A WEAK L^2 ESTIMATE FOR A MAXIMAL DYADIC SUM OPERATOR ON $\mathbb{R}^n$

MALABIKA PRAMANIK AND ERIN TERWILLEGER

Abstract. Lacey and Thiele have recently obtained a new proof of Carleson's theorem on almost everywhere convergence of Fourier series. This paper is a generalization of their techniques (known broadly as time-frequency analysis) to higher dimensions. In particular, a weak-type (2,2) estimate is derived for a maximal dyadic sum operator on $\mathbb{R}^n$, $n > 1$. As an application one obtains a new proof of Sjölin's theorem on weak L^2 estimates for the maximal conjugated Calderón-Zygmund operator on $\mathbb{R}^n$.

1. Introduction

In 1966, Carleson [1] proved his celebrated theorem on almost everywhere convergence of Fourier series of square integrable functions on $\mathbb{R}$. This was followed by a new proof given by C. Fefferman [2] in 1973. The techniques used by C. Fefferman have become known as time-frequency analysis and have found wide application in harmonic analysis in recent years. In particular, Lacey and Thiele [3], [4] have refined and extended these ideas in their pioneering work on the bilinear Hilbert transform on $\mathbb{R}$. In 2000, they obtained a new proof of Carleson's theorem [5] in which these techniques play a crucial role. These powerful techniques stem from interaction of extremely deep ideas which include delicate orthogonality estimates, combinatorics, and quasi-orthogonal decompositions well-localized in both time and space. It is the goal of this paper to extend the techniques of time-frequency analysis of [5] to higher dimensions.

The main result of this paper is a weak-type L^2 estimate for a maximal dyadic sum operator in $\mathbb{R}^n$, $n > 1$. In dimension one, this operator may be thought of as a linearized and discretized version of the Carleson operator $\mathcal{C}$,

$$\mathcal{C}f(x) := \sup_N \left| \int_{-\infty}^{N} \widehat{f}(\xi) e^{2\pi i x \xi} \, d\xi \right|.$$

Date: November 5, 2004.

1991 *Mathematics Subject Classification.* Primary 42B20, 42B25; Secondary 42B10, 47B38 .

The main point of Lacey and Thiele's proof [5] is to show that the discretized operator satisfies a weak-type (2,2) estimate. This in turn implies a similar estimate for $\mathcal{C}$, which is a key ingredient in proving that the Fourier series of a square-integrable function on the circle converges almost everywhere.

We introduce a higher dimensional analogue of the linearized and discretized Carleson operator and adapt the methodology of Lacey and Thiele to prove that this operator maps $L^2(\mathbb{R}^n)$ to $L^{2,\infty}(\mathbb{R}^n)$. One of the distinguishing aspects of our proof is the introduction of an ordering of points in $\mathbb{R}^n$ which allows us to organize the higher dimensional rectangles and thus control the large sums that appear in the operator. Unlike the situation in dimension one, the mapping property above does not lead to an almost everywhere convergence result in higher dimensions. However, it gives as a corollary a result of Sjölin [6] on the weak L^2 boundedness of the maximal conjugated Calderón-Zygmund operator on $\mathbb{R}^n$.

The proof is divided into seven sections. The first section explains the notation and terminology and gives the statement of the main theorem. The second section lists the main ingredients of the proof and the argument that binds them together. The subsequent four sections are devoted to proving the different lemmas needed in the main argument. The final section provides a new proof of Sjölin's theorem as an application of our main result.

2. Main Theorem

Time-frequency analysis provides the crucial set of ideas in the recent progress made in the understanding of Carleson's theorem. In this type of analysis one heavily uses the structure of dyadic intervals. A dyadic interval has the form $[m2^k, (m+1)2^k)$ where k and m are integers and k is called the scale. A dyadic cube $I \subset \mathbb{R}^n$ is of the form

$$\prod_{j=1}^{n} I^j = \prod_{j=1}^{n} [m_j 2^k, (m_j+1)2^k),$$

where k and m_j are integers for all $j = 1, 2, \cdots, n$. We easily see that the n-dimensional volume is given by $|I| = 2^{nk}$. Let $c(I) = (c(I^1), \cdots, c(I^n))$ denote the center of I, and for $a > 0$, aI will denote the cube with the same center as I and whose volume is $a^n|I|$.

Consider the time-frequency plane in $2n$ dimensional space with points $(\mathbf{x}, \boldsymbol{\xi})$, where $\mathbf{x}$ denotes the time coordinate in $\mathbb{R}^n$ and $\boldsymbol{\xi}$ denotes the frequency coordinate in $\mathbb{R}^n$. A "rectangle" in the time-frequency plane is the cross product of a dyadic cube from the time plane and a dyadic cube from the frequency plane. To be more precise, for a rectangle, p, the projection onto the time plane will be denoted by I_p, and its projection onto the frequency plane will be denoted by ω_p. We will denote by $\mathbf{D}$ the set of rectangles $p = I_p \times \omega_p$ such that $|I_p||\omega_p| = 1$. An element of $\mathbf{D}$ will be called a *tile*.

As mentioned earlier, it is important for the higher dimensional version of our time-frequency analysis to introduce an ordering in $\mathbb{R}^n$ that will play a role analogous to the linear ordering on $\mathbb{R}$. This is especially relevant in a certain selection scheme used in Section 5 in analogy with the work of Lacey and Thiele. Although the choice of ordering is not unique (we will mention an alternative in Section 5), we find it convenient to work with the lexicographical order defined as follows. Given $\mathbf{a} = (a_1, a_2, \cdots, a_n), \mathbf{b} = (b_1, b_2, \cdots, b_n) \in \mathbb{R}^n$,

$$\mathbf{a} < \mathbf{b} \iff \begin{cases} a_1 < b_1 \\ a_1 = b_1 \,,\, a_2 < b_2 \\ \quad\vdots \\ a_1 = b_1 \,,\, a_2 = b_2 \,,\, \ldots \,,\, a_{n-1} = b_{n-1} \,,\, a_n < b_n \,, \end{cases}$$

where the right hand side above is to be read with Boolean "or" standard.

For a tile p with $\omega_p = \omega_p^1 \times \omega_p^2 \times \ldots \times \omega_p^n$, we can divide each dyadic interval ω_p^j into two parts. In other words, for $j = 1, 2, \ldots, n$, we get

$$\omega_p^j = (\omega_p^j \cap (-\infty, c(\omega_p^j)) \cup (\omega_p^j \cap [c(\omega_p^j), \infty)).$$

Then ω_p can be decomposed into 2^n subcubes formed from all combinations of cross products of these half intervals. We number these subcubes using the lexicographical order on the centers and denote the subcubes by $\omega_{p(i)}$ for $i = 1, 2, \ldots, 2^n$. A tile p is then the union of 2^n *semi-tiles* given by $p(i) = I_p \times \omega_{p(i)}$ for $i = 1, 2, \ldots, 2^n$.

Let us define translation, modulation, and dilation operators by

$$\begin{aligned} T_{\mathbf{y}} f(\mathbf{x}) &:= f(\mathbf{x} - \mathbf{y}) \\ M_{\boldsymbol{\eta}} f(\mathbf{x}) &:= f(\mathbf{x}) e^{2\pi i \boldsymbol{\eta} \cdot \mathbf{x}} \\ D_\lambda^q f(\mathbf{x}) &:= \lambda^{-n/q} f(\lambda^{-1} \mathbf{x}), \lambda > 0. \end{aligned}$$

Note that if we set $q = 2$, these operators are isometries on $L^2(\mathbb{R}^n)$. We fix a Schwartz function ϕ such that $\widehat{\phi}$ is real, nonnegative, supported in the cube $[-1/10, 1/10]^n$ and equal to 1 on the cube $[-9/100, 9/100]^n$. For a tile $p \in \mathbf{D}$ and $\mathbf{x} \in \mathbb{R}^n$ we define

$$\phi_p(\mathbf{x}) = M_{c(\omega_{p(1)})} T_{c(I_p)} D^2_{|I_p|^{1/n}} \phi(\mathbf{x}). \tag{2.1}$$

Using the following definition of the Fourier transform

$$\widehat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \mathbf{x} \cdot \boldsymbol{\xi}} \, d\mathbf{x},$$

one easily can see that

$$\widehat{\phi_p}(\boldsymbol{\xi}) = T_{c(\omega_{p(1)})} M_{-c(I_p)} D^2_{|\omega_p|^{1/n}} \widehat{\phi}(\boldsymbol{\xi}). \tag{2.2}$$

Equation (2.1) tells us that for each p the function ϕ_p is well localized in time with most of its mass in I_p while equation (2.2) tells us that $\widehat{\phi_p}$ is supported in $\frac{1}{5}\omega_{p(1)}$. Note also that the ϕ_p have the same $L^2(\mathbb{R}^n)$ norm.

Let m be a multiplier in $C^\infty(\mathbb{R}^n \setminus \{\mathbf{0}\})$ which is homogeneous of degree 0, and define

$$\left(\psi_p^{\boldsymbol{\zeta}}\right)\widehat{}\,(\boldsymbol{\xi}) = m(\boldsymbol{\xi} - \boldsymbol{\zeta})\widehat{\phi_p}(\boldsymbol{\xi}),$$

where $\boldsymbol{\zeta}$ is contained in $\omega_{p(r)}$ for some fixed $r \in \{2, 3, \ldots, 2^n\}$. Note that we have the following fact for all $\boldsymbol{\zeta} \in \omega_{p(r)}$:

$$|\psi_p^{\boldsymbol{\zeta}}(\mathbf{x})| \le C_\nu |I_p|^{-1/2}\Big(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\Big)^{-\nu},$$

where ν is a large integer whose value may vary at different places in the proof. To see this fact we write

$$\begin{aligned}
\psi_p^{\boldsymbol{\zeta}}(\mathbf{x}) &= \int_{\mathbb{R}^n} e^{2\pi i \boldsymbol{\xi}\cdot\mathbf{x}} \widehat{\phi_p}(\boldsymbol{\xi}) m(\boldsymbol{\xi} - \boldsymbol{\zeta})\, d\boldsymbol{\xi} \\
&= \int_{\mathbb{R}^n} e^{2\pi i \boldsymbol{\xi}\cdot\mathbf{x}} T_{c(\omega_{p(1)})} M_{-c(I_p)} D^2_{|\omega_p|^{1/n}} \widehat{\phi}(\boldsymbol{\xi}) m(\boldsymbol{\xi} - \boldsymbol{\zeta})\, d\boldsymbol{\xi} \\
&= e^{2\pi i c(I_p)\cdot c(\omega_{p(1)})} \int_{\mathbb{R}^n} e^{2\pi i \boldsymbol{\xi}\cdot(\mathbf{x} - c(I_p))} |I_p|^{1/2} \widehat{\phi}\big(|I_p|^{1/n}(\boldsymbol{\xi} - c(\omega_{p(1)}))\big) m(\boldsymbol{\xi} - \boldsymbol{\zeta})\, d\boldsymbol{\xi}.
\end{aligned}$$

Making the change of variable

$$\boldsymbol{\xi'} = |I_p|^{1/n}(\boldsymbol{\xi} - c(\omega_{p(1)}))$$

in the above integral, we obtain

$$\begin{aligned}
\psi_p^{\boldsymbol{\zeta}}(\mathbf{x}) &= e^{2\pi i c(I_p)\cdot c(\omega_{p(1)})} \int_{\mathbb{R}^n} e^{2\pi i \big(\boldsymbol{\xi'}|I_p|^{-1/n} + c(\omega_{p(1)})\big)\cdot(\mathbf{x} - c(I_p))} |I_p|^{-1/2} \widehat{\phi}(\boldsymbol{\xi'}) \times \\
&\qquad\qquad m\big(\boldsymbol{\xi'}|I_p|^{-1/n} + c(\omega_{p(1)}) - \boldsymbol{\zeta}\big)\, d\boldsymbol{\xi'} \\
&= e^{2\pi i \mathbf{x}\cdot c(\omega_{p(1)})} \int_{\mathbb{R}^n} e^{2\pi i \boldsymbol{\xi'}\cdot(\mathbf{x} - c(I_p))|I_p|^{-1/n}} |I_p|^{-1/2} \widehat{\phi}(\boldsymbol{\xi'}) \\
&\qquad\qquad m\big(\boldsymbol{\xi'} + |I_p|^{1/n}(c(\omega_{p(1)}) - \boldsymbol{\zeta})\big)\, d\boldsymbol{\xi'},
\end{aligned}$$

where the second equality is obtained using the fact that m is homogeneous of degree 0. Since $\boldsymbol{\zeta} \in \omega_{p(r)}$ and $\boldsymbol{\xi} \in \frac{1}{2}\omega_{p(1)}$, we have

$$|\boldsymbol{\zeta} - \boldsymbol{\xi}| \gtrsim |\omega_p|^{1/n},$$

from which it follows that

$$\boldsymbol{\xi'} + |I_p|^{1/n}(c(\omega_{p(1)}) - \boldsymbol{\zeta})) \gtrsim 1.$$

Thus all the derivatives of $m(\boldsymbol{\xi'} + |I_p|^{1/n}(c(\omega_{p(1)}) - \boldsymbol{\zeta}))$ are bounded. A standard integration by parts argument finishes the proof.

Using the following definition for the inner product

$$\langle f, g\rangle = \int_{\mathbb{R}^n} f(\mathbf{x})\overline{g(\mathbf{x})}\, d\mathbf{x}\,,$$

given $\boldsymbol{\zeta} \in \mathbb{R}^n$ and $f \in L^2(\mathbb{R}^n)$, we define an operator

$$B^r_{\boldsymbol{\zeta}} f(\cdot) = \sum_{p\in\mathbf{D}} \langle f, \phi_p\rangle \psi_p^{\boldsymbol{\zeta}}(\cdot) 1_{\omega_{p(r)}}(\boldsymbol{\zeta}).$$

Theorem 1. *There exists a constant C, depending only on dimension, so that for all $f \in L^2(\mathbb{R}^n)$ and $r \in \{2, 3, \dots, 2^n\}$*

$$\|\sup_{\boldsymbol{\zeta}\in\mathbb{R}^n} |B^r_{\boldsymbol{\zeta}} f|\|_{L^{2,\infty}(\mathbb{R}^n)} \le C\|f\|_{L^2(\mathbb{R}^n)}. \tag{2.3}$$

To prove the theorem, we will work with a linearized version of the operator. Consider a measurable function $\mathbf{x} \to \mathbf{N}(\mathbf{x}) = (N_1(\mathbf{x}), N_2(\mathbf{x}), \dots, N_n(\mathbf{x}))$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ and define a linear operator

$$B^r_{\mathbf{N}}(\mathbf{x}) := B^r_{\mathbf{N}(\mathbf{x})}(\mathbf{x}) = \sum_{p\in\mathbf{D}} \langle f, \phi_p\rangle \psi_p^{\mathbf{N}(\mathbf{x})}(\mathbf{x})(1_{\omega_{p(r)}} \circ \mathbf{N})(\mathbf{x}).$$

To prove (2.3) it will suffice to show that there exists a constant $C > 0$ such that for all $f \in L^2(\mathbb{R}^n)$

$$\sup_{\mathbf{N}:\mathbb{R}^n\to\mathbb{R}^n} \|B^r_{\mathbf{N}} f\|_{L^{2,\infty}(\mathbb{R}^n)} \le C\|f\|_{L^2(\mathbb{R}^n)}, \tag{2.4}$$

where the supremum is taken over all measurable functions $\mathbf{N}$ on $\mathbb{R}^n$.

By duality we will show that the adjoint operator

$$g \to \sum_{p\in\mathbf{D}} \langle (1_{\omega_{p(r)}} \circ \mathbf{N})\psi_p^{\mathbf{N}}, g\rangle \phi_p$$

maps $L^{2,1}(\mathbb{R}^n)$ into $L^2(\mathbb{R}^n)$ with bounds independent of the measurable function $\mathbf{N}$. Since $L^{2,1}(\mathbb{R}^n)$ is a Lorentz space, it suffices to show that the dual operator maps $L^{2,1}(\mathbb{R}^n) \cap \{1_E : E \subset \mathbb{R}^n, E \text{ measurable }, |E| < \infty\}$ into $L^2(\mathbb{R}^n)$. Hence, we need to show

$$\|\sum_{p\in\mathbf{D}} \langle (1_{\omega_{p(r)}} \circ \mathbf{N})\psi_p^{\mathbf{N}}, 1_E\rangle \phi_p\|_{L^2(\mathbb{R}^n)} \le C|E|^{1/2}. \tag{2.5}$$

By duality, (2.5) is equivalent to

$$|\sum_{p\in\mathbf{D}} \langle (1_{\omega_{p(r)}} \circ \mathbf{N})\psi_p^{\mathbf{N}}, 1_E\rangle \langle \phi_p, f\rangle| \le C|E|^{1/2}, \tag{2.6}$$

for all Schwartz functions f with L^2 norm one. We will further restrict the sum to an arbitrary finite subset $\mathbf{P}$ of $\mathbf{D}$.

