

A MODULATION INVARIANT CARLESON EMBEDDING THEOREM OUTSIDE LOCAL L^2

FRANCESCO DI PLINIO AND YUMENG OU

ABSTRACT. The article [6] develops a theory of Carleson embeddings in outer L^p spaces for the wave packet transform

$$F_\phi(f)(u, t, \eta) = \int f(x) e^{i\eta(u-x)} \phi\left(\frac{u-x}{t}\right) \frac{dx}{t}, \quad (u, t, \eta) \in \mathbb{R} \times (0, \infty) \times \mathbb{R}$$

of functions $f \in L^p(\mathbb{R})$, in the $2 \leq p \leq \infty$ range referred to as local L^2 . In this article, we formulate a suitable extension of this theory to exponents $1 < p < 2$, answering a question posed in [6]. The proof of our main embedding theorem involves a refined multi-frequency Calderón-Zygmund decomposition in the vein of [5, 15]. We apply our embedding theorem to recover the full known range of L^p estimates for the bilinear Hilbert transforms [10] without reducing to discrete model sums or appealing to generalized restricted weak-type interpolation.

1. INTRODUCTION

We are concerned with the *continuous wave packet transform* of $f : \mathbb{R} \rightarrow \mathbb{C}$

$$(1.1) \quad F_\phi(f)(u, t, \eta) = f * \phi_{t,\eta}(u), \quad (u, t, \eta) \in \mathcal{Z} := \mathbb{R} \times (0, \infty) \times \mathbb{R}$$

where, for $t > 0$ and $\eta \in \mathbb{R}$,

$$\phi_{t,\eta} = \text{Mod}_\eta \text{Dil}_t^1 \phi, \quad \phi_{t,\eta}(x) := e^{i\eta x} \frac{1}{t} \phi\left(\frac{x}{t}\right)$$

is the *wave packet* at scale t and frequency η , and ϕ is a nondegenerate real valued Schwartz function with compact frequency support. The coefficients $F_\phi(f)$, whose arguments parametrize the group of symmetries of the class of modulation invariant singular integrals, provide an efficient description of the action of operators from this class. However, due to the overdetermination of the family $\{\phi_{t,\eta} : t \in (0, \infty), \eta \in \mathbb{R}\}$, an effective control of the coefficient norms of $F_\phi(f)$ in terms of the norms of f cannot be obtained via immediate orthogonality considerations, unlike the case of the *wavelet transform* of f where the modulation parameter η is fixed.

A fundamental modulation invariant singular integral is the (family of) trilinear form(s) on Schwartz functions

$$\Lambda_{\vec{\beta}}(f_1, f_2, f_3) = \int_{\mathbb{R}} \text{p.v.} \int_{\mathbb{R}} f_1(x - \beta_1 t) f_2(x - \beta_2 t) f_3(x - \beta_3 t) \frac{dt}{t} dx$$

2000 *Mathematics Subject Classification.* Primary: 42B20. Secondary: 42B25.

Key words and phrases. Wave packet transform, Carleson embedding theorems, bilinear Hilbert transform, outer measure theory.

FDP was partially supported by the National Science Foundation under the grant NSF-DMS-1500449. YO was partially supported by the National Science Foundation under the grant NSF-DMS 0901139.

parametrized by unit vectors $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$, perpendicular to $\vec{1} = (1, 1, 1)$ and non-degenerate in the sense that

$$(1.2) \quad \delta_0 = \inf_{j \neq k} |\beta_j - \beta_k| > 0.$$

The adjoint bilinear operators $T_{\vec{\beta}}$ to $\Lambda_{\vec{\beta}}$ are the *bilinear Hilbert transforms*, see for instance [9, 10] and the monograph [18]. As noted in [6], for a suitable choice of $\widehat{\text{supp } \phi}$, the form $\Lambda_{\vec{\beta}}$ is a nontrivial linear combination of the integral of the pointwise product with a trilinear form involving the wave packet transforms:

$$(1.3) \quad \begin{aligned} V_{\vec{\beta}}(f_1, f_2, f_3) &:= \int_{\mathcal{Z}} \left(\prod_{j=1}^3 G_j(u, t, \eta) \right) du dt d\eta, \\ G_j(u, t, \eta) &:= F_{\phi}(f_j)(u, t, \alpha_j \eta + \beta_j t^{-1}) \end{aligned}$$

where $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ is a unit vector orthogonal to $\vec{\beta}$ and $(1, 1, 1)$. Therefore, Hölder-type estimates for $\Lambda_{\vec{\beta}}$ are an immediate consequence of the corresponding bounds for $V_{\vec{\beta}}$. Throughout the article, we refer to exponents triples (p_1, p_2, p_3) with

$$-\infty < p_1, p_2, p_3 \leq \infty, \quad \sum_{j=1}^3 \frac{1}{p_j} = 1$$

as Hölder triples of exponents. The article [6] proves Hölder-type $L^{p_1} \times L^{p_2} \times L^{p_3}$ -bounds for the form $V_{\vec{\beta}}$ in the *local* L^2 range $2 < p_1, p_2, p_3 < \infty$. These estimates recover L^p bounds for the bilinear Hilbert transform of [9] in the same restricted range, bypassing the discretization procedures which are ubiquitous in the analysis of modulation invariant singular integrals; see [5, 9, 10, 14], the monographs [13, 18] and references therein.

The proof of [6] is articulated in two distinct and complementary steps. The first one is the application of an *outer Hölder inequality* (see Lemma 3.1) to bound the trilinear integral in (1.3) by the product of certain *outer* L^{p_j} -norms of each G_j , whose definition we postpone to Section 2. Here, we remark that this step does not rely in any way on G_j being the wave packet transform of f_j . On the contrary, the nature of the G_j is fundamental in the second and final step, which is the proof of an inequality of the type

$$(1.4) \quad \|G_j\|_{L^{p_j}(\mathcal{Z}, \sigma, s)} \leq C_{\vec{\beta}, p_j} \|f_j\|_{p_j}, \quad 2 < p_j \leq \infty$$

where the norm appearing on the left hand side is the same outer L^{p_j} -norm coming into play in the outer Hölder inequality. Estimates of this type are referred to as *Carleson embedding theorems*. The motivation is that the $p_j = \infty$ case of this inequality is essentially the same as

$$|F_{\phi}(f_j)(u, t, \eta)|^2 \frac{du dt}{t}$$

being a Carleson measure on $Z = \mathbb{R} \times (0, \infty)$ when $f_j \in L^{\infty}$, uniformly in $\eta \in \mathbb{R}$.

Carleson embeddings of the type (1.4) do not hold, at least in this form, for $1 < p_j < 2$. The main purpose of this article is to develop a suitable extension of (1.4) outside of the local L^2 range treated in [6], answering one of the questions posed in that article. A loose

description of our main result, which is precisely stated in Theorem 1, Section 2, is that for all $1 < p < 2$

$$\|\tilde{G}_j\|_{L^q(\mathcal{Z}, \sigma, s)} \leq C_{\tilde{\beta}, p, q} \|f_j\|_p, \quad p' < q \leq \infty$$

where $\tilde{G}_j$ is the restriction of G_j outside a suitable exceptional subset of $\mathcal{Z}$ depending only on the maximal p -averages of f_j . The proof involves a multi-frequency Calderón-Zygmund lemma which is an evolution of the approaches of [5, 15].

Via a simple modification of the outer Hölder inequality argument, the Carleson embeddings of Theorem 1 yield estimates for $V_{\tilde{\beta}}(f_1, f_2, f_3)$, and thus for $\Lambda_{\tilde{\beta}}(f_1, f_2, f_3)$ when f_3 is a suitably restricted subindicator, and f_1, f_2 are unrestricted functions, corresponding to a *direct proof* of the weak-type estimate

$$(1.5) \quad T_{\beta} : L^{p_1}(\mathbb{R}) \times L^{p_2}(\mathbb{R}) \rightarrow L^{\frac{p_1 p_2}{p_1 + p_2}, \infty}(\mathbb{R})$$

for the adjoint(s) to Λ_{β} . This is an improvement from the generalized restricted weak-type paradigm of [14], where all input functions in Λ_{β} are taken to be subindicator, resulting in the (weaker) version of (1.5) for subindicators. Our approach thus recovers, via bilinear weak-type interpolation rather than generalized restricted weak-type interpolation of e.g. [14] or [18, Chapter 3], the strong-type bound for the bilinear Hilbert transforms T_{β} in the full range

$$(1.6) \quad 1 < p_1, p_2 \leq \infty, \quad \frac{2}{3} < \frac{p_1 p_2}{p_1 + p_2} < \infty.$$

of expected exponents first established in [10]. A precise statement and proof can be found in Theorem 2, Section 3.

We expect Theorem 1, or suitable strengthenings thereof, to impact on other related questions. The first concerns uniform bounds for the family $\Lambda_{\tilde{\beta}}$ as the degeneracy parameter δ_0 from (1.2) approaches zero. In spite of the significant progress made in the articles [7, 11, 17], it is still open whether uniform bounds holds in the full expected range of exponents p_1, p_2 common to both the generic and degenerate case. The article [16] proves the full range of uniform bounds for a discrete model of $\Lambda_{\tilde{\beta}}$, relying on a multi-frequency decomposition argument in the simple discrete setting, where the bad part has trivial contribution. We do not explicitly track the dependence on δ in our estimates. However, the argument for Theorem 1 is currently ill-conditioned with respect to this parameter, and a more refined treatment is needed for uniform bounds.

The recent article [4] develops an outer L^p theory for functions taking values in UMD Banach spaces, which is then used for the proof of multilinear multiplier theorems in the UMD-valued setting. In view of possible applications to Banach-valued multilinear singular integrals with modulation invariance, it is reasonable to ask whether a form of Theorem 1 extends to UMD Banach space valued functions (or interpolation UMD, as in [8]). In an appendix, we sketch how to obtain such an extension for Hilbert space valued functions.

Notation. We normalize the Fourier transform of a Schwartz function $f \in \mathcal{S}(\mathbb{R})$ as

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x} dx.$$

Unless otherwise mentioned, the constants denoted by the letter C , as well as those implied by the almost inequality signs, are meant to be absolute. Their value may vary between

occurrences without explicit mention. With $C_{a_1, \dots, a_n}$, we denote a generic constant depending on the parameters $a_1, \dots, a_n$ which may also differ at each occurrence.

We write $c(I)$ for the center of the interval $I \subset \mathbb{R}$ and write $|I|$ for its length. For $\kappa > 0$, we denote by κI the interval centered at $c(I)$ with $|\kappa I| = \kappa |I|$. We say that J is a finitely overlapping collection of intervals if for all intervals $I \subset \mathbb{R}$

$$(1.7) \quad \sum_{J \subset I} |J| \leq C |I|.$$

We will write

$$M_p f(x) = \sup_{\substack{I \subset \mathbb{R} \text{ interval} \\ x \in I}} \left(\int_I |f(y)|^p \frac{dy}{|I|} \right)^{\frac{1}{p}}$$

for the p -Hardy-Littlewood maximal operator, for $1 \leq p < \infty$ and omit the subscript when $p = 1$. If $\phi \in \mathcal{S}(\mathbb{R})$, we will write $\mathcal{S}_N(\phi)$ for the Schwartz norms of ϕ of order N .

Acknowledgments. We are indebted to Christoph Thiele for suggesting the problem of formulating an outer measure embedding theorem for the wave packet transform outside local L^2 . We are also grateful to the anonymous referee for his generous work which greatly helped with improving the clarity of the presentation.

2. MAIN RESULTS

We begin by recalling from [6] the fundamental definitions concerning the outer L^p spaces appearing in the statement of our main Carleson embedding theorem.

2.1. Outer L^p spaces from tents. Let Z be a metric space and $\mathcal{T}$ be a subcollection of the Borel subsets of Z . Following the terminology of [6, Definition 2.1], we refer to any function $\sigma : \mathcal{T} \rightarrow [0, \infty]$ as a *premeasure*. The *outer measure* μ generated by σ is

$$Z \supset A \mapsto \mu(A) := \inf_{\mathcal{T}'} \sum_{T \in \mathcal{T}'} \sigma(T)$$

where the infimum is taken over all *countable* subcollections $\mathcal{T}'$ of $\mathcal{T}$ which cover the set A . Let $\mathcal{B}(Z)$ denote complex, Borel-measurable functions on Z . A *size* is a map

$$s : \mathcal{B}(Z) \rightarrow [0, \infty]^{\mathcal{T}}$$

that satisfies for any $F, G \in \mathcal{B}(Z)$ and $T \in \mathcal{T}$

$$(2.1) \quad \begin{aligned} |F| \leq |G| &\implies s(F)(T) \leq s(G)(T), \\ s(\lambda F)(T) &= |\lambda| s(F)(T), \quad \forall \lambda \in \mathbb{C}, \\ s(F + G)(T) &\leq c_s s(F)(T) + c_s s(G)(T), \quad \text{for some fixed } c_s \geq 1. \end{aligned}$$

We refer to the triple (Z, σ, s) as an *outer measure space*. Notice that the generating collection $\mathcal{T}$ is implicitly encoded in the notation as the domain of σ . We now introduce *outer L^p spaces*. Let us first define the *outer (essential) supremum* of $F \in \mathcal{B}(Z)$ over the Borel-measurable subset $E \subset Z$ as

$$\text{outsup}_E s(F) := \sup_{T \in \mathcal{T}} s(F \mathbf{1}_E)(T).$$

Then, for each $\lambda > 0$, the *super level measure* is defined by

$$\mu(s(F) > \lambda) := \inf \left\{ \mu(E) : E \subset Z \text{ Borel, } \operatorname{outsup}_{Z \setminus E} s(F) \leq \lambda \right\}.$$

Finally, for $0 < p < \infty$ and $f \in \mathcal{B}(Z)$ we set

$$\|F\|_{L^p(Z, \sigma, s)} := \left(\int_0^\infty p \lambda^{p-1} \mu(s(F) > \lambda) d\lambda \right)^{\frac{1}{p}}, \quad \|F\|_{L^{p, \infty}(Z, \sigma, s)} := \sup_{\lambda > 0} \lambda (\mu(s(F) > \lambda))^{\frac{1}{p}},$$

and for $p = \infty$

$$\|F\|_{L^{\infty, \infty}(Z, \sigma, s)} = \|F\|_{L^\infty(Z, \sigma, s)} := \operatorname{outsup}_Z s(F) = \sup_{T \in \mathcal{T}} s(F)(T).$$

We are concerned with two interrelated concrete examples of outer L^p spaces.

2.1.1. Tents. Our base metric space will be the upper half-space $Z = \mathbb{R} \times (0, \infty)$. Let $I \subset \mathbb{R}$ be an interval centered at $c(I)$. The *tent* over I is defined as

$$(2.2) \quad T(I) = \{(u, t) \in Z : 0 < t < |I|, |u - c(I)| < |I| - t\}.$$

An outer measure μ on Z , with $\{T(I) : I \subset \mathbb{R}\}$ as the generating collection is then defined via the premeasure $\sigma(T(I)) = |I|$. For $F \in \mathcal{B}(Z)$, we define the *sizes*

$$(2.3) \quad \begin{aligned} s_q(F)(T(I)) &:= \left(\frac{1}{|I|} \int_{T(I)} |F(u, t)|^q \frac{du dt}{t} \right)^{\frac{1}{q}}, \quad 0 < q < \infty, \\ s_\infty(F)(T(I)) &:= \sup_{(u, t) \in T(I)} |F(u, t)|, \end{aligned}$$

and denote by $L^p(Z, \sigma, s_q)$, $L^{p, \infty}(Z, \sigma, s_q)$ the corresponding strong and weak outer L^p space. Note that the case $p = \infty$ fits into the scale of *tent spaces* originating in [2].

