SINGULAR MAXIMAL FUNCTIONS AND RADON TRANSFORMS NEAR L^1

ANDREAS SEEGER   TERENCE TAO   JAMES WRIGHT

May 12, 2002

ABSTRACT. We show that some singular maximal functions and singular Radon transforms satisfy a weak type $L \log\log L$ inequality. Examples include the maximal function and Hilbert transform associated to averages along a parabola. The weak type inequality yields pointwise convergence results for functions which are locally in $L \log\log L$.

1.Introduction

Let Σ be a compact smooth hypersurface of $\mathbb{R}^d$, and let μ be a compactly supported smooth density on Σ, *i.e.*

$$\mu = \chi d\sigma$$

where $\chi \in C_0^\infty(\mathbb{R}^d)$ and $d\sigma$ is the surface carried measure on Σ.

Unless stated otherwise we shall always make the following

Curvature Assumption. *The Gaussian curvature does not vanish to infinite order on Σ.*

We consider a group of dilations on $\mathbb{R}^d$, given by $t^P = \exp(P \log t)$, $t > 0$, and we assume that P is a $d \times d$ matrix whose eigenvalues have positive real part. For $k \in \mathbb{Z}$ we set $\delta_k = 2^{kP}$ and define the measure μ_k by

$$\langle \mu_k, f \rangle = \langle \mu, f(\delta_k \cdot) \rangle. \tag{1.1}$$

We shall consider the convolutions $\mu_k * f$ and study the behavior of the maximal function

$$\mathcal{M}f(x) = \sup_{k \in \mathbb{Z}} |\mu_k * f(x)| \tag{1.2}$$

and some related singular integrals. By a rescaling we may assume that the measure μ is supported in the unit ball $\{x : |x| \leq 1\}$.

The first complete L^p bounds ($1 < p < \infty$) for a class of such operators (Hilbert transforms on curves) seems to be due to Nagel, Rivière and Wainger [9]. A classical reference is the article by Stein and Wainger [17] containing many related results; see also the paper by Duoandikoetxea and Rubio de Francia [6] which contains general results for maximal functions and singular integrals generated by singular measures, with decay assumptions on the Fourier transform. Concerning the behavior on L^1 it is presently not known even for the special classes considered here whether the maximal operator $\mathcal{M}$ is of weak type $(1, 1)$, *i.e.*

1991 *Mathematics Subject Classification.* 42B20.

The first author is supported in part by a grant from the National Science Foundation. The second author is a Clay Prize fellow and is supported by the Sloan and Packard foundations.

whether it maps L^1 to the Lorentz space $L^{1,\infty}$. This question had been raised in [17]. For some 'flat' cases counterexamples are in [3], but these do not seem to apply in the case of our curvature assumption.

We shall examine the behavior of the maximal function on spaces "near" L^1. Two results in this direction are known: Christ and Stein [4] showed by an extrapolation argument that if f is supported in a cube Q and $f \in L\log L(Q)$ then the maximal function $\mathcal{M}f$ belongs to $L^{1,\infty}$ (again under substantially weaker finite type assumptions). Moreover Christ [2] showed that the lacunary spherical maximal function maps the standard Hardy space $H^1(\mathbb{R}^d)$ to $L^{1,\infty}$, and that maximal functions and Hilbert transforms associated to a parabola in $\mathbb{R}^2$ map the appropriate Hardy space with respect to nonisotropic dilations to $L^{1,\infty}$. Weak L^1 (see also Grafakos [8] and our recent paper [12] for related results). For the two operators associated to the parabola (t,t^2) it is also known ([11]) that they map the smaller product-type Hardy space $H^1_{\text{prd}}(\mathbb{R}\times\mathbb{R})$ to the smaller Lorentz space $L^{1,2}$.

We recall that for f to belong to a Hardy space H^1 a rather substantial cancellation condition has to be satisfied. If locally the cancellation is missing one has a restriction on the size of f; more precisely if a function $f \in H^1$ is single signed in an open ball then f belongs to $L\log L(K)$ for all compact subsets K of this ball. This can be deduced from the maximal function characterization of H^1 and the fact that $f_0 \in L\log L(q_0)$ if f_0 is supported on the cube q_0 and the appropriate variant of the Hardy-Littlewood maximal function of f_0 belongs to $L^1(q_0)$, see [15, §I.5.2 (c)]. Here we are interested in the behavior in Orlicz spaces near L^1 without assuming additional cancellation conditions.

Our main result is that the maximal operator acts well on $L\log\log L$ and the global version satisfies weak type $L\log\log L$ inequalities. We first give a

Definition. *Let $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ be a convex function and let T be an operator mapping simple functions on $\mathbb{R}^d$ to measurable functions. T is of weak type $\Phi(L)$ if there is a constant C so that the inequality*

$$\big|\{x \in \mathbb{R}^d : |Tf(x)| > \alpha\}\big| \le \int \Phi\Big(\frac{C|f(x)|}{\alpha}\Big)dx \tag{1.3}$$

holds for all $\alpha > 0$.

Abusing the notation slightly we shall say that T is of weak type $L\log\log L$ if there is a constant C so that the inequality (1.3) holds with $\Phi(t) = t\log\log(e^2+t)$.

Theorem 1.1. *The maximal operator $\mathcal{M}$ is of weak type $L\log\log L$.*

We also prove a related theorem on singular convolution operators with kernels supported on hypersurfaces (assuming our finite type curvature assumption).

