BANACH-VALUED MULTILINEAR SINGULAR INTEGRALS WITH MODULATION INVARIANCE

FRANCESCO DI PLINIO, KANGWEI LI, HENRI MARTIKAINEN, AND EMIL VUORINEN

Abstract. We prove that the class of trilinear multiplier forms with singularity over a one dimensional subspace, including the bilinear Hilbert transform, admit bounded L^p-extension to triples of intermediate UMD spaces. No other assumption, for instance of Rademacher maximal function type, is made on the triple of UMD spaces. Among the novelties in our analysis is an extension of the phase-space projection technique to the UMD-valued setting. This is then employed to obtain appropriate single tree estimates by appealing to the UMD-valued bound for bilinear Calderón-Zygmund operators recently obtained by the same authors.

1. Introduction and main results

Let $\mathcal{X}_k, k = 1, 2, 3$, be Banach spaces with a trilinear contraction $\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \to \mathbb{C}$ which we denote by $(e_1, e_2, e_3) \mapsto e_1e_2e_3 = \prod_{k=1}^{3} e_k$. To a multiplier m defined on the orthogonal complement Γ of $(1, 1, 1) \in \mathbb{R}^3$, we may associate the trilinear form

$$\Lambda_m(f_1, f_2, f_3) = \int_\Gamma m(\xi) \left(\prod_{k=1}^{3} \widehat{f_k}(\xi_k) \right) \mathrm{d}\xi \tag{1.1}$$

acting on functions $f_k \in \mathcal{S}(\mathbb{R}) \otimes \mathcal{X}_k, k = 1, 2, 3$, where the former is the Schwartz class. This article is concerned with multipliers m whose singularity lies on a one-dimensional subspace perpendicular to a unit vector $\beta \in \Gamma$ which is nondegenerate in the sense that

$$\Delta_\beta := \min_{j \neq k} |\beta_j - \beta_k| > 0, \tag{1.2}$$

(F. Di Plinio) Department of Mathematics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130-4899, USA

(K. Li) Center for Applied Mathematics, Tianjin University, Weijin Road 92, 300072 Tianjin, China

(H. Martikainen) Department of Mathematics and Statistics, University of Helsinki, P.O.B. 68, FI-00014 University of Helsinki, Finland

(E. Vuorinen) Centre for Mathematical Sciences, University of Lund, P.O.B. 118, 22100 Lund, Sweden

E-mail addresses: francesco.diplinio@wustl.edu, kangwei.nku@gmail.com, henri.martikainen@helsinki.fi, j.e.vuorin@gmail.com.

2010 *Mathematics Subject Classification.* 42B20.

Key words and phrases. bilinear Hilbert transform, singular integrals, modulation invariance, UMD spaces, phase-space projections, interpolation spaces.

F. Di Plinio has been partially supported by the National Science Foundation under the grants NSF-DMS-1800628 and NSF-DMS-2000510.

H. Martikainen was supported by the Academy of Finland through the grants 294840, 306901 and 327271 and by the three-year research grant 75160010 of the University of Helsinki. He is a member of the Finnish Centre of Excellence in Analysis and Dynamics Research supported by the Academy of Finland (project No. 307333).

E. Vuorinen was supported by the Jenny and Antti Wihuri Foundation.

and satisfies for all multi-indices α

$$\sup_{\xi\in\Gamma}\Big(\operatorname{dist}(\xi,\beta^\perp)\Big)^{\alpha}\Big|\partial_\alpha m(\xi)\Big| \lesssim_\alpha 1. \tag{1.3}$$

Assumption (1.3) is a $\beta^\perp$-modulation invariant version of the Coifman-Meyer condition. This class includes the bilinear Hilbert transform with parameter β, whose dual trilinear multiplier form may be obtained by choosing

$$m(\xi) = \operatorname{sgn}(\xi\cdot\beta).$$

The (adjoint form to the) bilinear Hilbert transform

$$\mathrm{BHT}_\beta(f_1,f_2,f_3) = \int_{\mathbb{R}} \mathrm{p.v.}\int \prod_{j=1}^{3} f_j(x-\beta_j t)\,\frac{\mathrm{d}t}{t}\mathrm{d}x$$

was first introduced by Calderón within the context of L^p estimates for the first commutator of the Cauchy integral along Lipschitz curves. The celebrated articles of Lacey and Thiele [27, 28] contain the first proof of L^p estimates for the bilinear Hilbert transform, while more general multipliers of the class (1.3) were treated by Muscalu, Tao and Thiele [29].

1.1. **Main results.** In this article, we prove that the trilinear multiplier forms (1.1), where m is a multiplier of the class (1.3), admit L^p-bounded extensions to triples of intermediate UMD Banach spaces. This class of UMD spaces first appears in the survey work of Rubio de Francia [32] with focus on the Banach function space case, and has subsequently been considered by Hytönen and Lacey in the articles [18, 19] in the context of modulation-invariant operators. We repeat ths definition below and send to [24] and references therein for background and generalities on UMD Banach spaces.

Let $2\le q\le\infty$ and $\mathcal{X}_0,\mathcal{X}_1$ be a couple of compatible Banach spaces, with $\mathcal{X}_0$ being a UMD space and $\mathcal{X}_1$ being a Hilbert space. We say that the Banach space $\mathcal{X}$ is *q-intermediate* UMD if

$$\mathcal{X} = [\mathcal{X}_0,\mathcal{X}_1]_{\frac{2}{q}}$$

namely, $\mathcal{X}$ is the complex interpolation of a UMD Banach space with a Hilbert space. Such Banach space $\mathcal{X}$ is automatically a UMD space. Notice that $\mathcal{X}$ is q-intermediate UMD if and only if its Banach dual $\mathcal{X}'$ is.

The precise statement of our main result is as follows.

1.4. **Theorem.** *Let $\mathcal{X}_j$, $j=1,2,3$, be Banach spaces with Banach duals $\mathcal{Y}_j=\mathcal{X}_j'$ and suppose that each $\mathcal{X}_j$ is $q_{\mathcal{X}_j}$-intermediate* UMD. *Assume that*

$$\rho = \sum_{j=1}^{3}\frac{1}{q_{\mathcal{X}_j}} - 1 > 0. \tag{1.5}$$

Let σ be any permutation of $\{1,2,3\}$, m be a multiplier satisfying (1.3) *and $T_{m,\sigma}$ denote the adjoint bilinear operator to* (1.1) *acting on pairs of $\mathcal{X}_{\sigma(1)},\mathcal{X}_{\sigma(2)}$-valued functions. Then*

$$\left\|T_{m,\sigma}(f_{\sigma(1)},f_{\sigma(2)})\right\|_{L^{\frac{p_1p_2}{p_1+p_2}}(\mathbb{R};\mathcal{Y}_{\sigma(3)})} \lesssim \left\|f_{\sigma(1)}\right\|_{L^{p_1}(\mathbb{R};\mathcal{X}_{\sigma(1)})}\left\|f_{\sigma(2)}\right\|_{L^{p_2}(\mathbb{R};\mathcal{X}_{\sigma(2)})}$$

whenever

$$1 < p_1, p_2 \le \infty, \qquad (p_1, p_2) \neq (\infty, \infty), \qquad \left(\frac{1}{p_1}, \frac{1}{p_2}\right) \in \mathrm{int}(\mathcal{H}). \tag{1.6}$$

Here $\mathcal{H}$ is the hexagon with vertices A, B, C, D, E, F as follows:

$$\begin{aligned}
A &: \left(\tfrac{1}{q_{\mathcal{X}_1}} - \rho q_{\mathcal{X}_3}, \tfrac{1}{q_{\mathcal{X}_2}}\right), & D &: \left(\tfrac{1}{q_{\mathcal{X}_1}} + \rho q_{\mathcal{X}_1} - \rho, \tfrac{1}{q_{\mathcal{X}_2}}\right),\\
B &: \left(\tfrac{1}{q_{\mathcal{X}_1}}, \tfrac{1}{q_{\mathcal{X}_2}} - \rho q_{\mathcal{X}_3}\right), & E &: \left(\tfrac{1}{q_{\mathcal{X}_1}}, \tfrac{1}{q_{\mathcal{X}_2}} + \rho q_{\mathcal{X}_2} - \rho\right),\\
C &: \left(\tfrac{1}{q_{\mathcal{X}_1}} + \rho q_{\mathcal{X}_1} - \rho, \tfrac{1}{q_{\mathcal{X}_2}} - \rho q_{\mathcal{X}_1}\right), & F &: \left(\tfrac{1}{q_{\mathcal{X}_1}} - \rho q_{\mathcal{X}_2}, \tfrac{1}{q_{\mathcal{X}_2}} + \rho q_{\mathcal{X}_2} - \rho\right).
\end{aligned}$$

The proof of Theorem 1.4 relies on the main energy and tree lemmata of Section 3 and is outlined in Subsection 3.4. We note in passing that if condition (1.5) holds, the range $\mathrm{int}(\mathcal{H})$ is nonempty and in particular contains the region

$$q_{\mathcal{X}_k} < p_k < \infty, \; k = 1,2,3, \qquad p_3 := \left(\tfrac{p_1 p_2}{p_1 + p_2}\right)'$$

which is the analogue of the local L^2 range for the scalar case, see [27]. In addition, we point out that $\mathrm{int}(\mathcal{H})$ may contain quasi-Banach pairs (p_1, p_2), that is, pairs with $\frac{p_1 p_2}{p_1 + p_2} < 1$. This is easier to see by particularizing Theorem 1.4 to the case

$$\mathcal{X}_1 = \mathcal{X}, \; \mathcal{X}_2 = \mathcal{X}', \; \mathcal{X}_3 = \mathbb{C},$$

as in the following corollary. Herein, quasi-Banach estimates are available if $2 < q < 3$.

1.7. **Corollary.** *Let $\mathcal{X}$ be a q-intermediate* UMD *space and define the trilinear contraction*

$$(\mathfrak{x}, \phi, \lambda) \in \mathcal{X} \times \mathcal{X}' \times \mathbb{C} \mapsto \lambda\phi(\mathfrak{x}).$$

Let m be a multiplier satisfying (1.3) *and T_m denote the adjoint bilinear operator to* (1.1) *acting on pairs of $\mathcal{X}, \mathcal{X}'$-valued functions.*

Suppose that $2 \le q \le 3$. *Then*

$$\left\|T_m(f_1, f_2)\right\|_{L^{\frac{p_1 p_2}{p_1 + p_2}}(\mathbb{R})} \lesssim \left\|f_1\right\|_{L^{p_1}(\mathbb{R};\mathcal{X})} \left\|f_2\right\|_{L^{p_2}(\mathbb{R};\mathcal{X}')} \tag{1.8}$$

whenever

$$1 + \frac{(q-1)(q-2)}{q(5-q)-2} < p_1, p_2 \le \infty, \qquad \frac{2}{3}\left(1 + \frac{q-2}{5-q}\right) < \frac{p_1 p_2}{p_1 + p_2} < \frac{q}{q-2}. \tag{1.9}$$

If $3 < q < 4$, *then* (1.8) *holds true if, in addition to* (1.9), *the condition*

$$\frac{q^2 - 3q + 1}{q} < \min_{(u,v) \in \{(1,2),(2,1)\}} \left\{\frac{q-1}{p_u} + \frac{q-2}{p_v}\right\}$$

is verified.

Theorem 1.4 and Corollary 1.7 further the rather recent line of research on the extension of singular operators with modulation invariance properties to UMD Banach spaces *without any UMD Banach function space structure,* or lattice structure altogether: a prototypical example are noncommutative L^p spaces such as the reflexive Schatten-von Neumann subclasses of the algebra of bounded operators on a Hilbert space. This line of research was initiated by Hytönen and Lacey in their proof of boundedness of the Carleson maximal partial Fourier sum operator for intermediate UMD spaces in the Walsh [19] and Fourier

[18] setting; see also [21] for Walsh-Carleson variation norm bounds. Subsequently, the same authors and Parissis [20] proved the analogue of Theorem 1.4 for the Walsh model of the bilinear Hilbert transform. In fact, the range of exponents int($\mathcal{H}$) is the same as the one obtained therein for the Walsh model, see [20, Theorem 9.3]. Results in the vein of [20] were recently reproved by Amenta and Uraltsev in [3] as a byproduct of novel Banach-valued outer L^p space embeddings for the Walsh wave packet transform.

The theory of UMD-valued linear singular integrals of Calderón-Zygmund type is rooted in the works by Burkholder [6] and Bourgain [5] among others, and has been extensively developed since then, see for instance [7, 15–17, 22, 25, 26, 36] and the monograph [24]. Recent advances have concerned the UMD extension of multilinear Caldéron-Zygmund operators [11–13]. The above mentioned references deal with generic UMD spaces, as opposed to lattices, and thus develop fundamentally different techniques from those of the classical vector-valued theory of e.g. Benedek, Calderón and Panzone, Fefferman-Stein, Rubio de Francia, which are strictly tied to A_p-type weighted norm inequalities. In a similar contrast, the present article combines novel technical tools in UMD-valued time frequency analysis to the UMD interpolation space idea of [18] in order to deal with *multilinear* modulation invariant operators on *non-lattice* UMD spaces, which are out of reach for typical lattice-based techniques.

Nevertheless, a systematic function space-valued theory for (1.1) is quite recent. The first proof of ℓ^p-valued bounds for the bilinear Hilbert transform in a wide range of exponents is due to P. Silva [33]. In [33], those estimates have been employed to obtain bounds for the biparameter bilinear operator obtained by tensoring the bilinear Hilbert transform with a Coifman-Meyer multiplier. Several extensions and refinements of [33] have since appeared, see e.g. [1, 4, 8, 9]. In general, as Corollary 1.7 demonstrates, Theorem 1.4 is outside the scope of the above references, although it does imply a strict subset of the ℓ^p estimates of [33]. We send to [3, 20] for a detailed discussion of this point.

However, to stress the difference with the results of [33] and followups, we would like to showcase here a further application of Theorem 1.4 to a triple of non-function, non-lattice UMD Banach spaces, in addition to that of Corollary 1.7. In the corollary that follows we denote by S^p, $1 \le p < \infty$ the p-th Schatten-von Neumann class, namely, the subspace of the von Neumann algebra $\mathcal{B}(H)$ of linear bounded operators on a separable Hilbert space defined by the norm

$$\|A\|_{S^p} = \|s_n(A)\|_{\ell^p(n\in\mathbb{N})}$$

where $\{s_n(A) : n \in \mathbb{N}\}$ is the sequence of singular values of A, that is eigenvalues of the Hermitian operator $|A| = \sqrt{A^*A}$. Notice that the classes S^p are increasingly nested with p and that the trilinear form

$$(A_1, A_2, A_3) \in S^{t_1} \times S^{t_2} \times S^{t_3} \mapsto \mathrm{trace}(A_1A_2A_3) \tag{1.10}$$

is a contraction provided that

$$\sum_{k=1}^{3} \frac{1}{t_k} \ge 1. \tag{1.11}$$

1.12. **Corollary.** *Suppose that the exponents* $1 < t_1, t_2, t_3 < \infty$ *satisfy*

$$\rho = \sum_{k=1}^{3} \frac{1}{\max\{t_k, (t_k)'\}} - 1 > 0.$$

Let σ *be a permutation of* $\{1, 2, 3\}$*, and* m *be a multiplier satisfying* (1.3)*. Then the corresponding adjoint bilinear operator* $T_{m,\sigma}$ *maps*

$$T_{m,\sigma} : L^{p_1}(\mathbb{R}; S^{t_{\sigma(1)}}) \times L^{p_2}(\mathbb{R}; S^{t_{\sigma(2)}}) \to L^{\frac{p_1 p_2}{p_1+p_2}}(\mathbb{R}; S^{t'_{\sigma(3)}})$$

boundedly for all p_1, p_2 *specified by* (1.6)*.*

Corollary 1.12 is obtained from Theorem 1.4 by noticing that S^p, $1 < p < \infty$ is intermediate UMD of exponent q for all $q > \max\{p, p'\}$. Similar statements may be obtained for more general tuples of non-commutative spaces $L^p(\mathcal{A})$ with the property that $\|A\|_{L^p(\mathcal{A})} \leq \|A\|_{L^q(\mathcal{A})}$ for $p > q$, so that (1.10) is a trilinear contraction in the range (1.11). We send to [31] for comprehensive definitions and background: a quick harmonic analyst-friendly introduction is given in [11, Section 3].

1.2. **Techniques of proof and novelties.** The standard proofs of L^p-bounds for the scalar-valued versions of the forms Λ_m in (1.1) are articulated in roughly three separate moments. The first is to realize that the forms (1.1) lie in the convex hull of suitable discretized model versions, the so-called *tri-tile forms*, displaying the same modulation and translation invariance properties of the condition (1.3): this step extends *verbatim* to the vector-valued case. We may thus focus on L^p-bounds for the model sums.

An essential step of the proof is the decomposition of the model operators into (discretized) multipliers which are adapted to a certain fixed top frequency and localized in space to a top interval. These *tree model sums* are essentially trilinear Calderón-Zygmund forms. The contribution of each tree is then controlled by localized space-frequency norms of the involved functions, the so-called *energies* (or sizes). This bound is referred to as *tree estimate*.

In the vector-valued case, this second step has to be adapted in a nontrivial and novel fashion. First of all, the vector-valued *energies*, introduced in (3.7), (3.10) must be defined in terms of local q-norms of (linear) tree operators rather than simply ℓ^2 sums of wavelet coefficients coming from each tree. We do so by means of a technical modification of the approach in [18]. Second, and most important, we obtain an effective tree estimate by replacing the involved functions with vector-valued phase-space projections to the space-frequency support of the tree. This extension of the scalar-valued phase-space projections of e.g. [10, 30] to UMD spaces, which may be of independent interest, is carried out in Proposition 4.3, and is the main technical novelty of the article. The tree model sum acts on the phase-space projections roughly as a trilinear CZ multiplier operator, and the L^p-norms of the constructed projections are controlled by the corresponding energies. These observations may be used in conjunction with the L^p-bound for UMD extensions of bilinear CZ operators, recently obtained by the authors of this paper in [11], to produce the tree estimate of Lemma 3.16.

Finally, the recomposition of the bounds obtained for each tree into a global estimate relies on almost-orthogonality considerations. To export this almost-orthogonality to the vector-valued scenario, we rely, as in previous literature [3, 18, 20], on the q-intermediate property of the involved spaces $\mathcal{X}$. This step is carried out in Lemma 3.13. As every known example of UMD space is q-intermediate for some q, this assumption may seem harmless. However, unlike the linear setting of [18], it is the combined q-intermediate type of the three spaces that introduces the restriction (1.5) and influences the range $\mathrm{int}(\mathcal{H})$ in Theorem 1.4. Further investigation on the necessity and on possible weakening of the q-intermediate assumptions are left for future work.

Plan of the paper. Section 2 contains the preliminary material needed to define the model tri-tile forms. Section 3 presents the outline of the proof of Theorem 1.4: in particular, the definitions of trees, vector-valued energies as well as the statement of the energy and tree lemmata, respectively Lemma 3.13 and 3.16. Section 4 contains the proof of the tree Lemma 3.16 via the reduction to the phase-space projection Proposition 4.3. The proof of the latter proposition is developed in Section 5. Section 6 contains the proof of the energy Lemma 3.13, while Lemma 3.12 is proved in the concluding Section 7. We include some of the pre-existing results of space-frequency analysis, adapted to the framework we work with, in an Appendix at the end of the article. We include the proof or the proof sketch whenever (small) adaptations are required, but claim no originality.

Remark. In the final stages of preparation of the present manuscript, the authors learned of the work [2] by Amenta and Uraltsev. These authors obtain a simultaneous and independent version of Theorem 1.4, focused on the bilinear Hilbert transform in the Banach range of exponents, under the same intermediate space condition (1.5). Interestingly, the methods employed in [2] are rather different from ours: the use of phase-space projections and of the UMD Calderón-Zygmund estimates from [11] is replaced by outer embeddings for the vector-valued wave packet transform involving telescoping (defect) energies.

The authors want to thank Alex Amenta and Gennady Uraltsev for sharing their preprint and for interesting discussions on the subject. They are also grateful to the very generous referees for the careful reading and the numerous suggestions which helped improving the quality of the article.

2. Space-frequency model sums

2.1. **Notation.** While our estimates are valid in any ambient space $\mathbb{R}^d$, we work with $d = 1$ to avoid unnecessary notational proliferation. However, we adopt d-dimensional terminology and notation whenever possible. For instance, we write $B_r(x) = \{y \in \mathbb{R} : |y - x| < r\}$ and simply B_r in place of $B_r(0)$. Whenever possible, spatial and frequency 1-dimensional cubes are indicated respectively by I, ω. The center and sidelength of a 1-dimensional cube I are respectively denoted by $c(I), \ell(I)$. We use the Japanese bracket notation $\langle x \rangle = \sqrt{1 + |x|^2}$.

If m is a bounded function on $\mathbb{R}$, we denote both the corresponding $L^2(\mathbb{R})$-bounded Fourier multiplier operator and its trivial extension to $L^2(\mathbb{R}) \otimes \mathcal{X}$ for any Banach space $\mathcal{X}$ as

$$T_m f(x) = \int \widehat{f}(\xi) m(\xi) \mathrm{e}^{ix\xi} \, \mathrm{d}\xi, \qquad x \in \mathbb{R}.$$

When $\mathcal{X}$ is a Banach space, we keep denoting by T the trivial extension $T \otimes \mathrm{Id}_{\mathcal{X}}$ of a linear operator T.

2.2. **Frequency-localized indicators.** Indicator functions, e.g. of intervals, possess perfect localization in space but poor frequency decay. We intend to define frequency-localized approximations of indicator functions by weakening such spatial localization to polynomial decay.

For this reason, we introduce suitable normalized classes of frequency-localized functions adapted to an interval $I \subset \mathbb{R}$. For a large positive integer N and $\delta, C > 0$ we say $\chi \in X_I(N,\delta,C)$ if $\chi \in \mathcal{S}(\mathbb{R})$ satisfies

$$\operatorname{supp}\widehat{\chi} \subset B_{\delta\ell(I)^{-1}}; \tag{2.1}$$

$$\chi \text{ real-valued}, \quad \left\langle \frac{x-c(I)}{\ell(I)} \right\rangle^{-N} \le \chi(x) \le C \left\langle \frac{x-c(I)}{\ell(I)} \right\rangle^{-N}, \qquad x \in \mathbb{R}. \tag{2.2}$$

If ψ instead satisfies (2.1) and

$$|\psi(x)| \le C \left\langle \frac{x-c(I)}{\ell(I)} \right\rangle^{-N}, \qquad x \in \mathbb{R}$$

in place of the more stringent (2.2), we say that $\psi \in \Psi_I(N,\delta,C)$. Obviously, we have the inclusion $X_I(N,\delta,C) \subset \Psi_I(N,\delta,C)$. It is important to notice that if I, I' are A-comparable intervals, that is $I \subset AI', I' \subset AI$ and $\chi \in X_I(N,\delta,C)$, then $c\chi \in X_{I'}(N,\delta',C')$ as well, for suitable constant c and values of δ', C' depending only on the comparability constant A and on N, δ, C. A similar statement applies to the classes $\Psi_I(N,\delta,C)$.

Suitable frequency-supported approximate indicators to $E \subset \mathbb{R}$ may be constructed as follows. For a fixed large positive integer N and $\delta > 0$, construct $\eta \in \mathcal{S}(\mathbb{R})$ satisfying

$$\widehat{\eta}(0) = 1, \qquad \operatorname{supp}\widehat{\eta} \subset B_\delta, \qquad \langle x \rangle^{-N} \le \eta(x) \lesssim_{N,\delta} \langle x \rangle^{-N} \quad \forall x \in \mathbb{R}.$$

We rescale η at frequency scale 2^j, $\eta_j := 2^j\eta(2^j\cdot)$, and for a positive integer J which we keep implicit in the notation of the left hand side, $E \subset \mathbb{R}$, $j \in \mathbb{R}$ we introduce

$$\chi_{E,j} = \mathbf{1}_E * \eta_{Jj}$$

whose frequency support is contained in $B_{\delta 2^{Jj}}$. The function $\chi_{E,j}$ is an approximate indicator in the sense that

$$\left|\chi_{E,j}(x) - \mathbf{1}_E(x)\right| \lesssim \left\langle \frac{\operatorname{dist}(x,\partial E)}{2^{-Jj}} \right\rangle^{-N+1}, \qquad x \in \mathbb{R}, \tag{2.3}$$

where ∂E is the topological boundary of E. This estimate is easily checked arguing separately in each case $x \in E$, $x \in \mathbb{R} \setminus E$. When $E = I$ is an interval with $\ell(I) = 2^{-Jj}$, the function $c\chi_{I,j}$ belongs to $X_I(N,\delta,C)$ for a suitable constant c depending only on the parameters (N,δ,C). We reserve for this case the simplified notation

$$\chi_I := \chi_{I,j} = \mathbf{1}_I * \eta_{Jj}. \tag{2.4}$$

It is important to notice that if I, I' are A-comparable intervals, that is $I \subset AI', I' \subset AI$ and $\chi \in X_I(N,\delta,C)$, then $\chi \in X_{I'}(N',\delta',C')$ as well, for suitable values of N', δ', C' depending only on the comparability constant A and on N, δ, C. A similar statement applies to the classes $\Psi_I(N,\delta,C)$.