2.1.2. Generalized tents. The outer measure space of generalized tents is based on the upper 3-half space $\mathcal{Z} = Z \times \mathbb{R}$. Let α be a set of parameters obeying the restrictions

$$(2.4) \quad \alpha = (\alpha, \beta, \delta), \quad 0 < \alpha \leq 1, \quad 0 \leq |\beta| < \frac{3}{4}, \quad 2^{-16} \delta_0 \leq \delta \leq 2^{-8} \delta_0,$$

where δ_0 is a small fixed constant. In fact, throughout the article we will work with δ_0 being the same as in (1.2). Let $I \subset \mathbb{R}$ be an interval, $T(I)$ be as in (2.2) and $\xi \in \mathbb{R}$. The corresponding *generalized tent* with parameters α , and its *lacunary part*, are defined by

$$\begin{aligned} T_\alpha(I, \xi) &:= \{(u, t, \eta) \in \mathcal{Z} : (u, t) \in T_\alpha(I), |\alpha(\eta - \xi) + \beta t^{-1}| \leq t^{-1}\}, \\ T_\alpha^\ell(I, \xi) &:= \{(u, t, \eta) \in T_\alpha(I, \xi) : t|\xi - \eta| > \delta\}. \end{aligned}$$

Outer measures ν_α on $\mathcal{Z}$, with $\{T_\alpha(I, \xi) : I \subset \mathbb{R}, \xi \in \mathbb{R}\}$ as the generating collection are then defined via the premeasure $\sigma_\alpha(T_\alpha(I, \xi)) = |I|$. For $F \in \mathcal{B}(\mathcal{Z})$, we define the sizes

$$(2.5) \quad \begin{aligned} s_2(F)(T(I, \xi)) &:= \left(\frac{1}{|I|} \int_{T_\alpha^\ell(I, \xi)} |F(u, t, \eta)|^2 du dt d\eta \right)^{\frac{1}{2}}, \\ s_\infty(F)(T(I, \xi)) &:= \sup_{(u, t, \eta) \in T_\alpha(I, \xi)} |F(u, t, \eta)|, \\ s &:= s_2 + s_\infty. \end{aligned}$$

We denote by $L^p(\mathcal{Z}, \sigma_\alpha, s)$, $L^{p,\infty}(\mathcal{Z}, \sigma_\alpha, s)$ and similarly for the other sizes, the corresponding strong and weak outer L^p spaces.

2.2. A Carleson embedding below local L^2 . Let $\phi \in \mathcal{S}(\mathbb{R})$ with $\text{supp } \widehat{\phi} \subset [-2^{-8}\delta, 2^8\delta]$, where δ refers to (2.4). In [6], it was proved that the wave packet transform $F_\phi(f)$ defined in (1.1) enjoys the local L^2 Carleson embeddings

$$(2.6) \quad \|F_\phi(f)\|_{L^q(\mathcal{Z}, \sigma_\alpha, s)} \leq C_{\alpha, \phi, q} \|f\|_q, \quad 2 < q \leq \infty,$$

$$(2.7) \quad \|F_\phi(f)\|_{L^{2,\infty}(\mathcal{Z}, \sigma_\alpha, s)} \leq C_{\alpha, \phi, 2} \|f\|_2,$$

with $C_{\alpha, \phi, q} = C_q \alpha^{-C} \mathcal{S}_C(\phi)$. Our main result is a suitable extension of (2.6) below local L^2 .

Theorem 1. *Let $f \in \mathcal{S}(\mathbb{R})$, $\lambda > 0$, $1 < p < 2$. Define*

$$(2.8) \quad \begin{aligned} I_{f, \lambda, p} &:= \text{maximal dyadic intervals } I \text{ s.t. } I \subset \{x \in \mathbb{R} : M_p(Mf)(x) > \lambda \|f\|_p\}, \\ E_{f, \lambda, p} &:= \bigcup_{I \in I_{f, \lambda, p}} T(3I) \times \mathbb{R} \subset \mathcal{Z}. \end{aligned}$$

Then

$$(2.9) \quad \left\| F_\phi(f) \mathbf{1}_{\mathcal{Z} \setminus E_{f, \lambda, p}} \right\|_{L^q(\mathcal{Z}, \sigma_\alpha, s)} \leq C_{p, q, \alpha, \phi} \lambda^{1 - \frac{p}{q}} \|f\|_p \quad \forall p' < q \leq \infty,$$

The constant $C_{p, q, \alpha, \phi}$ depends on p, q, α (polynomially) and ϕ only.

We use Theorem 1, coupled with the outer Hölder inequality, to extend the L^p estimates for $V_{\vec{\beta}}$ to exponents outside local L^2 .

Theorem 2. *Let (p_1, p_2, p_3) be a Hölder tuple with p_1, p_2 within the range (1.6). For all $A \subset \mathbb{R}$ of finite measure, and all $f_1, f_2 \in \mathcal{S}(\mathbb{R})$, there exists $\tilde{A} \subset A$ with $|\tilde{A}| \leq 2|\tilde{A}|$ and*

$$|V_{\vec{\beta}}(f_1, f_2, f_3)| \leq C_{\vec{\beta}, p_1, p_2} \|f_1\|_{p_1} \|f_2\|_{p_2} |A|^{\frac{1}{p_3}} \quad \forall f_3 \in \mathcal{S}(\mathbb{R}), |f_3| \leq \mathbf{1}_{\tilde{A}}.$$

Relying on the observation above (1.3), the same estimate of Theorem 2 holds for the forms $\Lambda_{\vec{\beta}}$. Equivalently, the adjoints $T_{\vec{\beta}}$ to $\Lambda_{\vec{\beta}}$ satisfy (1.5) for all p_1, p_2 within the range (1.6). Finally, since the range (1.6) is open, multilinear interpolation upgrades (1.5) to the strong-type estimate.

Remark 2.1. The proof of Theorem 1 contains a modulation invariant version of the Calderón-Zygmund decomposition. Loosely speaking, the argument proceeds by selecting the generalized tents on which $F_\phi(f)$ has large size, and decomposing f into the sum of a *good part* $g \in L^2$ and a *bad part* b . The bad part is devised to have moments of rather high order vanishing when integrated against the selected frequencies. This extra cancellation renders the size of $F_\phi(b)$ very small, forcing $F_\phi(g)$ to have large size. At this point, since $g \in L^2$, we get to use the weak- L^2 version of the Carleson embedding (2.7), or rather its proof. Unlike [5, 15] and the similar projection argument used in the Walsh context (see [3] and references therein), the frequencies on which f is projected are intrinsic to f itself.

2.3. Proofs and structure of the article. In the upcoming Section 3, we prove Theorem 2. The main steps of the proof of Theorem 1 are postponed to Section 7. The proof relies on the material of Section 6, which contains somewhat refined Carleson embedding theorems for the wave packet transform of f at a fixed frequency ξ . The preliminary Sections 4 and 5 are respectively dedicated to some properties of outer L^p spaces and to some results on adapted systems of wave packets, which are used in Section 6.

3. ESTIMATING THE TRILINEAR FORM $V_{\vec{\beta}}$ VIA CARLESON EMBEDDINGS

This section is dedicated to the proof of Theorem 2. Before the main argument, we state two preliminary results. The following lemma, which summarizes the discussion in [6, p. 46], is the above mentioned Hölder inequality involving outer L^p norms on the upper 3-half space.

Lemma 3.1. *Let $\{\vec{\alpha}, \vec{\beta}\}$ be an orthonormal basis of $(1, 1, 1)^\perp$, δ_0 be as in (1.2), and $\alpha_j := (\alpha_j, \beta_j, 2^{-9}\delta_0)$, $j = 1, 2, 3$. Then*

$$\int_{\mathcal{Z}} \left(\prod_{j=1}^3 |F_j(u, t, \alpha_j \eta + \beta_j t^{-1})| \right) du dt d\eta \leq C_{\delta_0} \prod_{j=1}^3 \|F_j\|_{L^{p_j}(\mathcal{Z}, \sigma_{\alpha_j, s})}$$

for all $F_j \in \mathcal{B}(\mathcal{Z})$, $j = 1, 2, 3$, and Hölder tuples (p_1, p_2, p_3) with $1 \leq p_1, p_2, p_3 \leq \infty$.

We will also need a localized version of (2.6) to be applied in conjunction with Theorem 1. The proof is given in Section 7.

Proposition 3.2. *Let $f \in \mathcal{S}(\mathbb{R})$ and $K > 1$ be fixed. Assume J is a finitely overlapping collection of intervals, see (1.7), such that*

$$(3.1) \quad \text{supp } f \cap KJ = \emptyset \quad \forall J \in J.$$

Then, for all $R > 1$ and $2 < q \leq \infty$

$$(3.2) \quad \|F_\phi(f) \mathbf{1}_{E_J}\|_{L^q(\mathcal{Z}, \sigma_{\alpha, s})} \leq C_{\alpha, \phi, R, q} K^{-R} \|f\|_q, \quad E_J := \bigcup_{J \in J} T(3J) \times \mathbb{R} \subset \mathcal{Z}.$$

Proof of Theorem 2. The implicit constants in this proof are allowed to depend on $p_1, p_2, \phi, \vec{\beta}$ without explicit mention. We deal in detail with the harder case $1 < p_1, p_2 < 2$. When either or both $p_1 > 2, p_2 > 2$, the estimate of the theorem is obtained by a similar (in fact, simpler) argument where (2.6) is used in place of the main estimate (2.9) for the corresponding function f_j . The case where either or both $p_1 = 2, p_2 = 2$ then follows by interpolation, see for instance [13]. We will use later that, since $1/p'_1 + 1/p'_2 > 1/2$ we can find a tuple q_1, q_2, q_3 satisfying

$$\frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{q_3} = 1, \quad q_1 > p'_1, \quad q_2 > p'_2, \quad q_3 > 2.$$

which is fixed from now on. By linearity in f_1, f_2 and horizontal scaling invariance, that is by possibly replacing f_j with $|A|^{-\frac{1}{p_j}} f_j(|A|^{-1} \cdot)$ we may work with

$$\|f_1\|_{p_1} = \|f_2\|_{p_2} = |A| = 1.$$

We begin the actual proof. To construct $\tilde{A}$, we find C_0 large enough such that each

$$H_j := \{x \in \mathbb{R} : M_{p_j}(Mf_j)(x) > C_0\}, \quad j = 1, 2$$

has measure less than $1/4$ and set $\tilde{A} = A \cap (H_1 \cup H_2)^c$. Let us define, with reference to the notation of (2.8), and for $k = 0, 1, \dots$,

$$I_k := \bigcup_{j=1,2} I_{f_j, 2^{10+k}C_0, p_j}, \quad E_k := \bigcup_{j=1,2} E_{f_j, 2^{10+k}C_0, p_j}, \quad \widetilde{E}_k = E_k \setminus E_{k+1}.$$

Let f_3 be supported on $\tilde{A}$ and bounded by 1 be fixed from now on. For reasons of space, let us write G_j , $j = 1, 2, 3$ as in the second line of (1.3). We have, with $z = (u, t, \eta)$

$$(3.3) \quad |\mathcal{V}_{\tilde{\beta}}(f_1, f_2, f_3)| \leq \int_{\mathcal{Z}} \left(\prod_{j=1}^3 |G_j \mathbf{1}_{\mathcal{Z} \setminus E_0}(z)| \right) dz + \sum_{k=0}^{\infty} \int_{\mathcal{Z}} \left(\prod_{j=1}^3 |G_j \mathbf{1}_{\widetilde{E}_k}(z)| \right) dz.$$

We will now prove that (3.3) $\lesssim 1$. Let $\alpha_j = (\alpha_j, \beta_j, \delta)$, $j = 1, 2, 3$, be our parameter vectors with $\tilde{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ chosen as in Lemma 3.1. By Theorem 1, estimate (2.9) for q_j , $j = 1, 2$ we have the embeddings

$$(3.4) \quad \|F_{\phi}(f_j) \mathbf{1}_{\mathcal{Z} \setminus E_k}\|_{L^{q_j}(\mathcal{Z}, \sigma_{\alpha_j, s})} \leq \|F_{\phi}(f_j) \mathbf{1}_{\mathcal{Z} \setminus E_{f_j, 2^{10+k}C_0, p_j}}\|_{L^{q_j}(\mathcal{Z}, \sigma_{\alpha_j, s})} \lesssim 2^k$$

To control $F_{\phi}(f_3)$ we use (2.6) and (3.2), with $R = 10$ say, and obtain the estimates

$$(3.5) \quad \begin{aligned} \|F_{\phi}(f_3)\|_{L^{q_3}(\mathcal{Z}, \sigma_{\alpha_3, s})} &\lesssim \|f_3\|_{q_3} \leq |A|^{\frac{1}{q_3}} = 1, \\ \|F_{\phi}(f_3) \mathbf{1}_{E_k}\|_{L^{q_3}(\mathcal{Z}, \sigma_{\alpha_3, s})} &\lesssim 2^{-10k} \|f_3\|_{q_3} \leq 2^{-10k} \end{aligned}$$

We remark that the assumptions of (3.2) are met since the support of f_3 does not intersect $3 \cdot 2^k I$ for all $I \in I_k$. Using Lemma 3.1, and later (3.4) and the first estimate of (3.5) the first term in (3.3) is bounded by

$$\left(\prod_{j=1}^2 \|F_{\phi}(f_j) \mathbf{1}_{\mathcal{Z} \setminus E_0}\|_{L^{q_j}(\mathcal{Z}, \sigma_{\alpha_j, s})} \right) \|F_{\phi}(f_3)\|_{L^{q_3}(\mathcal{Z}, \sigma_{\alpha_3, s})} \lesssim 1.$$

Again using Lemma 3.1, and later (3.4), trivial estimates and the second bound of (3.5) the second term in (3.3) is controlled by

$$\begin{aligned} &\prod_{j=1}^3 \|F_{\phi}(f_j) \mathbf{1}_{\widetilde{E}_k}\|_{L^{q_j}(\mathcal{Z}, \sigma_{\alpha_j, s})} \\ &\leq \left(\prod_{j=1}^2 \|F_{\phi}(f_j) \mathbf{1}_{\mathcal{Z} \setminus E_{k+1}}\|_{L^{q_j}(\mathcal{Z}, \sigma_{\alpha_j, s})} \right) \|F_{\phi}(f_3) \mathbf{1}_{E_k}\|_{L^{q_3}(\mathcal{Z}, \sigma_{\alpha_3, s})} \lesssim 2^{-8k} \end{aligned}$$

This is summable in k , completing the estimation of (3.3) and the proof of the Theorem. $\square$

4. ELEMENTARY PROPERTIES OF OUTER L^p

In this short section, we collect, without proofs, some basic properties of outer L^p spaces which we use in our analysis. The following proposition summarizes [6, Proposition 3.1, 3.3].