Let μ_k be as in (1.1) and assume that in addition

$$\int d\mu = 0. \tag{1.4}$$

For Schwartz functions f define the singular integral operator (or singular Radon transform) T by

$$Tf(x) = \sum_{k\in\mathbb{Z}} \mu_k * f. \tag{1.5}$$

Theorem 1.2. *T extends to an operator which is of weak type $L\log\log L$.*

1.3 Remarks and examples.

1.3.1. Theorem 1.1 implies an estimate on the Orlicz space $\Phi(L)(Q_0)$ where Q_0 is a unit cube and the norm on $\Phi(L)$ is given by $\|f\|_{\Phi(L)} = \inf\{\alpha > 0 : \int_{Q_0} \Phi(|f(x)|/\alpha)dx \le 1\}$. Consider the local maximal operator

$$\mathcal{M}_{\mathrm{loc}} f(x) = \sup_{k<C} |\mu_k * [f\chi_{Q_0}](x)|;$$

then $\mathcal{M}_{\mathrm{loc}}$ maps $L\log\log L(Q_0)$ to $L^{1,\infty}$. To see this we may assume that $\|f\|_{L\log\log L(Q_0)} = 1$. Then the estimate

$$|\{x \in Q_0 : \mathcal{M}_{\mathrm{loc}} f > \alpha\}| \lesssim \alpha^{-1}$$

is trivial for $\alpha < 1$ while for $\alpha > 1$ it follows from the better estimate (1.3).

We note that conversely the better estimate $|\{x \in \mathbb{R}^n : \mathcal{M}_{\mathrm{loc}} f > \alpha\}| \lesssim \int \Phi(C|f(x)|/\alpha)$ can be deduced from the $L\log\log L(Q_0) \to L^{1,\infty}$ boundedness by the Orlicz space variant of Stein's theorem [14]. Then the global variant of Theorem 1.1 follows by scaling and limiting arguments.

1.3.2. Similarly if we assume the cancellation condition (1.4) then the local singular Radon transform $\sum_{k<C} \mu_k * [f\chi_{Q_0}](x)$ maps $L\log\log L(Q_0)$ to $L^{1,\infty}$.

1.3.3. Suppose that $\int d\mu = 1$ and suppose that the measurable function f belongs locally to $L\log\log L$; i.e. $\int_K |f(x)|\log\log(e^2 + |f(x)|)dx < \infty$ for every compact set K. Then $\lim_{k\to-\infty} \mu_k * f(x) = f(x)$ almost everywhere.

This follows by a standard argument. Observe that we have $\int \alpha^{-1}|f(x)|\log\log(e^2 + \alpha^{-1}|f(x)|)dx < \infty$, for every $\alpha > 0$. Fix $\alpha > 0$ and let

$$\Omega_\alpha(f) = \{x : \limsup_{k\to-\infty} \mu_k * f(x) - \liminf_{k\to-\infty} \mu_k * f(x) > \alpha\}.$$

Given $\varepsilon > 0$ we show that $|\Omega_\alpha(f)| < \varepsilon$. One can find a bounded function h with compact support so that $\int \Phi(2C|f-h|/\alpha)dx \le \varepsilon$ and since $\mu_k * h \to h$ almost everywhere we see that $\Omega_{\alpha/2}(h)$ has measure zero. Moreover $|\Omega_\alpha(f)| \le |\Omega_{\alpha/2}(f-h)| + |\Omega_{\alpha/2}(h)|$ and by Theorem 1.1 we see that $\Omega_{\alpha/2}(f-h)$ and thus $\Omega_\alpha(f)$ has measure $< 2\varepsilon$. Since ε was arbitrary we see that $\Omega_\alpha(f)$ has measure zero; thus $\cup_m \Omega_{2^{-m}}(f)$ has measure zero and the result on pointwise convergence follows.

1.3.4. Examples of Theorem 1.1 include the lacunary spherical maximal operator where $\mu_k * f$ is the average of f over the sphere of radius 2^k centered at x (for the early L^p results see [1], [5]). The sphere may be replaced by any smooth compact hypersurface for which the curvature vanishes of finite order only, and the isotropic dilations may be replaced by nonisotropic ones. We remark that the proof of Theorem 1.1 for isotropic dilations is much less technical, see the expository note [13].

1.3.5. Other examples of Theorem 1.1 concern the averages along a parabola

$$\mathcal{P}_r f(x) = \frac{1}{r}\int_0^r f(x_1 - t, x_2 - t^b)dt$$

or higher dimensional versions for paraboloids $(t', |t'|^b)$, $b \ne 1$. Again if f belongs locally to $L\log\log L$ then $\lim_{r\to 0} \mathcal{P}_r f(x) = f(x)$ almost everywhere.

1.3.6. Similarly Theorem 1.2 can be used to deduce the weak type $L\log\log L$ inequality for the Hilbert transform

$$\mathcal{H}f(x) = p.v. \int_{-\infty}^{\infty} f(x_1 - t, x_2 - t^b)\frac{dt}{t}.$$

We give a brief outline of the paper. The main novelty in this paper is a stopping time argument based on the quantities of *thickness* Θ_n and *length* Λ_n associated to a density $v(x)dx$ (depending on an additional parameter n). Basically, the point is that the length $\Lambda_n[v]$ is used to control the size of an exceptional set while the thickness $\Theta_n[v]$ is used to control the L^2 norm of an essential part of the maximal function outside of the exceptional set, for suitable choices of v. The quantities of length and thickness are complementary in some sense; this and other basic properties are discussed in §2. In §3 we include preliminary and standard arguments from Calderón Zygmund theory. These arguments can be skipped by the experts; they may be used to reprove the known $L \log L$ estimates. In §4 we describe the stopping time argument based on length and thickness. The proof of the weak-type $L \log\log L$ inequality for the maximal operator is given in §5. The bounds for the singular Radon transforms are discussed in §6.