Now for all integers j we have the identity

$$\sum_{p\in\mathbf{P}} |\langle (1_{\omega_{p(r)}} \circ \mathbf{N})\psi_p^{\mathbf{N}}, 1_E\rangle\langle\phi_p, f\rangle| =$$
$$2^{-\frac{jn}{2}} \sum_{u\in P(j)} |\langle (1_{\omega_{u(r)}} \circ \mathbf{N}_j)\psi_u^{\mathbf{N}_j}, 1_{2^j\otimes E}\rangle\langle\phi_u, 2^{-\frac{jn}{2}} f(2^{-j}(\cdot))\rangle|,$$

where for any set A we define $2^j \otimes A = \{2^j\mathbf{y} = (2^j y_1, 2^j y_2, \ldots, 2^j y_n) : \mathbf{y} \in A\}$, $\mathbf{N}_j(\mathbf{x}) = 2^{-j}\mathbf{N}(2^{-j}\mathbf{x})$, and $\mathbf{P}(j) = \{(2^j \otimes I_p) \times (2^{-j} \otimes \omega_p) : p \in \mathbf{P}\}$. By picking j so that $1 \leq 2^{jn}|E| \leq 2$, we can absorb $|E|$ into the constant on the right hand side of (2.6). Finally we note that the left hand side of (2.6) can be rewritten so that the estimate we need to show now becomes

$$\sum_{p\in\mathbf{P}} |\langle 1_{E\cap\mathbf{N}^{-1}[\omega_{p(r)}]}, \psi_p^{\mathbf{N}}\rangle\langle\phi_p, f\rangle| \leq C, \tag{2.7}$$

for all Schwartz functions f with L^2 norm one, measurable functions $\mathbf{N}$, measurable sets E with $|E| \leq 1$, and all finite subsets $\mathbf{P}$ of $\mathbf{D}$. For the rest of the paper we fix f, $\mathbf{N}$, and E in this manner. By $\mathbf{N}^{-1}[\omega_{p(r)}]$ we mean $\{\mathbf{x} : \mathbf{N}(\mathbf{x}) \in \omega_{p(r)}\}$.

3. Main Argument

We now set up some tools that we will use throughout the rest of the paper. Define a partial order $<$ on the set of tiles $\mathbf{D}$ by setting

$$p < p' \qquad \Longleftrightarrow \qquad I_p \subset I_{p'} \quad \text{and} \quad \omega_{p'} \subset \omega_p.$$

We have the property that if two tiles $p, p' \in \mathbf{D}$ intersect, then either $p < p'$ or $p' < p$. To see this, observe that dyadic cubes have the property that if two of them intersect, then one is contained in the other. This extends from the same property for dyadic intervals in dimension one. Now, suppose two tiles p and p' in $\mathbf{D}$ intersect, and without loss of generality let $|I_p| \leq |I_{p'}|$. Then p and p' intersect in both the time and frequency components, i.e. $I_p \cap I_{p'} \neq \emptyset$, $\omega_p \cap \omega_{p'} \neq \emptyset$. From size considerations, one obtains that $I_p \subset I_{p'}$ and $\omega_{p'} \subset \omega_p$, hence $p < p'$. A consequence of this property is that for a finite set of tiles $\mathbf{P}$, all maximal elements of $\mathbf{P}$ under $<$ must be disjoint sets.

A finite set of tiles $\mathbf{T}$ is called a *tree* if there exits a tile $t \in \mathbf{D}$ such that $p < t$ for all $p \in \mathbf{T}$. We call t the *top* of the tree $\mathbf{T}$ and denote it by $p_{\mathbf{T}} = I_{\mathbf{T}} \times \omega_{\mathbf{T}}$. Note that the top is unique but not necessarily an element of the tree. Another useful observation is that any finite set of tiles $\mathbf{P}$ can be written as a union of trees. Consider all maximal elements of $\mathbf{P}$ under $<$. Then a nonmaximal element $p \in \mathbf{P}$ must be less than, under "$<$", some maximal element $t \in \mathbf{P}$ which places p in the tree with top t. For $i \in \{1, 2, \ldots, 2^n\}$, we call a tree an *$i$-tree* denoted by $\mathbf{T}^i$ if

$$\omega_{\mathbf{T}(i)} \subset \omega_{p(i)}$$

for all $p \in \mathbf{T}^i$. Observe that any tree can be written as the disjoint union of i-trees. Also for fixed i_0, and $p, p' \in \mathbf{T}^{i_0}$ the subcubes $\omega_{p(i)}$ and $\omega_{p'(i)}$ are pairwise disjoint and disjoint from $\omega_{\mathbf{T}(i)}$ for all $i \in \{1, 2, \ldots, 2^n\} \setminus \{i_0\}$.

For $p \in \mathbf{D}$, define the *mass* of $\{p\}$ as

$$\mathcal{M}(\{p\}) = \sup_{\substack{u \in \mathbf{D} \\ p < u}} \int_{E \cap \mathbf{N}^{-1}[\omega_u]} \frac{|I_u|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_u)|}{|I_u|^{1/n}}\right)^{10n}} \, d\mathbf{x}.$$

We can then define the mass of a finite set of tiles $\mathbf{P}$ to be

$$\mathcal{M}(\mathbf{P}) = \sup_{p \in \mathbf{P}} \mathcal{M}(\{p\}).$$

Note that the mass of any set of tiles is at most one since by a change of variables

$$\mathcal{M}(\mathbf{P}) \leq \int_{\mathbb{R}^n} \frac{1}{(1 + |\mathbf{x}|)^{10n}} \, d\mathbf{x} \leq 1.$$

The *energy*, depending on r, of a finite set of tiles $\mathbf{P}$ is defined

$$\mathcal{E}(\mathbf{P}) = \sup_{\mathbf{T}^r \in \mathbf{P}} \Big(|I_{\mathbf{T}^r}|^{-1} \sum_{p \in \mathbf{T}^r} |\langle f, \phi_p \rangle|^2 \Big)^{1/2}.$$

Recall that $r \in \{2, 3, \ldots, 2^n\}$ is fixed and f is a fixed Schwartz function of $L^2(\mathbb{R}^n)$ norm one. The following three lemmata will provide the main steps in proving the theorem, and their proofs will be shown in the next four sections of the paper.

Lemma 1. *There exists a constant C_1 such that for any finite set of tiles $\mathbf{P}$ there is a subset $\mathbf{P}'$ of $\mathbf{P}$ such that*

$$\mathcal{M}(\mathbf{P} \setminus \mathbf{P}') \leq \frac{1}{4} \mathcal{M}(\mathbf{P}) \tag{3.1}$$

and $\mathbf{P}'$ is the union of trees $\mathbf{T}_j$ satisfying

$$\sum_j |I_{\mathbf{T}_j}| \leq \frac{C_1}{\mathcal{M}(\mathbf{P})}. \tag{3.2}$$

Lemma 2. *There exists a constant C_2 such that for any finite set of tiles $\mathbf{P}$ there is a subset $\mathbf{P}''$ of $\mathbf{P}$ such that*

$$\mathcal{E}(\mathbf{P} \setminus \mathbf{P}'') \leq \frac{1}{2} \mathcal{E}(\mathbf{P}) \tag{3.3}$$

and $\mathbf{P}''$ is the union of trees $\mathbf{T}_j$ satisfying

$$\sum_j |I_{\mathbf{T}_j}| \leq \frac{C_2}{\mathcal{E}(\mathbf{P})^2}. \tag{3.4}$$

Lemma 3. *(The Tree Inequality) There exists a constant C_3 such that for all trees $\mathbf{T}$*

$$\sum_{p\in\mathbf{T}} |\langle 1_{E\cap\mathbf{N}^{-1}[\omega_{p(r)}]}, \psi_p^{\mathbf{N}}\rangle\langle\phi_p, f\rangle| \le C_3|I_{\mathbf{T}}|\mathcal{E}(\mathbf{T})\mathcal{M}(\mathbf{T}). \tag{3.5}$$

We will now prove (2.7), and hence Theorem 1, assuming the three lemmata. In the argument below set

$$C_0 = C_1 + C_2.$$

Given a finite set of tiles $\mathbf{P}$, find a very large integer m_0 such that $\mathcal{E}(\mathbf{P}) \le 2^{m_0 n}$ and $\mathcal{M}(\mathbf{P}) \le 2^{2m_0 n}$. We construct by decreasing induction a sequence of pairwise disjoint sets $\mathbf{P}_{m_0}$, $\mathbf{P}_{m_0-1}$, $\mathbf{P}_{m_0-2}$, $\mathbf{P}_{m_0-3}$, ... such that

$$\bigcup_{j=-\infty}^{m_0} \mathbf{P}_j = \mathbf{P}$$

and such that the following properties are satisfied

(1) $\mathcal{E}(\mathbf{P}_j) \le 2^{(j+1)n}$ for all $j \le m_0$.
(2) $\mathcal{M}(\mathbf{P}_j) \le 2^{(2j+2)n}$ for all $j \le m_0$.
(3) $\mathcal{E}\big(\mathbf{P}\setminus(\mathbf{P}_{m_0}\cup\cdots\cup\mathbf{P}_j)\big) \le 2^{jn}$ for all $j \le m_0$.
(4) $\mathcal{M}\big(\mathbf{P}\setminus(\mathbf{P}_{m_0}\cup\cdots\cup\mathbf{P}_j)\big) \le 2^{2jn}$ for all $j \le m_0$.
(5) $\mathbf{P}_j$ is a union of trees $\mathbf{T}_{jk}$ such that $\sum_k |I_{\mathbf{T}_{jk}}| \le C_0 2^{-2jn}$ for all $j \le m_0$.

Assume momentarily that we have constructed a sequence $\mathbf{P}_j$ as above. Then to obtain estimate (2.7) we use (1), (2), (5), the observation that the mass is always bounded by 1, and Lemma 3 to obtain

$$\begin{aligned}
&\sum_{s\in\mathbf{P}} \big|\langle 1_{E\cap\mathbf{N}^{-1}[\omega_{p(r)}]}, \psi_p^{\mathbf{N}}\rangle\langle f,\phi_p\rangle\big| \\
\le&\sum_j \sum_{p\in\mathbf{P}_j} \big|\langle 1_{E\cap\mathbf{N}^{-1}[\omega_{p(r)}]}, \psi_p^{\mathbf{N}}\rangle\langle f,\phi_p\rangle\big| \\
\le&\sum_j \sum_k \sum_{p\in\mathbf{T}_{jk}} \big|\langle 1_{E\cap\mathbf{N}^{-1}[\omega_{p(r)}]}, \psi_p^{\mathbf{N}}\rangle\langle f,\phi_p\rangle\big| \\
\le& C_3 \sum_j \sum_k |I_{\mathbf{T}_{jk}}|\,\mathcal{E}(\mathbf{T}_{jk})\,\mathcal{M}(\mathbf{T}_{jk}) \\
\le& C_3 \sum_j \sum_k |I_{\mathbf{T}_{jk}}|\,2^{(j+1)n}\,\min(1, 2^{(2j+2)n}) \\
\le& C_3 \sum_j C_0 2^{-2jn} 2^{(j+1)n}\,\min(1, 2^{(2j+2)n}) \\
\le& 2^{3n} C_0 C_3 \sum_j \min(2^{jn}, 2^{-jn}) \le C_n\,.
\end{aligned}$$

This proves estimate (2.7).

It remains to construct a sequence of disjoint sets $\mathbf{P}_j$ satisfying (1)-(5). We start our induction at $j = m_0$ by setting $\mathbf{P}_{m_0} = \emptyset$. Then (1), (2), and (5) are clearly satisfied, while

$$\mathcal{E}(\mathbf{P} \setminus \mathbf{P}_{m_0}) = \mathcal{E}(\mathbf{P}) \leq 2^{m_0 n}$$
$$\mathcal{M}(\mathbf{P} \setminus \mathbf{P}_{m_0}) = \mathcal{M}(\mathbf{P}) \leq 2^{2m_0 n}\,,$$

hence (3) and (4) are also satisfied for $\mathbf{P}_{m_0}$.

Suppose we have selected pairwise disjoint sets $\mathbf{P}_{m_0}, \mathbf{P}_{m_0-1}, \dots, \mathbf{P}_m$ for some $m < m_0$ such that (1)-(5) are satisfied for all $j \in \{m_0, m_0-1, \dots, m\}$. We will construct a set of tiles $\mathbf{P}_{m-1}$ disjoint from all the sets already constructed such that (1)-(5) are satisfied for all $j = m-1$. This procedure is given by decreasing induction. We will need to consider the following four cases.

Case 1. $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{(m-1)n}$ and $\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{2(m-1)n}$.

In this case set $\mathbf{P}_{m-1} = \emptyset$ and observe that (1)-(5) trivially hold.

Case 2. $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) > 2^{(m-1)n}$ and $\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{2(m-1)n}$.

Use Lemma 2 to find a subset $\mathbf{P}_{m-1}$ of $\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)$ such that

$$\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}_{m-1})\big) \leq \frac{1}{2}\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq \frac{1}{2} 2^{mn} \tag{3.6}$$

and $\mathbf{P}_{m-1}$ is a union of trees (whose set of tops we denote by $\mathbf{P}^*_{m-1}$) such that

$$\sum_{t \in \mathbf{P}^*_{m-1}} |I_t| \leq C_2 \mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big)^{-2} \leq C_2 2^{-2(m-1)n}. \tag{3.7}$$

Then (3.6) gives (3) and (3.7) gives (5) for $j = m-1$. Since

$$\mathcal{E}(\mathbf{P}_{m-1}) \leq \mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{mn} = 2^{((m-1)+1)n}\,,$$

estimate (1) is satisfied for $j = m-1$. Also by our induction hypothesis we have

$$\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}_{m-1})\big) \leq \mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{2(m-1)n}\,,$$

hence (4) is satisfied for $j = m-1$. Finally $\mathbf{P}_{m-1}$ is contained in $\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)$ and hence its mass is at most the mass of the latter which is trivially bounded by $2^{(2(m-1)+2)n}$, thus (2) is also satified for $j = m-1$.

Case 3. $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{(m-1)n}$ and $\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) > 2^{2(m-1)n}$.

In this case we repeat the argument in case 2 with the roles of the mass and energy reversed. Precisely, use Lemma 1 to find a subset $\mathbf{P}_{m-1}$ of the set $\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)$ such that

$$\mathcal{M}\big(\mathbf{P}\setminus(\mathbf{P}_{m_0}\cup\cdots\cup\mathbf{P}_n\cup\mathbf{P}_{m-1})\big) \leq \frac{1}{4}\mathcal{M}\big(\mathbf{P}\setminus(\mathbf{P}_{m_0}\cup\cdots\cup\mathbf{P}_m)\big) \leq \frac{1}{4}2^{2mn} \tag{3.8}$$

and $\mathbf{P}_{m-1}$ is a union of trees (whose set of tops we denote by $\mathbf{P}^*_{m-1}$) such that

$$\sum_{t\in\mathbf{P}^*_{m-1}} |I_t| \leq C_1\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big)^{-1} \leq C_1 2^{-2(m-1)n}. \tag{3.9}$$

Then (3.8) gives (4) and (3.9) gives (5) for $j = m-1$. By induction we have

$$\mathcal{M}(\mathbf{P}_{m-1}) \leq \mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{2mn} = 2^{(2(m-1)+2)n},$$

thus (2) is satisfied for $j = m-1$. Finally (1) and (3) follow from the inclusion $\mathbf{P}_{m-1} \subset \mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)$ and the assumption $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{(m-1)n}$. This concludes the proof of (1)-(5) for $j = m-1$.

Case 4. $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) > 2^{(m-1)n}$ and $\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) > 2^{2(m-1)n}$.

This is the most difficult case since it involves elements from both of the previous cases. We start by using Lemma 1 to find a subset $\mathbf{P}'_{m-1}$ of the set $\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)$ such that

$$\mathcal{M}\big(\mathbf{P}\setminus(\mathbf{P}_{m_0}\cup\cdots\cup\mathbf{P}_m\cup\mathbf{P}'_{m-1})\big) \leq \frac{1}{4}\mathcal{M}\big(\mathbf{P}\setminus(\mathbf{P}_{m_0}\cup\cdots\cup\mathbf{P}_m)\big) \leq \frac{1}{4}2^{2mn} \tag{3.10}$$

and $\mathbf{P}'_{m-1}$ is a union of trees (whose set of tops we denote by $(\mathbf{P}'_{m-1})^*$) such that

$$\sum_{t\in(\mathbf{P}'_{m-1})^*} |I_t| \leq C_1\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big)^{-1} \leq C_1 2^{-2(m-1)n}. \tag{3.11}$$

We now consider the following two subcases of case 4.

Subcase 4(a). $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1})\big) \leq 2^{(m-1)n}$

In this subcase, we set $\mathbf{P}_{m-1} = \mathbf{P}'_{m-1}$. Then (3) is automatically satisfied for $j = m-1$ and also (5) is satisfied in view of (3.11). By the inductive hypothesis we have $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{mn} = 2^{((m-1)+1)n}$ and also $\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{2mn} = 2^{(2(m-1)+2)n}$. Since $\mathbf{P}_{m-1}$ is contained in $\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)$ the same estimates hold for $\mathcal{E}(\mathbf{P}_{m-1})$ and $\mathcal{M}(\mathbf{P}_{m-1})$, thus (1) and (2) also hold for $j = m-1$. Finally (4) for $j = m-1$ follows from (3.5) since $\mathbf{P}'_{m-1} = \mathbf{P}_{m-1}$.

Subcase 4(b). $\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1})\big) > 2^{(m-1)n}$

Here we use Lemma 2 one more time to find a subset $\mathbf{P}''_{m-1}$ of the set $\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1})$ such that

$$\begin{aligned}&\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1} \cup \mathbf{P}''_{m-1})\big)\\ \leq &\frac{1}{2}\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1})\big)\end{aligned} \tag{3.12}$$

and $\mathbf{P}''_{m-1}$ is a union of trees (whose set of tops we denote by $(\mathbf{P}''_{m-1})^*$) such that

$$\sum_{t \in (\mathbf{P}''_{m-1})^*} |I_t| \leq C_2 \mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1})\big)^{-2} \leq C_2 2^{-2(m-1)n}. \tag{3.13}$$

We set $\mathbf{P}_{m-1} = \mathbf{P}'_{m-1} \cup \mathbf{P}''_{m-1}$ and we observe that $\mathbf{P}_{m-1}$ is disjoint from all the previously selected $\mathbf{P}_j$'s. Since by the induction hypothesis the last term in (3.12) is bounded by $\frac{1}{2}\mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq \frac{1}{2} 2^{mn}$, the first term in (3.12) is also bounded by $2^{(m-1)n}$, thus (3) holds for $j = m-1$. Likewise, since

$$\begin{aligned}\mathcal{E}(\mathbf{P}_{m-1}) &\leq \mathcal{E}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{mn} = 2^{((m-1)+1)n}\\ \mathcal{M}(\mathbf{P}_{m-1}) &\leq \mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m)\big) \leq 2^{2mn} = 2^{(2(m-1)+2)n},\end{aligned}$$

(1) and (2) are satified for $j = m-1$. Since

$$\mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}_{m-1})\big) \leq \mathcal{M}\big(\mathbf{P} \setminus (\mathbf{P}_{m_0} \cup \cdots \cup \mathbf{P}_m \cup \mathbf{P}'_{m-1})\big)$$

(3.10) implies that (4) is satisfied for $j = m-1$. Now each of $\mathbf{P}'_{m-1}$ and $\mathbf{P}''_{m-1}$ is given as a union of trees, thus the same is true for $\mathbf{P}_{m-1}$. The set of tops of all of these trees, call it $(\mathbf{P}_{m-1})^*$, is contained in the union of the set of tops of the trees in $\mathbf{P}'_{m-1}$ and the trees in $\mathbf{P}''_{m-1}$, i.e in $(\mathbf{P}'_{m-1})^* \cup (\mathbf{P}''_{m-1})^*$. This implies that

$$\sum_{t \in (\mathbf{P}_{m-1})^*} |I_t| \leq \sum_{t \in (\mathbf{P}'_{m-1})^*} |I_t| + \sum_{t \in (\mathbf{P}''_{m-1})^*} |I_t| \leq (C_1 + C_2) 2^{-2(m-1)n} = C_0 2^{-2(m-1)n}$$

in view of (3.11) and (3.13). This proves (5) for $j = m-1$ and concludes the inductive step $j = m-1$. The construction of the $\mathbf{P}_j$'s is now complete.