Let now ω be a *frequency* interval. The class $M_\omega(N)$ will consist of those smooth functions m with $\operatorname{supp} m \subset \omega$ and adapted to ω of order N, in the sense that

$$\sup_{|\alpha|\le N} \sup_{\xi\in\mathbb{R}} \ell(\omega)^\alpha \left|\partial_\xi^\alpha m(\xi)\right| \le 1.$$

As customary, we will work with *tiles* $t = I_t \times \omega_t \subset \mathbb{R}\times\mathbb{R}$, namely the cartesian product of intervals in $\mathbb{R}$ of reciprocal length, to specify space-frequency localizations. Mimicking rank 1 projections in a Hilbert space, we may define classes of multiplier operators adapted to each tile t as follows. Whenever $\psi \in \Psi_{I_t}(N,\delta,C), m \in M_{\omega_t}(N)$, the operator

$$S_t f(x) = \psi(x) T_m f(x) = \int_{\mathbb{R}} \psi(x) m(\xi) \widehat{f}(\xi) \mathrm{e}^{ix\cdot\xi}\,\mathrm{d}\xi$$

is said to belong to the class $\mathbb{S}_t(N,\delta,C)$ of t-localized operators. From the rapid decay of the kernel of T_m and the adaptedness of ψ it follows that

$$\left|S_t f(x)\right|_{\mathcal{X}} \lesssim \left\langle \frac{\mathrm{dist}(x, I_t)}{\ell(I_t)} \right\rangle^{-100} M(|f|_{\mathcal{X}})(x), \qquad x \in \mathbb{R}.$$

Here $\mathcal{X}$ may be any Banach space, not necessarily UMD.

2.5. *Remark.* In the remainder of the paper, the values N, δ, C will be kept implicit and dropped from the notation for $X_I(N,\delta,C)$, $\Psi_I(N,\delta,C)$, $M_\omega(N)$ and $\mathbb{S}_t(N,\delta,C)$ whenever they vary within the fixed range

$$10^6 \le N \le \bar{N}, \quad C \le \bar{C}, \quad 2^{-3J} \le \delta \le 2^{-2J},$$

where J is a large integer depending on the nondegeneracy parameter Δ_β from (1.2) as specified in Appendix A.2. Therefore, the reader is warned that the precise values of these parameters may vary from line to line without explicit mention. In addition, if the function χ is such that $c\chi \in X_I$ for $c > 0$ varying in a fixed range depending on the parameters $\bar{N}, \bar{C}, J$, we abuse notation and write $\chi \in X_I$ instead. An advantageous example of usage for this convention is that whenever $\chi \in X_I$, the functions $\chi^m \in X_I$ as well for small values of $m \in \mathbb{N}$. We keep a similar convention for the other adapted classes.

The one place where we do not keep the parameters N, C implicit is in the definitions of the maximal energies (3.10) and their related quantities.

In our arguments we will make use of a form of Bernstein's inequality involving approximate indicators, in particular functions of the classes X_I described above. This is a known phenomenon in the literature, see e.g. [30, Lemma 5.4]; we give the proof as we are in the vector-valued context.

2.6. **Lemma.** *Let $R > 0$, $\mathcal{X}$ be a Banach space and f be an $\mathcal{X}$-valued function on $\mathbb{R}$ with*

$$\mathrm{supp}\,\widehat{f} \subset B_R.$$

Let $w : \mathbb{R} \to (0,\infty)$ be essentially constant at scale R^{-1}, namely

$$A^{-1}\langle R|x-y|\rangle^{-100} \le \frac{w(x)}{w(y)} \le A\langle R|x-y|\rangle^{100}, \qquad x, y \in \mathbb{R}$$

for some positive constant A. Then for all $0 < \alpha \le 1$

$$\|wf\|_{L^\infty(\mathbb{R};\mathcal{X})} \lesssim_{A,\alpha} R^\alpha \|wf\|_{L^{\frac{1}{\alpha}}(\mathbb{R};\mathcal{X})}.$$

Proof. Let ϕ be a smooth nonnegative function with $\phi = 1$ on B_R and $\phi = 0$ off B_{2R}. Notice that $|\widehat{\phi}(x)| \lesssim R\langle R|x|\rangle^{-200}$, for all $x \in \mathbb{R}$. Then $f = f * \widehat{\phi}$, and

$$\begin{aligned} |w(x)f(x)|_{\mathcal{X}} &\lesssim_{\alpha,A} R \int \frac{|w(y)f(y)|_{\mathcal{X}}}{\langle R|x-y|\rangle^{100}} \mathrm{d}y \\ &\le R^{\alpha} \|wf\|_{L^{\frac{1}{\alpha}}(\mathbb{R};\mathcal{X})} \left(\int \frac{R\mathrm{d}y}{\langle R|x-y|\rangle^{100}} \right)^{1-\alpha} \lesssim R^{\alpha} \|wf\|_{L^{\frac{1}{\alpha}}(\mathbb{R};\mathcal{X})} \end{aligned}$$

as claimed. The proof is complete. □

We will apply the lemma above to $w = \chi \in X_I$ for values $R \sim \ell(I)$.

2.3. **Tri-tiles and Rank 1 forms.** Trilinear multiplier forms of the type (1.1) admit a discretization in term of *tri-tiles*. There are several well-known versions of this discretization procedure, with origins rooted in Lacey and Thiele [27, 28] in the case of the bilinear Hilbert transforms. Working in the generality of the multipliers (1.1), whose singularity lies on the line $\Gamma' = \Gamma \cap \beta^{\perp}$, we choose to rely on the procedure described in [29, Section 5], where the multiplier (1.1) is decomposed in frequency via a partition of unity subordinated to a Whitney cover of $\mathbb{R}^3 \setminus \Gamma'$ by cubes $Q = Q_1 \times Q_2 \times Q_3 \in \mathbf{Q}$ coming from finitely many shifted dyadic grids in $\mathbb{R}^3$. Details are given in Subsection A.1 of the Appendix.

We say that the ordered triple of tiles $P = (P_1, P_2, P_3)$ is a *tri-tile* if

$$I_{P_1} = I_{P_2} = I_{P_3} =: I_P.$$

In accordance to the uncertainty principle, tri-tiles specify the space-frequency essential support of single scale multiplier forms, spatially concentrated on I_P and frequency supported on the *frequency cube*

$$Q_P = \omega_{P_1} \times \omega_{P_2} \times \omega_{P_3}. \tag{2.7}$$

Reflecting the invariance of condition (1.3) under the one-parameter family of translations along Γ', the collections of tri-tiles which are relevant to us are those whose frequency cubes come from the cover $\mathbf{Q}$, which is designed to be invariant under the one-parameter family of translations along Γ'. For this reason, we refer to these collections of tri-tiles as *rank 1 collections*: below, a formal definition is given.

In what follows, if J is a positive integer, and $\mathcal{G}$ is a subcollection of a dyadic system (or grid) on $\mathbb{R}$, we say that $\mathcal{G}$ is a J-separated dyadic grid if

$$\begin{aligned} &I, I' \in \mathcal{G},\ \ell(I) < \ell(I') \implies 2^J \ell(I) \le \ell(I'), \\ &I \ne I' \in \mathcal{G},\ \ell(I) = \ell(I') \implies \mathrm{dist}(I, I') \ge 2^{J+10}\ell(I). \end{aligned}$$

We say that $\mathbb{P}_1$ is a rank 1 collection of tri-tiles with parameters $K > 0$ and $J \in \mathbb{N}$, with $K \gg 1, 2^J \ge K^{10}$, if the following properties hold.

a. The frequency boxes $Q_P = \omega_{P_1} \times \omega_{P_2} \times \omega_{P_3}$ belong to the collection $\mathbf{Q}$ for all $P \in \mathbb{P}_1$.
b. The collection $\mathcal{I} = \{I_P : P \in \mathbb{P}_1\}$ is the collection of all dyadic intervals on $\mathbb{R}$ whose sidelengths are of the form $\ell(I) = 2^{jJ+u}$ for some $j \in \mathbb{Z}$, where u is any fixed integer in $\{0, \dots, J-1\}$. The collections $\Omega_k = \{\omega_{P_k} : P \in \mathbb{P}_1\}$, $k = 1, 2, 3$ are J-separated dyadic grids. Furthermore, there exist additional J-separated collections of dyadic intervals $\overline{\Omega_k}$, $k = 1, 2, 3$, such that for each $P \in \mathbb{P}_1$ and $k = 1, 2, 3$ we may find

$\overline{\omega_{P_k}} \in \overline{\Omega_k}$ with $K^4\omega_{P_k} \subset \overline{\omega_{P_k}}$ and $\ell(\omega_{P_k}) \geq K^{-5}\ell(\overline{\omega_{P_k}})$. Thus both the intervals $\omega_{P_k}, P \in \mathbb{P}_1$ and their K^4 dilates have good dyadic properties.

c. If $P \neq P' \in \mathbb{P}_1$ are such that $I_P = I_{P'}$ then $\omega_{P_j} \cap \omega_{P'_j} = \varnothing$ for each $j \in \{1,2,3\}$.

d. The rank 1. properties r8. to r10. hold for all choices of $P, P' \in \mathbb{P}_1$. These properties are stated in Section A.2 of the Appendix in terms of certain approximate order relations among tri-tiles. Here we give the equivalent explicit description

$$\begin{aligned} &\operatorname{dist}(Q_P, \Gamma) \leq 2\ell(Q_P);\\ &P, P' \in \mathbb{P}_1,\ I_P \subset I_{P'},\ \omega_{P'_k} \subset 5\omega_{P_k} \text{ for some } k \in \{1,2,3\}\\ \implies\ &\omega_{P'_\kappa} \subset K\omega_{P_\kappa} \setminus 3\omega_{P_\kappa} \text{ for all } \kappa \in \{1,2,3\} \setminus \{k\}. \end{aligned}$$

The singular multiplier forms Λ_m from (1.1) then lie in the convex hull of the *tri-tile forms*

$$\Lambda_{\mathbb{P}}(f_1, f_2, f_3) = \sum_{P \in \mathbb{P}} \int_{\mathbb{R}} \prod_{k=1}^{3} \chi_{I_P}(x) T_{m_{P_k}} f_k(x)\, \mathrm{d}x, \tag{2.10}$$

where $\mathbb{P}$ is the (finite) subset of a rank 1 collection of tri-tiles $\mathbb{P}_1$ whose spatial intervals $\{I_P : P \in \mathbb{P}\}$ are contained in a fixed but arbitrary J-dyadic interval and whose frequency box set $\{Q_P : P \in \mathbb{P}\}$ is finite but arbitrary, $\chi_{I_P} \in X_{I_P}$ has been defined in (2.4) of Subsection 2.2 and $m_{P_k} \in M_{\omega_{P_k}}(N), P \in \mathbb{P}$ satisfies the consistency condition

$$Q_P = Q_{P'} \implies m_{P_k} = m_{P'_k},\ k = 1,2,3, \tag{2.11}$$

referring to (2.7). Therefore, to bound Λ_m it suffices to bound $\Lambda_{\mathbb{P}}$ uniformly. A detailed proof of these statements is given in [29, Section 5], see also [30,35]. A summary of proof is given in the Appendix. The remainder of the article will be devoted to the proof of such uniform bounds for $\Lambda_{\mathbb{P}}$. Note that the operators

$$f \mapsto \chi_{I_P} T_{m_{P_k}} f$$

belong to the class $\mathbb{S}_{P_k}$, for $k = 1,2,3$.

2.12. *Remark.* As most of the analysis in this paper is an extension to UMD spaces of the phase space projection technique introduced in [30], we take a moment to explain how our choice of discretization relates to the one developed in the latter reference. The main point of [30] is obtaining bounds uniform in the degeneracy parameter (1.2), which may be identified with the absolute ratio between largest and smallest eigenvalues of a map L rescaling the subspace $\mathbb{R}(1,1,1)$ to the singular line Γ' of m. In [30], the multiplier m is pulled back to $m(L\cdot)$ and then discretized with a partition of unity subordinated to the finitely overlapping cover

$$\mathbf{Q} = \left\{Q := L^{-1}\tilde{Q} : \tilde{Q} \in \tilde{\mathbf{Q}}\right\},$$

where $\tilde{\mathbf{Q}}$ is instead a Whitney decomposition of $\mathbb{R}^3 \setminus \mathbb{R}(1,1,1)$ into cubes. The Whitney pieces of the multiplier are thus adapted to boxes Q whose sidelengths are scaled by the eigenvalues of the map L. When the degeneracy parameter Δ_β in (1.2) is $\gtrsim 1$, the scaling map L may be neglected by suitable finite splittings and $Q, \tilde{Q}$ may be conflated; the model of [30] essentially coincides with our discretization. The reader who wants to compare our analysis to that of [30] is encouraged to do so.

3. Proof of Theorem 1.4: tree and energy estimates

In this section, after devising the necessary definitions in our context, we present the statements of the three main lemmas, which may then be combined to prove Theorem 1.4 in a standard fashion.

We first introduce *trees*, roughly speaking, collections of tri-tiles sitting at a common frequency and spatially localized to an interval. Then we define *tree operators*, that is modulated Calderón-Zygmund localized operators associated to each tree. These are used to define the *energy* of a certain $\mathcal{X}$-valued function with respect to a set of tri-tiles $\mathbb{P}$: this is a sort of localized maximal $L^q(\mathbb{R};\mathcal{X})$-norm of tree operators coming from $\mathbb{P}$.

Finally, we state the main steps in the proof of Theorem 1.4. The first is the *energy lemma*, which allows us to decompose any given collection of tri-tiles into unions of trees of controlled energy for each function f_k of bounded spatial support. The second is the *tree lemma*, which provides a bound of the tri-tile form (2.10) when $\mathbb{P}$ is a tree. The proof of this lemma is one of the main novelties of this article, as it relies on a combination of the multilinear UMD CZ theory of [11] with newly developed phase-space projections adapted to the vector-valued setting.

3.1. **Trees.**

In the definitions below, an important role is played by the singular line of m in (1.3), namely $\Gamma' = \Gamma \cap \beta^{\perp}$. Let $\mathbb{P}$ be a rank 1 collection of tritiles. Rank 0 subcollections of $\mathbb{P}$, whose associated forms $\Lambda_{\mathbb{P}}$ are discretized multilinear CZ type multipliers, are called *trees*. We work with a specific notion of tree which satisfies certain additional properties along the lines of [30, Section 4].

Recall from Subsection 2.3 that if $P \in \mathbb{P}$, $\overline{\omega_{P_k}}$ is an interval approximating the dilate $K^4\omega_{P_k}$ and coming from a fixed J-separated dyadic grid. We say that the nonempty collection $\mathsf{T} \subset \mathbb{P}$ is a tree having $(I_\mathsf{T}, \xi_\mathsf{T})$ as top data if the following conditions hold.

a. I_T is a J-dyadic interval in $\mathbb{R}$ and

$$I_P \subset I_\mathsf{T} \qquad \forall P \in \mathsf{T},\ k = 1,2,3.$$

b. $\xi_\mathsf{T} = ((\xi_\mathsf{T})_1, (\xi_\mathsf{T})_2, (\xi_\mathsf{T})_3) \in \Gamma'$ and

$$(\xi_\mathsf{T})_k \in \overline{\omega_{P_k}} \qquad \forall P \in \mathsf{T},\ k = 1,2,3.$$

It is convenient to denote by $\mathbf{j}_\mathsf{T} = \{j \in \mathbb{Z} : \ell(Q_P) = 2^{jJ} \text{ for some } P \in \mathsf{T}\}$, the frequency scales appearing in T. Then

$$\mathsf{T} = \bigcup_{j \in \mathbf{j}_\mathsf{T}} \mathsf{T}(j), \qquad \mathsf{T}(j) = \{P \in \mathsf{T} : \ell(Q_P) = 2^{Jj}\}. \tag{3.1}$$

We also take the opportunity here to observe that trees constructed via greedy selection processes, such as the ones in the proof of Lemma 3.13 below and explicitly defined in Section A.5 of the Appendix, satisfy the following additional properties.

g1. The *frequency localization sets* $\mathbf{Q}_\mathsf{T} = \{Q_P : P \in \mathsf{T}\}$ are such that

$$Q, Q' \in \mathbf{Q}_\mathsf{T},\ \ell(Q) = \ell(Q') \implies Q = Q';$$

namely, there is only one frequency localization for each J-dyadic scale.

g2. The *spatial localization sets*

$$E_{Q,\mathsf{T}} = \bigcup \{I_P : P \in \mathsf{T}, Q_P = Q\}, \qquad Q \in \mathbf{Q}_\mathsf{T}$$

are nested, that is

$$Q, Q' \in \mathbf{Q}_{\mathsf{T}},\ \ell(Q) \le \ell(Q') \implies E_{Q,\mathsf{T}} \supset E_{Q',\mathsf{T}}.$$

Furthermore, a family of sets $\{E_j : j \in \mathbb{Z}\}$ with the properties that

g3. $\tilde{E}_{j+1} \subset \tilde{E}_j\ \forall j \in \mathbb{Z}$, and $Q \in \mathbf{Q}_{\mathsf{T}}, \ell(Q) = 2^{Jj} \implies E_{Q,\mathsf{T}} \subset \tilde{E}_j$

and with useful smoothing properties may be constructed as detailed in Section A.5 of the Appendix. By virtue of these observations, we may rely on g1. to g3. when proving the phase-space projection estimates of Proposition 4.3. We send to Section A.5 of the Appendix for the proofs of properties g1. to g3. and more detailed statements.

3.2. **Tree operators.** We now introduce two special types of trees with different frequency localization properties. We say that the tree T is k-lacunary for a certain index $k \in \{1, 2, 3\}$ if

$$\{2\omega_{P_k} : P \in \mathsf{T}\} \text{ are a pairwise disjoint collection.} \tag{3.2}$$

3.3. *Remark.* Let T be a k-lacunary tree for a certain $k \in \{1, 2, 3\}$, and suppose that $\mathsf{T}_{\mathsf{in},k} := \{P \in \mathsf{T} : (\xi_{\mathsf{T}})_k \in 2\omega_{P_k}\}$ is nonempty. From (3.2) and property g1. we immediately see that $\mathsf{T}_{\mathsf{in},k} = \mathsf{T}(j_{\mathsf{in},k})$ for some $j_{\mathsf{in},k} \in \mathbf{j}_{\mathsf{T}}$. Notice that $\mathsf{T}_{\mathsf{in},k}$ is also a k-lacunary tree with the same top data as T. If $\mathsf{T}_{\mathsf{in},k} = \varnothing$, we set $j_{\mathsf{in},k} = -\infty$ for unifying purposes.

3.4. *Remark.* A consequence of properties r9. and r10. of rank 1 collections is that each tree T can be written as the disjoint union

$$\mathsf{T} = \bigcup_{\substack{A \subset \{1,2,3\} \\ \#A \ge 2}} \mathsf{T}_A, \tag{3.5}$$

where each T_A is a tree with the same top data as T and has the additional property (3.2) for $k \in A$, while

$$3\omega_{P_\kappa} \ni (\xi_{\mathsf{T}})_\kappa, \qquad \forall \kappa \in B := \{1, 2, 3\} \setminus A. \tag{3.6}$$

We prove this claim in the Appendix, Section A.4.

We then introduce tree operators associated to k-lacunary trees. For our purposes here, we need a more refined object than the usual, e.g. appearing in [9, 18, 27, 28], fully discretized tree operator

$$f \mapsto \sum_{P \in \mathsf{T}} \langle f, \varphi_{P_k} \rangle \varphi_{P_k}, \qquad k \in A,$$

where φ_{P_k} is a wave packet adapted to the tile P_k. Let T be a k-lacunary tree. A (scalar) tree operator of k-th type is the linear operator

$$S_{\mathsf{T}} f = \sum_{P \in \mathsf{T}} S_{P_k} f,$$

where each $S_{P_k} \in \mathbb{S}_{P_k}$, $P \in \mathsf{T}$. When $\xi_{\mathsf{T}} = 0$, the defined tree operator is a pseudo-differential operator with symbol

$$a(x, \xi) = \sum_{P \in \mathsf{T}} \psi_{P_k}(x) m_{P_k}(\xi),$$

where each $\psi_{P_k} \in \Psi_{I_P}$ and $m_{P_k} \in M_{\omega_{P_k}}$. A routine computation relying on the space-frequency localization of S_{P_k} verifies that this symbol is uniformly of class $S^0_{1,1}$. Further, as the intervals $\{\omega_{P_k} : P \in \mathsf{T}\}$ are pairwise disjoint, S_T is uniformly $L^2(\mathbb{R})$ bounded. Relying on these two observations, we gather that S_T is an $L^2(\mathbb{R})$-bounded Calderón-Zygmund operator, see for instance the discussion at [34, p. 271]. Therefore, S_T satisfies uniform $L^q(\mathbb{R}; \mathcal{X})$ bounds, $1 < q < \infty$, as well as $L^\infty(\mathbb{R}; \mathcal{X}) \to \mathrm{BMO}(\mathbb{R}; \mathcal{X})$ estimates, whenever $\mathcal{X}$ is a UMD Banach space (see [14, 16]). In fact, by modulation invariance, we may remove the $\xi_\mathsf{T} = 0$ assumption and conclude that tree operators S_T are uniformly $L^q(\mathbb{R}; \mathcal{X})$ bounded, when $1 < q < \infty$.

The definition of k-overlapping tree is very simple. We say that a tree T with top data $(I_\mathsf{T}, \xi_\mathsf{T})$ is k-overlapping for some $k = 1, 2, 3$ if

$$(\xi_\mathsf{T})_k \in 3\omega_{P_k}, \qquad \forall P \in \mathsf{T}.$$

Notice that the collection of a single tri-tile $\{P\}$ may be made into a k-lacunary tree by picking as top data (I, ξ) whenever I is a J-dyadic interval with $I_P \subset I$ and $\xi_k \in K\omega_{P_k} \setminus 2\omega_{P_k}$. However, a single tri-tile $\{P\}$ may also be made into a k-overlapping tree by picking as top data (I, ξ) whenever I is a J-dyadic interval with $I_P \subset I$ and $\xi_k \in 3\omega_{P_k}$. We will use the latter observation in the next definition.

3.3. **Energy and energy lemma.** This definition is a re-elaboration of [18, Section 8]. Let $q \geq 2$ and f be a $\mathcal{X}$-valued function. If T is a k-lacunary tree with top data $(I_\mathsf{T}, \xi_\mathsf{T})$ we define

$$\|f\|_{\mathsf{lac};\mathsf{T},k,q} = \sup \frac{1}{|I_\mathsf{T}|^{\frac{1}{q}}} \left\| S_\mathsf{T} f \right\|_{L^q(\mathbb{R};\mathcal{X})}, \tag{3.7}$$

where the supremum is taken over all possible choices of type k tree operators $S_\mathsf{T} = \sum_{P \in \mathsf{T}} S_{P_k}$ normalized to satisfy $S_{P_k} \in \mathbb{S}_{P_k}(10^4, \delta, 1)$. We give an analogous definition for k-overlapping trees. If T is a k-overlapping tree with top data $(I_\mathsf{T}, \xi_\mathsf{T})$, we define the corresponding tile $t_{\mathsf{T},k}$

$$t_{\mathsf{T},k} = I_\mathsf{T} \times \omega_{\mathsf{T},k}, \qquad \omega_{\mathsf{T},k} := \Big[(\xi_\mathsf{T})_k - [2\ell(I_\mathsf{T})]^{-1}, (\xi_\mathsf{T})_k + [2\ell(I_\mathsf{T})]^{-1} \Big). \tag{3.8}$$

and

$$\begin{aligned} \|f\|_{\mathsf{ov};\mathsf{T},k,q} &:= \sup_{m \in \widetilde{M}_{\omega_{\mathsf{T},k}}} \sup_{\zeta \in X_{I_\mathsf{T}}(10^4, \delta, \bar{C})} \frac{1}{|I_\mathsf{T}|^{\frac{1}{q}}} \left\| \zeta T_m f \right\|_{L^q(\mathbb{R};\mathcal{X})}, \\ \widetilde{M}_{\omega_{\mathsf{T},k}} &:= \{ m \in M_{\omega_{\mathsf{T},k}} : m((\xi_\mathsf{T})_k) = 0 \}. \end{aligned} \tag{3.9}$$

Notice that, referring to (3.9), $\zeta T_m \in \mathbb{S}_{t_{\mathsf{T},k}}$, but we require the extra property that m vanishes at $(\xi_\mathsf{T})_k$. We also need to define maximal versions: for each set $\mathbb{P}$ of tri-tiles,

$$\begin{aligned} \mathsf{eng}_{\mathsf{lac};k}(f)(\mathbb{P}; q) &:= \sup_{\substack{\mathsf{T} \subset \mathbb{P} \\ \mathsf{T}\, k\text{-lacunary}}} \|f\|_{\mathsf{lac};\mathsf{T},k,q}, \\ \mathsf{eng}_{\mathsf{ov};k}(f)(\mathbb{P}; q) &:= \sup_{\substack{\mathsf{T} \subset \mathbb{P} \\ \mathsf{T}\, k\text{-overlapping}}} \|f\|_{\mathsf{ov};\mathsf{T},k,q}, \\ \mathsf{eng}_k(f)(\mathbb{P}; q) &:= \max \left\{ \mathsf{eng}_{\mathsf{lac};k}(f)(\mathbb{P}; q), \mathsf{eng}_{\mathsf{ov};k}(f)(\mathbb{P}; q) \right\}. \end{aligned} \tag{3.10}$$

For instance, we have[1]

$$\Big\|\sum_{P\in\mathsf{T}} \chi_{I_P} T_{m_{P_k}} f\Big\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim |I_{\mathsf{T}}|^{\frac{1}{q}} \mathsf{eng}_{\mathsf{lac};k}(f)(\mathbb{P};q)$$

when χ_{I_P} and m_{P_k} are as in (2.10), whenever $\mathsf{T}\subset\mathbb{P}$ is a k-lacunary tree, as $f\mapsto c\chi_{I_P}T_{m_{P_k}}f$ belongs to $\mathbb{S}_{P_k}(10^4,\delta,1)$ for a uniform constant c.

We briefly explain the usage we make of the second term in (3.10) with a lemma.

