instead, and thus the details are omitted. The first term in (32) is the main term. To deal with it define the sets

$$\beta_Z := \mathrm{Conv}\left(\bigcup\{\alpha\in\boldsymbol{\omega} : \ell_\alpha < (\ell_Z)^{-1}\}\right), \quad \gamma_Z := \mathrm{Conv}\left(\bigcup\{\alpha\in\boldsymbol{\omega} : \ell_\alpha \leq (\ell_Z)^{-1}\}\right).$$

Using the Whitney property of $\boldsymbol{\omega}$, there exist positive constants c_1, c_2, c_3, c_4 with $c_2 - c_1 \simeq 1$, $c_4 - c_3 \simeq 1$ such that if $\omega = [a,b)$,

$$\beta_Z = (a, a + c_1(\ell_Z)^{-1}) \cup (b - c_3(\ell_Z)^{-1}, b) \subset \gamma_Z = (a, a + c_2(\ell_Z)^{-1}) \cup (b - c_4(\ell_Z)^{-1}, b).$$

Therefore, we may choose a smooth function ψ_Z with the properties that

$$|Z|\|\chi_Z^{-9}\psi_Z\|_\infty \lesssim 1, \quad \mathbf{1}_{\beta_Z} \leq \widehat{\psi_Z} \leq \mathbf{1}_{\gamma_Z}.$$

In particular, $\widehat{\psi_Z} = 1$ on ω_s whenever $s \in \mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(3I)$ and $\ell_{I_s} > \ell_Z$, given that in this case $\omega_s \subset \beta_Z$, while $\widehat{\psi_Z} = 0$ on ω_s whenever $s \in \mathbb{S}^{\boldsymbol{\omega}}_{\subseteq}(3I)$ and $\ell_{I_s} < \ell_Z$, given that instead $\omega_s \cap \gamma_Z = \varnothing$. Hence, the main term of (32) equals $P_{3I,\omega} * \psi_Z$, up to removal of the tiles at scale $\ell_{I_s} = \ell_Z$, on whose frequency intervals $\widehat{\psi_Z}$ is not necessarily equal to zero or one. For the details, let $\mathbb{S}(Z,\omega)$ be the set of tiles with $I_s \subset 3I$, $\ell_{I_s} = \ell_Z$, and $\omega_s \in \boldsymbol{\omega}$. The spatial intervals of the tiles $\mathbb{S}(Z,\omega)$ are contained in $3I$, pairwise disjoint and of the same scale ℓ_Z, so that for $x \in Z$

$$\left| \sum_{s\in\mathbb{S}(Z,\omega)} \langle f,\varphi_s\rangle\vartheta_s(x) \right| = \left| \sum_{s\in\mathbb{S}(Z,\omega)} \langle H_\omega(f\chi_I^9), \chi_I^{-9}\varphi_s\rangle\vartheta_s(x) \right| \lesssim \inf_{3Z} \mathrm{M}[H_\omega(f\chi_I^9)]. \tag{34}$$

Then

$$\left|\sum_{\substack{s\in\mathbb{S}^{\omega}_{\subseteq}(3I)\\ \ell_{I_s}>\ell_Z}}\langle f,\varphi_s\rangle\vartheta_s\right| = \left|\left(P_{3I,\omega}f-\sum_{s\in\mathbb{S}(Z,\omega)}\langle f,\varphi_s\rangle\vartheta_s\right)*\psi_Z\right|$$

$$\lesssim \inf_{3Z}\mathrm{M}[P_{3I,\omega}f]+\inf_{3Z}\mathrm{M}[H_\omega(f\chi_I^9)].$$

Together with (33) and (34), this completes the proof of the lemma. □

4.3. Proof of Proposition 4.1. Fix an instance of Ω, Φ, and let f be a fixed compactly supported function in $L^2(\mathbb{R})$. It is possible to choose $Q\in\mathcal{D}$ with the property that $\operatorname{supp} f\subset(1+3^{-1})Q$. A first lemma takes care of the tails

$$P^{\Phi}_{\mathsf{out},\omega}f := P^{\Phi}_{\mathbb{R},\omega}f-\mathbf{1}_{5Q}P^{\Phi}_{3Q,\omega}f,\quad \omega\in\Omega.$$

Lemma 4.5. *With $\mathcal{R}(Q)$ as in* (5)*, there holds*

$$\|P^{\Phi}_{\mathsf{out},\omega}f\|_{\ell^2(\omega\in\Omega)}\lesssim T_{2,\mathcal{R}(Q)}f+\mathrm{M}_2 f \tag{35}$$

pointwise almost everywhere.

Proof: In this proof, Φ is fixed and thus omitted from superscripts. First of all, note that

$$P_{\mathsf{out},\omega}=[P_{\mathbb{R},\omega}f-P_{3Q,\omega}]+\mathbf{1}_{\mathbb{R}\setminus 5Q}P_{3Q,\omega}$$

and the summand outside the square bracket, that is the non-local part of $P_{3Q,\omega}$, is immediately controlled by (ii) of Lemma 4.2. Therefore, it suffices to control the difference $P_{\mathbb{R},\omega}f-P_{3Q,\omega}$, which by (4) satisfies

$$|P_{\mathbb{R},\omega}f-P_{3Q,\omega}f|\le\sum_{|m|\le1}\sum_{k\ge1}\left|\sum_{s\in\mathbb{S}^{\omega}(Q^{(k,m)})}\langle f,\varphi_s\rangle\vartheta_s\right|+\sum_{|m|=2,3}\sum_{k\ge0}|P_{Q^{(k,m)},\omega}f|$$

$$=:\sum_{|m|\le1}\sum_{k\ge1}U_{m,k,\omega}+\sum_{|m|=2,3}\sum_{k\ge0}V_{m,k,\omega}.$$