4. Proof of lemma 1

Given a finite set of tiles $\mathbf{P}$, set $\mu = \mathcal{M}(\mathbf{P})$ and define

$$\mathbf{P}' = \{p \in \mathbf{P} : \mathcal{M}(\{p\}) > \frac{1}{4}\mu\}.$$

Clearly $\mathcal{M}(\mathbf{P} \setminus \mathbf{P}') \leq \frac{1}{4}\mu$, thus it remains to show that $\mathbf{P}'$ satisfies (3.2). By definition of the mass, for each $p \in \mathbf{P}'$ there is a tile $u(p) = u \in \mathbf{D}$ such that

$$\int_{E \cap \mathbf{N}^{-1}[\omega_u]} \frac{|I_u|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_u)|}{|I_u|^{1/n}}\right)^{10n}} \, d\mathbf{x} > \frac{\mu}{4}. \tag{4.1}$$

Set $\mathbf{U} = \{u(p) : p \in \mathbf{P}'\}$, and let $\mathbf{U}_{\max}$ be the subset of $\mathbf{U}$ containing all maximal elements of $\mathbf{U}$ under the partial order on tiles. As observed earlier, the tiles in $\mathbf{U}$ can be grouped into trees with tops in $\mathbf{U}_{\max}$. Now $\mathbf{U}$ is not necessarily a subset of $\mathbf{P}'$, but each $u \in \mathbf{U}$ is associated to a $p \in \mathbf{P}'$ as described above. In particular, if p is a maximal element in $\mathbf{P}'$, then there exists a $u \in \mathbf{U}$ with $p < u$ such that (4.1) holds. If this u is not in $\mathbf{U}_{\max}$ then there exists $u' \in \mathbf{U}_{\max}$ with $u < u'$. We must then have u' associated to another $p' \in \mathbf{P}'$ which implies, by maximality of p, that $p' < p$. Hence for each maximal element $p \in \mathbf{P}'$ there exists a unique element $u \in \mathbf{U}_{\max}$ with $p < u$, and there is at most one such maximal element for each $u \in \mathbf{U}_{\max}$. Therefore, we will show

$$\sum_{u \in \mathbf{U}_{\max}} |I_u| \leq C_1 \mu^{-1}, \tag{4.2}$$

which implies (3.2). Now we will rewrite (4.1) as

$$\frac{2^n - 1}{2^{n+2}} \mu \sum_{k=0}^{\infty} 2^{-kn} < \sum_{k=0}^{\infty} \int_{\substack{E \cap \mathbf{N}^{-1}[\omega_u] \\ \cap (2^k I_u \setminus 2^{k-1} I_u)}} \frac{|I_u|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_u)|}{|I_u|^{1/n}}\right)^{10n}} \, d\mathbf{x},$$

where we set $\frac{1}{2} I_u = \emptyset$. This estimate holds for all $u \in \mathbf{U}$, so in particular for every $u \in \mathbf{U}_{\max}$ there exists a $k \geq 0$ such that

$$\begin{aligned} \frac{2^n - 1}{2^{n+2}} \mu |I_u| 2^{-kn} &< \int_{\substack{E \cap \mathbf{N}^{-1}[\omega_u] \\ \cap (2^k I_u \setminus 2^{k-1} I_u)}} \frac{1}{\left(1 + \frac{|\mathbf{x} - c(I_u)|}{|I_u|^{1/n}}\right)^{10n}} \, d\mathbf{x} \\ &\leq \frac{|E \cap \mathbf{N}^{-1}[\omega_u] \cap (2^k I_u \setminus 2^{k-1} I_u)|}{C(\sqrt{n})^{10n} 2^{10kn}}, \end{aligned}$$

where the second inequality above follows from the fact that $\dfrac{|\mathbf{x} - c(I_u)|}{|I_u|^{1/n}} \sim \sqrt{n} 2^k$ for $\mathbf{x} \in (2^k I_u \setminus 2^{k-1} I_u)$. Here and throughout the paper C denotes a constant depending only on dimension and whose value may change at different places in the proof. Now we define for $k \geq 0$

$$\mathbf{U}_k = \{u \in \mathbf{U}_{\max} : C\mu |I_u| 2^{9kn} < |E \cap \mathbf{N}^{-1}[\omega_u] \cap 2^k I_u|\}.$$

Since $\mathbf{U}_{\max} = \bigcup_{k=0}^{\infty} \mathbf{U}_k$, if we show that

$$\sum_{u\in\mathbf{U}_k} |I_u| \le C2^{-8kn}\mu^{-1}, \tag{4.3}$$

summing over $k \in \mathbf{Z}^+$ gives us estimate (4.2).

We now concentrate on showing estimate (4.3). Fix $k \ge 0$ and select an element $v_0 \in \mathbf{U}_k$ so that $|I_{v_0}|$ is largest possible. Then select an element $v_1 \in \mathbf{U}_k \setminus \{v_0\}$ such that the enlarged rectangle $(2^k I_{v_1}) \times \omega_{v_1}$ is disjoint from $(2^k I_{v_0}) \times \omega_{v_0}$ and $|I_{v_1}|$ is largest possible. Continuing by induction, at the j-th step we select an element $v_j \in \mathbf{U}_k \setminus \{v_0, \dots, v_{j-1}\}$ so that $(2^k I_{v_j}) \times \omega_{v_j}$ is disjoint from the enlarged rectangles of previously selected tiles and $|I_{v_j}|$ is largest possible. Since we have a finite set of tiles, this process will terminate, and we will have the set of selected tiles in $\mathbf{U}_k$, which we will call $\mathbf{V}_k$.

Next we make some key observations about the tiles. First, note that elements of $\mathbf{U}_k$ are maximal in $\mathbf{U}$ and therefore disjoint. Second, for any $u \in \mathbf{U}_k$ there exits a selected tile $v \in \mathbf{V}_k$ with $|I_u| \le |I_v|$ and such that the enlarged rectangles of u and v intersect. We will associate u to this v. Third, if u and u' are both associated to the same v, then I_u and $I_{u'}$ are disjoint. Indeed, $(2^k I_u) \times \omega_u$ intersects $(2^k I_v) \times \omega_v$ which means $2^k I_u \cap 2^k I_v \ne \emptyset$ and $\omega_u \cap \omega_v \ne \emptyset$. This implies, together with the fact that $|I_u| \le |I_v|$, that $\omega_u \supset \omega_v$. Similarily $\omega_{u'} \supset \omega_v$. Therefore, one of ω_u and $\omega_{u'}$ contains the other. But u and u' are disjoint, thus I_u is disjoint from $I_{u'}$. Finally , all tiles $u \in \mathbf{U}_k$ associated to a particular $v \in \mathbf{V}_k$ satisfy $I_u \subset 2^{k+2} I_v$.

From the observations above and the definition of $\mathbf{U}_k$, we have

$$\begin{aligned}\sum_{u\in\mathbf{U}_k} |I_u| &\le \sum_{v\in\mathbf{V}_k} \sum_{\substack{u\in\mathbf{U}_k \\ u \text{ assoc } v}} |I_u| \\ &= \sum_{v\in\mathbf{V}_k} \Bigg| \bigcup_{\substack{u\in\mathbf{U}_k \\ u \text{ assoc } v}} I_u \Bigg| \le \sum_{v\in\mathbf{V}_k} 2^{(k+2)n} |I_v| \\ &\le C\mu^{-1} 2^{-9kn} 2^{(k+2)n} \sum_{v\in\mathbf{V}_k} |E \cap \mathbf{N}^{-1}[\omega_v] \cap 2^k I_v| \\ &\le C2^{2n}\mu^{-1}2^{-8kn}|E| \le C2^{2n}\mu^{-1}2^{-8kn},\end{aligned}$$

where we have used that for $v \in \mathbf{V}_k$, the enlarged rectangles are disjoint, and therefore so are the subsets $E \cap \mathbf{N}^{-1}[\omega_v] \cap 2^k I_v$ of E.

5. Proof of lemma 2

We begin by fixing a finite set of tiles $\mathbf{P}$ and $r = 2^n$. This choice of r ensures that $\boldsymbol{\eta} < \boldsymbol{\xi}$ in the lexicographical order for all $\boldsymbol{\eta} \in \omega_{p(1)}$ and $\boldsymbol{\xi} \in \omega_{p(r)}$. For $r \ne 2^n$ the proof goes through by a suitable permutation of the coordinates of $\mathbb{R}^n$ which changes the coordinate that takes precedence in the lexicographical order. Here we note that we can be less precise by taking any linear functional L that separates $\omega_{p(1)}$ and $\omega_{p(r)}$ in any given cube ω_p. Then we let $\boldsymbol{\eta} < \boldsymbol{\xi}$ if $L(\boldsymbol{\eta}) < L(\boldsymbol{\xi})$. In particular, we can take L to be the projection onto the appropriate axis so that the usual linear ordering on $\mathbb{R}$ is relevant. Let ϵ denote $\mathcal{E}(\mathbf{P})$. Define for $\mathbf{T}'$ a 2^n-tree

$$\Delta(\mathbf{T}') = \left(|I_{\mathbf{T}'}|^{-1} \sum_{p \in \mathbf{T}'} |\langle f, \phi_p \rangle|^2 \right)^{1/2}.$$

Consider all 2^n-trees $\mathbf{T}'$ contained in $\mathbf{P}$ which satisfy

$$\Delta(\mathbf{T}') \geq \frac{1}{2}\epsilon. \tag{5.1}$$

Among these select a 2^n-tree $\mathbf{T}'_1$ such that $c(\omega_{\mathbf{T}'_1})$ is minimal in the lexicographical order. Let $\mathbf{T}_1$ be the set of all $p \in \mathbf{P}$ such that $p < p_{\mathbf{T}'_1} = p_{\mathbf{T}_1}$. In other words $\mathbf{T}_1$ is the maximal tree containing $\mathbf{T}'_1$ with the same top as $\mathbf{T}'_1$. Now consider all 2^n-trees contained in $\mathbf{P} \setminus \mathbf{T}_1$ and select a 2^n-tree $\mathbf{T}_2$, such that $c(\omega_{\mathbf{T}'_2})$ is minimal. Let $\mathbf{T}_2$ be the set of all $p \in \mathbf{P}$ such that $p < p_{\mathbf{T}'_2} = p_{\mathbf{T}_2}$. Continue inductively to obtain a finite sequence of pairwise disjoint 2^n-trees

$$\mathbf{T}'_1\,,\, \mathbf{T}'_2\,,\, \ldots\,,\, \mathbf{T}'_q$$

and pairwise disjoint trees

$$\mathbf{T}_1\,,\, \mathbf{T}_2\,,\, \ldots\,,\, \mathbf{T}_q$$

where $p_{\mathbf{T}'_j} = p_{\mathbf{T}_j}$, $\mathbf{T}'_j \subset \mathbf{T}_j$, and the $\mathbf{T}'_j$ satisfy (5.1). Let

$$\mathbf{P}'' = \bigcup_{j=1}^{q} \mathbf{T}_j\ ,$$

then clearly

$$\mathcal{E}(\mathbf{P} \setminus \mathbf{P}'') \leq \frac{1}{2}\epsilon\ .$$

Thus we need to show that $\mathbf{P}''$ satisfies condition (3.4) of Lemma 2, i.e.

$$\sum_{j=1}^{q} |I_{\mathbf{T}_j}| \leq \frac{C_2}{\epsilon^2}. \tag{5.2}$$

Since the trees $\mathbf{T}'_j$ satisfy (5.1),

$$\begin{aligned}
&\frac{1}{4}\epsilon^2 \sum_j |I_{\mathbf{T}_j}| \\
&\leq \sum_j \sum_{p\in\mathbf{T}'_j} |\langle f,\phi_p\rangle|^2 \\
&= \sum_j \sum_{p\in\mathbf{T}'_j} \langle f,\phi_p\rangle\overline{\langle f,\phi_p\rangle} \\
&= \langle f, \sum_j \sum_{p\in\mathbf{T}'_j} \overline{\langle f,\phi_p\rangle}\phi_p\rangle \\
&\leq \|f\|_{L^2(\mathbb{R}^n)} \Big\| \sum_j \sum_{p\in\mathbf{T}'_j} \langle f,\phi_p\rangle\phi_p \Big\|_{L^2(\mathbb{R}^n)} \\
&\leq \Big\| \sum_j \sum_{p\in\mathbf{T}'_j} \langle f,\phi_p\rangle\phi_p \Big\|_{L^2(\mathbb{R}^n)}.
\end{aligned} \tag{5.3}$$

Letting $\mathbf{U} = \bigcup_j \mathbf{T}'_j$, we will show that

$$\Big\| \sum_{p\in\mathbf{U}} \langle f,\phi_p\rangle\phi_p \Big\|_{L^2(\mathbb{R}^n)} \leq C\Big(\epsilon^2 \sum_j |I_{\mathbf{T}_j}|\Big)^{1/2}, \tag{5.4}$$

which, together with (5.3), will give us (5.2). The square of the left hand side of (5.4) can be estimated by

$$\sum_{\substack{p,u\in\mathbf{U}\\ \omega_p=\omega_u}} |\langle f,\phi_p\rangle\langle f,\phi_u\rangle\langle\phi_p,\phi_u\rangle| + 2 \sum_{\substack{p,u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |\langle f,\phi_p\rangle\langle f,\phi_u\rangle\langle\phi_p,\phi_u\rangle|. \tag{5.5}$$

Here we have used that $\langle\phi_p,\phi_u\rangle = 0$ unless $\omega_{p(1)}$ intersects $\omega_{u(1)}$ which implies that either $\omega_p = \omega_u$ or $\omega_{u(1)}$ contains ω_p or $\omega_{p(1)}$ contains ω_u. We are then able to utilize the symmetry in p and u to combine the off diagonal terms. We estimate $|\langle f,\phi_p\rangle|$ and $|\langle f,\phi_u\rangle|$ by the larger one and bound the first term in (5.5) by

$$\begin{aligned}
&\sum_{p\in\mathbf{U}} |\langle f,\phi_p\rangle|^2 \sum_{\substack{u\in\mathbf{U}\\ \omega_p=\omega_u}} |\langle\phi_p,\phi_u\rangle| \\
&\leq \sum_{p\in\mathbf{U}} |\langle f,\phi_p\rangle|^2 \sum_{\substack{u\in\mathbf{U}\\ \omega_p=\omega_u}} C \frac{\min\left(\frac{|I_u|}{|I_p|},\frac{|I_p|}{|I_u|}\right)^{1/2}}{\left(1+\frac{|c(I_p)-c(I_u)|}{\max(|I_p|,|I_u|)^{1/n}}\right)^{10n}} \\
&\leq \sum_{p\in\mathbf{U}} |\langle f,\phi_s\rangle|^2 \sum_{\substack{u\in\mathbf{U}\\ \omega_p=\omega_u}} C \int_{I_u} \frac{1}{|I_p|}\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{1/n}}\right)^{-10n} d\mathbf{x}
\end{aligned}$$

$$\begin{aligned}
&\le C \sum_{p\in\mathbf{U}} |\langle f,\phi_p\rangle|^2 \\
&= C \sum_j \sum_{p\in\mathbf{T}'_j} |I_{\mathbf{T}_j}||I_{\mathbf{T}_j}|^{-1}|\langle f,\phi_p\rangle|^2 \\
&\le C \sum_j |I_{\mathbf{T}_j}|\epsilon^2,
\end{aligned}$$

where we have used that for $p \in \mathbf{U}$, the I_u for which $\omega_p = \omega_u$ are pairwise disjoint.