2. Length and thickness

In this section let v be an integrable nonnegative function which vanishes in the complement of a dyadic cube q. Dyadic cubes are supposed to be 'half-open', i.e. of the form $\prod_{i=1}^d [n_i 2^m, (n_i+1)2^m)$ where $n_i, m \in \mathbb{Z}$.

We define a dyadic version of a *one-dimensional Hausdorff content* or simply *length* $\lambda(E)$ to be

$$\lambda(E) := \inf_{\mathcal{Q}} \sum_{Q \in \mathcal{Q}} l(Q) \tag{2.1}$$

where $\mathcal{Q}$ ranges over all finite collections $\mathcal{Q}$ of dyadic cubes with $E \subset \bigcup_{Q\in\mathcal{Q}} Q$, and $l(Q)$ denotes the sidelength of Q. Note that this definition differs from the usual definition of a one-dimensional Hausdorff measure as $\lambda(E) \le l(Q)$ if E is contained in the dyadic cube Q.

Given $n \in \mathbb{Z}$ we denote by $E_n[v]$ the conditional expectation of v, for the σ-algebra generated by dyadic cubes of sidelength 2^{-n}; thus

$$E_n[v](x) = \sum_Q \chi_Q(x)|Q|^{-1} \int_Q v(y)dy$$

where of course the sum runs over all dyadic cubes of sidelength 2^{-n}. We also define

$$\mathfrak{S}_n(v) = \{x : E_n[v](x) \neq 0\}. \tag{2.2}$$

Notice that $v(x) = 0$ for almost every $x \in \mathbb{R}^d \setminus \mathfrak{S}_n[v]$ since v is nonnegative. Now define

$$\Lambda_n[v] = \lambda(\mathfrak{S}_n(v)). \tag{2.3}$$

Note that $\mathfrak{S}_n(v)$ is a union of dyadic cubes of length 2^{-n} and therefore the infimum in the definition of λ becomes a minimum; *i.e.* there is a collection $\mathcal{Q}$ of dyadic cubes covering the set $\mathfrak{S}_n(v)$ so that $\Lambda_n[v] = \sum_{Q\in\mathcal{Q}} l(Q)$. Here the cubes in $\mathcal{Q}$ have to be chosen to be of sidelength at least 2^{-n}.

Next we define the *thickness* of v to be the quantity

$$\Theta_n[v] := \sup_Q \frac{1}{l(Q)} \int_Q v(x)dx \tag{2.4}$$

where Q ranges over all dyadic cubes of sidelength $l(Q) \ge 2^{-n}$. Clearly, if v vanishes off a dyadic cube q it is sufficient to only consider dyadic subcubes of q in (2.4).

We note that the restriction to dyadic cubes in the definition of length and thickness is convenient but not essential. Since every cube of sidelength 2^L ($L \in \mathbb{Z}$) is contained in a union of 2^d dyadic cubes of sidelength 2^L we observe that

$$\begin{aligned} \Theta_n[v(\cdot + a)] &\le 2^d \Theta_n[v] \\ \Lambda_n[v(\cdot + a)] &\le 2^d \Lambda_n[v]. \end{aligned} \tag{2.5}$$

The quantities of length and thickness are complementary. Namely, it is immediate from the definitions of Λ_n and Θ_n that

$$\int v(x)dx \le \Lambda_n[v]\Theta_n[v]. \tag{2.6}$$

The bound (2.6) can be attained, for instance if v is the characteristic function of a dyadic box. It would be desirable to have a converse to (2.6), with bounded constants, but this generally does not hold as the following example shows. Let E_n be the union of $n+1$ rectangles R_ν, parallel to the coordinate axes, with dimensions $(2^{-\nu}, 1)$ so that the left lower endpoint of R_ν has coordinates $(\nu, 0)$, $\nu = 0, \dots, n$. Let $v_n = \chi_{E_n}$. Then $\Lambda_n[v_n] = n+1$, $\int v_n(x)dx < 2$ and $\Theta_n[v_n] = 1$; thus the converse of (2.3) fails with a uniform constant.

However we shall show that v can be efficiently decomposed into a sum of functions for which a converse of (2.6) does hold. The main result needed to achieve this is

Proposition 2.1. *Let q be a dyadic cube with $l(q) \ge 2^{-n}$. Suppose that v is a bounded nonnegative measurable function supported in q. Then there exists a decomposition*

$$v = g + h$$

with nonnegative functions g and h and g, h vanish in the complement of the set $\mathfrak{S}_n(v) \subset q$; moreover the inequalities

$$\Lambda_n[h] \le \frac{1}{2}\Lambda_n[v] \tag{2.7}$$

and

$$\Lambda_n[v]\Theta_n[g] \le 8 \int g(x)dx \tag{2.8}$$

hold.

In particular we see from (2.7/8) that the function g satisfies

$$\Lambda_n[g]\Theta_n[g] \le 8 \int g(x)dx,$$

thus a converse to (2.6).

We shall first prove a technical result which states that for each dyadic cube one may construct a function v_I from v so that v_I has 'controlled' thickness and 'large' integral.