3.11. **Lemma.** *Fix $k\in\{1,2,3\}$. Let I be a J-dyadic interval with $\ell(I)=2^{-Jj}$, $\xi\in\Gamma'$ and ω be an interval centered at ξ_k and satisfying*

$$2^{Jj}\le \ell(\omega)\le 2^{J(j+5)}.$$

Suppose that the tri-tile P is such that

$$I_P\subset 10I,\qquad \ell(I_P)\le 2^{-J(j+10)},\qquad \xi_k\in 3\omega_{P_k}.$$

Then

$$\sup_{\chi\in X_I}\sup_{m\in\widetilde{M}_\omega}\big\|\chi T_m f\big\|_{L^q(\mathbb{R};\mathcal{X})}\lesssim |I|^{\frac{1}{q}}\mathsf{eng}_{\mathsf{ov};k}(f)(\{P\};q).$$

Proof. Let $\chi\in X_I$, $m\in\widetilde{M}_\omega$. Let I' be the unique J-dyadic interval with $2^{-J}<\ell(\omega)\ell(I')\le 1$ that contains I_P, which must exist because of the relations between lengths. Then $\mathrm{dist}(I,I')\le 10\ell(I)$ and $\ell(I)\sim_J \ell(I')$, whence $\chi\in X_{I'}$ up to constants. For this reason we may as well assume $I=I'$. In this case, the collection $\{P\}$ is a k-overlapping tree T with top data (I,ξ), as $\omega\subset\omega_{\mathsf{T},k}$ and $\ell(\omega)\ge\frac{1}{2}\ell(\omega_{\mathsf{T},k})$, and the claim follows. □

The next lemma is a variation of e.g. [18, Corollary 9.6] .

3.12. **Lemma.** *Let $\mathbb{P}$ be a finite collection of tri-tiles. Then*

$$\mathsf{eng}_k(f)(\mathbb{P};q)\lesssim \sup_{P\in\mathbb{P}}\inf_{I_P}\mathrm{M}(|f|_{\mathcal{X}}).$$

Although the arguments of [18] may be adapted to the context of Lemma 3.12, we provide a more direct proof in Section 7.

In the last main lemma, the quantitative assumption of $\mathcal{X}_k$ being an interpolation space is used. We could alternatively bring forth definitions akin to the tile-type of a Banach space in [18,20,21], which is a formal consequence of our intermediate UMD assumption, but for simplicity and lack of examples we give up on this additional formal generality.

3.13. **Lemma.** *Suppose $\mathcal{X}$ is $q_{\mathcal{X}}$-intermediate* UMD *and let $f\in L^\infty(\mathbb{R};\mathcal{X})$ be subordinated to the finite measure set F, namely $|f|_{\mathcal{X}}\le \mathbf{1}_F$. Fix $q>q_{\mathcal{X}}$ and let $\mathbb{P}$ be a finite set of tri-tiles. Then $\mathbb{P}=\mathbb{P}^{\mathsf{low}}\cup\mathbb{P}^{\mathsf{hi}}$ with the property that*

$$\mathsf{eng}_k(f)(\mathbb{P}^{\mathsf{low}};q_{\mathcal{X}})\le 2^{-1}\mathsf{eng}_k(f)(\mathbb{P};q_{\mathcal{X}}) \tag{3.14}$$

and that $\mathbb{P}^{\mathsf{hi}}$ is a union of trees $\mathsf{T}\in\mathbf{T}$ with the property that

$$\sum_{\mathsf{T}\in\mathbf{T}}|I_{\mathsf{T}}|\lesssim_q \left[\mathsf{eng}_k(f)(\mathbb{P};q_{\mathcal{X}})\right]^{-q}|F|. \tag{3.15}$$

[1]The point of this example is to explain how the normalization in the definitions (3.7), (3.9) plays a role in relation to Remark 2.5.

The proof of Lemma 3.13 is a revisitation of the steps leading to [18, Proposition 8.4] and is postponed to Section 6. Note that Lemma 3.13 is the only main step of the proof of Theorem 1.4 where a $q_{\mathcal{X}}$-intermediate assumption is used.

The final main tool of the proof of Theorem 1.4 is a bound on the forms (2.10) in terms of energy parameters in the particular case where the collection $\mathbb{P}$ is a tree.

3.16. **Lemma.** *Let $\mathcal{X}_k$, $k = 1, 2, 3$ be* UMD *spaces and*

$$2 \leq q_1, q_2, q_3 < \infty, \qquad \sum_{k=1}^{3} \frac{1}{q_k} \geq 1. \tag{3.17}$$

Let T *be a tree. With reference to* (2.10) *there holds*

$$|\Lambda_{\mathsf{T}}(f_1, f_2, f_3)| \lesssim |I_{\mathsf{T}}| \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k)$$

uniformly over all choice of tri-tile forms Λ_{T}.

The proof of Lemma 3.16 uses a novel vector-valued version of the phase-space projection technique of [10,30] in conjunction with [11, Theorem 1.2] and is given in Section 4.

3.4. **Proof of Theorem 1.4.** We are now ready to compile the proof of the main theorem. A standard combination of Lemmata 3.16, 3.13 and 3.12 yields a range of restricted weak type estimates for the forms (2.10): the elementary procedure is identical to that leading to [20, Corollary 9.2]. These estimates then entail Theorem 1.4 by standard multilinear restricted weak type interpolation, see e.g. [35]. This deduction is the same as that of [20, Theorem 9.3] from [20, Corollary 9.2]. We omit the details.

4. Phase space projections and the proof of the tree lemma

We develop phase-space projections in the vector-valued context and combine them with the bounds for vector-valued extensions of bilinear CZ operators to prove Lemma 3.16. The following treatment is an adaptation of the construction made in [30, Sec. 7 and 8]. Our arguments are more involved due to the vector-valued nature of the involved functions. However, we take advantage of a significant simplification in that no uniformity issues are considered: in the language of [30] the indices $\mathbf{m}_i$ are all zero. Uniform estimates in the vector-valued context will be the object of future work.

In the main proposition of this section, we make use of Littlewood-Paley projections as follows. The operator T_j stands for a Fourier multiplier whose symbol Φ_j is real, even and

$$\operatorname{supp} \Phi_j \subset (-2^{J(j+2)}, 2^{J(j+2)}), \quad \Phi_j(\xi) = 1 \text{ on } [-2^{J(j+2)-1}, 2^{J(j+2)-1}]. \tag{4.1}$$

Then the projections $S_j := T_j - T_{j-1}$ are Fourier multiplier with symbol Ψ_j satisfying

$$\operatorname{supp} \Psi_j \subset \{\xi : 2^{J(j+1)-1} \leq |\xi| \leq 2^{J(j+2)}\}, \quad \Psi_j(\xi) = 1 \text{ for } 2^{J(j+1)} \leq |\xi| \leq 2^{J(j+2)-1}. \tag{4.2}$$

The projections S_j appear also in Lemma 4.12 below.

4.3. **Proposition** (phase-space projections)**.** *For $k = 1, 2, 3$, let $\mathcal{X}_k$ be a* UMD *space and $q_k \in [2, \infty)$. Let* T *be a tree with the following properties:*

i. $\xi_{\mathsf{T}} = 0$;

ii. T *is k-lacunary for* $k \in A$, *in the sense that* (3.2) *holds, and not* κ*-lacunary for* $\kappa \in B$, *in the sense of* (3.6), *with* $A \cup B = \{1,2,3\}$ *disjoint union and* $\#A \in \{2,3\}$*;*

iii. *the separation of scales condition*

$$\inf\{|j - j'| : j, j' \in \mathbf{j}_{\mathsf{T}}, j \neq j'\} \geq 10$$

holds true.

Choose $\{S_{P_k} \in \mathbb{S}_{P_k}, P \in \mathsf{T}, k \in \{1,2,3\}\}$. *Then there are linear operators* Π_k *with the following properties.*

a. *If* $p \geq q_k$, *there holds*

$$\|\Pi_k f\|_{L^p(\mathbb{R};\mathcal{X}_k)} \lesssim |I_{\mathsf{T}}|^{\frac{1}{p}} \mathsf{eng}_k(f)(\mathsf{T}; q_k); \tag{4.4}$$

b. *if* $k \in A$, *for all* $j \in \mathbf{j}_{\mathsf{T}}$, *with reference to Remark 3.3 and to* (4.1), *we have the equality*

$$\sum_{P \in \mathsf{T}(j)} S_{P_k} f = S_j(\Pi_k f), \qquad j \neq j_{\mathsf{in},k}.$$

c. *If* $p \geq q_k$, $k \in B$, $j_0 \in \mathbf{j}_{\mathsf{T}}$ *and* $\ell(I_0) = 2^{-Jj_0}$,

$$\left\| \mathbf{1}_{I_0} \sum_{P \in \mathsf{T}(j_0)} S_{P_k}(f - \Pi_k f) \right\|_{L^p(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{p}} \mathsf{eng}_k(f)(\mathsf{T}; q_k) \int_{\mathbb{R}} \chi_{I_0}(x) \mu_{j_0}(x) \frac{\mathrm{d}x}{|I_0|}. \tag{4.5}$$

where μ_j *is defined in* (A.11). *Furthermore*

$$\left\| \mathbf{1}_{I_0} \sum_{P \in \mathsf{T}(j_0)} S_{P_k} \Pi_k f \right\|_{L^p(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{p}} \mathsf{eng}_k(f)(\mathsf{T}; q_k). \tag{4.6}$$

While the operators Π_k *depend on the choice of* $\{S_{P_k} \in \mathbb{S}_{P_k}, P \in \mathsf{T}, k \in \{1,2,3\}\}$, *the estimates above are uniform over such choice.*

The proof of Proposition 4.3 is postponed to the next section. Herein, we proceed to showing how this proposition may be coupled with the main result of [11] to obtain the tree estimate we claimed in Lemma 3.16. The next subsection contains some preliminaries, while the main line of argument, namely the proof of estimate (4.11), is deployed in Subsection 4.2.

4.1. **Preliminaries.** We begin with a preliminary localized single scale estimate for tree operators which will be of use towards Lemma 3.16 as well as in Section 5.

4.7. **Lemma.** *Let* T *be any tree,* I_0 *be a J-dyadic interval with* $\ell(I_0) = 2^{-j_0 J}$, $\psi_{I_0} \in \Psi_{I_0}$, $S_{P_k} \in \mathbb{S}_{P_k}$ *for each* $P \in \mathsf{T}(j_0)$. *Then for any* $k \in \{1,2,3\}$,

$$\left\| \psi_{I_0} \sum_{P \in \mathsf{T}(j_0)} S_{P_k} f \right\|_{L^p(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{p}} \mathsf{eng}_k(f)(\mathsf{T}; q_k), \qquad q_k \leq p \leq \infty. \tag{4.8}$$

Proof. Arguing by interpolation, it suffices to prove the extremal cases. Notice also that we may assume $\psi_{I_0} \in X_{I_0}$ by possibly replacing $\psi_{I_0} \in \Psi_{I_0}$ with a pointwise majorant in X_{I_0}.

Case $p = q_k$. Let $n \in \mathbb{N}$. By virtue of J-dyadicity and property g1. of greedily constructed trees, there are at most two $P \in \mathsf{T}(j_0)$ such that $\operatorname{dist}(I_P, I_0) = n2^{-j_0 J}$. Fix such a P. It then suffices to estimate

$$\left\|\psi_{I_0} S_{P_k} f\right\|_{L^{q_k}(\mathbb{R};X_k)} \lesssim \langle n\rangle^{-100} |I_0|^{\frac{1}{q_k}} \mathsf{eng}_k(f)(\mathsf{T}; q_k).$$

Write $S_{P_k} f = \psi T_m f$. Then $\widetilde{\psi} := \langle n\rangle^{100} \psi_{I_0} \psi \in \Psi_{I_P}$ and the estimate in the last display simply follows from the definition of $\mathsf{eng}_k(f)(\mathsf{T}; q_k)$.

Case $p = \infty$. The function we are estimating has frequency support in a ball of radius $O(2^{j_0 J})$. Then this case follows from the case $p = q_k$ and a straightforward application of Lemma 2.6. □

We then particularize the definition (2.10) to the case where $\mathbb{P}$ is our tree T. By translation and scaling invariance, we may reduce Lemma 3.16 to the case $I_{\mathsf{T}} = [0, 1)$. By invariance with respect to modulations along the subspace Γ', we may also reduce to the case $\xi_{\mathsf{T}} = 0$. Notice that, referring to (2.7), (4.1), property g1. ensures that we have

$$Q_P = Q_{P'} =: Q_j \qquad P, P' \in \mathsf{T}(j).$$

Consequently, in view of (2.11), $m_{P_k} = m_{P'_k} := m_{j,k}$ for all $P, P' \in \mathsf{T}(j), k = 1, 2, 3$. Therefore we may set for $j \in \mathbf{j}_{\mathsf{T}}$, referring to (2.4)

$$\begin{aligned} \tilde{\chi}_j &:= \sum_{P \in \mathsf{T}(j)} \chi_{I_P} \\ \tilde{\pi}_{j,k} &:= T_{m_{j,k}}, \qquad k = 1, 2, 3 \end{aligned}$$

and rewrite, and subsequently estimate, (2.10) for $\mathsf{T} = \mathbb{P}$ as

$$\left|\Lambda_{\mathsf{T}}(f_1, f_2, f_3)\right| = \left|\sum_{j \in \mathbf{j}_{\mathsf{T}}} \int_{\mathbb{R}} \tilde{\chi}_j \prod_{k=1}^{3} \tilde{\pi}_{j,k} f_k\right| \le \sum_{j \in \mathbf{j}_{\mathsf{T}}} \left|\int_{\mathbb{R}} \tilde{\chi}_j \prod_{k=1}^{3} \tilde{\pi}_{j,k} f_k\right|. \tag{4.9}$$

We may turn $\tilde{\chi}_j$ into $(\tilde{\chi}_j)^3$ by virtue of the bound

$$\sum_{j \in \mathbf{j}_{\mathsf{T}}} \left|\int_{\mathbb{R}} \tilde{\chi}_j(x) \prod_{k=1}^{3} \tilde{\pi}_{j,k} f_k(x) - \prod_{k=1}^{3} \tilde{\chi}_j(x) \tilde{\pi}_{j,k} f_k(x)\, \mathrm{d}x\right| \lesssim \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \tag{4.10}$$

whose proof is given at the end of this Subsection. As the error in (4.10) is acceptable for the estimate of Lemma 3.16, we have reduced the tree Lemma 3.16 to proving that

$$\left|\sum_{j \in \mathbf{j}_{\mathsf{T}}} \varepsilon_j \int_{\mathbb{R}} \prod_{k=1}^{3} \tilde{\chi}_j \tilde{\pi}_{j,k} f_k\right| \lesssim \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \tag{4.11}$$

uniformly over choices of unimodular coefficients $\{\varepsilon_j : j \in \mathbf{j}_{\mathsf{T}}\}$ which is the core of the argument, and is left for the next subsection.

The final preliminary result is a Hölder type estimate for the vector-valued extension of a classical trilinear paraproduct form. Such estimate is a particular case of the main result of [11] and depends only on the UMD property of the spaces involved. The proof is postponed to the end of this Subsection.

4.12. **Lemma.** *Let $\{p_k : k = 1,2,3\}$ be a Hölder tuple of exponents with $1 < p_k < \infty$ for all $k = 1,2,3$. Let $\mathcal{X}_k$ be* UMD *spaces with a trilinear contraction $\prod_{k=1}^3 \mathcal{X}_k \to \mathbb{C}$. Let $g_k \in (L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})) \otimes \mathcal{X}_k$, for $k = 1,2,3$. Then*

$$\left| \int \sum_{j \in \mathbf{j}_\mathsf{T}} \varepsilon_j (\tilde{\pi}_{j,1} g_1)(S_j g_2)(S_j g_3) \right| \lesssim \prod_{k=1}^3 \|g_k\|_{L^{p_k}(\mathbb{R};\mathcal{X}_k)}.$$

Proof of (4.10). This is analogous to [30, Lemma 7.3]. First of all, we bound the single scale pieces of (4.9). Relying on (3.17), we may find a Hölder tuple p_1, p_2, p_3 with $q_k \le p_k < \infty$. If $I \in \mathcal{D}_j$, we may pick[2] $\overline{P} \in \mathsf{T}(j)$ and $\zeta \in X_{I_{\overline{P}}}$ such that $\zeta^{12} \gtrsim \tilde{\chi}_j$ on I. In the display below, we couple this with Hölder inequality followed by Lemma 4.7 with $I_0 = I_{\overline{P}}$, $\psi_{I_0} = \zeta$, and the operators $S_{P_k} = \zeta \tilde{\pi}_{j,k}$ when $P = \overline{P}$ and $S_{P_k} = 0$ otherwise. We obtain

$$\int_I (\tilde{\chi}_j)^{\frac{1}{2}} \prod_{k=1}^3 |\tilde{\pi}_{j,k} f_k| \lesssim \prod_{k=1}^3 \|\zeta^2 \tilde{\pi}_{j,k} f_k\|_{p_k} \lesssim |I| \prod_{k=1}^3 \mathsf{eng}_k(f_k)(\mathsf{T}; q_k).$$

The left hand side of (4.10) is then controlled by

$$\begin{aligned} &\sum_{j \in \mathbf{j}_\mathsf{T}} \sum_{I \in \mathcal{D}_j} \int_I \left| \tilde{\chi}_j - (\tilde{\chi}_j)^3 \right| \prod_{k=1}^3 |\tilde{\pi}_{j,k} f_k| \\ &\lesssim \prod_{k=1}^3 \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \times \left(\sum_{j \in \mathbf{j}_\mathsf{T}} \sum_{I \in \mathcal{D}_j} |I| \left\| \left| 1 - (\tilde{\chi}_j)^2 \right| (\tilde{\chi}_j)^{\frac{1}{2}} \right\|_{L^\infty(I)} \right). \end{aligned} \tag{4.13}$$

Here we use (2.3) and the fact that $E_{Q^j,\mathsf{T}}$ is a union of disjoint intervals of $\mathcal{D}_j$ to obtain the estimate

$$\left| \tilde{\chi}_j - \mathbf{1}_{E_{Q^j,\mathsf{T}}} \right| \lesssim \left\langle \frac{\operatorname{dist}(x, \partial E_{Q^j,\mathsf{T}})}{2^{-Jj}} \right\rangle^{-N+1},$$

leading to the following bound [3] for the bracketed term in (4.13):

$$\lesssim \sum_{j \in \mathbf{j}_\mathsf{T}} \sum_{I \in \mathcal{D}_j} \int_I \left\langle \frac{\operatorname{dist}(x, \partial E_{Q^j,\mathsf{T}})}{2^{-Jj}} \right\rangle^{-\frac{N}{2}} \mathrm{d}x \lesssim \sum_{j \in \mathbf{j}_\mathsf{T}} 2^{-Jj} \# \partial E_{Q^j,\mathsf{T}} \lesssim |I_\mathsf{T}| = 1, \tag{4.14}$$

having used (A.6) in the last step. Combining the last display with (4.13) yields exactly (4.10). □

Proof of Lemma 4.12. Recall that $m_{j,1}$, the symbol of $\tilde{\pi}_{j,1}$, is adapted and supported in $(Q^j)_1$, which is a dyadic interval of length 2^{Jj} and such that $2^{\frac{J}{2}}(Q^j)_1$ contains the origin. Thus $m_{j,1}$

---

[2] Pick $\overline{P} \in \mathsf{T}(j)$ such that $\operatorname{dist}(I, I_P)$ is minimal.

[3] Here, the estimate is obtained by writing $|\tilde{\chi}_j - (\tilde{\chi}_j)^3| (\tilde{\chi}_j)^{-\frac{1}{2}} = |1 - \tilde{\chi}_j||1 + \tilde{\chi}_j|(\tilde{\chi}_j)^{\frac{1}{2}} \lesssim |1 - \tilde{\chi}_j|(\tilde{\chi}_j)^{\frac{1}{2}}$ and observing that when $x \in E_{Q^j,\mathsf{T}}$ we have $|1 - \tilde{\chi}_j|(\tilde{\chi}_j)^{\frac{1}{2}} \lesssim |\tilde{\chi}_j - \mathbf{1}_{E_{Q^j,\mathsf{T}}}| \lesssim \left\langle 2^{Jj} \operatorname{dist}(x, \partial E_{Q^j,\mathsf{T}}) \right\rangle^{-N+1}$, while when $x \in \mathbb{R} \setminus E_{Q^j,\mathsf{T}}$, $|1 - \tilde{\chi}_j|(\tilde{\chi}_j)^{\frac{1}{2}} \lesssim (\chi_j)^{\frac{1}{2}} \lesssim \left\langle 2^{Jj} \operatorname{dist}(x, \partial E_{Q^j,\mathsf{T}}) \right\rangle^{-N/2}$.

vanishes outside $|\xi| \leq 2^{J(j+\frac{1}{2})}$. The symbol Ψ_j of S_j is supported on $2^{J(j+1)-1} \leq |\xi| \leq 2^{J(j+2)}$. Let $g_k \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ be scalar functions. Then Plancherel's equality yields

$$\int \sum_{j \in \mathsf{j}_\mathsf{T}} \varepsilon_j (\tilde{\pi}_{j,1} g_1)(S_j g_2)(S_j g_3) = \langle O(g_1, g_2), \overline{g_3} \rangle, \tag{4.15}$$

where O is the bilinear Fourier multiplier operator

$$O(g_1, g_2)(x) = \int_{\mathbb{R}\times\mathbb{R}} \widehat{g_1}(\xi_1)\widehat{g_2}(\xi_2) m(\xi_1, \xi_2) e^{2\pi i x(\xi_1+\xi_2)} \mathrm{d}\xi_1 \mathrm{d}\xi_2, \qquad x \in \mathbb{R},$$

$$m(\xi_1, \xi_2) := \sum_{j \in \mathsf{j}_\mathsf{T}} \varepsilon_j m_{j,1}(\xi_1) \Psi_j(\xi_2) \Psi_j(-\xi_1 - \xi_2).$$

The support and smoothness conditions on $m_{j,1}, \Psi_j$ imply that m satisfies the Coifman-Meyer condition and thus O is a bilinear CZ operator. We may then use [11, Theorem 1.1] to conclude that O extends to a bounded bilinear operator

$$L^{p_1}(\mathbb{R}; \mathcal{X}_1) \times L^{p_2}(\mathbb{R}; \mathcal{X}_2) \to L^{p_3'}(\mathbb{R}; \mathcal{X}_3').$$

As (4.15) continues to hold for $g_k \in (L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})) \otimes \mathcal{X}_k$, the vector-valued bound of the above display and duality complete the proof of the lemma. □

4.2. **Proof of Lemma 3.16, estimate** (4.11). We keep using the local notation $\mathcal{D}_j$ for the collection of all J-dyadic intervals of length 2^{-Jj}. By the condition (3.17), we may find a Hölder tuple p_1, p_2, p_3 with $q_k \leq p_k < \infty$. To apply Proposition 4.3 it is useful to keep in mind the equalities

$$\int_{\mathbb{R}} \prod_{k=1}^{3} \tilde{\chi}_j \tilde{\pi}_{j,k} f_k = \int_{\mathbb{R}} \prod_{k=1}^{3} \left(\sum_{P \in \mathsf{T}(j)} S_{P_k} f_k \right), \qquad j \in \mathsf{j}_\mathsf{T}$$

having called $S_{P_k} \in \mathbb{S}_{P_k}$ the operator $f \mapsto \chi_{I_P} T_{m_{j,k}} f_k$. Recall that A stands for the lacunary components and B the non-lacunary components.

We first take care of the case where $\mathsf{T}_{\mathsf{in},k} \neq \emptyset$ for some $k \in A$. To do so, it is convenient to introduce the polynomial cutoff

$$\gamma_\mathsf{T}(x) = \left\langle \frac{x - c(I_\mathsf{T})}{\ell(I_\mathsf{T})} \right\rangle^{100}, \qquad x \in \mathbb{R}.$$

The key idea is that $u_P := c\gamma_\mathsf{T} \chi_{I_P} \in cX_{I_P}$ for all $P \in \mathsf{T}$. We then estimate the contribution of the $j = j_{\mathsf{in},k}$ scale as follows: choosing $\psi_I \in cX_I$ with $(\psi_I)^3 \geq 1$ on I for each $I \in \mathcal{D}_j$,

$$\begin{aligned} \left| \int_{\mathbb{R}} \prod_{k=1}^{3} \tilde{\chi}_j \tilde{\pi}_{j,k} f_k \right| &\lesssim \sum_{I \in \mathcal{D}_j} \left| \int_I \gamma_\mathsf{T}^{-3} \prod_{k=1}^{3} \left(\psi_I \sum_{P \in \mathsf{T}(j)} u_P T_{m_{j,k}} f_k \right) \right| \\ &\leq \sum_{I \in \mathcal{D}_j} \|\gamma_\mathsf{T}^{-3}\|_{L^\infty(I)} \prod_{k=1}^{3} \left\| \psi_I \sum_{P \in \mathsf{T}(j)} u_P T_{m_{j,k}} f_k \right\|_{L^{p_k}(\mathbb{R}; \mathcal{X}_k)} \\ &\lesssim \sum_{I \in \mathcal{D}_j} \left\langle \frac{\mathrm{dist}(I, I_\mathsf{T})}{\ell(I_T)} \right\rangle^{-80} |I| \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \lesssim |I_\mathsf{T}| \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \end{aligned}$$

provided we have chosen a Hölder tuple p_k with $q_k \le p_k < \infty$, and having used Lemma 4.7 in the passage to the last line.