Using Cauchy-Schwarz, the second term in (5.5) can be estimated by

$$\begin{aligned}
&2\sum_j \sum_{p\in\mathbf{T}'_j} |\langle f,\phi_p\rangle| \sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |\langle f,\phi_u\rangle||\langle\phi_p,\phi_u\rangle| \\
&\le 2\sum_j \left\{\sum_{p\in\mathbf{T}'_j}|\langle f,\phi_p\rangle|^2\right\}^{1/2} \left\{\sum_{p\in\mathbf{T}'_j}\left(\sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |\langle f,\phi_u\rangle||\langle\phi_p,\phi_u\rangle|\right)^2\right\}^{1/2} \\
&\le 2\sum_j |I_{\mathbf{T}_j}|^{1/2}\Delta(\mathbf{T}'_j) \left\{\sum_{p\in\mathbf{T}'_j}\left(\sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |\langle f,\phi_u\rangle||\langle\phi_o,\phi_u\rangle|\right)^2\right\}^{1/2} \\
&\le 2\epsilon\sum_j |I_{\mathbf{T}_j}|^{1/2} \left\{\sum_{p\in\mathbf{T}'_j}\left(\sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |\langle f,\phi_u\rangle||\langle\phi_p,\phi_u\rangle|\right)^2\right\}^{1/2} .
\end{aligned}$$

To complete the proof, we need to show that the expression inside the curly brackets is bounded by $C\epsilon^2|I_{\mathbf{T}_j}|$. Since for a single tile u

$$\mathcal{E}(\{u\}) = \left(|I_u|^{-1}|\langle f,\phi_u\rangle|^2\right)^{1/2} = |I_u|^{-1/2}|\langle f,\phi_u\rangle| \le \epsilon,$$

we get that

$$\sum_{p\in\mathbf{T}'_j}\left(\sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |\langle f,\phi_u\rangle||\langle\phi_p,\phi_u\rangle|\right)^2 \le \epsilon^2 \sum_{p\in\mathbf{T}'_j}\left(\sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |I_u|^{1/2}|\langle\phi_p,\phi_u\rangle|\right)^2 .$$

Thus we now need to show that

$$\sum_{p\in\mathbf{T}^r_j}\left(\sum_{\substack{u\in\mathbf{U}\\ \omega_p\subset\omega_{u(1)}}} |I_u|^{1/2}|\langle\phi_p,\phi_u\rangle|\right)^2 \le C|I_{\mathbf{T}_j}|. \tag{5.6}$$

To prove this, we will need the following lemma.

Lemma 4. *Let $p \in \mathbf{T}'_j$ and $u \in \mathbf{T}'_k$. Then if $\omega_p \subset \omega_{u(1)}$, we have $I_u \cap I_{\mathbf{T}_j} = \emptyset$. If $u \in \mathbf{T}'_k, v \in \mathbf{T}'_l, u \neq v, \omega_p \subset \omega_{u(1)}$, and $\omega_p \subset \omega_{v(1)}$ for some fixed $p \in \mathbf{T}'_j$, then $I_u \cap I_v = \emptyset$.*

Proof. Since $\omega_p \subset \omega_{u(1)}$ and $\mathbf{T}'_j$ and $\mathbf{T}'_k$ are 2^n-trees, $\mathbf{T}'_j$ and $\mathbf{T}'_k$ are not the same tree. Otherwise $\omega_p \subset \omega_{u(2^n)}$. We know that $\omega_{\mathbf{T}'_j} \subset \omega_p \subset \omega_{u(1)}$, which implies that $c(\omega_{\mathbf{T}'_j})$ is contained in $\omega_{u(1)}$. We also have $\omega_{\mathbf{T}'_k} \subset \omega_{u(2^n)}$, which implies that $c(\omega_{\mathbf{T}'_k})$is contained in $\omega_{u(2^n)}$. Therefore $c(\omega_{\mathbf{T}'_j}) \leq c(\omega_{\mathbf{T}'_k})$ in the lexicographical order which means that $\mathbf{T}'_j$ was chosen before $\mathbf{T}'_k$ in the original selection process. Now suppose $I_u \cap I_{\mathbf{T}_j} \neq \emptyset$. Then either $I_u \subset I_{\mathbf{T}_j}$ or $I_u \supset I_{\mathbf{T}_j}$, however $\omega_{\mathbf{T}_j} \subset \omega_u$ implies that $I_u \subset I_{\mathbf{T}_j}$. Thus we have $\omega_{\mathbf{T}_j} \subset \omega_u$ and $I_u \subset I_{\mathbf{T}_j}$ which says that u belongs to the tree $\mathbf{T}_j$. However, $u \in \mathbf{T}_k$ and thus was chosen from $\mathbf{P} \setminus \mathbf{T}_j$, which gives a contradiction. Thus $I_u \cap I_{\mathbf{T}_j} = \emptyset$.

Next suppose that $u \in \mathbf{T}'_k, v \in \mathbf{T}'_l, u \neq v$, and $\omega_p \subset (\omega_{u(1)} \cap \omega_{v(1)})$ for some fixed $p \in \mathbf{T}'_j$. We have three cases to consider: (a) $\omega_u \subset \omega_{v(1)}$ which means $I_v \cap I_{\mathbf{T}_k} = \emptyset$ and thus $I_v \cap I_u = \emptyset$, (b) $\omega_v \subset \omega_{u(1)}$ which means $I_u \cap I_{\mathbf{T}_l} = \emptyset$ and thus $I_v \cap I_u = \emptyset$, and (c) $\omega_v = \omega_u$ which tells us $|I_u| = |I_v|$, thus I_u and I_v are disjoint since u and v don't coincide. $\square$

We now return to estimate (5.6). Observe that Lemma 4 tells us that for the tiles $u \in \mathbf{U}$ appearing in the interior sum of (5.6), the I_u are pairwise disjoint and contained in $(I_{\mathbf{T}_j})^c$. Thus we have

$$
\begin{aligned}
&\sum_{p \in \mathbf{T}'_j} \Bigg(\sum_{\substack{u \in \mathbf{U} \\ \omega_p \subset \omega_{u(1)}}} |I_u|^{1/2} |\langle \phi_p, \phi_u \rangle| \Bigg)^2 \\
\leq & C \sum_{p \in \mathbf{T}'_j} \Bigg(\sum_{\substack{u \in \mathbf{U} \\ \omega_p \subset \omega_{u(1)}}} |I_u|^{1/2} \left(\frac{|I_p|}{|I_u|} \right)^{1/2} \int_{I_u} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, d\mathbf{x} \Bigg)^2 \\
\leq & C \sum_{p \in \mathbf{T}'_j} |I_p| \Bigg(\sum_{\substack{u \in \mathbf{U} \\ \omega_p \subset \omega_{u(1)}}} \int_{I_u} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, d\mathbf{x} \Bigg)^2 \\
\leq & C \sum_{p \in \mathbf{T}'_j} |I_p| \Bigg(\int_{(I_{\mathbf{T}_j})^c} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, d\mathbf{x} \Bigg)^2 \\
\leq & C \sum_{p \in \mathbf{T}'_j} |I_p| \int_{(I_{\mathbf{T}_j})^c} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, d\mathbf{x}
\end{aligned}
$$

$$\leq C \sum_{k=0}^{\infty} 2^{-kn} |I_{\mathbf{T}_j}| \sum_{\substack{p \in \mathbf{T}_j' \\ |I_p| = 2^{-kn}|I_{\mathbf{T}_j}|}} \int_{(I_{\mathbf{T}_j})^c} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, d\mathbf{x}.$$

The proof of Lemma 2 will be complete if we can show that

$$\sum_{\substack{p \in \mathbf{T}_j' \\ |I_p| = 2^{-kn}|I_{\mathbf{T}_j}|}} \int_{(I_{\mathbf{T}_j})^c} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, d\mathbf{x} \lesssim 2^{k(n-1)},$$

thus allowing the sum in k to converge. Throughout the paper, $A \lesssim B$ means that A is less than or equal to B up to a constant depending only on dimension. The first observation we have is that

$$\int_{(I_{\mathbf{T}_j})^c} \frac{|I_p|^{-1}}{\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n}} \, dx \lesssim \left(\frac{\operatorname{dist}((I_{\mathbf{T}_j})^c, I_p)}{|I_p|^{1/n}} \right)^{-9n}.$$

To see this, note that by a change of variables, it suffices to let the center of $I_{\mathbf{T}_j}$ be at the origin. Also note that we have the inequality

$$\left(1 + \frac{|\mathbf{x} - c(I_p)|}{|I_p|^{1/n}}\right)^{10n} \geq \prod_{i=1}^{n} \left(1 + \frac{|x_i - c(I_p)_i|}{|I_p|^{1/n}}\right)^{10}.$$

Therefore, the integral above is bounded by a constant times

$$\prod_{i=1}^{n} \left(\int_{|x_i| > \frac{1}{2}|I_{\mathbf{T}_j}|} \frac{|I_p|^{-1/n}}{\left(1 + \frac{|x_i - c(I_p)_i|}{|I_p|^{1/n}}\right)^{10}} \, dx_i \right)$$
$$\lesssim \prod_{i=1}^{n} \left(\frac{\operatorname{dist}((I_{\mathbf{T}_j})^c, I_p)}{|I_p|^{1/n}} \right)^{-9} \lesssim \left(\frac{\operatorname{dist}((I_{\mathbf{T}_j})^c, I_p)}{|I_p|^{1/n}} \right)^{-9n},$$

where we have used that $|x_i - c(I_p)_i| \geq \operatorname{dist}((I_{\mathbf{T}_j})^c, I_p)$ for all $i = 1, 2, \ldots, n$. Now we need to sum over p for a fixed scale k. Consider an $n-1$ dimensional face of $I_{\mathbf{T}_j}$ and fix a cube I_p whose face is contained in the face of $I_{\mathbf{T}_j}$. We allow the remaining coordinate to vary and sum over those I_p in this "column". In a fixed column, the distances from $(I_{\mathbf{T}_j})^c$ to each I_p sum as additive multiples of $|I_p|^{1/n}$. For each face, there are $2^{k(n-1)}$ such columns.

Thus

$$\begin{aligned}&\sum_{\substack{p\in\mathbf{T}'_j\\|I_p|=2^{-kn}|I_{\mathbf{T}_j}|}}\left(\frac{\operatorname{dist}((I_{\mathbf{T}_j})^c,I_p)}{|I_p|^{1/n}}\right)^{-9n}\\&\lesssim 2^{k(n-1)}\times(\#\text{of faces})\sum_{m=0}^{\infty}\frac{1}{m^{9n}}\\&\lesssim 2^{k(n-1)}.\end{aligned}$$

6. Proof of Lemma 3 - The Tree Inequality

Let $\mathcal{J}$ be the collection of all maximal dyadic cubes J such that $3J$ does not contain any I_p with $p\in\mathbf{T}$. Then $\mathcal{J}$ is a partition of $\mathbb{R}^n$.

We can write the left hand side of (3.5) as follows, where the terms α_p are phase factors of modulus 1 which make up for the absolute value signs in (3.5):

$$||\sum_{p\in\mathbf{T}}\alpha_p\langle f,\phi_p\rangle\psi_p^{\mathbf{N}}1_{E_{2p}}||_1\le\mathcal{K}_1+\mathcal{K}_2$$

where

$$E_{2p}:=E\cap\mathbf{N}^{-1}[\omega_{p(2^n)}],\tag{6.1}$$

$$\mathcal{K}_1:=\sum_{J\in\mathcal{J}}\sum_{p\in\mathbf{T}\,;\,|I_p|\le 2^n|J|}||\langle f,\phi_p\rangle\psi_p^{\mathbf{N}}1_{E_{2p}}||_{L^1(J)},\tag{6.2}$$

$$\mathcal{K}_2:=\sum_{J\in\mathcal{J}}\left\|\sum_{p\in\mathbf{T}\,;\,|I_p|>2^n|J|}\alpha_p\langle f,\phi_p\rangle\psi_p^{\mathbf{N}}1_{E_{2p}}\right\|_{L^1(J)}.\tag{6.3}$$

Let

$$\epsilon=\mathcal{E}(\mathbf{T})\quad\text{and}\quad\mu=\mathcal{M}(\mathbf{T}).$$

We begin with $\mathcal{K}_1$. For every $p\in\mathbf{T}$, $\{p\}$ is a 2^n-tree contained in $\mathbf{T}$, and therefore

$$|\langle f,\phi_p\rangle|\le\epsilon|I_p|^{\frac{1}{2}}.$$

This implies that

$$\begin{aligned}
\mathcal{K}_1 &\le C\epsilon \sum_{J\in\mathcal{J}} \sum_{\substack{p\in\mathbf{T}\\|I_p|\le 2^n|J|}} |I_p| \int\limits_{J\cap E\cap\mathbf{N}^{-1}[\omega_p]} \frac{|I_p|^{-1}}{\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{20n}}\,d\mathbf{x}\\
&\le C\epsilon \sum_{J\in\mathcal{J}} \sum_{\substack{p\in\mathbf{T}\\|I_p|\le 2^n|J|}} |I_p| \left(\int\limits_{J\cap E\cap\mathbf{N}^{-1}[\omega_p]} \frac{|I_p|^{-1}}{\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{10n}}\,d\mathbf{x}\right)\left(\sup_{\mathbf{x}\in J}\frac{1}{\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{10n}}\right)\\
&\le C\epsilon\mu \sum_{J\in\mathcal{J}} \sum_{\substack{p\in\mathbf{T}\\|I_p|\le 2^n|J|}} |I_p| \sup_{\mathbf{x}\in J}\frac{1}{\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{10n}}\\
&\le C\epsilon\mu \sum_{J\in\mathcal{J}} \sum_{k\,:\,2^{kn}\le 2^n|J|} 2^{kn} \sum_{\substack{p\in\mathbf{T}\\|I_p|=2^{kn}}} \frac{1}{\left(1+\frac{\mathrm{dist}(J,I_p)}{2^k}\right)^{5n}}\frac{1}{\left(1+\frac{\mathrm{dist}(J,I_p)}{2^k}\right)^{5n}},
\end{aligned}$$

where we have used that for $\mathbf{x}\in J$,

$$|\mathbf{x}-c(I_p)| \ge \mathrm{dist}(J,c(I_p)) \ge \mathrm{dist}(J,I_p)+\frac{2^k}{2},$$

hence

$$\frac{|\mathbf{x}-c(I_p)|}{2^k} \ge \frac{\mathrm{dist}(J,c(I_p))}{2^k} \ge \frac{\mathrm{dist}(J,I_p)}{2^k}+\frac{1}{2}.$$

For all $p\in\mathbf{T}$ with $|I_p|=2^{kn}$, the I_p are pairwise disjoint and contained in $I_{\mathbf{T}}$. Therefore $\mathrm{dist}(J,I_p)\ge\mathrm{dist}(J,I_{\mathbf{T}})$ and $|I_{\mathbf{T}}|^{-\frac{1}{n}}\le 2^{-k}$, which gives

$$\frac{1}{\left(1+\frac{\mathrm{dist}(J,I_p)}{2^k}\right)^{5n}} \le \frac{1}{\left(1+\frac{\mathrm{dist}(J,I_{\mathbf{T}})}{|I_{\mathbf{T}}|^{\frac{1}{n}}}\right)^{5n}}. \tag{6.4}$$

We also need the following lemma.

Lemma 5. *For $J\in\mathcal{J}$ such that $2^{kn}\le 2^n|J|$,*

$$\sum_{p\in\mathbf{T}}\frac{1}{\left(1+\frac{\mathrm{dist}(J,I_p)}{2^k}\right)^{5n}} \le C(n),$$

where $C(n)$ is independent of J,k and $\mathbf{T}$.

Proof : We first observe that $\mathrm{dist}(J,I_p)$ and $\mathrm{dist}(c(J),I_p)$ are of comparable size. The inequality $\mathrm{dist}(J,I_p)\le\mathrm{dist}(c(J),I_p)$ is clear. To see the other inequality, note that $|I_p|\le 2^n|J|=|2J|$ implies that I_p is disjoint from $3J$, since $3J$ does not contain any I_p. Thus we have

$$\begin{aligned}\operatorname{dist}(I_p, c(J)) &\le \operatorname{dist}(I_p, J) + \operatorname{dist}(\partial J, c(J)) \\ &\le \operatorname{dist}(I_p, J) + \frac{\sqrt{n}}{2}\operatorname{dist}(I_p, J) \\ &\left(1+\frac{\sqrt{n}}{2}\right)\operatorname{dist}(I_p, J).\end{aligned}$$

Hence it suffices to replace $\operatorname{dist}(J, I_p)$ by $\operatorname{dist}(c(J), I_p)$. Let $\mathbf{x}_0 = c(J)$ and decompose $\mathbb{R}^n$ as follows :

$$\mathbb{R}^n = \bigcup_{m=1}^{\infty} \mathcal{O}_m,$$

where

$$\begin{aligned}\mathcal{O}_1 &:= B(\mathbf{x}_0, 3\sqrt{n}2^k), \\ \mathcal{O}_m &:= B(\mathbf{x}_0, 3m\sqrt{n}2^k)\backslash B(\mathbf{x}_0, 3(m-1)\sqrt{n}2^k).\end{aligned}$$

Let

$$\begin{aligned}S_1 &:= \left\{p \in \mathbf{T} \,:\, |I_p| = 2^{kn}, I_p \cap B\left(\mathbf{x}_0, 3\sqrt{n}2^k\right) \neq \emptyset\right\}, \\ S_m &:= \left\{p \in \mathbf{T} \,:\, |I_p| = 2^{kn}, I_p \cap \mathcal{O}_m \neq \emptyset, I_p \cap \left(\cup_{i=1}^{m-1}\mathcal{O}_i\right) = \emptyset\right\}, \quad m \ge 2.\end{aligned}$$

Since the diameter of I_p is $\sqrt{n}2^k$, I_p will not intersect three annuli, so each p in the sum is contained in exactly one S_m. In order to estimate the number of tiles in S_m, we consider the volume of the corresponding annulus $\mathcal{O}_m$. Now,

$$\operatorname{volume}(\mathcal{O}_m) = (3\sqrt{n})^n 2^{kn}\left(m^n - (m-1)^n\right) = C_n 2^{kn} m^{n-1}.$$

Since the I_p-s are disjoint, there are $C_n m^{n-1}$ cubes I_p of size 2^{kn} in the set S_m. Also for $p \in S_m$,

$$3(m-1)\sqrt{n}2^k \le \operatorname{dist}(\mathbf{x}_0, I_p) \le 3m\sqrt{n}2^k.$$

Thus,

$$\sum_{\substack{s\in\mathbf{T} \\ |I_p|=2^{kn}}} \frac{1}{\left(1+\frac{\operatorname{dist}(\mathbf{x}_0,I_p)}{2^k}\right)^{5n}} \le \sum_{m=1}^{\infty} \frac{m^{n-1}}{(1+m)^{5n}} \le \sum_{m=1}^{\infty} \frac{1}{m^6} < \infty.$$

$\square$

Using (6.4) and the lemma, we have that $\mathcal{K}_1$ is bounded by

$$\begin{aligned} & C\epsilon\mu \sum_{J\in\mathcal{J}} \sum_{kn=-\infty}^{\log_2 2^n|J|} \frac{2^{kn}}{\left(1+\frac{\text{dist}(J,I_{\mathbf{T}})}{|I_{\mathbf{T}}|^{\frac{1}{n}}}\right)^{5n}} \\ \le & C\epsilon\mu \sum_{J\in\mathcal{J}} \frac{|J|}{\left(1+\frac{\text{dist}(J,I_{\mathbf{T}})}{|I_{\mathbf{T}}|^{\frac{1}{n}}}\right)^{5n}} \\ \le & C\epsilon\mu \sum_{J\in\mathcal{J}} \int_J \frac{d\mathbf{x}}{\left(1+\frac{|\mathbf{x}-c(I_{\mathbf{T}})|}{|I_{\mathbf{T}}|^{\frac{1}{n}}}\right)^{5n}} \\ \le & C\epsilon\mu|I_{\mathbf{T}}|. \end{aligned}$$

This completes the estimate of $\mathcal{K}_1$.

