VARIATIONAL CARLESON EMBEDDINGS INTO THE UPPER 3-SPACE.

GENNADY URALTSEV

Abstract. In this paper we formulate embedding maps into time-frequency space related to the Carleson operator and its variational counterpart. We prove bounds for these embedding maps by iterating the outer measure theory of \cite{DT15}. Introducing iterated outer $L^p$ spaces is a main novelty of this paper.

1. Introduction

In this paper we consider the Carleson Operator

\begin{equation}
C_c f(z) := \int_{c(z)}^{+\infty} \hat{f}(\xi)e^{i\xi z}d\xi,
\end{equation}

with $c: \mathbb{R} \to \mathbb{R}$ a Borel-measurable stopping function. The Variational Carleson Operator studied by Oberlin et al. in \cite{Obe+12} is given by:

\begin{equation}
V^r C_c f(z) = \left( \sum_{k \in \mathbb{Z}} \left| C_{c_{k+1}} f(z) - C_{c_k} f(z) \right|^r \right)^{1/r},
\end{equation}

where $c: \mathbb{Z} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a stopping sequence of Borel-measurable functions such that $c_k(z) \leq c_{k+1}(z)$ for all $z \in \mathbb{R}$ and $k \in \mathbb{Z}$. The boundedness on $L^p(\mathbb{R})$ with $p \in (1, \infty)$ of these operators, uniformly with respect to the stopping functions $c$ and $c'$, implies the famous Carleson Theorem on the almost everywhere convergence of the Fourier integral for functions in $L^p(\mathbb{R})$. The main technique for bounding these operators were first introduced by Carleson in his paper \cite{Car66} on the convergence of Fourier series for $L^2([-\pi/2, \pi/2])$ functions and is often referred to as time-frequency analysis.

The purpose of this paper is to discuss embedding maps into time-frequency space $X = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ relevant to (1.1) and (1.2). In Theorems 1.1, 1.2, and 1.3 we show the boundedness properties of these embedding maps in terms of appropriately defined norms. Generally speaking an embedding map is a representation of a function by another function defined on the symmetry group of the problem at hand. The appropriate norms for dealing with these embedded functions are the outer measure $L^p$ norms introduced in \cite{DT15} in the context of the Bilinear Hilbert Transform, an operator with the same symmetries as (1.1) and (1.2).

Theorem 1.2 is an extension of the result of \cite{DT15} to $1 < p < 2$. For our proof we introduce iterated, or semi-direct product, outer measure $L^p$ spaces and incorporate the idea by Di Plinio and Ou \cite{DPO15} of using multi-frequency Calderón-Zygmund theory from \cite{NOT10}. The embedding Theorems 1.1 and 1.3 are somewhat dual to 1.2 for the purpose of bounding the bilinear form associated to (1.1) and (1.2) respectively.

In \cite{Obe+12} the operator (1.2) has been shown to be bounded for $p \in (1, \infty)$ and $r \in (2, p')$. The proof in the range $p \in (2, r)$ requires only theorems that make use of non-iterated outer measure spaces of \cite{DT15}. While initially introduced
only to address the range \( p \in (r', 2] \), iterated outer measure spaces surprisingly provide a direct proof in the complete range \( p \in [r, \infty) \), and hereby explain ad-hoc interpolation techniques used in [Obe+12].

The advantage of reasoning in terms of embedding maps is also attested by the recent developments in [CDPO16] that prove sharp weighted bounds for the Bilinear Hilbert Transform using the embedding from [DPO15]. This is done by dominating the trilinear form associated to the operator by sparse forms following the approach of Lacey [Lac15]. In a similar spirit, sparse domination and weighted boundedness results for the Variational Carleson Operator are forthcoming in a paper by the Di Plinio, Do, and the author, that make use of the embedding maps of the present paper.

We also point out the recent paper [DMT16] in which Do, Muscalu, and Thiele use outer measure \( L^p \) spaces to provide variational bounds for bilinear Fourier inversion integrals, that are bilinear versions of (1.2).

On a historical note, we point out Hunt’s extension [Hun68] to \( L^p \) with \( p \in (1, \infty) \) of Carleson’s pointwise almost-everywhere convergence result [Car66] for Fourier series of functions on \( L^2([-\pi/2, \pi/2]) \). Carleson’s and Hunt’s results depend on a fine analysis of the properties of a function on the torus. In [Fe73] Fefferman concentrated on proving the same result by a careful study of the operator (1.1). The wave-packet representation for the operator that is crucial for making use of embedding maps appeared in [LT00] that provides a more symmetric approach encompassing the aforementioned two ideas. This approach inspired both [Obe+12] and the present paper.

Finally, we emphasize that we formulate an embedding map into the time-frequency space parameterized by continuous parameters, in the vein of [DT15]. This allows us to avoid model-sum operators and averaging procedures ubiquitous in other works in time-frequency analysis. Furthermore, such a formulation proves to be more versatile and in particular the results of the present paper imply all the bounds for the discretized model used in [Obe+12].

1.1. The Carleson operator. For simplicity we begin by discussing the Carleson operator (1.1) that is a specific instance of (1.2) for \( r = +\infty \). The operator is given pointwise by the Fourier multiplier operator associated to the multiplier \( \mathbb{1}_{c(z), +\infty}(\xi) \) applied to \( f \). This can be expressed in terms of a wavelet frame centered at frequency \( c(z) \) using a continuous Littlewood-Paley decomposition:

\[
C_c f(z) = \int_{\mathbb{R}^+} \int_{\mathbb{R}} f * \psi_{\eta,t} \ast \psi_{\eta,t}(z) \chi(t(\eta - c(z))) \, dt \, d\eta
\]  

(1.3)

where

\[
\psi_{\eta,t}(z) := t^{-1} e^{i\eta z} \psi\left(\frac{z}{t}\right)
\]  

(1.4)

with \( \psi \in S(\mathbb{R}) \) a suitably normalized, non-negative, even, generating wavelet with Fourier transform \( \hat{\psi} \) supported in a small ball \( B_0 \). We use the notation \( B_r(x) := (x - r, x + r) \) to denote a ball of radius \( r \) centered at \( x \), while if \( x = 0 \) we omit it.
by simply writing $B_r$. The non-negative cutoff function $\chi$ satisfies
\begin{equation}
\chi \in C^\infty_c(B_r(d)) \quad B_r(d) \subset (b, +\infty) \quad \int \chi = 1.
\end{equation}

Given two functions $f, a \in S(\mathbb{R})$ set
\begin{align}
F(y, \eta, t) &:= f \ast \psi_{\eta, t}(y) \\
A(y, \eta, t) &:= \int_{\mathbb{R}} a(z) \psi_{\eta, t}(y-z) \chi(t(\eta - c(z))) \, dz.
\end{align}

The arguments of the above functions are points of the time-frequency space $X = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ that parameterizes the defining symmetries of the class of operators defined by (1.1) i.e. translation of the function, translation of its Fourier transform, and dilation. The outer measure $L^p$ spaces allow one to deal with the overderminicncy of the wave-packets.

The wave packet representation (1.3) gives the inequality
\begin{equation}
\left| \int_{\mathbb{R}} C_s f(z) a(z) \, dz \right| \leq \left| \int_{\mathbb{R}} F(y, \eta, t) A(y, \eta, t) \, dy \, d\eta \, dt \right|.
\end{equation}

By duality the bound of the operator (1.1) on $L^p(\mathbb{R})$ follows from bounds on $L^p(\mathbb{R}) \times L^{p'}(\mathbb{R})$ of the bilinear form on the left hand side of the previous display.

The abstract framework of outer measure $L^p$ spaces provides us with the Hölder type bound
\begin{equation}
\left| \int_{\mathbb{R}} F(y, \eta, t) A(y, \eta, t) \, dy \, d\eta \, dt \right| \lesssim \|F\|_{L^p(\mathbb{R})} \|A\|_{L^{p'}(\mathbb{R})}.
\end{equation}

with $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{q} + \frac{1}{q'} = 1$. Appearing on the right are the iterated outer $L^p$ quasi-norms that we elaborate on in Section 2.

The embedding maps defined via equations (1.4) and (1.6), that we call “mass” and “energy” embeddings for historical reasons (compare with [LT00]), satisfy the quasi-norms that we elaborate on in Section 2.

\begin{enumerate}
\item[(1.10)] $\|A\|_{L^{p'}(\mathbb{R})} \lesssim \|a\|_{L^p}$, \\
\item[(1.11)] $\|F\|_{L^p(\mathbb{R})} \lesssim \|f\|_{L^p}$.
\end{enumerate}

**Theorem 1.1** (Mass embedding bounds). For any $p' \in (1, \infty)$, $q' \in (1, \infty)$, and for any function $a \in L^{p'}(\mathbb{R})$ the bounds (1.10) for the embedding (1.7) hold with a constant independent of the Borel measurable function $c: \mathbb{R} \to \mathbb{R}$.

**Theorem 1.2** (Energy embedding bounds). For any $p \in (1, \infty)$, $q \in (\max(2; p'), \infty)$, and for any $f \in L^p(\mathbb{R})$ the bounds (1.11) for the embedding (1.6) hold.

Theorem 1.1 follows as a corollary of Theorem 1.3 below while Theorem 1.2 will be proven in Section 3.

The boundedness of the Carleson Operator on $L^p(\mathbb{R})$ follow as a result of the above discussion. Indeed for any $p, p' \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{p'} = 1$ one can find $q, q' \in (1, \infty)$ such that $\frac{1}{q} + \frac{1}{q'} = 1$ and bounds (1.10) and (1.11) hold.

We remark that iterated outer measure spaces are used to address the case $p \in (1, 2)$. In Section 6 we show that if $p \in (2, \infty)$ a the non-iterated version of outer measure $L^p$ spaces are sufficient to prove $L^p$ boundedness of (1.1).

1.2. The variational Carleson operator. The operator (1.2), introduced and studied in [Obe+12], is bounded on $L^p(\mathbb{R})$ for $r \in (2, \infty)$ and $p \in (r', \infty)$. The above paper also shows that this range is sharp in the sense that that strong $L^p$ bounds do not hold outside this range (see Figure 1).
By duality it is sufficient to prove the bilinear a priori bound
\begin{equation}
\left| \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} a_k(z) \int_{\xi_k(z)} f^k e^{i\xi z} d\xi \right| \leq \left\| f \right\|_{L^p} \left\| a \right\|_{L^{p'}(r')}.
\end{equation}
with a constant independent of the stopping sequence $c$. For the above expression to make sense we require that $f \in S(\mathbb{R})$ while $a \in L^{p'}(l'^\prime)$ i.e. $z \mapsto a(z) = (a_k(z))_{k \in \mathbb{Z}}$ is a function on $\mathbb{R}$ such that for every $z \in \mathbb{R}$ its value is the sequences $a(z) = (a_k(z))_{k \in \mathbb{Z}} \in l'^\prime(\mathbb{Z})$. The function $a$ is Borel measurable in Bochner sense and
\[ \left\| a \right\|_{L^{p'}(l'^\prime)} := \left( \int_{\mathbb{R}} \left\| a(z) \right\|_{l'^\prime}^p dz \right)^{1/p'} < \infty. \]

Analogously to (1.8), the left hand side of (1.12) admits a wave-packet domination
\begin{equation}
\left| \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} a_k(z) \int_{t_k(z)} f^k(x) e^{i\xi z} d\xi \right| \leq \iint_X |F(y, \eta, t)\mathcal{A}(y, \eta, t)| d\eta dt.
\end{equation}
where the embedding map $a \mapsto \mathcal{A}$ is given by
\begin{equation}
\mathcal{A}(y, \eta, t) := \sup_{\Psi} \left| \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} a_k(z) \Psi_{y,\eta,t}^{k}(z) \Psi_{c,\eta,t}^{k}(z) dz \right|.
\end{equation}
The supremum above is taken over all possible choices of left or right truncated wave packets $\Psi_{y,\eta,t}^{c,\eta,t}$. A left truncated wave packet $\Psi_{y,\eta,t}^{c,\eta,t}$ at $(y, \eta, t) \in \mathcal{X}$ is a $S(\mathbb{R})$ function parameterized by $c_- < c_+ \in \mathbb{R} \cup \{-\infty\}$. The parameterization satisfies the properties below. The following three functions of the variable $z$
\begin{equation}
e^{-i\eta(y+tz)}t\Psi_{y,\eta,t}^{c_-}(y+tz)
t^{-1}\partial_{c_-} \left( e^{-i\eta(y+tz)} \Psi_{y,\eta,t}^{c_-}(y+tz) \right)
t^{-1}\partial_{c_+} \left( e^{-i\eta(y+tz)} \Psi_{y,\eta,t}^{c_+}(y+tz) \right)
\end{equation}
are bounded in $S(\mathbb{R})$ uniformly for all $(y, \eta, t) \in \mathcal{X}$ and $c_- < c_+ \in \mathbb{R}$. For some constant $b > 0$ the functions $\Psi_{y,\eta,t}^{c_-}$ satisfy
\begin{equation}
\text{spt } \widehat{\Psi}_{y,\eta,t}^{c_-} \subset B_{t^{-1}b}(\eta).
\end{equation}
For some constants $d, d', d'' > 0$, and $\varepsilon > 0$ it holds that
\begin{equation}
\Psi_{y,\eta,t}^{c_-} \neq 0 \quad \text{only if} \quad \begin{cases} t(\eta - c_-) \in B_{d}(d) \\
(\xi - \eta) > d' > 0
\end{cases}
\end{equation}
\begin{equation}
\Psi_{y,\eta,t}^{c_-} = \Psi_{y,\eta,t}^{c_- + \varepsilon} \quad \text{if} \quad t(c_+ - \eta) > d'' > d' > 0.
\end{equation}
The wave packet $\Psi_{y,\eta,t}^{c_-}$ is right truncated if $\Psi_{y,\eta,t}^{-c_+,c_-}$ is left truncated.

The main result of this paper is the following bounds for the embedding (1.14) that are analogous to the bounds (1.14).

**Theorem 1.3** (Variational mass embedding bounds). For any $r' \in [1, 2)$, $p' \in [1, \infty]$, and $q' \in (r', \infty]$ and any function $a \in L^{p'}(l'^\prime)$ the function $\mathcal{A}$ defined by (1.14) satisfies the bounds
\begin{equation}
\left\| a \right\|_{L^{p'}(l'^\prime)} \lesssim \left\| a \right\|_{L^{p'}(r')} \quad \text{if} \quad p' \in (1, \infty] \quad q' \in (r', \infty].
\end{equation}
furthermore the weak endpoint bounds
\begin{equation}
\left\| a \right\|_{L^{p'}(l'^\prime)} \lesssim \left\| a \right\|_{L^{p'}(r')} \quad \text{if} \quad p' \in (1, \infty], \quad q' \in (r', \infty].
\end{equation}
\[
\|A\|_{L^{1,\infty}\rightarrow L^{1,\infty}(S_m)} \lesssim \|a\|_{L^1(r')} 
\]
hold. All the above inequalities hold with constants independent of the stopping sequence \(c\) appearing in (1.14).

We refer to Section 2 for the description of the outer measure structure on \(\mathcal{X}\) and for the precise definition of the iterated outer measure \(L^p\) norms appearing on the left hand sides.

**Corollary 1.4** (Boundedness of the variational Carleson operator \([\text{Obe+12}]\)). The operator (1.2) defined pointwise for \(f \in S(\mathbb{R})\) extends to a bounded operator on \(L^p(\mathbb{R})\) for \(r \in (2, \infty)\) and \(p \in (r', \infty)\).

Given Theorem 1.3 the above can be obtained analogously as for the operator (1.1). For \(p \) and \(r \) set \(\frac{1}{q} = 1 - \frac{1}{p}\), \(\frac{1}{q'} = 1 - \frac{1}{r}\), and choose \(q\) and \(q'\) so that \(\frac{1}{p} + \frac{1}{q} = 1\) and the bounds (1.11) and (1.19) hold. Using the outer measure Hölder inequality (1.9) with the variational embedded function \(A\) in lieu of \(\Lambda\) and the wave-packet representation (1.13) we obtain the required bound (1.12).

Theorem 1.1 follows from from Theorem 1.3 when \(r = \infty\) by formally setting

\[
a_k(z) = \begin{cases} a(z) & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases} \\
c_k(z) = \begin{cases} -\infty & \text{if } k < 0 \\ c(z) & \text{if } k = 0 \\ +\infty & \text{if } k > 0 \end{cases}
\]

In particular the term \(\tilde{\psi}_{\eta,t}(y-z)\chi(t(\eta-c_-))\) appearing in (1.7) are left truncated wave packets with respect to the parameters \(c_-\) and \(c_+ = +\infty\).

### 1.3. Structure of the paper.

The rest of this paper is organized as follows. In Section 2 we define the outer measure structure on \(\mathcal{X}\). We then recall properties of outer measure \(L^p\) spaces and generalize them to the iterated construction. In addition we illustrate a limiting argument for maps to outer measure \(L^p\) spaces that allows to consider the bounds (1.10), (1.11), and (1.19) as a-priori estimates. We also prove interpolation inequalities that allow us to restrict the proof only to the the weak endpoints of the above bounds. Finally, we formulate the abstract outer Hölder inequality and an outer Radon-Nikodym Lemma that imply inequality (1.19).