Lemma 2.2. *Let $\gamma > 0$. For any dyadic cube I of sidelength $\ge 2^{-n}$, there exists a (possibly empty) collection $\mathcal{Q}[I]$ of disjoint dyadic cubes of sidelength $\ge 2^{-n}$ contained in I, and a measurable function v_I such that*

$$0 \le v_I(x) \le v\chi_I(x) \tag{2.9}$$

for all $x \in \mathbb{R}^d$,

$$\Theta_n[v_I] \le 2\gamma \tag{2.10}$$

and

$$2\int v_I(x)dx \ge 2\gamma \sum_{Q\in\mathcal{Q}[I]} l(Q) + \int_{I\setminus\bigcup_{Q\in\mathcal{Q}[I]}Q} v(x)dx. \tag{2.11}$$

Proof. We prove this by induction on the sidelength of I. We first assume that $l(I) = 2^{-n}$. Notice that in this case we have

$$\Theta_n[v\chi_I] = \frac{1}{l(I)} \int_I v(x)dx.$$

We distinguish two cases. First if $\Theta_n[v\chi_I] \le 2\gamma$ we choose $v_I = v\chi_I$ and take for $\mathcal{Q}[I]$ the empty collection. Clearly (2.9), (2.10), (2.11) are satisfied.

Next if $\Theta_n[v\chi_I] > 2\gamma$ we may choose a measurable function v_I which vanishes outside I such that $0 \le v_I(x) \le v\chi_I(x)$ for all $x \in \mathbb{R}^d$ and

$$\gamma \le \frac{1}{l(I)} \int_I v_I(x)dx \le 2\gamma. \tag{2.12}$$

Clearly $\Theta_n[v_I] \le 2\gamma$. For $\mathcal{Q}[I]$ we take the singleton collection $\{I\}$ and (2.11) is satisfied because of the first inequality in (2.12).

Now fix a dyadic cube I with $l(I) > 2^{-n}$ and suppose that the lemma has been proven for all proper dyadic subcubes I' of sidelength at least 2^{-n}. Partition I into 2^d subcubes $I_1, \dots, I_{2^d}$ of sidelength $\frac{1}{2}l(I)$. By the induction hypothesis, we may construct collections $\mathcal{Q}[I_j]$ and measurable functions v_{I_j} for $j = 1, \dots, 2^d$ satisfying the properties of the lemma relative to I_j.

To prove the assertion for I we again distinguish two cases. First suppose that

$$\sum_{j=1}^{2^d} \int v_{I_j}(x)dx \le 2\gamma l(I). \tag{2.13}$$

In this case we simply define $v_I(x) := \sum_{j=1}^{2^d} v_{I_j}(x)$ and $\mathcal{Q}[I] := \bigcup_{j=1}^{2^d} \mathcal{Q}[I_j]$. Then by the induction hypothesis

$$2\int v_I(x)dx = \sum_{j=1}^{2^d} 2\int v_{I_j}(x)dx \ge \sum_{j=1}^{2^d} \Big[2\gamma \sum_{Q \in \mathcal{Q}[I_j]} l(Q) + \int_{I_j \setminus \cup_{Q \in \mathcal{Q}[I_j]} Q} v(x)dx\Big]$$

which is equal to the right hand side of (2.11). From (2.13) it follows that

$$\frac{1}{l(I)} \int v_I(x)dx \le 2\gamma$$

and if Q is a proper dyadic subcube of I then $Q \subset I_j$ for some j and

$$\frac{1}{l(Q)} \int_Q v_I(x)dx = \frac{1}{l(Q)} \int_Q v_{I_j}(x)dx \le 2\gamma$$

by the induction hypothesis. Altogether (2.10) follows in case (2.13).

Now suppose that

$$\sum_{j=1}^{2^d} \int v_{I_j}(x)dx > 2\gamma l(I). \tag{2.14}$$

In this case we can find a function v_I so that $v_I(x) \le \sum_{j=1}^{2^d} v_{I_j}(x)$ and

$$\gamma l(I) \le \int v_I dx \le 2\gamma l(I). \tag{2.15}$$

We then take for $\mathcal{Q}[I]$ the singleton set $\{I\}$. Then (2.11) is immediate by (2.15). Clearly also by (2.15) $\frac{1}{l(I)}\int v_I(x)dx \le 2\gamma$. As above we can use the induction hypothesis to see that if Q is a proper dyadic subcube, thus contained in an I_j, we have $\frac{1}{l(Q)}\int_Q v_I(x)dx \le \frac{1}{l(Q)}\int_Q v_{I_j}(x)dx \le 2\gamma$, thus altogether (2.10) also holds in this case. □

Proof of Proposition 2.1. We define the *critical thickness* $\vartheta_n(v)$ to be the largest non-negative number γ such that the inequality

$$\gamma\Lambda_n[v] \le 2\gamma\sum_{Q\in\mathcal{Q}} l(Q) + \int_{q\setminus\bigcup_{Q\in\mathcal{Q}}Q} v(x)dx \tag{2.16}$$

holds for all finite collections $\mathcal{Q}$ of dyadic cubes of sidelength 2^{-n} (here the empty collection is admitted). Equivalently, one can define $\vartheta_n(v)$ by

$$\vartheta_n(v) := \inf_{\mathcal{Q}} \frac{\int_{q\setminus\bigcup_{Q\in\mathcal{Q}}Q} v(x)dx}{(\Lambda_n[v]-2\sum_{Q\in\mathcal{Q}} l(Q))_+}. \tag{2.17}$$

Observe that since v vanishes in the complement of q and since all cubes have sidelength at least 2^{-n} we are in effect taking the infimum over a finite set of collections, each consisting of a finite number of cubes, so that this infimum becomes a minimum, and (2.16) holds with $\gamma = \vartheta_n(v)$.