Now we consider $\mathcal{K}_2$ defined by (6.3). We can assume that the summation runs only over those $J \in \mathcal{J}$ for which there exists a $p \in \mathbf{T}$ with $2^n|J| < |I_p|$. Then we have $J \subset 3I_{\mathbf{T}}$ and $2^n|J| < |I_{\mathbf{T}}|$ for all J occurring in the sum.

Let us fix a dyadic cube $J \in \mathcal{J}$ and observe that the set

$$G_J = J \cap \bigcup_{p\in\mathbf{T}\,:\,|I_p|>2^n|J|} E_{2p}$$

has measure at most $C\mu|J|$. To see this, let J' be the unique dyadic cube which contains J and $|J'| = 2^n|J| < |I_{\mathbf{T}}|$. By maximality of J, $3J'$ contains I_{p_0} for some $p_0 \in \mathbf{T}$. There are two cases to consider. Case (a): I_{p_0} is the dyadic cube that is formed from taking the unique double of each side of J' which is also dyadic. In this case $|I_{p_0}| = 2^n|J'|$ and we set $p_0 = p' < I_{\mathbf{T}} \times \omega_{\mathbf{T}}$. Case (b): I_{p_0} is contained in one of the dyadic cubes of size $|J'|$ contained in $3J'$. Since $|J'| = 2^n|J| < |I_{\mathbf{T}}|$, the dyadic cube which contains I_{p_0} is contained in $I_{\mathbf{T}}$. In this case there exists a tile p' with $|I_{p'}| = |J'|$ so that $I_{p_0} \subset I_{p'} \subset I_{\mathbf{T}}$. In both cases we have a tile p' such that $p_o < p' < I_{\mathbf{T}} \times \omega_{\mathbf{T}}$ and $|\omega_{p'}|$ is either $2^{-n}|J|^{-1}$ or $2^{-2n}|J|^{-1}$. We claim that

$$\bigcup_{p\in\mathbf{T}\,:\,|I_p|>2^n|J|} E_{2p} \subset E \cap \mathbf{N}^{-1}[\omega_{p'}].$$

To see this, let us choose $p \in \mathbf{T}$ such that $|I_p| > 2^n|J|$. Then $|\omega_p| < 2^{-n}|J|^{-1}$, which means $|\omega_p| < |\omega_{p'}|$. But $\omega_{\mathbf{T}} \subset \omega_p \cap \omega_{p'}$, which leads us to conclude $\omega_p \subset \omega_{p'}$. Recalling that

$$E_{2p} = \left\{\mathbf{x} : \mathbf{N}(\mathbf{x}) \in \omega_{p(2^n)}\right\} \cap E$$

now completes the proof of the claim.
The above claim implies that $G_J \subset J' \cap E \cap \mathbf{N}^{-1}[\omega_{p'}]$. Therefore,

$$|G_J| \leq |J' \cap E \cap \mathbf{N}^{-1}\left[\omega_{p'}\right]| = \int\limits_{E\cap \mathbf{N}^{-1}\left[\omega_{p'}\right]} 1_{J'}(\mathbf{x})\, d\mathbf{x}.$$

Since

$$1_{J'}(\mathbf{x}) \leq C\left(1+\frac{|\mathbf{x}-c(I_{p'})|}{|I_{p'}|^{\frac{1}{n}}}\right)^{-\nu},$$

and $\text{mass}(\{p\}) \leq \mu$, we get $|G_J| \leq C\mu|J|$.

Let $\mathbf{T}_2$ be the 2^n-tree of all $p \in \mathbf{T}$ such that $\omega_{\mathbf{T}(2^n)} \subset \omega_{p(2^n)}$ and let $\mathbf{T}_1 = \mathbf{T}\backslash\mathbf{T}_2$. Define for $j = 1, 2$,

$$F_{jJ} := \sum_{p\in\mathbf{T}_j\,:\,|I_P|>2^n|J|} \alpha_P \langle f, \phi_p\rangle \psi_p^{\mathbf{N}} 1_{E_{2p}}.$$

First we consider F_{1J}. We have

$$\begin{aligned}|F_{1J}(\mathbf{x})| &\leq \sum_{p\in\mathbf{T}_1\,:\,|I_p|>2^n|J|} |\langle f, \phi_p\rangle||\psi_p^{\mathbf{N}}(\mathbf{x})|1_{E_{2p}}(\mathbf{x}) \\ &\leq C\epsilon \sum_{p\in\mathbf{T}_1\,:\,|I_p|>2^n|J|} \left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{-\nu} 1_{E_{2p}}(\mathbf{x}).\end{aligned}$$

We will sum the expression on the right hand side of the above inequality in two steps. First let us construct

$\mathcal{I} := \left\{\omega\,:\,\text{there exists } p \in \mathbf{T_1} \text{ such that } |I_p| > 2^n|J| \text{ and } \omega_{p(2^n)} = \omega\right\}$, and $\mathcal{P}_\omega := \left\{p \in \mathbf{T}_1\,:\,|I_p| > 2^n|J|, \omega_{p(2^n)} = \omega\right\}$, for $\omega \in \mathcal{I}$.

This means that the sum estimating F_{1J} may be written as

$$\sum_{\omega\in\mathcal{I}}\sum_{p\in\mathcal{P}_\omega} \left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{-\nu} 1_{E_\omega}(\mathbf{x}),$$

where

$$E_\omega := E \cap \{\mathbf{x}\,:\,\mathbf{N}(\mathbf{x}) \in \omega\}.$$

Now note that for $p \in \mathbf{T}_1$, the semitiles $I_p \times \omega_{p(2^n)}$ are disjoint. In particular, for $p, p' \in \mathcal{P}_\omega$, $p \neq p'$, one has $I_p \cap I_{p'} = \emptyset$. Therefore,

$$\sum_{p\in\mathcal{P}_\omega} \left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{-\nu} \leq C.$$

The proof of this fact is similar to that of Lemma 5. This implies that

$$|F_{1J}(\mathbf{x})| \leq C\epsilon \sum_{\omega\in\mathcal{I}} 1_{E_\omega}(\mathbf{x}) = C\epsilon 1_{\bigcup E_\omega}(\mathbf{x}).$$

Here we have used the fact that the ω-s in $\mathcal{I}$ are disjoint. This yields

$$||F_{1J}(\mathbf{x})||_{L^1(J)} \leq C\epsilon \int_J 1_{\bigcup E_\omega}(\mathbf{x})\, d\mathbf{x} = C\epsilon |G_J| \leq C\epsilon\mu |J|,$$

This estimate, summed over the disjoint $J \subset 3I_{\mathbf{T}}$ yields the desired bound.

To complete the proof of (3.5) we estimate $F_{2J}(\mathbf{x})$. Fix $\mathbf{x}$ and assume that $F_{2J}(\mathbf{x})$ is not zero. Since the cubes $\omega_{p(2^n)}$ with $p \in \mathbf{T}_2$ are all nested and $E_{2p} = \left\{\mathbf{x} : \mathbf{N}(\mathbf{x}) \in \omega_{p(2^n)}\right\} \cap E$, there is a largest cube ω_+ of the form ω_p with $p \in \mathbf{T}_2$, $\mathbf{x} \in E_{2p}$ and $|I_p| > 2^n |J|$. Similarly there is a smallest cube which we call ω_s satisfying the above properties. Let us define $\omega_- = \omega_{s(2^n)}$. Then $\mathbf{x} \in E_{2p}$ for some $p \in \mathbf{T}$ with $|I_p| > 2^n|J|$ if and only if $|\omega_-| < |\omega_p| \leq |\omega_+|$. Fix $\boldsymbol{\xi}_0 \in \omega_{\mathbf{T}}$. We can now write $F_{2J}(\mathbf{x})$ as

$$F_{2J}(\mathbf{x}) = \sum_{\substack{p \in \mathbf{T}_2 \\ |\omega_-| < |\omega_p| \leq |\omega_+|}} \alpha_p \langle f, \phi_p \rangle \psi_p^{\mathbf{N}}(\mathbf{x}),$$

which may be decomposed as

$$\sum_{\substack{p \in \mathbf{T}_2 \\ |\omega_-| < |\omega_p| \leq |\omega_+|}} \alpha_p \langle f, \phi_p \rangle \left(e^{2\pi i \boldsymbol{\xi}_0 \cdot (\cdot)} K(\cdot) * \phi_p(\cdot)\right)(\mathbf{x}) +$$

$$\left(\left[\left(e^{2\pi i \mathbf{N}(\mathbf{x}) \cdot (\cdot)} - e^{2\pi i \boldsymbol{\xi}_0 \cdot (\cdot)} \right) K(\cdot) \right] * \sum_{\substack{p \in \mathbf{T}_2 \\ |\omega_-| < |\omega_p| \leq |\omega_+|}} \alpha_p \langle f, \phi_p \rangle \phi_p(\cdot) \right)(\mathbf{x})$$

$$= \sum_{p \in \mathbf{T}_2} \alpha_p \langle f, \phi_p \rangle \left(\psi_p^{\boldsymbol{\xi}_0} * \left(M_{c(\omega_+)} D^1_{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}} \phi - M_{c(\omega_-)} D^1_{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}} \phi \right) \right)(\mathbf{x})$$
$$+ \Bigg\{ \left[\left(e^{2\pi i \mathbf{N}(\mathbf{x}) \cdot (\cdot)} - e^{2\pi i \boldsymbol{\xi}_0 \cdot (\cdot)} \right) K(\cdot) \right] *$$
$$\Bigg[\sum_{p \in \mathbf{T}_2} \alpha_p \langle f, \phi_p \rangle \left(\phi_p * \left(M_{c(\omega_+)} D^1_{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}} \phi - M_{c(\omega_-)} D^1_{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}} \phi \right) \right) \Bigg] \Bigg\}(\mathbf{x}).$$

We claim that the last equality follows from the geometry of the supports of the Fourier transforms of the two convolving functions. More specifically, $\widehat{\phi_p}$ is supported on $\frac{1}{5}\omega_{p(1)}$ while

$$\left(M_{c(\omega_\pm)} D^1_{\frac{1}{6}|\omega_\pm|^{-\frac{1}{n}}} \phi \right)\widehat{\ }(\boldsymbol{\xi}) = \begin{cases} 1 & \text{if } \boldsymbol{\xi} \in \omega_\pm \\ 0 & \text{if } \boldsymbol{\xi} \notin \left(1 + \frac{1}{5}\right) \omega_\pm. \end{cases}$$

Therefore

$$\left(M_{c(\omega_+)} D^1_{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}} \phi - M_{c(\omega_-)} D^1_{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}} \phi \right)\widehat{\ }(\boldsymbol{\xi}) = \begin{cases} 1 & \text{if } \boldsymbol{\xi} \in \omega_+ \setminus \left(1 + \frac{1}{5}\right) \omega_- \\ 0 & \text{if } \boldsymbol{\xi} \in \omega_- \cup \left(\left(1 + \frac{1}{5}\right) \omega_+\right)^c. \end{cases}$$

For those $p \in \mathbf{T}_2$ such that $|\omega_-| < |\omega_p| \leq |\omega_+|$, we have $\frac{1}{5}\omega_{p(1)} \subset \omega_+ \setminus \left(1+\frac{1}{5}\right)\omega_-$. Conversely, for $p \in \mathbf{T}_2$ such that $|\omega_p| > |\omega_+|$ or $|\omega_p| \leq |\omega_-|$, we have $\frac{1}{5}\omega_{p(1)} \subset \omega_- \cup \left(\left(1+\frac{1}{5}\right)\omega_+\right)^c$. This concludes the proof of the claim.

The expression for F_{2J} may therefore be written as

$$(6.5) \quad \left(G_1 * \left(M_{c(\omega_+)} D^1_{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\phi - M_{c(\omega_-)} D^1_{\frac{1}{6}\omega_-|^{-\frac{1}{n}}}\phi\right)\right)(\mathbf{x}) +$$
$$\left[\left(e^{2\pi i \mathbf{N}(\mathbf{x})\cdot(\cdot)} - e^{2\pi i \boldsymbol{\xi}_0\cdot(\cdot)}\right) K(\cdot) * G_2(\cdot) * \left(M_{c(\omega_+)} D^1_{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\phi - M_{c(\omega_-)} D^1_{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}}\phi\right)\right](\mathbf{x}),$$

where

$$\begin{aligned} G_1(\mathbf{x}) &= \sum_{p\in\mathbf{T}_2} \alpha_p \langle f, \phi_p\rangle \psi_p^{\boldsymbol{\xi}_0}(\mathbf{x}) \\ &= \sum_{p\in\mathbf{T}_2} \alpha_p \langle f, \phi_p\rangle \left(e^{2\pi i \boldsymbol{\xi}_0\cdot(\cdot)} K(\cdot) * \phi_p(\cdot)\right)(\mathbf{x}), \\ G_2(\mathbf{x}) &= \sum_{p\in\mathbf{T}_2} \alpha_p \langle f, \phi_p\rangle \phi_p. \end{aligned}$$

Claim :

$$(6.6) \qquad |F_{2J}(\mathbf{x})| \leq C\left(\sup_{J\subset I} \frac{1}{|I|}\int_I |G_1(\mathbf{z})|\, d\mathbf{z} + \sup_{J\subset I} \frac{1}{|I|}\int_I |G_2(\mathbf{z})|\, d\mathbf{z}\right),$$

where the suprema above are taken over all cubes I containing J.
The proof of the claim is given in the next section. One should recognize the claim as a slight variant of the classical inequality

$$T^*\bar{f} \lesssim M(T\bar{f}) + M(\bar{f}),$$

where

$$Tg = \left(e^{2\pi i \boldsymbol{\xi}_0\cdot(\cdot)} K(\cdot)\right) * g, \quad \bar{f} = G_2(x),$$

T^* is the maximal operator corresponding to T, and M denotes the Hardy-Littlewood maximal function. In the rest of this section we show how the proof of Theorem 1 may be completed using (6.6).

We observe that the right hand side of the above expression is constant on J and that $F_{2J}1_J$ is supported on G_J, which is of measure less than or equal to $C\mu|J|$. Hence,

$$\begin{aligned}
&\sum_{J\in\mathcal{J}} ||F_{2J}||_{L^1(J)} \\
\leq& C\mu \sum_{J\in\mathcal{J}\,:\,J\subset 3I_{\boldsymbol{T}}} |J| \left(\sup_{J\subset I}\frac{1}{|I|}\int_I |G_1(\mathbf{z})|\,d\mathbf{z} + \sup_{J\subset I}\frac{1}{|I|}\int_I |G_2(\mathbf{z})|\,d\mathbf{z}\right) \\
\leq& C\mu \left(\left|\left|M\left(\sum_{p\in\mathbf{T_2}}\alpha_p\langle f,\phi_p\rangle\psi_p^{\boldsymbol{\xi}_0}\right)\right|\right|_{L^1(3I_{\mathbf{T}})} + \left|\left|M\left(\sum_{p\in\mathbf{T_2}}\alpha_p\langle f,\phi_p\rangle\phi_p\right)\right|\right|_{L^1(3I_{\mathbf{T}})}\right) \\
\leq& C\mu|I_{\mathbf{T}}|^{\frac{1}{2}} \left(\left|\left|\sum_{p\in\mathbf{T_2}}\alpha_p\langle f,\phi_p\rangle\psi_p^{\boldsymbol{\xi}_0}\right|\right|_{L^2(\mathbb{R}^n)} + \left|\left|\sum_{p\in\mathbf{T_2}}\alpha_p\langle f,\phi_p\rangle\phi_p\right|\right|_{L^2(\mathbb{R}^n)}\right).
\end{aligned}$$

Here we have used the L^2 boundedness of the Hardy-Littlewood maximal function M. We would now like to show that $||G_1||_{L^2}$ and $||G_2||_{L^2}$ are bounded above by a constant multiple of $\epsilon|I_T|^{\frac{1}{2}}$.

$$\begin{aligned}
||G_1||_{L^2}^2 &= \sum_{p,p'\in\mathbf{T_2}} \alpha_p\alpha_{p'}\langle f,\phi_p\rangle\,\overline{\langle f,\phi_{p'}\rangle}\,\langle\psi_p^{\boldsymbol{\xi}_0},\psi_{p'}^{\boldsymbol{\xi}_0}\rangle \\
&= \sum_{\substack{p,p'\in\mathbf{T_2}\\ \omega_p\neq\omega_p'}} \alpha_p\alpha_{p'}\langle f,\phi_p\rangle\,\overline{\langle f,\phi_{p'}\rangle}\,\langle m(\cdot-\boldsymbol{\xi}_0)\widehat{\phi_p}, m(\cdot-\boldsymbol{\xi}_0)\widehat{\phi_{p'}}\rangle \\
&\qquad + \sum_{\substack{p,p'\in\mathbf{T_2}\\ \omega_p=\omega_p'}} \alpha_p\alpha_{p'}\langle f,\phi_p\rangle\,\overline{\langle f,\phi_{p'}\rangle}\,\langle\psi_p^{\boldsymbol{\xi}_0},\psi_{p'}^{\boldsymbol{\xi}_0}\rangle.
\end{aligned}$$