Replacing T by $\mathsf{T} \setminus (\bigcup_{k\in A} \mathsf{T}_{\mathsf{in},k})$ we may now assume that $\mathsf{T}_{\mathsf{in},k} = \varnothing$ for all $k \in A$, so that $j_{\mathsf{in},k} \notin \mathbf{j}_\mathsf{T}$ for all $k \in A$. We handle the easy case where $B = \varnothing$. Applying part b. of Proposition 4.3 to each f_k, for each $j \in \mathbf{j}_\mathsf{T}$, we have

$$\sum_{j\in\mathbf{j}_\mathsf{T}} \varepsilon_j \int_\mathbb{R} \prod_{k=1}^3 \tilde{\chi}_j \tilde{\pi}_{j,k} f_k = \sum_{j\in\mathbf{j}_\mathsf{T}} \varepsilon_j \int_\mathbb{R} \prod_{k=1}^3 S_j(\Pi_k f_k).$$

For $k = 1, 2, 3$ let $\sigma_k = \{\sigma_{j,k} : j \in \mathbf{j}_\mathsf{T}\}$ be a sequence of i.i.d. random variables which take the values $1, -1$ with equal probability. We denote the expectation with respect to σ_k by $\mathbb{E}^k$. Using [11, Lemma 4.1], Hölder's inequality, L^{p_k}-bounds for the $\mathcal{X}_k$-valued randomized square function (as $\mathcal{X}_k$ is UMD, the space $\mathrm{Rad}(\mathcal{X}_k)$ is also UMD and the randomized square function is a $\mathrm{Rad}(\mathcal{X}_k)$-valued CZ operator, [23]), and subsequently part a. of Proposition 4.3, there holds

$$\begin{aligned} &\left| \sum_{j\in\mathbf{j}_\mathsf{T}} \varepsilon_j \int_\mathbb{R} \prod_{k=1}^3 S_j(\Pi_k f_k) \right| \\ &\lesssim \prod_{k=1}^3 \left(\mathbb{E}^k \int_\mathbb{R} \left| \sum_{j\in\mathbf{j}_\mathsf{T}} \sigma_{j,k} S_j(\Pi_k f_k) \right|_{\mathcal{X}_k}^{p_k} \right)^{\frac{1}{p_k}} \lesssim \prod_{k=1}^3 \|\Pi_k f_k\|_{L^{p_k}(\mathbb{R};\mathcal{X}_k)} \lesssim \prod_{k=1}^3 \mathsf{eng}_k(f_k)(\mathsf{T}; q_k), \end{aligned}$$

which is the claim (4.11).

We turn to the harder case where $\#A = 2$. By symmetry we may work with $B = \{1\}$. We use Proposition 4.3 to bound the left hand side of (4.11) by $\mathsf{MAIN} + \mathsf{ERR}_1 + \mathsf{ERR}_2$ where

$$\begin{aligned} \mathsf{MAIN} &:= \left| \sum_{j\in\mathbf{j}_\mathsf{T}} \varepsilon_j \int_\mathbb{R} \tilde{\pi}_{j,1}(\Pi_1 f_1) \prod_{k=2}^3 \tilde{\chi}_j \tilde{\pi}_{j,k} f_k \right| \\ \mathsf{ERR}_1 &:= \sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathcal{D}_j} \int_I |\tilde{\chi}_j \tilde{\pi}_{j,1}(f_1 - \Pi_1 f_1)|_{\mathcal{X}_1} \prod_{k=2}^3 \left|\tilde{\chi}_j \tilde{\pi}_{j,k} f_k\right|_{\mathcal{X}_k}, \\ \mathsf{ERR}_2 &:= \sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathcal{D}_j} \int_I \left[\left|1 - \tilde{\chi}_j\right| (\tilde{\chi}_j)^{\frac{1}{3}} \right] |\tilde{\chi}_j \tilde{\pi}_{j,1}(\Pi_1 f_1)|_{\mathcal{X}_1} \prod_{k=2}^3 \left| (\tilde{\chi}_j)^{\frac{1}{3}} \tilde{\pi}_{j,k} f_k \right|_{\mathcal{X}_k}, \end{aligned}$$

the second and third of which are error terms.

We first handle the error terms: via Hölder's inequality with exponents p_k, and a combination of (4.5) for the $\mathcal{X}_1$ with Lemma 4.7 for the $\mathcal{X}_k$ factors, $k = 2, 3$, we achieve the estimates

$$\mathsf{ERR}_1 \lesssim \left(\prod_{k=1}^3 \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \right) \sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathcal{D}_j} \int \chi_I \mu_j \lesssim \left(\prod_{k=1}^3 \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \right) \sum_{j\in\mathbf{j}_\mathsf{T}} \int \mu_j.$$

Then expanding out μ_j (for the definition of μ_j, see (A.11)), we have

$$\sum_{j\in\mathbf{j}_\mathsf{T}} \int \mu_j \lesssim \sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{y\in\partial\tilde{E}_j} \int (1 + 2^{Jj}|x-y|)^{-100} dx \lesssim \sum_{j\in\mathbf{j}_\mathsf{T}} 2^{-Jj} \#\partial\tilde{E}_j \lesssim |I_\mathsf{T}| = 1, \tag{4.16}$$

where we have used (A.10) in the last inequality. This shows that ERR_1 complies with the right hand side of (4.11). The second error term is bounded proceeding as in the proof of (4.10): namely, splitting the integral over with $I \in \mathcal{D}_j$ and applying Hölder's inequality (4.10) followed by the single scale estimates with $I_0 = I$ (4.6) for $\tilde{\chi}_j \tilde{\pi}_{j,1}(\Pi_1 f_1)$ and Lemma 4.7 for $(\tilde{\chi}_j)^{\frac{1}{3}} \tilde{\pi}_{j,k} f_k$, $k = 1, 2$. The resulting estimate is

$$\mathsf{ERR}_2 \lesssim \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k) \times \left(\sum_{j \in \mathbf{j}_\mathsf{T}} \sum_{I \in \mathcal{D}_j} |I| \left\| \left|1 - \tilde{\chi}_j\right| (\tilde{\chi}_j)^{\frac{1}{3}} \right\|_{L^\infty(I)} \right) \lesssim \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k),$$

where the bracketed term has been bounded with the same procedure leading to (4.14) above.

We move to the main term. Using part b. of the Proposition, we recognize that

$$\mathsf{MAIN} = \left| \sum_{j \in \mathbf{j}_\mathsf{T}} \varepsilon_j \int (\tilde{\pi}_{j,1} \Pi_1 f_1)(S_j \Pi_2)(S_j \Pi_3) \right| \lesssim \prod_{j=1}^{3} \|\Pi_k f_k\|_{L^{p_k}(\mathbb{R}; \mathcal{X}_k)} \lesssim \prod_{k=1}^{3} \mathsf{eng}_k(f_k)(\mathsf{T}; q_k)$$

having used Lemma 4.12 for the first bound, and (4.4) for the second. This completes the proof of Lemma 3.16.

5. Proof of Proposition 4.3

In all cases below, the index $k \in \{1, 2, 3\}$ is fixed and we avoid mentioning it whenever possible. For instance, we write q for q_k, $\mathcal{X}$ for $\mathcal{X}_k$ and α for α_k, where $\alpha_k = 1 - \frac{1}{q_k}$. In accordance with this policy, we will use the notation $\omega(j) := (Q^j)_k$, for $j \in \mathbf{j}_\mathsf{T}$ where Q^j is the (at most) unique element of $\mathbf{Q}_\mathsf{T}$ with sidelength 2^{Jj}. We have also fixed the collection $\{S_{P_k} \in \mathbb{S}_{P_k}, P \in \mathsf{T}\}$ and will write below $S_{P_k} f = \zeta_P T_{m_P} f$ for a fixed choice of $\zeta_P \in \Psi_{I_P}$ and $m_P \in M_{\omega(j)}$, when $P \in \mathsf{T}(j)$.

5.1. Proof of Proposition 4.3, a. and b. parts: lacunary case.

Referring to Remark 3.3, a consequence of k-lacunarity of T is that $0 = (\xi_\mathsf{T})_k \in \overline{\omega(j)} \setminus 2\omega(j)$ whenever $j \neq j_{\mathsf{in},k}$. Thus, the Fourier transforms of the functions

$$\Pi_{k,j} f = \sum_{P \in \mathsf{T}(j)} S_{P_k} f \qquad j \neq j_{\mathsf{in},k} \tag{5.1}$$

are supported in the disjoint intervals $\{\xi : 2^{J(j+1)-1} \leq |\xi| \leq 2^{J(j+2)}\}$ where the symbol of S_j is constant equal to one, cf. (4.2): this is because of the Fourier support of $\zeta_P \in \Psi_{I_P}$ and $m_P \in M_{\omega(j)}$ for $P \in \mathsf{T}(j)$. In the lacunary case, the definition of Π_k is then very simple, namely referring to (5.1)

$$\Pi_k := \sum_{j \in \mathbf{j}_\mathsf{T} \setminus \{j_{\mathsf{in},k}\}} \Pi_{k,j}$$

and the equality in b. is immediate from the above considerations , while the estimate in a. for $p = q_k$ is immediate from the definition, as Π_k is itself a tree operator. We now prove the estimate

$$\|\Pi_k f\|_{\mathrm{BMO}(\mathbb{R}; \mathcal{X}_k)} \lesssim \mathsf{eng}_k(f)(\mathsf{T}; q_k)$$

and a. for the other values of p will follow by interpolation.

Fix a J-dyadic interval I. We first bound the contribution of the large scales: set $\mathsf{T}^+ = \{P \in \mathsf{T} : \ell(I_P) > \ell(I)\}$. Then if $P \in \mathsf{T}^+$ with $\ell(I_P) = 2^v \ell(I)$ and $\mathrm{dist}(I_P, I) \sim 2^n \ell(I_P)$, $v, n \in \mathbb{N}$, the Poincaré inequality yields

$$\begin{aligned}\mathrm{osc}_I\left(S_{P_k} f\right) &:= \frac{1}{|I|}\left\|S_{P_k} f - \fint_I (S_{P_k} f)\right\|_{L^1(I;\mathcal{X})} \\ &\lesssim \frac{1}{|I|^{\frac{1}{q}}}\left\|\ell(I)\nabla S_{P_k} f\right\|_{L^q(I;\mathcal{X})} \le 2^{-\alpha v}\frac{1}{|I_P|^{\frac{1}{q}}}\left\|\ell(I_P)\nabla(S_{P_k} f)\right\|_{L^q(I;\mathcal{X})}.\end{aligned}$$

We write $\widetilde{m_P}(\xi) := \ell(I_P)\xi m_P(\xi)$, so that

$$\begin{aligned}|\ell(I_P)\nabla(S_{P_k} f)|_{\mathcal{X}} = |\zeta_P T_{\widetilde{m_P}} f + (\ell(I_P)\nabla\zeta_P) T_{m_P} f|_{\mathcal{X}} &\le \widetilde{\chi_{I_P}}\left(\left|\widetilde{\chi_{I_P}} T_{\widetilde{m_P}} f\right|_{\mathcal{X}} + \left|\widetilde{\chi_{I_P}} T_{m_P} f\right|_{\mathcal{X}}\right) \\ &:= \widetilde{\chi_{I_P}}\left(|S_{P_k,1} f|_{\mathcal{X}} + |S_{P_k,2} f|_{\mathcal{X}}\right)\end{aligned}$$

for a suitable choice of $\widetilde{\chi_{I_P}} \in X_{I_P}$ so that the domination of the last display holds. Observe that with this choice $S_{P_k,u}, u = 1, 2$, belong to the class $\mathbb{S}_{P_k}$ and are thus single scale tree operators, whence

$$\frac{1}{|I_P|^{\frac{1}{q}}}\left\|S_{P_k,u}\right\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim \mathsf{eng}_k(f)(\mathsf{T}; q).$$

Using the bounds $\|\widetilde{\chi_{I_P}}\|_{L^\infty(I)} \lesssim 2^{-100n}$, we have proved that

$$\mathrm{osc}_I\left(S_{P_k} f\right) \lesssim 2^{-\alpha v - 100n}\mathsf{eng}_k(f)(\mathsf{T}; q).$$

which is summable over $P \in \mathsf{T}^+$, that is over $v, n \in \mathbb{N}$ as claimed.

We move to handling the small scales, that is $\mathsf{T}^- = \{P \in \mathsf{T} : \ell(I_P) \le \ell(I)\}$. We may partition T^- as the union of $\mathsf{T}^{-,0} = \{P \in \mathsf{T}^- : I_P \subset 3I\}$ and $\mathsf{T}^{-,n} = \{P \in \mathsf{T}^- : I_P \subset (2^{n+1}+1)I \setminus (2^n+1)I\}$ for $n \ge 1$. We may choose $\widetilde{\chi_I} \in X_I$ so that the estimate

$$\mathrm{osc}_I\left(\sum_{P \in \mathsf{T}^-} S_{P_k} f\right) \le \frac{1}{|I|^{\frac{1}{q}}}\sum_{n \ge 0}\|g_n\|_{L^q(\mathbb{R};\mathcal{X})} \qquad g_n = \sum_{P \in \mathsf{T}^{-,n}} \widetilde{\chi_I} S_{P_k} f \tag{5.2}$$

holds. We now estimate each term appearing in the last summation over n. Fix $P \in \mathsf{T}^{-,n}$ for a moment and notice that $\mathrm{dist}(I, I_P) \sim 2^n \ell(I)$. Writing again $S_{P_k} f = \zeta_P T_{m_P} f$, define

$$\widetilde{\zeta_P} := 2^{100n}\widetilde{\chi_I}\zeta_P, \qquad \widetilde{S_{P_k}} f = \widetilde{\zeta_P} T_{m_{P_k}} f.$$

We claim that the function $\widetilde{\zeta_P}$ belongs to X_{I_P}. Indeed, the decay condition (2.2) for $\widetilde{\zeta_P}$ is easy to verify, with the additional 2^{100n} factor being allowed by virtue of the previously observed separation between I, I_P. The frequency support condition (2.1) for $\widetilde{\zeta_P}$ derives from the fact that the Fourier support of $\widetilde{\chi_I}$ has an equal or smaller scale than the Fourier support of ζ_P. Then we notice that $\mathsf{T}^{-,n}$ is a tree with top data $(I^n, 0) := ((2^{n+1}+1)I, 0)$ and contained in T, whence

$$\|g_n\|_{L^q(\mathbb{R};\mathcal{X})} \le 2^{-100n}\left\|\sum_{P \in \mathsf{T}^{-,n}}\widetilde{S_{P_k}} f\right\|_{L^q(\mathbb{R};\mathcal{X})} \le 2^{-100n}|I^n|^{\frac{1}{q}}\mathsf{eng}_k(f)(\mathsf{T}; q) \le 2^{-99n}|I|^{\frac{1}{q}}\mathsf{eng}_k(f)(\mathsf{T}; q),$$

where the second bound holds because the operator inside the norm is a tree operator. Summation of the above bounds over n yields the required control for the left hand side of (5.2). This estimate completes the proof of a. and b. parts of the Proposition.

5.2. **Proof of Proposition 4.3, a. and c. part: non-lacunary case.** We keep the convention of writing q for $q_{\mathcal{X}_k}$ and $\mathcal{X}$ for $\mathcal{X}_k$.

Let $j \in \mathbf{j}_{\mathsf{T}}$. In this proof, the sets $\tilde{E}_j$, the collections Ω_j, the intervals I_j^ℓ, I_j^r refer to Section A.5 of the Appendix, to which we send for a detailed definition. There is no loss in generality with assuming that $\inf \mathbf{j}_{\mathsf{T}} = 0$, this corresponds to the normalization $\ell(I_{\mathsf{T}}) = 1$. Define for $x \in \tilde{E}_0$, $j(x) = \max\{j \in \mathbf{j}_{\mathsf{T}} : x \in \tilde{E}_j\}$. The scale $2^{-j(x)J}$ is the smallest spatial scale relevant for x. It is logical to choose $2^{j(x)J}$ as the frequency scale for the cutoff at x, motivating the definition of

$$\widetilde{\Pi_k} f := \mathbf{1}_{\tilde{E}_0} T_{j(x)} f = \mathbf{1}_{\tilde{E}_0} T_0 f + \sum_{j \geq 1} \mathbf{1}_{\tilde{E}_j} S_j f, \tag{5.3}$$

where the nestedness of $\tilde{E}_j$ and telescoping have been used to get the second equality. The construction of the actual phase-space projection operator Π_k is made by suitably modifying $\widetilde{\Pi_k}$ and begins now.

Fix a scale $j \in \mathbf{j}_{\mathsf{T}}$ and a connected component $I = [x_I^\ell, x_I^r] \in \Omega_j$. The perturbation of $g_j = \mathbf{1}_I S_j f$ is made by adding and subtracting two auxiliary pieces at spatial scale 2^{-Jj} which kill the mean value of g_j: details follow.

Recall from [30, Lemma 4.12] that I_j^ℓ [resp. I_j^r] are intervals of length $2^{-2}\ell(I)$ whose right endpoint [resp. left endpoint] sits to the left of x_I^ℓ [resp. to the right of x_I^r] at a distance of $2^{-2}\ell(I)$. These intervals are well separated, see [30, Lemma 4.12] over $I \in \Omega_j, j \in \mathbf{j}_{\mathsf{T}}$. Introduce bump functions $\phi_{I,j}^\ell$ [resp. $\phi_{I,j}^r$] adapted to and supported on I_j^ℓ [resp. I_j^r] with normalization

$$\int \phi_{I,j}^\ell = \int \phi_{I,j}^r = 2^{-Jj}.$$

Decomposing $\mathbf{1}_I(x) = H_I^\ell(x) + H_I^r(x) := H(x - x_I^\ell) - H(x - x_I^r)$, where H stands for Heaviside function, we introduce the $\mathcal{X}$-valued coefficients

$$c_{I,j}^\star := 2^{Jj} \int H_I^\star S_j f, \qquad \star \in \{\ell, r\}.$$

5.4. **Lemma.** *Let $I \in \Omega_j$ and $\star \in \{\ell, r\}$. We have the estimate*

$$|c_{I,j}^\star|_{\mathcal{X}} \lesssim \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.5}$$

Proof. To fix ideas we work with $\star = \ell$. Before we start, we recall that $\tilde{E}_j$, and hence I, is a union of J-dyadic intervals of length 2^{-jJ}, see Lemma A.7. Therefore there exists a J-dyadic interval I' of length 2^{-jJ} whose left endpoint coincides with x_I^ℓ. Then Lemma A.8 yields the existence of $P \in \mathsf{T}$ with $I_P \subset 10I'$ and $0 = (\xi_{\mathsf{T}})_k \in 3\omega_{P_k}$; as $k \in B$, the latter fact is read from (3.6). In particular, this shows that there exists $P \in \mathsf{T}$ with $I_P \subset 10I'$. This last fact will be of use later.

We then prove that there exists $\zeta \in \Psi_{I'}$ such that

$$|c_{I,j}^\ell|_{\mathcal{X}} \lesssim \frac{1}{|I'|} \int |\zeta^2(x) S_j f(x)|_{\mathcal{X}} \, \mathrm{d}x. \tag{5.6}$$

Inequality (5.6) is proved in the same fashion as [30, eq. (65)]; we adjust the details to our setup. Recall that S_j is a smooth Littlewood-Paley projection with support specified by (4.2). Let ϕ be a bump function whose Fourier transform is bounded by 1, equals 1 on the support of S_j and vanishes outside of $2^{(j-1)J} \le |\xi| \le 2^{j(J+3)}$, and let Φ be its antiderivative. Then

$$c^{\ell}_{I,j} = 2^{Jj} \int (H^{\ell}_I * \phi) S_j f.$$

As $(H^{\ell}_I * \phi)(\cdot) = \pm\Phi(\cdot - x^{\ell}_I)$, and the latter function belongs to $\Psi_{I'}$, the desired estimate follows with $\zeta := c\Phi(\cdot - x^{\ell}_I)$.

We finally turn to the proof of (5.5), where the previously found $P \in \mathsf{T}$ with $I_P \subset 10I'$ will play a role. As both I_P, I' are J-dyadic, the case $\ell(I_P) > \ell(I')$ is forbidden. Thus we are in either of the cases below. Suppose first that $\ell(I_P) = \ell(I')$. Then the intervals I_P and I' are comparable, so that $\zeta \in X_{I_P}$, and $S_j \in M_{\omega_{P_k}}$ by construction. Therefore, ζS_j is a tree operator adapted to the k-lacunary tree $(P, I_P, c(Q_P))$ and

$$|c^{\ell}_{I,j}|_{\mathcal{X}} \le \frac{1}{|I'|} \left\|\zeta S_j f\right\|_{L^q(\mathcal{X})} \lesssim \mathsf{eng}_{\mathsf{lac},k}(f)(\{P\}; q) \le \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.7}$$

Suppose instead that $\ell(I_P) = 2^{-Jj_P} < 2^{-Jj}$. Then by separation of scales property iii., it must be $j_P > j + 10$. We already know that $0 \in 3\omega_{P_k}$, and the multiplier of S_j vanishes at 0 and is supported on the interval ω centered at 0 and of length $2^{J(j+5)}$. Lemma 3.11 then applies with I' in place of I, $\chi = \zeta$, $T_m = S_j$ and $\xi = \xi_{\mathsf{T}} = 0$ whenever $\chi \in X_I$, yielding

$$|c^{\ell}_{I,j}|_{\mathcal{X}} \le \frac{1}{|I'|} \left\|\zeta S_j f\right\|_{L^q(\mathcal{X})} \lesssim \mathsf{eng}_{\mathsf{ov},k}(f)(\{P\}; q) \le \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.8}$$

In both cases, we have reached (5.5). This completes the proof. □

With Lemma 5.4 in hand, we are able to define the phase-space projection operator: with reference to (5.3),

$$\Pi_k f := \widetilde{\Pi_k} f - \sum_{j \in \mathbf{j}_{\mathsf{T}}} \sum_{I \in \Omega_j} \sum_{\star \in \{\ell, r\}} c^{\star}_{I,j} \phi^{\star}_{I,j}. \tag{5.9}$$

5.2.1. *Proof of Proposition 4.3, part a. for $k \in B$.* It suffices by interpolation to prove estimate (4.4) for $p = q$ together with the endpoint

$$\|\Pi_k f\|_{L^\infty(\mathbb{R};\mathcal{X})} \lesssim \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.10}$$

Proof of (5.10). First of all, by virtue of the separation properties of the support of the $\phi^{\star}_{I,j}$ over $I \in \Omega_j$, $j \in \mathbf{j}_{\mathsf{T}}$ we have recalled earlier, and of the second bound in Lemma 5.4,

$$\left\| \sum_{j \in \mathbf{j}_{\mathsf{T}}} \sum_{I \in \Omega_j} \sum_{\star \in \{\ell, r\}} c^{\star}_{I,j} \phi^{\star}_{I,j} \right\|_{L^\infty(\mathbb{R};\mathcal{X})} \lesssim \mathsf{eng}_k(f)(\mathsf{T}; q).$$

Hence, it suffices to prove an L^∞ bound on $\widetilde{\Pi_k}$. Fix $x \in \tilde{E}_0$ and set $j = j(x)$. By construction of $j(x)$, there is an interval $I' \subset \tilde{E}_j$ of length 2^{-Jj} containing x, and by construction of $\mathbf{I}_{\mathsf{T}}$ there is a tile $P \in \mathsf{T}$ with $I_P \subset 10I'$, see [30, Lemma 4.11]. For a suitable choice of $\zeta_{I'} \in X_{I'}$, we then have

$$|\widetilde{\Pi_k} f(x)|_{\mathcal{X}} \lesssim |\zeta_{I'} T_j f(x)|_{\mathcal{X}} \lesssim |I'|^{-\frac{1}{q}} \|\zeta_{I'} T_j f\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim \mathsf{eng}_k(f)(\{P\}; q) \le \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.11}$$

We have used Lemma 2.6 in the second inequality and argued exactly like in (5.7) if $\ell(I_P) = \ell(I)$. However, in the case $\ell(I_P) < \ell(I)$, in the appeal to Lemma 3.11 we must be a bit more careful and take $\xi \in \Gamma'$ such that $\xi_k = \pm 2^{J(j+2)}$ instead of $\xi = 0$, due to T_j being in general not vanishing at 0. This is no harm because $\ell(\omega_{P_k}) \geq 2^{(j+10)J}$, whence in both cases $\xi_k \in 3\omega_{P_k}$. This completes the proof of (5.10). □

Proof of (4.4) *for* $p = q$. First of all, using the disjointness of $I \in \Omega_j, j \in \mathbf{j}_\mathsf{T}$ we estimate the $L^q(\mathbb{R};\mathcal{X})$-norm of the part involving the $\phi^\star_{I,j}$ by

$$\left(\sum_{j\in\mathbf{j}_\mathsf{T}} \#\Omega_j 2^{-Jj}\right)^{\frac{1}{q}} \left(\sup_{\star\in\{\ell,r\}} \sup_{\substack{j\in\mathbf{j}_\mathsf{T}\\ I\in\Omega_j}} |c^\star_{I,j}|_\mathcal{X}\right) \lesssim |I_\mathsf{T}|^{\frac{1}{q}} \mathsf{eng}_k(f)(\mathsf{T};q),$$

where the first factor is bounded directly by [30, Lemma 4.12] while the second is (5.5) from Lemma 5.4. We are then left with proving

$$\|\widetilde{\Pi_k} f\|^q_{L^q(\mathbb{R};\mathcal{X})} \lesssim |I_\mathsf{T}| \mathsf{eng}_k(T)(f;q)^q. \tag{5.12}$$

To prove (5.12) we recall that the sets $\tilde{E}_j$ are decreasing in j and each is a union of disjoint intervals $I \in \mathbf{I}_j$ with $\ell(I) = 2^{-Jj}$ [30, Lemma 4.10]. Thus, the sets $E_I = I \cap (\tilde{E}_j \setminus \tilde{E}_{j+1}), I \in \mathbf{I}_j$ are a disjoint cover of each $\tilde{E}_j \setminus \tilde{E}_{j+1}$, and the latter sets are also pairwise disjoint and cover the support of $\widetilde{\Pi_k} f$. Furthermore, we see from [30, Lemma 4.10] that for each $I \in \mathbf{I}_j$ we may find $I' \subset \tilde{E}_j \setminus \tilde{E}_{j+1}$ with $I' \subset I$ and $\ell(I') = 2^{-J}\ell(I)$, hence $|E_I| \geq 2^{-J}|I|$ and

$$\sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathbf{I}_j} |I| \lesssim \sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathbf{I}_j} |E_I| \lesssim |I_\mathsf{T}|. \tag{5.13}$$

As $\widetilde{\Pi_k} f(x) = T_j f$ for $x \in \tilde{E}_j \setminus \tilde{E}_{j+1}$, the left hand side of (5.12) is controlled by

$$\sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathbf{I}_j} \|\mathbf{1}_{E_I} T_j f\|^q_{L^q(\mathbb{R};\mathcal{X})} \lesssim \sum_{j\in\mathbf{j}_\mathsf{T}} \sum_{I\in\mathbf{I}_j} \|\zeta_I T_j f\|^q_{L^q(\mathbb{R};\mathcal{X})}, \qquad \zeta_I \in X_I.$$

By virtue of the last display and of (5.13), it suffices to show that

$$\|\zeta_I T_j f\|^q_{L^q(\mathbb{R};\mathcal{X})} \lesssim |I| \mathsf{eng}_k(f)(\mathsf{T};q)^q, \qquad \forall j \in \mathbf{j}_\mathsf{T}, I \in \mathbf{I}_j.$$

Fix such j, I. We now appeal to [30, Lemma 4.11] to find $P \in \mathsf{T}$ with $I_P \subset 10I$ and the last display follows by similar arguments as (5.11), completing the proof of (4.4). □

5.2.2. *Proof of Proposition 4.3, part c.* We begin the proof by using the single scale estimate of Lemma 4.7. In fact (4.6) follows immediately from (4.5), (4.8) and the fact that μ_j is uniformly bounded. So it remains to prove (4.5). As usual, we prove the extremal cases. In fact, it suffices to prove the case $p = q$, as the case $p = \infty$ may then be recovered from Lemma 2.6.