Clearly $\vartheta_n(v) \le \Lambda_n[v]^{-1}\int v(x)dx$. Observe also that $\vartheta_n(v) > 0$ since $\int_{q\setminus\bigcup_{Q\in\mathcal{Q}}Q} v(x)dx$ is positive whenever $\sum_{Q\in\mathcal{Q}} l(Q) \le \Lambda_n[v]/2$.

We can now find a finite collection $\mathcal{Q}_1$ of dyadic cubes in q, of sidelength at least 2^{-n}, so that

$$\vartheta_n(v)\Lambda_n[v] = 2\vartheta_n(v)\sum_{Q\in\mathcal{Q}_1} l(Q) + \int_{E_*} v(x)dx \tag{2.18}$$

where

$$E_* := q\setminus\bigcup_{Q\in\mathcal{Q}_1} Q. \tag{2.19}$$

We claim that

$$\Theta_n[v\chi_{E_*}] \le 2\vartheta_n(v). \tag{2.20}$$

Indeed, suppose for contradiction that there existed a dyadic cube Q' such that

$$\int_{E_*\cap Q'} v(x)dx > 2\vartheta_n(v)l(Q'). \tag{2.21}$$

By (2.21) and $\vartheta_n(v) > 0$ we have $|E_*\cap Q'| > 0$ which implies that $Q' \notin \mathcal{Q}_1$. If we apply (2.16) to the collection $\mathcal{Q}_1\cup\{Q'\}$ we obtain

$$\vartheta_n(v)\Lambda_n[v] \le 2\vartheta_n(v)\Big(l(Q') + \sum_{Q\in\mathcal{Q}_1} l(Q)\Big) + \int_{E_*\setminus Q'} v(x)dx,$$

but by (2.18) this implies

$$\int_{E_*} v(x)dx \le 2\vartheta_n(v)l(Q') + \int_{E_*\setminus Q'} v(x)dx$$

contradicting (2.21). This proves (2.20).

We shall now invoke Lemma 2.2 with $\gamma = \vartheta_n(v)$ and $I = q$, thus finding a function v_q and a collection $\mathcal{Q}[q]$ obeying the properties in the lemma. We define

$$g(x) = v(x)\chi_{E_*}(x) + v_q(x)\chi_{q\setminus E_*}(x)$$

and

$$h(x) = \big(v(x) - v_q(x)\big)\chi_{q\setminus E_*}(x).$$

Observe that g and h are nonnegative functions. To show (2.7) we use that $\Lambda_n[h] \le \lambda(q \setminus E_*)$ since the latter set is a union of dyadic cubes of sidelength 2^{-n}. Thus we observe

$$\Lambda_n[h] \le \sum_{Q\in\mathcal{Q}_1} l(Q) \le \frac{1}{2}\Lambda_n[v],$$

by (2.18). This gives (2.7).

To show (2.8) we use that $v_q \le v$ and observe that by (2.11)

$$\int g(x)dx \ge \int v_q(x)dx \ge \frac{1}{2}\Big(2\vartheta_n(v) \sum_{Q\in\mathcal{Q}[q]} l(Q) + \int_{q\setminus\cup_{Q\in\mathcal{Q}[q]}Q} v(x)dx\Big),$$

since now $\gamma = \vartheta_n(v)$. By (2.16) we thus see that

$$\int g(x)dx \ge \frac{1}{2}\Lambda_n[v]\vartheta_n(v).$$

By (2.20) and (2.10)

$$\Theta_n[g] \le \Theta_n[v\chi_{E_*}] + \Theta_n[v_q] \le 2\vartheta_n(v) + 2\vartheta_n(v) = 4\vartheta_n(v),$$

we see that $\Theta_n[g] \le 8\Lambda_n[v]^{-1}\int g(x)dx$ which is (2.8). $\square$

Remark. There are analogues of Proposition 2.1 where for $0 < \beta < d$ the length $\lambda(E)$ is replaced by the β-dimensional Hausdorff content

$$\lambda_\beta(E) = \inf_{\mathcal{Q}} \sum_{Q\in\mathcal{Q}} l(Q)^\beta$$

where again $\mathcal{Q}$ ranges over all finite collections $\mathcal{Q}$ of dyadic cubes with $E \subset \cup_{Q\in\mathcal{Q}}Q$. Then if we define $\Lambda_{\beta,n}(v) = \lambda_\beta(\mathfrak{S}_n(v))$ and the β-thickness by

$$\Theta_{\beta,n}[v] := \sup_Q \frac{1}{l(Q)^\beta}\int_Q v(x)dx$$

then an assertion analogous to Proposition 2.1 holds true. The proof requires only notational changes.

In what follows it will be convenient to extend the definition of length and thickness to not necessarily nonnegative functions, and we simply put

$$\Lambda_n[f] := \Lambda_n[|f|], \quad \Theta_n[f] := \Theta[|f|].$$

Proposition 2.1 can be applied iteratively. This leads to

Proposition 2.3. *Suppose that f is integrable and vanishes in the complement of dyadic cube of length* 1. *Set $h_0(x) = f(x)$. For $m \geq 1$ we may decompose*

$$f = h_m + \sum_{\nu=1}^{m} g_\nu$$

almost everywhere, so that the following properties hold.

(i) $h_m(x)$ and the $g_\nu(x)$ are nonnegative if and only if f is nonnegative, and $h_m(x)$ and the $g_\nu(x)$ are nonpositive if and only if f is nonpositive.

(ii) $\Theta_n[g_\nu]\Lambda_n[h_{\nu-1}] \leq 8\int |g_\nu(x)|dx$.

(iii) $\Lambda_n[h_m] \leq 2^{-m}\Lambda_n[f]$.