In Section 3 we prove the wave-packet domination bound (1.13). In particular it is shown that one can choose both the geometric parameters of the outer measure space (see Section 2) and the parameters of the truncated wave-packets in a compatible way i.e. so that both Theorems 1.2 and 1.3 as well as the conditions (1.10), (1.11) hold. This is done by providing a wave-packet representation for multipliers of the form \(\mathbb{1}_{[c,c]}\) with \(c_- < c_+ \in \mathbb{R} \cup \{+\infty\}\). For any stopping sequence \(c\) this yields an embedded function \(A_c(y, \eta, t)\) so that

\[
\int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} a_k(z) \int_{R_k(z)} \hat{f}(\xi) e^{it\xi} d\xi dz = \iint_{\mathbb{R}} F(y, \eta, t) A_c(y, \eta, t) dy dt dt.
\]

The embedded function \(A_c\) is pointwise dominated by \(\Lambda\) and the map \(a \mapsto A_c\) is shown to be linear. Furthermore the same procedure shows that the inequality in (1.3) is actually an equality i.e.

\[
\int_{\mathbb{R}} C_c f(z) a(z) dz = \iint_{\mathbb{R}^2} F(y, \eta, t) A(y, \eta, t) dy dt dt.
\]

In Section 4 we introduce an auxiliary embedding map for which we show iterated outer measure bounds. The crucial result is given by the covering Lemma 1.3 that allows one to control the measure of super-level sets of this embedding map and by a projection Lemma 1.5 that implies iterated bounds.
In Section 5 we actually prove Theorem 1.3 by showing the auxiliary embedding map of Section 4 dominates the embedding (1.14) in terms of sizes.

Finally, in Section 6 we show that bound (1.11) holds; this follows from an adaptation of the results of [DPO15]. We also remark how in the case $p \in (2, r)$ a non-iterated version of outer measure $L^p$ spaces is enough to obtain $L^p$ bounds for (1.2) and thus for (1.1) with $p \in (2, \infty)$.

1.4. Notation. We quickly recall some useful notation.

We say that $A(x) \lesssim B(x)$ if there exists a constant $C > 0$ such that $A(x) \leq C B(y)$ for all $x, y$ in the domains of $A$ and $B$ respectively. Unless otherwise specified the constant $C > 0$ is absolute. We may emphasize the dependence on a specific parameter $p$ by writing $A(x) \lesssim_p B(y)$. We write $A(x) \approx B(y)$ if $A(x) \lesssim B(y)$ and $A(x) \gtrsim B(y)$.

We denote open and close Euclidean balls of $\mathbb{R}$ as $B_r(x) := (x-r, x+r)$ $B_r := (−r, +r)$ $\overline{B}_r(x) := [x-r, x+r]$ $\overline{B}_r := [−r, +r]$.

We indicate by $\mathbb{1}_\Theta$ the characteristic function of the set $\Theta$ i.e.

$$\mathbb{1}_\Theta(x) := \begin{cases} 1 & \text{if } x \in \Theta \\ 0 & \text{if } x \notin \Theta \end{cases}$$

For an arbitrary large $N > 0$ we introduce the smooth bump function

$$W(z) := (1 + |z|^2)^{-N/2}, \quad W_t(z) := t^{-1} W \left( \frac{z}{t} \right).$$

We define

$$\int_{B_r(x)} f(z) dz := \frac{1}{2r} \int_{B_r(x)} f(z) dz.$$

The operators $M$ and $M_p$ are the Hardy-Littlewood maximal function i.e.

$$Mf(z) := \sup_{t \in \mathbb{R}^+} \int_{B_t(z)} |f(z')| dz'$$

$$M_p f(z) := \sup_{t \in \mathbb{R}^+} \left( \int_{B_t(z)} |f(z')|^p dz' \right)^{1/p}.$$ 

Given a function $\varphi \in S(\mathbb{R})$ we obtain its frequency translates and dilates by setting

$$\varphi_{n,t}(z) := t^{-1} e^{inz} \varphi \left( \frac{z}{t} \right).$$

The stopping sequence $\mathbf{c}$ will denote a Borel measurable function defined on $\mathbb{R}$ with values in increasing sequences in $\mathbb{R} \cup +\infty$ i.e.

$$z \mapsto \mathbf{c}(z) = (c_k(z))_{k \in \mathbb{Z}} \quad -\infty < \cdots \leq c_{k-1}(z) \leq c_k(z) \leq c_{k+1}(z) \leq \cdots \leq +\infty.$$ 

Similarly $\mathbf{a}$ will denote a Borel Bochner-measurable function on $\mathbb{R}$ with values in $L^{r'}$ i.e.

$$z \mapsto \mathbf{a}(z) = (a_k(z))_{k \in \mathbb{Z}} \in L^{r'}.$$ 

We use the notation $L^p(S)$ and $L^q L^p(S)$ to denote (iterated) outer measure $L^p$ spaces. The (outer-) measure of the space is omitted from the notation. We distinguish the above from $L^p$ that are classical Lebesgue spaces. In the case of $L^p$ spaces on $\mathbb{R}$ the measure is the Lebesgue measure; when necessary we may emphasize the measure $\mathcal{L}$ on the space by writing $L^p(\mathcal{L}).$
2. Outer measures on the time-frequency space

We begin the description of the outer measure on the time-frequency space $\mathcal{X}$ by introducing a family of distinguished generating sets. The tent $T(x, \xi, s) \subset \mathcal{X}$ indexed by the top point $(x, \xi, s) \in \mathcal{X}$ is the set

\begin{equation}
T(x, \xi, s) := T^{(i)}(x, \xi, s) \cup T^{(e)}(x, \xi, s)
\end{equation}

\begin{align}
T^{(i)}(x, \xi, s) &:= \{(y, \eta, t) : |y - x| < s, t(\eta - \xi) \in \Theta^{(i)}, t < s\} \\
T^{(e)}(x, \xi, s) &:= \{(y, \eta, t) : |y - x| < s, t(\eta - \xi) \in \Theta^{(e)}, t < s\}
\end{align}

where

\begin{equation}
\Theta = (\alpha^-, \alpha^+) \\
\Theta^{(i)} = (\beta^-, \beta^+), \quad \Theta^{(e)} = \Theta \setminus \Theta^{(i)}
\end{equation}

are geometric intervals such that $0 \in \Theta^{(i)} \subseteq \Theta$ i.e. $\alpha^- \leq \beta^- < 0 < \alpha^+ \leq \beta^+$. We refer to $T^{(i)}$ and $T^{(e)}$ as the interior and exterior parts of the tent $T$. To define the iterated outer measure structure we introduce strips $D(x, s) \subset \mathcal{X}$ as

\begin{equation}
D(x, s) := \{(y, \eta, t) : |y - x| < s, t < s\}
\end{equation}

We now define the outer measures $\mu$ and $\nu$ by introducing the pre-measures $\overline{\mu}$, and $\overline{\nu}$ on the generating sets

\begin{equation}
\overline{\mu}(T(x, \xi, s)) := s \\
\overline{\nu}(D(x, s)) := s
\end{equation}

The specific values of the geometric intervals $\Theta, \Theta^{(i)}$, and $\Theta^{(e)}$ in (2.1) are often inessential. However, the freedom of choosing appropriate parameters was shown to be important in [DT15]. Theorem 1.3 holds as long as

\begin{equation}
B_b \subset \Theta^{(i)} \subset \overline{B}_d, \quad B_{d'}(d) \cup B_{d'}(-d) \subset \Theta^{(e)}
\end{equation}

with $b, d, d'$, and $\varepsilon$ appearing in (1.16), (1.17), and (1.18). As a matter of fact, if one were to consider only left truncated wave-packets in (1.14) then Theorem 1.3 would hold as long as

\begin{equation}
B_b \subset \Theta^{(e)} - d' < \beta^- \quad B_{d'}(d) \subset \Theta^{(e)} \cap \mathbb{R}^+ = [\beta^+, \alpha^+].
\end{equation}

Theorem 1.1 holds as long as satisfies

\begin{equation}
B_b \subset \Theta^{(i)} \quad \text{spt } \chi \subset \Theta^{(e)} \cap \mathbb{R}^+ = [\beta^+, \alpha^+].
\end{equation}

Theorem 1.2 holds as long as $B_b \subset \Theta^{(i)}$. From now on we will allow all our implicit constants to depend on $\Theta$ and $\Theta^{(i)}$.
The outer measure of an arbitrary subset \( E \subset \mathcal{X} \) are obtained via a covering procedure using countable unions of generating sets i.e.

\[
\nu(E) := \inf \left\{ \sum_{n \in \mathbb{N}} \nu(D_n) : E \subset \bigcup_{D_n \in \mathcal{D}} D_n \right\}
\]

and similarly for \( \mu \) using \( \pi \) and the family \( \mathcal{I} \). We say that \( \nu \) and \( \mu \) are generated by the pre-measures \( (\nu, \mathcal{D}) \) and \( (\pi, \mathcal{T}) \) respectively. We call an outer measure space a pair \( (\mathcal{X}, \mu) \) of a separable complete measure space \( \mathcal{X} \) and an outer measure \( \mu : 2^\mathcal{X} \to \mathbb{R}^+ \cup \{+\infty\} \). We will henceforth suppose that \( \mu \) is generated by pre-measures \( (\pi, \mathcal{T}) \) where \( \mathcal{T} \) is a collection of subsets \( T \subset \mathcal{X} \) that we assume to be Borel measurable.

The final ingredient we need for introducing outer measure \( L^p \) spaces is a notion of how large a function on \( \mathcal{X} \) is. We call a size any quasi-norm \( \| \cdot \|_S \) on Borel functions on \( \mathcal{X} \) i.e. a positive functional that satisfies the following properties.

**Monotonicity:** for any Borel function \( G_1 \) and \( G_2 \)

\[
|G_1| \leq |G_2| \implies \|G_1\|_S \leq \|G_2\|_S.
\]

**Positive homogeneity:** for all Borel functions \( G \)

\[
\|\lambda G\|_S = |\lambda|\|G\|_S \quad \forall \lambda \in \mathbb{C}.
\]

**Quasi-triangle inequality:** for any sequence of Borel functions \( G_k \) and for some quasi-triangle constant \( c_s \geq 1 \)

\[
\sum_{k=0}^\infty \|G_k\|_S \leq \sum_{k=0}^\infty c_s^{k+1}\|G_k\|_S
\]

We define the \( S, \mu \)- super-level outer measure as

\[
\mu (\|G\|_S > \lambda) := \inf \left\{ \mu(E_\lambda) : \|G 1_{\mathcal{X}\setminus E_\lambda}\|_S \leq \lambda \right\}
\]

where the lower bound is taken over Borel subset \( E_\lambda \) of \( \mathcal{X} \). The outer-\( L^p \) quasi-norms for \( p \in (0, \infty] \) are give by

\[
\|G\|_{L^p(S)} := \int_{\lambda \in \mathbb{R}^+} p\lambda^p \mu (\|G\|_S > \lambda) \frac{d\lambda}{\lambda},
\]

weak outer \( L^p \) quasi-norms are similarly given by

\[
\|G\|_{L^p, \infty(S)} := \sup_{\lambda \in \mathbb{R}^+} p\lambda^p \mu (\|G\|_S > \lambda).
\]

The outer \( L^p \) spaces are subspaces of Borel functions on \( \mathcal{X} \) for which the above norms are finite. The expressions defining outer \( L^p \) quasi-norms are based on the super-level set representation of the Lebesgue integral, however the expression \( \mu (\|G\|_S > \lambda) \) that appears in lieu of the classical \( \mu (\{x : |g(x)| > \lambda\}) \) cannot always be interpreted as a measure of a specific set. Generally speaking, \( L^p \) spaces for \( p \in (0, \infty) \) are interpolation spaces between the size quasi-norm and the outer measure of the support of a function.

Using a slight abuse of notation we say that a size \( \| \cdot \|_S \) is generated by \( \| \cdot \|_{S(T), \mathcal{T}} \) where \( \| \cdot \|_{S(T)} \) are sizes indexed by generating sets \( T \in \mathcal{T} \) and in particular

\[
\|G\|_S := \sup_{T \in \mathcal{T}} \|G\|_{S(T)}.
\]

The construction of iterated outer \( L^p \) spaces is based on using localized versions of outer \( L^q \) quasi-norms as sizes themselves. Notice that outer \( L^q \) norms are quasi-norms since they too satisfy the quasi-triangle inequality. Given a size \( S \) and a
generating pre-measure \((\mathcal{N}, \mathcal{D})\), outer \(L^q(S)\) sizes are generated by \((L^q(S)(D), \mathcal{D})\) where
\[
\|G\|_{L^q(S)(D)} := \frac{\|F 1_D\|_{L^q(S)}}{\nu(D)^{1/q}}
\]
so \(\|G\|_{L^q(S)} := \sup_{D\in\mathcal{D}} \|G\|_{L^q(S)(D)}\). Consequently we construct iterated outer \(L^p\) spaces as
\[
\|G\|_{L^p_{\mathcal{N}}(S)} := \int_{\tau\in\mathbb{R}^+} p \tau^p \nu \left( \|G\|_{L^q(S)} > \tau \right) \frac{d\tau}{\tau}.
\]
To deal with embedded functions \(F\) and \(A\) from \([1.6]\) and \([1.14]\) we introduce the respective sizes \(\|\cdot\|_S\) and \(\|\cdot\|_{S_m}\) that are generated by \((S_c(T), \tau)\) and \((S_m(T), \tau)\) respectively. The two families of “local” sizes \(S_c(T)\) and \(S_m(T)\) are given by
\[
\|F\|_{S_c(T)} := \frac{\|F 1_{T(\cdot)}\|_2}{\mu(T)^{1/2}} + \|F\|_{L^\infty(T)}
\]
\[
\|A\|_{S_m(T)} := \frac{\|A 1_{T(\cdot)}\|_2}{\mu(T)^{1/2}} + \|A\|_{L^1(T)}
\]
\[
\|F\|_{S_m(T)} := \frac{\|F 1_{\Theta}\|_2}{\mu(T)^{1/2}} + \|F\|_{L^\infty(T)}
\]
\[
\|A\|_{S_m(T)} := \frac{\|A 1_{\Theta}\|_2}{\mu(T)} + \|A\|_{L^1(T)}.
\]
Here \(L^2, L^\infty,\) and \(L^1\) refer to classical Lebesgue \(L^p\) norms on \(X\) with respect to the Borel measure \(d\eta d\mu dt\). The local sizes \(\|\cdot\|_{S_c(T)}\) coincide with the ones introduced on the upper 3-space in \([DT15]\) while \(\|\cdot\|_{S_m(T)}\) are dual to the former in an appropriate sense.

We conclude the construction of outer measure \(L^p\) spaces with a useful remark about the specific geometric properties of coverings with tents \(\mathbb{T}\). For any tent \(T(x, \xi, s)\) we define its \(R\)-enlargement with \(R > 1\) as
\[
RT(x, \xi, s) := \bigcup_{|\xi' - \xi| < Rs^{-1}} T(x, \xi', Rs).
\]
Notice that \(\mu(RT) \lesssim R^3 \mu(T)\) with a constant that depends on the geometric intervals \([2.2]\) but not on \(R\). As a matter of fact the set \(RT\) can be covered by a finite collection of tents \(T(x, \xi_i, Rs)\) by choosing \(\xi_i\) such that
\[
\bigcup_i \{\eta : Rs(\eta - \xi_i) \in \Theta\} \supset \bigcup_{|\xi' - \xi| < Rs^{-1}} \{\eta : Rs(\eta - \xi') \in \Theta\}
\]
The number of points \(\xi_i\) needed to do this is bounded up to a constant factor by \(R^2\) and thus \(\mu(RT) \lesssim R^3 \mu(T)\).

2.1. Properties of outer measure \(L^p\) spaces. We recall some important properties of outer measure \(L^p\) spaces and elaborate on how they carry over to iterated outer-measure spaces. Generally \(X\) may be any locally compact complete metric space; in our case \(X = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+\) with
\[
\text{dist} \left( (y, \eta, t); (y', \eta', t') \right) = t^{-1} |y - y'| + |t\eta - \eta'| + \left| \log \frac{t}{t'} \right|.
\]
2.1.1. Dominated convergence. While outer measure \(L^p\) spaces fall into the class of quasi-Banach spaces, we record only some functional properties that are useful for our applications.

Recall that the quasi-triangle inequality for sizes \([2.11]\) holds for both finite and infinite sums. Given an outer measure space \((X, \mu)\) and a size \(\|\cdot\|_S\), the outer
measure $L^p$ quasi-norms also satisfy the quasi-triangle inequality:

\begin{equation}
(2.19) \quad \left\| \sum_{k=0}^{\infty} G_k \right\|_{L^p(S)} \lesssim c_s \left( \sum_{k=0}^{\infty} c_s^{k+1} \right) \| G_k \|_{L^p(S)}
\end{equation}

for any $c_s' > c_s$ where $c_s$ is the quasi-triangle constant of the size $S$. As a matter
of fact, for any $\lambda > 0$ and for every $k$ choose $E_{\lambda,k}$ such that

\[ \| G_k I_{C \setminus E_{\lambda,k}} \|_S \leq \lambda \quad \text{and} \quad \| G_k \|_{L^p(S)} \lesssim \int_{R^+} p \lambda^p \mu(E_{\lambda,k}) \frac{d\lambda}{\lambda} \]

and set $E_{\lambda} = \bigcup_{k=0}^{\infty} E_{\lambda c_s' - 1, k}$ so that using the quasi-triangle inequality for $\| \cdot \|_S$

one has

\[ \mu(E_{\lambda}) \leq \sum_{k=0}^{\infty} \mu \left( E_{\lambda c_s' - 1, k} \right) \quad \text{and} \quad \| \sum_{k=0}^{\infty} G_k I_{E_{\lambda}} \|_S \leq \lambda \frac{c_s}{(c_s' - c_s)} \]

Thus

\[ \left\| \sum_{k=0}^{\infty} G_k \right\|^p_{L^p(S)} \leq p \left( \frac{c_s}{c_s' - c_s} \right)^p \int_{R^+} \lambda^p \mu(E_{\lambda}) \frac{d\lambda}{\lambda} \]

\[ \leq \left( \frac{c_s}{c_s' - c_s} \right)^p \sum_{k=0}^{\infty} c_s^{p(k+1)} \| G_k \|^p_{L^p(S)} \]

If $p \geq 1$ then this concludes the proof. Otherwise for any $\varepsilon > 0$ one has

\[ \left\| \sum_{k=0}^{\infty} G_k \right\|^p_{L^p(S)} \lesssim \varepsilon, p, c_s \sum_{k=0}^{\infty} (1 + \varepsilon)^{k+1} c_s^{p(k+1)} \| G_k \|^p_{L^p(S)} \]

but since $c_s' > c_s$ was arbitrary this also allows us to conclude.