Applying respectively (i) and (iii) of Lemma 4.2 yields for all $k\ge0$ the pointwise estimates

$$\|U_{m,k,\omega}\|_{\ell^2(\omega\in\Omega)}\lesssim\langle f\rangle_{2,\widetilde{Q^{(k)}}}\chi^9_{\widetilde{Q^{(k)}}},\quad \|V_{m,k,\omega}\|_{\ell^2(\omega\in\Omega)}\lesssim\langle f\rangle_{2,\widetilde{Q^{(k)}}}\chi^9_{\widetilde{Q^{(k)}}}, \tag{36}$$

the first of which holds uniformly over $k\ge1$, $|m|\le1$, while the second holds uniformly over $k\ge0$, $|m|\in\{2,3\}$. For the second control let $\tau\in\{\pm2,\pm3\}$ and apply Lemma 4.2(iii) with $J=Q^{(k,\tau)}$ together with the fact that $\langle f\rangle_{2,7Q^{(k,\tau)}}\simeq\langle f\rangle_{2,\widetilde{Q^{(k)}}}$. The desired estimate follows since

$$\operatorname{supp} f\subset\widetilde{Q^{(k)}},\quad |7Q^{(k,\tau)}|\simeq|\widetilde{Q^{(k)}}|.$$

Now fix m and a point $x\in\mathbb{R}$. Summing (36) up over k and splitting according to whether or not $x\in\widetilde{Q^{(k)}}$ entails the claim of the lemma. □

As $T_{2,\mathcal{R}(Q)}$ is a sparse operator and $\mathrm{M}_2 f$ obeys a sparse bound of the type (24), it remains to control

$$\mathbf{1}_{5Q} P^{\Phi}_{3Q,\omega} = \sum_{\substack{|m|\le 1\\ |j|\le 2}} \mathbf{1}_{Q^{(0,j)}} P^{\Phi}_{Q^{(0,m)},\omega}.$$

Applying Lemma 4.2(ii) we gather that

$$|j-m|>1 \Longrightarrow \|\mathbf{1}_{Q^{(0,j)}} P^{\Phi}_{Q^{(0,m)},\omega} f\|_{\ell^2(\omega\in\Omega)} \lesssim \mathrm{M}_2 f. \tag{37}$$

If $|j-m|\le 1$, we instead have

$$\mathbf{1}_{Q^{(0,j)}} P^{\Phi}_{Q^{(0,m)},\omega} = \mathbf{1}_J P^{\Phi^{j,m}}_{3J,\omega}, \quad \omega\in\Omega, \tag{38}$$

having set $J=Q^{(0,j)}$ and having constructed $\Phi^{j,m}=\{\varphi_s^{j,m},\vartheta_s : s\in\mathbb{S}\}$ from $\Phi=\{\varphi_s,\vartheta_s : s\in\mathbb{S}\}$ as follows: $\varphi_s^{j,m}=\varphi_s$ if $I_s\subset Q^{(0,m)}$ and $\varphi_s^{j,m}=0$ otherwise. Applying Lemma 4.3 to each right hand side of (38) and combining with (35)–(37) completes the proof of Proposition 4.1.

4.4. Rough square functions: proof of Theorem C. Using [**12**, equation (2.10)], we learn that

$$|T_\omega f| \lesssim \sum_{j=1}^{285} |P^{\Phi_j}_{\mathbb{R},\omega,\boldsymbol{\omega}_j} f|, \quad \omega\in\Omega,$$

where, for each $j\in\{1,\dots,285\}$, Φ_j is a suitable collection of wave packets, $\boldsymbol{\omega}_j$ is a collection of dyadic intervals satisfying properties (i)–(iii) of Subsection 4.1, and both Φ_j and $\boldsymbol{\omega}_j$ are constructed on the fixed shifted grid $\mathcal{D}^{k_j}$, for some $k_j\in\{0,1,2\}$. Proposition 4.1 then immediately implies the conclusion of Theorem C.

5. The Walsh case

In this section, we will prove Theorem D. The strategy of proof involves a Walsh version of the wave packet square function W^Ω. An $L^2(w)$-quantitative relation between this object and the Walsh–Rubio de Francia square function is provided by the Chang–Wilson–Wolff inequality, [**40**, Theorem 3.4].

5.1. The setting for the Walsh model. Let $\mathbb{T}=[0,1)$ be the 1-torus. For $k=0,1,2,\dots$, define the *Walsh function* W_{2^k}

$$W_{2^k}(x) := \mathrm{sign}(\sin(2^{k+1}\pi x)), \quad x\in\mathbb{T}.$$

Now, for $n\in\mathbb{N}$, write the binary expansion of n as

$$n=\sum_{k=0}^{\infty} n_k 2^k,$$

where $n_k\in\{0,1\}$ for every $k=0,1,2,\dots$, and define the n-th *Walsh function* W_n as

$$W_n(x) := \prod_{k=0}^{\infty} W_{2^k}(x)^{n_k}, \quad x\in\mathbb{T}. \tag{39}$$