Proof of (4.5) *for* $p = q$. In the proofs that follow, we use the local notation

$$S_{P_k} g = \zeta_{I_P} T_{m_{P_k}} g, \qquad Og = \sum_{P\in\mathsf{T}(j_0)} S_{P_k} g,$$

where $\zeta_{I_P} \in X_{I_P} = X_{I_P}(2N, \delta, C)$ and $m_{P_k} \in M_{\omega_{P_k}}$. Notice that O is a tree operator and thus is bounded on $L^q(\mathbb{R};\mathcal{X}_k)$, but it is also pointwise bounded by maximal averages and thus bounded on $L^\infty(\mathbb{R};\mathcal{X}_k)$.

Recall that Π_k is defined in (5.9). The first step in the proof proper is to notice that

$$O(f - \Pi_k f) = O(T_{j_0} f - \Pi_k f),$$

leading to the key decomposition

$$T_{j_0}f - \Pi_k f =$$

$$\mathbf{1}_{\mathbb{R}\setminus\tilde{E}_{j_0}} T_{j_0} f \tag{5.14}$$

$$- \mathbf{1}_{\mathbb{R}\setminus\tilde{E}_{j_0}} \widetilde{\Pi_k} f \tag{5.15}$$

$$+ \sum_{\star\in\{\ell,r\}} \mathbf{1}_{\mathbb{R}\setminus\tilde{E}_{j_0}} \sum_{j\le j_0} \sum_{I\in\Omega_j} c^{\star}_{I,j}\phi^{\star}_{I,j} \tag{5.16}$$

$$- \sum_{\star\in\{\ell,r\}} \sum_{j>j_0} \sum_{I\in\Omega_j} (H^{\star}_I S_j f - c^{\star}_{I,j}\phi^{\star}_{I,j}). \tag{5.17}$$

cf. [30, eqs. (77)-(82)]. We now have to estimate the four contributions separately, and, as in [30], distinguish the local case $5I_0 \cap \tilde{E}_{j_0} \neq \varnothing$ from the complementary nonlocal case: for clarity, we first present the local case, and at the end of the proof we elaborate on the sketch provided in [30, p. 295] and unify the two cases: see Remark 5.24 below.

We first estimate the contribution of $g = (5.14) - (5.15) + (5.16)$. Using the decay at scale $\ell(I_0)$ of the kernel of O together with the L^∞ bounds (5.11), (5.10), (5.5)

$$\|\mathbf{1}_{I_0} O(\mathbf{1}_{\mathbb{R}\setminus 3I_0} g)\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \sum_{j\le j_0} 2^{-J|j-j_0|} \left\langle \frac{\mathrm{dist}(I_0, \partial\tilde{E}_j)}{\ell(I_0)} \right\rangle^{-100} \mathsf{eng}_k(f)(\mathsf{T}; q), \tag{5.18}$$

which is acceptable for (4.5). Further, if $\mathbf{1}_{3I_0} g$ is nonzero, then I_0 is close to the boundary of $\tilde{E}_{j_0}$. The integral term in the right hand side of (4.5) is $O(1)$ and we may just aim for the estimate

$$\|O(\mathbf{1}_{3I_0} g)\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \|O\|_{L^q(\mathbb{R};\mathcal{X}_k)} \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.19}$$

Although the O-norm appearing here is $O(1)$, we choose to keep this constant in evidence for later use.

We begin the proof of (5.19). We argue separately for each summand of g. First of all, we bound the contribution of (5.14). Appealing to [30, Lemma 4.11], we learn that there exists $P \in \mathsf{T}$ such that $I_P \subset 10I_0$. Hence, for suitable choice of $\zeta_{I_0} \in X_{I_0}$, arguing via Lemma 3.11 as in the proof of (5.11),

$$\|\mathbf{1}_{3I_0} T_{j_0} f\|_{L^q(\mathbb{R};\mathcal{X}_k)} \le \|\zeta_{I_0} T_{j_0} f\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \mathsf{eng}_k(f)(\{P\}, q) \le |I_0|^{\frac{1}{q}} \mathsf{eng}_k(f)(\mathsf{T}; q). \tag{5.20}$$

This makes the contribution of (5.14) acceptable for (5.19). To control the contribution of (5.15) we note that $(\mathbb{R} \setminus \tilde{E}_{j_0}) \cap 3I_0$ is the union of at most three intervals I_1 of length $\ell(I_0)$, on which $\widetilde{\Pi_k} f$ coincides with $T_{j_0-1} f$. On each of these intervals, by the same argument used for (5.20),

$$\|\mathbf{1}_{I_1} T_{j_0-1} f\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \mathsf{eng}_k(f)(\mathsf{T}; q) \tag{5.21}$$

which is acceptable. Finally, from the last claim of [30, Lemma 4.12] we gather that $I^{\star}_j \cap 3_{I_0} \neq \varnothing$ for at most $O(1)$ intervals $I \in \Omega_j$ with $j \le j_0$. Therefore

$$\|\mathbf{1}_{3I_0}(5.16)\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \sup_{I,j,\star} |c_{I,j}| \lesssim |I_0|^{\frac{1}{q}} \mathsf{eng}_k(f)(\mathsf{T}; q)$$

by (5.5), and we have proved (5.19). This finishes the control of terms (5.14) to (5.16).

To complete the proof of (4.5), we are left with estimating the small spatial scales term (5.17). Using the triangle inequality and the definition of μ_{j_0}, it will suffice to prove that for each fixed $\star \in \{\ell, r\}$, $j > j_0, I \in \Omega_j$ there holds

$$\|\mathbf{1}_{I_0} O(G_I)\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \mathsf{eng}_k(f)(\mathsf{T}; q) \int \frac{\zeta_{I_0}(x)}{|I_0|} 2^{-\frac{(j-j_0)}{100}} \langle 2^{Jj}|x - x_I^\star|\rangle^{-100} \, \mathrm{d}x,$$
$$G_I := H_I^\star S_j f - c_{I,j}^\star \phi_{I,j}^\star.$$

As they will be kept fixed below, we have omitted $\star$ and j from the G_I notation for simplicity. Let $n \in \mathbb{N}$ be the least integer such that $2^n I_0 \cap I_j^\star \neq \varnothing$. A direct computation of the right hand side and the fact that $\chi_{I_0} \in X_{I_0}$ tells us that the above bound is equivalent to the estimate

$$\|\mathbf{1}_{I_0} O(G_I)\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim |I_0|^{\frac{1}{q}} \mathsf{eng}_k(f)(\mathsf{T}; q) 2^{-\frac{(j-j_0)}{100}} 2^{-Jj} 2^{-100n}. \tag{5.22}$$

The final stretch of the proof will be to establish (5.22). As the frequency support of O is localized near 2^{j_0}, we gather that $O[(T_{j-1}H_I^\star)(S_j f)] = 0$. This means we may replace G_I by

$$F_I = G_I - (T_{j-1}H_I^\star)(S_j f) = [(1 - T_{j-1})H_I^\star] S_j f - c_{I,j}^\star \phi_{I,j}^\star.$$

As both G_I and $F_I - G_I$ have mean zero, F_I also does. Letting Φ_I be the antiderivative of F_I which vanishes at $\pm\infty$, we have

$$OF_I(x) = 2^{Jj_0} \widetilde{O} \Phi_I(x),$$

where

$$\widetilde{O} g(x) = \sum_{P \in \mathsf{T}(j_0)} \int \zeta_{I_P}(x) 2^{-Jj_0} \xi m_{P_k}(\xi) \widehat{g}(\xi) \mathrm{e}^{ix\xi} \, \mathrm{d}\xi.$$

Note that $\xi \mapsto 2^{-Jj_0} \xi m_{P_k}(\xi)$ belongs to $M_{\omega_{P_k}}$ and let u_P be the Fourier transform of the latter function. Let g be a scalar function and $\tilde{g}_N = \langle 2^{j_0 J}(\cdot - c(I_0))\rangle^{-N} g$. If $P \in \mathsf{T}(j_0)$, $|c(I_0) - c(I_P)| \gtrsim n\ell(I_0)$ and $x \in I_0$ we have

$$|\zeta_{I_P}(x)| \left[|g| * |u_P|\right](x) \leq \frac{1}{\langle n \rangle^{10}} \int |u_P(y)| \left\langle 2^{j_0 J}(x - c(I_0) - y)\right\rangle^N |\tilde{g}_N(x - y)| \, \mathrm{d}y \lesssim \frac{1}{\langle n \rangle^{10}} \mathrm{M}(\tilde{g}_N)(x)$$

by virtue of the rapid decay of u_P at scale $2^{-j_0 J}$. Summing up over $P \in \mathsf{T}(j_0)$, and thus over $n \in \mathbb{N}$, we may thus estimate (5.22) by

$$\|\mathbf{1}_{I_0} O(G_I)\|_{L^q(\mathbb{R};\mathcal{X}_k)} \lesssim \|\zeta_{I_0} \Phi_I\|_{L^q(\mathbb{R};\mathcal{X}_k)}.$$

for a suitable choice of $\zeta_{I_0} \in X_{I_0}$. An estimate on $|\Phi_I(x)|_{\mathcal{X}}$ compatible with the right hand side of (5.22) may be produced, cf. [30, p. 298], once we establish the pointwise bound

$$|F_I(x)|_{\mathcal{X}} \lesssim \mathsf{eng}_k(f)(\mathsf{T}; q) \langle 2^{Jj}|x - x_I^\star|\rangle^{-100}. \tag{5.23}$$

The last step towards (5.22), and therefore (4.5), is to prove (5.23). The contribution of $c_{I,j}^\star \phi_{I,j}^\star$ is controlled by virtue of the decay of $\phi_{I,j}^\star$ and (5.5). We turn to controlling the summand $[(1 - T_{j-1})H_I^\star] S_j f$. First we recall that by construction of $I \in \Omega_j$, $I_j^\star$ and by [30, Lemma 4.11], we may find a dyadic interval I' with $\ell(I') = 2^{-Jj}$ adjacent to one of the endpoints of I, and $P \in \mathsf{T}$ such that $I_P \subset 10I'$. Pick $\chi \in X_{I'}$ with decay parameter N (for instance). If $\ell(I_P) = \ell(I)$, we argue as in (5.7). Otherwise, as in (5.8) we may appeal to Lemma 3.11 for $S = \chi S_j$, I' in place of I, $\xi = \xi_{\mathsf{T}} = 0$, so that

$$\|\chi S_j f\|_{L^\infty(\mathbb{R};X)} \leq \frac{1}{|I'|^{\frac{1}{q}}} \|\chi S_j f\|_{L^q(\mathbb{R};X)} \lesssim \mathsf{eng}_k(f)(\{P\};q) \leq \mathsf{eng}_k(f)(\mathsf{T};q).$$

As $\chi(x) \gtrsim \langle 2^{Jj}|x - x_{I_j^\star}|\rangle^{-N}$, we have

$$|S_j f(x)| \lesssim \langle 2^{Jj}|x - x_{I_j^\star}|\rangle^{N} \mathsf{eng}_k(f)(\mathsf{T};q).$$

Integrating repeatedly by parts the high frequency function $[(1-T_{j-1})H_I^\star]$, we may bound it pointwise by factors of $\lesssim_N \langle 2^{Jj}|x - x_I^\star|\rangle^{-N-100}$, compensating the polynomial growth of the last display and yielding an acceptable right hand side for (5.23) which is finally proved. The proof of (4.5) is finally complete. □

5.24. *Remark* (The nonlocal case of (4.5)). The local/nonlocal cases can be unified by introduction of the parameter

$$Z = \text{least nonnegative integer such that } I_0 \pm Z\ell(I_0) \cap \tilde{E}_{j_0} \neq \emptyset.$$

Comparing with what we did to obtain (5.20), and to [30, Lemma 4.11], we learn that there exists $P \in \mathsf{T}$ such that $I_P \subset 10 \cdot ZI_0$, whence $\{P\}$ is a k-overlapping tree with top data $(10 \cdot ZI_0, \xi_{\mathsf{T}})$. Applying Lemma 3.11 with this top data yields a Z^N loss in e. g. estimates (5.20), (5.21). However, as we are concerned with estimates for $1_{I_0} O(T_{j_0} f - \Pi_k f)$, we may replace O by the operator $g \mapsto \tilde{O} g = \chi_{I_0} O g$, where $\zeta_{I_0} \in X_{I_0}(2N, \delta, C)$ and $\zeta_{I_0} \geq \mathbf{1}_{I_0}$. The separation between $\tilde{E}_{j_0}$ and I_0 yields that

$$\|\tilde{O}\|_{L^q(\mathbb{R};X_k)} \lesssim Z^{-2N}, \tag{5.25}$$

and the same additional decay factor is gained in the kernel estimates for $\tilde{O}$. Replacing O by $\tilde{O}$ in (5.18), (5.19) and taking (5.25) into account offsets the loss introduced in (5.20), (5.21).

6. Proof of Lemma 3.13

The first two paragraphs of this section are devoted to certain almost orthogonality estimates in the Hilbert space case, respectively for k-lacunary and k-overlapping trees. These are then extended to q-intermediate UMD spaces by interpolation, along the lines of [18] in Subsection 6.3. The proof of Lemma 3.13 is given in the concluding subsection.

6.1. **The L^2-orthogonality estimates: k-lacunary trees.** We begin with a definition. We say that a family of trees $\mathsf{T} \in \mathbf{T}$ is lac;*k-strongly disjoint* with parameter $1 \leq \theta \lesssim 2^J$ if

i. each T is a k-lacunary tree;
ii. if $\mathsf{T}, \mathsf{T}' \in \mathbf{T}$, $\mathsf{T} \neq \mathsf{T}'$, then

$$P \in \mathsf{T},\, P' \in \mathsf{T}',\, \ell(\omega_P) \leq \ell(\omega_{P'}),\, 10\theta\omega_{P_k} \cap 10\theta\omega_{P'_k} \neq \emptyset \implies I_{P'} \cap I_{\mathsf{T}} = \emptyset.$$

The rationale behind this definition is that, if the consequence of the above implication failed, the tri-tile P' would qualify to be in a suitable completion of the tree T. In what follows, we work with the parameter $\theta = 1$, as the general case $1 \leq \theta \lesssim 2^J$ may be handled by finite splitting. Tree operators associated to families of lac;k-strongly disjoint trees give rise to an L^2 almost orthogonality estimate: this is well known, and extends to the case of Hilbert space valued functions, as detailed in the next lemma. This lemma is a transposition of [18, Proposition 6.1] to our context. It is convenient in what follows to introduce the single tile version of the energy parameters. To do so, for each interval I and for each tri-tile P we introduce the functions

$$u_I := \left\langle \frac{|x - c(I)|}{\ell(I)} \right\rangle^{-10}, \qquad u_P := u_{I_P} \tag{6.1}$$

and also define

$$\|f\|_{P,k,q} = \sup_{m_{P_k} \in M_{P_k}} |I_P|^{-\frac{1}{q}} \left\| u_P T_{m_{P_k}} f \right\|_{L^q(\mathbb{R};\mathcal{X})}. \tag{6.2}$$

For uniformity, we gave the definitions above for a generic $1 \le q \le \infty$. However, Lemma 2.6 shows the upper bound $\|f\|_{P,k,\infty} \lesssim \|f\|_{P,k,1}$, and it follows that $\|f\|_{P,k,p} \sim_{p,q} \|f\|_{P,k,q}$ for all $1 \le p, q \le \infty$. Below, we will only use the value $q = 2$ in (6.2). Note the trivial bounds

$$\|f\|_{P,k,q} \lesssim |I_P|^{-\frac{1}{q}} \|u_P\|_q \sup_{m_{P_k} \in M_{P_k}} \|T_{m_{P_k}} f\|_{L^\infty(\mathbb{R};\mathcal{X})} \lesssim \|f\|_{L^\infty(\mathbb{R};\mathcal{X})}, \tag{6.3}$$

$$\sup_{S_{P_k} \in \mathbb{S}_{P_k}} \|S_{P_k} f\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim |I_P|^{\frac{1}{q}} \|f\|_{P,k,q}. \tag{6.4}$$

6.5. **Lemma.** *Let $\mathcal{X}$ be a Hilbert space and $\mathbf{T}$ be a collection of* $\mathsf{lac};k$*-strongly disjoint trees, and define $\mathbb{T} = \bigcup\{\mathsf{T} : \mathsf{T} \in \mathbf{T}\}$. There holds*

$$\left\| \sqrt{|I_\mathsf{T}|} \mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T};2) \right\|_{\ell^2(\mathsf{T} \in \mathbf{T})} \lesssim \|f\|_{L^2(\mathbb{R};\mathcal{X})} + \left(\|f\|_{L^\infty(\mathbb{R};\mathcal{X})} \left[\sum_{\mathsf{T} \in \mathbf{T}} |I_\mathsf{T}| \right]^{\frac{1}{2}} \right)^{\frac{1}{3}} \|f\|_{L^2(\mathbb{R};\mathcal{X})}^{\frac{2}{3}}. \tag{6.6}$$

Before entering the proof, we detail the almost orthogonality of the single tile operators within a k-lacunary tree.

6.7. **Lemma.** *Let $\mathcal{X}$ be a Hilbert space and T be a k-lacunary tree. Then*

$$\sqrt{|I_\mathsf{T}|} \mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T};2) \lesssim \sqrt{\sum_{P \in \mathsf{T}} |I_P| \|f\|_{P,k,2}^2}.$$

Proof. By modulation invariance, it suffices to take care of the case $\xi_\mathsf{T} = 0$. Choose a tree operator $S_\mathsf{T} = \sum_{P \in \mathsf{T}} S_{P_k}$ that nearly achieves the supremum in $\mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T};2)$ and write $S_{P_k} g = \zeta_P T_{m_{P_k}} g$. From the disjointness of the frequency supports, we have that, referring to (4.1)

$$\langle S_{P_k} f, S_{P'_k} f \rangle \neq 0 \implies P, P' \in \mathsf{T}(j).$$

For $n \in \mathbb{Z}$ denote by P^{+n} the unique (if it exists) tri-tile $P' \in \mathsf{T}(j)$ with $I_{P'} = I_P + n\ell(I_P)$. Then define

$$\widetilde{\zeta_P} := \frac{\zeta_P}{u_P}, \qquad \widetilde{S_{P_k}} g := \widetilde{\zeta_P} T_{m_{P_k}} g. \tag{6.8}$$

It is immediate to see that $\widetilde{\zeta_P} \in \Psi_{I_P}$ as multiplying by the correctly scaled polynomial $\frac{1}{u_P}$ does not change the frequency support neither significantly alters the rapid decay of ζ_P, hence $\tilde{S}_{P_k}$ belongs to $\mathbb{S}_{P_k}$. Therefore

$$\begin{aligned}|I_{\mathsf{T}}|\mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T};2)^2 &\lesssim \|S_{\mathsf{T}}f\|^2_{L^2(\mathbb{R};\mathcal{X})} \lesssim \sum_{j\in \mathsf{j}_{\mathsf{T}}}\sum_{P\in\mathsf{T}(j)}\sum_{n\in\mathbb{Z}}\int S_{P_k}f\overline{S_{P_k^{+n}}f} \\ &\leq \sum_{j\in \mathsf{j}_{\mathsf{T}}}\sum_{P\in\mathsf{T}(j)}\sum_{n\in\mathbb{Z}}\int |\widetilde{S_{P_k}}f||\widetilde{S_{P_k^{+n}}}f|u_P u_{P^{+n}} \\ &\lesssim \sum_{j\in \mathsf{j}_{\mathsf{T}}}\sum_{P\in\mathsf{T}(j)}\sum_{n\in\mathbb{Z}}\langle n\rangle^{-10}\left(\|\widetilde{S_{P_k}}f\|^2_{L^2(\mathbb{R};\mathcal{X})} + \|\widetilde{S_{P_k^{+n}}}f\|^2_{L^2(\mathbb{R};\mathcal{X})}\right) \\ &\lesssim \sum_{P\in\mathsf{T}}\|\widetilde{S_{P_k}}f\|^2_{L^2(\mathbb{R};\mathcal{X})} \lesssim \sum_{P\in\mathsf{T}}|I_P|\|f\|^2_{P,k,2}\end{aligned}$$

and this proves the claimed inequality. We have used $\|u_P u_{P^{+n}}\|_\infty \lesssim \langle n\rangle^{-10}$ to pass to the second line and (6.4) in the last estimate. □

Proof of Lemma 6.5. Let us choose the scaling $\|f\|_{L^2(\mathbb{R};\mathcal{X})} = 1$. From Lemma 6.7, we may bound the quantity

$$S := \sqrt{\sum_{P\in\mathbb{T}}|I_P|\|f\|^2_{P,k,2}}$$

in place of the left hand side of (6.6). Then

$$S^2 \sim \sum_{P\in\mathbb{T}}\sqrt{|I_P|}\|f\|_{P,k,2}\|u_P T_{m_{P_k}}f\|_{L^2(\mathbb{R};\mathcal{X})}.$$

having linearized the suprema in $\|f\|_{P,k,2}$ with a suitable choice $m_{P_k} \in M_{P_k}$, $P \in \mathbb{T}$. From now on, as m_{P_k} and k are fixed, we simply write T_P in place of $T_{m_{P_k}}$. Defining the $\mathcal{X}$-valued function

$$v_P = \sqrt{|I_P|}(u_P)^2 T_P f$$

we have the identity

$$\sqrt{|I_P|}\|u_P T_P f\|_{L^2(\mathbb{R};\mathcal{X})} = \langle T_P^* v_P, f\rangle, \qquad P\in\mathbb{T}$$

and the pointwise estimate

$$|v_P(x)|_{\mathcal{X}} \leq u_P(x)\sqrt{|I_P|}\|u_P T_P f\|_\infty \lesssim u_P(x) \tag{6.9}$$

coming from Lemma 2.6 with $R = (\ell(I_P))^{-1}, w = u_P, T_P f$ in place of f. These considerations lead to the estimate

$$S^2 \sim \left\langle \sum_{P\in\mathbb{T}}\|f\|_{P,k,2}T_P^* v_P, f\right\rangle \leq \left\|\sum_{P\in\mathbb{T}}\|f\|_{P,k,2}T_P^* v_P\right\|_{L^2(\mathbb{R};\mathcal{X})}. \tag{6.10}$$

Define now

$$\mathbb{T}_<(P) := \{P' \in \mathbb{T} : \omega_{P_k} \subsetneq \omega_{P'_k}\}, \qquad \mathbb{T}_=(P) := \{P' \in \mathbb{T} : \omega_{P'_k} = \omega_{P_k}\}.$$

Frequency support considerations applied to the inner products $\langle T_P^* v_P, T_{P'}^* v_{P'}\rangle$ then lead to the chain of inequalities

$$\begin{aligned}
&\Big\| \sum_{P\in\mathbb{T}} \|f\|_{P,k,2} T_P^* v_P \Big\|_{L^2(\mathbb{R};\mathcal{X})}^2 \\
&= \sum_{P\in\mathbb{T}} \sum_{P'\in\mathbb{T}_=(P)} \|f\|_{P,k,2}\|f\|_{P',k,2}\langle T_P^* v_P, T_{P'}^* v_{P'}\rangle \\
&+ 2\sum_{P\in\mathbb{T}} \sum_{P'\in\mathbb{T}_<(P)} \|f\|_{P,k,2}\|f\|_{P',k,2}\langle T_P^* v_P, T_{P'}^* v_{P'}\rangle := S_1 + 2S_2.
\end{aligned} \tag{6.11}$$

We first treat S_1. Note that if $P' \in \mathbb{P}_=(P)$ then $P' = P^{+n}$ for some $n \in \mathbb{Z}$, see the line before (6.8) for a definition. The decay of v_P (6.9) and the kernel estimate for T_P^* guarantee the pointwise bound

$$|T_P^* v_P|_{\mathcal{X}} \lesssim u_P \tag{6.12}$$

whence

$$|\langle T_P^* v_P, T_{P^{+n}}^* v_{P^{+n}}\rangle| \lesssim |I_P|\langle n\rangle^{-10}, \qquad n \in \mathbb{Z}.$$

Therefore, we control

$$S_1 \lesssim \sum_{P\in\mathbb{P}} \sum_{n\in\mathbb{Z}} \langle n\rangle^{-10}\left(|I_P|\|f\|_{P,k,2}^2 + |I_{P^{+n}}|\|f\|_{P^{+n},k,2}^2\right) \lesssim S^2 \tag{6.13}$$

using the definition of S. We turn to S_2. Notice that if $P' \in \mathbb{T}_<(P)$, then $\ell(I_{P'}) < \ell(I_P)$. Relying on (6.12) again

$$|\langle T_P^* v_P, T_{P'}^* v_{P'}\rangle| \lesssim |I_{P'}|\left\langle \frac{\operatorname{dist}(I_P, I_{P'})}{\ell(I_P)}\right\rangle^{-10} \lesssim \|\mathbf{1}_{I_{P'}} u_P\|_1, \qquad P' \in \mathbb{T}_<(P). \tag{6.14}$$