Similarly,

$$\begin{aligned}
||G_2||_{L^2}^2 &= \sum_{p,p'\in\mathbf{T_2}} \alpha_p\alpha_{p'}\langle f,\phi_p\rangle\,\overline{\langle f,\phi_{p'}\rangle}\,\langle\phi_p,\phi_{p'}\rangle \\
&= \sum_{\substack{p,p'\in\mathbf{T_2}\\ \omega_p\neq\omega_p'}} \alpha_p\alpha_{p'}\langle f,\phi_p\rangle\,\overline{\langle f,\phi_{p'}\rangle}\,\langle\widehat{\phi_p},\widehat{\phi_{p'}}\rangle \\
&\qquad + \sum_{\substack{p,p'\in\mathbf{T_2}\\ \omega_p=\omega_p'}} \alpha_p\alpha_{p'}\langle f,\phi_p\rangle\,\overline{\langle f,\phi_{p'}\rangle}\,\langle\phi_p,\phi_{p'}\rangle.
\end{aligned}$$

But $\langle\widehat{\phi_p},\widehat{\phi_{p'}}\rangle = 0$ for $p,p'\in\mathbf{T}_2, \omega_p\neq\omega_{p'}$, since $\widehat{\phi_p}$ and $\widehat{\phi_{p'}}$ have disjoint supports in this case. Therefore we only need to consider the second sum in the expressions for $||G_1||_{L^2}^2$ and $||G_2||_{L^2}^2$. Our pointwise bounds imply that

$$\langle\phi_p,\phi_{p'}\rangle, \langle\psi_p^{\boldsymbol{\xi}_0},\psi_{p'}^{\boldsymbol{\xi}_0}\rangle \lesssim |I_p|^{-\frac{1}{2}}|I_{p'}|^{-\frac{1}{2}}\int\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{-\nu}\left(1+\frac{|\mathbf{x}-c(I_{p'})|}{|I_{p'}|^{\frac{1}{n}}}\right)^{-\nu}d\mathbf{x},$$

so it is enough to estimate the right hand side above for p, p' satisfying $\omega_p = \omega'_p$. Upon simplification this reduces to

$$
\begin{aligned}
&|I_p|^{-1}\int_{\mathbb{R}^n}\left(1+\frac{|\mathbf{x}-c(I_p)|}{|I_p|^{\frac{1}{n}}}\right)^{-\nu}\left(1+\frac{|\mathbf{x}-c(I_{p'})|}{|I_{p'}|^{\frac{1}{n}}}\right)^{-\nu}d\mathbf{x}\\
&=\int_{\mathbb{R}^n}(1+|\mathbf{y}|)^{-\nu}\left(1+\left|\mathbf{y}+\frac{c(I_p)-c(I_p)}{|I_p|^{\frac{1}{n}}}\right|\right)^{-\nu}d\mathbf{y}\\
&\lesssim\left(1+\left|\frac{c(I_p)-c(I_{p'})}{|I_p|^{\frac{1}{n}}}\right|\right)^{-\nu}.
\end{aligned}
$$

With this estimate, the proof that $||G_1||_{L^2}$ and $||G_2||_{L^2}$ are less than $C\epsilon|I_{\mathbf{T}}|^{\frac{1}{2}}$ is similar to an argument outlined in the proof of Lemma 2. One needs to follow the proof of the estimate of the first term of (5.5) to complete the proof of Lemma 3, given the claim.

7. Proof of Claim (6.6)

Let us estimate the first term in the expression (6.5). We denote by $\tilde{c}$ any one of the two constants $\frac{1}{6}|\omega_+|^{-\frac{1}{n}}$ or $\frac{1}{6}|\omega_-|^{-\frac{1}{n}}$. By translation invariance, let I_- and I_+ be the unique dyadic cubes of the form

$$
I_{\pm}=\prod_{j=1}^{n}[0,|\omega_{\pm}|^{-\frac{1}{n}}).
$$

For a dyadic cube $I=\prod_{j=1}^{n}[0,2^k)$ and $\mathbf{r}=(r_1,r_2,\cdots,r_n)\in\mathbb{Z}^n$, $I+\mathbf{r}$ will denote the unique dyadic cube of the form

$$
[r_1 2^k,(r_1+1)2^k)\times\cdots\times[r_n 2^k,(r_n+1)2^k).
$$

Now,

$$
\begin{aligned}
&\int|G_1(\mathbf{y})|\tilde{c}^{-n}\left|\phi\left(\frac{\mathbf{x}-\mathbf{y}}{\tilde{c}}\right)\right|d\mathbf{y}\\
\leq&\tilde{c}^{-n}\int\limits_{|\mathbf{x}-\mathbf{y}|\leq\tilde{c}}|G_1(\mathbf{y})|\,d\mathbf{y}+\tilde{c}^{-n}\sum_{j\geq1}2^{-j\nu}\int\limits_{|\mathbf{x}-\mathbf{y}|\sim\tilde{c}2^j}|G_1(\mathbf{y})|\,d\mathbf{y}\\
\leq&\tilde{c}^{-n}\int\limits_{\substack{|x_i-y_i|\leq\tilde{c}\\1\leq i\leq n}}|G_1(\mathbf{y})|\,d\mathbf{y}+\sum_{j\geq1}2^{-j}\frac{1}{(\tilde{c}2^j)^n}\int\limits_{\substack{|x_i-y_i|\leq\tilde{c}2^j\\1\leq i\leq n}}|G_1(\mathbf{y})|\,d\mathbf{y}\\
\lesssim&\sup_{I\,:\,J\subset I}\frac{1}{|I|}\int_I|G_1(\mathbf{z})|\,d\mathbf{z}.
\end{aligned}
$$

Here we have used the fact that since $\mathbf{x} \in J$ and $|J| < |I_+| < |I_-|$, we have $J \subset \{\mathbf{y} : |x_i - y_i| \leq \tilde{c}$ for all $i, 1 \leq i \leq n\}$.

We denote by $\mathcal{B}$ the second term in (6.5).

$$\mathcal{B} = \int_{(\mathbf{y},\mathbf{z}) \in \mathbb{R}^n \times \mathbb{R}^n} \left(e^{2\pi i \mathbf{N}(\mathbf{x}) \cdot \mathbf{y}} - e^{2\pi i \boldsymbol{\xi}_0 \cdot \mathbf{y}}\right) K(\mathbf{y}) \times \left[e^{2\pi i c(\omega_+) \cdot (\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}\right)^{-n} \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\right) - e^{2\pi i c(\omega_-) \cdot (\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}\right)^{-n} \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}}\right)\right] G_2(\mathbf{x}-\mathbf{z})\, d\mathbf{y} d\mathbf{z}.$$

To estimate $\mathcal{B}$ we write it as

$$\mathcal{B} = \mathcal{B}_1 - \mathcal{B}_2 + \mathcal{B}_3 + \mathcal{B}_4,$$

where

$$\mathcal{B}_1 := \int_{\mathbf{z} \in I_+} \int_{\mathbf{y} \in \mathbb{R}^n} \left(e^{2\pi i \mathbf{N}(\mathbf{x}) \cdot \mathbf{y}} - e^{2\pi i \boldsymbol{\xi}_0 \cdot \mathbf{y}}\right) K(\mathbf{y}) e^{2\pi i c(\omega_+) \cdot (\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}\right)^{-n} \times \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\right) G_2(\mathbf{x}-\mathbf{z})\, d\mathbf{y}\, d\mathbf{z},$$

$$\mathcal{B}_2 := \int_{\mathbf{z} \in I_-} \int_{\mathbf{y} \in \mathbb{R}^n} \left(e^{2\pi i \mathbf{N}(\mathbf{x}) \cdot \mathbf{y}} - e^{2\pi i \boldsymbol{\xi}_0 \cdot \mathbf{y}}\right) K(\mathbf{y}) e^{2\pi i c(\omega_-) \cdot (\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}\right)^{-n} \times \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}}\right) G_2(\mathbf{x}-\mathbf{z})\, d\mathbf{y}\, d\mathbf{z},$$

$$\mathcal{B}_3 := \sum_{\mathbf{r} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} \int_{\mathbf{z} \in I_- + \mathbf{r}} \int_{\mathbf{y} \in \mathbb{R}^n} \left(e^{2\pi i \mathbf{N}(\mathbf{x}) \cdot \mathbf{y}} - e^{2\pi i \boldsymbol{\xi}_0 \cdot \mathbf{y}}\right) K(\mathbf{y}) \times \left[e^{2\pi i c(\omega_+) \cdot (\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}\right)^{-n} \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\right) - e^{2\pi i c(\omega_-) \cdot (\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}\right)^{-n} \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_-|^{-\frac{1}{n}}}\right)\right] G_2(\mathbf{x}-\mathbf{z})\, d\mathbf{y} d\mathbf{z},$$

and

$$\mathcal{B}_4 := \int_{\mathbf{z}\in I_-\setminus I_+}\int_{\mathbf{y}\in\mathbb{R}^n}\left(e^{2\pi i\mathbf{N}(\mathbf{x})\cdot\mathbf{y}} - e^{2\pi i\boldsymbol{\xi}_0\cdot\mathbf{y}}\right)K(\mathbf{y})e^{2\pi i c(\omega_+)\cdot(\mathbf{z}-\mathbf{y})}\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}\right)^{-n}\times$$
$$\phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\right)G_2(\mathbf{x}-\mathbf{z})\,d\mathbf{y}\,d\mathbf{z}.$$

We have the following bound for $\mathcal{B}_1$:

$$|\mathcal{B}_1| \le \left[\int_{\mathbf{z}\in I_+}\int_{\mathbf{y}\in I_+} + \sum_{\mathbf{m}\in\mathbb{Z}^n\setminus\{0\}}\int_{\mathbf{z}\in I_+}\int_{\mathbf{y}\in I_++\mathbf{m}}\right]$$
$$\left|\left(e^{2\pi i\mathbf{N}(\mathbf{x})\cdot\mathbf{y}} - e^{2\pi i\boldsymbol{\xi}_0\cdot\mathbf{y}}\right)K(\mathbf{y})\right||\omega_+|\left|\phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}\,|\omega_+|^{-\frac{1}{n}}}\right)\right|\,d\mathbf{y}|G_2(\mathbf{x}-\mathbf{z})|\,d\mathbf{z}$$
$$\lesssim \int_{\mathbf{z}\in I_+}|\omega_+|\int_{\mathbf{y}\in I_+}|\mathbf{N}(\mathbf{x})-\boldsymbol{\xi}_0||\mathbf{y}|^{1-n}d\mathbf{y}|G_2(\mathbf{x}-\mathbf{z})|\,d\mathbf{z} \quad +$$
$$\sum_{\mathbf{m}\in\mathbb{Z}^n\setminus\{0\}}\int_{\mathbf{z}\in I_+}|\omega_+|\int_{\mathbf{y}\in I_++\mathbf{m}}|\mathbf{y}|^{-n}|\mathbf{m}|^{-\nu}\,d\mathbf{y}|G_2(\mathbf{x}-\mathbf{z})|\,d\mathbf{z}$$
$$\lesssim |\omega_+||\omega_-|^{\frac{1}{n}}|I_+|^{\frac{1}{n}}\int_{\mathbf{z}\in I_+}|G_2(\mathbf{x}-\mathbf{z})|\,d\mathbf{z} \quad +$$
$$\sum_{\mathbf{m}\in\mathbb{Z}^n\setminus\{0\}}|\omega_+|\left(|\mathbf{m}|\,|I_+|^{\frac{1}{n}}\right)^{-n}|\mathbf{m}|^{-\nu}|I_+|\int_{\mathbf{z}\in I_+}|G_2(\mathbf{x}-\mathbf{z})|\,d\mathbf{z}$$
$$\lesssim \left[\left(\frac{|\omega_-|}{|\omega_+|}\right)^{\frac{1}{n}} + \sum_{\mathbf{m}\in\mathbb{Z}^n\setminus\{0\}}|\mathbf{m}|^{-n-\nu}\right]\sup_{I\,:\,J\subset I}\frac{1}{|I|}\int_I|G_2(\mathbf{z})|\,d\mathbf{z}$$
$$\lesssim \sup_{I\,:\,J\subset I}\frac{1}{|I|}\int_I|G_2(\mathbf{z})|\,d\mathbf{z},$$

for large ν. The treatment for $\mathcal{B}_2$ is similar.

Next we estimate $\mathcal{B}_3$. Note that it suffices to consider $|\mathbf{r}| \ge 2$. The case $|\mathbf{r}| < \mathbf{2}$ follows with a similar argument as in the treatment of $\mathcal{B}_1$.

$$\mathcal{B}_3 = \sum_{\mathbf{r}\in\mathbb{Z}^n\setminus\{0\}}\int_{\mathbf{z}\in I_-+\mathbf{r}}\int_{\boldsymbol{\xi}\in\mathbb{R}^n}\left(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0)\right)e^{2\pi i\mathbf{z}\cdot\boldsymbol{\xi}}\times$$
$$\left[\widehat{\phi}\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_+))\right) - \widehat{\phi}\left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_-))\right)\right]d\boldsymbol{\xi}|G_2(\mathbf{x}-\mathbf{z})|\,d\mathbf{z}.$$

Let us write

$$e^{2\pi i\mathbf{z}\cdot\boldsymbol{\xi}} = |\mathbf{z}|^{-2N}(L_{\boldsymbol{\xi}})^N\left(e^{2\pi i\mathbf{z}\cdot\boldsymbol{\xi}}\right),$$

where N is a large positive integer and $L_{\boldsymbol{\xi}}$ is a suitably chosen differential operator. Then,

$$\mathcal{B}_3 = \sum_{\mathbf{r}\in\mathbb{Z}^n\setminus\{0\}} \int_{\mathbf{z}\in I_-+\mathbf{r}} |\mathbf{z}|^{-2N} \int_{\boldsymbol{\xi}\in\mathbb{R}^n} e^{2\pi i\mathbf{z}\cdot\boldsymbol{\xi}} L_{\boldsymbol{\xi}}^N \Big[(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0)) \times$$
$$\left(\widehat{\phi}\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_+)) \right) - \widehat{\phi}\left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_-)) \right) \right) \Big] \, d\boldsymbol{\xi} |G_2(\mathbf{x}-\mathbf{z})| \, d\mathbf{z}.$$

For simplicity, let us consider only those terms where $(L_{\boldsymbol{\xi}})^N$ is applied to either one of the terms

$$(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0)) \text{ or}$$
$$\left(\widehat{\phi}\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_+)) \right) - \widehat{\phi}\left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_-)) \right) \right).$$

The analysis for the other terms is similar. First let us look at the inner integral

$$\int_{\boldsymbol{\xi}\in\mathbb{R}^n} e^{2\pi i\mathbf{z}\cdot\boldsymbol{\xi}} L_{\boldsymbol{\xi}}^N \left[(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0)) \right] \times \left[\left(\widehat{\phi}\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_+)) \right) - \widehat{\phi}\left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_-)) \right) \right) \right] d\boldsymbol{\xi}. \tag{7.1}$$

We observe that

$$\left| L_{\boldsymbol{\xi}}^N \left[(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0)) \right] \right| \lesssim \left(|\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})|^{-2N} + |\boldsymbol{\xi}-\boldsymbol{\xi}_0|^{-2N} \right),$$

and that the integrand is supported on $(1+\frac{1}{5})\omega_+ \setminus \omega_-$. Also, given ω_+ and ω_-, there exists a unique sequence of nested intervals

$$\omega_- \subset \omega_{p_1} \subset \omega_{p_2} \subset \cdots \subset \omega_{p_M} = \omega_+, \quad |\omega_{p_{i+1}}| = 2^n |\omega_{p_i}|.$$

It is not difficult to see that if $\boldsymbol{\xi} \in (1+\frac{1}{5})\omega_{p_{i+1}} \setminus \omega_{p_i}$ and $\mathbf{N}(\mathbf{x}), \boldsymbol{\xi}_0 \in \omega_- \subset \omega_{p_i}$, then

$$|\mathbf{N}(\mathbf{x})-\boldsymbol{\xi}|, |\boldsymbol{\xi}-\boldsymbol{\xi}_0| \gtrsim |\omega_{p_i}|^{\frac{1}{n}}, \quad 1 \le i \le M-1.$$

This implies that the expression in (7.1) is bounded by a constant multiple of

$$\sum_{i=1}^{M-1} |\omega_{p_i}| |\omega_{p_i}|^{-\frac{2N}{n}} \lesssim |\omega_-|^{1-\frac{2N}{n}}.$$

Next we consider the inner integral

$$\int_{\boldsymbol{\xi}\in\mathbb{R}^n} e^{2\pi i\mathbf{z}\cdot\boldsymbol{\xi}} \left[(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0)) \right] \times L_{\boldsymbol{\xi}}^N \left[\left(\widehat{\phi}\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_+)) \right) - \widehat{\phi}\left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_-)) \right) \right) \right] d\boldsymbol{\xi} \tag{7.2}$$

$$
\begin{aligned}
= \int_{\boldsymbol{\xi}\in\mathbb{R}^n} & e^{2\pi i \mathbf{z}\cdot\boldsymbol{\xi}} \left[(m(\boldsymbol{\xi}-\mathbf{N}(\mathbf{x})) - m(\boldsymbol{\xi}-\boldsymbol{\xi}_0))\right] \times \\
& \Big[|\omega_+|^{-\frac{2N}{n}} \left(\frac{1}{6}\right)^{2N} \left(L_{\boldsymbol{\xi}}^N \widehat{\phi}\right)\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_+))\right) - \\
& \qquad |\omega_-|^{-\frac{2N}{n}} \left(\frac{1}{6}\right)^{2N} \left(L_{\boldsymbol{\xi}}^N \widehat{\phi}\right)\left(\frac{1}{6}|\omega_-|^{-\frac{1}{n}}(\boldsymbol{\xi}-c(\omega_-))\right)\Big]\, d\boldsymbol{\xi}.
\end{aligned}
$$

Therefore the expression in (7.2) is bounded above by a constant multiple of

$$
|\omega_+|^{1-\frac{2N}{n}} + |\omega_-|^{1-\frac{2N}{n}} \lesssim |\omega_-|^{1-\frac{2N}{n}}.
$$

All the other terms originating from the integration by parts yield the same bound. Also note that for $\mathbf{z} \in I_- + \mathbf{r}$, we have $|\mathbf{z}|^{-2N} \sim (|\mathbf{r}||I_-|)^{-2N}$ since $|\mathbf{r}| \geq 2$. Therefore choosing N large enough, we obtain

$$
\begin{aligned}
|\mathcal{B}_3| &\lesssim \sum_{\mathbf{r}\in\mathbb{Z}^n\setminus\{0\}} \int_{\mathbf{z}\in I_-+\mathbf{r}} |\mathbf{z}|^{-2N} |\omega_-|^{1-\frac{2N}{n}} \left|G_2(\mathbf{x}-\mathbf{z})\right|\, d\mathbf{z} \\
&\lesssim \sum_{\mathbf{r}\in\mathbb{Z}^n\setminus\{0\}} |\omega_-|^{1-\frac{2N}{n}} \left(|\mathbf{r}||I_-|^{\frac{1}{n}}\right)^{-2N} \int_{\mathbf{z}\in I_-+\mathbf{r}} \left|G_2(\mathbf{x}-\mathbf{z})\right|\, d\mathbf{z} \\
&\lesssim \sum_{\mathbf{r}\in\mathbb{Z}^n\setminus\{0\}} |\mathbf{r}|^{n-2N} \frac{1}{\left(|\mathbf{r}||I_-|^{\frac{1}{n}}\right)^n} \int_{\mathbf{z}\in C|\mathbf{r}|I_-} \left|G_2(\mathbf{x}-\mathbf{z})\right|\, d\mathbf{z} \\
&\lesssim \sup_{I\,:\,J\subset I} \frac{1}{|I|} \int_I |G_2(\mathbf{z})|\, d\mathbf{z},
\end{aligned}
$$

since $J \subset C|\mathbf{r}|I_-$.