This fact is crucial to be able to use localized outer $L^p$ quasi-norms as sizes themselves. Furthermore we deduce the following domination property.

**Corollary 2.1.** Suppose that $G$ is a Borel function on $\mathbb{K}$ and $|G| \leq \limsup_{n \to \infty} |G_n|$ pointwise on $\mathbb{K}$ for some sequence of Borel functions $G_n$ that satisfy

\[ \| G_{n+1} - G_n \|_{L^p(S)} \leq C c_s^{-n} \| G_0 \|_{L^p(S)} \quad \text{for some} \quad c_s' > c_s. \]

Then

\[ \| G \|_{L^p(S)} \lesssim_{C, p, c_s, c_s} \| G_0 \|_{L^p(S)}. \]

This follows from (2.19) and from the monotonicity properties of sizes and thus of outer $L^p$ quasi-norms.

Using this property we will restrict ourselves to proving bounds (1.10), (1.11), and (1.19) for a dense class of functions. In particular we will always consider the functions in play to be smooth and rapidly decaying. For example, given a function $a \in L^p(l^r)$ one may always choose a sequence of approximating functions $a^{(n)} \in C^\infty_c(l^r)$ such that

\[ \| a^{(n)} \|_{L^p(l^r)} \lesssim \| a \|_{L^p(l^r)} \]

\[ \| a^{(n+1)} - a^{(n)} \|_{L^p(l^r)} \lesssim 2^{-N} \| a \|_{L^p(l^r)} \]

for an arbitrary $N > 1$. Considering the sequence embedded functions $\Delta_n$ associated to $a^{(n)}$ via (1.14), the pointwise relation $\mathbf{\Delta} = \lim_n \Delta_n$ clearly holds. Corollary 2.1 applied to $\Delta_n$ allows us to conclude that if bounds of Theorem 1.3 hold for the functions $a^{(n)}$ they also hold for $a$. Thus we can restrict to proving the bounds as a priori estimates i.e. we can restrict to showing that they hold for a dense class of functions $a$. The same can be done for the energy embedding bounds of Theorem 1.2.
2.1.2. Hölder and Radon-Nikodym inequalities. We now illustrate the abstract outer measure results from which inequality (1.19) follows. The first two statements relate to general outer measure spaces and are similar to what was obtained in [DT15].

**Lemma 2.2** (Radon-Nikodym domination). Consider \((\mathcal{X}, \mu)\) an outer measure space with \(\mu\), generated by \((\mathcal{F}, \mathcal{T})\) as in (2.3), endowed with a size \(|| \cdot ||_S\) generated by \(|| \cdot ||_{S(T)}\), \(\mathcal{T}\). Suppose that the generating family \(\mathcal{T}\) consists of Borel sets and satisfies the covering condition i.e. \(\mathcal{X} = \bigcup_{i \in \mathbb{N}} T_i\) for some countable sub-collection \(T_i \in \mathcal{T}\).

If \(L\) is a positive Borel measure on \(\mathcal{X}\) such that

\[
(2.20) \quad \int_T |G(P)|d\mathcal{L}(P) \leq C \|G\|_{S(T)} \mathcal{P}(T) \quad \forall T \in \mathcal{T}
\]

and for any Borel function \(G\) and

\[
(2.21) \quad \mu(E) = 0 \implies L(E) = 0 \quad \forall E \subset \mathcal{X} \text{ Borel}
\]

then for any Borel function \(G\) the bound

\[
(2.22) \quad \left| \int_\mathcal{X} G(P)d\mathcal{L}(P) \right| \lesssim \|G\|_{L^1(S)}
\]

holds.

The proof of this Lemma is similar to the one in [DT15].

**Proof.** Suppose \(\|G\|_{L^1(S)} < \infty\), otherwise there is nothing to prove. For each \(k \in \mathbb{Z}\) let \(E_{2k}\) be a Borel set such that

\[
\|G\|_{L^1(S)} \lesssim \sum_{k=-\infty}^{+\infty} 2^k \mu(E_{2k}).
\]

Set

\[
E_{2k} := \bigcup_{l=k}^{+\infty} E_{2l}^\prime \quad \Delta E_{2k} := E_{2k-1} \setminus E_{2k} \quad E_0 = \bigcup_{k=-\infty}^{+\infty} E_{2k} \quad E_{\infty} = \bigcap_{k=-\infty}^{+\infty} E_{2k}.
\]

We have

\[
\left| \int_\mathcal{X} G(P)d\mathcal{L}(P) \right| \leq \int_{E_0} |G(P)|d\mathcal{L}(P) + \sum_{k=-\infty}^{+\infty} \int_{\Delta E_{2k}} |G(P)|d\mathcal{L}(P) + \int_{E_{\infty}} |G(P)|d\mathcal{L}(P),
\]

where

\[
\|G\|_{L^1(S)} \lesssim \sum_{k=-\infty}^{+\infty} 2^k \mu(\Delta E_{2k}).
\]

For every \(k\) there exists a countable covering \(\bigcup_{l \in \mathbb{N}} T_{k,l} \supset \Delta E_{2k}\) such that

\[
\sum_{l \in \mathbb{N}} \mathcal{P}(T_{k,l}) \leq 2\mu(\Delta E_{2k}).
\]

For each \(k \in \mathbb{Z}\) apply (2.20) to obtain

\[
\int_{\Delta E_{2k}} |G(P)|d\mathcal{L}(P) \leq \sum_{l \in \mathbb{N}} \int_{T_{k,l}} |G(P)|\mathbf{1}_{\Delta E_{2k}}(P)d\mathcal{L}(P)
\]

\[
\leq \|G\|_{L^1(S)} \|S\| \sum_{l \in \mathbb{N}} \mathcal{P}(T_{k,l}) \leq 2^{k+1} \mu(\Delta E_{2k})
\]

Thus

\[
\sum_{k=-\infty}^{+\infty} \int_{\Delta E_{2k}} |G(P)|d\mathcal{L}(P) \lesssim \|G\|_{L^1(\mathcal{X}, \mu, S)}.
\]
The term $\int_{\mathcal{K}\setminus E_0} |G(P)|d\mathcal{L}(P)$ vanishes because we may represent $\mathcal{K} = \bigcup_{i\in\mathbb{N}} T_i$. Using (2.20) and the monotonicity of sizes we have
\[
\int_{\mathcal{K}\setminus E_0} |G(P)|d\mathcal{L}(P) \leq \sum_{i\in\mathbb{N}} \int_{T_i} |G(P)|1_{\mathcal{K}\setminus E_0}(P)d\mathcal{L}(P) \\
\leq \sum_{i\in\mathbb{N}} \|G1_{\mathcal{K}\setminus E_0}\|_{S(T_i)} \mathcal{P}(T_i) = 0.
\]

The term $\int_{E^+\infty} |G(P)|d\mathcal{L}(P)$ also vanishes since
\[
\mu(E_{2k}) \leq \sum_{i=k}^{\infty} \mu(E_{2i}) \lesssim 2^{-k}\|G\|_{L^1(S)}
\]
and thus $\mu(E_{+\infty}) = 0$ and $\mathcal{L}(E_{+\infty}) = 0$ by (2.21). This concludes the proof. \qed

The proof of the following outer measure Hölder inequality can be found in [DT15].

**Proposition 2.3** (Outer Hölder inequality). Let $(\mathcal{K}, \mu)$ be an outer measure space endowed with three sizes $\|\cdot\|_S$, $\|\cdot\|_S'$, and $\|\cdot\|_{S''}$ such that for any Borel functions $F$ and $A$ on $\mathcal{K}$ the product estimate for sizes
\[
(2.23) \quad \|FA\|_S \lesssim \|F\|_{S'}|A|_{S''}
\]
holds. Then for any Borel functions $F$ and $A$ on $\mathcal{K}$ the following outer Hölder inequality holds:
\[
(2.24) \quad \|FA\|_{L^p(S)} \leq 2\|F\|_{L^{p'}(S')}|A|_{L^{p''}(S'')}
\]
for any triple $p, p', p'' \in (0, \infty]$ of exponents such that $\frac{1}{p} + \frac{1}{p'} = \frac{1}{p} = \frac{1}{q}$.

The above two statement can be easily extended to iterated outer measure spaces. Suppose from now on that $\mathcal{K}$ is endowed with two outer measures $\nu$ and $\mu$, the former generated by a pre-measure $(\mathcal{V}, \mathcal{D})$ as described in (2.8). Given a size $\|\cdot\|_S$ we introduce local $L^q(S)$ sizes as described by (2.14) and the corresponding iterated outer $L^pL^q(S)$ quasi-norms as described in (2.15).

**Corollary 2.4** (Outer Hölder inequality for iterated outer measure spaces). Let $(\mathcal{K}, \mu)$ be an outer measure space endowed with three sizes $\|\cdot\|_S$, $\|\cdot\|_S'$, and $\|\cdot\|_{S''}$ satisfying the assumptions of Proposition 2.3. Then given any two triples of exponents $p, p', p'' \in (0, \infty)$ and $q, q', q'' \in (0, \infty]$ such that $\frac{1}{p} + \frac{1}{p'} = \frac{1}{p} = \frac{1}{q}$ and $\frac{1}{q} + \frac{1}{q'} = \frac{1}{q}$ the iterated Hölder bounds
\[
(2.25) \quad \|FA\|_{L^pL^q(S)} \lesssim \|F\|_{L^{p'}L^{q'}(S')}|A|_{L^{p''}L^{q''}(S'')}
\]
hold for any Borel functions $F$ and $A$ on $\mathcal{K}$.

As a matter of fact the inequality
\[
(2.26) \quad \|FA\|_{L^pL^q(S)} \lesssim \|F\|_{L^{p'}L^{q'}(S')}|A|_{L^{p''}L^{q''}(S'')}
\]
holds for localized $L^q(S)$ sizes satisfy the inequality by Proposition 2.3 applied to the defining expression (2.14). Thus the local $L^q(S)$ sizes themselves satisfy the conditions of Hölder inequality and the statement of the above Corollary follows.

The Radon-Nikodym Lemma 2.2 can also be generalized to iterated outer measure $L^p$ spaces.

**Corollary 2.5** (Iterated Radon-Nikodym domination). Consider $(\mathcal{K}, \mu)$ an outer measure space with a size $\|\cdot\|_S$ and a Borel measure $\mathcal{L}$ that satisfy the conditions of Lemma 2.2 and let $\nu$ be a measure generated by $(\mathcal{V}, \mathcal{D})$. Suppose that $\mathcal{D}$ also
satisfies the covering condition of Lemma 2.2. Then the iterated Radon-Nikodym domination
\[ \left| \int_X G(P) d\mathcal{L}(P) \right| \lesssim \|G(P)\|_{L^1(S)} \]
holds.

As a matter of fact for any Borel function \( G \) the inequality
\[ \int_D |G(P)| \mathcal{L}(P) \lesssim \|G\|_{L^1(S)(D)\mathcal{T}(D)} \]
follows from (2.14) and Lemma 2.2. Thus the outer measure space \((X, \mu)\) and the family of local sizes \( \| \cdot \|_{L^1(S)} \) satisfy the conditions of Lemma 2.2 and the statement of the Corollary follows.

Using the above properties one can deduce inequality (1.9): introduce the size
\[ \|G\|_{S^1} := \sup_{T \in \mathbb{T}} \|G\|_{S^1(T)} = \sup_{T \in \mathbb{T}} \frac{\|G\|_{L^1}}{\mu(T)} \]
so that the sizes \( \| \cdot \|_{S^1}, \| \cdot \|_{S_{\nu}}, \text{ and } \| \cdot \|_{S_{\mu}} \) satisfy the product estimate (2.23). It follows from the iterated Hölder inequality (2.24) that
\[ \|FA\|_{L^p(S^1)} \lesssim \|F\|_{L^p(S_{\nu}(S))} \|A\|_{L^p(S_{\mu}(S))} \]
for conjugate exponents \( \frac{1}{p} + \frac{1}{q} = 1 \) and \( \frac{1}{q} + \frac{1}{\nu} = 1 \). Furthermore we may apply 2.5 to \((X, \nu)\) with the local size \( \| \cdot \|_{L^1(S^1)} \) so (1.9) follows.

2.1.3. Interpolation. Here we recall some interpolation properties of outer measure \( L^p \) spaces from [DT15] and extend them to iterated outer measure \( L^p \) spaces.

The proof of the following Propositions can be found in [DT15].

**Proposition 2.6** (Logarithmic convexity of \( L^p \) norms). Let \((X, \mu)\) be an outer measure space with size \( \| \cdot \|_S \) and let \( G \) be a Borel function on \( X \). For every \( \theta \in (0, 1) \) and for \( \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \) with \( p_0, p_1 \in (0, \infty], p_0 \neq p_1 \) the inequality
\[ \|G\|_{L^p(S)} \leq C_{\theta, p_0, p_1} \|G\|_{L^{p_0, \infty}(S)}^{1-\theta} \|G\|_{L^{p_1, \infty}(S)}^{\theta} \]
holds.

The following straightforward remarks are useful to be able to compare outer measure spaces with differing sizes.

**Remark 2.7** (Monotonicity of outer \( L^p \) spaces). Consider an outer measure space \((X, \mu)\) with two sizes \( \| \cdot \|_S \) and \( \| \cdot \|_{S'} \). Suppose that given two Borel functions \( G \) and \( G' \) on \( X \) we have that
\[ \|G\|_{X\times E} \lesssim \|G\|_{X\setminus E} \]
for any \( E = \bigcup_{n \in \mathbb{N}} T_n \) that is countable union of generating sets \( T_n \in \mathbb{T} \). Then
\[ \|G\|_{L^p(S)} \lesssim \|G\|_{L^p(S')} \]
for all \( p \in (0, \infty) \) and for iterated spaces
\[ \|G\|_{L^p(S)} \lesssim \|G\|_{L^p(S')} \]
for all \( p, q \in (0, \infty) \). Similar statements hold for weak spaces.

**Remark 2.8** (Interpolation of sizes). Given an outer measure space \((X, \mu)\) with two sizes \( \| \cdot \|_S \) and \( \| \cdot \|_{S'} \), define the sum size as \( \| \cdot \|_{S + S'} := \| \cdot \|_S + \| \cdot \|_{S'} \). Then the following inequality holds for any Borel function \( G \) and for any \( p \in (0, \infty] \)
\[ \|G\|_{L^p(S + S')} \leq 2 \left( \|G\|_{L^p(S)} + \|G\|_{L^p(S')} \right). \]
The proofs of the above remarks consists of simply applying the definition of outer measure $L^p$ quasi-norms and as such are left to the reader.

As a consequence of the above properties, given a function $G$ the following inequality holds:

$$\|G\|_{L^{p,q}(S)} \leq C_{q,q_0,q_1} \left(\|G\|_{L^{p,q_0,\infty}(S)} + \|G\|_{L^{p,q_1,\infty}(S)}\right)$$

for all $q_0, q_1 \in (1, \infty]$ and $q \in (q_0, q_1)$.

Finally we state a version of the Marcinkiewicz interpolation for maps into outer measure $L^p$ spaces

**Proposition 2.9** (Marcinkiewicz interpolation). Let $(Y, \mathcal{L})$ be a classical measure space, $(X, \mu)$ be an outer measure space with size $\|\cdot\|_S$ and assume $1 \leq p_0 < p_1 \leq \infty$. Let $T$ an operator that maps $L^{p_0}(Y, \mathcal{L}) + L^{p_1}(Y, \mathcal{L})$ to Borel function on $X$ so that

- **Scaling:** $|T(\lambda f)| = |\lambda T(f)|$ for all $f \in L^{p_0}(Y, \mathcal{L}) + L^{p_1}(Y, \mathcal{L})$ and $\lambda \in \mathbb{R}$;
- **Quasi sub-additivity:** $|T(f + g)| \leq C(|T(f)| + |T(g)|)$ for all $f, g \in L^{p_0}(Y, \mathcal{L}) + L^{p_1}(Y, \mathcal{L})$;
- **Boundedness:**
  - $\|T(f)\|_{L^{p_0,\infty}(S)} \leq C_1 \|f\|_{L^{p_0}(Y, \mathcal{L})}$ \quad $\forall f \in L^{p_0}(Y, \mathcal{L})$
  - $\|T(g)\|_{L^{p_1,\infty}(S)} \leq C_2 \|g\|_{L^{p_1}(Y, \mathcal{L})}$ \quad $\forall g \in L^{p_1}(Y, \mathcal{L})$.