We claim that the intervals $\{I_{P'} : P' \in \mathbb{T}_<(P)\}$ are pairwise disjoint and do not intersect $I_{\mathsf{T}(P)}$ where $\mathsf{T}(P)$ is the unique tree in $\mathbb{T}$ where P belongs. The argument is standard, see e.g. [18, 30], but we reproduce it for completeness. Due to k-lacunarity of $\mathsf{T}(P)$, we have $\mathbb{T}_<(P) \cap \mathsf{T}(P) = \emptyset$. Therefore, the condition ii. in the definition of the strongly disjoint trees forces $I_{P'} \cap I_{\mathsf{T}(P)} = \emptyset$ for all $P' \in \mathbb{T}_<(P)$. Furthermore, if $P', P'' \in \mathbb{T}_<(P)$ it follows that $10\omega_{P'_k} \cap 10\omega_{P''_k} \neq \emptyset$. If P', P'' belong to distinct trees $\mathsf{T}', \mathsf{T}''$, then condition ii. forces $I_{P'} \cap I_{\mathsf{T}''} = \emptyset$. If P', P'' belong to the same tree T' then k-lacunarity forces $\omega_{P'_k} = \omega_{P''_k}$. In both cases $I_{P'} \cap I_{\mathsf{T}''} = \emptyset$. Using the bound (6.3), estimate (6.14), the trivial estimate $\|\mathbf{1}_{I_{P'}} u_P\|_1 \leq \|u_P\|_1 \lesssim |I_P|$, Cauchy-Schwarz, disjointness and separation from $I_{\mathsf{T}(P)}$ of $\{I_{P'} : P' \in \mathbb{T}_<(P)\}$, and we obtain

$$\begin{aligned}
S_2 &\lesssim \|f\|_\infty \sum_{P\in\mathbb{T}} \|f\|_{P,k,2} \sum_{P'\in\mathbb{T}_<(P)} \|\mathbf{1}_{I_{P'}} u_P\|_1 \leq \|f\|_\infty \sum_{P\in\mathbb{T}} \sqrt{|I_P|}\|f\|_{P,k,2}\|\mathbf{1}_{\mathbb{R}\setminus I_{\mathsf{T}}} u_P\|_1^{\frac{1}{2}} \\
&\lesssim \|f\|_\infty \Big(\sum_{P\in\mathbb{T}} |I_P|\|f\|_{P,k,2}^2\Big)^{\frac{1}{2}} \Big(\sum_{\mathsf{T}\in\mathbf{T}}\sum_{P\in\mathsf{T}} \|\mathbf{1}_{\mathbb{R}\setminus I_{\mathsf{T}}} u_P\|_1\Big)^{\frac{1}{2}} \lesssim \|f\|_\infty \Big(\sum_{\mathsf{T}\in\mathbf{T}} |I_{\mathsf{T}}|\Big)^{\frac{1}{2}} S;
\end{aligned} \tag{6.15}$$

we omitted some of the details, see e.g. [18, Proposition 6.1]. Summarizing (6.10), (6.11), (6.13), (6.15)

$$S^2 \lesssim \sqrt{S_1 + 2S_2} \lesssim \left(S^2 + \|f\|_\infty \left(\sum_{\mathsf{T}\in\mathbf{T}} |I_\mathsf{T}| \right)^{\frac{1}{2}} S \right)^{\frac{1}{2}}$$

which yields the claimed bound. The details can be read from [18, Proposition 6.1], hence we omit them. □

6.2. **The L^2-orthogonality estimates: k-overlapping trees.** We begin with some additional definitions related to the top data $(I_\mathsf{T}, \xi_\mathsf{T})$ of a k-overlapping tree T. Recall the notation (3.8) for $t_{\mathsf{T},k}$ and denote

$$\omega_{\mathsf{T},k,+} = \Big[(\xi_\mathsf{T})_k, (\xi_\mathsf{T})_k + 2^{-1}\ell(\omega_{\mathsf{T},k})\Big), \qquad \omega_{\mathsf{T},k,-} = \Big[(\xi_\mathsf{T})_k - 2^{-1}\ell(\omega_{\mathsf{T},k}), (\xi_\mathsf{T})_k\Big).$$

Consider a finitely overlapping cover of $\omega_{\mathsf{T},k,\pm}$ by intervals $\{\omega_{\mathsf{T},k,(\sigma,\tau),\pm} : \sigma \in \mathbb{N}, \tau = 1,\ldots,2^6\}$, with the properties that

$$\frac{\ell(\omega_{\mathsf{T},k,(\sigma,\tau),\pm})}{\ell(\omega_{\mathsf{T},k})} = 2^{-\sigma-5}, \qquad 2^{-\sigma} \le \frac{\operatorname{dist}\Big(\omega_{\mathsf{T},k,(\sigma,\tau),\pm}, (\xi_\mathsf{T})_\kappa\Big)}{\ell(\omega_{\mathsf{T},k})} \le 2^{1-\sigma} \qquad \forall \sigma \in \mathbb{N}, \tau = 1,\ldots,2^6.$$

The role of the parameter τ is to refine the Whitney decomposition of $\omega_{\mathsf{T},k,\pm}$ so that the 10-dilates of the Whitney intervals stay in the half-line $\pm\xi > \pm(\xi_\mathsf{T})_k$.

We say that a family of trees $\mathsf{T} \in \mathbf{T}$ is $\mathsf{ov};k$*-strongly disjoint* of type $\bullet \in \{+,-\}$ if

i. each T is a k-overlapping tree;
ii. if $\mathsf{T}, \mathsf{T}' \in \mathbf{T}$, $\mathsf{T} \ne \mathsf{T}'$, then for all $\sigma \in \mathbb{N}, \tau = 1,\ldots,2^6$ there holds

$$10\omega_{\mathsf{T},k,(\sigma,\tau),\bullet} \cap 10\omega_{\mathsf{T}',k,(\sigma,\tau),\bullet} \ne \varnothing \implies I_\mathsf{T} \cap I_{\mathsf{T}'} = \varnothing.$$

The analogue of Lemma 6.16 in the overlapping setup is the following. In the statement, we find convenient to denote $H_{\mathsf{T},\pm}$ the frequency restriction of f to the half-line $\pm\xi > (\xi_\mathsf{T})_k$.

6.16. **Lemma.** *Let $\mathcal{X}$ be a Hilbert space and $\mathbf{T}$ be a collection of $\mathsf{ov};k$-strongly disjoint trees of type $\pm$. There holds*

$$\left\| \sqrt{|I_\mathsf{T}|}\|H_{\mathsf{T},\pm}f\|_{\mathsf{ov};\mathsf{T},k,2} \right\|_{\ell^2(\mathsf{T}\in\mathbf{T})} \lesssim \|f\|_{L^2(\mathbb{R};\mathcal{X})} + \left(\|f\|_{L^\infty(\mathbb{R};\mathcal{X})} \left[\sum_{\mathsf{T}\in\mathbf{T}} |I_\mathsf{T}| \right]^{\frac{1}{2}} \right)^{\frac{1}{3}} \|f\|^{\frac{2}{3}}_{L^2(\mathbb{R};\mathcal{X})}. \tag{6.17}$$

The proof of Lemma 6.16 is similar to that of the lacunary case, with some modifications. As it is nonstandard, we provide the complete argument below. For the sake of definiteness we work in the + case. Let us again fix the scaling $\|f\|_{L^2(\mathbb{R};\mathcal{X})} = 1$. By linearizing the supremum in each $\|H_{\mathsf{T},\pm}f\|_{\mathsf{ov};\mathsf{T},k,2}$ we realize we need to estimate

$$S := \sqrt{\sum_{\mathsf{T}\in\mathbf{T}} \left\| u_{I_\mathsf{T}} T_{m_\mathsf{T}} f \right\|^2_{L^2(\mathbb{R};\mathcal{X})}}$$

for an extremizing choice of multiplier $m_\mathsf{T} \in \widetilde{M}_{\omega_{\mathsf{T},k}}$ whose support lies in the right half of $\omega_{\mathsf{T},k}$. For simplicity, we redefine

$$\|f\|_\mathsf{T} := \frac{\left\| u_{I_\mathsf{T}} T_{m_\mathsf{T}} f \right\|_{L^2(\mathbb{R};\mathcal{X})}}{\sqrt{|I_\mathsf{T}|}}$$

and note for future use that $\|f\|_{\mathsf{T}} \lesssim \|f\|_\infty$. Arguing as in the previous subsection, we obtain that

$$S^2 \sim \left\langle \sum_{\mathsf{T}\in\mathbf{T}} \|f\|_{\mathsf{T}} T^*_{m_{\mathsf{T}}} v_{\mathsf{T}}, f \right\rangle \leq \left\| \sum_{\mathsf{T}\in\mathbf{T}} \|f\|_{\mathsf{T}} T^*_{m_{\mathsf{T}}} v_{\mathsf{T}} \right\|_{L^2(\mathbb{R};\mathcal{X})}, \tag{6.18}$$

where the $\mathcal{X}$-valued functions v_{T} satisfy the bound

$$|v_{\mathsf{T}}|_{\mathcal{X}} \lesssim u_{I_{\mathsf{T}}}. \tag{6.19}$$

Using a smooth partition of unity subordinated to the cover $\{\omega_{\mathsf{T},k,(\sigma,\tau),+} : \sigma \in \mathbb{N}, \tau = 1,\ldots,2^6\}$ of the support of m_{T}, and the fact that m_{T} vanishes at $(\xi_{\mathsf{T}})_k$, we may then decompose

$$T_{m_{\mathsf{T}}} = \sum_{\sigma\in\mathbb{N}} \sum_{\tau=1}^{2^6} 2^{-\sigma} T_{m_{\mathsf{T},\sigma,\tau}},$$

where the multiplier $m_{\mathsf{T},\sigma,\tau}$ is adapted to and supported on $\omega_{\mathsf{T},k,(\sigma,\tau),+}$. We will prove the estimate

$$U_{\sigma,\tau} := \left\| \sum_{\mathsf{T}\in\mathbf{T}} \|f\|_{\mathsf{T}} T^*_{m_{\mathsf{T},\sigma,\tau}} v_{\mathsf{T}} \right\|_{L^2(\mathbb{R};\mathcal{X})} \lesssim 2^{\frac{\sigma}{2}} \left(S^2 + \|f\|_\infty \left(\sum_{\mathsf{T}\in\mathbf{T}} |I_{\mathsf{T}}| \right)^{\frac12} S \right)^{\frac12} \tag{6.20}$$

uniformly over σ, τ which combined with the triangle inequality and (6.18) returns (6.17) via standard manipulations. As τ does not play any role in the argument below and takes 64 values, we fix a value and omit it from the notation. Squaring (6.20) gives

$$U_\sigma^2 = \sum_{\mathsf{T}\in\mathbf{T}} \|f\|^2_{\mathsf{T}} \|T_{m_{\mathsf{T},\sigma}} v_{\mathsf{T}}\|_2^2 + 2 \sum_{\mathsf{T}\in\mathbf{T}} \sum_{\mathsf{T}'\in\mathbf{T}(\mathsf{T})} \|f\|_{\mathsf{T}} \|f\|_{\mathsf{T}'} \langle T^*_{m_{\mathsf{T},\sigma}} v_{\mathsf{T}}, T^*_{m_{\mathsf{T}',\sigma}} v_{\mathsf{T}'} \rangle, \tag{6.21}$$

where $\mathbf{T}(\mathsf{T}) := \{\mathsf{T}' \in \mathbf{T} : \ell(I_{\mathsf{T}'}) \leq \ell(I_{\mathsf{T}}), \omega_{\mathsf{T},k,\sigma,+} \cap \omega_{\mathsf{T}',k,\sigma,+} \neq \varnothing\}$. The first term in (6.21) is $\lesssim S^2$, as (6.19) and L^2-boundedness tells us that $\|T_{m_{\mathsf{T},\sigma}} v_{\mathsf{T}}\|_2^2 \lesssim |I_{\mathsf{T}}|$.

We move to the second term. Suppose $\mathsf{T}' \in \mathbf{T}(\mathsf{T})$. The $O(2^{-\sigma}\ell(\omega_{\mathsf{T}}))$-frequency localization of $T_{m_{\mathsf{T}',\sigma}} T^*_{m_{\mathsf{T},\sigma}}$ and the fact that v_{T} is localized on I_{T} entail

$$\left| \langle T^*_{m_{\mathsf{T},\sigma}} v_{\mathsf{T}}, T^*_{m_{\mathsf{T}',\sigma}} v_{\mathsf{T}'} \rangle \right| \lesssim \langle u_{2^\sigma I_{\mathsf{T}}}, u_{I_{\mathsf{T}'}} \rangle \lesssim \|\mathbf{1}_{I_{\mathsf{T}'}} u_{2^\sigma I_{\mathsf{T}}}\|_1.$$

We then notice that if $\mathsf{T}' \in \mathbf{T}(\mathsf{T})$ then $\omega_{\mathsf{T},k,\sigma,+} \subset 10\omega_{\mathsf{T}',k,\sigma,+}$, which together with ii. in the definition of $\mathsf{ov};k$-strong disjointness of type + tells us that the intervals $\{I_{\mathsf{T}'} : \mathsf{T}' \in \mathbf{T}(\mathsf{T})\}$ are pairwise disjoint. Proceeding as in (6.15), we then bound the second term in (6.21) by

$$\|f\|_\infty \sum_{\mathsf{T}\in\mathbf{T}} \|f\|_{\mathsf{T}} \sum_{\mathsf{T}'\in\mathbf{T}(\mathsf{T})} \|\mathbf{1}_{I_{\mathsf{T}'}} u_{2^\sigma I_{\mathsf{T}}}\|_1 \leq \|f\|_\infty \sum_{\mathsf{T}\in\mathbf{T}} \|f\|_{\mathsf{T}} \|u_{2^\sigma I_{\mathsf{T}}}\|_1 \lesssim 2^\sigma \|f\|_\infty \left(\sum_{\mathsf{T}\in\mathbf{T}} |I_{\mathsf{T}}| \right)^{\frac12} S, \tag{6.22}$$

where we used pairwise disjointness of $\{I_{\mathsf{T}'} : \mathsf{T}' \in \mathbf{T}(\mathsf{T})\}$ in the first step. Chaining (6.22) with (6.21) yields the claimed bound for (6.20), and finishes the proof of Lemma 6.16.

6.3. **Transporting almost orthogonality to intermediate spaces.** In the previous subsections, we have shown that the definitions (3.10), (6.2) lead to Hilbert space valued orthogonality estimates for families of strongly disjoint trees. The point is that the definitions (3.10), (6.2) are of maximal nature and involve more general operators than the rank 1 projections $f \mapsto \langle f, \varphi_{P_k}\rangle \varphi_{P_k}$ of [18], namely operators of the class $\mathbb{S}_{P_k}$. It is because of this additional generality that we had to reproduce, with small changes, the classical TT^* arguments of [18].

Now that our version of [18, Prop. 6.1], namely Lemma 6.5 is in place, the interpolation arguments of [18, Section 7] may be perused *mutatis mutandis*, leading to the following almost orthogonality estimate for interpolation spaces.

6.23. **Proposition.** *Let* $2 \le p < \infty$ *and* $\mathcal{X} = [\mathcal{Y}_0, \mathcal{Y}_1]_{\frac{2}{p}}$ *be the complex interpolation space of a* UMD *space* $\mathcal{Y}_0$ *and a Hilbert space* $\mathcal{Y}_1$*. Then for all* $0 < \alpha \le 1$ *the inequality*

$$\left\| |I_{\mathsf{T}}|^{\frac{1}{p}} \mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T};p) \right\|_{\ell^p(\mathsf{T}\in\mathbf{T})} \lesssim_\alpha \|f\|_{L^p(\mathbb{R};\mathcal{X})} + \left(\|f\|_{L^\infty(\mathbb{R};\mathcal{X})} \left[\sum_{\mathsf{T}\in\mathbf{T}} |I_{\mathsf{T}}| \right]^{\frac{1}{p}} \right)^{1-\alpha} \|f\|^{\alpha}_{L^p(\mathbb{R};\mathcal{X})} \tag{6.24}$$

holds uniformly over all collections $\mathbf{T}$ *of* $\mathsf{lac};k$*-strongly disjoint trees while the inequality*

$$\left\| |I_{\mathsf{T}}|^{\frac{1}{p}} \|H_{\mathsf{T},\pm} f\|_{\mathsf{ov};\mathsf{T},k,p} \right\|_{\ell^p(\mathsf{T}\in\mathbf{T})} \lesssim_\alpha \|f\|_{L^p(\mathbb{R};\mathcal{X})} + \left(\|f\|_{L^\infty(\mathbb{R};\mathcal{X})} \left[\sum_{\mathsf{T}\in\mathbf{T}} |I_{\mathsf{T}}| \right]^{\frac{1}{p}} \right)^{1-\alpha} \|f\|^{\alpha}_{L^p(\mathbb{R};\mathcal{X})} \tag{6.25}$$

holds uniformly over all collections $\mathbf{T}$ *of* $\mathsf{ov};k$*-strongly disjoint trees of type* $\pm$.

Proof. We first prove (6.24). The first step of the proof consists of deducing the case $p = 2$ of (6.24) from Lemma 6.5. This is accomplished following step by step the proof of [18, Proposition 6.6]. The second step consists in the deduction of an endpoint at $p = \infty$, which is

$$\left\| \mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T};\star) \right\|_{\ell^\infty(\mathsf{T}\in\mathbf{T})} \lesssim \|f\|_{L^\infty(\mathbb{R};\mathcal{X})} \tag{6.26}$$

having denoted

$$\|f\|_{\mathsf{lac};k,\mathsf{T},\star} := \sup \left\| \mathrm{Mod}_{-\xi_{\mathsf{T}}} S_{\mathsf{T}} f \right\|_{\mathrm{BMO}(\mathbb{R};\mathcal{X})}, \quad \mathsf{eng}_{\mathsf{lac};k}(f)(\mathbb{P};\star) := \sup_{\substack{\mathsf{T}'\subset\mathbb{P} \\ \mathsf{T}'\ k-\text{lacunary}}} \|f\|_{\mathsf{lac};k,\mathsf{T}',\star}$$

where Mod_ξ stands for modulation by ξ, and as usual the first supremum is taken over all possible choices of type k tree operators S_{T}. The estimate (6.26) is an immediate consequence of the uniform estimate for demodulated tree operators

$$\mathrm{Mod}_{-(\xi_{\mathsf{T}})_k} S_{\mathsf{T}} \mathrm{Mod}_{(\xi_{\mathsf{T}})_k} : L^\infty(\mathbb{R};\mathcal{X}) \to \mathrm{BMO}(\mathbb{R};\mathcal{X}),$$

which holds by virtue of the fact that each operator $\mathrm{Mod}_{-(\xi_{\mathsf{T}})_k} S_{\mathsf{T}} \mathrm{Mod}_{(\xi_{\mathsf{T}})_k}$ is a Calderón-Zygmund operator. Finally, the proof of the proposition is obtained by complex interpolation of the case $q = 2$ of (6.24) with (6.26). Details are given in [18, Proposition 7.3].

The proof of (6.25) is similar, the only difference being the endpoint inequality

$$\left\| \|H_{\mathsf{T},\pm} f\|_{\mathsf{ov};\mathsf{T},k,\star} \right\|_{\ell^\infty(\mathsf{T}\in\mathbf{T})} \lesssim \|f\|_{L^\infty(\mathbb{R};\mathcal{X})}, \tag{6.27}$$

where, in analogy with (3.9)

$$\|g\|_{\mathsf{ov};\mathsf{T},k,\star} := \sup_{S\in\mathbb{S}_{t_{\mathsf{T},k}}} \left\|\mathrm{Mod}_{-\xi_{\mathsf{T}}} Sg\right\|_{\mathrm{BMO}(\mathbb{R};\mathcal{X})}.$$

The bound (6.27) is easily established: if $S \in \mathbb{S}_{t_{\mathsf{T},k}}$ the composition $\mathrm{Mod}_{-(\xi_{\mathsf{T}})_k} SH_{\mathsf{T},\pm}\mathrm{Mod}_{(\xi_{\mathsf{T}})_k}$ is a Calderón-Zygmund operator, so that

$$\left\|\mathrm{Mod}_{-(\xi_{\mathsf{T}})_k} SH_{\mathsf{T},\pm} f\right\|_{\mathrm{BMO}(\mathbb{R};\mathcal{X})} \lesssim \left\|\mathrm{Mod}_{-(\xi_{\mathsf{T}})_k} f\right\|_{L^\infty(\mathbb{R};\mathcal{X})} = \left\|f\right\|_{L^\infty(\mathbb{R};\mathcal{X})}$$

as claimed. The proof of Proposition 6.23 is thus complete. □

6.4. **The proof proper of Lemma 3.13.** The proof is iterative in nature. One additional remark necessary here is that the selected trees come from a greedy selection process, as described in Section A.5 of the Appendix, and therefore satisfy properties g1. to g3. appearing in Subsection 3.1.

For the proof, write $p = q_{\mathcal{X}}$, $\lambda := \mathsf{eng}_k(f)(\mathbb{P};p)$ and let $\alpha \in (0,1)$ be chosen so that $q = p/\alpha$. We start by excising high k-lacunary energies. Performing an iterative algorithm analogous to [29, Lemma 7.7], we decompose

$$\mathbb{P} := \mathbb{P}^{\mathsf{low,lac}} \cup \mathbb{P}^{\mathsf{hi,lac}},$$

where

$$\mathsf{eng}_{\mathsf{lac},k}(f)(\mathbb{P}^{\mathsf{low,lac}};p) \leq \frac{\lambda}{2} \tag{6.28}$$

and $\mathbb{P}^{\mathsf{hi,lac}} = \bigcup\{\mathsf{T} : \mathsf{T} \in \mathbf{T}^{\mathsf{hi,lac}}\}$, where $\mathbf{T}^{\mathsf{hi,lac}}$ is a family of greedily, in the sense of Subsection A.5 of the Appendix, selected trees with the following property: for each T there exists a k-lacunary tree $\mathsf{T}' \subset \mathsf{T}$ with same top data as T, and the family $\mathbf{T}'^{,\mathsf{hi,lac}} = \{\mathsf{T}' : \mathsf{T} \in \mathbf{T}^{\mathsf{hi,lac}}\}$ consists of $\mathsf{lac};k$-strongly disjoint trees with

$$\mathsf{eng}_{\mathsf{lac};k}(f)(\mathsf{T}';p) \gtrsim \lambda.$$

Using the first part of Proposition 6.23 in the second inequality,

$$\begin{aligned} \lambda^p \sum_{\mathsf{T}\in\mathbf{T}^{\mathsf{hi,lac}}} |I_{\mathsf{T}}| &\lesssim \left\| |I_{\mathsf{T}}|^{\frac{1}{p}} \mathsf{eng}_k(f)(\mathsf{T};p)\right\|^p_{\ell^p(\mathsf{T}\in\mathbf{T}'^{,\mathsf{hi,lac}})} \\ &\lesssim \|f\|^p_{L^p(\mathbb{R};\mathcal{X})} + \|f\|^{p(1-\alpha)}_{L^\infty(\mathbb{R};\mathcal{X})}\left(\sum_{\mathsf{T}\in\mathbf{T}^{\mathsf{hi,lac}}}|I_{\mathsf{T}}|\right)^{1-\alpha}\|f\|^{\alpha p}_{L^p(\mathbb{R};\mathcal{X})} \lesssim |F| + \left(\sum_{\mathsf{T}\in\mathbf{T}^{\mathsf{hi,lac}}}|I_{\mathsf{T}}|\right)^{1-\alpha}|F|^{\alpha}. \end{aligned} \tag{6.29}$$

Dividing into cases depending on whether the $|F|$ summand in the last line is larger or not than the $|F|^\alpha$ one,

$$\sum_{\mathsf{T}\in\mathbf{T}^{\mathsf{hi,lac}}} |I_{\mathsf{T}}| \lesssim \max\{\lambda^{-p},\lambda^{-q}\}|F| \lesssim \lambda^{-q}|F|. \tag{6.30}$$

In the last comparison we have used that $q > p$ and

$$\lambda \lesssim \sup_{P\in\mathbb{P}} \inf_{I_P} \mathrm{M}(|f|_{\mathcal{X}}) \lesssim 1,$$

a consequence of Lemma 3.12.

We then excise from $\mathbb{P}^{\mathsf{low,lac}}$ the high overlapping energies. As this piece is less standard, we produce an explicit iterative algorithm, which is in fact the chaining of two subsequent similar iterative procedures, one for each type of k-overlapping tree. First, we run the following greedy selection algorithm:

INIT. $\mathbb{P}^{\mathsf{stc}} := \mathbb{P}^{\mathsf{low,lac}}, \mathbf{T}^{\mathsf{hi,ov,+}} = \emptyset, \mathbf{T}'^{\mathsf{,hi,ov,+}} = \emptyset$.