The functions $\{W_n : n\in\mathbb{N}\}$ are the characters of the Walsh group $(\mathbb{T},\oplus)$, where $\oplus$ stands for addition of binary digits without carry, cf. [**11**] for an introduction from the harmonic analysis viewpoint. For an interval $\omega = [k, m) \subset \mathbb{R}$ with $k, m \in \mathbb{N}$, recall the definition of Walsh frequency projection

$$H_\omega f(x) := \sum_{n=0}^{\infty} \mathbf{1}_\omega(n) \langle f, W_n \rangle W_n(x), \quad x \in \mathbb{T}.$$

For a collection $\Omega = \{\omega\}_{\omega\in\Omega}$ of pairwise disjoint intervals with endpoints in $\mathbb{N}$, recall that the Walsh–Rubio de Francia square function is defined as

$$T^\Omega f(x) := \left(\sum_{\omega\in\Omega} |H_\omega f(x)|^2 \right)^{\frac{1}{2}}, \quad x \in \mathbb{T}.$$

Below we describe the time-frequency model for this square function. We say that $s = I_s \times \omega_s \subseteq \mathbb{T}\times[0,\infty)$ is a tile if I_s and ω_s are dyadic intervals satisfying $|I_s|\cdot|\omega_s| = 1$. Then, for every tile s we can find an integer $n = n(s) \in \mathbb{N}$ so that

$$s = I_s \times \omega_s = I_s \times \frac{1}{|I_s|}[n, n+1).$$

Letting $\mathbb{S}$ denote the universe of all tiles thus defined, the notations (15), (16) will be used in exactly the same way below. Given a tile $s \in \mathbb{S}$, we define the L^2-normalized wave packet associated to s by

$$\varphi_s(x) := \frac{1}{|I_s|^{\frac{1}{2}}} W_{n(s)} \left(\frac{x}{\ell_{I_s}} \right) \mathbf{1}_{I_s}(x), \quad x \in \mathbb{T}.$$

Observe that the Haar functions arise as a special case of these wave packets by taking $s = I \times \frac{1}{|I|}[1, 2)$ with I dyadic subinterval of $\mathbb{T}$, namely

$$h_I(x) := \frac{1}{|I|^{\frac{1}{2}}} W_1 \left(\frac{x}{\ell_I} \right) \mathbf{1}_I(x), \quad x \in \mathbb{T}.$$

Then, the wave packet square function for the Walsh model is defined by

$$W^\Omega f(x) := \left(\sum_{s\in\mathbb{S}^\Omega} |\langle f, \varphi_s \rangle|^2 \frac{\mathbf{1}_{I_s}(x)}{|I_s|} \right)^{\frac{1}{2}}, \quad x \in \mathbb{T}.$$

5.2. The Walsh wave packet square function. Describing the relation between T^Ω and its wave packet model requires some preliminaries. For $\omega \in \Omega$, denote by $\boldsymbol{\omega}$ the collection of maximal dyadic intervals in ω. Imagining the frequency intervals as living on the vertical real axis, denote by $\boldsymbol{\omega}^u$ the collection of dyadic intervals $\omega \in \boldsymbol{\omega}$ which are the upper half of their dyadic parent, and by $\boldsymbol{\omega}^d := \boldsymbol{\omega} \setminus \boldsymbol{\omega}^u$ those which are the lower half of their parent. Then set

$$\Omega^\star := \bigcup_{\sigma\in\{u,d\}} \bigcup_{\omega\in\Omega} \boldsymbol{\omega}^\sigma.$$

The following lemma is a consequence of the Chang–Wilson–Wolff inequality, [**5, 40**].

Lemma 5.1. *Let $w \in A_{\infty,\mathcal{D}}$. Then, the following inequality holds:*

$$\|T^{\Omega} f\|_{L^2(w)} \lesssim [w]_{A_{\infty,\mathcal{D}}}^{\frac{1}{2}} \|W^{\Omega^\star} f\|_{L^2(w)}.$$

Proof: Let $\omega = [k, m)$ be an interval with endpoints in $\mathbb{N}$, and write

$$H_\omega f = \sum_{n=0}^{m-1} \langle f, W_n \rangle W_n - \sum_{n=0}^{k-1} \langle f, W_n \rangle W_n = \sum_{\sigma \in \{u,d\}} \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^\sigma}} \langle f, \varphi_s \rangle \varphi_s;$$

for the last identity see [**38**, Section 8.1] (also [**21**, p. 995]). By symmetry, only consider the case $\sigma = d$ and study the operator

$$T_d^{\Omega} f := \left(\sum_{\omega \in \Omega} \left| \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} \langle f, \varphi_s \rangle \varphi_s \right|^2 \right)^{\frac{1}{2}}.$$

Note that for each fixed $\omega \in \Omega$, the frequency components of the tiles in $\mathbb{S}^{\boldsymbol{\omega}^d}$ form a Whitney decomposition of ω with respect to the right endpoint of ω. Using this fact, which in time-frequency terminology says that $\mathbb{S}^{\boldsymbol{\omega}^d}$ is a tree, together with [**20**, Lemma 2.2] yields the identity

$$\left| \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} \langle f, \varphi_s \rangle \varphi_s \right| = \left| \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} \langle f, W_{n(s)} h_{I_s} \rangle h_{I_s} \right|, \tag{40}$$