(iv) If $m \geq n$ then $g_{m+1} = h_m$, $h_{m+1} = 0$.

Proof. We first extend the statement of Proposition 2.1 to not necessarily nonnegative functions, in the obvious way. We simply decompose $|f| = \tilde{g} + \tilde{h}$ as in Proposition 2.1, and then define $g(x) = \tilde{g}(x)\text{sign}\,(f(x))$, and $h(x) = \tilde{h}(x)\text{sign}\,(f(x))$. We can then iterate this procedure (decomposing in the second step the function $|h| = \tilde{g}_2 + \tilde{h}_2$ etc.) and obtain the above decomposition so that statements (i), (ii), (iii) hold.

Also observe that if $\Lambda_n[|h|] \leq 2^{-n}$ then $\mathfrak{S}_n[h]$ is contained in a dyadic cube of sidelength 2^{-n} and we thus know that $\Theta_n[|h|]\Lambda_n[|h|] = \int |h(x)|dx$. This implies statement (iv). □

We now describe how the quantities of length and thickness are used in certain convolution estimates involving the measure μ and appropriate localizations μ^n. To define the localization we choose a C^∞ function ϕ with compact support in $\{x : |x| \leq 1/2\}$ such that $\int \phi(x)dx = 1$ and such that

$$\int \phi(x)(P(x) - P(0))dx = 0$$

for all polynomials of degree $\leq d$. Set $\phi_n(x) = 2^{nd}\phi(2^n x)$ and let

$$\mu^n = \phi_n * \mu. \tag{2.22}$$

Lemma 2.4. *Let f be supported on a set of diameter at most* 10. *Then*

$$\text{meas}(\text{supp }(\mu^n * f)) \lesssim \Lambda_n[f].$$

Proof. Note that if Q is a cube with center x_Q and sidelength $l(Q)$ with $2^{-n} \leq l(Q) \leq 100$ and f_Q is supported in Q then $\mu^n * f_Q$ is supported on the x_Q-translate of a tubular neighborhood of Σ of width $O(l(Q))$, thus on a set of measure $O(l(Q))$. The assertion follows by working with an efficient cover of the support of f arising from the definition of Λ_n. □

The quantity $\Theta_n[f]$ can be used to estimate the L^2 norm of the support $\mu^n * f$ provided that one has a lower bound for the curvature. To make this precise we first prove a slight variant of an observation in [7].

Lemma 2.5. *Let ψ be a real valued C^∞ function on $[-1,1]^d$, so that $\sup_{|\alpha|\leq 3} |\partial^\alpha \psi(x)| \leq A_3$; here $A_3 \leq 1$. Suppose $|\det \psi''(y_0)| \geq \beta$ and $Q \subset [-1,1]^{d-1}$ is a $d-1$ dimensional cube of sidelength $\varepsilon_1\beta$, containing y_0, here $\varepsilon_1 \leq [10(d-1)^4 A_3]^{-1}$.*

Let χ be a C^∞ function supported on Q so that the inequalities $\|\partial^\alpha \chi\|_\infty \leq c_\alpha(\varepsilon_1\beta)^{-|\alpha|}$ hold. Define the measure ν by

$$\langle \nu, f\rangle = \int \chi(y')f(y', \psi(y'))dy'$$

and define the reflection $\langle\widetilde{\nu}, f\rangle = \langle\nu, f(-\cdot)\rangle$.

Then there are constants C_α *so that*

$$|\partial_x^\alpha[\nu * \widetilde{\nu}](x)| \le C_\alpha \beta^{d-3-2|\alpha|}|x|^{-1-|\alpha|}.$$

Proof. We assume that $d \ge 3$ but after notational modification the proof applies also to the case $d = 2$. Since $\nu * \tilde{\nu}$ does not change if we translate the measure we may assume that $y_0 = 0$.

We compute

$$\begin{aligned}\langle\nu * \widetilde{\nu}, f\rangle &= \iint f(x-y)d\nu(x)d\nu(y)\\ &= \sum_k \int \chi(u'+y')\chi(u')\zeta_k(u')f(u', \psi(y'+u') - \psi(y'))dy'du' := \sum_k I_k(f)\end{aligned}$$

where the ζ_k form a partition of unity on the unit sphere in $\mathbb{R}^{d-1}$ which is extended to a homogeneous function of degree 0. We assume that the restriction of ζ_k to the unit sphere is supported on a set of diameter $\le \varepsilon_1\beta$ and the summation is over $O((\varepsilon_1\beta)^{1-d})$ terms. The ζ_k satisfy the natural estimates

$$|\partial^\alpha \zeta_k(u')| \le C_\alpha(\varepsilon_1\beta)^{-|\alpha|}|u'|^{-|\alpha|}.$$

Note that in the integral defining I_k the variables u' and y' are restricted to a ball of radius $\lesssim \varepsilon_1\beta$ and u' is further restricted to a sector with solid angle $\varepsilon_1\beta$.

Now note that by $|\partial^2_{x_ix_j}\psi| \le A_3$, $|\det\psi''(0)| \ge \beta$ and Cramer's rule we have

$$|u'| \le \beta^{-1}(d-1)^2 A_3^{d-1}|\psi''(0)u'|. \tag{2.23}$$

We now pick a unit vector $\theta_k \in \operatorname{supp}\zeta_k$.