Then for all $f \in L^{p_0}(Y, \mathcal{L}) \cap L^{p_1}(Y, \mathcal{L})$ we have

$$\|T(f)\|_{L^{p_0}(S)} \leq \theta \|p_0, p_1, C_1^{1-\theta}C_2^{\theta}\|f\|_{L^{p_0}(Y, \mathcal{L})}$$

with $\theta \in [0, 1]$ and $\frac{1}{p_0} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

3. **Wave-packet decomposition**

The main object of this section is to show inequality (1.13) i.e. the domination of the linearized variational Carleson operator via embedding maps. The following procedure follows the general scheme for obtaining (1.8) and (1.9).

**Lemma 3.1.** Consider any fixed parameters $d > b > 0$, $0 < d' < d - 2b$, $d'' > d + b$, and a small enough $\epsilon > 0$ appearing in properties (1.10), (1.11), and (1.12). There exists a choice of truncated left and right wave packets $\Psi_{0,0,t}^{c_-,c_+}$ and $\Psi_{0,0,t}^{c_-,c_+}$ such that for all $c_- < c_+ \in \mathbb{R} \cup \{\infty\}$ the expansion

$$\mathbb{1}_{(c_-,c_+)}(\xi) = \int_{\mathbb{R} \times \mathbb{R}^+} \left(\widehat{\Psi}_{0,0,t}^{c_-,c_+}(\xi) + \widehat{\Psi}_{0,0,t}^{c_-,c_+}(\xi)\right) dt$$

holds where the integral converges in locally uniformly for $\xi$ in $(c_-, c_+)$.  

**Proof.** Let $\varphi \in S(\mathbb{R})$ and $\chi \in C^\infty_c(\mathbb{R})$ be two non-negative functions such that for $\epsilon > 0$ small enough, to be determined later the following holds

$$\text{spt } \varphi \subset B_b \quad \text{spt } \chi \subset B_\epsilon(d) \subset (b, +\infty) \quad \int_{\mathbb{R} \times \mathbb{R}^+} \widehat{\varphi}(\tilde{t} - \tilde{\eta}) \chi(\tilde{\eta}) d\tilde{\eta} d\tilde{t}$$

A change of variable $\tilde{\eta} = t\eta$ and $\tilde{t} = \frac{t}{|\eta|}$, gives:

$$\mathbb{1}_{(0, +\infty)}(\xi) = \int_{\mathbb{R} \times \mathbb{R}^+} \varphi_{\eta,t}(\xi) \chi(t\eta) dt d\eta$$

with $\varphi_{\eta,t}(z) := e^{i\eta z} t^{-1} \varphi\left(\frac{z}{t}\right)$. Let $\gamma \in C^\infty_c([0,1 + \epsilon])$ so that

$$\gamma(t) = 1 \text{ for } t \in [0, (1 + \epsilon)^{-1}] \quad \gamma(t) + \theta(1/t) = 1 \text{ for } t \in \mathbb{R}^+.$$
Such a function can be constructed by taking \( \tilde{\gamma} \) to satisfy the first two conditions and by setting \( \gamma(t) := \frac{\tilde{\gamma}(t)}{\gamma(t) + \tilde{\gamma}(1/T)} \). Let us then set

\[
\beta(\xi) := \int_{\mathbb{R} \times \mathbb{R}^+} \gamma(t') \hat{\varphi}_{\eta', t'}(\xi) \chi(t' \eta') \, dt'
\]

so that

\[
\beta(t \xi) = \int_{\mathbb{R} \times \mathbb{R}^+} \gamma(t' / t) \hat{\varphi}_{\eta', t'}(\xi) \chi(t' \eta') \, dt'.
\]

Using (3.3) one obtains

\[
\mathbb{1}_{(c_-, c_+)}(\xi) = \int_{\mathbb{R} \times \mathbb{R}^+} \hat{\varphi}_{\eta, t}(\xi) \chi(t(\eta - c_-)) \hat{\varphi}_{\eta', t'}(\xi) \chi(t'(c_+ - \eta')) \, dt' \, d\eta' dt,
\]

so the representation (3.1) holds with

\[
\omega_{\eta, t}(\xi) := \chi(t(\eta - c_-)) \hat{\varphi}_{\eta, t}(\xi) \beta(t(c_+ - \eta))
\]

\[\hat{\omega}_{\eta, t}(\xi) := \chi(t(c_+ - \eta)) \hat{\varphi}_{\eta, t}(\xi) \beta(t(\xi - c_+)).\]

It remains to check that \( \Psi_{c_-, c_+}^\xi \) are left truncated wave packets. By symmetry it will follow that \( \Psi_{c_-, c_+}^\xi \) is a right truncated wave packet. First of all (1.16) holds according to (3.5) since \( \text{spt} \hat{\varphi}_{\eta, t}(\xi) \subset B_{\mathbb{R} \times \mathbb{R}^+}(\eta) \).

Notice that

\[
\beta(\xi) = 1 \text{ on } \{(d + \varepsilon + b)(1 + \varepsilon), +\infty\}.
\]

As a matter of fact the integrand in (3.4) is non-zero only if \( t'(\xi - \eta') \in B_0 \) and \( t' \eta' \in B_2(d) \) so \( t' \xi \in B_{c_+ b}(d) \). This shows that

\[
\xi \leq \frac{d - \varepsilon - b}{1 + \varepsilon} \implies t' > 1 + \varepsilon \text{ or } t' < 0 \implies \gamma(t') = 0 \implies \beta(\xi) = 0
\]

\[
\xi \geq (d + \varepsilon + b)(1 + \varepsilon) \implies t' < (1 + \varepsilon)^{-1} \implies \gamma(t') = 1 \implies \beta(\xi) = 1
\]

where the last equality follows from (3.3).

We now check that (1.17) holds. It follows from (3.5) that \( \Psi_{\eta, t}^{c_-, c_+} \theta(\xi) \) vanishes unless \( \chi(t(\eta - c_-)) \neq 0 \) i.e. unless \( t(\eta - c_-) \in B_0(d) \). Also \( \Psi_{\eta, t}^{c_-, c_+} \theta(\xi) = 0 \) unless \( t(\xi - \eta) > -b \) and \( t(c_+ - \xi) > \frac{d - \varepsilon - b}{1 + \varepsilon} \) i.e. unless \( t(c_+ - \eta) > \frac{d + \varepsilon + b}{1 + \varepsilon} \). As long as \( 0 < d' < d + 2b \) one can choose \( \varepsilon > 0 \) small enough for (1.17) to hold.

We now check that (1.18) holds. We have that \( \beta(t(c_+ - \eta)) = 1 \) if \( t(c_+ - \xi) > (d + \varepsilon + b)(1 + \varepsilon) \) and we know that \( \hat{\varphi}_{\eta, t}(\xi) \neq 0 \) only if \( t(\xi - \eta) \in B_0 \) thus if \( t(c_+ - \eta) > (d + \varepsilon + b)(1 + \varepsilon) + b \) then

\[
\hat{\Psi}_{\eta, t}^{c_-, c_+} = \chi(t(\eta - c_-)) \hat{\varphi}_{\eta, t}(\xi) = \hat{\Psi}_{\eta, t}^{c_-, c_+}.
\]

so (1.18) holds as long as \( d' > d + 2b \) and \( \varepsilon > 0 \) is chosen small enough.

We now need to check the smoothness conditions (1.15). We must show that the functions

\[
\hat{\Psi}_{\eta, t}^{c_-, c_+} \left( \frac{\xi + \eta}{t} \right) \quad t^{-1} \partial_{\xi} \hat{\Psi}_{\eta, t}^{c_-, c_+} \left( \frac{\xi + \eta}{t} \right) \quad t^{-1} \partial_{\eta} \hat{\Psi}_{\eta, t}^{c_-, c_+} \left( \frac{\xi + \eta}{t} \right)
\]

are all uniformly bounded in \( S(\mathbb{R}) \) for all \( \eta, t \in \mathbb{R} \times \mathbb{R}^+ \) and \( c_- < c_+ \in \mathbb{R} \). Clearly

\[
\Psi_{\eta, t}^{c_-, c_+} \left( \frac{\xi + \eta}{t} \right) = \chi(t \eta - tc_-) \hat{\varphi}_{\eta, t}(\xi + \eta) \beta(tc_+ - \xi + \eta)
\]

and the claim follows.
Corollary 3.2. Let us fix a set of parameters $d', d'' > 0$ with $d'' > \max(d'; d)$ and $3d > d'$. Then for any $\varepsilon > 0$ small enough there exists $b > 0$ such that there exists a choice of left and right truncated wave packets $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ and $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ such that (3.1) holds for all $c_+ < c_+ < R \cup \{+\infty\}$.

Proof. If $d'' > d > d' > 0$ then let us choose $\varepsilon > 0$ and $b > 0$ small enough so that the conditions for Lemma 3.1 hold. Then the Lemma provides us with wave packets $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ and $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ such that (3.1) holds as required.

Suppose now that $3d > d' > 2d$ and $d'' > d'$ and consider the set of parameters $\tilde{d}', \tilde{d}', \tilde{d}', \tilde{d}' > 0$ given by

$$
\tilde{d} = \varepsilon \quad \tilde{b} = b - \delta \quad \tilde{d}' = d + \delta \quad \tilde{d}' = d' - \delta 
$$

for some $d > \delta > 0$. We need to check that the above parameters satisfy the assumptions of Lemma 3.1 that will give us the left and right truncated wave-packets $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ and $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ for which (1.10), (1.17), and (1.18) hold with these modified parameters. As long as $2\delta < \tilde{d} - \tilde{d}'$, setting $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t} := \tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ and $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t} := \tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ will provide us with the required wave-packets so that (3.1) holds.

Set $b = \frac{d' - d}{2(1 - 3\varepsilon)}$ and $\delta = (1 - \varepsilon)b$ so that $\tilde{b} = \varepsilon b$ with $\varepsilon > 0$ small enough for the subsequent inequalities to hold. All the abovementioned conditions hold since

$$
\tilde{d} - \varepsilon - \tilde{b} = d + \delta - \varepsilon - b + \delta - 2\delta = d - \varepsilon - b = d - \varepsilon - \frac{d' - d}{2(1 - 3\varepsilon)} > 0
$$

$$
\tilde{b} = \varepsilon b > 0
$$

$$
\tilde{d}' = d' + \delta - \delta = d' + \frac{1 - \varepsilon - d' - d}{2} > 0
$$

$$
\tilde{d}' = 2b - \tilde{d}' = d - d' - 2b + 4\delta = d - d' + 2(1 - 2\varepsilon)b = (d' - d) \left( \frac{1 - 2\varepsilon}{1 - 3\varepsilon} - 1 \right) > 0
$$

$$
\tilde{d}' - \tilde{d}' = 2b = d'' - d - 2b + d'' - d' > d'' - d' + (d' - d) \left( \frac{1 - 1}{1 - 3\varepsilon} \right) > 0.
$$

This concludes the proof. \□

As a consequence we obtain the following representation Lemma.

Lemma 3.3. Let us fix a set of parameters $d', d', d > 0$ with $d'' > \max(d'; d)$ and $3d > d'$. For any $\varepsilon > 0$ small enough there exists $b > 0$ such that for any $f \in S(\mathbb{R})$ and $c_- < c_+ < R \cup \{+\infty\}$ the expansion

$$
(3.7) \quad \int_{c_-}^{c_+} \hat{f}(\xi) e^{i\xi \eta} d\xi = \int_{\mathbb{R}} \int_{\mathbb{R}} f * \psi_{y, \eta, t}(y) \left( \tilde{\Psi}^{c_-, c_+}_{y, \eta} \psi_{y, \eta, t}(z) + \tilde{\Psi}^{c_-, c_+}_{y, \eta} \psi_{y, \eta, t}(z) \right) dy \eta dt
$$

holds. Here $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ and $\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}$ are some left and right truncated wave packets for which properties (1.15), (1.16), (1.17), and (1.18) hold with the parameters above.

The function $\psi_{y, \eta, t}$ is obtained from some $\psi \in S(\mathbb{R})$ as in (1.4); we also have

$$
\text{spt} \tilde{\psi} \in B(1 + \varepsilon)b \quad \text{with}(1 + \varepsilon)b < d - \varepsilon.
$$

Proof. Let us choose $\psi \in S(\mathbb{R})$ such that $\text{spt} \tilde{\psi} \in B(1 + \varepsilon)b$ and $\tilde{\psi} = 1$ on $B_b$ so that

$$
\tilde{\Psi}^{c_-, c_+}_{y, \eta, t}(\xi) = \tilde{\psi}_{y, \eta, t}(\xi) \tilde{\Psi}^{c_-, c_+}_{y, \eta, t}(\xi) \quad \tilde{\Psi}^{c_-, c_+}_{y, \eta, t}(\xi) = \tilde{\psi}_{y, \eta, t}(\xi) \tilde{\Psi}^{c_-, c_+}_{y, \eta, t}(\xi)
$$
and let us set \( \Psi_{y,\eta,t}^{c_{e}+,f}(z) = \Psi_{0,\eta,t}^{c_{e}+,f}(z-y) \) and \( \tilde{\Psi}_{y,\eta,t}^{c_{e}+,f}(z) = \Psi_{0,\eta,t}^{c_{e}+,f}(z-y) \). It follows that
\[
\iint_{\mathbb{R}} f * \psi_{\eta,t}(y) \left( \Psi_{y,\eta,t}^{c_{e}+,f}(z) + \tilde{\Psi}_{y,\eta,t}^{c_{e}+,f}(z) \right) dy d\eta dt
= \int_{\mathbb{R} \times \mathbb{R}^{+}} f * \psi_{\eta,t}(y) \left( \Psi_{0,\eta,t}^{c_{e}+,f} + \tilde{\Psi}_{0,\eta,t}^{c_{e}+,f} \right)(z) dy d\eta dt
= \mathcal{F}^{-1} \left( \iint_{\mathbb{R} \times \mathbb{R}} \hat{f}(\xi) \tilde{\psi}_{\eta,t}(\xi) \left( \tilde{\Psi}_{0,\eta,t}^{c_{e}+,f}(\xi) + \tilde{\Psi}_{0,\eta,t}^{c_{e}+,f}(\xi) \right) \right) = \mathcal{F}^{-1} \left( \hat{f}(\xi) 1_{c_{e}+}(\xi) \right)
\]
as required, where \( \mathcal{F}^{-1} \) is the inverse Fourier transform. \( \square \)

As a corollary of the above Lemma we have the following pointwise wave-packet representation for the linearized variational Carleson operator:
\[
\sum_{k \in \mathbb{Z}} a_k(z) \iint_{c_k(z)} \hat{f}(\xi) e^{\xi z} d\xi
= \sum_{k \in \mathbb{Z}} \iint_{\mathbb{R}} f * \psi_{\eta,t}(y) \left( \Psi_{y,\eta,t}^{c_k(z),c_{k+1}(z),f}(z) + \tilde{\Psi}_{y,\eta,t}^{c_k(z),c_{k+1}(z),f}(z) \right) a_k(z) dy d\eta dt.
\]
Setting
\[
\Psi_{\epsilon,\eta,t}(y,\eta,t) := \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} \left( \Psi_{y,\eta,t}^{c_k(z),c_{k+1}(z),f}(z) + \tilde{\Psi}_{y,\eta,t}^{c_k(z),c_{k+1}(z),f}(z) \right) a_k(z) dz
\]
gives (1.22). We also remark that if \( c \) and \( a \) are as in (1.21) then the above construction reduces to the one described by (1.13), (1.6) and (1.7), thus showing (1.23).

Finally notice that if we fix \( \Theta = (\alpha^{-},\alpha^{+}) = (-1,1) \) and set \( d < d' < 3d \), then for every \( \epsilon > 0 \) small enough we may apply Lemma 3.3 to obtain the parameter \( b > 0 \) and wave-packets \( \Psi_{y,\eta,t}^{c_{e}+,f}(z) \), \( \tilde{\Psi}_{y,\eta,t}^{c_{e}+,f}(z) \), and \( \psi_{\eta,t} \). Supposing that \( \epsilon > 0 \) is small enough so that \( d+\epsilon < \alpha^{+} = 1 \) we can find \( (1+\epsilon)b < \beta^{+} < d - \epsilon \) and set \( \Theta^{(\epsilon)} = (\beta^{-},\beta^{+}) = (-\beta^{+},\beta^{+}) \). Thus there exists a set of parameters \( \alpha^{-} < \beta^{-} < \beta^{+} < \alpha^{+} \) such that (1.13) holds and (2.1) is satisfied so that Theorem 1.3 and Theorem 1.2 hold.

4. THE AUXILIARY EMBEDDING MAP

In this section we introduce an auxiliary embedding map used to control the embedded function \( \Psi \). The bounds with the same exponents as in (1.19) hold for the auxiliary embedded function \( \mathcal{M} \) with \( S^{\infty} \) in lieu of \( S_{\alpha} \). However it is technically easier to control the super-local outer measure \( \mu \left( \|\mathcal{M}\|_{S^{\infty}} > \lambda \right) \) of the auxiliary embedded function \( \mathcal{M} \). A crucial covering Lemma implies non-iterated outer \( L^{p'} \) space bounds for \( \mathcal{M} \) while a locality property and a projection Lemma allows for the extension to iterated outer \( L^{p'} \) \( L^{p'} \) spaces.