where $n(\omega) \in \mathbb{N}$ depends on $\omega \in \Omega$. Thus

$$\begin{aligned} \left\| \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} \langle f, \varphi_s \rangle \varphi_s \right\|_{L^2(w)}^2 &= \left\| \sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} \langle f W_{n(s)}, h_{I_s} \rangle h_{I_s} \right\|_{L^2(w)}^2 \\ &\lesssim [w]_{A_{\infty,\mathcal{D}}} \left\| \left(\sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} |\langle f W_{n(s)}, h_{I_s} \rangle|^2 \frac{\mathbf{1}_{I_s}}{|I_s|} \right)^{\frac{1}{2}} \right\|_{L^2(w)}^2 \\ &= [w]_{A_{\infty,\mathcal{D}}} \left\| \left(\sum_{s \in \mathbb{S}^{\boldsymbol{\omega}^d}} |\langle f, \varphi_s \rangle|^2 \frac{\mathbf{1}_{I_s}}{|I_s|} \right)^{\frac{1}{2}} \right\|_{L^2(w)}^2, \end{aligned}$$

where we have used the Chang–Wilson–Wolff inequality in the form of [**40**, Theorem 3.4] to pass to the second line, and identity (40) again for the last equality. Note that the right hand side of the identity above is easily seen to be bounded by $[w]_{A_{\infty,\mathcal{D}}} \|W^{\Omega^\star} f\|_{L^2(w)}^2$ and the proof is complete. $\square$

Because of Lemma 5.1, Theorem D is reduced to the following proposition. Here we can drop the restriction $I_s \subset \mathbb{T}$ in the tiles s in the definition of W^{Ω} and just work with tiles with $I_s \subset \mathbb{R}$, which we implicitly assume below.

Proposition 5.2. *Let Ω be a collection of pairwise disjoint intervals with endpoints in $\mathbb{N}$. Then*

$$\|W^{\Omega} f\|_{L^p(w)} \lesssim [w]_{A_{\frac{p}{2},\mathcal{D}}}^{1/2} \|f\|_{L^p(w)}, \quad 2 \le p < \infty,$$

with implicit constants depending only on p. These bounds are sharp in terms of the power of the appearing weight characteristic.

Proof: We begin with the case $p=2$ and we prove that for any non-negative locally integrable function w, the following stronger estimate holds:

$$\int |W^\Omega f|^2 w \lesssim \int |f|^2 \mathrm{M}_{\mathcal{D}} w.$$

This clearly implies the conclusion for $p=2$. Here we recall that $\mathrm{M}_{\mathcal{D}}$ stands for the dyadic maximal operator. To that end, we make the qualitative assumption $w \in L^1$, which will be removed momentarily, and use a layer-cake decomposition to prove an $L^2(w)$-bound for the operator W^Ω. More precisely, for $w \in A_1$ and $f \in L^2(\mathbb{T})$, say, write

$$\begin{aligned}\|W^\Omega f\|^2_{L^2(w)} &= \int_{\mathbb{T}} \sum_{s\in\mathbb{S}^\Omega} |\langle f,\varphi_s\rangle|^2 \frac{\mathbf{1}_{I_s}(x)}{|I_s|} w(x)\,\mathrm{d}x \\ &= \sum_{s\in\mathbb{S}^\Omega} |\langle f,\varphi_s\rangle|^2 \int_0^\infty \mathbf{1}_{\{\frac{w(I_s)}{|I_s|}>\lambda\}}\,\mathrm{d}\lambda = \int_0^\infty \sum_{\substack{s\in\mathbb{S}^\Omega\\ \frac{w(I_s)}{|I_s|}>\lambda}} |\langle f,\varphi_s\rangle|^2\,\mathrm{d}\lambda.\end{aligned}$$

Now, let

$$\mathcal{R}_\lambda \coloneqq \left\{ I_s : s\in\mathbb{S}^\Omega,\ \frac{w(I_s)}{|I_s|} > \lambda \right\},$$

and denote by $\mathcal{R}^\star_\lambda$ the collection of maximal elements of $\mathcal{R}_\lambda$. Then,

$$\|W^\Omega f\|^2_{L^2(w)} = \int_0^\infty \sum_{I^\star\in\mathcal{R}^\star_\lambda} \sum_{s\in\mathbb{S}^\Omega_{\subseteq}(I^\star)} |\langle f,\varphi_s\rangle|^2\,\mathrm{d}\lambda. \tag{41}$$

The rightmost term in the display above can be estimated by using the local orthogonality of the Walsh wave packets in the form

$$\sum_{s\in\mathbb{S}^\Omega_{\subseteq}(I)} |\langle f,\varphi_s\rangle|^2 \le \int_I |f(x)|^2\,\mathrm{d}x \quad \forall I\in\mathcal{D}.$$

Applying this in (41) we get

$$\|W^\Omega f\|^2_{L^2(w)} \le \int_0^\infty \sum_{I^\star\in\mathcal{R}^\star_\lambda} \int_{I^\star} |f(x)|^2\,\mathrm{d}x\,\mathrm{d}\lambda = \int_0^\infty \int_{\bigcup_{I^\star\in\mathcal{R}^\star_\lambda} I^\star} |f(x)|^2\,\mathrm{d}x\,\mathrm{d}\lambda.$$

Recall that $\mathrm{M}_{\mathcal{D}}$ stands for the dyadic maximal function, and observe that

$$\bigcup_{I^\star\in\mathcal{R}^\star_\lambda} I^\star \subset \{x : \mathrm{M}_{\mathcal{D}} w(x) > \lambda\}.$$

Thus,

$$\|W^\Omega f\|^2_{L^2(w)} \le \int_0^\infty \int_{\mathrm{M}_{\mathcal{D}} w(x)>\lambda} |f(x)|^2\,\mathrm{d}x\,\mathrm{d}\lambda = \int_{\mathbb{T}} |f(x)|^2 \mathrm{M}_{\mathcal{D}} w(x)\,\mathrm{d}x,$$

which is the estimate we want to prove for $p=2$. We can now drop the assumption $w\in L^1$, for example by a monotone convergence argument, yielding the same inequality for arbitrary locally integrable non-negative functions w. Finally the estimate above and the definition of dyadic A_1-weights readily yields

$$\|W^\Omega f\|^2_{L^2(w)} \le [w]_{A_{1,\mathcal{D}}} \|f\|^2_{L^2(w)},$$

which is the conclusion of the proposition for $p=2$.