Let

$$\mathfrak{v}_k = \frac{\psi''(0)\theta_k}{|\psi''(0)\theta_k|}$$

and let $\mathfrak{v}_{k,2}, \dots, \mathfrak{v}_{k,d-1}$ be an orthonormal basis of the orthogonal complement of $\mathbb{R}\mathfrak{v}_k$, and with $t'' = (t_2, \dots, t_{d-1})$ define $\mathfrak{w}_k(t'') = \sum_{i=2}^{d-1} t_i\mathfrak{v}_{k,i}$. Now write $y' = \mathfrak{w}_k(t'') + t_1\mathfrak{v}_k$ and we get

$$I_k(f) = \int_{t''}\int_{u'}\int_{t_1} \chi(u')\zeta_k(u')\chi(u' + \mathfrak{w}_k(t'') + t_1\mathfrak{v}_k)f(u', \Psi_k(t_1, t'', u'))dt_1du'dt''$$

where

$$\begin{aligned}\Psi_k(t_1, t'', u')) &= \psi(\mathfrak{w}_k(t'') + t_1\mathfrak{v}_k + u') - \psi(\mathfrak{w}_k(t'') + t_1\mathfrak{v}_k)\\ &= \Big\langle u', \int_0^1 \nabla\psi(\mathfrak{w}_k(t'') + t_1\mathfrak{v}_k + su')ds\Big\rangle.\end{aligned}$$

We wish to change variables in the inner t_1-integral. Observe that

$$\begin{aligned}\frac{d}{dt}\Psi_k(t_1, t'', u') =& |u'|\langle\theta_k, \psi''(0)\mathfrak{v}_k\rangle\\ &+ |u'|\int_0^1 \big\langle\theta_k, \big[\psi''(\mathfrak{w}_k(t'') + t_1\mathfrak{v}_k + su') - \psi''(0)\big]\mathfrak{v}_k\big\rangle ds\\ &+ |u'|\int_0^1 \big\langle\frac{u'}{|u'|} - \theta_k, \psi''(\mathfrak{w}_k(t'') + t_1\mathfrak{v}_k + su')\mathfrak{v}_k\big\rangle ds\\ =& |u'||\psi''(0)\theta_k| + e_1(t_1, t'', u') + e_2(t_1, t'', u')\end{aligned} \tag{2.24}$$

where by our assumption on the third derivatives the error term e_1 is bounded by $2(d-1)^2A_3\varepsilon_1\beta|u'|$, and since $u'\in \text{supp}\,\zeta_k$ the error term e_2 is bounded by $(d-1)^2A_3\varepsilon_1\beta|u'|$. The main term is $|u'||\psi''(0)\theta_k|\ge |u'|\beta(d-1)^{-2}A_3^{1-d}$ and thus the derivative $\partial_t\Psi_k$ is single signed and of size $\approx \beta|u'|$. Therefore we may perform the change of variables $t_1\mapsto u_d=\Psi_k(t_1,t'',u')$ with inverse $t_1^k(u_d;u',t'')$ and obtain

$$\langle \nu*\widetilde{\nu},f\rangle=\sum_k\iiint f(u',u_d)H_k(u',u_d,t'')du_d du' dt'' \tag{2.25}$$

where

$$H_k(u',u_d,t'')=\frac{\zeta_k(u')\chi(u')\chi(u'+\mathfrak{w}_k(t'')+t_1\mathfrak{v}_k)}{|\partial_t\Psi_k(t_1^k(u_d;u',t''),t'',u')|}.$$

We have the estimate

$$|H_k(u',s,t'')|\lesssim \beta^{-1}|u'|^{-1}$$

and $H_k(u',u_d,t'')$ vanishes if $|u'|\ge C|u_d|$ or $|u'/|u'|-\theta_k|\ge\varepsilon_1\beta$ or $|t''|\ge\beta$. Integrating in t'' yields a factor of $O(\beta^{d-2})$ and since $\sum_k\zeta_k(u')=O(1)$ we obtain the claimed estimate for $\alpha=0$. The estimates for the derivatives follow by a straightforward examination of the derivatives of $t_1^k(u_d;u',t'')$ and applications of the chain rule. We omit the details. □

Now let ϕ_n be as in (2.22).

Lemma 2.6. *There is a small constant ε_1 depending only on Σ so that the following holds for $\beta\le 1$.*

Let $\chi\in C_0^\infty$ is supported on a set of diameter $\varepsilon_1\beta$ and suppose that the support of χ contains a point P on Σ where the Gaussian curvature satisfies $|K(P)|\ge\beta$. Let $\nu^n=\phi_n*\mu$. Suppose that f is supported on a set of diameter 1. Then

$$\|\widetilde{\nu}^n*\nu^n*f\|_\infty\lesssim\beta^{d-3}(1+n)\Theta_n[f].$$

Proof. After localization and a change of variable we may reduce to the situation of Lemma 2.5.

Notice that $|\nu^n(x)|\lesssim 2^n$ since ν is a density on a hypersurface. By Lemma 2.5 we have

$$|\widetilde{\nu}^n*\nu^n*f(x)|\lesssim\beta^{d-3}\int\min\{2^n,\frac{1}{|x-y|}\}|f(y)|dy$$

and we observe that

$$\int_{|x-y|\le 2^{-n}}2^n|f(y)|dy\le 2^d\Theta_n[f]$$

and

$$\int_{2^{-\ell}\le|x-y|\le 2^{-\ell+1}}\frac{1}{|x-y|}|f(y)|dy\le 2^{d+1}\Theta_n[f],\quad 0\le\ell\le n.$$