WHILE the collection of k-overlapping trees (T', I, ξ) with $\mathsf{T}' \subset \mathbb{P}^{\mathsf{stc}}$ and the property that

$$\|H_{\mathsf{T}',+} f\|_{\mathsf{ov};\mathsf{T}',k,p} > \frac{\lambda}{4} \tag{6.31}$$

is nonempty choose within such collection a k-overlapping tree (T', I, ξ) with the property that

$$+\xi_k \quad \text{is maximal.}$$

Then let T be the maximal, with respect to inclusion, tree contained in $\mathbb{P}^{\mathsf{stc}}$ with the same top data (I, ξ) as T'. At the end of this proof we will refer to T as the *completion* of the tree T'. Then set

$$\mathbb{P}^{\mathsf{stc}} := \mathbb{P}^{\mathsf{stc}} \setminus \mathsf{T}, \qquad \mathbf{T}^{\mathsf{hi,ov,+}} := \mathbf{T}^{\mathsf{hi,ov,+}} \cup \{\mathsf{T}\}, \qquad \mathbf{T}'^{\mathsf{,hi,ov,+}} := \mathbf{T}'^{\mathsf{,hi,ov,+}} \cup \{\mathsf{T}'\}. \tag{6.32}$$

When the algorithm terminates, set $\mathbb{P}^{\mathsf{low,+}} := \mathbb{P}^{\mathsf{stc}}$. Subsequently, perform again the above iterative algorithm, replacing the initialization step by

INIT. $\mathbb{P}^{\mathsf{stc}} := \mathbb{P}^{\mathsf{low,+}}, \mathbf{T}^{\mathsf{hi,ov,-}} = \emptyset, \mathbf{T}'^{\mathsf{,hi,ov,-}} = \emptyset$.

and replacing $+$ by $-$ in (6.31) to (6.32). Once this second algorithm has terminated, we finally set

$$\mathbb{P}^{\mathsf{low}} := \mathbb{P}^{\mathsf{stc}}, \qquad \mathbf{T} := \mathbf{T}^{\mathsf{hi,lac}} \cup \mathbf{T}^{\mathsf{hi,ov,+}} \cup \mathbf{T}^{\mathsf{hi,ov,-}}.$$

We notice that in view of the last two iterative algorithms, if (T', I, ξ) is a k-overlapping tree with $\mathsf{T}' \subset \mathbb{P}^{\mathsf{low}}$, it must be

$$\|f\|_{\mathsf{ov};\mathsf{T}',k,p} \le \|H_{\mathsf{T}',+} f\|_{\mathsf{ov};\mathsf{T}',k,p} + \|H_{\mathsf{T}',-} f\|_{\mathsf{ov};\mathsf{T}',k,p} \le \frac{\lambda}{4} + \frac{\lambda}{4} = \frac{\lambda}{2},$$

where the second inequality holds because the algorithms terminated without T' being selected in either. Therefore

$$\mathsf{eng}_{\mathsf{ov},k}(f)(\mathbb{P}^{\mathsf{low}}; p) \le \frac{\lambda}{2}$$

and, also in view of (6.28) and the inclusion $\mathbb{P}^{\mathsf{low}} \subset \mathbb{P}^{\mathsf{low,lac}}$, the small energy estimate on $\mathbb{P}^{\mathsf{low}}$; (3.14) is proved. We will show at the end of the proof that the families $\mathbf{T}'^{\mathsf{,hi,ov,\pm}}$ are of $\mathsf{ov}; k$-strongly disjoint of type $\pm$. The estimate on the sum of the tree tops for $\mathbf{T}^{\mathsf{hi,ov,\pm}}$ is proved analogously to what we did in (6.29), but appealing to the second part of Proposition 6.23 instead. This consideration, together with the estimate (6.30), yields the counting function bound (3.15).

We prove the claim that that the families $\mathbf{T}'^{\mathsf{,hi,ov,\pm}}$ thus selected are $\mathsf{ov}; k$-strongly disjoint of type $\pm$. For the sake of definiteness, we work in the $+$ case and for simplicity write $\mathbf{T} = \mathbf{T}'^{\mathsf{,hi,ov,+}}$. Suppose $\mathsf{T}, \mathsf{T}' \in \mathbf{T}$ are such that the intersection $10\omega_{\mathsf{T},k,(\sigma,\tau),+} \cap 10\omega_{\mathsf{T}',k,(\sigma,\tau),+} \neq \emptyset$. We need to show that $I_{\mathsf{T}} \cap I_{\mathsf{T}'} = \emptyset$.

Case $\ell(I_{\mathsf{T}}) = \ell(I_{\mathsf{T}'})$. In this case, the assumption $10\omega_{\mathsf{T},k,(\sigma,\tau),+} \cap 10\omega_{\mathsf{T}',k,(\sigma,\tau),+} \neq \emptyset$ yields $|(\xi_{\mathsf{T}})_k - (\xi_{\mathsf{T}'})_k| \lesssim 20\ell(\omega_{\mathsf{T},k}) = 20\ell(\omega_{\mathsf{T}',k})$. Suppose by contradiction $I_{\mathsf{T}} \cap I_{\mathsf{T}'} \neq \emptyset$. In this case by dyadicity $I_{\mathsf{T}} = I'_{\mathsf{T}}$. By symmetry suppose that T has been selected first. As trees are nonempty, we may find $P' \in \mathsf{T}'$ such that $I_P \subset I_{\mathsf{T}'} = I_{\mathsf{T}}$ and $(\xi_{\mathsf{T}'})_k \in 3\omega_{\mathsf{T}',k}$. The latter property and $\ell(\omega_{P'_k}) \geq \ell(\omega_{\mathsf{T}',k})$ implies that $(\xi_{\mathsf{T}})_k \in \overline{\omega_{P'_k}}$ for all k. Therefore the tri-tile P' qualifies to be in the completion of T and hence was not available when T' was selected. Contradiction.

Case $\ell(I_{\mathsf{T}}) > \ell(I_{\mathsf{T}'})$. First we prove by contradiction that[4] $(\xi_{\mathsf{T}})_k > (\xi_{\mathsf{T}'})_k$, so that T must have been selected before T'. Suppose $(\xi_{\mathsf{T}})_k < (\xi_{\mathsf{T}'})_k$ then $\sup 10\omega_{\mathsf{T}',k,(\sigma,\tau),+} \leq (\xi_{\mathsf{T}})_k + 2^{-\sigma+1}\ell(\omega_{\mathsf{T},k})$ while $\inf 10\omega_{\mathsf{T}',k,(\sigma,\tau),+} \geq (\xi_{\mathsf{T}'})_k + 2^{-\sigma-1}\ell(\omega_{\mathsf{T}',k})$, which is a contradiction, as $\ell(\omega_{\mathsf{T}',k}) \geq 2^J\ell(\omega_{\mathsf{T},k})$ due to J-dyadicity. Now, assume again for contradiction purposes that $I_{\mathsf{T}} \cap I_{\mathsf{T}'} \neq \emptyset$. We may then find $P' \in \mathsf{T}'$ such that $I_{P'} \subset I_{\mathsf{T}'}$ and $(\xi_{\mathsf{T}'})_k \in 3\omega_{\mathsf{T}',k}$. The latter property and $\ell(\omega_{P'_k}) \geq \ell(\omega_{\mathsf{T}',k}) \geq |(\xi_{\mathsf{T}})_k - (\xi_{\mathsf{T}'})_k|/20$ imply that $(\xi_{\mathsf{T}})_k \in \overline{\omega_{P'_k}}$ for all k. Again, we have showed that P' qualified to be in the completion of T and hence was not available when T' was selected. Contradiction. This case completes the proof of $\mathsf{ov};k$-strong disjointness of type $\pm$ for the families $\mathbf{T}^{',\mathsf{hi},\mathsf{ov},\pm}$ and thus the proof of Lemma 3.13.

7. Proof of Lemma 3.12

Notation. Throughout this proof, if I is a J-dyadic interval, we write $I^{+v} = I + v\ell(I)$ for $v \in \mathbb{Z}$ to denote the v-th translate of I. Further, we introduce the local notation

$$\gamma_I(x) := \left\langle \frac{x - c(I)}{\ell(I)} \right\rangle^{100}, \qquad x \in \mathbb{R}. \tag{7.1}$$

The polynomial γ_I will be used to apply the so-called localization trick. As we perform this a few times in the proof, we isolate the related notation here. If T is a k-lacunary tree and S_{T} a tree operator of the order appearing in the definition of (3.10), we write

$$\widetilde{S_{\mathsf{T}}}g := \sum_{P \in \mathsf{T}} \tilde{S}_{P_k}g, \quad \tilde{S}_{P_k}g := \gamma_{I_{\mathsf{T}}} S_{P_k} g. \tag{7.2}$$

It is immediate to verify that $\tilde{S}_{P_k} \in \mathbb{S}_{P_k}$ for all $P \in \mathsf{T}$, so that $\widetilde{S_{\mathsf{T}}}$ is also a tree operator albeit of a slightly different order. This difference is inconsequential for our analysis and we do not keep track of it in the notation.

The overlapping term. We first deal with the overlapping part of the energy which is much easier. Let $\mathsf{T} \subset \mathbb{P}$ be a k-overlapping tree with top data $(I_{\mathsf{T}}, \xi_{\mathsf{T}})$ extremizing $\mathsf{eng}_{\mathsf{ov};k}(f)(\mathbb{P}; q)$ and

$$\zeta \in X_{I_T}, \qquad m \in \widetilde{M}_{\omega_{\mathsf{T},k}} \subset M_{\omega_{\mathsf{T},k}}$$

be the corresponding data extremizing $\|f\|_{\mathsf{ov};\mathsf{T},k,q}$. By the rapid decay of ζ at scale $\ell(I_{\mathsf{T}})$

$$\frac{\|\zeta T_m f\|_{L^q(\mathcal{X})}}{|I_{\mathsf{T}}|^{\frac{1}{q}}} \lesssim \sup_{\sigma \geq 0} 2^{-100\sigma} \|F * |K|\|_{L^\infty(2^\sigma I_{\mathsf{T}})}, \tag{7.3}$$

[4]This claim is best proved by picture.

where, locally, $F = |f|_{\mathcal{X}}$ and K stands for the inverse Fourier transform of m. From the localization of m, $|K| \lesssim \frac{1}{|I_{\mathsf{T}}|\gamma_{I_{\mathsf{T}}}}$, so that if $x \in 2^\sigma I_{\mathsf{T}}$ we have

$$F * |K|(x) \lesssim \sup_\tau 2^{-90\tau} \inf_{y \in B_{2^\tau \ell(I_{\mathsf{T}})}(x)} \mathrm{M}F(y) \lesssim \inf_{y \in B_{\ell(I_{\mathsf{T}})}(x)} \mathrm{M}F(y) \lesssim 2^\sigma \inf_{y \in I_{\mathsf{T}}} \mathrm{M}F(y).$$

Combining the last display with (7.3),

$$\mathsf{eng}_{\mathsf{ov};k}(f)(\mathbb{P};q) \lesssim \frac{\|\zeta T_m f\|_{L^q(\mathcal{X})}}{|I_{\mathsf{T}}|^{\frac{1}{q}}} \lesssim \inf_{I_{\mathsf{T}}} \mathrm{M}F \leq \inf_{I_P} \mathrm{M}F, \tag{7.4}$$

where we used that T contains at least one tri-tile P, and $I_P \subset I_{\mathsf{T}}$. This completes the handling of the overlapping part of the energy.

The lacunary term. It remains to estimate the lacunary component. The proof strategy is an adaptation of [18, Section 9]: indeed, with (7.4) at hand, the bound of Lemma 3.12 is an immediate consequence of the estimate (7.5) below. Having fixed a k-lacunary tree T, there holds

$$\left\|S_{\mathsf{T}} f\right\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim_q \lambda |I_{\mathsf{T}}|^{\frac{1}{q}}, \qquad \lambda := \sup_{I \in \mathcal{I}} \inf_I \mathrm{M}(|f|_{\mathcal{X}}), \quad \mathcal{I} := \{I_P : P \in \mathsf{T}\}, \qquad 1 < q < \infty. \tag{7.5}$$

The estimate is uniform over tree operators S_{T}.

By modulation invariance of (7.5), we may reduce to treating the case $\xi_{\mathsf{T}} = 0$. Then, estimate (7.5) will be obtained as a consequence of the next lemma.

7.6. **Lemma.** *Let T be a k-lacunary tree with $\xi_{\mathsf{T}} = 0$, S_{T} be a tree operator and λ be the same as in* (7.5)*. For each J-dyadic interval $K \subset \mathbb{R}$ there exists a constant a_K with the property that*

$$\left\|\mathbf{1}_K (S_{\mathsf{T}} f - a_K)\right\|_{L^{1,\infty}(\mathbb{R};\mathcal{X})} \lesssim \lambda |K| \tag{7.7}$$

with bound independent of K, S_{T} and T. In particular, if $\ell(K) \geq \ell(I_{\mathsf{T}})$ we may take $a_K = 0$.

We use Lemma 7.6 to finish the proof of (7.5). Fix a tree operator S_{T}. Then, referring to (7.2),

$$\left\|S_{\mathsf{T}} f\right\|_{L^q(\mathbb{R};\mathcal{X})} = \left\|\gamma_{I_{\mathsf{T}}}^{-1} \widetilde{S_{\mathsf{T}}} f\right\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim \sum_{v \in \mathbb{Z}} \langle v \rangle^{-100} \left\|\mathbf{1}_{I_{\mathsf{T}}^{+v}} \widetilde{S_{\mathsf{T}}} f\right\|_{L^q(\mathbb{R};\mathcal{X})}. \tag{7.8}$$

But, Lemma 7.6 applied to $\tilde{T}$ together with the John-Strömberg inequality yields the two estimates

$$\left\|\mathbf{1}_{I_{\mathsf{T}}^{+v}} \widetilde{S_{\mathsf{T}}} f\right\|_{L^{1,\infty}(\mathbb{R};\mathcal{X})} \lesssim \lambda |I_{\mathsf{T}}|, \qquad \left\|\widetilde{S_{\mathsf{T}}} f\right\|_{\mathrm{BMO}(\mathbb{R};\mathcal{X})} \lesssim \lambda,$$

which together with the John-Nirenberg inequality tell us that

$$\left\|\mathbf{1}_{I_{\mathsf{T}}^{+v}} \widetilde{S_{\mathsf{T}}} f\right\|_{L^q(\mathbb{R};\mathcal{X})} \lesssim \lambda |I_{\mathsf{T}}|^{\frac{1}{q}}. \tag{7.9}$$

A combination of (7.9) and (7.8) finally yields (7.5).

Proof of Lemma 7.6. We fix a tree operator and use the local notation

$$S_{\mathsf{T}} f = \sum_{I \in \mathcal{I}} S_I f$$

where $S_I = S_{P_k} \in \mathbb{S}_{P_k}$ for the unique tri-tile $P \in \mathsf{T}$ with $I_P = I$.

We begin the proof with the definition of the constant a_K. This constant comes from the large scales contribution on K, that is the intervals

$$\mathcal{I}_{\mathrm{low}} = \{I \in \mathcal{I} : \ell(I) > \ell(K)\}.$$

For $n \in \mathbb{N}$ let $K^{(n)}$ be the n-th J-dyadic parent of K. Then if $I \in \mathcal{I}_{\mathrm{low}}$, it must be $I = K^{(n)+v}$ for some $n \in \mathbb{N}, v \in \mathbb{Z}$. We define

$$a_K = \sum_{n\geq 1}\sum_{v\in\mathbb{Z}} S_{K^{(n)+v}} f(c(K)) \tag{7.10}$$

where we have simply set $S_{K^{(n)+v}} = 0$ if $K^{(n)+v} \notin \mathcal{I}_{\mathrm{low}}$. Clearly, the second claim now follows from the first, as $\mathcal{I}_{\mathrm{low}}$ is empty, whence a_K is zero, when $\ell(K) \geq \ell(I_{\mathsf{T}})$.

We continue with the proof of (7.7). We claim that

$$\mathbf{1}_K \sum_{n\geq 1}\sum_{v\in\mathbb{Z}} \left|S_{K^{(n)+v}} f(c(K)) - S_{K^{(n)+v}} f(x)\right|_{\mathcal{X}} \lesssim \lambda. \tag{7.11}$$

Indeed, denoting by $F = |f|_{\mathcal{X}}$, by $u_{n,v}$ the kernel of $T_{K^{(n)+v}}$ and by $\chi_{n,v} = \chi_{K^{(n)+v}}$ for simplicity, and using the kernel estimates for $u_{n,v}$ and the extra decay in v, we have for $x \in K$

$$\begin{aligned}
&\left|S_{K^{(n)+v}} f(c(K)) - S_{K^{(n)+v}} f(x)\right| \\
&\leq \left|\chi_{n,v}(x) - \chi_{n,v}(c(K))\right| (F * |u_{n,v}|(x)) + \int_x^{c(K)} F * \left|Du_{n,v}\right|(z)\,\mathrm{d}z \\
&\lesssim \langle v\rangle^{-100} 2^{-n} \inf_{K^{(n)+v}} \mathrm{M}F \leq \langle v\rangle^{-100} 2^{-n}\lambda,
\end{aligned}$$

which is summable over v, n in (7.11). The last estimate follows from the membership of $K^{(n)+v}$ to $\mathcal{I}$.

We now come to the small scales. We first deal with the contribution of the intervals

$$\mathcal{I}^{\mathrm{high}}_{n,v} = \{I \in \mathcal{I} : \ell(I) = 2^{-n}\ell(K), I \subset K^{+v}\}, \qquad n \geq 0,\ v \in \mathbb{Z},\ |v| > 1.$$

Notice that this excludes the intervals $\mathcal{I}^{\mathrm{high}} = \{I \in \mathcal{I} : I \subset 3K\}$ which will be handled as the main term. The $\mathcal{I}^{\mathrm{high}}_{n,v}$ are tail terms: in fact, with the same notations as before, if $x \in K$ and $I \in I^{\mathrm{high}}_{n,v}$

$$|\chi_I(x)| (F * |u_{n,v}|(x)) \lesssim (v2^n)^{-100} \sum_{t\geq 0} 2^{-100t} \langle F\rangle_{[x-2^{t+1}\ell(I), x+2^{t+1}\ell(I)]}\,. \tag{7.12}$$

As, for $x \in K$,

$$\langle F\rangle_{[x-2^{t+1}\ell(I), x+2^{t+1}\ell(I)]} \lesssim \begin{cases} \dfrac{v\ell(K)}{2^t\ell(I)} \langle F\rangle_{[x-2^{10}v\ell(K), x+2^{10}v\ell(K)]} \leq (v2^n) \inf\limits_{x\in I} \mathrm{M}F(x) & 2^{t+1}\ell(I) \leq v\ell(K) \\ \inf\limits_{x\in I} \mathrm{M}F(x) & 2^{t+1}\ell(I) > v\ell(K) \end{cases}$$

we obtain by summation of (7.12) that

$$\mathbf{1}_K \sum_{n\geq 0}\sum_{|v|\geq 2}\sum_{I\in\mathcal{I}^{\mathrm{high}}_{n,v}} \left|S_I f\right|_{\mathcal{X}} \lesssim \lambda. \tag{7.13}$$

We are left to estimate the contribution of $\mathcal{I}^{\mathrm{high}}$. The union of the intervals $\mathcal{I}^{\mathrm{high}}$ is contained in $3K$. By possibly splitting $\mathcal{I}^{\mathrm{high}}$ into three collections and replacing $I \in \mathcal{I}^{\mathrm{high}}$

with the corresponding smoothing interval from one of three shifted dyadic grids, so that the union is still contained in $18K$, we can achieve the property that if $I, L \in \mathcal{I}^{\text{high}}$ and $I \subset L$ then $3I \subset L$.

Let now $L \in \mathcal{L}$ be the collection of those $L \in \mathcal{I}^{\text{high}}$ which are maximal with respect to inclusion and $\mathcal{I}(L) = \{I \in \mathcal{I}^{\text{high}} : I \subsetneq L\}$. First we remove the tops. It is immediate to bound

$$\sum_{L\in\mathcal{L}} \left\|S_L f\right\|_{L^1(\mathbb{R};\mathcal{X})} \lesssim \sum_{L\in\mathcal{L}} |L| \inf_L \mathrm{M}F \lesssim \lambda|K|. \tag{7.14}$$

We estimate one more tail term. For $n \geq 1$ let $\mathcal{I}^n(L) = \{I \in \mathcal{I}(L) : \ell(I) = 2^{-n}\ell(L)\}$. For each $I \in \mathcal{I}^n(L)$, let z_I be the least nonnegative integer z such that $(I \pm z\ell(I)) \cap (\mathbb{R} \setminus L) \neq \varnothing$. As $3I \subset L$, we have $z_I \geq 1$. Furthermore for each integer $z \geq 1$, there are at most two intervals $I \in \mathcal{I}^n(L)$ with $z_I = z$. As for $x \in \mathbb{R} \setminus L$ we have $\operatorname{dist}(x, I) \geq \ell(I)$, there holds

$$\begin{aligned}\mathbf{1}_{\mathbb{R}\setminus L}(x)|S_I f(x)|_X &\lesssim \left\langle \tfrac{\operatorname{dist}(x,I)}{\ell(I)} \right\rangle^{-100} \sup_{s\gtrsim\ell(I)} \frac{1}{|B_s(x)|} \int_{B_s(x)} F \lesssim \left\langle \tfrac{\operatorname{dist}(x,I)}{\ell(I)} \right\rangle^{-99} \inf_I \mathrm{M}F \\ &\lesssim \left\langle \tfrac{\operatorname{dist}(x,I)}{\ell(I)} \right\rangle^{-90} z_I^{-9}\lambda.\end{aligned}$$

Integrating over $\mathbb{R} \setminus L$ the last display for each I-summand, we have

$$\sum_{n\geq1} \sum_{I\in\mathcal{I}^n(L)} \left\|\mathbf{1}_{\mathbb{R}\setminus L} S_I f\right\|_{L^1(\mathbb{R};\mathcal{X})} \lesssim \lambda \sum_{n\geq1} \sum_{I\in\mathcal{I}^n(L)} z_I^{-9}|I| \lesssim \lambda \sum_{n\geq1} \sum_{z\geq1} z^{-9} 2^{-n}|L| \lesssim \lambda|L|,$$

whence

$$\sum_{L\in\mathcal{L}} \sum_{I\in\mathcal{I}(L)} \left\|\mathbf{1}_{\mathbb{R}\setminus L} S_I f\right\|_{L^1(\mathbb{R};\mathcal{X})} \lesssim \lambda \sum_{L\in\mathcal{L}} |L| \lesssim \lambda|K|. \tag{7.15}$$

We are left to estimate the main term. Using disjointness of the supports of the summands below

$$\left\| \sum_{L\in\mathcal{L}} \mathbf{1}_L \sum_{I\in\mathcal{I}(L)} S_I f \right\|_{L^{1,\infty}(\mathbb{R};\mathcal{X})} \leq \sum_{L\in\mathcal{L}} \left\| \sum_{I\in\mathcal{I}(L)} S_I f \right\|_{L^{1,\infty}(\mathbb{R};\mathcal{X})}. \tag{7.16}$$

To estimate each summand on the right hand side of the last display, we use the localization trick. Referring to (7.1), set $\bar{S}_I g := S_I(\gamma_L g)$. We then have

$$\begin{aligned}\sum_{L\in\mathcal{L}} \left\| \sum_{I\in\mathcal{I}(L)} S_I f \right\|_{L^{1,\infty}(\mathbb{R};\mathcal{X})} &= \sum_{L\in\mathcal{L}} \left\| \sum_{I\in\mathcal{I}(L)} \bar{S}_I(\gamma_L^{-1} f) \right\|_{L^{1,\infty}(\mathbb{R};\mathcal{X})} \\ &\lesssim \sum_{L\in\mathcal{L}} \left\|\gamma_L^{-1} f\right\|_{L^1(\mathbb{R};\mathcal{X})} \lesssim \sum_{L\in\mathcal{L}} |L| \inf_L \mathrm{M}F \lesssim \lambda|K|\end{aligned} \tag{7.17}$$

as each $\bar{S}_I \in \mathbb{S}_P$ where $P \in \mathsf{T}$ is the unique tri-tile with $I_P = I$, and therefore each L-th summand on the right hand side of the first line is a Calderón-Zygmund operator. We achieve (7.7) by putting together (7.10), (7.11), (7.13), (7.14), (7.15), (7.16) and (7.17). The proof of the lemma is then complete. □

Appendix A. Details on the space-frequency analysis of (1.1)

In this Appendix, we collect a few well-known procedures and results of space-frequency analysis that we have used throughout the article. The frequency discretization of (1.3) presented in Subsection A.1 is classical and reflects the treatment of [29] and its expanded version [35]. In the subsequent paragraphs, we describe explicitly the approximate order relations, borrowed from [29], characterizing the rank 1 collections of tri-tiles defined in Subsection 2.3, and explain the spatial discretization leading to the model sums (2.10). The notion of tree given in Subsection 3.1 is different from that of [29]: in Subsection A.4 we explain how the k-trees in [29] fit into our definition and also prove the lacunarity claim (3.5). Finally, in Subsection A.5 we develop, in parallel with the treatment in [30], the properties g1. to g3. of greedily selected trees we have used to construct phase-space projections.