The auxiliary embedding map associates to \( a \in C_c^{\infty}(l'^{r'}) \) the function on \( \mathcal{X} \) given by
\[
\mathcal{M}(y,\eta,t) := \int_{\mathbb{R}} \left( \sum_{k \in \mathbb{Z}} |a_k(z)|^{r'} 1_{\Theta}(l(\eta - c_k(z))) \right)^{1/r'} W_{\epsilon}(z-y) dz
\]
where the bump function \( W \) is as in (1.21).

**Proposition 4.1** (Bounds on the auxiliary embedding map \( \mathcal{M} \)). For any \( r' \in [1,\infty] \), \( p' \in (1,\infty] \), and \( q' \in (r',\infty] \) and for any function \( a \in L^{p'}(l'^{r'}) \) the function \( \mathcal{M} \) defined by (1.1) satisfies the bounds
\[
\|\mathcal{M}\|_{L^{p'}(S^{\infty})} \lesssim \|a\|_{L^{p'}(l'^{r'})}
\]
where $S^\infty(\mathcal{M}) := \sup_{(y, \eta, t) \in \mathbb{R}} \mathcal{M}(y, \eta, t)$. Furthermore the weak endpoint bounds

$$
\|\mathcal{M}\|_{L^{p'}(x', r')} \lesssim \|a\|_{L^{p'}(x')} \quad p' \in (1, \infty],
$$

$$
\|\mathcal{M}\|_{L^{q'}(x', r')} \lesssim \|a\|_{L^{q'}(x')} \quad q' \in (r', \infty],
$$

hold. All the above inequalities hold as long as $N > N_0$ holds.

We may make two reductions to prove the above bounds. First of all one can substitute $W_t(z)$ by a normalized characteristic function of a ball. As a matter of fact set

$$
\mathcal{M}_R(y, \eta, t) := \int_{B_R(y)} \left( \sum_{k \in \mathbb{Z}} |a_k(z)|^{r'} \mathbb{1}_{\Theta}(t(\eta - \xi_k(z))) \right)^{1/r'} dz
$$

so that $\mathcal{M}(y, \eta, t) \lesssim \sum_{n \in \mathbb{N}} R^{-N_n} \mathcal{M}_{R^n}(y, \eta, t)$. Thus it is sufficient to prove that the bounds (4.2) hold for $\mathcal{M}_R$ with a constant that grows at most as $R^{N'}$ for some $N' > 0$ as $R \to \infty$. The bounds for $\mathcal{M}$ follow by quasi-subadditivity as remarked in Section 2.1.1 as long as $N > N'$. For the second reduction split $\Theta = \Theta^+ \cup \Theta^-$ into $\Theta^+ := \Theta \cap [0, +\infty]$ and $\Theta^- := \Theta \cap [-\infty, 0]$. Set

$$
\mathcal{M}^+_R(y, \eta, t) := \int_{B_R(y)} \left( \sum_{k \in \mathbb{Z}} |a_k(z)|^{r'} \mathbb{1}_{\Theta^+}(t(\eta - \xi_k(z))) \right)^{1/r'} dz
$$

so that $\mathcal{M}_R \leq \mathcal{M}^+_R + \mathcal{M}^-_R$. Thus it will suffice to prove the sharpness of the bounds (4.2) only for $\mathcal{M}^+_R$.

We begin by introducing the concept of disjoint tents relative to the embedding (1.2) and record an important covering lemma.

**Definition 4.2** ($Q^+$-disjointness). Let $Q > 0$. We say two tents $T(x, \xi, s)$ and $T(x', \xi', s')$ are $Q^+$-disjoint if either

$$
B_Q(x) \cap B_Q(x') = \emptyset \quad \text{or} \quad \{c : s(\xi - c) \in \Theta^+\} \cap \{c : s'(\xi' - c) \in \Theta^+\} = \emptyset.
$$

Notice that if a sequence of tents $T(x_l, \xi_l, s_l)_{l \in \mathbb{N}}$ are pairwise $Q^+$-disjoint, with $Q \geq R$, then for every $z \in \mathbb{R}

$$
\left| \sum_{l \in \mathbb{N}} \mathbb{1}_{\Theta^+}(s_l(\xi_l - \xi_k(z))) \mathbb{1}_{B_R} \left( \frac{x_l - z}{s_l} \right) \right| \leq 1
$$

and the bound

$$
\sum_{l \in \mathbb{N}} s_l \|\mathcal{M}^+_R(x_l, \xi_l, s_l)\|_{L^{r'}(x', r')} \leq \sum_{l \in \mathbb{N}} s_l \int_{B_{R}(x_l)} \sum_{k \in \mathbb{Z}} |a_k(z)|^{r'} \mathbb{1}_{\Theta^+}(t(\xi_l - \xi_k(z))) dz

\leq (2R)^{-1} \int_{\mathbb{R}} \|a(z)\|_{L^{r'}(x', r')}^{r'} dz = (2R)^{-1} \|a\|_{L^{r'}(x', r')}^{r'}
$$

holds.

What follows is a covering lemma. We remark that this is the only instance where we require smoothness and rapid decay assumptions on $a$.

**Lemma 4.3.** Let $a \in C^\infty_c(x', r')$. If $Q > R > R_0$ for some $R_0 > 0$ depending on $\Theta$ the super level set

$$
E_{\lambda,R} := \{(x, \xi, s) : \mathcal{M}^+_R(x, \xi, s) \geq \lambda\}
$$

admits a finite covering $\bigcup_{l=1}^{L} 3Q^2 T_l \supset E_{\lambda,R}$ with tents $Q^+$-disjoint tents $T_l = T(x_l, \xi_l, s_l)$ centered at points $(x_l, \xi_l, s_l) \in E_{\lambda,R}$. 

Proof. Introduce the relation $\prec$ between points of $\mathbb{X}$ such that $(x, \xi, s) \prec (x', \xi', s')$ if $B_{Qs}(x) \cap B_{Qs'}(x') \neq \emptyset$, $s(\xi - \xi') \in \Theta$ and $s' > Qs$. We say $(x, \xi, s)$ is maximal in a set $P \subset \mathbb{X}$ if there is no $(x', \xi', s') \in P$ such that $(x, \xi, s) \prec (x', \xi', s')$. Notice that $E_{\lambda,R}$ is $(x,t)$-bounded in the sense that for some $C > 1$ large enough

$$E_{\lambda,R} \subset B_C(0) \times \mathbb{R} \times (0,C)$$

holds. As a matter of fact $M_t$ below since $Q$ contradicts the selection condition.

Given Proposition 4.4.

Inductively construct a covering starting with an empty collection of tents $T^0 = \emptyset$. At the $l$th step consider the points in the set

$$\bigcup_{T \in T^{l-1}} 3Q^2T$$

and select from it a point $(x_l, \xi_l, s_l)$ that is maximal with respect to the relation $\prec$ and set and $T^l = T^{l-1} \cup \{T(x_l, \xi_l, s_l)\}$. We claim that at each step of the algorithm all the selected tent $T(x_l, \xi_l, s_l)$ are pairwise $Q^+$-disjoint. Reasoning by contradiction, suppose that two tents $T(x_l, \xi_l, s_l)$ and $T(x_{l'}, \xi_{l'}, s_{l'})$ with $l < l'$ are not $Q^+$-disjoint, then $B_{Qs_l}(x_l) \cap B_{Qs_{l'}}(x_{l'}) \neq \emptyset$ and there also exists a $c \in \mathbb{R}$ such that $s_l(\xi_l - c) \in \Theta^+$ and $s_{l'}(\xi_{l'} - c) \in \Theta^+$. Recall that $\Theta^+ = [0, \lambda^+]$ so

$$\xi_l - s_l^{-1} \alpha^+ \leq c \leq \xi_l \quad \xi_{l'} - s_{l'}^{-1} \alpha^+ \leq c \leq \xi_{l'}$$

If $s_l \geq Qs_l$ one would have

$$-s_l^{-1} Q^{-1} \alpha^+ \leq -s_l^{-1} \alpha^+ \leq \xi_l - \xi_{l'} \leq s_l^{-1} \alpha^+$$

and thus $s_l(\xi_l - \xi_{l'}) \in \Theta$ as long as $\alpha^+ \leq -R_0^{-1} \alpha^+$. This contradicts the maximality of $(x_l, \xi_l, s_l)$ that was chosen before $(x_{l'}, \xi_{l'}, s_{l'})$. On the other hand if $s_l < Qs_l$ then

$$-s_l^{-1} \alpha^+ \leq \xi_{l'} - \xi_l \leq s_l^{-1} \alpha^+$$

and, as long as $Q \geq R_0 \geq \lambda^+$, this implies that $(x_{l'}, \xi_{l'}, s_{l'}) \in 3Q^2 T(x_l, \xi_l, s_l)$ contradicting the selection condition.

Finally notice that the selection algorithm terminates after finitely many steps since at every step $\|M^+_R\|_{L^{p'}(\mathbb{R}^d)} \lesssim_R \|a\|_{L^{p'}(\mathbb{R}^d)}$ holds having chosen $Q \geq R$, since $s_l$ are bounded from below since $M^+_{R_l}(x_l, \xi_l, s_l) \geq \lambda$. Thus $E_{\lambda} \subset \bigcup_{l=1}^L 3Q^2 T_l$. \hfill \Box

A consequence of the above Lemma are non-iterated bounds for $M^+_R$.

Proposition 4.4. Given $a \in L^{p'}(\mathbb{R}^d)$ with $p' \in (r', \infty]$ the bound

$$\|M^+_R\|_{L^{p'}(\mathbb{R}^d)} \lesssim_R \|a\|_{L^{p'}(\mathbb{R}^d)}$$

holds. Furthermore the weak endpoint bound

$$\|M^+_R\|_{L^{p'}(\mathbb{R}^d)} \lesssim_R \|a\|_{L^{p'}(\mathbb{R}^d)}$$

holds. All the above bounds hold with a constant that grows at most polynomially in $R$ as $R \to \infty$ and is independent of the stopping sequence $\epsilon$ appearing in (4.4).

The bound (4.7) for $p = \infty$ is straightforward:

$$M^+_R(y, \eta, t) = \int_{B_{R}(y)} \left( \sum_{k \in \mathbb{Z}} |a_k(z)|^{p'} 1_{\Theta^+} (t(\eta - c_k(z))) \right)^{1/p'} dz$$

$$\leq \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} |a_k(z)|^{r'} \tau^{-1} M^+_R \left( \frac{z - y}{t} \right) dz \lesssim \|a\|_{L^{p'}(\mathbb{R}^d)}.$$
It is sufficient to show bound (4.8) so that will follow for \( p \in (r', \infty) \) by interpolation. In particular to obtain (4.8) we will show that given \( \lambda > 0 \) the bound on the measure of the super-level set
\[
\mu(E_{\lambda,R}) \lesssim R^{-r'} \|a\|_{L^{r'}(l^{r'})}^r
\]
holds. It is sufficient to consider the covering provided by Lemma 4.3 with \( Q = R \).

Since \( (x_l, \xi_l, s_l) \in \tilde{E}_{\lambda,R} \) and the covering \( T = \mathcal{T} = \{ T(x_l, \xi_l, s_l) \}_{l \in L} \) consists of disjoint tents, the bound (4.5) gives
\[
\lambda' \sum_{l=1}^L s_l \leq (2R)^{-1} \|a\|_{L^{r'}(l^{r'})}^r.
\]

Since \( E_{\lambda,R} \subset \bigcup_{l=1}^L 3R^2 T(x_l, \xi_l, s_l) \) one deduces
\[
\mu(E_{\lambda,R}) \leq \sum_{l=1}^L \mu(T(\xi_l, x_l, s_l)) \leq \sum_{l=1}^L s_l \leq \frac{\|a\|_{L^{r'}(l^{r'})}}{\lambda'}
\]
where the implied constant grows polynomially in \( R \) as required.

The proof of (4.1) relies on a locality property and a strip projection lemma.

**Lemma 4.5 (Locality of \( \mathcal{M}_R^+ \)).** Consider a strip \( D = D(x, s) \) and a function \( a \in L^1_{\text{loc}}(l') \) with
\[
\text{dist} (\text{spt } a; B_s(x)) > Rs
\]
then for all \((y, \eta, t) \in D(x, s)\) we have
\[
\mathcal{M}_R^+ 1_{D(x,s)} = 0.
\]

**Proof.** The statement follows directly from the definition (4.4) of the embedding. As a matter of fact if \((y, \eta, t) \in D_{x,s}\) then \( B_{R}(y) \subset B_{sR}(y) \) and \( B_{sR}(y) \cap \text{spt } a = \emptyset \). \( \square \)

**Lemma 4.6 (Mass projection for \( \mathcal{M}_R^+ \)).** Fix any collection of pairwise disjoint strips \( D(\zeta_m, \tau_m), m \in \{1, \ldots, M\} \) and any finite collection of disjoint tents
\[
T(x_l, \xi_l, s_l) \not\subset \bigcup_{m=1}^M D(\zeta_m, 3\tau_m), \quad l \in \{1, \ldots, L\}
\]
with \( Q > 2R > 2 \). Given a function \( a \in L^1_{\text{loc}}(l') \) and a stopping sequence \( \epsilon \) there exists a function \( \tilde{a} \in L^1_{\text{loc}}(l') \) and a new stopping sequence \( \tilde{\epsilon} \) such that
\[
\|\tilde{a}(z)\|_{l'} \lesssim \int_{B_{\tau_m}(\zeta_m)} \|a(z)\|_{l'} dz \quad \forall z \in B_{\tau_m}(\zeta_m) \quad \forall m \in \{1, \ldots, M\}
\]
\[
\tilde{a}_k(z) = a_k(z) \quad \forall z \notin \bigcup_{m=1}^M B_{\tau_m}(\zeta_m)
\]
and
\[
\mathcal{M}_R^+(x_l, \xi_l, s_l) \geq \mathcal{M}_R^+(x_l, \xi_l, s_l) \quad \forall l \in \{1, \ldots, L\}.
\]
where \( \mathcal{M}_R^+ \) is the embedded function as given by expression (4.4) associated to \( \tilde{a} \) with the stopping sequence \( \tilde{\epsilon} \).

**Proof.** Let us order the tents \( T(x_l, \xi_l, s_l) \) so that \( \xi_l \leq \xi_{l'} \) if \( l < l' \). For every strip \( D(\zeta_m, \tau_m) \) let
\[
\mathbb{I}_m := \{ l \in \{1, \ldots, L\} : D(x_l, Rs_l) \cap D(\zeta_m, \tau_m) \neq \emptyset \}.
\]
Set
\[ \tilde{\alpha}_k(z) = \begin{cases} a_k(z) & \text{if } z \notin \bigcup_m B_{r_m}(\zeta_m) \\ \int_{B_{r_m}(\zeta_m)} \left( \sum_{j \in Z} |a_j(z)|^{r'} 1_{\Theta^+} (s_l(\xi - c_j(z))) \right)^{1/r'} dz & \text{if } z \in B_{r_m}(\zeta_m) \text{ and } k \in \mathbb{1}_m \\ 0 & \text{if } z \in B_{r_m}(\zeta_m) \text{ and } k \notin \mathbb{1}_m \end{cases} \]

\[ \tilde{\epsilon}_k(z) = \begin{cases} \epsilon_k(z) & \text{if } z \notin \bigcup_m B_{r_m}(\zeta_m) \\ \xi_k & \text{if } z \in B_{r_m}(\zeta_m) \text{ and } k \in \{1, \ldots, L\} \\ \xi_l & k < 1 \\ \xi_l & k > L. \end{cases} \]

The expressions above are well defined since \( D(\zeta_m, \tau_m) \) are pairwise disjoint.

The bound (4.10) follows by the Minkowski inequality. For \( z \in B_{r_m}(\zeta_m) \) one has
\[ ||\tilde{a}(z)||_{l'} = \left( \sum_{k \in Z} \left( \int_{B_{r_m}(\zeta_m)} \left( \sum_{j \in Z} |a_j(z)|^{r'} 1_{\Theta^+} (s_l(\xi_k - c_j(z))) \right)^{1/r'} dz \right)^{1/r'} \right)^{1/r'} \leq \int_{B_{r_m}(\zeta_m)} ||a(z)||_{l'}, \]
where the last inequality holds since the tents \( T(x_l, \xi_l, s_l) \) are \( Q^+ \)-disjoint.