For $p > 2$, the $L^p(w)$-estimates of the proposition follow easily by extrapolation, using for example [**16**, Corollary 4.2]. Finally note that these bounds are sharp, in terms of the exponents of the $A_{p,\mathcal{D}}$-weight. Indeed, as in [**33**], any of the bounds in the conclusion of the theorem implies that the unweighted L^p-norms of the martingale square function grow like $p^{1/2}$, which is best possible: any better exponent on $[w]_{A_p,\mathcal{D}}$ would imply a stronger, and false, p growth of the martingale square function as $p \to \infty$. □

5.3. Optimality in Theorem D. The only thing remaining to show in order to complete the proof of Theorem D is the optimality of the exponent 1 appearing in the exponent of the weight constant in the estimate

$$\|T^\Omega f\|_{L^p(w)} \lesssim [w]_{A_{p/2,\mathcal{D}}} \|f\|_{L^p(w)}, \quad 2 \le p < \infty.$$

For this we can just choose Ω consisting of a single interval of the form $[0, b+1)$ with $b \in \mathbb{N}$. We claim that an estimate of the form

$$\sup_{b\in\mathbb{N}} \Big\| \sum_{n\in[0,b]} \langle f, W_n\rangle W_n \Big\|_{L^p(w)} \le C_p \|f\|_{L^p(w)}$$

implies that martingale transforms are bounded on L^p with constant at most C_p. This observation together with the considerations in [**33**] will imply again the claimed sharpness. In order to verify the observation above, we quote from [**13**, equation (5.7)] the equality

$$\left|\sum_{n=0}^{b} \langle f, W_n\rangle W_n\right| = \left|\sum_{I\in\mathcal{D}} \varepsilon_{I,b} \langle f W_b, h_I\rangle h_I\right|,$$

where $\varepsilon_{I,b} \in \{0,1\}$ is a sequence depending on b only. The right hand side is a Haar martingale transform of fW_b. This shows that we can recover any Haar martingale transform of f by suitable choice of b, which is the promised claim and completes the proof of the optimality of the exponents in Theorem D.

6. Proof of Theorem E

The proof of Theorem E begins with relating T^Ω to the intrinsic wave packet square function W^Ω defined in (21) via a version of the Chang–Wilson–Wolff inequality.

Lemma 6.1. *For all $w \in A_\infty$ there holds*

$$\|T^\Omega\|_{L^2(w)} \lesssim [w]_{A_\infty}^{\frac12} \|G^\Omega\|_{L^2(w)} \lesssim [w]_{A_\infty}^{\frac12} \|W^\Omega\|_{L^2(w)}.$$

Proof: The second inequality in the conclusion of the lemma is an application of (22). The first inequality follows from an application of the Chang–Wilson–Wolff inequality [**5**], for example in the form elaborated by Lerner in [**29**, Theorem 2.7]. □

The next step is to establish a sufficient condition for $L^2(w)$-boundedness of W^Ω based on the super-level sets of $\mathrm{M}w$; see (42) below. We will later show that (42) is satisfied by radially decreasing, even A_1-weights. Turning to the former task, let $w \in A_1$ and $\lambda > 0$, and let $\mathcal{R}_\lambda$ denote a collection of dyadic intervals such that $\frac{w(R)}{|R|} > \lambda$. Note that

$$\bigcup_{R\in\mathcal{R}_\lambda} R \subseteq \{\mathrm{M}_{\mathcal{D}} w > \lambda\},$$

where $\mathrm{M}_{\mathcal{D}}$ denotes the dyadic maximal function. Arguing as in Section 5 we can reduce the $L^2(w)$-boundedness of the intrinsic wave packet square function W^Ω defined in (21) to the estimate

$$\int_0^\infty \sum_{\substack{t\in\mathbb{S}^\Omega\\ R_t\in\mathcal{R}_\lambda}} |R_t|\langle f,\varphi_t\rangle^2 \lesssim_w \int |f|^2 w, \tag{42}$$

for an arbitrary choice of $\varphi_t\in\Psi_t$ for each $t\in\mathbb{S}^\Omega$. This is because

$$\begin{aligned}\|W^\Omega f\|^2_{L^2(w)} &\lesssim \int_{\mathbb{R}} \sum_{t\in\mathbb{S}^\Omega} |R_t||\langle f,\varphi_t\rangle|^2 \frac{\mathbf{1}_{R_t}(x)}{|R_t|} w(x)\,\mathrm{d}x\\ &= \sum_{t\in\mathbb{S}^\Omega} |R_t||\langle f,\varphi_t\rangle|^2 \int_0^\infty \mathbf{1}_{\{\frac{w(R_t)}{|R_t|}>\lambda\}}\,\mathrm{d}\lambda = \int_0^\infty \sum_{\substack{t\in\mathbb{S}^\Omega\\ R_t\in\mathcal{R}_\lambda}} |R_t||\langle f,\varphi_t\rangle|^2\,\mathrm{d}\lambda.\end{aligned}$$

Thus, our goal is to prove (42). We will use the following definition.