Proposition 4.1. *Let (Z, σ, s) be an outer measure space, $F, G \in \mathcal{B}(Z)$, and $0 < p \leq \infty$. Then*

$$\begin{aligned}
 (4.1) \quad & |F| \leq |G| \implies \|F\|_{L^p(Z, \sigma, s)} \leq \|G\|_{L^p(Z, \sigma, s)}, \\
 & \|\lambda F\|_{L^p(Z, \sigma, s)} = |\lambda| \|F\|_{L^p(Z, \sigma, s)}, \quad \forall \lambda \in \mathbb{C}, \\
 & \|F\|_{L^p(Z, \lambda \sigma, s)} = \lambda^{1/p} \|F\|_{L^p(Z, \sigma, s)}, \quad \forall \lambda > 0, \\
 & \|F + G\|_{L^p(Z, \sigma, s)} \leq 2c_s (\|F\|_{L^p(Z, \sigma, s)} + \|G\|_{L^p(Z, \sigma, s)}), \\
 (4.2) \quad & \|F\|_{L^p(Z, \sigma, s)} \leq c_{p, p_1, p_2} (\|F\|_{L^{p_1, \infty}(Z, \sigma, s)})^\theta (\|F\|_{L^{p_2, \infty}(Z, \sigma, s)})^{1-\theta}, \\
 & 0 < p_1 < p < p_2 \leq \infty, \quad \frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}.
 \end{aligned}$$

And identical statements hold for the spaces $L^{p, \infty}$.

Remark 4.2. Iterating (4.1), one obtains the quasi-triangle inequality

$$(4.3) \quad \left\| \sum_{k=0}^K F_k \right\|_{L^p(Z, \sigma, s)} \leq \sum_{k=0}^K (2c_s)^{(k+1)} \|F_k\|_{L^p(Z, \sigma, s)}.$$

Note that the constant c_s in (4.1) is the same appearing in (2.1) and can be taken equal to 1 in our concrete cases.

We also record the Marcinkiewicz interpolation theorem in outer L^p , [6, Proposition 3.5]

Proposition 4.3. *Let (X, ν) be a measure space, (Z, σ, s) be an outer measure space and T be a quasi-sublinear operator mapping functions in $L^{p_1}(X, \nu)$ and $L^{p_2}(X, \nu)$ into $\mathcal{B}(Z)$, for some $0 < p_1 < p_2 \leq \infty$. Assume that for all $f \in L^{p_1}(X, \nu) + L^{p_2}(X, \nu)$,*

$$\|T(f)\|_{L^{p_j, \infty}(Z, \sigma, s)} \leq B_j \|f\|_{L^{p_j}(X, \nu)}, \quad j = 1, 2.$$

Then, for $0 < \theta < 1$,

$$\|T(f)\|_{L^p(Z, \sigma, s)} \leq C_{\theta, p_1, p_2} B_1^\theta B_2^{1-\theta} \|f\|_{L^p(X, \nu)}, \quad \frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}.$$

Remark 4.4 (Scaling properties of estimate (2.9)). The structure of the wave packet transform (1.1) and of the outer $L^q(\mathcal{Z}, \sigma_a, s)$ spaces yield a useful horizontal scaling property for estimate (2.9). Fix $1 \leq p < \infty$, $f \in L^p(\mathbb{R})$ and $\kappa > 0$ and write, referring to (2.8),

$$f_\kappa := \text{Dil}_{\kappa^{-p}}^p f = \kappa f(\kappa^p \cdot), \quad F := F_\phi(f) \mathbf{1}_{\mathcal{Z} \setminus E_{f, \lambda, p}}, \quad F_\kappa := F_\phi(f_\kappa) \mathbf{1}_{\mathcal{Z} \setminus E_{f_\kappa, \kappa \lambda, p}}$$

Then $\|f_\kappa\|_p = \|f\|_p$ and

$$(4.4) \quad \|F\|_{L^q(\mathcal{Z}, \sigma_a, s)} = \kappa^{\frac{p}{q}-1} \|F_\kappa\|_{L^q(\mathcal{Z}, \sigma_a, s)}$$

and similarly for outer weak L^q norms. To see this, introduce the bijection

$$T_\kappa : \mathcal{Z} \rightarrow \mathcal{Z}, \quad T_\kappa(u, t, \eta) = (\kappa^{-p} u, \kappa^{-p} t, \kappa^p \eta)$$

Then [6, Proposition 3.2] applied to both T_κ and its inverse yields the equality of norms

$$\|G \circ T_\kappa\|_{L^q(\mathcal{Z}, \sigma_a, s)} = \kappa^{\frac{p}{q}} \|G\|_{L^q(\mathcal{Z}, \sigma_a, s)}$$

and similarly for weak norms. We omit the straightforward verification of the assumptions of the proposition for T_κ . Horizontal scaling of the definition of (2.8) yields

$$(u, t, \eta) \in E_{f, \lambda, p} \iff T_\kappa(u, t, \eta) \in E_{f_\kappa, \kappa \lambda, p}$$

while the structure of (1.1) entails

$$F_\phi(f)(u, t, \eta) = \kappa^{-1} F_\phi(f_\kappa)(T_\kappa(u, t, \eta)),$$

and we conclude that $F = \kappa^{-1} F_\kappa \circ T_\kappa$, whence (4.4).

5. ADAPTED SYSTEMS

Let $\xi \in \mathbb{R}$ and $0 < T \leq \infty$. A family

$$\Phi^\xi := \{\phi_t : t \in (0, T)\} \subset \mathcal{S}(\mathbb{R})$$

is said to be a ξ -adapted system with adaptation constants $A_N = A_N(\Phi^\xi)$ if

$$(5.1) \quad \sup_{n, m \leq N} \sup_{t \in (0, T)} \sup_{x \in \mathbb{R}} t^{m+1} \left(1 + \left|\frac{x}{t}\right|\right)^n |(e^{-i\xi \cdot} \phi_t(\cdot))^{(m)}(x)| \leq A_N$$

for all nonnegative integers N . The ξ -adapted system Φ^ξ is said to have mean zero if

$$(5.2) \quad \widehat{\phi_t}(\xi) = 0 \quad \forall t \in (0, T).$$

A prime example of ξ -adapted system [resp. with mean zero] is obtained by dilation and modulation of a single $\phi \in \mathcal{S}(\mathbb{R})$ [resp. with mean zero]

$$(5.3) \quad \phi_t := \text{Mod}_\xi \text{Dil}_t^1 \phi, \quad \phi_t(x) := e^{i\xi x} \frac{1}{t} \phi\left(\frac{x}{t}\right), \quad t \in (0, \infty).$$

The remainder of the section is occupied by three lemmata, each of which plays a significant role in our investigation. First, we record the following well-known principle of the Littlewood-Paley theory.

Lemma 5.1. *Let Φ^ξ be a ξ -adapted system with mean zero. Then*

$$\left(\int_{\mathbb{R} \times (0, \infty)} |f * \phi_t(u)|^2 \frac{du dt}{t} \right) \lesssim A_3(\Phi^\xi) \|f\|_2.$$

Proof. By replacing f with $f(\cdot) e^{-ix\xi}$, we can reduce to the case $\xi = 0$. A consequence of (5.1)-(5.2) is that

$$|\widehat{\phi_t}(\xi)| \lesssim A_3 \min\{t|\xi|, (1+t|\xi|)^{-1}\}.$$

This entails

$$\sup_\xi \int_0^\infty |\widehat{\phi_t}(\xi)|^2 \frac{dt}{t} \lesssim A_3,$$

from which the lemma follows by two applications of Plancherel. $\square$

The next two lemmata deal with the decomposition of the functions belonging to a ξ -adapted system into a compactly supported part and an exponentially small remainder.

Lemma 5.2. *Let Φ^ξ be a ξ -adapted system [resp. with mean zero]. Let $K \geq 1$ and $Q, N \in \mathbb{N}$ be given. There exist two ξ -adapted systems [resp. with mean zero] $\Psi^\xi = \{\psi_t\}$, $\Upsilon^\xi = \{v_t\}$, such that*

$$(5.4) \quad \begin{aligned} \phi_t &= \psi_t + K^{-Q} v_t \quad \forall t \in (0, T), \\ A_N(\Psi^\xi), A_N(\Upsilon^\xi) &\leq C_{Q,N} A_{N+Q+1}(\Phi^\xi) \quad N \in \mathbb{N}, \\ \text{supp } \psi_t &\subset [-Kt, Kt] \quad \forall t \in (0, T). \end{aligned}$$

Proof. We only prove the mean zero case, which is more involved. By replacing ϕ_t with $e^{-i\xi\cdot}\phi_t(\cdot)$, we can reduce to the case $\xi = 0$ and write Φ instead of Φ^0 . Note that, for any fixed value $t_0 > 0$, the system

$$\{\widetilde{\phi}_t := \frac{1}{t_0}\phi_{\frac{t}{t_0}}\left(\frac{\cdot}{t_0}\right) : t \in (0, t_0 T)\}$$

is 0-adapted with the same constants as Φ , it suffices to produce the decomposition (5.4) for $t = 1$. We write ϕ in place of ϕ_1 , and similarly for ψ_1, v_1 .

Let β be a smooth function satisfying

$$|\beta| \leq 1, \quad \beta \equiv 1 \text{ on } \left[-\frac{1}{2}, \frac{1}{2}\right], \quad \text{supp } \beta \subset [-1, 1] \quad \int \beta = 1.$$

Let $K > 1$. Define $\beta_K = \beta(\cdot/K)$. We record the obvious facts

$$(5.5) \quad \begin{aligned} \text{supp } \beta_K &\subset [-K, K], \quad \text{supp } (\mathbf{1}_{\mathbb{R}} - \beta_K) \subset \left[-\frac{K}{2}, \frac{K}{2}\right]^c, \\ \sup_{x \in \mathbb{R}} (1 + |x|)^n \left| \beta_K^{(m)}(x) \right| &\leq C_{n,m} K^{n-m} \end{aligned}$$

with constant $C_{n,m}$ depending only on $m, n \in \mathbb{N}$ through β 's derivatives. Define $I_K := \int_{\mathbb{R}} \beta_K \phi \, dx$. In view of the first line of (5.5), since ϕ has mean zero, we have for all nonnegative integers Q

$$(5.6) \quad \begin{aligned} |I_K| &= \left| \int_{\mathbb{R}} (\mathbf{1}_{\mathbb{R}} - \beta_K) \phi \, dx \right| \leq \int_{|x| > \frac{K}{2}} |\phi| \, dx \\ &\leq A_{Q+1}(\Phi) \int_{|x| > \frac{K}{2}} (1 + |x|)^{-(Q+1)} \, dx \leq C_Q A_{Q+1}(\Phi^\xi) K^{-Q}. \end{aligned}$$

Fix now K and Q . The decomposition (5.4) is achieved by setting

$$\begin{aligned} K^{-Q} v &:= \frac{I_K}{K} \beta_K + \eta_K, \quad \eta_K := (\mathbf{1}_{\mathbb{R}} - \beta_K) \phi, \\ \psi &:= \beta_K \phi - \frac{I_K}{K} \beta_K \end{aligned}$$

Indeed, ψ is obviously supported in $[-K, K]$. Exploiting (5.5), and (5.6) with $Q+N$ in place of Q , we have for $m, n \leq N$

$$(5.7) \quad \sup_{x \in \mathbb{R}} (1 + |x|)^n \left| \frac{I_K}{K} \beta_K^{(m)}(x) \right| \leq C_{Q,N} A_{Q+N+1}(\Phi) K^{-Q}.$$

Furthermore, for $m, n \leq N$, since $\beta_K^{(j)}$ is supported in $|x| > K/2$ for $j \geq 1$,

$$(5.8) \quad \begin{aligned} \left| \eta_K^{(m)}(x) \right| &\leq |(\mathbf{1}_{\mathbb{R}} - \beta_K)(x)| |\phi^{(m)}(x)| + \sum_{j=1}^m \binom{m}{j} |\beta_K^{(j)}(x)| |\phi^{(m-j)}(x)| \\ &\leq C_{Q,N} A_{N+Q}(\Phi) K^{-Q} (1 + |x|)^{-n}. \end{aligned}$$

Combining (5.7)-(5.8), we obtain the required adaptation bound for v , and the one for ψ follows by comparison. $\square$

Iterating the proof of the previous lemma we obtain the following decomposition of a ξ -adapted system Φ^ξ into ξ -adapted systems with compact support. The statement is a slightly more precise version of [12, Lemma 3.1], to which we send for the proof details.

Lemma 5.3. *Let Φ^ξ be a ξ -adapted system [resp. with mean zero] and $Q \in \mathbb{N}$. There exists ξ -adapted systems [resp. with mean zero] $\Psi^{\xi,k} = \{\psi_{t;k}\}$, $k = 0, 1, 2, \dots$ such that*

$$(5.9) \quad \begin{aligned} \phi_t &= \sum_{k \geq 0} 2^{-Qk} \psi_{t;k} \quad \forall t \in (0, T), \\ A_N(\Psi^{\xi,k}) &\leq C_{Q,N} A_{Q+N+1}(\Phi^\xi) \quad N \in \mathbb{N}, \\ \text{supp } \psi_{t;k} &\subset [-2^k t, 2^k t] \quad \forall t \in (0, T). \end{aligned}$$

6. CARLESON EMBEDDINGS FOR TENTS REVISITED

This section is dedicated to Carleson embedding theorems in the outer $L^p(Z, \sigma, s_q)$ spaces previously defined, for the function

$$(6.1) \quad \Phi^\xi(f)(u, t) = f * \phi_t(u), \quad (u, t) \in Z$$

where $\Phi^\xi = \{\phi_t : t \in (0, \infty)\}$ is a ξ -adapted system. Throughout this section, the absolute constant C appearing in $A_C(\Phi^\xi)$ is moderate; $C = 20$ would suffice.

6.1. Global-type Carleson embeddings. The upcoming Carleson embedding theorems are very close in spirit to those given in [6, Section 4]. The slight differences with [6] is that we do not work with dilates of a single function ϕ as in (5.3), and that our ϕ_t need not be compactly supported. Albeit these additions are minor, they will allow us to obtain generalized L^∞ Carleson embeddings in Section 7 by simply averaging the results of this section.

Proposition 6.1. *Let Φ^ξ be a ξ -adapted system. Then*

$$(6.2) \quad \|\Phi^\xi(f)\|_{L^p(Z, \sigma, s_\infty)} \lesssim A_C(\Phi^\xi) \|f\|_p, \quad 1 < p \leq \infty,$$

$$(6.3) \quad \|\Phi^\xi(f)\|_{L^{1,\infty}(Z, \sigma, s_\infty)} \lesssim A_C(\Phi^\xi) \|f\|_1.$$

If, in addition, Φ^ξ has mean zero,

$$(6.4) \quad \|\Phi^\xi(f)\|_{L^p(Z, \sigma, s_2)} \lesssim A_C(\Phi^\xi) \|f\|_p, \quad 1 < p \leq \infty,$$

$$(6.5) \quad \|\Phi^\xi(f)\|_{L^{1,\infty}(Z, \sigma, s_2)} \lesssim A_C(\Phi^\xi) \|f\|_1.$$

After a couple of remarks, we provide the details of proof for (6.4) and (6.5). The proofs of (6.2) and (6.3) can be then easily readapted (and are almost immediate to begin with). In the upcoming remark, we reduce to compact support of $\phi_t \in \Phi^\xi$ by means of Lemma 5.3.