It remains to show (4.11). Since \( T(x_l, \xi_l, s_l) \not\subset D(\zeta_m, 3\tau_m) \) for any \( m \) we have that
\[ B_{R_s}(x_l) \cap B_{r_m}(\zeta_m) \neq \emptyset \implies D(\zeta_m, \tau_m) \subset D(x_l, 2R_{s_l}) \]
so set
\[ \mathfrak{M}_l = \left\{ m : D(\zeta_m, \tau_m) \subset D(x_l, 2R_{s_l}) \right\}. \]

Using the definitions of \( \tilde{a} \) and \( \tilde{\epsilon} \) we obtain
\[ \mathfrak{M}^+_l(x_l, \xi_l, s_l) = \int_{B_{R_{s_l}}(x_l)} \left( \sum_{k \in Z} |\tilde{a}_k(z)|^{r'} 1_{\Theta^+} (s_l(\xi_l - \tilde{\epsilon}_k(z))) \right)^{1/r'} dz \gtrsim (4R_{s_l})^{-1} \int_{B_{R_{s_l}}(x_l)} \left( \sum_{k \in Z} |\tilde{a}_k(z)|^{r'} 1_{\Theta^+} (s_l(\xi_l - \tilde{\epsilon}_k(z))) \right)^{1/r'} dz + (4R_{s_l})^{-1} \sum_{m \in \mathfrak{M}_l} \int_{B_{r_m}(\zeta_m)} \left( \sum_{k \in Z} |\tilde{a}_k(z)|^{r'} 1_{\Theta^+} (s_l(\xi_l - \xi_k)) \right)^{1/r'} dz. \]

Using the fact that \( T(x_l, \xi_l, s_l) \) are \( Q^+ \)-disjoint with \( Q > 2R \) we obtain that \( s_l(\xi_l - \xi_k) \in \Theta^+, \ z \in B_{r_m}(\zeta_m), \) and \( \tilde{a}_k(z) \neq 0 \) only if \( k = l; \) thus
\[ \sum_{m \in \mathfrak{M}_l} \int_{B_{r_m}(\zeta_m)} \left( \sum_{k \in \mathbb{1}_m} |\tilde{a}_k(z)|^{r'} 1_{\Theta^+} (s_l(\xi_l - \xi_k)) \right)^{1/r'} dz = \sum_{m \in \mathfrak{M}_l} \int_{B_{r_m}(\zeta_m)} \tilde{a}_l(z) dz = \sum_{m \in \mathfrak{M}_l} \int_{B_{r_m}(\zeta_m)} \left( \sum_{j \in Z} |a_j(z)|^{r'} 1_{\Theta^+} (s_l(\xi_l - c_j(z))) \right)^{1/r'} dz. \]

This allows us to conclude that
\[ \mathfrak{M}^+_l(x_l, \xi_l, s_l) \gtrsim (4R_{s_l})^{-1} \int_{B_{R_{s_l}}(x_l)} \left( \sum_{k \in Z} |\tilde{a}_k(z)|^{r'} 1_{\Theta^+} (s_l(\xi_l - \tilde{\epsilon}_k(z))) \right)^{1/r'} dz = \mathfrak{W}_l^+(x_l, \xi_l, s_l). \]
We now have all the tools to prove \( \text{(1.2)} \) for \( \| \|_{R} \). We proceed by interpolation, as described in 2.1.3 between the four (weak) endpoints

\[
(p', q') \in \{ (\infty, \infty), (\infty, r'), (1, \infty), (1, r') \}.
\]

Proof of bounds 4.2 for \( \| \|_{R} \).

The bound for \( (p', q') = (\infty, \infty) \) follows directly from \( \text{(1.2)} \) with \( p' = \infty \).

The bound for \( (p', q') = (\infty, r') \) follows from the locality property \( \text{(1.3)} \). We must show that for any strip \( D(x, s) \) one has

\[
\| \|_{R} \|_{L^{r'}(\infty, \infty)} \lesssim R \nu(D(x, s)) \| a \|_{L^{r'}(\infty, \infty)}
\]

but due to locality and \( \text{(1.3)} \) we have that

\[
\| \|_{R} \|_{L^{r'}(\infty, \infty)} \lesssim R \| a \|_{L^{r'}(\infty, \infty)}
\]

as required.

The bound for \( (p', q') = (1, \infty) \) makes use of the Mass Projection Lemma 4.6. We need to show that for every \( \omega > 0 \) there exists \( \K \subseteq \K \) such that

\[
\nu(K) \lesssim \omega^{-1} \| a \|_{L^{r}(\infty, \infty)}
\]

and the union of intervals \( \K = \bigcup_{m=1}^{M} B_{\tau_{m}}(\zeta_{m}) \). Let

\[
K_{\omega} := \bigcup_{m=1}^{M} D(\zeta_{m}, 9\tau_{m}) \implies \nu(K_{\omega}) \lesssim \omega^{-1} \| a \|_{L^{r}(\infty, \infty)}
\]

by the weak \( L^{1} \) bound on the Hardy-Littlewood maximal function.

For any tent \( T(y, \eta, t) \not\subseteq D(\zeta_{m}, 3\tau_{m}) \) apply Lemma 4.6 with respect to the the strips \( D(\zeta_{m}, 3\tau_{m}) \) and the one tent \( T(\xi, x, s) \). By construction we obtain a function \( \tilde{a} \) such that \( \| \tilde{a} \|_{L^{\infty}(r')} \lesssim \omega \). Using the statement of the Lemma and bound \( \text{(1.7)} \) we have

\[
\| \tilde{a} \|_{L^{\infty}(r')} \lesssim \omega
\]

as required.

The proof of the case \( (p', q') = (1, r') \) goes along the same lines. Let us suppose, without loss of generality, that \( a \in C_{c}^{\infty}(r') \) We need to show that for every \( \omega > 0 \) there exists \( \K \subseteq \K \) such that

\[
\nu(K) \lesssim \omega^{-1} \| a \|_{L^{r}(\infty, \infty)}
\]

Choose \( \K = \bigcup_{m=1}^{M} D(\zeta_{m}, 9\tau_{m}) \) as before. Let \( \lambda > 0 \) and set

\[
E_{\lambda, R} = E_{\lambda, R} \cap (D(x, s) \setminus K_{\omega}) \quad E_{\lambda, R} = \{ (y, \eta, t) : \| \tilde{a} \|_{R} > \lambda \}.
\]

The Covering Lemma 4.3 can be applied to \( E_{\lambda, R} \) with \( Q > 2R \) sufficiently large yielding a covering \( \{ T(x_{i}, \xi_{i}, s_{i}) \}_{i \in \{1, \ldots, L\}} \) such that \( \bigcup_{i=1}^{L} 3Q_{i}^{2}T(x_{i}, \xi_{i}, s_{i}) \supseteq E_{\lambda, R} \) with the tents \( T(x_{i}, \xi_{i}, s_{i}) \) that are pairwise \( Q^{2} \)-disjoint. Now apply the Mass Projection Lemma 4.6 with respect to the strips \( D(\zeta_{m}, 3\tau_{m}) \) and the tents \( T(\xi_{i}, x_{i}, s_{i}) \). The resulting \( \tilde{a} \) satisfies \( \| \tilde{a} \|_{L^{\infty}(r')} \lesssim \omega \) while

\[
\| \tilde{a} \|_{L^{\infty}(r')} \geq \| \tilde{a} \|_{L^{\infty}(r')} \geq \lambda.
\]
Using the bound (4.5) and the locality property 4.5 we have that
\[ µ(\mathcal{E}_{\lambda,R}) \lesssim R \sum_{l=1}^{L} s_l \lesssim \lambda^{-r'} \| \tilde{a} \mathcal{B}_{2^n}(x) \|_{L^{r'}(l')} \lesssim s \omega^{-r'} \lambda^{-r'} . \]
This concludes the proof. □

5. Proof of Theorem 1.3

In the previous section the bounds (4.2) were shown to hold for the auxiliary embedding \( M \). To prove Theorem 1.3 it is sufficient to show that the values of \( M \) control \( \| \cdot \|_{S_m} \). More specifically we require the following result.

Proposition 5.1. Given any union of strips \( K \) and a union of tents \( E \) such that
\[ M(y, \eta, t) \leq \lambda \quad \forall (y, \eta, t) \in \mathcal{X} \setminus (K \cup E) \]
the bound \( \| \mathbb{A} \mathbb{1}_{\mathcal{X}\setminus(K\cup E)} \|_{S_m} \lesssim \lambda \) holds.

Assuming that the above statement holds, Theorem 1.3 follows by the monotonicity property of outer \( L^p \) sizes 2.7.

The above proposition follows from showing that the required bound holds for all local sizes: \( \| \mathbb{A} \mathbb{1}_{\mathcal{X}\setminus(K\cup E)} \|_{S_m(T)} \lesssim \lambda \). The proof is divided into two parts relative to showing \( L^1 \)-type bounds over \( T^{(i)} \) and \( L^2 \) type bounds over \( T^{(e)} \) (see (2.17)).

The former part uses crucial disjointness properties related to the conditions (1.17) on the truncated wave packets.

The latter part depends on the fact that the sizes over a single tent \( T \) resembles an \( L^2 \) estimate for variational truncation of the Hilbert transform or of a square function in the spirit of [JSW08]. We will elaborate on this variational estimate in Lemma 5.2 in the following part on technical preliminaries.

The proof also involves a crucial stopping time argument. Similarly to the rest of the paper we avoid discretization and formulate a continuous version of this argument that we isolate Lemma 5.5 below.

5.1. Technical preliminaries. The following variational truncation bounds are a slightly modified version of the results appearing in [JSW08].

Lemma 5.2 (Variational truncations of singular integral operators [JSW08]). For any function \( H \in L^p(\mathbb{R}) \) and \( \sigma \in [0, \infty) \) let us define the variational truncation operator
\[ \mathcal{V}^\sigma_r H(z) = \sup_{\sigma < t_1 < \cdots < t_k < \cdots} \left( \sum_k | H \ast \Upsilon_{t_{k+1}}(z) - H \ast \Upsilon_{t_k}(z) |^r \right)^{1/r} \]
where \( \Upsilon \in S(\mathbb{R}) \), \( \int_\mathbb{R} \Upsilon(z)dz = 1 \) and \( \Upsilon_t(z) := t^{-1} \Upsilon \left( \frac{z}{t} \right) \).

If \( r > 2 \) and for any \( p \in (1, \infty) \), above operator satisfies the bounds
\[ \| \mathcal{V}^\sigma_r H \|_{L^p} \lesssim_{r,p} \| H \|_{L^p} \]
and if \( \sigma > 0 \) then
\[ \mathcal{V}^\sigma_r H(z) \lesssim_{r,p} \int_{B_r(z)} M \left( \mathcal{V}^\sigma_r H \right)(z')dz' \]
where \( M \) is the Hardy-Littlewood maximal function. The implicit constants are allowed to depend on \( \Upsilon \).

We record some useful properties of so called convex regions of tents.
Definition 5.3 (Convex regions). A convex region of a tent is a subset $\Omega \subset T(x, \xi, s)$ of a tent of the form
\begin{equation}
\Omega := \bigcup_{\theta \in \Theta} \Omega_{\theta} := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\}.
\end{equation}
for some function $\sigma_{\theta}(y) : \Theta \times B_s(x) \to [0, s]$.

Proof. Fix $y$ and follow that $|\epsilon| > 0$ was arbitrary and one can invert the role of $y''$ and $y$ in the above reasoning we obtain that $|\tilde{\sigma}(y'') - \tilde{\sigma}(y)| \leq \frac{|y'' - y|}{L}$ as required.

Given any tent $T \in \mathbb{T}$, any collection of strips $\mathcal{D}$, and any collection of tents $\mathcal{T}$, the set
\[ \Omega = T \setminus \left( \bigcup_{D \in \mathcal{D}} D \cup \bigcup_{T \in \mathcal{T}} T \right) \subset T \]
is a convex region of the tent $T$. With the next lemma we show that the bound \((5.1)\) on a convex regions can be extended to larger regions with scale bound $\sigma$ that is Lipschitz in the space variable.

Lemma 5.4 (Lipschitz convex regions). Let $T(x, \xi, s) \in \mathbb{T}$ be a tent and $\Omega = \bigcup_{\theta \in \Theta} \Omega_{\theta} \subset T(x, \xi, s)$ be a convex region as defined in \((5.5)\) and let us fix a constant $L > 2$. For every $\theta \in \Theta$ such that the bound \(5.1\) holds for $\Omega_{\theta} = \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\}$ and thus for any $y$ such that $\sigma_{\theta}(y) \in \Theta$.

\begin{align}
\forall (y, \eta, t) &\in \mathbb{R}^3, \quad \forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
\end{align}
\begin{align}
\forall \theta \in \Theta, \quad \theta^t \in (\theta, \lambda) \quad \text{such that} \quad \Omega := \{(y, \xi + \theta t^{-1}, t) : t > \sigma_{\theta}(y)\} \neq \emptyset,
thus $sW_s(x - y)W_s(z - y) \lesssim_s W_v(z - y')$ for all $z \in \mathbb{R}$

In the case that $t \in [2s, 3s]$ there also exists $y' \in B_s(x)$ and $t' \in (\sigma(y'), s)$ such that $t' \in (t/2, t)$ since $\Omega_y \neq \emptyset$. It follows from (5.9) that $|y' - y| > 2Ls$ so $|x - y| \approx_{L} |y' - y|$ so for all $z \in \mathbb{R}$

$$sW_s(x - y)W_s(z - y) \lesssim_{L} sW_s(y' - y)W_s(z - y) \lesssim_{L} W_v(z - y').$$

Thus, since in both cases $(y', \xi, \theta t^{-1}, t') \in \Omega$ we have by the definition (4.1) of $\mathbb{M}$ that

$$sW_s(x - y)\mathbb{M}(y, \xi, \theta t^{-1}, t) \lesssim_{L} \mathbb{M}(y', \xi, \theta t^{-1}, t) \leq \lambda$$

as required. \hfill \Box

The next technical lemma will be used as a continuous stopping time argument. It relates the Lipschitz assumption on enlarged convex regions of the previous statement with a crucial measurability estimate.

**Lemma 5.5** (Continuous stopping time). Let $\sigma : \mathbb{R} \rightarrow \mathbb{R}^+$ be a Lipschitz function with Lipschitz constant $L^{-1} < 1$. Then the function

$$\rho_\sigma(z) := \int_{\mathbb{R}} \frac{1}{2\sigma(x)} |B_{\sigma(x)}(z - x)dx$$

satisfies $1 + \frac{|z|}{\rho_\sigma(z)} \leq \rho_\sigma(z) < 1 + \frac{1}{\rho_\sigma(z)}$ and in particular for any non-negative function $h(z)$ the bounds

$$\int_{\mathbb{R}} h(z)dz \approx_{L} \int_{\mathbb{R}} \int_{B_{\sigma(z)}(x)} h(z)dz dx.$$

hold.

**Proof.** Since $\sigma$ is $L^{-1}$-Lipschitz, for any $z \in \mathbb{R}$ we have that

$$B_{(1 + L^{-1})^{-1} \sigma(z)} \subseteq \{x : z \in B_{\sigma(x)}(x)\} \subseteq B_{(1 + L^{-1})^{-1} \sigma(z)}.$$

By the same reason on $\{x : z \in B_{\sigma(z)}(x)\}$ we have that

$$(1 + L^{-1})^{-1} \sigma(z) \leq \sigma(x) \leq (1 + L^{-1})^{-1} \sigma(z).$$

The conclusion follows. \hfill \Box

### 5.2. Proof of Proposition 5.1

Let $T = T(x, \xi, s)$ be a tent and suppose that $K$ and $E$ are as in 5.1. Since the statement of Proposition 5.1 is invariant under time and frequency translations, we may assume, without loss of generality, that $T$ is centered at the origin i.e. $T = T(0, 0, s)$. If $T \setminus (K \cup E) = \emptyset$ there is nothing to prove. Let us set

$$\Theta_* = \{\theta \in \Theta : \exists(y, \theta t^{-1}, t) \in T(0, 0, s) \setminus (K \cup E)\},$$

$$\Theta_*^{(i)} := \Theta^{(i)} \cap \Theta_* \quad \Theta_*^{(e)} := \Theta^{(e)} \cap \Theta_*.$$

For $\theta \in \Theta_*$, using Lemma 5.4 we may assume that there exists a $L^{-1}$-Lipschitz function $\sigma_\theta : \mathbb{R} \rightarrow [0, 2s]$, with $L > 4$ sufficiently large to be chosen later, that satisfies condition (5.6) such that

$$sW_s(y)\mathbb{M}(y, \theta t^{-1}, t) \lesssim L \forall y \in \mathbb{R}, \theta \in \Theta, t \in (\sigma(y), 3s).$$

Let us set $\Omega = \bigcup_{\theta \in \Theta_\Theta} \Omega_\Theta$, $\Omega^{(i)} = \bigcup_{\theta \in \Theta^{(i)}} \Omega_\Theta$, and $\Omega^{(e)} = \bigcup_{\theta \in \Theta^{(e)}} \Omega_\Theta$ with

$$\Omega_\Theta = \begin{cases} \{(y, \theta t^{-1}, t) \in T(0, 0, s) : t > \sigma_\Theta(y) \} \quad \theta \in \Theta_* \\ \emptyset \quad \text{otherwise.} \end{cases}$$
We need to show that
\[ \| \mathbb{1}_{\mathcal{K}} \|_{S_m(T)} \leq \| \mathbb{1}_{\mathcal{I}} \|_{S_m(T)} \leq \lambda \quad \forall T \in \mathbb{T} \]
or equivalently (see (2.17)) that
\[ \| \mathbb{1}_{\mathcal{I}} \|_{S^1(T(\mathcal{I}))} \lesssim \lambda \quad \text{and} \quad \| \mathbb{1}_{\mathcal{I}} \|_{S^2(T(\mathcal{I}))} \lesssim \lambda. \]

In this proof all our implicit constants depend on the choice of \( L \).

Let us fix a choice of left truncated wave packets \( \Psi_{y,\eta,t}^{c,y}(z) \) in the defining expression (1.14). We will show that the statement holds in this case. The proof for right truncated wave packets is symmetric.