Definition 6.2. Let $\mathcal{R}$, $\mathcal{R}^*$ be collections of intervals. We say that $\mathcal{R}$ is *subordinate* to $\mathcal{R}^*$ if for every $R\in\mathcal{R}$ there exists $R^*\in\mathcal{R}^*$ such that $R\subseteq R^*$.

The canonical example of a collection $\mathcal{R}^*$ to which $\mathcal{R}$ is subordinate is the collection of its maximal elements. However, other choices are possible. Now we assume that for each $\lambda>0$ the collection $\mathcal{R}_\lambda$ is subordinate to $\mathcal{R}^*_\lambda$. Then, in view of Lemma 3.1 we have the chain of inequalities

$$\int_0^\infty \sum_{R_t\in\mathcal{R}_\lambda} |R_t||\langle f,\varphi_t\rangle|^2 = \int_0^\infty \sum_{R^*\in\mathcal{R}^*_\lambda} \sum_{\substack{R_t\in\mathcal{R}_\lambda\\ R_t\subseteq R^*}} |R_t||\langle f,\varphi_t\rangle|^2 \lesssim \int_{\mathbb{R}} |f|^2 \int_0^\infty \sum_{R^*\in\mathcal{R}^*_\lambda} \chi^9_{R^*}.$$

Thus, a sufficient condition for the desired $L^2(w)$-boundedness (42) is that for a.e. $x\in\mathbb{R}$ there holds

$$\int_0^\infty \sum_{R^*\in\mathcal{R}^*_\lambda} \chi^9_{R^*}(x)\,\mathrm{d}\lambda \lesssim w(x), \tag{43}$$

where $\mathcal{R}^*_\lambda$ is such that $\mathcal{R}_\lambda$ is subordinate to $\mathcal{R}^*_\lambda$ for every $\lambda>0$. By considering a single interval R and taking $\lambda<w(R)/|R|$ we readily see that (43) implies the A_1 condition.

6.1. $L^2(w)$-boundedness for even and radially decreasing A_1-weights.

We can show the sufficient condition (43) for even and radially decreasing weights $w\in A_1$, i.e. $w(x)=w_0(|x|)$ for some $w_0\colon[0,\infty)\to[0,\infty)$ and w_0 decreasing. The proof proceeds by verifying the sufficient condition (43). In doing so we also provide the promised generalization of Theorem E of the previously known results for $w(x):=|x|^{-\alpha}\in A_1$ to even radially decreasing A_1-weights on the real line.

Let $R=[a,b]$ be an interval belonging to $\mathcal{R}_\lambda$, which we recall is the collection of intervals such that $\frac{w(R)}{|R|}>\lambda$. Without loss of generality, assume that $|a|<|b|$ so that $R\subseteq[-|b|,|b|]$. Since the weight w is even and decreasing we have that

$$\lambda<\frac{w(R)}{|R|}\le[w]_{A_1}\inf_{x\in R} w(x)=[w]_{A_1}w_0(|b|)\Longleftrightarrow w_0(|b|)>\frac{\lambda}{[w]_{A_1}}.$$

Since w is decreasing the last inequality implies the existence of some $b_\lambda=b_\lambda([w]_{A_1},w)>0$ with $w_0(b_\lambda)>\lambda/[w]_{A_1}$ such that $|b|\le b_\lambda$. That is, denoting $R_\lambda:=[-b_\lambda,b_\lambda]$ we have that $R \subseteq R_\lambda$ and all intervals $R \in \mathcal{R}_\lambda$ are subordinate to the collection $\{R_\lambda\}$ for every $\lambda > 0$. However,

$$\int_0^\infty \chi_{R_\lambda}^9 \,\mathrm{d}\lambda \lesssim \sum_{\tau\geq 0} 2^{-18\tau} \int_0^\infty \mathbf{1}_{\{|x|\leq 2^\tau b_\lambda\}}(x)\,\mathrm{d}\lambda,$$

where we have used the decay of $\chi_{R_\lambda}^9$. Observe that

$$|x| \leq 2^\tau b_\lambda \iff w_0\left(\frac{|x|}{2^\tau}\right) \geq w_0(b_\lambda) > \frac{\lambda}{[w]_{A_1}}.$$

Thus,

$$\int_0^\infty \chi_{R_\lambda}^9 \,\mathrm{d}\lambda \lesssim \sum_{\tau\geq 0} 2^{-18\tau} \int_0^{[w]_{A_1} w_0(\frac{|x|}{2^\tau})} \mathrm{d}\lambda = \sum_{\tau\geq 0} 2^{-18\tau}[w]_{A_1} w_0\left(\frac{|x|}{2^\tau}\right).$$

Finally, note that $\frac{|x|}{2^\tau} \leq |x|$, so that

$$\begin{aligned} w_0\left(\frac{|x|}{2^\tau}\right) &= \inf_{(0,\frac{|x|}{2^\tau})} w \leq \frac{w([0,|x|/2^\tau])}{|x|/2^\tau} \leq \frac{w([0,|x|])}{|x|}\cdot 2^\tau \\ &\leq [w]_{A_1} 2^\tau \inf_{(0,|x|)} w = 2^\tau [w]_{A_1} w_0(|x|). \end{aligned}$$

Using this in the previous inequality yields

$$\int_0^\infty \chi_{R_\lambda}^9 \,\mathrm{d}\lambda \lesssim [w]_{A_1} \sum_{\tau\geq 0} 2^{-18\tau} 2^\tau [w]_{A_1} w_0(|x|) \lesssim [w]_{A_1}^2 w(x).$$

This shows that even and radially decreasing A_1-weights satisfy the sufficient condition (43) and thus, combined with Lemma 6.1, Theorem E is proved.