Remark 6.2. We claim that the estimates of Proposition 6.1 can be obtained by proving the corresponding version under the assumption that $\phi_t \in \Phi^\xi$ is supported in $[-Kt, Kt]$, with a bound depending polynomially on K (and on the adaptation constants of Φ^ξ). For instance, (6.4)-(6.5) will follow from

$$(6.6) \quad \|\Phi^\xi(f)\|_{L^p(Z, \sigma, s_2)} \lesssim K^2 A_3(\Phi^\xi) \|f\|_p, \quad 1 < p \leq \infty,$$

$$(6.7) \quad \|\Phi^\xi(f)\|_{L^{1,\infty}(Z, \sigma, s_2)} \lesssim K^2 A_3(\Phi^\xi) \|f\|_1$$

under the above compact support assumption. Let us detail how to derive (6.4). We apply the decomposition of Lemma 5.3 for $Q = 6$ say, and, referring to the notation therein, obtain the decomposition

$$(6.8) \quad \Phi^\xi(f) = \sum_{k \geq 0} 2^{-6k} \Psi^{\xi;k}(f)$$

Applying the triangle inequality (4.3), noting that $\text{supp } \psi_{t;k} \subset [-2^k t, 2^k t]$, and using (6.6)

$$\begin{aligned} \|\Phi^\xi(f)\|_{L^p(Z, \sigma, s_2)} &\lesssim \sum_{k \geq 0} 2^{-4k} \|\Psi^{\xi;k}(f)\|_{L^p(Z, \sigma, s_2)} \lesssim \|f\|_p \sum_{k \geq 0} 2^{-2k} A_3(\Psi^{\xi;k}) \\ &\lesssim A_C(\Phi^\xi) \|f\|_p \end{aligned}$$

which is (6.4).

Remark 6.3. Given a ξ -adapted system [resp. with mean zero] $\Phi^\xi = \{\phi_t\}$, the system $\Phi := \{e^{-i\xi \cdot} \phi_t(\cdot)\}$ is zero-adapted with same adaptation constants. Since

$$|\Phi^\xi(f)| = |\Phi(f(\cdot)e^{i\xi \cdot})|,$$

we can reduce to the case $\xi = 0$ when proving (6.6)-(6.7). This is merely for notational convenience.

Proof of (6.6)-(6.7). We begin by proving the case $p = \infty$ of (6.6). Let I be any interval. We need to prove that

$$s_2(\Phi(f))(T(I))^2 = \frac{1}{|I|} \int_{T(I)} |\Phi(f)(u, t)|^2 \frac{dudt}{t} \lesssim K^4 A_3(\Phi)^2 \|f\|_\infty^2.$$

Due to $\text{supp } \phi_t \subset [-Kt, Kt]$, $\Phi(f 1_{\mathbb{R} \setminus 3KI})(u, t) = 0$ whenever $(u, t) \in T(I)$. Thus using Lemma 5.1,

$$\begin{aligned} \int_{T(I)} |\Phi(f)(u, t)|^2 \frac{dudt}{t} &= \int_{T(I)} |\Phi(f 1_{3KI})(u, t)|^2 \frac{dudt}{t} \leq \int_Z |\Phi(f 1_{3KI})(u, t)|^2 \frac{dudt}{t} \\ &\lesssim A_3(\Phi)^2 \|f 1_{3KI}\|_2^2 \lesssim A_3(\Phi)^2 K |I| \|f\|_\infty^2, \end{aligned}$$

which is (more than) what we sought.

Next we will prove (6.7), and the remaining cases of (6.6) will follow from outer interpolation. We can assume, by linearity, that $\|f\|_1 = 1$. Given $\lambda > 0$, let $I \in \mathcal{I}$ be the collection of maximal dyadic intervals such that I is contained in $\{Mf > \lambda\}$. By the maximal theorem, the set $E = \bigcup_{I \in \mathcal{I}} T(3I)$ has outer measure $\mu(E) \lesssim \lambda^{-1}$. It will then suffice to show that

$$(6.9) \quad \sup_J s_2(\Phi(f) 1_{Z \setminus E})(T(J)) \lesssim A_3(\Phi) K^2 \lambda$$

Let

$$f = g + b = g + \sum_{I \in \mathcal{I}} b_I$$

be the Calderón-Zygmund decomposition of f at level λ . Since $\|g\|_\infty \lesssim \lambda$, by the previous part of the proof we get

$$(6.10) \quad \sup_J s_2(\Phi(g) 1_{Z \setminus E})(T(J)) \leq \sup_J s_2(\Phi(g))(T(J)) \lesssim A_3(\Phi) K \lambda.$$

We now turn to bounding the contribution of b . Let $(u, t) \notin T(3I)$. By virtue of the support condition on ϕ_t , $b_I * \phi_t(u) = 0$ unless $t \geq K^{-1}|I|$. Letting B_I denote the compactly, disjointly supported primitives of b_I , we recall that

$$\|B_I\|_\infty \lesssim \lambda|I|.$$

Also, letting $\psi_t = (\phi_t)'$, we integrate by parts to obtain

$$(6.11) \quad |b * \phi_t(u)| = \left| \sum_{Kt \geq |I|} B_I * \psi_t(u) \right| \leq \left\| \sum_{Kt \geq |I|} t^{-1} B_I \right\|_\infty \|t\psi_t\|_1 \lesssim A_3(\Phi)K\lambda, \quad (u, t) \notin E.$$

Thus

$$(6.12) \quad \sup_J s_\infty(\Phi(b)\mathbf{1}_{Z \setminus E})(T(J)) \lesssim KA_3(\Phi)\lambda.$$

Fix now $I \in \mathcal{I}$. Notice that $\Phi(b_I) = 0$ on $T(J)$ unless $I \subset 3KJ$. By the same support considerations and integrating by parts

$$\begin{aligned} \int_{T(J) \setminus E} |b_I * \phi_t(u)| \frac{dudt}{t} &\leq \int_{Kt \geq |I|} \int_{|u-c(I)| \leq 2Kt} \|B_I\|_1 \|\psi_t\|_\infty \frac{dudt}{t} \\ &\lesssim KA_3(\Phi)\lambda|I|^2 \int_{Kt \geq |I|} \frac{dt}{t^2} \lesssim K^2 A_3(\Phi)\lambda|I|, \end{aligned}$$

so that, summing over $I \subset 3KJ$ (which are pairwise disjoint, we obtain

$$s_1(\Phi(b)\mathbf{1}_{Z \setminus E})(T(J)) \lesssim K^3 A_3(\Phi)\lambda.$$

Interpolating the above bound with (6.12) it follows that

$$(6.13) \quad \sup_J s_2(\Phi(b)\mathbf{1}_{Z \setminus E})(T(J)) \lesssim A_3(\Phi)K^2\lambda,$$

which, combined with (6.10), completes the proof of (6.9), and, in turn, of (6.7). $\square$

6.2. Localized embeddings. Let $\mathcal{I}$ be a collection of intervals. The tent over $\mathcal{I}$ is defined as

$$(6.14) \quad E_{\mathcal{I}} = \bigcup_{I \in \mathcal{I}} T(3I).$$

For technical reasons, we need to define $E_{\mathcal{I}}$ using the slightly enlarged tents $T(3I)$. We warn the reader that this creates a discrepancy between the tent over an interval I , namely $T(I)$, and the tent over the family $\mathcal{I} = \{I\}$, which is $E_{\mathcal{I}} = T(3I)$. The next three results are L^∞ Carleson embeddings outside or inside $E_{\mathcal{I}}$ -type sets.

Lemma 6.4. *Let Φ^ξ be a ξ -adapted system. For $f \in \mathcal{S}(\mathbb{R})$, define*

$$(6.15) \quad \mathcal{I}_{f, \lambda, 1} := \{\text{maximal dyadic intervals } I \subset \{x \in \mathbb{R} : Mf(x) > \lambda\}\}.$$

and, referring to (6.14), write $E_{f, \lambda, 1} := E_{\mathcal{I}_{f, \lambda, 1}}$. Then

$$\left\| \Phi^\xi(f) \mathbf{1}_{Z \setminus E_{f, \lambda, 1}} \right\|_{L^\infty(Z, \sigma, s_\infty)} \lesssim A_C(\Phi^\xi)\lambda.$$

If furthermore Φ^ξ has mean zero

$$(6.16) \quad \left\| \Phi^\xi(f) \mathbf{1}_{Z \setminus E_{f,\lambda,1}} \right\|_{L^\infty(Z,\sigma,s_2)} \lesssim A_C(\Phi^\xi) \lambda.$$

Proof. Perusal of the proof of (6.9), relying on the Calderón-Zygmund decomposition. The extension to the non-compactly supported case is then obtained via the same decomposition (6.8) recalled in Remark 6.2. $\square$

Lemma 6.5. *Let Φ^ξ be a ξ -adapted system, $K \geq 1$, and Q be a positive integer. Let J be a finitely overlapping collection of intervals and $f \in S(\mathbb{R})$ be such that*

$$\text{supp } f \cap 27KJ = \emptyset \quad \forall J \in J.$$

Then, referring to (6.14) for E_J ,

$$(6.17) \quad \left\| \Phi^\xi(f) \mathbf{1}_{E_J} \right\|_{L^\infty(Z,\sigma,s_\infty)} \leq C_Q A_{CQ}(\Phi^\xi) K^{-Q} \sup_{J \in J} \inf_{x \in J} Mf(x),$$

and furthermore, if Φ^ξ has mean zero,

$$(6.18) \quad \left\| \Phi^\xi(f) \mathbf{1}_{E_J} \right\|_{L^\infty(Z,\sigma,s_2)} \leq C_Q A_{CQ}(\Phi^\xi) K^{-Q} \sup_{J \in J} \inf_{x \in J} Mf(x),$$

Proof. We prove (6.18), the proof of (6.17) being exactly the same. First of all, we notice that

$$(6.19) \quad \left\| \Phi^\xi(f) \mathbf{1}_{E_J} \right\|_{L^\infty(Z,\sigma,s_2)} = \sup_{L \subset \mathbb{R}} s_2(\Phi^\xi(f) \mathbf{1}_{E_J})(T(L)) \leq C \sup_{J \in J} \sup_{L \subset 9J} s_2(\Phi^\xi(f))(T(L)).$$

This is because if $T(L)$ intersects E_J nontrivially, and there is no $J \in J$ with $L \subset 9J$, it must be that $3J \subset 3L$ whenever $J \in J$ is such that $3J \cap L \neq \emptyset$. Therefore

$$E_J \cap T(L) \subset \bigcup_{J \in J: J \subset 3L} T(3J),$$

whence

$$(s_2(\Phi^\xi(f) \mathbf{1}_{E_J})(T(L)))^2 \leq \frac{1}{|L|} \sum_{J \in J: J \subset 3L} |3J| (s_2(\Phi^\xi(f))(T(3J)))^2$$

which, by finite overlap of $J \in J$ is less than C times the supremum on the right hand side of (6.19). We will now bound the right hand side of (6.19). Fix $J \in J$ and $L \subset 9J$, and let

$$\lambda = 9 \inf_{x \in J} Mf(x).$$

Construct $I_{f,\lambda,1}$ and $E_{f,\lambda,1}$ as in (6.15). Since $27J \cap \text{supp } f = \emptyset$, it is easy to see that

$$9J \cap 3H = \emptyset \quad \text{for all intervals } H \text{ with } \int_H |f| dx > \lambda |H|,$$

which in turn implies

$$(6.20) \quad T(L) \cap E_{f,\lambda,1} = \emptyset.$$

From Lemma 5.2, there exists two ξ -adapted systems with mean zero $\Psi^\xi = \{\psi_t\}$, $\Upsilon^\xi = \{v_t\}$, such that for all t

$$\phi_t = \psi_t + K^{-Q} v_t, \quad \text{supp } \psi_t \subset [-Kt, Kt].$$

Since $\text{dist}(L, \text{supp } f) \geq 9K|J| \geq K|L|$, the support of $(u, t) \mapsto \Psi^\xi(f)(u, t)$ does not intersect $T(L)$. Therefore

$$\Phi^\xi(f)(u, t) = K^{-Q} \Upsilon^\xi(f)(u, t) = K^{-Q} \Upsilon^\xi(f) \mathbf{1}_{Z \setminus E_{f, \lambda, 1}}(u, t), \quad (u, t) \in T(L),$$

where the second equality is a consequence of (6.20). By (6.16),

$$s_2(\Phi^\xi(f))(T(L)) \leq K^{-Q} \left\| \Upsilon^\xi(f) \mathbf{1}_{Z \setminus E_{f, \lambda, 1}} \right\|_{L^\infty(Z, \sigma, s_2)} \lesssim K^{-Q} A_C(\Upsilon^\xi) \lambda \leq C_Q A_{CQ}(\Phi^\xi) K^{-Q} \lambda,$$

relying on (5.9) to get the last equality. This proves the claimed bound of (6.19). $\square$

The final lemma of this section is a strengthening of Lemma 6.4 in the case where the function involved is a sum of Calderón-Zygmund atoms with Q vanishing moments. This produces an exponential gain in Q .

Lemma 6.6. *Let $\xi \in \mathbb{R}$ and let L be a countable collection of intervals. Let $b \in L^1_{\text{loc}}(\mathbb{R})$ be such that*

$$(6.21) \quad b = \sum_{L \in L} b_L, \quad \text{supp } b_L \subset L, \quad \sup_{I \subset \mathbb{R}} \frac{1}{|I|} \sum_{L \subset I} \|b_L\|_1 =: \lambda < \infty$$

$$(6.22) \quad \int b_L(x) x^j e^{-i\xi x} dx = 0 \quad \forall j = 0, \dots, 2(Q+1).$$

Let $K > 3$ be given and denote by $KL = \{KL : L \in L\}$. Let Φ^ξ be a ξ -adapted system [resp. with mean zero]. Then, for $p = \infty$ [resp. $p = 2$],

$$(6.23) \quad \left\| \Phi^\xi(b) \mathbf{1}_{Z \setminus E_{KL}} \right\|_{L^\infty(Z, \sigma, s_p)} \leq C_L C_Q A_{CQ}(\Phi^\xi) K^{-Q} \lambda.$$

Remark 6.7. Note that (6.21), with $C\lambda$ in place of λ , follows if it is known that

$$(6.24) \quad \text{supp } b_L \subset L, \quad \sup_{L \in L} \frac{1}{|L|} \|b_L\|_1 \leq \lambda, \quad \sup_{I \in \mathbb{R}} \frac{1}{|I|} \sum_{\substack{L \in L \\ L \subset I}} |L| \leq C.$$

Remark 6.8. Before the actual proof, we record that, as a consequence of (6.21), there holds

$$I \not\subset \tilde{E} := \bigcup_{L \in L} 3L \implies \|b \mathbf{1}_I\|_1 \leq \sum_{\substack{L \in L \\ L \subset 3I}} \|b_L\|_1 \leq 3\lambda |I|$$

which in turn implies that $\{x \in \mathbb{R} : Mb(x) > 3\lambda\} \subset \tilde{E}$. Namely, referring to the notations below (6.15) and in (6.14), is that if $I \in \mathbf{I}_{b, 3\lambda, 1}$ then $3I$ is covered by the intervals $\{9L : L \in L\}$ whence

$$(6.25) \quad E_{b, 3\lambda, 1} \subset E_{3L} \subset E_{KL}.$$

Proof of Lemma 6.6. We prove the mean zero case. The proof of the other one is actually within this case. To lessen the notational burden, one may reduce to the case $\xi = 0$ by the same argument as in Remark 6.3 (i. e. replacing $b(\cdot)$ by $b(\cdot)e^{i\xi \cdot}$), and simply call Φ the 0-adapted system with mean zero.