It therefore remains to analyze $\mathcal{B}_4$. Since $\mathbf{z} \in I_- \setminus I_+$, there exists $\mathbf{r} \in \mathbb{Z}^n \setminus \{0\}$, $1 \le |\mathbf{r}| \lesssim \left(\frac{|\omega_+|}{|\omega_-|}\right)^{\frac{1}{n}}$ such that $\mathbf{z} \in I_+ + \mathbf{r}$. Now,

$$
\begin{aligned}
&\left| \int_{\mathbf{y}\in\mathbb{R}^n} \left(e^{2\pi i \mathbf{N}(\mathbf{x})\cdot\mathbf{y}} - e^{2\pi i \boldsymbol{\xi}_0\cdot\mathbf{y}}\right) K(\mathbf{y}) e^{2\pi i c(\omega_+)\cdot(\mathbf{z}-\mathbf{y})} \left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}\right)^{-n} \phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\right) d\mathbf{y} \right| \\
&\le \sum_{\mathbf{m}\in\mathbb{Z}^n} \left| \int_{\mathbf{y}\in I_+ + \mathbf{m}} \left(e^{2\pi i \mathbf{N}(\mathbf{x})\cdot\mathbf{y}} - e^{2\pi i \boldsymbol{\xi}_0\cdot\mathbf{y}}\right) K(\mathbf{y}) e^{2\pi i c(\omega_+)\cdot(\mathbf{z}-\mathbf{y})} \frac{\phi\left(\frac{\mathbf{z}-\mathbf{y}}{\frac{1}{6}|\omega_+|^{-\frac{1}{n}}}\right)}{\left(\frac{1}{6}|\omega_+|^{-\frac{1}{n}}\right)^n} d\mathbf{y} \right| \\
&\le \int_{\mathbf{y}\in I_+} |\omega_-|^{\frac{1}{n}} |\mathbf{y}|^{1-n} |\omega_+| |\mathbf{r}|^{-\nu} d\mathbf{y} \quad + \\
&\qquad \left[\left(\sum_{\mathbf{m}\neq 0\,:\, |\mathbf{m}-\mathbf{r}|\le \frac{|\mathbf{r}|}{2}} + \sum_{\mathbf{m}\neq 0\,:\, |\mathbf{m}-\mathbf{r}|> \frac{|\mathbf{r}|}{2}} \right) \left(|\mathbf{m}||I_+|^{\frac{1}{n}}\right)^{1-n} |\omega_-|^{\frac{1}{n}} \min\left(1, |\mathbf{m}-\mathbf{r}|^{-\nu}\right) \right] \\
&\le |\omega_-|^{\frac{1}{n}} |\omega_+| |I_+|^{\frac{1}{n}} |\mathbf{r}|^{-\nu} + \left(|\mathbf{r}||I_+|^{\frac{1}{n}}\right)^{1-n} |\omega_-|^{\frac{1}{n}} + \sum_{\mathbf{m}\,:\, |\mathbf{m}-\mathbf{r}|>\frac{|\mathbf{r}|}{2}} |\mathbf{m}-\mathbf{r}|^{-\nu} |\omega_-|^{\frac{1}{n}} |I_+|^{\frac{1}{n}(1-n)} \\
&\lesssim \left(|\mathbf{r}||I_+|^{\frac{1}{n}}\right)^{1-n} |\omega_-|^{\frac{1}{n}} \\
&\lesssim |\mathbf{z}|^{1-n} |\omega_-|^{\frac{1}{n}}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|\mathcal{B}_4| &\lesssim |\omega_-|^{\frac{1}{n}} \int_{|I_+|^{\frac{1}{n}} \lesssim |\mathbf{z}| \lesssim |I_-|^{\frac{1}{n}}} |\mathbf{z}|^{1-n} |G_2(\mathbf{x}-\mathbf{z})| d\mathbf{z} \\
&\lesssim g * |G_2|(\mathbf{x}),
\end{aligned}
$$

where

$$
g(\mathbf{z}) = h(|\mathbf{z}|) = |\omega_-|^{\frac{1}{n}} |\mathbf{z}|^{1-n} 1_{|I_+|^{\frac{1}{n}} \lesssim |\mathbf{z}| \lesssim |I_-|^{\frac{1}{n}}}(\mathbf{z}).
$$

We observe that $g \in L^1, ||g||_1 \le C$ and g is radially decreasing. Let us approximate g from below by g_γ defined as follows,

$$
g_\gamma(\mathbf{x}) = \begin{cases} 0 & \text{if } |\mathbf{x}| \le |I_+|^{\frac{1}{n}} \\ h\left(|I_+|^{\frac{1}{n}} + k\gamma\right) & \text{if } (k-1)\gamma + |I_+|^{\frac{1}{n}} < |\mathbf{x}| \le k\gamma + |I_+|^{\frac{1}{n}},\ 1 \le k \le k_0 \\ 0 & \text{if } |\mathbf{x}| > |I_-|^{\frac{1}{n}} \end{cases}
$$

where $k_0\gamma = |I_-|^{\frac{1}{n}} - |I_+|^{\frac{1}{n}}$.
We can write g_γ as

$$g_\gamma = -h\left(|I_+|^{\frac{1}{n}} + \gamma\right) 1_{B\left(0;|I_+|^{\frac{1}{n}}\right)} \quad + \sum_{k=1}^{k_0} \left(h\left(|I_+|^{\frac{1}{n}} + k\gamma\right) - h\left(|I_+|^{\frac{1}{n}} + (k+1)\gamma\right)\right) 1_{B\left(0;|I_+|^{\frac{1}{n}}+k\gamma\right)}.$$

Therefore,

$$|g_\gamma * |G_2|| \leq h(|I_+|^{\frac{1}{n}} + \gamma)|G_2| * 1_{B\left(0;|I_+|^{\frac{1}{n}}\right)} \quad + \sum_k \left(h\left(|I_+|^{\frac{1}{n}} + k\gamma\right) - h\left(|I_+|^{\frac{1}{n}} + (k+1)\gamma\right)\right) |G_2| * 1_{B\left(0;|I_+|^{\frac{1}{n}}+k\gamma\right)},$$

which in turn is bounded by

$$\left(\sup_{I\,:\,J\subset I} \frac{1}{|I|}\int_I |G_2(\mathbf{z})|d\mathbf{z}\right) \times \left[h\left(|I_+|^{\frac{1}{n}} + \gamma\right)\left|B\left(0;|I_+|^{\frac{1}{n}}\right)\right| \quad + \sum_k \left(h\left(|I_+|^{\frac{1}{n}} + k\gamma\right) - h\left(|I_+|^{\frac{1}{n}} + (k+1)\gamma\right)\right)\left|B\left(0;|I_+|^{\frac{1}{n}} + k\gamma\right)\right|\right]$$

$$\leq \left(\sup_{I\,:\,J\subset I} \frac{1}{|I|}\int_I |G_2(\mathbf{z})|d\mathbf{z}\right) ||\tilde{g_\gamma}||_1.$$

Here $\tilde{g_\gamma}$ is given by

$$\tilde{g_\gamma} = h\left(|I_+|^{\frac{1}{n}} + \gamma\right) 1_{B\left(0;|I_+|^{\frac{1}{n}}\right)} \quad + \sum_k \left(h\left(|I_+|^{\frac{1}{n}} + k\gamma\right) - h\left(|I_+|^{\frac{1}{n}} + (k+1)\gamma\right)\right) 1_{B\left(0;|I_+|^{\frac{1}{n}}+k\gamma\right)}.$$

In other words,

$$\tilde{g_\gamma}(\mathbf{x}) = \begin{cases} 2h\left(|I_+|^{\frac{1}{n}} + \gamma\right) & \text{if } 0 \leq |\mathbf{x}| \leq |I_+|^{\frac{1}{n}} \\ h\left(|I_+|^{\frac{1}{n}} + k\gamma\right) & \text{if } |I_+|^{\frac{1}{n}} + (k-1)\gamma < |\mathbf{x}| < |I_+|^{\frac{1}{n}} + k\gamma \\ 0 & \text{if } |\mathbf{x}| > |I_-|^{\frac{1}{n}} \end{cases}$$

Therefore,

$$||\tilde{g_\gamma}||_1 \lesssim ||g||_1 + |\omega_-|^{\frac{1}{n}} \int_{|\mathbf{y}|\leq|I_+|^{\frac{1}{n}}} |\mathbf{y}|^{1-n} d\mathbf{y} \lesssim 1,$$

since $|\omega_-| < |\omega_+|$. Thus,

$$g_\gamma * |G_2|(\mathbf{x}) \lesssim \sup_{I\,:\,J\subset I} \frac{1}{|I|}\int_I |G_2(\mathbf{z})|\, d\mathbf{z}.$$

Letting $\gamma \to 0$ and applying the dominated convergence theorem now yields the desired bound for $\mathcal{B}_4$.

8. An Application

As an immediate application of the weak L^2 mapping property of the maximal dyadic sum operator, we obtain a new proof of Sjölin's theorem [6] on a weak-type (2,2) estimate for the maximal conjugated Calderón-Zygmund operator on $\mathbb{R}^n$, $n > 1$.

Theorem 2. *Let*

$$K(\mathbf{x}) = \Omega\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right)|\mathbf{x}|^{-n}$$

be a Calderón-Zygmund kernel in $\mathbb{R}^n$ *with* $\Omega \in C^\infty(S^{n-1})$. *Let*

$$Bf = f * K,$$

and

$$\mathcal{C}f(\mathbf{x}) = \sup_{\boldsymbol{\xi}\in\mathbb{R}^n}\left|\left(e^{2\pi i\boldsymbol{\xi}\cdot(\cdot)}Be^{-2\pi i\boldsymbol{\xi}\cdot(\cdot)}f\right)\right|(\mathbf{x}).$$

Then,

$$||\mathcal{C}f||_{L^{2,\infty}} \leq C||f||_{L^2},$$

with a constant C *independent of* f.

Proof : The proof again follows techniques similar to those used by Lacey and Thiele [5] in proving Carleson's theorem on almost everywhere convergence of Fourier series. Following [5], we introduce the operators

$$A_{\boldsymbol{\eta}}f := \sum_{p\in\mathbf{D}}\langle f,\phi_p\rangle\phi_p 1_{\omega_{p(2^n)}}(\boldsymbol{\eta})$$

$$Af := \lim_{N\to\infty}\frac{1}{K_N}\int_{K_N\times[0,1]} M_{-\boldsymbol{\eta}}T_{-\mathbf{y}}D^2_{2^{-\kappa}}A_{2^{-\kappa}\boldsymbol{\eta}}D^2_{2^\kappa}T_{\mathbf{y}}M_{\boldsymbol{\eta}}f\,d\mathbf{y}\,d\boldsymbol{\eta}\,d\kappa,$$

where K_N is any increasing sequence of rectangles filling out $\mathbb{R}^n\times\mathbb{R}^n$. For any Schwartz function f and any $\mathbf{x}\in\mathbb{R}^n$, the limit representing $Af(\mathbf{x})$ exists by the argument given by Lacey and Thiele.

Note that by rotation invariance, it is enough to prove Theorem 2 when the multiplier is supported on a nonempty open cone in $\mathbb{R}^n$.

Lemma 6. *There exists a nonempty open cone* $\tilde{K}_0$ *with vertex at the origin,*

$$\tilde{K}_0 \subset \{\boldsymbol{\xi} = (\xi_1,\xi_2,\cdots,\xi_n);\xi_i \leq 0 \text{ for all } i\},$$

such that for all $\boldsymbol{\xi}\in\tilde{K}_0$,

$$(Af)\widehat{\ }(\boldsymbol{\xi}) = c\widehat{f}(\boldsymbol{\xi}),$$

where c *is a constant independent of* f.

Proof:

$$
\begin{aligned}
(Af)\widehat{\ }(\boldsymbol{\xi}) &= \lim_{N\to\infty}\frac{1}{|K_N|}\int_{((\mathbf{y},\boldsymbol{\eta}),\kappa)\in K_N\times[0,1]}\left(M_{-\boldsymbol{\eta}}T_{-\mathbf{y}}D^2_{2^{-\kappa}}A_{2^{-\kappa}\boldsymbol{\eta}}D^2_{2^{\kappa}}T_{\boldsymbol{y}}M_{\boldsymbol{\eta}}f\right)\widehat{\ }(\boldsymbol{\xi})\,d\mathbf{y}\,d\boldsymbol{\eta}\,d\kappa \\
&= \lim_{N\to\infty}\frac{1}{|K_N|}\int_{K_N\times[0,1]}\sum_{p\in\mathbf{D}}\langle f, M_{-\boldsymbol{\eta}}T_{-\mathbf{y}}D^2_{2^{-\kappa}}\phi_p\rangle\left(M_{-\boldsymbol{\eta}}T_{-\mathbf{y}}D^2_{2^{-\kappa}}\phi_p\right)\widehat{\ }(\boldsymbol{\xi})\,d\mathbf{y}\,d\boldsymbol{\eta}\,d\kappa \\
&= \lim_{N\to\infty}\frac{1}{|K_N|}\int_{K_N\times[0,1]}\sum_{p\in\mathbf{D}}\int_{\mathbb{R}^n}\widehat{f}(\boldsymbol{\xi'})2^{\frac{-\kappa n}{2}}\overline{\widehat{\phi}_p\left(2^{-\kappa}(\boldsymbol{\xi'}+\boldsymbol{\eta}\right)}e^{-2\pi i\mathbf{y}\cdot(\boldsymbol{\xi'}+\boldsymbol{\eta})}d\boldsymbol{\xi'} \\
&\qquad \times 2^{\frac{-\kappa n}{2}}\widehat{\phi_p}\left(2^{-\kappa}(\boldsymbol{\xi}+\boldsymbol{\eta})\right)e^{2\pi i\mathbf{y}\cdot(\boldsymbol{\xi}+\boldsymbol{\eta})}1_{\omega_{p(2^n)}}(2^{-\kappa}\boldsymbol{\eta})\,d\mathbf{y}\,d\boldsymbol{\eta}\,d\kappa \\
&= \lim_{N\to\infty}\frac{1}{|K_N|}\int_{K_N\times[0,1]}\sum_{p\in\mathbf{D}}\int_{\mathbb{R}^n}\widehat{f}(\boldsymbol{\xi'})2^{-\kappa n+\alpha n}\overline{\widehat{\phi}\left(2^{\alpha-\kappa}(\xi'+\boldsymbol{\eta})-\left(\mathbf{l}+\frac{1}{4}\right)\right)} \\
&\qquad \times e^{-2\pi i\left(\mathbf{m}+\frac{1}{2}\right)\cdot\left(2^{\alpha-\kappa}(\boldsymbol{\xi'}+\boldsymbol{\eta})-\left(\mathbf{l}+\frac{1}{4}\right)\right)}\widehat{\phi}\left(2^{\alpha-\kappa}(\boldsymbol{\xi}+\boldsymbol{\eta})-\left(\mathbf{l}+\frac{1}{4}\right)\right) \\
&\qquad \times e^{2\pi i\left(\mathbf{m}+\frac{1}{2}\right)\cdot\left(2^{\alpha-\kappa}(\boldsymbol{\xi}+\boldsymbol{\eta})-\left(\mathbf{l}+\frac{1}{4}\right)\right)}1_{\omega_{p(2^n)}}(2^{-\kappa}\eta)\,d\mathbf{y}\,d\boldsymbol{\eta}\,d\kappa\,d\boldsymbol{\xi'},
\end{aligned}
$$

where we have expressed I_p and ω_p as

$$
I_p = \prod_{i=1}^n\left[m_i 2^\alpha, (m_i+1)2^\alpha\right), \quad \omega_p = \prod_{i=1}^n\left[l_i 2^{-\alpha}, (l_i+1)2^{-\alpha}\right)
$$

with

$$
\mathbf{m} = (m_1, m_2, \cdots, m_n),\ \mathbf{l} = (l_1, l_2, \cdots l_n)\in\mathbb{Z}^n.
$$