References

[1] N. Accomazzo, F. Di Plinio, P. Hagelstein, I. Parissis, and. L. Roncal, Directional square functions, *Anal. PDE* **16(7)** (2023), 1651–1699. DOI: 10.2140/apde.2023.16.1651.

[2] C. Benea, Vector-valued extensions for singular bilinear operators and applications, Thesis (Ph.D.)-Cornell University (2015). Available on https://ecommons.cornell.edu.

[3] F. Bernicot, D. Frey, and S. Petermichl, Sharp weighted norm estimates beyond Calderón–Zygmund theory, *Anal. PDE* **9(5)** (2016), 1079–1113. DOI: 10.2140/apde.2016.9.1079.

[4] G. Brocchi, A sparse quadratic $T(1)$ theorem, *New York J. Math.* **26** (2020), 1232–1272.

[5] S. Y. A. Chang, J. M. Wilson, and T. H. Wolff, Some weighted norm inequalities concerning the Schrödinger operators, *Comment. Math. Helv.* **60(2)** (1985), 217–246. DOI: 10.1007/BF02567411.

[6] J. M. Conde-Alonso, A. Culiuc, F. Di Plinio, and Y. Ou, A sparse domination principle for rough singular integrals, *Anal. PDE* **10(5)** (2017), 1255–1284. DOI: 10.2140/apde.2017.10.1255.

[7] J. M. Conde-Alonso, F. Di Plinio, I. Parissis, and M. N. Vempati, A metric approach to sparse domination, *Ann. Mat. Pura Appl. (4)* **201(4)** (2022), 1639–1675. DOI: 10.1007/s10231-021-01174-7.

[8] J. M. Conde-Alonso and G. Rey, A pointwise estimate for positive dyadic shifts and some applications, *Math. Ann.* **365(3-4)** (2016), 1111–1135. DOI: 10.1007/s00208-015-1320-y.

[9] D. Cruz-Uribe, J. M. Martell, and C. Pérez, Sharp weighted estimates for classical operators, *Adv. Math.* **229(1)** (2012), 408–441. DOI: 10.1016/j.aim.2011.08.013.

[10] A. Culiuc, F. Di Plinio, and Y. Ou, Domination of multilinear singular integrals by positive sparse forms, *J. Lond. Math. Soc. (2)* **98(2)** (2018), 369–392. DOI: 10.1112/jlms.12139.

[11] C. Demeter, A guide to Carleson's theorem, *Rocky Mountain J. Math.* **45(1)** (2015), 169–212. DOI: 10.1216/RMJ-2015-45-1-169.

[12] F. Di Plinio and A. Fragkos, The weak-type Carleson theorem via wave packet estimates, *Trans. Amer. Math. Soc.* **378(3)** (2025), 1551–1592. DOI: 10.1090/tran/9303.

[13] F. Di Plinio and A. K. Lerner, On weighted norm inequalities for the Carleson and Walsh–Carleson operator, *J. Lond. Math. Soc. (2)* **90(3)** (2014), 654–674. DOI: 10.1112/jlms/jdu049.

[14] C. Domingo-Salazar, M. Lacey, and G. Rey, Borderline weak-type estimates for singular integrals and square functions, *Bull. Lond. Math. Soc.* **48(1)** (2016), 63–73. DOI: 10.1112/blms/bdv090.

[15] J. Duoandikoetxea, *Fourier Analysis*, Translated and revised from the 1995 Spanish original by David Cruz-Uribe, Grad. Stud. Math. **29**, American Mathematical Society, Providence, RI, 2001. DOI: 10.1090/gsm/029.

[16] J. Duoandikoetxea, Extrapolation of weights revisited: New proofs and sharp bounds, *J. Funct. Anal.* **260(6)** (2011), 1886–1901. DOI: 10.1016/j.jfa.2010.12.015.

[17] D. Frey and Z. Nieraeth, Weak and strong type A_1-A_∞ estimates for sparsely dominated operators, *J. Geom. Anal.* **29(1)** (2019), 247–282. DOI: 10.1007/s12220-018-9989-2.

[18] R. Garg, L. Roncal, and S. Shrivastava, Quantitative weighted estimates for Rubio de Francia's Littlewood–Paley square function, *J. Geom. Anal.* **31(1)** (2021), 748–771. DOI: 10.1007/s12220-019-00297-x.

[19] T. P. Hytönen, The sharp weighted bound for general Calderón–Zygmund operators, *Ann. of Math. (2)* **175(3)** (2012), 1473–1506. DOI: 10.4007/annals.2012.175.3.9.

[20] T. P. Hytönen and M. T. Lacey, Pointwise convergence of Walsh–Fourier series of vector-valued functions, *Math. Res. Lett.* **25(2)** (2018), 561–580. DOI: 10.4310/MRL.2018.v25.n2.a11.

[21] T. P. Hytönen, M. T. Lacey, and I. Parissis, A variation norm Carleson theorem for vector-valued Walsh–Fourier series, *Rev. Mat. Iberoam.* **30(3)** (2014), 979–1014. DOI: 10.4171/RMI/804.