Comparing the definitions (1.14) and (1.1) for \( \mathbb{A} \) and \( \mathbb{M} \) respectively, it follows from the bound \( \| \Psi_{y,\eta,t}^{c,y}(z) \|_{S^1(\mathcal{I})} \leq W_t(z - y) \) that
\[ (5.11) \quad \mathbb{A}(y, \eta, t) \lesssim \mathbb{M}(y, \eta, t), \quad \| \mathbb{A} \|_{S^1(T(\mathcal{I}))} := \sup_{(y, \eta, t) \in \mathbb{I}} \mathbb{A}(y, \eta, t) \lesssim \lambda. \]

This implies
\[ (5.12) \quad \frac{1}{s} \iint_{(y, \eta, t) \in \mathbb{I}} |\mathbb{A}(y, \eta, t)|^2 \leq \frac{1}{s} \iint_{(y, \eta, t) \in \mathbb{I}} |\mathbb{A}(y, \eta, t)| d\eta dy dt \]
and thus we may assume that \( \alpha^- = \beta^- < 0 < \beta^+ < \alpha^+ \) and we can reduce to showing
\[ (5.13) \quad \| \mathbb{A} \|_{S^1(T(\mathcal{I}))} \lesssim \lambda \quad \text{and} \quad \| \mathbb{A} \|_{S^2(T(\mathcal{I}))} \lesssim \lambda. \]

5.2.1. Proof of the first inequality of (5.13).

It holds that
\[ \| \mathbb{A} \|_{S^1(T(\mathcal{I}))} \approx \iint_{\mathbb{I}} \left| \int_{\Theta(1)} \sum_{k \in \mathbb{Z}} \mathbf{a}_k(z) \Psi_{y,\eta,t}^{k,(\mathcal{I}),\epsilon_{k+1}(\mathcal{I})}(z) d\eta \right| d\eta dy dt \]
\[ \leq \frac{1}{s} \int_{\Theta(1)} \int_{y \in \mathbb{B}_s} \int_{t = \sigma_k(y)} \int_{z \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |\mathbf{a}_k(z)| \left| \Psi_{y,\eta,t}^{k,(\mathcal{I}),\epsilon_{k+1}(\mathcal{I})}(z) \right| d\eta d\eta dy dt. \]

According to (1.17) the wave-packet \( \Psi_{y,\eta,t}^{k,(\mathcal{I}),\epsilon_{k+1}(\mathcal{I})}(z) \) vanishes unless \( \theta - tc_k(z) \in B_s(d) \) and \( tc_{k+1}(z) - \theta > d^- \). Since \( \Theta(1) \subset [-d^-, d^-] \), the integrand vanishes unless \( \epsilon_k(z) < 0 < \epsilon_{k+1}(z) \). Let \( k^*_0 \in \mathbb{Z} \) be the index, if it exists, such that this inequality holds and set \( \alpha^*(z) := \mathbf{a}_{k^*_0}(z), \beta^*(z) = \epsilon_{k^*_0}(z) \). If no such index exists simply set \( \alpha^*(z) = 0 \).

Using that given \( t < s \) and \( y \in \mathbb{B}_s \) one has
\[ \left| \Psi_{y,\eta,t}^{k,(\mathcal{I}),\epsilon_{k+1}(\mathcal{I})}(z) \right| \lesssim sW_s(z) t W_t(z - y)^2 \leq W_t(z - y) \]
and using the statement of Lemma 5.5 we have that
\[ \| \mathbb{A} \|_{S^1(T(\mathcal{I}))} \lesssim \frac{1}{s} \int_{\Theta(1)} \int_{y \in \mathbb{B}_s} \int_{t = \sigma_k(y)} \int_{z \in \mathbb{R}} \int_{x \in \mathbb{B}_s} |\alpha^*(z)| sW_s(z) t W_t(y - z)^2 \times 1_{B_s(d)}(\theta - tc^*(z)) d\eta d\eta dy \times \frac{dt}{t} d\eta d\theta = I + II \]
where
\[ I := \frac{1}{s} \int_{\Theta(1)} \int_{x \in \mathbb{R}} \int_{z \in \mathbb{B}_s} |\alpha^*(z)| \int_{y \in \mathbb{B}_s} W_t(y - z) dy \times \frac{dt}{t} d\eta d\theta \]
We bound
\[ \|sW\|_\infty \]
This concludes the proof for the first bound of (5.13).

Since the inmost integral vanishes unless \( t > (1 - 2/L)\sigma_\theta(x) \) we obtain
\[
I \lesssim \int_{\theta \in \Theta^{(c)}} \int_{x \in \mathbb{R}} W_s(x) \int_{y \in B_s} \left(1^{-(2/L)\sigma_\theta(x)}\right) W_t(y-x) d\theta dt dy dx
\]
\[
\lesssim \int_{\theta \in \Theta^{(c)}} \int_{x \in \mathbb{R}} W_s(x) \int_{y \in B_s}\left(1^{-(2/L)\sigma_\theta(x)}\right) t W_t(y-x) \frac{t}{2\sigma_\theta(x)} W_t(y-x) d\theta dt dy dx.
\]
Since the inmost integral vanishes unless \( |y-x| > 2\sigma_\theta(x) \), we have that
\[
\int_{t=\sigma_\theta(x)}^{1-(2/L)\sigma_\theta(x)} \frac{t}{2\sigma_\theta(x)} W_t(y-x) \frac{dt}{t} \lesssim W_{\sigma_\theta(x)}(y-x)
\]
and so using (5.10) we obtain
\[
II \lesssim \lambda \int_{\theta \in \Theta^{(c)}} \int_{x \in \mathbb{R}} W_s(x) \int_{y \in B_s} W_{\sigma_\theta(x)}(y-x) d\theta dy dx
\]
\[
\lesssim \lambda \int_{\theta \in \Theta^{(c)}} \int_{x \in \mathbb{R}} W_s(x) d\theta dy dx \lesssim \lambda.
\]
This concludes the proof for the first bound of (5.13).

5.2.2. Proof of the second inequality of (5.13). As noted in (5.12) we may suppose that \( \Theta^{(c)} = [\beta^+, \alpha^+] \); for ease of notation set \( \tilde{\Omega}^{(c)} = \Omega \cap T^{(c)} \) so the required quantity to bound becomes \( \|A \chi_{\tilde{\Omega}^{(c)}}\|_{L^1(T^{(c)})} \). We concentrate on showing the dual bound
\[
(5.16) \quad \frac{1}{s} \left| \iint_{x \in \mathbb{R}} h(y, \eta, t) \left( \sum_{k \in \mathbb{Z}} a_k(z) \Psi_\theta^{(c)}(x)_{k, \psi_{\eta, \theta, t}}(z) \right) dz d\eta dt \right| \lesssim \lambda \frac{||h(y, \eta, t)||_{L^2}}{s^{1/2}}.
\]
for any \( h \in C_c^\infty(\Omega^{(c)}) \) where the \( \| \cdot \|_{L^2} \) is the classical Lebesgue \( L^2 \) norm relative to the measure \( dy \rho dt \). A change of variables and the Minkowski inequality give
\[
\frac{1}{s} \left| \int_R \int_{\Omega^{(c)}} h(y, \eta, t) \int_{k \in Z} a_k(z) \Psi_{s, \eta, t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) dz \ dy \rho dt \right|
\leq \frac{1}{s} \int_R \sum_{k \in Z} |a_k(z)| \left| \int_{\Omega^{(c)}} h(y, \eta, t) \Psi_{s, \eta, t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) dy \rho dt \right| dz
\leq \frac{1}{s} \int_{\theta \in \Theta^{(c)}} \int_R \sum_{k \in Z} |a_k(z)| \left| \int_R \left| h(y, t^{-1} \theta, t) \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) \right| dy \rho dt \right| dz d\theta.
\]
On the other hand the Hölder inequality gives that
\[
\int_{\theta \in \Theta^{(c)}} \| h(y, t^{-1} \theta, t) \|_{L^2(\rho dt, t)} \ d\theta = \int_{\theta \in \Theta^{(c)}} \left( \int_{y \in B_s} \int_{t = \sigma(y)}^s |h(y, t^{-1} \theta, t)|^2 dy \rho \right)^{\frac{1}{2}} \ d\theta
\lesssim \| h(y, \eta, t) \|_{L^2(\rho dt, t)}^s
\]
where \( \| \cdot \|_{L^2(\rho dt, t)} \) is the classic Lebesgue \( L^2 \) norm with respect to the measure \( dy \rho \). Thus (5.16) follows by showing
\[
\frac{1}{s} \int_R \sum_{k \in Z} |a_k(z)| \left| \int_R \left| h(y, t^{-1} \theta, t) \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) \right| dy \rho dt \right| dz \lesssim \lambda \| h(y, \eta, t) \|_{L^2(\rho dt, t)}^s
\]
with a constant uniform in \( \theta \in \Theta^{(c)} \). For sake of notation from now on we will omit the dependence on \( \theta \) by writing
\[
h(y, t) := h(y, t^{-1} \theta, t) \quad \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) := \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) \quad \sigma(x) := \sigma(x).
\]
Using the above notation and Lemma (5.7) we write
\[
\frac{1}{s} \int_R \sum_{k \in Z} |a_k(z)| \left| \int_{y \in B_s} \left| \int_{t = \sigma(y)}^s h(y, t^{-1} \theta, t) \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) dy \rho \right| dt \right| dz \lesssim I + II
\]
where
\[
I := \frac{1}{s} \int_{x \in B_{2L}} \int_{y \in B_{\sigma(x)}(x)} \sum_{k \in Z} |a_k(z)| \left| \int_{y \in B_s} \left( \int_{t = \sigma(y)}^s h(y, t) \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) dy \rho \right) dt \right| dz dx
\]
\[
II := \frac{1}{s} \int_{x \in R} \int_{y \in B_{\sigma(x)}(x)} \sum_{k \in Z} |a_k(z)| \left| \int_{y \in B_s} \left( \int_{t = \sigma(y)}^s h(y, t) \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) dy \rho \right) dt \right| dz dx
\]
We start with bounding \( I \). Suppose that \( L > 1 \) is chosen large enough so that (5.15) holds and recall that \( \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) = 0 \) unless \( \theta - t \epsilon_k(z) \in B_{\epsilon}(d) \). We thus have
\[
I = \frac{1}{s} \int_{x \in B_{2L}} \int_{y \in B_{\sigma(x)}(x)} \sum_{k \in Z} |a_k(z)| \left| \int_{y \in B_s} \left( \int_{t = \sigma(y)}^s h(y, t) \Psi_{s, \theta t}^{\epsilon_k(z), \epsilon_{k+1}(z)}(z) dy \rho \right) dt \right| dz dx
\]
\[
\lesssim \frac{1}{s} \int_{x \in B_{2L}} \int_{B_{\sigma(x)}(x)} \left( \sum_{k \in Z} |a_k(z)| r \| \theta - \sigma(x) \epsilon_k(z) \|^{1/r} \right)^{1/r} \| \epsilon_k(z) \| \mathcal{H}_x(z) dx dz
\]
\[ \leq \frac{1}{s} \int_{x \in B_{2L_s}} M(x, \theta \sigma(x)^{-1}, \sigma(x)) \sup_{z \in B_{\sigma(x)}(x)} \mathcal{H}_x(z) \, dx \leq \frac{\lambda}{s} \int_{x \in B_{2L_s}} \sup_{z \in B_{\sigma(x)}(x)} \mathcal{H}_x(z) \, dx, \]

where

\[ \mathcal{H}_x(z) := \left( \sum_{k \in \mathbb{Z}} \left| \int_{y \in B_s} \int_{t = (1 - \frac{k}{2})\sigma(x)}^s h(y, t) \psi_{y,t}^{\alpha_k}(z) \, dy \right|^r \right)^{1/r}. \]

We claim that

\[ \mathcal{H}_x(z) \lesssim \mathcal{V}_{\sigma(x)}^\mathcal{H} H_x(z) + \mathcal{E}_{\sigma(x)}(z) \quad \text{and} \quad H_T(z) := \int_{t = 0}^T \int_{y \in B_s} h(y, t) \psi_{y,t}^{0,\infty}(z) \, dy \, dt \]

\[ \mathcal{E}_{\sigma(x)}(z) := \left( \int_{(1 - 2/L_s)\sigma(x)}^s |h^*(z, t)|^2 \frac{dt}{t} \right)^{1/2} \quad \text{and} \quad h^*(z, t) := \int_R |h(y, t)| W_t(z - y) \, dy \]

with \( \mathcal{V}_{\sigma(x)}^\mathcal{H} \) defined in Lemma 5.2 and that

\[ \| \mathcal{E}_0 \|_{L^2} \lesssim \| h \|_{L^2(\mathbb{R}^d \times \mathbb{R}^+)} \sup_{z \in B_{\sigma(x)}(x)} \mathcal{E}_{\sigma(x)}(z) \lesssim \int_{B_{2\sigma(x)}(x)} \mathcal{E}_0(z) \, dz \]

\[ \| H_x \|_{L^2} \lesssim \| h \|_{L^2(\mathbb{R}^d \times \mathbb{R}^+)} \sup_{z \in B_{\sigma(x)}(x)} \mathcal{H}_x(z) \, dx \lesssim \int_{B_{2\sigma(x)}(x)} M^\mathcal{V} H_x(z) \, dz. \]

This would provide us with the required bounds for \( I \). As a matter of fact, according to Lemma 5.5 and 5.2 we have that

\[ I \lesssim \frac{\lambda}{s^{1/2}} \int_{x \in B_{2L_s} B_{2\sigma(x)}(x)} \left( \| M^\mathcal{V} H_x(z) + \mathcal{E}_{\sigma(x)}(z) \|_{L^2} \right) \, dz \lesssim_R \frac{\lambda}{s^{1/2}} \left( \| M^\mathcal{V} H_x \|_{L^2} + \| \mathcal{E}_0 \|_{L^2} \right) \]

\[ \lesssim \frac{\lambda}{s^{1/2}} \left( \| H_x \|_{L^2} + \| \mathcal{E}_0 \|_{L^2} \right) \lesssim \lambda \frac{\| h \|_{L^2(\mathbb{R}^d \times \mathbb{R}^+)}^2}{s^{1/2}} \]

as required.

The first bound of (5.19) follows by the Young inequality and Fubini:

\[ \| \mathcal{E}_0 \|_{L^2} \leq \int_{\mathbb{R} \times \mathbb{R}^+} \left( \int_R |h(y, t)| W_t(z - y) \, dy \right)^2 \, dz \, dt \]

\[ \leq \int_{\mathbb{R}^+} \int_R |h(y, t)|^2 \, dy \left( \int_R W_t(z - y) \, dy \right)^2 \frac{dt}{t} \lesssim \| h \|_{L^2(\mathbb{R}^d \times \mathbb{R}^+)}^2. \]

The second bound follows from the fact that for small enough \( \varepsilon > 0 \) and as long as \( |z - z'| < \varepsilon t \) the bound

\[ |h^*(z, t) - h^*(z', t)| \leq \int_R |h(y, t)| W_t(z - y) - W_t(z' - y) \, dy \]

\[ \leq 2^{-100} \int_R |h(y, t)| W_t(z - y) \, dy = 2^{-100} h^*(z, t) \]

holds so similarly

\[ \left| \mathcal{E}_{\sigma(x)}(z) - \mathcal{E}_{\sigma(x)}(z') \right| \lesssim 2^{-100} \mathcal{E}_{\sigma(x)}(z) \]

as long as \( |z - z'| < \varepsilon \sigma(x) \) for some sufficiently small \( \varepsilon > 0 \).

\[ \int_{B_{\sigma(x)}(z)} \mathcal{E}_0(z') \, dz' \gtrsim \int_{B_{\sigma(x)}(z)} \mathcal{E}_{\sigma(x)}(z') \, dz' \gtrsim \varepsilon \int_{B_{\sigma(x)}(z)} \mathcal{E}_{\sigma(x)}(z) \, dz' = \mathcal{E}_{\sigma(x)}(z) \]

and the claim follows.
The first bound of (5.20) uses standard oscillatory integral techniques: notice that for \( t > t' \) one has
\[
\left| \int_{\mathbb{R}} \Psi_{y,t}(z) \Psi_{y',t'}(z) dz \right| \leq \frac{t'}{t} W_t(y - y')
\]
so
\[
\int |H_{t}(z)|^2 dz \lesssim 2 \int_{t=0}^{t'} \int_{y \in B_t} \int_{y' \in B_t} |h(y, t)| |h(y', t')| W_t(y - y') dy dy' dt dt'
\]
\[
\lesssim \|h\|_{L_t^2(y,y')}^2.
\]
The second bound follows directly from Lemma 5.2.