Interpreting the sum

$$
\sum_{\mathbf{m}\in\mathbb{Z}^n} e^{2\pi i\left(\mathbf{m}+\frac{1}{2}\right)\cdot 2^{\alpha-\kappa}\left(\boldsymbol{\xi}-\boldsymbol{\xi'}\right)}
$$

in the sense of distributions, we find that

(8.1)

$$
\begin{aligned}
(Af)\widehat{\ }(\boldsymbol{\xi}) = c\lim_{N\to\infty}\frac{1}{|K_N|}\int_{K_N\times[0,1]}\sum_{\mathbf{l}\in\mathbb{Z}^n,\alpha\in\mathbb{Z}}\widehat{f}(\boldsymbol{\xi})\left|\widehat{\phi}\left(2^{\alpha-\kappa}(\boldsymbol{\xi}+\boldsymbol{\eta})-\left(\mathbf{l}+\frac{1}{4}\right)\right)\right|^2 \\
\times 1_{\omega_{p(2^n)}}(2^{-\kappa}\boldsymbol{\eta})\,d\mathbf{y}\,d\boldsymbol{\eta}\,d\kappa.
\end{aligned}
$$

Now, $\widehat{\phi}\left(2^{\alpha-\kappa}(\boldsymbol{\xi}+\boldsymbol{\eta})-\left(\mathbf{l}+\frac{1}{4}\right)\right)$ is supported on

$$
\begin{aligned}
&\left\{(\boldsymbol{\xi},\boldsymbol{\eta})\,:\,-\frac{1}{10}\le 2^{\alpha-\kappa}(\xi_i+\eta_i)-\left(l_i+\frac{1}{4}\right)\le\frac{1}{10}\text{ for all } i,\ 1\le i\le n\right\} \\
=&\left\{(\boldsymbol{\xi},\boldsymbol{\eta})\,:\,\left(l_i+\frac{3}{20}\right)2^{\kappa-\alpha}\le\xi_i+\eta_i\le\left(l_i+\frac{7}{20}\right)2^{\kappa-\alpha}\text{ for all } i,\ 1\le i\le n\right\}.
\end{aligned}
$$

Also,

$$1_{\omega_{p(2^n)}}(2^{-\kappa}\boldsymbol{\eta}) \neq 0 \Leftrightarrow \left(l_i + \frac{1}{2}\right) 2^{-p} \leq 2^{-\kappa}\eta_i \leq (l_i+1)\, 2^{-\alpha}$$
$$\Leftrightarrow \left(l_i + \frac{1}{2}\right) 2^{\kappa-\alpha} \leq \eta_i \leq (l_i+1)\, 2^{\kappa-\alpha}.$$

Therefore, the integrand in the right hand side of (8.1) is supported in

$$\bigcup_{\alpha\in\mathbb{Z}} \left\{ -\frac{17}{20} 2^{\kappa-\alpha} \leq \xi_i \leq -\frac{3}{20} 2^{\kappa-\alpha} \text{ for all } i,\ 1 \leq i \leq n \right\}.$$

Moreover, if

$$-\frac{13}{20} 2^{\kappa-\alpha} \leq \xi_i \leq -\frac{7}{20} 2^{\kappa-\alpha} \text{ for all } i, \quad 1 \leq i \leq n, \tag{8.2}$$

then

$$\left\{ \boldsymbol{\eta} \,:\, \left(l_i + \frac{3}{20}\right) 2^{\kappa-\alpha} \leq \xi_i + \eta_i \leq \left(l_i + \frac{7}{20}\right) 2^{\kappa-\alpha} \text{ for all } i,\ 1 \leq i \leq n \right\}$$
$$\subset \left\{ \boldsymbol{\eta} \,:\, \left(l_i + \frac{1}{2}\right) 2^{\kappa-\alpha} \leq \eta_i \leq (l_i+1)\, 2^{\kappa-\alpha} \right\}.$$

Note that there exists a nonempty open cone $\tilde{K}_0$ with vertex at the origin such that for all $\boldsymbol{\xi} \in \tilde{K}_0$, there exist $\alpha \in \mathbb{Z}$ and $\kappa \in [0,1]$ satisfying (8.2). In the sequel we work with such a cone.

For $\boldsymbol{\xi} \in \tilde{K}_0$ and choosing $K_N = [-N,N]^n \times [-N,N]^n$, we have

$$\begin{aligned}
(Af)\widehat{\ }(\boldsymbol{\xi}) &= c\widehat{f}(\boldsymbol{\xi}) \times \\
&\lim_{N\to\infty} \frac{1}{|K_N|} \int_{K_N\times[0,1]} \sum_{\mathbf{l}\in\mathbb{Z}^n,\alpha\in\mathbb{Z}} \left| \widehat{\phi}\left(2^{\alpha-\kappa}\left(\boldsymbol{\xi}+\boldsymbol{\eta}\right) - \left(\mathbf{l} + \frac{1}{4}\right)\right)\right|^2 d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa \\
&= c\widehat{f}(\boldsymbol{\xi}) \times \\
&\lim_{N\to\infty} \frac{1}{N^n} \int_{(\boldsymbol{\eta},\kappa)\in[-N,N]^n\times[0,1]} \sum_{\substack{\mathbf{l}\in\mathbb{Z}^n, p\in\mathbb{Z} \\ |\xi_i|\sim 2^{\kappa-\alpha} \\ \eta_i \sim l_i 2^{\kappa-\alpha}\ \forall i}} \left| \widehat{\phi}\left(2^{\alpha-\kappa}\left(\boldsymbol{\xi}+\boldsymbol{\eta}\right) - \left(\mathbf{l} + \frac{1}{4}\right)\right)\right|^2 d\boldsymbol{\eta}\, d\kappa \\
&= c\widehat{f}(\boldsymbol{\xi}) \lim_{N\to\infty} N^{-n} \int_0^1 \sum_{\substack{\mathbf{l}\in\mathbb{Z}^n,\alpha\in\mathbb{Z} \\ |\xi_i|\sim 2^{\kappa-\alpha} \\ -N\lesssim l_i 2^{\kappa-\alpha} \lesssim N\ \forall i}} \left(2^{\kappa-\alpha}\right)^n d\kappa \\
&= c\widehat{f}(\boldsymbol{\xi}) \lim_{N\to\infty} N^{-n} \int_0^1 \sum_{\substack{\alpha\in\mathbb{Z} \\ |\xi_i|\sim 2^{\kappa-\alpha}\ \forall i}} \left(N 2^{-\kappa+\alpha}\right)^n \left(2^{\kappa-\alpha}\right)^n d\kappa
\end{aligned}$$

$$= c\widehat{f}(\boldsymbol{\xi}) \int_0^1 \sum_{\substack{\alpha\in\mathbb{Z} \\ |\xi_i|\sim 2^{\kappa-\alpha}\ \forall i}} 1\, d\kappa$$
$$= c\widehat{f}(\boldsymbol{\xi}).$$

This completes the proof of the lemma.

Now, let m be the multiplier associated with the Calderón-Zygmund kernel K. Then $m \in C^\infty(\mathbb{R}^n\setminus\{\mathbf{0}\})$ and is homogeneous of degree 0. Suppose further, without loss of generality, that m is supported on the cone $\tilde{K}_0$ described earlier. We may reduce the problem to this case via a partition of unity and by invoking rotation invariance.

Recalling that

$$Bf = f * K,$$

and

$$\mathcal{C}f(\mathbf{x}) = \sup_{\boldsymbol{\eta}} |M_{\boldsymbol{\eta}} B M_{-\boldsymbol{\eta}} f|(\mathbf{x}),$$

we define, for $\boldsymbol{\zeta} \in \mathbb{R}^n$,

$$B_{\boldsymbol{\zeta}} f := \sum_{p\in\mathbf{D}} \langle f, \phi_p\rangle \psi_p^{\boldsymbol{\zeta}} 1_{\omega_{p(2^n)}}(\boldsymbol{\zeta}),$$

where

$$\left(\psi_p^{\boldsymbol{\zeta}}\right)\widehat{\ }(\boldsymbol{\xi}) = m(\boldsymbol{\xi}-\boldsymbol{\zeta})\widehat{\phi_p}(\boldsymbol{\xi}).$$

Using Lemma 6, it is not hard to see that

$$\begin{aligned} Bf :&= \lim_{N\to\infty} \frac{1}{|K_N|} \int_{K_N\times[0,1]} M_{-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} B_{2^{-\kappa}\boldsymbol{\eta}} D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}} f\, d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa \\ &= \lim_{N\to\infty} \frac{1}{|K_N|} \int_{K_N\times[0,1]} \sum_{p\in\mathbf{D}} \Big[\langle f, M_{-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \phi_p\rangle \times \\ &\qquad \left(M_{-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \psi_p^{2^{-\kappa}\boldsymbol{\eta}}\right) \left(1_{\omega_{p(2^n)}}(2^{-\kappa}\boldsymbol{\eta})\right)\Big]\, d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa. \end{aligned}$$

In fact,

$$\begin{aligned} (Bf)\widehat{\ }(\boldsymbol{\xi}) &= \lim_{N\to\infty} \frac{1}{|K_N|} \int_{K_N\times[0,1]} \sum_{p\in\mathbf{D}} \Big[\langle f, M_{-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \phi_p\rangle \times \\ &\qquad T_{-\boldsymbol{\eta}} M_{\mathbf{y}} D^2_{2^\kappa} \left(\psi_p^{2^{-\kappa}\boldsymbol{\eta}}\right)\widehat{\ }(\boldsymbol{\xi}) \left(1_{\omega_{p(2^n)}}(2^{-\kappa}\boldsymbol{\eta})\right)\Big]\, d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa \\ &= \lim_{N\to\infty} \frac{1}{|K_N|} \int_{K_N\times[0,1]} \sum_{p\in\mathbf{D}} \Big[\langle f, M_{-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \phi_p\rangle \times \\ &\qquad T_{-\boldsymbol{\eta}} M_{\mathbf{y}} D^2_{2^\kappa} \left[m\left(\boldsymbol{\xi} - \boldsymbol{\eta}2^{-\kappa}\right) \widehat{\phi_p}(\boldsymbol{\xi})\right] \left(1_{\omega_{p(2^n)}}(2^{-\kappa}\boldsymbol{\eta})\right)\Big]\, d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa. \end{aligned}$$

Since

$$T_{-\boldsymbol{\eta}} M_{\mathbf{y}} D^2_{2^\kappa} \left[m\left(\boldsymbol{\xi} - \boldsymbol{\eta}2^{-\kappa}\right) \widehat{\phi_p}(\boldsymbol{\xi})\right] = m(\boldsymbol{\xi}) T_{-\boldsymbol{\eta}} M_{\mathbf{y}} D^2_{2^\kappa} \widehat{\phi_p}(\boldsymbol{\xi}),$$

we get that

$$(Bf)\widehat{\ }(\boldsymbol{\xi}) = m(\boldsymbol{\xi})(Af)\widehat{\ }(\boldsymbol{\xi}) = c\, m(\boldsymbol{\xi})\widehat{f}(\boldsymbol{\xi}),$$

where the last equality follows from the claim and the fact that $\operatorname{supp} m \subset \tilde{K}_0$.

Now, for $f \in C_0^\infty(\mathbb{R}^n)$ and $\boldsymbol{\zeta} \in \mathbb{R}^n$

$$\begin{aligned}
&M_{\boldsymbol{\zeta}} B M_{-\boldsymbol{\zeta}} f(\mathbf{x}) \\
&= \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} M_{\boldsymbol{\zeta}} M_{-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} B_{2^{-\kappa}\boldsymbol{\eta}} D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}} M_{-\boldsymbol{\zeta}} f(\mathbf{x})\, d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa \\
&= \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} M_{\boldsymbol{\zeta}-\boldsymbol{\eta}} T_{-\mathbf{y}} D^2_{2^{-\kappa}} B_{2^{-\kappa}\boldsymbol{\eta}} D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}-\boldsymbol{\zeta}} f(\mathbf{x})\, d\mathbf{y}\, d\boldsymbol{\eta}\, d\kappa \\
&= \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l^{\boldsymbol{\zeta}}\times[0,1]} M_{-\boldsymbol{\eta}'} T_{-\mathbf{y}} D^2_{2^{-\kappa}} B_{2^{-\kappa}(\boldsymbol{\eta}'+\boldsymbol{\zeta})} D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f(\mathbf{x})\, d\mathbf{y}\, d\boldsymbol{\eta}'\, d\kappa \\
&= \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} M_{-\boldsymbol{\eta}'} T_{-\mathbf{y}} D^2_{2^{-\kappa}} B_{2^{-\kappa}(\boldsymbol{\eta}'+\boldsymbol{\zeta})} D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f(\mathbf{x})\, d\mathbf{y}\, d\boldsymbol{\eta}'\, d\kappa,
\end{aligned}$$

where $K_l^{\boldsymbol{\zeta}} := \boldsymbol{\zeta} + K_l$. The last equality follows from the fact that for $f \in C_0^\infty(\mathbb{R}^n)$ and any fixed $\boldsymbol{\zeta}$, the integrand gets arbitrarily small on the domain $K_l^{\boldsymbol{\zeta}} \triangle K^l$ as $l \to \infty$. The details are left to the interested reader.

Therefore,

$$\begin{aligned}
&\sup_{\boldsymbol{\zeta}} |M_{\boldsymbol{\zeta}} B M_{-\boldsymbol{\zeta}} f(\mathbf{x})| \\
&\le \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} \left| M_{-\boldsymbol{\eta}'} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \left[\sup_{\boldsymbol{\zeta}} B_{\boldsymbol{\zeta}}\right] D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f(\mathbf{x}) \right| d\mathbf{y}\, d\boldsymbol{\eta}'\, d\kappa.
\end{aligned}$$

We recall the following fact about the weak L^2 norm : there exist universal constants C_1, C_2 such that

$$C_1 \sup_E \frac{|\langle f, 1_E\rangle|}{|E|^{\frac{1}{2}}} \le ||f||_{L^{2,\infty}} \le C_2 \sup_E \frac{|\langle f, 1_E\rangle|}{|E|^{\frac{1}{2}}},$$

where the supremum is over all measurable sets E with finite Lebesgue measure. This implies

$$\begin{aligned}
&\left|\left| \sup_{\boldsymbol{\zeta}} |M_{\boldsymbol{\zeta}} B M_{-\boldsymbol{\zeta}} f| \right|\right|_{L^{2,\infty}} \\
&\le C \sup_E \frac{1}{|E|^{\frac{1}{2}}} \\
&\qquad \times \int_E \left[\lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} \left| M_{-\boldsymbol{\eta}'} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \left(\sup_{\boldsymbol{\zeta}} B_{\boldsymbol{\zeta}}\right) D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f(\mathbf{x}) \right| d\mathbf{y} d\boldsymbol{\eta}' d\kappa \right] d\mathbf{x} \\
&\le C \lim_{l\to\infty} \frac{1}{|K_l|}
\end{aligned}$$

$$\times \int_{K_l\times[0,1]} \left[\sup_E \frac{1}{|E|^{\frac{1}{2}}} \int_E \left| M_{-\boldsymbol{\eta}'} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \left(\sup_{\boldsymbol{\zeta}} B_{\boldsymbol{\zeta}} \right) D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f(\mathbf{x}) \right| d\mathbf{x} \right] d\mathbf{y}\, d\boldsymbol{\eta}'\, d\kappa$$

$$\leq C \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} \left|\left| |M_{-\boldsymbol{\eta}'} T_{-\mathbf{y}} D^2_{2^{-\kappa}} \left(\sup_{\boldsymbol{\zeta}} B_{\boldsymbol{\zeta}} \right) D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f| \right|\right|_{L^{2,\infty}} d\mathbf{y}\, d\boldsymbol{\eta}'\, d\kappa$$

$$\leq C \lim_{l\to\infty} \frac{1}{|K_l|} \int_{K_l\times[0,1]} \left|\left| \left(\sup_{\boldsymbol{\zeta}} B_{\boldsymbol{\zeta}} \right) D^2_{2^\kappa} T_{\mathbf{y}} M_{\boldsymbol{\eta}'} f \right|\right|_{L^{2,\infty}} d\mathbf{y}\, d\boldsymbol{\eta}'\, d\kappa,$$

since the weak L^2 norm is invariant under the translation, dilation and modulation operators defined in Section 2. The same invariance properties also hold true for the L^2 norm. In order to prove the weak L^2 bound for the Carleson operator, it therefore suffices to show that

$$||\sup_{\boldsymbol{\zeta}} B_{\boldsymbol{\zeta}} f||_{L^{2,\infty}} \leq C||f||_{L^2}, \tag{8.3}$$

which is the conclusion of Theorem 1.

9. Acknowledgements

Versions of the main result in this paper were obtained independently by the two authors as part of their dissertations at the University of California, Berkeley and the University of Missouri, Columbia respectively. The authors would like to thank their advisors Michael Christ and Loukas Grafakos for encouragement and many insightful discussions during preparation of the paper.

References

[1] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157.
[2] C. Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. **98** (1973), 551–571.
[3] M. Lacey and C. Thiele, *L^p estimates on the bilinear Hilbert transform for* $2 < p < \infty$, Ann. of Math. **146** (1997), 693–724.
[4] M. Lacey and C. Thiele, *On Calderón's conjecture*, Ann. of Math **149** (1999), 475–496.
[5] M. Lacey and C. Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Let. **7** (2000), 361–370.
[6] P. Sjölin,*Convergence almost everywhere of certain singular integrals and multiple Fourier series*, Arkiv For Matematik **9** (1971), 65–90.
[7] E.M. Stein, *Harmonic Analysis: Real variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton NJ 1993.

Malabika Pramanik, Department of Mathematics, University of Wisconsin, Madison, Madison, WI 53706
E-mail address: malabika@math.wisc.edu

Erin Terwilleger, Department of Mathematics, University of Missouri, Columbia, MO 65211, USA
E-mail address: mathgr60@math.missouri.edu