We begin the actual proof. Clearly, for (6.23) it suffices to estimate $s_2(\Phi(b) \mathbf{1}_{Z \setminus E_{KL}})(T(J))$ for those intervals J with $T(J) \not\subset E_{KL}$. Fix one such J from now on and let $\kappa = \sqrt{K}$. We

apply Lemma 5.2 with κ in place of K and $2Q$ replacing Q , to find adapted systems with mean zero $\Psi = \{\psi_t\}$, $\Upsilon = \{v_t\}$, such that

$$(6.26) \quad \phi_t = \psi_t + K^{-Q}v_t, \quad \text{supp } \psi_t \subset [-\kappa t, \kappa t].$$

Using (6.25) for the first inequality, we apply the estimate (6.16) from Lemma 6.4, and subsequently (5.9), yielding

$$(6.27) \quad s_2(\Upsilon(b)\mathbf{1}_{Z \setminus E_{KL}})(T(J)) \leq s_2(\Upsilon(b)\mathbf{1}_{Z \setminus E_{b,3\lambda,1}})(T(J)) \lesssim A_C(\Upsilon)\lambda \leq C_Q A_{CQ}(\Phi^\xi)\lambda.$$

The treatment of $\Psi(b)$ is very similar to the proof of (6.13), the difference being that Q integration by parts are performed. We make two important observations. The first is that, in view of the support of ψ_t being as in (6.26),

$$(6.28) \quad (u, t) \notin T(3KL) \implies \Psi(b_L)(u, t) = 0 \text{ unless } t \geq \kappa|L|, L \subset (u - 3\kappa t, u + 3\kappa t)$$

The second, by the same reasons, is that

$$(6.29) \quad (u, t) \in T(J) \implies \Psi(b_L)(u, t) = 0 \text{ unless } L \subset 9\kappa J$$

Iteratively define

$$b_{L,0} := b_L, \quad b_{L,j+1} := \int b_{L,j}(x) dx, \quad j = 0, \dots, 2Q+1, \quad B_L = b_{L,2(Q+1)}.$$

The moment conditions (6.22) ensure that B_L are supported on L and

$$(6.30) \quad \|B_L\|_\infty \leq |L|^{2Q+1} \|b_L\|_1.$$

Let now $\varphi_t = (\psi_t)^{(2Q+2)}$. Similarly to (6.11), we integrate by parts $2(Q+1)$ times to obtain

$$\begin{aligned} |\Psi(b)(u, t)| &\leq \sum_{\substack{t \geq \kappa|L| \\ L \subset (u-3\kappa t, u+3\kappa t)}} |B_L * \varphi_t(u)| \leq \sum_{\substack{t \geq \kappa|L| \\ L \subset (u-3\kappa t, u+3\kappa t)}} t^{-2(Q+1)} \|B_L\|_\infty \|t^{2(Q+1)} \varphi_t\|_1 \\ &\leq \kappa^{-2Q} (\kappa t)^{-1} \left(\sum_{L \subset (u-3\kappa t, u+3\kappa t)} \|b_L\|_1 \right) A_{2(Q+2)}(\Psi) \lesssim A_{2(Q+2)}(\Psi) \kappa^{-2Q} \lambda, \end{aligned}$$

whenever $(u, t) \notin E_{KL}$. Note that the restriction in the first sum is due to (6.28) and we have used it, together with (6.30), to pull out a κ^{-2Q-1} factor. The last inequality is obtained thanks to (6.21). Therefore

$$(6.31) \quad s_\infty(\Phi(b)\mathbf{1}_{Z \setminus E_{KL}})(T(J)) \lesssim A_{2(Q+2)}(\Psi) \kappa^{-2Q} \lambda \leq C_Q A_{CQ}(\Phi) K^{-Q} \lambda.$$

Note that (6.30) entails $\|B_L\|_1 \leq |L|^{2Q+2} \|b_L\|_1$. Making use of the observation (6.29), another integration by parts yields

$$\begin{aligned} \int_{T(J) \setminus E_{KL}} |b_L * \psi_t(u)| \frac{du dt}{t} &\leq \int_{t \geq \kappa|L|} \int_{|u-c(L)| \leq \kappa t} \|B_L\|_1 \|t^{2Q+3} \varphi_t\|_\infty \frac{du dt}{t^{2Q+4}} \\ &\lesssim A_{2(Q+2)}(\Psi) |L|^{2Q+2} \|b_L\|_1 \int_{t \geq \kappa|L|} \frac{\kappa dt}{t^{2Q+3}} \lesssim A_{2(Q+2)}(\Psi) \kappa^{-2Q-1} \|b_L\|_1. \end{aligned}$$

Dividing by $|J|$, summing over $L \subset 9\kappa J$, and using again the assumption (6.21) we obtain, also in view of (5.9), the inequality

$$(6.32) \quad s_1(\Psi(b)\mathbf{1}_{Z \setminus E_{KL}})(T(J)) \leq C_Q A_{CQ}(\Phi^\xi) K^{-Q} \lambda$$

An interpolation between (6.32) and (6.31) yields

$$(6.33) \quad s_2(\Psi(b)\mathbf{1}_{Z \setminus E})(T(J)) \leq C_Q A_{CQ}(\Phi^\xi) K^{-Q} \lambda$$

and combining (6.33) with (6.27) and the decomposition (6.26) finishes the proof of the claimed estimate. $\square$

Remark 6.9. Lemmata 6.4 and 6.5 are outer measure analogues of two usual time-frequency analysis lemmata, the John-Nirenberg inequality (see for example [18, Proposition 2.4.1]) and the localization trick, used for instance to estimate the contribution of the part of the model operator localized inside the exceptional set. Lemma 6.6, on the other hand, has no close discrete analogue. Its purpose is to control the bad part arising in the multi-frequency CZ decomposition which will be used in the proof of Theorem 1.

7. PROOFS OF THEOREM 1 AND PROPOSITION 3.2

In the upcoming subsection, we prove Proposition 3.2, which will follow directly from the embeddings of Section 6. The remainder of the section is devoted to the proof of Theorem 1. Subsection 7.2 contains the preliminary tools leading to the main argument in Subsection 7.3, which in turn reduces (2.9) from Theorem 1 to the multi-frequency Calderón-Zygmund Lemma 7.8. The proof of Lemma 7.8 is postponed to Section 8.

7.1. Proof of Proposition 3.2. We begin with a remark linking outer measure spaces on classical tents with their generalized counterpart. To highlight the dependence of this observation on the geometric parameters α , we go back to explicitly writing the subscript α till the end of the subsection. Let $F \in \mathcal{B}(Z)$ and define

$$F_{\xi, \theta} \in \mathcal{B}(Z), \quad F_{\xi, \theta}(u, t) := F\left(u, t, \xi + \frac{\theta - \beta}{\alpha t}\right).$$

Comparing the definitions of (2.5) and (2.3), a change of variables entails

$$\begin{aligned} s_2(F)(T_\alpha(I, \xi))^2 &= \alpha^{-1} \int_{\Theta_\alpha} \left(s_2(F_{\xi, \theta})(T_\alpha(I))\right)^2 d\theta, \quad \Theta_\alpha := \{|\theta| < 1, |\theta - \beta| > \alpha\delta\}, \\ s_\infty(F)(T_\alpha(I, \xi)) &= \sup_{|\theta| < 1} \left(s_\infty(F_{\xi, \theta})(T_\alpha(I))\right) \end{aligned}$$

in consequence of which, if $G \subset Z$ is a Borel set

$$(7.1) \quad \|F \mathbf{1}_{G \times \mathbb{R}}\|_{L^\infty(Z, \sigma_\alpha, s)} \leq \sup_{\xi \in \mathbb{R}} \left(\sup_{|\theta| < 1} \|F_{\xi, \theta} \mathbf{1}_G\|_{L^\infty(Z, \sigma_\alpha, s_\infty)} + \alpha^{-\frac{1}{2}} \sup_{\theta \in \Theta_\alpha} \|F_{\xi, \theta} \mathbf{1}_G\|_{L^\infty(Z, \sigma_\alpha, s_2)} \right).$$

In relation to (7.1), and recalling that $F_\phi(f)(u, t, \eta) = f * \phi_{t, \eta}(u)$ as well as the notation (6.1), one has the equality

$$(7.2) \quad (F_\phi(f))_{\xi, \theta} = \Phi^{\xi; \theta}(f), \quad \xi \in \mathbb{R}, |\theta| < 1$$

where

$$\Phi^{\xi; \theta} = \left\{ \phi_{t, \eta(\theta)} : t \in (0, \infty) \right\}, \quad \eta(\theta) := \xi + t^{-1} \frac{\theta - \beta}{\alpha}$$

are ξ -adapted systems with adaptation constants

$$A_N(\Phi^{\xi;\theta}) \leq C_N \alpha^{-CN} S_N(\phi).$$

In particular, these adaptation constants are uniform in $\theta \in (-1, 1)$. If one further restricts to Θ_α , we obtain a ξ -adapted system with mean zero as well.

We will now use these observations in the proof of Proposition 3.2. Assume f and J satisfy (3.1) for some $K > 1$. It suffices, by virtue of (2.6), to argue for large K . In view of (7.1) and (7.2), and the uniform adaptedness of $\Phi^{\xi;\theta}$, an application of the estimates (6.17) and (6.18) from Lemma 6.5 for a slightly smaller K and for some $Q > 1$, entail

$$(7.3) \quad \|F_\phi(f) \mathbf{1}_{E_J}\|_{L^\infty(\mathcal{Z}, \sigma_{\alpha, s})} \leq C_{\alpha, \phi, Q} K^{-Q} \sup_{J \in \mathcal{J}} \inf_{x \in J} Mf(x).$$

The supremum on the right hand side is obviously bounded by $\|f\|_\infty$, whence (3.2) for $q = \infty$, $Q = R$. To obtain the case $2 < q < \infty$, we use outer Marcinkiewicz interpolation between (7.3), with Q suitably chosen depending on R and q , and (2.7), on the linear operator

$$f \mapsto F_\phi(f \mathbf{1}_{(\bigcup_{J \in \mathcal{J}} KJ)^c}) \mathbf{1}_{E_J}.$$

This completes the proof of Proposition 3.2.

7.2. Proof of Theorem 1: preliminaries. First of all, we record two lemmata which are obtained respectively from Lemma 6.4 and 6.6 via an argument analogous to the one used for Proposition 3.2 and involving (7.1). The first is the $q = \infty$ easy endpoint of (2.9) from Theorem 1. The second is key to the estimation of the *bad part* leading to the embedding 2.9.

Lemma 7.1. *Let $f \in \mathcal{S}(\mathbb{R})$, $\lambda > 0$ and $1 \leq p < 2$, and refer to the notation of (2.8) for $E_{f, \lambda, p}$. Then*

$$(7.4) \quad \left\| F_\phi(f) \mathbf{1}_{\mathcal{Z} \setminus E_{f, \lambda, p}} \right\|_{L^\infty(\mathcal{Z}, \sigma_{\alpha, s})} \leq C_{\alpha, \phi} \lambda \|f\|_p$$

with $C_{\alpha, \phi} = C \alpha^{-C} S_C(\phi)$.

Lemma 7.2. *Let $\xi \in \mathbb{R}$, and b, L , and $\lambda > 0$ be as in Lemma 6.6. Let $K > 1$ be a given constant and denote*

$$(7.5) \quad E_{KL} := \bigcup_{L \in \mathcal{L}} T(3KL) \times \mathbb{R} \subset \mathcal{Z}.$$

Then

$$(7.6) \quad \sup_{J \subset \mathbb{R}} (F_\phi(b) \mathbf{1}_{\mathcal{Z} \setminus E})(T(J, \xi)) \leq C_{\alpha, \phi, Q} K^{-Q} \lambda,$$

with $C_{\alpha, \phi, Q} = C_Q \alpha^{-C_Q} S_{CQ}(\phi)$.

We will also need several components of the proof of (2.7), which in fact can be reconstructed by combining the definitions and lemmata that follow. Since α will be thought of as fixed throughout this paragraph, as well as for the next section, we omit it from the notation and write T in place of T_α , σ in place of σ_α , ν in place of ν_α . In the upcoming definitions, the constant $C_{\alpha, \phi}$ is allowed to depend on the parameter α , more precisely, polynomially in $|\alpha|^{-1}$, and on $S_C(\phi)$ for large enough C .

Remark 7.3 (Reduction to tents with discrete parameters). Let $\mathcal{D}$ be a finite union of dyadic grids on $\mathbb{R}$. We denote by $E_{\mathcal{D}}$ the collection of generalized tents $T(I, \xi)$ such that $I \in \mathcal{D}$ and $\xi \in \delta|I|^{-1}\mathbb{Z}$ for some small, fixed dyadic parameter δ . Momentarily, denote by $\mu_{\mathcal{D}}$ the outer measure generated by the collection $E_{\mathcal{D}}$ via the premeasure $\sigma_{\mathcal{D}}(T(I, \xi)) = |I|$. Following the reduction presented in [6, Lemma 5.2], there is a suitable choice of $\mathcal{D}$ yielding equivalence of outer L^p spaces

$$L^p(\mathcal{Z}, \sigma_{\mathcal{D}}, s) \sim L^p(\mathcal{Z}, \sigma, s), \quad L^{p,\infty}(\mathcal{Z}, \sigma_{\mathcal{D}}, s) \sim L^{p,\infty}(\mathcal{Z}, \sigma, s)$$

with absolute constant. Relying upon the above equivalence, we switch to working with the spaces $L^p(\mathcal{Z}, \sigma_{\mathcal{D}}, s)$ and their weak counterparts in place of $L^p(\mathcal{Z}, \sigma, s)$. However, we drop the $\mathcal{D}$, thus reverting to the original notation. There is no loss in generality in actually assuming that $\mathcal{D}$ is the standard dyadic grid.