A.1. **Frequency discretization.** Recall that $\Gamma' = \Gamma \cap \beta^{\perp}$ is the singular line of the multiplier m satisfying (1.3). We may extend m from $\Gamma \setminus \Gamma'$ to all of $\mathbb{R}^3 \setminus \Gamma'$ so that

$$\sup_{\xi \in \mathbb{R}^3 \setminus \Gamma'} \Big(\operatorname{dist}(\xi, \Gamma') \Big)^{\alpha} \big| \partial_\alpha m(\xi) \big| \lesssim_\alpha 1. \tag{A.1}$$

For $j \in \mathbb{Z}$, let $\mathcal{D}_j$ be the collection of cubes in $\mathbb{R}^3$ whose sides have length 2^j and whose centers lie on the lattice $2^{j-10}\mathbb{Z}^3$, and let $\mathcal{D} = \bigcup_{j \in \mathbb{Z}} \mathcal{D}_j$. Let $\mathbf{Q} \subset \mathcal{D}$ be a Whitney cover of $\mathbb{R}^3 \setminus \Gamma'$, namely $Q \in \mathbf{Q}$ if

$$Q \in \mathcal{D}, \qquad \operatorname{dist}(Q, \Gamma') \sim K\ell(Q), \tag{A.2}$$

where K is a large constant to be chosen later and the hidden constants in $\sim$ are absolute. As $\{\frac{1}{10}Q : Q \in \mathbf{Q}\}$ is a finitely overlapping Whitney cover of $\mathbb{R}^3 \setminus \Gamma'$, we may decompose

$$m = \sum_{Q \in \mathbf{Q}} m_Q,$$

where each m_Q is supported on $\frac{Q}{2}$ and such that (A.1) holds for $m = m_Q$. Hence, we have

$$\Lambda_m(f_1, f_2, f_3) = \sum_{Q \in \mathbf{Q}} \Lambda_{m_Q}(f_1, f_2, f_3).$$

By expanding each m_Q into its triple Fourier series on $Q = Q_1 \times Q_2 \times Q_3$ and using the rapid decay of its Fourier coefficients originating from (A.1), we learn that Λ_m belongs to the convex hull of forms

$$\Lambda(f_1, f_2, f_3) = \sum_{Q \in \mathbf{Q}} \int_{\mathbb{R}^3} \Big(\prod_{k=1}^{3} \widehat{f_k}(\xi_k) m_{Q_k}(\xi_k) \Big) \, \mathrm{d}\xi \tag{A.3}$$

with $m_{Q_k} \in M_{Q_k}$, $k = 1, 2, 3$. We now define $\mathbb{P}_1$ as the collection of all tri-tiles P whose frequency cube $Q_P \in \mathbf{Q}$ and whose spatial interval I_P is an interval from the standard dyadic grid; of course $\ell(I_P)$ is constrained to be reciprocal to the sidelength of Q_P. The next paragraph clarifies the rank properties of $\mathbb{P}_1$.

A.2. **Order relations.** We introduce a few approximate order relations between tiles. If t, τ are tiles, and K is the constant appearing in (A.2):

r1. $\tau < t$ if $I_\tau \subsetneq I_t$ and $\omega_t \subset 5\omega_\tau$;
r2. $\tau \le t$ if $\tau < t$ or $\tau = t$;
r3. $\tau \lesssim t$ if $I_\tau \subset I_t$ and $\omega_t \subset K\omega_\tau$;
r4. $\tau \lesssim' t$ if $\tau \lesssim t$ but it is not true that $\tau \le t$.

To shed light on r4. we remark that whenever $\tau \lesssim' t$, necessarily $3\omega_t \cap 3\omega_\tau = \varnothing$. We find useful to induce through r1. to r4. similar relations on tri-tiles $P \in \mathbb{P}_1$, as follows. If $P^a, P^b \in \mathbb{P}_1$ and $k = 1, 2, 3$ we say that

r5. $P^b \le_k P^a$ if $P^b_k \le P^a_k$;
r6. $P^b \lesssim_k P^a$ if $P^b_k \lesssim P^a_k$;
r7. $P^b \lesssim'_k P^a$ if $P^b_k \lesssim' P^a_k$.

Then [29, Lemma 6.2] may be summarized in the following two properties of $\mathbb{P}_1$. Firstly, whenever $P^a, P^b \in \mathbb{P}_1$

r8. $\mathrm{dist}(Q_P, \Gamma) \le 2\ell(Q_P)$;
r9. $P^b \le_k P^a$ for some $k \in \{1, 2, 3\} \implies P^b \lesssim_\kappa P^a$ for all $\kappa = 1, 2, 3$.

If in addition $\ell(I_{P^b}) \le 2^{-J}\ell(I_{P^a})$, we have

r10. $P^b \le_k P^a$ for some $k \in \{1, 2, 3\} \implies P^b \lesssim'_\kappa P^a$ for all $\kappa \in \{1, 2, 3\} \setminus k$

provided K is chosen sufficiently large depending on Δ_β from (1.2) and J is chosen such that $2^J \ge K^{10}$.

A.3. **Spatial discretization.** Following [30, Section 3], and referring to (2.4) for the definitions of χ_{I_P} we decompose any given form Λ appearing in (A.3) as

$$\Lambda(f_1, f_2, f_3) = \sum_{Q \in \mathbf{Q}} \sum_{\substack{P \in \mathbb{P}_1 \\ Q_P = Q}} \int_{\mathbb{R}} \prod_{k=1}^{3} \chi_{I_P}(x) T_{m_Q} f_k(x)\, \mathrm{d}x.$$

The discussion in Subsection A.2 shows that $\mathbb{P}_1$ may be decomposed into $O(2^{8J})$ rank 1 collections as above and ultimately reduces the proof of estimates for Λ_m from (1.1) to corresponding bounds for the model sums (2.10).

A.4. **Trees.** Let us fix a rank 1 collection $\mathbb{P}$. We have specified our notion of tree T in Subsection 3.1. Our definition of tree is less restrictive than the corresponding notion of k-tree in [29, Section 7]. The aim of this subsection is to expound this relationship, and subsequently to prove the claim leading to Remark 3.4.

First of all, let $k = 1, 2, 3$. We quote from [29] that a k-tree $\mathsf{T} \subset \mathbb{P}$ with top data the tri-tile $P_\mathsf{T} \in \mathbb{P}$ is a collection of tri-tiles satisfying $P \le_k P_\mathsf{T}$. We learn right away from r9. that $P \lesssim_\kappa P_\mathsf{T}$ for all $\kappa = 1, 2, 3$. In particular $Q_{P_\mathsf{T}} \subset KQ_P$ for all $P \in \mathsf{T}$ and $CKQ_{P_\mathsf{T}} \cap \Gamma' \neq \varnothing$ because of assumption a. of rank 1 collections and (A.2). Therefore, letting $I_\mathsf{T} = I_{P_\mathsf{T}}$, $\xi_\mathsf{T} \in \Gamma' \cap CKQ_{P_\mathsf{T}}$, we see at once that

$$I_P \subset I_\mathsf{T}, \qquad (\xi_\mathsf{T})_\kappa \in \overline{\omega_{P_\kappa}} \qquad \forall P \in \mathsf{T},\ \kappa = 1, 2, 3$$

as $\overline{\omega_{P_\kappa}}$ contains the dilate $K^4\omega_{P_\kappa}$. Hence, T is a tree with top data $(I_\mathsf{T}, \xi_\mathsf{T})$ according to the definition on Subsection 3.1.

In the same spirit, we move to the verification of Remark 3.4 by construction of a suitable splitting (3.5).

Let $\mathsf{T}_{\{2,3\}} = \{P \in \mathsf{T} : (\xi_\mathsf{T})_1 \in 3\omega_{P_1}\}$. Then for any $P, P' \in \mathsf{T}_{\{2,3\}}$ we have $3\omega_{P_1} \cap 3\omega_{P'_1} \neq \emptyset$. If β, β' are elements of $\{2\omega_{P_2} : P \in \mathsf{T}_{\{2,3\}}\}$, take $P, P' \in \mathsf{T}_{\{2,3\}}$ with $\beta = 2\omega_{P_2}, \beta' = 2\omega_{P'_2}$ and $\ell(\omega_{P_k}) \leq \ell(\omega_{P'_k})$. By g1. we must have $\ell(\omega_{P_k}) \leq 2^{-J}\ell(\omega_{P'_k})$, and the fact that $(\xi_\mathsf{T})_1 \in 3\omega_{P_1} \cap 3\omega_{P'_1}$ forces $3\omega_{P_1} \subset 5\omega_{P'_1}$. Let I'' be the unique J-dyadic interval with $\ell(I'') = \ell(I_P)$ and $I'' \supset I_{P'}$ and P'' be the unique tri-tile in $\mathbb{P}_1$ with $Q_{P''} = Q_P$ and $I_{P''} = I$. The above discussion shows that $P' \leq_1 P''$. Finally, by r10. we learn that that $P' \lesssim'_2 P''$. In particular $\omega_{P_2} = \omega_{P''_2} \not\subset 5\omega_{P'_2}$, which by virtue of $\ell(\omega_{P_2}) \leq 2^{-J}\ell(\omega_{P'_2})$ entails $\beta \cap \beta' = \emptyset$. Thus (3.2) holds true for $\mathsf{T} = \mathsf{T}_{\{2,3\}}$ and $k = 2$. An identical argument verifies (3.2) for $\mathsf{T} = \mathsf{T}_{\{2,3\}}$ and $k = 3$.

We now set $\mathsf{T}' := \mathsf{T} \setminus \mathsf{T}_{\{2,3\}}$, and define $\mathsf{T}_{\{1,3\}} = \{P \in \mathsf{T}' : (\xi_\mathsf{T})_2 \in 3\omega_{P_2}\}$. Repeating the argument for $\mathsf{T}_{\{2,3\}}$ shows that $\mathsf{T}_{\{1,3\}}$ enjoys (3.2) for $k = 1, 3$. We then set $\mathsf{T}'' := \mathsf{T}' \setminus \mathsf{T}_{\{1,3\}}$, and $\mathsf{T}_{\{1,2\}} = \{P \in \mathsf{T}'' : (\xi_\mathsf{T})_3 \in 3\omega_{P_3}\}$. Once again, $\mathsf{T}_{\{1,2\}}$ enjoys (3.2) for $k = 1, 2$. We are left with checking that $\mathsf{T}_{\{1,2,3\}} := \mathsf{T}'' \setminus \mathsf{T}_{\{1,3\}}$ enjoys (3.2) for $k = 1, 2, 3$. By construction,

$$(\xi_\mathsf{T})_k \in \overline{\omega_{P_k}} \setminus 3\omega_{P_k} \qquad \forall k = 1,2,3,\ P \in \mathsf{T}_{\{1,2,3\}}$$

which yields

$$\ell(\omega_{P_k}) \leq \operatorname{dist}(\omega_{P_k}, (\xi_\mathsf{T})_k) \leq K^6 \ell(\omega_{P_k}) \qquad \forall k = 1,2,3,\ P \in \mathsf{T}_{\{1,2,3\}}.$$

In particular $2\omega_{P_k} \subset \{2^{-1}\ell(\omega_{P_k}) \leq |\xi - (\xi_\mathsf{T})_k| \leq K^6\ell(\omega_{P_k})\}$, and g1. and separation of scales by $2^J > K^{10}$ leads to (3.2) for $\mathsf{T} = \mathsf{T}_{\{1,2,3\}}$ and $k = 1, 2, 3$. This completes the verification of Remark 3.4.

A.5. **Geometry of trees.** In this subsection we record some results on the geometry of trees described in [30, Section 4] that we have repeately used. We begin with the definition of maximal tree and of greedy selection process.

Given a top data (I, ξ) we call

$$\mathsf{T} = \mathsf{T}(I, \xi, \mathbb{P}) := \{P \in \mathbb{P} : I_P \subset I,\ \xi_k \in \overline{\omega_{P_k}} \quad \forall k = 1,2,3\}$$

is the maximal tree in $\mathbb{P}$ with top data (I, ξ). Note that if $\mathsf{T}' \subset \mathbb{P}$ has top data (I, ξ), then satisfies $\mathsf{T}' \subset \mathsf{T}(I, \xi, \mathbb{P})$.

If $\mathbb{P}$ is a finite subset of a collection of rank 1 tri-tiles $\mathbb{P}_1$ as specified in (2.10), a selection process consists of choosing a tree T_1 from $\mathbb{P}$ at step 1, and for $j \geq 1$, choosing a tree T_{j+1} from $\mathbb{P} \setminus \bigcup_{u=1}^{j} \mathsf{T}_u$. We say that the selection process is *greedy* if at each step j, the selected tree T_j is the maximal tree in $\mathbb{P} \setminus \bigcup_{u=1}^{j-1} \mathsf{T}_u$ for some top data (I, ξ), namely $\mathsf{T}_j = \mathsf{T}(I, \xi, \mathbb{P} \setminus \bigcup_{u=1}^{j-1} \mathsf{T}_u)$. Note that the selection procedures used in Subsection 6.4 for the proof of Lemma 3.13 are greedy.

A.4. **Lemma.** [30, Lemma 4.4] *Suppose that* T *is a tree constructed during a greedy selection process. Then property g1. holds. That is, the* frequency localization sets $\mathbf{Q}_\mathsf{T} = \{Q_P : P \in \mathsf{T}\}$ *are such that*

$$Q, Q' \in \mathbf{Q}_\mathsf{T},\ \ell(Q) = \ell(Q') \implies Q = Q'.$$

Proof. We recall the notation $\mathsf{T}(j) = \{P \in \mathsf{T} : \ell(Q_P) = 2^{Jj}\}$. We have that

$$2^{Jj} \leq \ell(\overline{\omega_{P_k}}) \leq K^5 2^{Jj} < 2^{(j+1)J} \qquad \forall\, k = 1,2,3,\ \forall P \in \mathsf{T}(j).$$

As the intervals $\overline{\omega_{P_k}}$ come from a J-separated grid, $\ell(\overline{\omega_{P_k}})$ is constant as $P \in \mathsf{T}(j)$ varies. As $(\xi_{\mathsf{T}})_k \in \overline{\omega_{P_k}}$ for all $P \in \mathsf{T}(j)$, this means that $\overline{\omega_{P_k}} = \overline{\omega_{P'_k}}$ for all $P, P' \in \mathsf{T}(j)$. As the intervals ω_{P_k}, $P \in \mathsf{T}(j)$ are $2^{(j+1)J+10}$-separated and $\overline{\omega_{P_k}} \subset K^6\omega_{P_k} \subset 2^J\omega_{P_k}$, we learn that $\omega_{P_k} = \omega_{P'_k}$ for all $P, P' \in \mathsf{T}(j)$ and $k = 1, 2, 3$. That is, $Q_P = Q_{P'}$ for all $P, P' \in \mathsf{T}(j)$ which was our claim. □

A.5. **Lemma.** [30, Lemma 4.7] *Suppose that* T *is a tree constructed during a greedy selection process. Then property g2. holds, that is, the* spatial localization sets

$$E_{Q,\mathsf{T}} = \bigcup\{I_P : P \in \mathsf{T}, Q_P = Q\}, \qquad Q \in \mathbf{Q}_{\mathsf{T}}$$

satisfy

$$Q, Q' \in \mathbf{Q}_{\mathsf{T}},\ \ell(Q) \le \ell(Q') \implies E_{Q,\mathsf{T}} \supset E_{Q',\mathsf{T}}.$$

Proof. By Lemma A.4 we may clearly assume $\ell(Q) < \ell(Q')$. We argue by contradiction. Suppose that there exists $P' \in \mathsf{T}$ with $Q_{P'} = Q'$ and $I_{P'} \not\subset E_{Q,\mathsf{T}}$. Pick any $P \in \mathsf{T}$ with $Q_P = Q$. Let P'' be the unique tri-tile with $Q_P = Q_{P''}$ and $I_P \subset I_{P''}$. By dyadicity, $\ell(I_P) \le \ell(I_{\mathsf{T}})$ and the fact that $I'_P \subset I_{P''} \cap I_{\mathsf{T}}$, we have $I_{P''} \subset I_{\mathsf{T}}$, which, together with $Q_P = Q_{P''}$, qualifies P'' to be in the maximal tree with top data $(I_{\mathsf{T}}, \xi_{\mathsf{T}})$. This means that P'' was selected in a tree T'' with top data $(I_{\mathsf{T}''}, \xi_{\mathsf{T}''})$ at an earlier stage than T. However $I_{P'} \subset I_{\mathsf{T}''}$, and we also have $\overline{\omega_{P''_k}} = \overline{\omega_{P_k}} \subset \overline{\omega_{P'_k}}$, for all k, as $\overline{\omega_{P_k}}, \overline{\omega_{P'_k}}$ are intervals from a dyadic grid both containing $(\xi_{\mathsf{T}})_k$. Therefore P' qualified to be in the maximal tree with top data $(I_{\mathsf{T}''}, \xi_{\mathsf{T}''})$, and would not have been available at the time of T being selected. Contradiction. □

A consequence of the nesting property of Lemma A.5 is the following estimate: if T is a tree constructed during a greedy selection process and Q^j is the unique $Q \in \mathbf{Q}_{\mathsf{T}}$ with $\ell(Q) = 2^{Jj}$,

$$\sum_{j \in \mathbf{j}_{\mathsf{T}}} 2^{-Jj} \# \partial E_{Q^j,\mathsf{T}} \lesssim |I_{\mathsf{T}}|. \tag{A.6}$$

The proof (as well as the statement, in fact) is identical to that of [30, Lemma 4.8], and for this reason we omit it. We continue with further notation to define a suitable smoothed out version of the sets $E_{Q,\mathsf{T}}$. Let T be a tree as defined in Subsection 3.1. Denote by $\mathbf{I}_{\mathsf{T}}$ the collection of J-dyadic intervals contained in I_T with the properties:

1) $3I$ does not contain any I_P with $P \in \mathsf{T}$;
2) The J-dyadic parent of I fails 1).

Define $\tilde{E}_j = \cup\{I \in \mathbf{I}_{\mathsf{T}} : \ell(I) < 2^{-Jj}\}$. Obviously $\tilde{E}_{j+1} \subset \tilde{E}_j$.

In the lemmata below, we also assume that T is a tree which has been constructed during a greedy selection process. As the proofs involve the spatial components only, they may be read word by word from the indicated reference and we do not repeat them.

A.7. **Lemma.** [30, Lemma 4.10] *Any two neighboring intervals in* $\mathbf{I}_{\mathsf{T}}$ *differ by at most a factor* 2^J *in length. Further, the set* $\tilde{E}_j$ *is a union of dyadic intervals of length* 2^{-Jj} *and contains* $E_{Q^j,\mathsf{T}}$ *if* $j \in \mathbf{j}_{\mathsf{T}}$.

A.8. **Lemma.** [30, Lemma 4.11] *If* I_0 *is a J-dyadic interval of length* 2^{-Jj_0} *such that* $3I_0 \cap \tilde{E}_{j_0} \ne \emptyset$, *then there is a tri-tile* $P \in \mathsf{T}$ *with* $|I_P| \le |I_0|$ *such that* $I_P \subset 10I_0$.

A.9. **Lemma.** [30, Lemma 4.12] *Let T be any tree. For each $j \in \mathbf{j}_\mathsf{T}$, let Ω_j be the collection of connected components of $\tilde{E}_j$. Then there holds*

$$\sum_{j\in\mathbf{j}_\mathsf{T}} 2^{-Jj}\#\Omega_j \lesssim |I_\mathsf{T}|. \tag{A.10}$$

For each $I \in \Omega_j$, let x_I^ℓ and x_I^r denote the left and right endpoints of I, and let I_j^ℓ and I_j^r denote the intervals

$$I_j^\ell := (x_I^\ell - 2^{-Jj-1}, x_I^\ell - 2^{-Jj-2}),\ I_j^r := (x_I^r + 2^{-Jj-2}, x_I^r + 2^{-Jj-1}).$$

Then the intervals I_j^ℓ are disjoint as j varies in the integers with $2^{-Jj} \leq |I_\mathsf{T}|$ and I varies in Ω_j. Moreover, if I_j^ℓ is an interval in the above collection, then the distance to the next interval $I'^\ell_{j'}$ is at least $2^{-J(j+2)}$. Similar statements hold for the I_j^r.

We conclude this Appendix with the definition of μ_j, which appears first in Proposition 4.3. For a fixed tree T, after construction of the sets $\tilde{E}_j$, set

$$\mu_j(x) := \sum_{j'\geq 0} 2^{-|j'-j|/100} \sum_{y\in\partial\tilde{E}_{j'}} (1 + 2^{Jj'}|x-y|)^{-100}. \tag{A.11}$$

The estimate involving $\sum\int \mu_j$ is proved in (4.16).

References

[1] Alex Amenta, Emiel Lorist, and Mark Veraar, *Fourier multipliers in Banach function spaces with UMD concavifications*, Trans. Amer. Math. Soc. **371** (2019), no. 7, 4837–4868. MR3934469

[2] Alex Amenta and Gennady Uraltsev, *The bilinear Hilbert transform on UMD spaces*, preprint arXiv:1909.06416 (2019).

[3] Alex Amenta and Gennady Uraltsev, *Banach-valued modulation invariant Carleson embeddings and outer-L^p spaces: the Walsh case*, J. Fourier Anal. Appl. **26** (2020), no. 4, Paper No. 53, 54. MR4115625

[4] Cristina Benea and Camil Muscalu, *Multiple vector-valued inequalities via the helicoidal method*, Anal. PDE **9** (2016), no. 8, 1931–1988. MR3599522

[5] Jean Bourgain, *Some remarks on Banach spaces in which martingale difference sequences are unconditional*, Ark. Mat. **21** (1983), no. 2, 163–168. MR727340

[6] Donald L. Burkholder, *A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions*, Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), 1983, pp. 270–286. MR730072

[7] Alejandro J. Castro and Tuomas Hytönen, *Bounds for partial derivatives: necessity of UMD and sharp constants*, Math. Z. **282** (2016), no. 3-4, 635–650. MR3473635

[8] David Cruz-Uribe and José María Martell, *Limited range multilinear extrapolation with applications to the bilinear Hilbert transform*, Math. Ann. **371** (2018), no. 1-2, 615–653. MR3788859

[9] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou, *Domination of multilinear singular integrals by positive sparse forms*, J. Lond. Math. Soc. (2) **98** (2018), no. 2, 369–392. MR3873113

[10] Ciprian Demeter and Christoph Thiele, *On the two-dimensional bilinear Hilbert transform*, Amer. J. Math. **132** (2010), no. 1, 201–256. MR2597511

[11] Francesco Di Plinio, Kangwei Li, Henri Martikainen, and Emil Vuorinen, *Multilinear singular integrals on non-commutative L^p spaces*, preprint, arXiv:1905.02139 (2019).

[12] Francesco Di Plinio, Kangwei Li, Henri Martikainen, and Emil Vuorinen, *Multilinear operator-valued Calderón-Zygmund theory*, J. Funct. Anal. **279** (2020), no. 8, 108666, 62. MR4109091

[13] Francesco Di Plinio and Yumeng Ou, *Banach-valued multilinear singular integrals*, Indiana Univ. Math. J. **67** (2018), no. 5, 1711–1763. MR3875242

[14] Tadeusz Figiel, *Singular integral operators: a martingale approach*, Geometry of Banach spaces (Strobl, 1989), 1990, pp. 95–110. MR1110189

[15] Stefan Geiss, Stephen Montgomery-Smith, and Eero Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. **362** (2010), no. 2, 553–575. MR2551497

[16] Timo S. Hänninen and Tuomas Hytönen, *Operator-valued dyadic shifts and the* $T(1)$ *theorem*, Monatsh. Math. **180** (2016), no. 2, 213–253. MR3502626

[17] Tuomas Hytönen, *On operator-multipliers for mixed-norm* $L^{\bar{p}}$ *spaces*, Arch. Math. (Basel) **85** (2005), no. 2, 151–155. MR2161804 (2006c:42011)

[18] Tuomas Hytönen and Michael T. Lacey, *Pointwise convergence of vector-valued Fourier series*, Math. Ann. **357** (2013), no. 4, 1329–1361. MR3124934

[19] Tuomas Hytönen and Michael T. Lacey, *Pointwise convergence of Walsh-Fourier series of vector-valued functions*, Math. Res. Lett. **25** (2018), no. 2, 561–580. MR3826835

[20] Tuomas Hytönen, Michael T. Lacey, and Ioannis Parissis, *The vector valued quartile operator*, Collect. Math. **64** (2013), no. 3, 427–454. MR3084406

[21] Tuomas Hytönen, Michael T. Lacey, and Ioannis Parissis, *A variation norm Carleson theorem for vector-valued Walsh-Fourier series*, Rev. Mat. Iberoam. **30** (2014), no. 3, 979–1014. MR3254998

[22] Tuomas Hytönen and Pierre Portal, *Vector-valued multiparameter singular integrals and pseudodifferential operators*, Adv. Math. **217** (2008), no. 2, 519–536. MR2370274 (2009e:42027)

[23] Tuomas Hytönen, Jan van Neerven, and Pierre Portal, *Conical square function estimates in UMD Banach spaces and applications to* H^∞*-functional calculi*, J. Anal. Math. **106** (2008), 317–351. MR2448989

[24] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 63, Springer, Cham, 2016. MR3617205

[25] Marius Junge and Quanhua Xu, *Representation of certain homogeneous Hilbertian operator spaces and applications*, Invent. Math. **179** (2010), no. 1, 75–118. MR2563760

[26] Nigel Kalton and Lutz Weis, *The* H^∞*-calculus and sums of closed operators*, Math. Ann. **321** (2001), no. 2, 319–345. MR1866491

[27] Michael Lacey and Christoph Thiele, L^p *estimates on the bilinear Hilbert transform for* $2 < p < \infty$, Ann. of Math. (2) **146** (1997), no. 3, 693–724. MR1491450

[28] Michael Lacey and Christoph Thiele, *On Calderón's conjecture*, Ann. of Math. (2) **149** (1999), no. 2, 475–496. MR1689336

[29] Camil Muscalu, Terence Tao, and Christoph Thiele, *Multi-linear operators given by singular multipliers*, J. Amer. Math. Soc. **15** (2002), no. 2, 469–496. MR1887641

[30] Camil Muscalu, Terence Tao, and Christoph Thiele, *Uniform estimates on multi-linear operators with modulation symmetry*, J. Anal. Math. **88** (2002), 255–309. Dedicated to the memory of Tom Wolff. MR1979774

[31] Gilles Pisier and Quanhua Xu, *Non-commutative* L^p*-spaces*, Handbook of the geometry of Banach spaces, Vol. 2, 2003, pp. 1459–1517. MR1999201

[32] José L. Rubio de Francia, *Martingale and integral transforms of Banach space valued functions*, Probability and Banach spaces (Zaragoza, 1985), 1986, pp. 195–222. MR875011

[33] Prabath Silva, *Vector-valued inequalities for families of bilinear Hilbert transforms and applications to bi-parameter problems*, J. Lond. Math. Soc. (2) **90** (2014), no. 3, 695–724. MR3291796

[34] Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR1232192 (95c:42002)

[35] Christoph Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics, vol. 105, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. MR2199086

[36] Lutz Weis, *Operator-valued Fourier multiplier theorems and maximal* L_p*-regularity*, Math. Ann. **319** (2001), no. 4, 735–758. MR1825406