[22] T. P. Hytönen and K. Li, Weak and strong A_p-A_∞ estimates for square functions and related operators, *Proc. Amer. Math. Soc.* **146(6)** (2018), 2497–2507. DOI: 10.1090/proc/13908.

[23] T. Hytönen and C. Pérez, Sharp weighted bounds involving A_∞, *Anal. PDE* **6(4)** (2013), 777–818. DOI: 10.2140/apde.2013.6.777.

[24] J.-L. Journé, Calderón–Zygmund operators on product spaces, *Rev. Mat. Iberoamericana* **1(3)** (1985), 55–91. DOI: 10.4171/RMI/15.

[25] S. Król, Fourier multipliers on weighted L^p spaces, *Math. Res. Lett.* **21(4)** (2014), 807–830. DOI: 10.4310/MRL.2014.v21.n4.a11.

[26] M. T. Lacey, *Issues related to Rubio de Francia's Littlewood–Paley Inequality*, NYJM Monogr. **2**, State University of New York, University at Albany, Albany, NY, 2007.

[27] M. Lacey and C. Thiele, A proof of boundedness of the Carleson operator, *Math. Res. Lett.* **7(4)** (2000), 361–370. DOI: 10.4310/MRL.2000.v7.n4.a1.

[28] A. K. Lerner, A simple proof of the A_2 conjecture, *Int. Math. Res. Not. IMRN* **2013(14)** (2013), 3159–3170. DOI: 10.1093/imrn/rns145.

[29] A. K. Lerner, Quantitative weighted estimates for the Littlewood–Paley square function and Marcinkiewicz multipliers, *Math. Res. Lett.* **26(2)** (2019), 537–556. DOI: 10.4310/MRL.2019.v26.n2.a7.

[30] A. K. Lerner and F. Nazarov, Intuitive dyadic calculus: The basics, *Expo. Math.* **37(3)** (2019), 225–265. DOI: 10.1016/j.exmath.2018.01.001.

[31] A. K. Lerner, S. Ombrosi, and C. Pérez, A_1 bounds for Calderón–Zygmund operators related to a problem of Muckenhoupt and Wheeden, *Math. Res. Lett.* **16(1)** (2009), 149–156. DOI: 10.4310/MRL.2009.v16.n1.a14.

[32] K. Li and W. Sun, Weak and strong type weighted estimates for multilinear Calderón–Zygmund operators, *Adv. Math.* **254** (2014), 736–771. DOI: 10.1016/j.aim.2013.12.027.

[33] T. Luque, C. Pérez, and E. Rela, Optimal exponents in weighted estimates without examples, *Math. Res. Lett.* **22(1)** (2015), 183–201. DOI: 10.4310/MRL.2015.v22.n1.a10.

[34] K. Moen, Sharp weighted bounds without testing or extrapolation, *Arch. Math. (Basel)* **99(5)** (2012), 457–466. DOI: 10.1007/s00013-012-0453-4.

[35] J. L. Rubio de Francia, Estimates for some square functions of Littlewood–Paley type, *Publ. Sec. Mat. Univ. Autònoma Barcelona* **27(2)** (1983), 81–108. DOI: 10.5565/PUBLMAT_27283_08.

[36] J. L. Rubio de Francia, A Littlewood–Paley inequality for arbitrary intervals, *Rev. Mat. Iberoamericana* **1(2)** (1985), 1–14. DOI: 10.4171/RMI/7.

[37] J. L. Rubio de Francia, Transference principles for radial multipliers, *Duke Math. J.* **58(1)** (1989), 1–19. DOI: 10.1215/S0012-7094-89-05801-8.

[38] C. Thiele, *Wave Packet Analysis*, CBMS Reg. Conf. Ser. Math. **105**, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. DOI: 10.1090/cbms/105.

[39] J. L. Torrea, J. García-Cuerva, J. Duoandikoetxea, and A. Carbery, The work of José Luis Rubio de Francia. I, II, III, IV, Special Conference in Honor of the Memory of José Luis Rubio de Francia (El Escorial, 1989), *Publ. Mat.* **35(1)** (1991), 9–25, 27–63, 65–80, 81–93. DOI: 10.5565/PUBLMAT_35191_01.

[40] M. Wilson, *Weighted Littlewood–Paley Theory and Exponential-Square Integrability*, Lecture Notes in Math. **1924**, Springer, Berlin, 2008. DOI: 10.1007/978-3-540-74587-7.

Francesco Di Plinio
Dipartimento di Matematica e Applicazioni, Università di Napoli, Via Cintia, Monte S. Angelo 80126 Napoli, Italy
E-mail address: francesco.diplinio@unina.it

Mikel Flórez-Amatriain
BCAM - Basque Center for Applied Mathematics, 48009 Bilbao, Spain
E-mail address: mflorez@bcamath.org

Ioannis Parissis
Departamento de Matemáticas, Universidad del País Vasco, Aptdo. 644, 48080 Bilbao, Spain and Ikerbasque, Basque Foundation for Science, Bilbao, Spain
E-mail address: ioannis.parissis@ehu.eus

Luz Roncal
BCAM - Basque Center for Applied Mathematics, 48009 Bilbao, Spain
Ikerbasque, Basque Foundation for Science, Bilbao, Spain, and
Universidad del País Vasco, Bilbao, Spain
E-mail address: lroncal@bcamath.org

Received on December 19, 2023.
Accepted on September 9, 2024.