Definition 7.4. A countable collection $\mathcal{T} = \{T_n\}$ of generalized tents is said to be ∞ -strongly disjoint if the following holds: there exist points $(u_n, t_n, \eta_n) \in T_n$ with $t_n \geq 2^{-3}|I_{T_n}|$, such that for all $g \in L^2(\mathbb{R})$, if

$$\rho := \inf_n F_{\phi}(g)(u_n, t_n, \eta_n) > 0$$

there holds

$$(7.7) \quad \rho^2 \sum_{T_n \in \mathcal{T}} |I_{T_n}| \leq C_{\alpha, \phi} \|g\|_2^2.$$

Definition 7.5. A countable collection $\mathcal{T}$ of generalized tents is said to be 2-strongly disjoint if the following holds. For each $T \in \mathcal{T}$ there exists a distinguished subset $T^* \subset T^\ell$ such that whenever $g \in L^2(\mathbb{R})$ satisfies

$$(7.8) \quad \sup_{T \in \mathcal{T}} \sup_{(u, t, \eta) \in T^*} |F_{\phi}(g)(u, t, \eta)| \leq C\rho,$$

$$C^{-1}\rho \leq \left(\frac{1}{|I_T|} \int_{T^*} |F_{\phi}(g)|^2 du dt d\eta \right)^{\frac{1}{2}} \leq C\rho \quad \forall T \in \mathcal{T}$$

for some $\rho > 0$, there holds

$$(7.9) \quad \rho^2 \sum_{T \in \mathcal{T}} |I_T| \leq C \sum_{T \in \mathcal{T}} \int_{T^*} |F_{\phi}(g)|^2 du dt d\eta \leq C_{\alpha, \phi} \|g\|_2^2.$$

Lemma 7.6 (Selection algorithm). Let $f \in \mathcal{S}(\mathbb{R})$ with compact frequency support and $E \subset \mathcal{Z}$ be given. Denote by $F := F_{\phi}(f)\mathbf{1}_{E^c}$ and assume that for some $\lambda > 0$

$$\sup_T s(F_{\phi}(f)\mathbf{1}_{E^c})(T) > \lambda.$$

Then, there exist an ∞ -strongly disjoint collection $\mathcal{T}_0 = \{\mathcal{T}_n\}$ with distinguished points $(u_n, t_n, \eta_n) \in \mathcal{T}_n$, and a 2-strongly disjoint collection $\mathcal{T}_1 = \{\mathcal{T}\}$ with distinguished subsets $\mathcal{T}^* \subset \mathcal{T}^\ell$ such that

$$(7.10) \quad \nu(s(F) > \lambda) \leq \sum_{\mathcal{T} \in \mathcal{T}_0 \cup \mathcal{T}_1} |I_{\mathcal{T}}|,$$

$$(7.11) \quad \inf_{\mathcal{T}_n \in \mathcal{T}_0} |F(u_n, t_n, \eta_n)| > \lambda,$$

$$(7.12) \quad \sup_{\mathcal{T} \in \mathcal{T}_1} \sup_{(u, t, \eta) \in \mathcal{T}^*} |F(u, t, \eta)| \leq \lambda,$$

$$(7.13) \quad 2^{-4}\lambda \leq \left(\frac{1}{|I_{\mathcal{T}}|} \int_{\mathcal{T}^*} |F|^2 du dt d\eta \right)^{\frac{1}{2}} \leq \lambda \quad \forall \mathcal{T} \in \mathcal{T}_1.$$

Proof. We will carry on explicitly the construction of $\mathcal{T}_0$, and prove that it is an ∞ -strongly disjoint collection, by perusing the first part of the argument of [6, Subsection 5.2], concerning the s_∞ part of s . Proceeding as in [6], we may select a (finite or countably infinite) collection of tents $\mathcal{T}_0 := \{\mathcal{T}_n\}$, whose union we denote by H , and $(u_n, t_n, \eta_n) \in \mathcal{T}_n \cap E^c$ such that

$$\inf_n F_\phi(f)(u_n, t_n, \eta_n) > \lambda, \quad \sup_{\mathcal{T}} s_\infty(F \mathbf{1}_{H^c})(\mathcal{T}) \leq \lambda.$$

The first property of the last display is (7.11). The second property ensures that

$$(7.14) \quad \nu(s_\infty(F) > \lambda) \leq \sum_{\mathcal{T} \in \mathcal{T}_0} |I_{\mathcal{T}}|.$$

The construction of the sequence $\{(u_n, t_n, \eta_n)\}$ ensures that, setting $\phi_n(\cdot) = \sqrt{t_n} \phi_{t_n, \eta_n}(u_n \cdot)$, there holds

$$\sup_n \sum_{k: t_k \leq t_n} \left(\frac{t_k}{t_n} \right)^{\frac{1}{2}} |\langle \phi_k, \phi_n \rangle| \leq C_{\alpha, \phi};$$

this is [6, eq. (5.11)]. Now, recalling that for any $g \in L^2(\mathbb{R})$, one has $\sqrt{t_n} F_\phi(g)(u_n, t_n, \eta_n) = \langle g, \bar{\phi}_k \rangle$ the exact same argument of [6] yields, for all $m \geq 0$,

$$\sum_{n: \rho 2^m \leq F_\phi(g)(u_n, t_n, \eta_n) < \rho 2^{m+1}} \rho^2 |I_{\mathcal{T}_n}| \leq C 2^{-2m} \sum_{n: \rho 2^m \leq F_\phi(g)(u_n, t_n, \eta_n) < \rho 2^{m+1}} |\langle g, \bar{\phi}_k \rangle|^2 \leq C_{\alpha, \phi} 2^{-2m} \|g\|_2^2$$

and the claim (7.7) follows by summing over m . We have thus proved that $\mathcal{T}_0$ is a ∞ -strongly disjoint collection. At this point, a (finite or countable) collection $\mathcal{T} \in \mathcal{T}_1$ of tents with distinguished subsets $\mathcal{T}^* \subset \mathcal{T}^\ell$ satisfying (7.12)-(7.13), and such that

$$(7.15) \quad \nu(s_2(F) > \lambda) \leq \sum_{\mathcal{T} \in \mathcal{T}_1} |I_{\mathcal{T}}|,$$

can be constructed by using the *second* part of [6, Subsection 5.2]. Notice that we finally obtain the estimate (7.10) by coupling (7.14) with (7.15). The same ideas used for $\mathcal{T}_0$ will yield that $\mathcal{T}_1$ is a 2-strongly disjoint collection, and this completes the proof of the lemma. $\square$

7.3. Proof of Theorem 1: main argument. Throughout this argument $1 < p < 2$, $q > p'$ are fixed. We set

$$(7.16) \quad \bar{\varepsilon} = \bar{\varepsilon}_{p, q} := \frac{1}{p'} - \frac{1}{q}, \quad \varepsilon := 2^{-10} \bar{\varepsilon}.$$

7.3.1. *Reduction to compact frequency support.* We show that Theorem (2.9) can be reduced to the case of f having compact frequency support, in which case the sets in (2.8) can be replaced by

$$(7.17) \quad \begin{aligned} \underline{I}_{f,\lambda,p} &= \max. \text{ dyad. int. } I \text{ s.t. } I \subset \{x \in \mathbb{R} : M_p f(x) > \lambda \|f\|_p\}, \\ \underline{E}_{f,\lambda,p} &= \bigcup_{I \in \underline{I}_{f,\lambda,p}} T(3I) \times \mathbb{R}. \end{aligned}$$

Note that, in particular, $\underline{E}_{f,\lambda,p}$ from (7.17) is contained in the corresponding version from (2.8).

Let $f \in \mathcal{S}(\mathbb{R})$, $\lambda > 0$ be fixed. Choose an increasing sequence of integers ξ_k with $\xi_0 = 0$ and with the property that, setting $A_0 = (-\xi_2, \xi_2)$, $A_k = \{\xi_k < |\xi| < \xi_{k+2}\}$ for $k \geq 1$, and $g_k := \mathcal{F}^{-1}(\widehat{f} \mathbf{1}_{A_k})$, there holds

$$\|g_k\|_p \leq C_q 2^{-10qk} \|f\|_p, \quad k \geq 0$$

Let ψ_k be a smooth partition of unity subordinated to $\{A_k\}$, and write $f_k := f * \widehat{\psi_k}$. Note that $\|f_k\|_p \leq C \|g_k\|_p \leq C_q 2^{-10qk} \|f\|_p$, and furthermore $|f_k| \leq Mf$ pointwise. This yields, comparing the definitions (2.8) and (7.17), that

$$\underline{E}_k := \underline{E}_{f_k, C_q^{-1} 2^{10qk} \lambda, p} \subset \underline{E}_{f,\lambda,p}$$

This observation, coupled with (4.3), and with the use of (2.9) for each f_k which has compact support in frequency, yields

$$\begin{aligned} \left\| F_\phi(f) \mathbf{1}_{\mathcal{Z} \setminus E_{f,\lambda,p}} \right\|_{L^q(\mathcal{Z}, \sigma_a, s)} &\leq \sum_{k=0}^{\infty} 2^{2k} \left\| F_\phi(f_k) \mathbf{1}_{\mathcal{Z} \setminus E_k} \right\|_{L^q(\mathcal{Z}, \sigma_a, s)} \\ &\leq C_{p,q,\alpha,\phi} \lambda^{1-\frac{p}{q}} \sum_{k=0}^{\infty} 2^{2k} 2^{10(q-p)k} \|f_k\|_p \leq C_{p,q,\alpha,\phi} \lambda^{1-\frac{p}{q}} \|f\|_p \end{aligned}$$

completing the reduction. In the argument that follows, we work with the definitions (7.17) in place of (2.8). For convenience, we drop the underline and return to the notation of (2.8).

7.3.2. *Main line of proof.* Let now $f \in \mathcal{S}(\mathbb{R})$ be fixed and have compact frequency support. By vertical scaling we can assume $\|f\|_p = 1$. To obtain estimate (2.9), it suffices to prove the analogue where $L^q(\mathcal{Z}, \sigma_a, s)$ is replaced by the corresponding outer $L^{q,\infty}$. The upgrade to outer L^q then follows from log-convexity of the norms of $F_\phi(f) \mathbf{1}_{\mathcal{Z} \setminus E_{f,\lambda,p}}$, see (4.2) from Proposition 4.1.

Furthermore, by the dilation invariance (4.4) described in Remark 4.4, it suffices to work with a single value of λ . We choose the value

$$(7.18) \quad \lambda = \lambda_{p,q,\alpha,\phi} := (2C_{\alpha,\phi})^{-1} 2^{-\frac{k_0}{q}} \leq 1,$$

where $C_{\alpha,\phi}$ is the constant appearing in (7.4) and 2^{k_0} is $2^{10} C_0 \varepsilon^{-1}$ times the greater of the two constants $C_{\alpha,\phi,Q}$ appearing in (7.3) and (7.6), for the choice $Q = 2^5 \varepsilon^{-1}$. This way, $k_0 = C_{p,q,\alpha,\phi}$. Here, C_0 is an absolute constant which will be specified later in the proof, see (8.13). We can use the already established (7.4) from Lemma 7.1 and obtain

$$\|F_\phi(f)\|_{L^\infty(\mathcal{Z} \setminus E_{f,\lambda,p}, \sigma_a, s)} \leq 2^{-\frac{k_0}{q}}.$$

The assertion of (2.9) can then be equivalently reformulated as follows: find $C_{p,q,\alpha,\phi}$ large enough such that

$$(7.19) \quad \sup_{k > k_0} 2^{-k} \nu \left(s(F) > 2^{-\frac{k}{q}} \right) \leq C_{p,q,\alpha,\phi}, \quad F := F_\phi(f) \mathbf{1}_{\mathcal{Z} \setminus E_{f,\lambda,p}}.$$

The collection of dyadic intervals

$$(7.20) \quad J = \left\{ J \in \mathcal{D} : \inf_{x \in J} M_p f(x) \leq \lambda \right\}$$

will play a role in our main line of proof, which begins now.

From now on, we fix $k > k_0$. The first step is the application of the selection Lemma 7.6, with λ replaced by $2^{-\frac{k}{q}}$, yielding an ∞ -strongly disjoint collection $\mathcal{T}_0$ and a 2-strongly disjoint collection $\mathcal{T}_1$, with the properties

$$(7.21) \quad \{I_T : T \in \mathcal{T}\} \subset J,$$

$$(7.22) \quad \nu_\alpha \left(s(F) > 2^{-\frac{k}{q}} \right) \leq \sum_{T \in \mathcal{T}} |I_T| =: 2^N,$$

$$(7.23) \quad \inf_{T \in \mathcal{T}_0} s_\infty(F)(T) > 2^{-\frac{k}{q}},$$

$$(7.24) \quad \sup_{T \in \mathcal{T}_1} \sup_{(u,t,\eta) \in T^*} |F(u,t,\eta)| \leq 2^{-\frac{k}{q}},$$

$$(7.25) \quad 2^{-\frac{k}{q}-4} \leq \left(\frac{1}{|I_T|} \int_{T^*} |F|^2 du dt d\eta \right)^{\frac{1}{2}} \leq 2^{-\frac{k}{q}} \quad \forall T \in \mathcal{T}_1.$$

We wrote for simplicity $\mathcal{T} = \mathcal{T}_0 \cup \mathcal{T}_1$. Property (7.21) follows by construction of $E_{f,\lambda,p}$, see (2.8). It is clear that (7.19) follows immediately if we prove that

$$(7.26) \quad 2^N \leq C_{p,q,\alpha,\phi} 2^k.$$

for a suitable constant. Of course, we can assume $N \geq k$ from now on, otherwise there is nothing to prove. By first passing to an arbitrary finite subcollection, and then by proving bounds independent on the cardinality of such subcollection, we can reduce to the case of $\mathcal{T}$ being finite. We will need to define the further (finite) parameters

$$(7.27) \quad 2^M := \max \left\{ 2^k, \sup_{J \in \mathcal{J}} \frac{1}{|J|} \sum_{T \in \mathcal{T}: I_T \subset J} |I_T| \right\}, \quad K = 2^{\varepsilon M}.$$

In particular, 2^M plays the role of BMO norm of the counting function in the next lemma. The proof (and in fact, the statement) is identical to that of [5, eq. (6.19)] and is thus omitted.

Lemma 7.7. *There exists a decomposition $\mathcal{T} = \mathcal{T}' \cup \mathcal{T}''$*

$$(7.28) \quad \left\| \sum_{T \in \mathcal{T}'} \mathbf{1}_{9KI_T} \right\|_\infty \leq 2^{9M},$$

$$(7.29) \quad \sum_{T \in \mathcal{T}''} |I_T| \leq 2^{-k}.$$

In view of (7.29), $\mathcal{T}''$ can be removed from $\mathcal{T}$ for all practical purposes related to (7.26). By doing this, (7.28) now holds for $\mathcal{T}$ in place of $\mathcal{T}'$. At this point, we enucleate the Calderón-Zygmund decomposition procedure in the following lemma. It is convenient to adopt the notations, for an interval $J \subset \mathbb{R}$,

$$\mathcal{T}_0(J) := \{T \in \mathcal{T}_0 : I_T \subset J\}, \quad \mathcal{T}_1(J) := \{T \in \mathcal{T}_1 : I_T \subset J\} \quad \mathcal{T}(J) = \mathcal{T}_0(J) \cup \mathcal{T}_1(J).$$

We think of the function g_J appearing in the Lemma below as the projection of f onto the time-frequency region corresponding to the generalized tents of $\mathcal{T}(J)$.