It remains to show inequality (5.18). Notice that
\[
\left| \nabla_{x \gamma} \int_{y \in B_t} \Psi_{y,t}(z) dz \right| \lesssim \left( \sum_{k \in \mathbb{Z}} \left( \int_{y \in B_t} |h(y, t)| \Psi_{y,t}^{\epsilon_k(z) \epsilon_{k+1}(z)}(z) dy \left| \frac{dt}{t} \right|^{1/r} \right)^{1/r}
\]
where for \( k \in \mathbb{Z} \) we set
\[
(5.21) \quad t_{k}(z) := \sup \left\{ t \in \left( (1 - 2/L) \sigma(x), \frac{1}{2} \sigma(x) \right) : \Psi_{y,t}^{\epsilon_k(z) \epsilon_{k+1}(z)}(z) \neq 0 \right\}
\]
\[
(5.22) \quad t_{k}^{+}(z) := \inf \left\{ t \in \left( (1 - 2/L) \sigma(x), \frac{1}{2} \sigma(x) \right) : \Psi_{y,t}^{\epsilon_k(z) \epsilon_{k+1}(z)}(z) \neq 0 \right\}.
\]
We have omitted writing the implicit dependence on \( x \in \mathbb{R} \) and we will simply ignore the indexes \( k \in \mathbb{Z} \) for which the above sets are empty. Notice that the intervals \([t_k^-(z), t_k^+(z)]\) are disjoint. According to the conditions (1.17) on the geometry of truncated wave packets the following bounds hold:
\[
(5.23) \quad t_{k}^{+}(z) \epsilon_{k}(z) \in B_{\epsilon}(\theta - d') \quad t_{k}^{-}(z) \epsilon_{k+1}(z) \geq \theta + d'.
\]
Using the smoothness conditions (1.15) on the wave packets and writing a Lagrange remainder term we have that
\[
(5.24) \quad \left| \Psi_{y,t}^{\epsilon_k(z) \epsilon_{k+1}(z)}(z) - \Psi_{y,t}^{\epsilon_k(z) \epsilon_{k+1}(z)}(z) \right| \leq \left( |t \epsilon_k(z)| + \max (d'' - \theta - t \epsilon_{k+1}(z), 0) \right) W_t(y - z)
\]
so the bound
\[
\mathcal{H}_r(z) \leq \mathcal{H}_{r,1}(z) + \mathcal{H}_{r,2}(z)
\]
holds. Notice that
\[
\int_{t_{k}^{+}(z)} t^2 |\epsilon_{k}(z)|^2 \frac{dt}{t} \leq \frac{|t^{+}_{k}(z) \epsilon_{k}(z)|^2}{2} \leq C_{\alpha^+}
\]
\[
\int_{t_{k}^{-}(z)} \max (d'' - \theta - t \epsilon_{k+1}(z), 0) \frac{dt}{t} \leq \int_{t^{+}_{k} \epsilon_{k+1}(z)} d'' - \theta \frac{dt}{t} \leq \int_{d'' + \theta} d'' - \theta \frac{dt}{t} \leq C_{d'', \alpha^+, \beta^+}
\]
for some constant \( C_{\alpha^+} \) and \( C_{d'', \alpha^+, \beta^+} \). Since \( r > 2 \), Cauchy-Schwartz gives
\[
\mathcal{H}_{r,2}(z) \leq \left( \sum_{k \in \mathbb{Z}} \int_{t_{k}^{-}(z)} \epsilon^2(z, t) \frac{dt}{t} \right)^{r} \lesssim \mathcal{E}_{\sigma}(z).
\]
This is consistent with (5.18). To estimate $H_{x,1}$: introduce a frequency cutoff $\tilde{T} \in S(\mathbb{R})$ such that

$$\tilde{T} \in C^\infty_c(B_{0+b}) \quad \tilde{T} \geq 0 \quad \tilde{T} = 1 \text{ on } B_0 \quad \Upsilon_\tau(z) := \tau^{-1} \Upsilon \left( \frac{z}{\tau} \right).$$

According to (1.16), $\hat{\Psi}^{0,+,\infty}_{y,t}$ is supported on $B_{1-t\theta}(t^{-1}\theta)$ so one has the following

$$\Psi^{0,+,\infty}_{y,t} * \Upsilon_\tau(z) = \Psi^{0,+,\infty}_{y,t}(z) \quad \text{ if } \frac{t}{\tau} \geq \frac{\theta + b}{\theta},$$

$$\Psi^{0,+,\infty}_{y,t} * \Upsilon_\tau(z) = 0 \quad \text{ if } \frac{t}{\tau} < \frac{\theta - b}{\theta}$$

$$|\Psi^{0,+,\infty}_{y,t} * \Upsilon_\tau(z)| \lesssim W_t(z - y) \quad \text{ if } \frac{\theta - b}{\theta + b} \leq \frac{t}{\tau} < \frac{\theta + b}{\theta}.$$

Thus

$$|H_x - H_x * \Upsilon_\tau(z)| \lesssim \int_{\tau^{\frac{\theta + b}{\theta}}/L}^\tau h^*(z,t) \frac{dt}{t}$$

so

$$H_{x,1}(z) \lesssim \left( \sum_{k \in Z} |H_x * \Upsilon_{t_k^+}(z) - H_x * \Upsilon_{t_k^-}(z)|^2 \right)^{1/2} + \left( \sum_{k \in Z} \left| \int_{t_k^+}^{t_k^-} h^*(z,t) \frac{dt}{t} \right|^2 \right)^{1/2}$$

$$\lesssim \left( \sum_{k \in Z} \left| \int_{t_k^+}^{t_k^-} h^*(z,t) \frac{dt}{t} \right|^2 \right)^{1/2} \lesssim \Upsilon_{\sigma(z)}(H_x(z) + E_{\sigma(z)}(z))$$

thus concluding the proof of (5.18) and the bound on the term $I$.

The estimate for the term $II$ can be done in a manner similar to the term $II$ in for the $S^1$ part of the size. Recall (5.15) so that in expression for $II$ one has that $z \in B_{\sigma(z)}(x)$, $|y - z| > \sigma(x) > t$, and $|x - y| \approx |y - z|$. We also have that $y \in B_s$ so

$$\Psi^{\tau_{k}(z),\tau_{k+1}(z)}_{y,t}(z) \lesssim sW_s(z) tW_t(z - y)^2$$

and $\Psi^{\tau_{k}(z),\tau_{k+1}(z)}_{y,t}(z) = 0$ unless $t\tau_k(z) < \theta < t\tau_{k+1}(z)$, thus

$$II \lesssim \frac{1}{s} \int_{x \in \mathbb{R}} W_s(x) \int_{y \in B_s} \int_{t = \sigma(y)}^{(2/L)\sigma(x)} W_t(y - x) \frac{dt}{t} \int_{B_{\sigma(z)}(x)} |a_k(z)|$$

$$\times 1_{\sigma(\theta - t\tau_k(z))} W_t(z - y) 1_{(\tau_{k},\tau_{k+1})(t^{-1}\theta)} dz \frac{dt}{t} dy dx$$

$$\lesssim \int_{x \in \mathbb{R}} W_s(x) \int_{y \in B_s} \int_{t = \sigma(y)}^{(2/L)\sigma(x)} \frac{t}{2\sigma(x)} W_t(y - x) \frac{dt}{t} dy dx$$

$$\times \int_{B_{\sigma(z)}(x)} \left( \sum_{k \in Z} |a_k(z)|^{t'/2} 1_{\sigma(\theta - t\tau_k(z))} \right)^{1/t'} W_t(y - z) dz \frac{dt}{t} dy dx$$

Since the inmost integral vanishes unless $|y - x| > 2\sigma(\theta)$, we have that

$$\left( \int_{t = \sigma(y)}^{(2/L)\sigma(x)} \frac{t}{2\sigma(x)} W_t(y - x) \frac{dt}{t} \right)^{1/2} \lesssim W_{\sigma}(y - x)$$

it follows that

$$II \lesssim \frac{1}{s} \int_{x \in \mathbb{R}} W_s(x) \int_{y \in B_s} W_{\sigma}(y - x) \left( \int_{t = 0}^{\tau} |h(y,t)|^2 \frac{dt}{t} \right)^{1/2} dy dx$$
32 GENNADY URAL'TSEV

\[ \lesssim \lambda \int_{x \in \mathbb{R}} W_s(x) M \left( \left( \int_{t=0}^s |h(\cdot, t)|^2 \frac{dt}{t} \right)^{1/2} \right)(x) dx \]

\[ \lesssim \lambda \|h\|_{L^2(\mathbb{R}^d; t^{1/2})} \]

This concludes the proof.

\[ \square \]

6. The energy embedding and non-iterated bounds

6.1. The energy embedding. Here we comment on how to deduce Theorem 1.2 from the result in [DPO15]. Let us fix a \( p \in (1, \infty) \) and \( q \in \left( \max(2, p'), +\infty \right) \) and without loss of generality let us suppose that \( \hat{f} \in C_c^\infty(\mathbb{R}) \). We will show that the weak versions of (1.11) holds i.e.

\[ \|F\|_{L^p, \infty \rightarrow L^q(S_e)} \lesssim \|f\|_{L^p}. \]  \hspace{1cm} (6.1)

By interpolation this would allow us to conclude the strong bounds of (1.11).

The paper [DPO15] deals with embeddings into the space \( X \) that they denote by \( Z \). The generating collection of tents that they make use of is described in Section 2.1.2 of that paper. Notice that the set of geometric parameters for the tents in the present paper (Section 6) is larger than the one in [DPO15] but a careful perusal of the proofs therein shows that the same statements hold for the extended range of parameters.

Let us recall the main statements from [DPO15].

**Theorem 6.1** (Theorem 1 of [DPO15]). Let \( f \in S(\mathbb{R}) \) with \( \hat{f} \in C_c^\infty \). Let \( p \in (1, 2) \) and consider the set \( \mathcal{I}_{f, \lambda, p} \) of maximal dyadic intervals contained in

\[ K_{f, \lambda, p} = \{x \in \mathbb{R} : M_p f(x) > \lambda \} \quad \text{and let} \quad K_{f, \lambda, p} := \bigcup_{B_\cdot(\cdot) \in \mathcal{I}_{f, \lambda, p}} D(\cdot, 3\tau). \]  \hspace{1cm} (6.2)

Then with \( q \in (p', \infty] \).

\[ \|F\|_{L^p, \infty \rightarrow L^q(S_e)} \lesssim \|f\|_{L^p} \lambda^{1-p/q} \|f\|_{L^p}^{p/q}. \]

We used the super level set of \( M_p f \) instead of the super level set of \( M_p(Mf) \) to define \( K_{f, \lambda, p} \). As mentioned in section 7.3.1 of [DPO15], the inner maximal function appears only in the reduction from the case with \( \hat{f} \) compactly supported to the case with a general \( f \in S(\mathbb{R}) \). By our assumptions we can effectively ignore this complication.

**Proposition 6.2** (Proposition 3.2 + equations (2.6) and (2.7) of [DPO15]). The estimate

\[ \|F 1_{D(x, s)}\|_{L^q(S_e)} \lesssim_{N, q} \left( 1 + \frac{\text{dist}(sptf; B_s(x))}{s} \right)^{-N} \|f\|_{L^q} \]

holds for all \( N > 0 \) and \( q \in (2, \infty] \).

**Lemma 6.3** (Equation (7.3) of [DPO15]). The estimate

\[ \|F 1_{D(x, s)}\|_{L^\infty(S_e)} \lesssim_{N} \left( 1 + \frac{\text{dist}(sptf; B_s(x))}{s} \right)^{-N} \inf_{z \in B_s(x)} Mf(z) \]

holds for any \( N > 0 \).
Corollary 6.4. Suppose that $spt f \cap B_{2s}(x) = \emptyset$ then

$$\|F 1_{D(x,s)}\|_{L^q(S_s)} \lesssim N_{p,q} \left(1 + \frac{\text{dist} (spt f; B_s(x))}{s}\right)^{-N} s^{-1/p} \|f\|_{L^p}$$

for all $p \in (1, 2), q > p'$, and $N > 0$.

Proof. If $spt f \cap B_{2s}(x) = \emptyset$ then $\inf_{x \in B_s(x)} M f(z) \lesssim s^{-1} \|f\|_{L^1}$. Using this fact and interpolating between the bounds from Proposition 6.2 and Lemma 6.3 we obtain the required inequality.

Fix $p \in (1, \infty]$ and $q \in (\max(p', 2), \infty]$ and let $\gamma \in (1, \min(p', 2))$ such that $q > \gamma$. We will now show that

$$\|F 1_K\|_{L^\gamma} \lesssim \lambda.$$

Since $\nu(K, \gamma) \lesssim \lambda^{-p} \|f\|_{L^p}$, this would prove (6.1).

Let us consider a strip $D(x, s) \in D$ and suppose that $D(x, s) \not\supset K, \gamma$, otherwise the estimate is trivial. We have $B_{2s}(x) \not\supset K, \gamma$. For an $N > 1$ large enough to be chosen later let us set

$$f(x) = f_0(x) + \sum_{k=1}^{\infty} f_k(x) = f(x) v\left(\frac{x - x_0}{5s}\right) + \sum_{k=1}^{\infty} f(x) \gamma\left(\frac{x - x_0}{5s2^{Nk}}\right),$$

where $\gamma(\cdot) = v(\cdot/2^N) - v(\cdot)$ with

$$v \in C_c^\infty(B_2) \quad v \geq 0 \quad v = 1 \text{ on } B_1.$$

Let $F_k$ be associated to $f_k$ via the embedding (1.6) and let $K, \gamma$ as in (6.2).

Since $K, \gamma \subset K, \gamma$ we have that $\|f_0\|_{L^\gamma} \lesssim \lambda$ and

$$\|F_0 1_{K\setminus K, \gamma} 1_{D(x,s)}\|_{L^\gamma(S_s)} \lesssim \lambda^{-1/\gamma} \|f_0\|_{L^\gamma} \lesssim \lambda^{1/\gamma}$$

by Theorem 6.1.

Since $K, \gamma \subset K, \gamma \not\supset B_{2s}(x)$ one has $\|f_k\|_{L^\gamma} \lesssim \nu(D(x, s))^1/\gamma 2^{Nk/\gamma}$ and by Corollary 6.3 we have that

$$\|F_k 1_{K\setminus K, \gamma} 1_{D(x,s)}\|_{L^\gamma(S_s)} \lesssim \nu(D(x, s))^1/\gamma 2^{-Nk/\gamma} \lesssim 2^{-Nk/\gamma} \lambda \lesssim 2^{-Nk} \lambda.$$

By quasi-subadditivity we can add up (6.4) and (6.5) to obtain

$$\|F 1_{K\setminus K, \gamma} 1_{D(x,s)}\|_{L^\gamma(S_s)} \lesssim \lambda.$$ 

Since $D(x, s)$ is arbitrary this implies (6.3).

6.2. Non-iterated bounds. We conclude by explaining that for $r \in (2, \infty]$ and $p \in (2, r)$ and simpler embedding bounds on the maps $f \mapsto F$ and $a \mapsto A$ are sufficient to prove boundedness on $L^p(\mathbb{R})$ of the Variational Carleson Operator (1.2) and thus also (1.1).

Hereafter we work with the non-iterated outer measure space $(\mathcal{X}, \mu)$. The energy embedding map satisfies the $L^p$ bounds

$$\|F\|_{L^p(S_s)} \lesssim \|f\|_{L^p} \quad p \in (2, \infty].$$

This follows directly from Proposition 6.2 by taking $s$ arbitrarily large.

Similarly, in Proposition 1.4, we have shown that the auxiliary embedding satisfies

$$\|\mathcal{M}\|_{L^{p'}(S_{s_0})} \lesssim \|a\|_{L^{p'}(U')} \quad p' \in (r', \infty].$$

and thus, by Proposition 5.1, we have that the variational mass embedding also satisfies such bounds:

$$\|\mathcal{A}\|_{L^{p'}(S_{s_0})} \lesssim \|a\|_{L^{p'}(U')} \quad p' \in (r', \infty].$$
It follows by the outer Hölder inequality \(2.3\) that
\[
\left| \iiint_X F(y, \eta, t) A(y, \eta, t) dy d\eta dt \right| \lesssim \| F \|_{L^p(S_\lambda)} \| A \|_{L^p(S_m)}.
\]
Using (6.6) and (6.8) and the wave-packet domination (1.13) it follows that (1.2) is bounded on \(L^p(\mathbb{R})\).

In conclusion we remark that the iterated outer-measure \(L^p\) spaces that were introduced provide an effective way of capturing the spatial locality property of the embedding maps. Both the proof of Theorem 1.3 and of Theorem 1.2 rely on first obtaining non-iterated bounds (see Propositions 4.4 and 6.2) and then using a locality lemma (see Lemmata 4.5 and 6.3) and a projection lemma (see Lemma 4.6 and Lemma 7.8 of [DPO15]) to bootstrap the full result.

References

[Car66] Lennart Carleson. “On convergence and growth of partial sums of Fourier series”. In: Acta Mathematica 116.1 (1966), pp. 135–157.
[CDPO16] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou. “Domination of multilinear singular integrals by positive sparse forms”. In: arXiv preprint arXiv:1603.05317 (2016).
[DMT16] Yen Do, Camil Muscalu, and Christoph Thiele. “Variational estimates for the bilinear iterated Fourier integral”. In: arXiv preprint arXiv:1605.00574 (2016).
[DPO15] Francesco Di Plinio and Yumeng Ou. “A modulation invariant Carleson embedding theorem outside local \(L^2\)”. In: arXiv preprint arXiv:1510.06433 (2015).
[DT15] Yen Do and Christoph Thiele. “\(L^p\) theory for outer measures and two themes of Lennart Carleson united”. In: Bulletin of the American Mathematical Society 52.2 (2015), pp. 249–296.
[Fef73] Charles Fefferman. “Pointwise convergence of Fourier series”. In: Annals of Mathematics (1973), pp. 551–571.
[Hun68] Richard A Hunt. “On the convergence of Fourier series”. In: Orthogonal expansions and their continuous analogues. 1968, pp. 235–255.
[JSW08] Roger L Jones, Andreas Seeger, and James Wright. “Strong variational and jump inequalities in harmonic analysis”. In: Transactions of the American Mathematical Society (2008), pp. 6711–6742.
[Lac15] Michael T Lacey. “An elementary proof of the \(A_2\) bound”. In: arXiv preprint arXiv:1501.05818 (2015).
[LT00] Michael Lacey and Christoph Thiele. “A proof of boundedness of the Carleson operator”. In: Mathematical Research Letters 7.4 (2000), pp. 361–370.
[NOT10] 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”. In: Mathematical Research Letters 17.3 (2010).
[Obe+12] Richard Oberlin et al. “A variation norm Carleson theorem”. In: Journal of the European Mathematical Society 14.2 (2012), pp. 421–464.