Lemma 7.8 (CZ decomposition). *Fix $J \in \mathcal{J}$. We can find $g^{(J)} \in L^2(\mathbb{R})$ with the properties*

$$(7.30) \quad \begin{aligned} & \text{supp } g^{(J)} \subset 9KJ \\ & \|g^{(J)}\|_2^2 \lesssim \varepsilon^{-1} \min \left\{ 2^{(1-\frac{2}{q})M} |J|, 2^{3\varepsilon M} 2^{(1-\frac{2}{p'})N} \right\} \end{aligned}$$

and

$$(7.31) \quad \begin{aligned} & \inf_{T_n \in \mathcal{T}_0} |F_\phi(g^{(J)})(u_n, t_n, \eta_n)| > 2^{-\frac{k}{q}-1}, \\ & \sup_{T \in \mathcal{T}_1(J)} \sup_{(u,t,\eta) \in T^*} |F_\phi(g^{(J)})(u, t, \eta)| \leq 2^{-\frac{k}{q}+1}, \\ & 2^{-\frac{k}{q}-5} \leq \left(\frac{1}{|I_T|} \int_{T^*} |F_\phi(g^{(J)})|^2 du dt d\eta \right)^{\frac{1}{2}} \leq 2^{-\frac{k}{q}+1} \quad \forall T \in \mathcal{T}_1(J). \end{aligned}$$

With Lemma 7.8 in hand, we are able to estimate

$$(7.32) \quad 2^M \leq C_{\alpha, \phi, \bar{\varepsilon}, q} 2^k.$$

Suppose $M > k$ (otherwise there is nothing to prove) and choose an interval J such that the supremum in (7.27) is nearly attained. We have, using the estimates of (7.31), property (7.9) of strongly disjoint collections for $\lambda = 2^{-\frac{k}{q}}$, and the first bound of (7.30) in the final estimate, that

$$2^{M-1} |J| \leq \sum_{T \in \mathcal{T}_0(J)} |I_T| + \sum_{T \in \mathcal{T}_1(J)} |I_T| \leq C_{\alpha, \phi} 2^{\frac{2k}{q}} \|g^{(J)}\|_2^2 \leq C_{\alpha, \phi} \varepsilon^{-1} 2^{\frac{2k}{q}} 2^{(1-\frac{2}{q})M} |J|$$

and the bound (7.32) follows by rearranging. Now, take J large enough such that $\mathcal{T}(J) = \mathcal{T}$. Such a J exists since $\mathcal{T}$ is finite. By the same token, but this time using the estimate in (7.30) which does not depend on $|J|$, and (7.32), we obtain

$$2^N = \sum_{T \in \mathcal{T}_0} |I_T| + \sum_{T \in \mathcal{T}_1} |I_T| \leq C_{\alpha, \phi} 2^{\frac{2k}{q}} \|g^{(J)}\|_2^2 \leq C_{\alpha, \phi, \bar{\varepsilon}, q} 2^{(\frac{2}{q}+3\varepsilon)k} 2^{(1-\frac{2}{p'})N} \leq C_{\alpha, \phi, \bar{\varepsilon}, q} 2^{\frac{2}{p'}k} 2^{(1-\frac{2}{p'})N},$$

and the estimate (7.26) follows by rearranging. The proof of (2.9) is complete up to the CZ Lemma 7.8.

8. PROOF OF THE CALDERÓN-ZYGMUND LEMMA 7.8

Before we enter the argument for Lemma 7.8, we state the following projection Lemma, which will be used in the construction of the functions $g^{(J)}$.

Lemma 8.1. *Let $f \in L^p(\mathbb{R})$, $1 \leq p < 2$. Let $L \subset \mathbb{R}$ be an interval and*

$$\lambda = |L|^{-\frac{1}{p}} \|f \mathbf{1}_L\|_p.$$

Let $\Xi \subset \mathbb{R}$ be a finite set and $Q \geq 0$ an integer. Then $f \mathbf{1}_L = g + b$ with $\text{supp } g, \text{supp } b \subset 3L$, and

$$\begin{aligned} \|g\|_2^2 &\lesssim \lambda^2 (Q \# \Xi)^{1-\frac{2}{p'}} |L|, \\ \int b(x) x^j e^{-i\xi x} dx &= 0 \quad \forall j = 0, \dots, Q, \xi \in \Xi. \end{aligned}$$

The proof of the lemma is a projection argument on the finite dimensional subspace of $L^2(3L)$ given by

$$\mathcal{V}_{\Xi, Q} = \text{span}\{x^j e^{-i\xi x} : j = 0, \dots, Q, \xi \in \Xi\}$$

which is translation invariant when viewed as a subspace of $L^2(L)$. We give a proof of the more general Hilbert space version in the appendix, Lemma A.1 therein.

8.1. Construction of $g^{(J)}$. Recalling, from (7.27), the definition of K , set

$$L := \max. \text{ dyad. int. } L \quad \text{s.t.} \quad \|f \mathbf{1}_L\|_p > 2^{10} K |L|^{\frac{1}{p}}, \quad \mathbb{L} = \bigcup_{L \in \mathbb{L}} L.$$

The intervals in $\mathbb{L}$ are pairwise disjoint. We have that

$$(8.1) \quad \begin{aligned} \|f \mathbf{1}_L\|_p &\lesssim K |L|^{\frac{1}{p}}, \\ \sup_{I \subset \mathbb{R}} \frac{1}{|I|} \sum_{L \in \mathbb{L}: 3L \subset I} |L| &\leq 1, \end{aligned}$$

$$(8.2) \quad E_{3KL} = \bigcup_{L \in \mathbb{L}} T(9KL) \times \mathbb{R} \subset E_{f, \lambda, p}.$$

Note that (8.1) means that $\{3L : L \in \mathbb{L}\}$ have finite overlap in the sense of (6.24). To obtain the third property, observe that each $L \in \mathbb{L}$ is contained in some $I \in \mathcal{I}_{f, \lambda, p}$, and that $\lambda \leq 1$, see (7.18). It must actually be that $9KL$ is contained in $3I$, since

$$\frac{|I|}{|L|} \geq \frac{2^{-3p} \|f \mathbf{1}_I\|_p^p}{2^{-10p} K^{-p} \|f \mathbf{1}_L\|_p^p} \geq 2^7 K.$$

By the same token, referring to definition (7.20), we see that

$$(8.3) \quad J \in \mathcal{J}, L \in \mathbb{L}, 3L \cap 3KJ \neq \emptyset \implies 3L \subset 9KJ.$$

Denote $\mathbb{L}(J) = \{L \in \mathbb{L} : 3L \cap 3KJ \neq \emptyset\}$. It will be convenient to observe that, using disjointness

$$(8.4) \quad \sum_{L \in \mathbb{L}(J)} |L| \lesssim \min \left\{ K |J|, (2^{10} K)^{-p} \|f\|_p^p \right\} \lesssim K \min \{|J|, 1\}.$$

From now on, $J \in \mathcal{J}$ is fixed. We come to the definition of the significant set of frequencies. Using the notation $T = T(I_T, \xi_T)$ for $T \in \mathcal{T}$, we set

$$(8.5) \quad \Xi_L := \{\xi : \exists T \in \mathcal{T}(J) \text{ with } \xi = \xi_T, 3L \subset 9KI_T\}.$$

A consequence of the bound (7.28) is that

$$(8.6) \quad \#\Xi_L \leq \inf_{x \in 3L} \sum_{T \in \mathcal{T}(J)} \mathbf{1}_{9KI_T}(x) \leq 2^{9M}.$$

We can use the first inequality in (8.6) above, together with disjointness of $L \in \mathcal{L}(J)$ to obtain

$$(8.7) \quad \sum_{L \in \mathcal{L}(J)} |L| \#\Xi_L \leq \left\| \sum_{T \in \mathcal{T}(J)} \mathbf{1}_{9KI_T} \right\|_1 \lesssim K \min \{2^M |J|, 2^N\}.$$

See (7.27) and (7.22) for the definition of M, N . For all $L \in \mathcal{L}(J)$, we apply Lemma 8.1 to $f = f \mathbf{1}_L$, suitably rescaled, with the choice $\Xi = \Xi_L$ and $Q = \lceil 2^5 \varepsilon^{-1} \rceil$. This choice of Q will be kept throughout the remainder of the proof. The functions g_L and b_L have the following properties:

$$(8.8) \quad f \mathbf{1}_L = g_L + b_L, \quad \text{supp } g_L, \text{supp } b_L \subset 3L,$$

$$(8.9) \quad \|g_L\|_2^2 \lesssim \varepsilon^{-1} K^2 (\#\Xi_L)^{1-\frac{2}{p'}} |L|,$$

$$(8.10) \quad \|b_L\|_1 \lesssim \varepsilon^{-1} K (\#\Xi_L)^{\frac{1}{2}-\frac{1}{p'}} |L| \lesssim \varepsilon^{-1} 2^{6M} |L|,$$

$$(8.11) \quad \int b_L(x) x^j e^{-i\xi x} dx = 0 \quad \forall j = 0, \dots, Q, \forall \xi \in \Xi_L.$$

The second inequality in (8.10) is a consequence of (8.6). We finally set

$$(8.12) \quad g^{(J)} := f \mathbf{1}_{3KJ \setminus \mathbb{L}} + \sum_{L \in \mathcal{L}(J)} g_L, \quad b := \sum_{L \in \mathcal{L}(J)} b_L.$$

Note that, taking advantage of (8.10) and of the finite overlap (8.1), one may argue as in Remark 6.8 to obtain that

$$\{x \in \mathbb{R} : Mb(x) \gtrsim \varepsilon^{-1} 2^{7M}\} \subset \bigcup_{L \in \mathcal{L}} 3L \subset \{x \in \mathbb{R} : M_p f(x) \geq 2^3 K\}.$$

Hence, recalling the definition of J ,

$$(8.13) \quad \sup_{I \in J} \inf_{x \in I} Mb(x) \lesssim \varepsilon^{-1} 2^{7M}.$$

We now choose C_0 to be the larger of the two implicit absolute constants in (8.10) and (8.13). It can be easily tracked that $C_0 \leq 2^{12}$. Recalling (7.21), we will use this later with $I = I_T$, $T \in \mathcal{T}$.

8.2. Proof of (7.30)-(7.31).

Proof of (7.30). Note that $\|f \mathbf{1}_{3KJ \setminus \mathbb{L}}\|_\infty \lesssim K$. Therefore

$$(8.14) \quad \|f \mathbf{1}_{3KJ \setminus \mathbb{L}}\|_2^2 \lesssim \min \{K^2 |J|, K^{2-p} \|f\|_p^p\} \leq 2^{2\varepsilon M} \min \{|J|, 1\}.$$

Relying on (8.9) and on the observation that $\{3L : L \in L_j\}$ have bounded overlap,

$$\begin{aligned} \left\| \sum_{L \in L(J)} g_L \right\|_2^2 &\lesssim \sum_{L \in L(J)} \|g_L\|_2^2 \lesssim \varepsilon^{-1} K^2 \left(\sum_{L \in L(J)} |L| \right)^{\frac{2}{p'}} \left(\sum_{L \in L(J)} |L| \# \Xi_L \right)^{1 - \frac{2}{p'}} \\ &\lesssim \varepsilon^{-1} K^3 \min \left\{ 2^{M(1 - \frac{2}{p'})} |J|, 2^{N(1 - \frac{2}{p'})} \right\} \lesssim \varepsilon^{-1} \min \left\{ 2^{M(1 - \frac{2}{q})} |J|, 2^{3\varepsilon M} 2^{N(1 - \frac{2}{p'})} \right\} \end{aligned}$$

where the penultimate estimate follows from (8.4) and (8.7), and the last by recalling the definition of ε in (7.16) and of K in (7.27). The proof of (7.30) is finished by combining the last display with (8.14). $\square$

Proof of (7.31). First of all, we note that

$$f = g^{(J)} + b + h, \quad h := f \mathbf{1}_{(\mathbb{L})^c \cap (3KJ)^c} + \sum_{\substack{L \in L \\ 3L \cap 3KJ = \emptyset}} f \mathbf{1}_L.$$

Also note that

$$\begin{aligned} |F_\phi(g^{(J)})| &\geq |F_\phi(g^{(J)})| \mathbf{1}_{\mathcal{Z} \setminus E_{f, \lambda, p}} \geq (|F_\phi(f)| - |F_\phi(b)| - |F_\phi(h)|) \mathbf{1}_{\mathcal{Z} \setminus E_{f, \lambda, p}} \\ &\geq F - (|F_\phi(b)| \mathbf{1}_{\mathcal{Z} \setminus E_{3KL}} + |F_\phi(h)|) \end{aligned}$$

where the inclusion (8.2) was used in the last inequality. Therefore, by comparison with (7.23)-(7.12), (7.31) will follow if we prove

$$(8.15) \quad \sup_{T \in \mathcal{T}(J)} s(F_\phi(h))(T) \leq 2^{-10k},$$

$$(8.16) \quad \sup_{T \in \mathcal{T}(J)} s(F_\phi(b) \mathbf{1}_{\mathcal{Z} \setminus E_{3KL}})(T) \leq 2^{-10k}.$$

We have used that $s \geq \max\{s_2, s_\infty\}$ and the obvious fact $s_2(H)(T) \geq (|I_T|^{-1} \int_{T^*} |H|^2)^{1/2}$.

By virtue of the fact that $\text{supp } h \cap 3KJ = \emptyset$, we can apply the intermediate estimate (7.3) to h with the collection J replaced by $\{J\}$, obtaining the chain of inequalities

$$\begin{aligned} (8.17) \quad &\sup_{T \in \mathcal{T}(J)} s(F_\phi(h))(T) \leq \|F_\phi(h) \mathbf{1}_{E_{\{J\}}}\|_{L^\infty(\mathcal{Z}, \sigma_a, s)} \\ &\leq C_{\alpha, \phi, Q} K^{-Q} \inf_{x \in J} Mh(x) \leq C_{\alpha, \phi, Q} K^{-Q} \inf_{x \in J} Mf(x) \\ &\leq C_{\alpha, \phi, Q} K^{-Q} \lambda \leq C_{\alpha, \phi, Q} 2^{-k\varepsilon Q} \leq C_{\alpha, \phi, Q} 2^{-32k} \leq 2^{-10k} \end{aligned}$$

from which (8.15) follows. We have used $h \leq f$ in the last step on the second line and (7.20) in the passage to the third line. The last step in (8.17) was obtained by recalling from (7.27) that $k \leq M$, that $Q = \lceil 2^5 \varepsilon^{-1} \rceil$ and that k_0 was chosen in order to have $C_{\alpha, \phi, Q} \leq 2^{k_0} \leq 2^k$.

We turn to (8.16). Fix a tent $T = T(I, \xi) \in \mathcal{T}(J)$. Define

$$L(I) = \{L \in L(J) : L \cap 3KI \neq \emptyset\}, \quad b^{\text{in}} := \sum_{L \in L(I)} b_L, \quad b^{\text{out}} = b - b^{\text{in}}.$$

The b^{out} part, supported outside $3KI$, is also handled via (7.3). Using below that $Mb^{\text{out}} \leq Mb$ pointwise, and later (8.13), we actually estimate the larger quantity

$$\begin{aligned} (8.18) \quad &s(F_\phi(b^{\text{out}}))(T) \leq \|F_\phi(b^{\text{out}}) \mathbf{1}_{E_{\{I\}}}\|_{L^\infty(\mathcal{Z}, \sigma_a, s)} \\ &\leq C_{\alpha, \phi, Q} K^{-Q} \inf_{x \in I} Mb(x) \leq C_0 \varepsilon^{-1} C_{\alpha, \phi, Q} 2^{M(7 - Q\varepsilon)} \leq 2^{-10k}. \end{aligned}$$

The last step has been obtained since $Q \geq 2^5 \varepsilon^{-1}$ and later using $M \geq k$ and the definition of k_0 (7.18). Now we have to control the b^{in} part. Observe that by our choice of Ξ_L (8.5), and relying on (8.3), $\xi \in \Xi_L$ for all $L \in L(I)$. Thus, taking also (8.8), (8.10) and (8.11) into account, we have

$$\begin{aligned} \text{supp } b_L &\subset 3L, & \|b_L\|_1 &\leq C_0 \varepsilon^{-1} 2^{6M} |L|, \\ \int b_L(x) x^j e^{-i\xi x} dx &= 0 & \forall j = 0, \dots, 2^5 \varepsilon^{-1}, \end{aligned}$$

for all $L \in L(I)$. In other words, also in view of the finite overlap (8.1), comparing with Remark 6.7, b^{in} satisfies the assumptions of Lemma 7.2 with $\lambda = 3C_0 \varepsilon^{-1} 2^{6M}$, L replaced by $\{3L : L \in L(I)\}$, and $Q = 2^5 \varepsilon^{-1}$. Applying the lemma, we obtain

$$s(F_\phi(b^{\text{in}}) \mathbf{1}_{\mathcal{Z} \setminus E_{3KL}})(T) \leq 3C_0 \varepsilon^{-1} C_{\alpha, \phi, Q} 2^{M(6-Q\varepsilon)} \leq 2^{-10k}.$$

The last inequality is obtained in the same fashion as (8.18). Combining (8.18) with the last display, we have finished the proof of (8.16), and, in turn, of (7.31). $\square$

APPENDIX A. REMARKS ON THE HILBERT SPACE VALUED SETTING

We would like to remark that our main theorem can be easily generalized to the setting of Hilbert space valued functions. Let $\mathcal{H}$ be a separable Hilbert space with orthonormal basis $\{h_n : n \in \mathbb{N}\}$. By $L^p(\mathbb{R}; \mathcal{H})$ we denote the usual $\mathcal{H}$ -valued Bochner spaces. The inner product associated to $\mathcal{H}$ is denoted as $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. Then for Schwartz functions $f_1, f_2 : \mathbb{R} \rightarrow \mathcal{H}$ and $f_3 : \mathbb{R} \rightarrow \mathbb{C}$, one can naturally define

$$(A.1) \quad V_{\vec{\beta}}(f_1, f_2, f_3) = \int_{\mathbb{R} \times (0, \infty) \times \mathbb{R}} \langle G_1, G_2 \rangle_{\mathcal{H}} G_3 \, du dt d\eta, \quad G_j(u, t, \eta) := F_\phi(f_j)(u, t, \alpha_j \eta + \beta_j t^{-1}).$$

where $\vec{\beta}$ is as before. We then have the following theorem, which is the Hilbert space valued analog of Theorem 2.

Theorem 3. *Let*

$$(A.2) \quad 1 < p_1, p_2 \leq \infty, \quad \frac{2}{3} < r := \frac{p_1 p_2}{p_1 + p_2} < \infty.$$

For all sets $A \subset \mathbb{R}$ of finite measure, and for all Schwartz functions $f_1, f_2 : \mathbb{R} \rightarrow \mathcal{H}$ we can find a subset $\tilde{A} \subset A$ such that $|A| \leq 2|\tilde{A}|$ and

$$|V_{\vec{\beta}}(f_1, f_2, f_3)| \leq C_{\vec{\beta}, p_1, p_2} \|f_1\|_{L^{p_1}(\mathbb{R}; \mathcal{H})} \|f_2\|_{L^{p_2}(\mathbb{R}; \mathcal{H})} |A|^{1-\frac{1}{r}} \quad \forall |f_3| \leq \mathbf{1}_{\tilde{A}}.$$

We remark that Theorem 3 implies the Hilbert space valued version of the full range of estimates for the bilinear Hilbert transform, which was not known before. The proof of Theorem 3 utilizes the Carleson embedding theorem in the $\mathcal{H}$ -valued setting, which can be easily generalized from the scalar valued case (Theorem 1) that has been presented, with the proper adaptation of sizes and outer L^p spaces. Roughly speaking, the $\mathcal{H}$ -adapted sizes are the same as in the scalar valued case but with the absolute value of F replaced by the

$\mathcal{H}$ norm. For example, in the case of generalized tents, let $F : \mathcal{Z} \rightarrow \mathcal{H}$ measurable, the sizes are defined as

$$\begin{aligned} s_2(F)(T(I, \xi)) &:= \left(\frac{1}{|I|} \int_{T_a^t(I, \xi)} \|F(u, t, \eta)\|_{\mathcal{H}}^2 du dt d\eta \right)^{\frac{1}{2}}, \\ s_\infty(F)(T(I, \xi)) &:= \sup_{(u, t, \eta) \in T_a(I, \xi)} \|F(u, t, \eta)\|_{\mathcal{H}}, \\ s &:= s_2 + s_\infty. \end{aligned}$$

Similar basic properties hold for outer L^p spaces in this setting, which we refer to [4] for details. In fact, the only part of the argument in the proof of Theorem 3 that might need some further explanation is the adapted version of Lemma 8.1, which we state as the following.

Lemma A.1. *Let $f \in L^p(\mathbb{R}; \mathcal{H})$, $1 \leq p < 2$. Let $L \subset \mathbb{R}$ be an interval and*

$$\lambda = |L|^{-\frac{1}{p}} \|f \mathbf{1}_L\|_{L^p(\mathbb{R}; \mathcal{H})}.$$

Let $\Xi \subset \mathbb{R}$ be a finite set and $Q \geq 0$ an integer. Then $f \mathbf{1}_L = g + b$ with $\text{supp } g, \text{supp } b \subset 3L$, and

$$\begin{aligned} \|g\|_{L^2(\mathbb{R}; \mathcal{H})}^2 &\lesssim \lambda^2 (Q \# \Xi)^{1 - \frac{2}{p'}} |L|, \\ \int b(x) x^j e^{-i\xi x} dx &= 0 \quad \forall j = 0, \dots, Q, \xi \in \Xi. \end{aligned}$$

This projection lemma extends the result of [1, 15] to the Hilbert space valued setting, as well as to a more general translation invariant finite dimensional subspace of $L^2(3L)$, which was unknown before even in the scalar-valued setting. We will need the following preliminary lemma in order to prove Lemma A.1.

Lemma A.2. *Let $v_j : \mathbb{R} \rightarrow \mathbb{C}$, $j = 1, \dots, N$ be continuous functions with the property that*

$$v \in V := \text{span}_{\mathbb{C}}\{v_j : j = 1, \dots, N\} \implies v(\cdot + t) \in V \quad \forall t \in \mathbb{R}.$$

Then, for all intervals $I \subset \mathbb{R}$

$$\sup_{x \in I} |v(x)| \leq \sqrt{\frac{N}{|I|} \int_{3I} |v(x)|^2 dx} \quad \forall v \in V, \forall R > 0.$$

Proof. By translation and scaling invariance of the assumptions and conclusions, we can reduce to the case $I = (-1/2, 1/2)$. Consider $V_I = \text{span}_{\mathbb{C}}\{v_j \mathbf{1}_I, j = 1, \dots, N\}$ as a linear subspace of $L^2(I)$ of dimension $n \leq N$. Let $\{w_1, \dots, w_n\}$ be an orthonormal basis of V_I . Since

$$\int_I \sum_{j=1}^n |w_j(x)|^2 dx = n$$

there exists $x_0 \in I$ such that

$$|I| \sum_{j=1}^n |w_j(x_0)|^2 \leq n \leq N.$$

Fix $v \in V_I$, and $x \in I$, and denote $v(\cdot) = \tilde{v}(\cdot + x_0 - x)\mathbf{1}_I(\cdot) \in V_I$. We then have

$$\begin{aligned} |v(x)|^2 &= |\tilde{v}(x_0)|^2 = \left| \sum_{j=1}^n \langle \tilde{v}, w_j \rangle w_j(x_0) \right|^2 \leq \left(\sum_{j=1}^n |w_j(x_0)|^2 \right) \left(\sum_{j=1}^n |\langle \tilde{v}, w_j \rangle|^2 \right) \\ &\leq \frac{N}{|I|} \int_I |\tilde{v}(t)|^2 dt \leq \frac{N}{|I|} \int_{3I} |v(t)|^2 dt \end{aligned}$$

which concludes the proof. $\square$

Proof of Lemma A.1. There is no loss in generality with assuming that $\mathcal{H}$ is finite dimensional and $\{h_n\}$ is finite with cardinality d . By scaling invariance, we can reduce to the case $|I| = 1$. The linear space

$$V = \text{span}_{\mathbb{C}} \{x \mapsto x^j e^{-i\xi_\ell x}, j = 0, \dots, Q, \ell = 1, \dots, \#\Xi\}, \quad N = (Q + 1)\#\Xi.$$

satisfies the assumptions of Lemma A.2 and let us use the same notation as in the lemma for V_I and its orthonormal basis $\{w_1, \dots, w_N\}$. Let

$$\mathcal{V} = \left\{ h = \sum_{j=1}^d u_j h_j, u_j \in V, h_j \in \mathcal{H} \right\}.$$

We claim that

$$(A.3) \quad \sup_{x \in I} \|h(x)\|_{\mathcal{H}} = \sup_{x \in I} \left(\sum_{j=1}^d |\langle h(x), h_j \rangle|^2 \right)^{\frac{1}{2}} \leq N^{\frac{1}{2}} \|h \mathbf{1}_{3I}\|_{L^2(\mathbb{R}; \mathcal{H})}$$

which implies by interpolation that

$$(A.4) \quad \|h \mathbf{1}_I\|_{L^{p'}(\mathbb{R}; \mathcal{H})} \leq N^{\frac{1}{2} - \frac{1}{p'}} \|h \mathbf{1}_{3I}\|_{L^2(\mathbb{R}; \mathcal{H})}$$

which is the one we will use. To prove (A.3), fix $x \in I$. Noting that $\langle h(\cdot), h_j \rangle = u_j \in V$, we apply Lemma A.2 and obtain

$$|\langle h(x), h_j \rangle|^2 \leq N \int_{3I} |\langle h(x), h_j \rangle|^2 dx$$

therefore

$$\sum_{j=1}^d |\langle h(x), h_j \rangle|^2 \leq N \int_{3I} \sum_{j=1}^d |\langle h(x), h_j \rangle|^2 dx = N \int_{3I} \|h(x)\|_{\mathcal{H}}^2 dx = N \|h \mathbf{1}_{3I}\|_{L^2(\mathbb{R}; \mathcal{H})}^2$$

as claimed. Using (A.4) and the assumption on f , $f \mathbf{1}_I$ defines a linear functional on the restriction of $\mathcal{V}$ to $3I$ viewed as a subspace of $L^2(3I; \mathcal{H})$, with norm bounded by $\lambda N^{\frac{1}{2} - \frac{1}{p'}}$. Then, using the Riesz representation theorem there exists $g \in L^2(3I; \mathcal{H})$ such that

$$\int_I \langle f(x), h \rangle_{\mathcal{H}} dx = \int_{3I} \langle g(x), h \rangle_{\mathcal{H}} dx \quad \forall h \in \mathcal{V}.$$

and satisfying the norm inequality. Defining the $\mathcal{H}$ -valued function $b = f\mathbf{1}_I - g$ whose support is in $3I$, we have

$$\int_{3I} \langle b(x), h_j \rangle_{\mathcal{H}} x^k e^{-i\xi_\ell x} dx = 0 \quad \forall j, k, \ell,$$

which implies that b satisfies the vanishing moments conditions. The proof is complete. $\square$

REFERENCES

- [1] Peter Borwein and Tamás Erdélyi, *Nikolskii-type inequalities for shift invariant function spaces*, Proc. Amer. Math. Soc. **134** (2006), no. 11, 3243–3246. MR 2231907 (2007e:41014)
- [2] R. R. Coifman, Y. Meyer, and E. M. Stein, *Some new function spaces and their applications to harmonic analysis*, J. Funct. Anal. **62** (1985), no. 2, 304–335. MR 791851 (86i:46029)
- [3] Francesco Di Plinio, *Weak- L^p bounds for the Carleson and Walsh-Carleson operators*, C. R. Math. Acad. Sci. Paris **352** (2014), no. 4, 327–331. MR 3186922
- [4] Francesco Di Plinio and Yumeng Ou, *Banach-valued multilinear singular integrals*, preprint arXiv:1506.05827.
- [5] Francesco Di Plinio and Christoph Thiele, *Endpoint bounds for the bilinear Hilbert transform*, Trans. Amer. Math. Soc. **368** (2016), no. 6, 3931–3972. MR 3453362
- [6] Yen Do and Christoph Thiele, *L^p theory for outer measures and two themes of Lennart Carleson united*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 2, 249–296. MR 3312633
- [7] Loukas Grafakos and Xiaochun Li, *Uniform bounds for the bilinear Hilbert transforms. I*, Ann. of Math. (2) **159** (2004), no. 3, 889–933. MR 2113017 (2006e:42011)
- [8] Tuomas P. Hytönen and Michael T. Lacey, *Pointwise convergence of vector-valued Fourier series*, Math. Ann. **357** (2013), no. 4, 1329–1361. MR 3124934
- [9] 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. MR 1491450 (99b:42014)
- [10] ———, *On Calderón’s conjecture*, Ann. of Math. (2) **149** (1999), no. 2, 475–496. MR 1689336 (2000d:42003)
- [11] Xiaochun Li, *Uniform bounds for the bilinear Hilbert transforms. II*, Rev. Mat. Iberoam. **22** (2006), no. 3, 1069–1126. MR 2320411 (2008c:42014)
- [12] Camil Muscalu, Jill Pipher, Terence Tao, and Christoph Thiele, *Multi-parameter paraproducts*, Rev. Mat. Iberoam. **22** (2006), no. 3, 963–976. MR 2320408 (2008b:42037)
- [13] Camil Muscalu and Wilhelm Schlag, *Classical and multilinear harmonic analysis. Vol. II*, Cambridge Studies in Advanced Mathematics, vol. 138, Cambridge University Press, Cambridge, 2013. MR 3052499
- [14] Camil Muscalu, Terence Tao, and Christoph Thiele, *Multi-linear operators given by singular multipliers*, J. Amer. Math. Soc. **15** (2002), no. 2, 469–496. MR 1887641 (2003b:42017)
- [15] Fedor Nazarov, Richard Oberlin, and Christoph Thiele, *A Calderón-Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain*, Math. Res. Lett. **17** (2010), no. 3, 529–545. MR 2653686 (2011d:42047)
- [16] Richard Oberlin and Christoph Thiele, *New uniform bounds for a Walsh model of the bilinear Hilbert transform*, Indiana Univ. Math. J. **60** (2011), no. 5, 1693–1712. MR 2997005
- [17] Christoph Thiele, *A uniform estimate*, Ann. of Math. (2) **156** (2002), no. 2, 519–563. MR 1933076 (2003i:47036)
- [18] ———, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics, vol. 105, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006. MR 2199086 (2006m:42073)

BROWN UNIVERSITY MATHEMATICS DEPARTMENT, BOX 1917, PROVIDENCE, RI 02912, USA
E-mail address: `fradipli@math.brown.edu` (F. Di Plinio)

BROWN UNIVERSITY MATHEMATICS DEPARTMENT, BOX 1917, PROVIDENCE, RI 02912, USA
E-mail address: `yumeng_ou@brown.edu` (Y. Ou)