The asserted estimate follows by summing over $\ell=0,\dots,n$. □

Finally we also need the behavior of the quantities of length and thickness under nonisotropic dilations. Here we will have to compare isotropic dilations to nonisotropic ones. Let $\tau=\mathit{trace}(P)$ and denote by λ_j the eigenvalues of P. Then we may choose positive constants a, A so that

$$a<\text{Re}(\lambda_j)<A<\tau. \tag{2.26}$$

Then there are positive constants $c_1\le C_1$ so that for all x

$$c_1t^a|x|\le|t^Px|\le C_1t^A|x|,\qquad t\ge 1. \tag{2.27}$$

Lemma 2.7. *Suppose that f is integrable and vanishes in the complement of a compact set.*

Then there is a constant C depending only on the dilation group and the dimension, so that

$$\Theta_n[f(\delta_j\cdot)] \le C2^{-j(\tau-A)}\Theta_n[f], \quad \text{if } j \ge 0 \tag{2.28}$$

and

$$\Lambda_n[f(\delta_{-m}\cdot)] \le C2^{Am}\Lambda_n[f], \quad \text{if } m \ge 0. \tag{2.29}$$

Proof. Let $j \ge 0$ and let Q be a dyadic cube of sidelength $l(Q) \ge 2^{-n}$. Then $\delta_j Q$ is contained in the union of at most 2^d dyadic cubes $\{q_i\}$, of sidelength $\approx 2^{jA}l(Q)$. Thus

$$\begin{aligned} l(Q)^{-1}\int_Q |f(\delta_j x)|dx &= 2^{-j\tau}l(Q)^{-1}\int_{\delta_j Q}|f(u)|du \\ &\le 2^{-j\tau}\sum_i C'(2^{-Aj}l(q_i))^{-1}\int_{q_i}|f(u)|du \le C'2^d 2^{-j(\tau-A)}\Theta_n[f]. \end{aligned}$$

If we take the supremum over all dyadic cubes we obtain (2.28).

Next let $m \ge 0$. Let $Q_1, \dots, Q_N$ be a cover of $\mathfrak{S}_n(|f|)$. Let Q_i^* be the double cube (dilated with respect to the center of Q_i).

Now $\mathfrak{S}_n(|f|) = \cup_{\nu=1}^{M_1} R_\nu$ where the R_ν are dyadic 2^{-n} cubes with center x_ν on which the expectation $E_n[|f|]$ does not vanish. Let $R^*_{\nu,m}$ be the union of dyadic cubes of sidelength 2^{-n} which intersect $\delta_m R_\nu$. Then $\mathfrak{S}_n(|f(\delta_{-m}\cdot)|)$ is contained in $\cup_{\nu=1}^{M_1} R^*_{\nu,m}$.

Since $m \ge 0$ each $R^*_{\nu,m}$ is contained in a 2-dilate of $\delta_m R_\nu$ relative to the center $\delta_m x_\nu$. Thus the union of the $R^*_{\nu,m}$ is contained in the union of the dilates $\delta_m Q_i^*$. Each $\delta_m Q_i^*$ is contained in no more than 4^d dyadic cubes of sidelength $2^{[mA+3]}l(Q_i)$. Consequently

$$\Lambda_n[f(\delta_{-m}\cdot)] \le C2^{Am}\sum_{i=1}^N l(Q_i).$$

If we work with an efficient cover of $\mathfrak{S}_n(|f|)$ we obtain (2.29). $\square$

3. Preliminary Calderón-Zygmund reductions

We shall begin with some reductions from standard Calderón-Zygmund theory. The estimates in this section together with a trivial L^1 estimate will only imply the known weak-type $L\log L$ inequality (see Corollary 3.1 below) but they apply to more general operators than those discussed in the introduction.

In this section we shall assume that the measure μ satisfies

$$|\widehat{\mu}(\xi)| \lesssim (1+|\xi|)^{-\gamma} \tag{3.1}$$

for some positive γ (without loss of generality $\gamma \le (d-1)/2$).

When estimating the singular integral operator (1.5) we shall assume the additional cancellation condition (1.4). We note that the original hypothesis of the curvature not vanishing to infinite order implies an estimate (3.1) for some $\gamma > 0$, by an application of van der Corput's lemma.

We shall apply a nonisotropic version of Calderón-Zygmund theory (see [10], [16]). Let ρ be a homogeneous distance function which satisfies $\rho(t^P x) = t\rho(x)$ for all x and $\rho(x) = 1$ if $|x| = 1$. If $x_0 \in \mathbb{R}^d$ and

$\rho_0 > 0$ then we set $B(x_0, \rho_0) = \{x : \rho(x - x_0) \le \rho_0\}$ and we refer to $B(x_0, \rho_0)$ as the ball with center x_0 and ρ_0 (see [17] for a discussion of such distance functions). Notice that

$$B(x_0, \rho_0) = \{x : |\rho_0^{-P}(x - x_0)| \le 1\}.$$

We note that $|x|^{1/a} \lesssim \rho(x) \lesssim |x|^{1/A}$ if $|x| \le 1$ and $|x|^{1/A} \lesssim \rho(x) \lesssim |x|^{1/a}$ if $|x| \ge 1$, see (2.26/27) above.

Let M_{HL} be the analogue of the Hardy-Littlewood maximal function associated to the family of these nonisotropic balls, *i.e.* $M_{HL}f(x) = \sup_{x \in B} |B|^{-1} \int_B |f(y)|dy$ where the supremum is taken over all balls $B = B(x_0, \rho_0)$ which contain x.

We now fix $\alpha > 0$ and define $\Omega = \{x : M_{HL}f > \alpha\}$ and thus

$$|\Omega| \lesssim \alpha^{-1} \|f\|_1.$$

By an analogue of the Lebesgue differentiation theorem we also know that $|f(x)| \le \alpha$ for all $x \in \mathbb{R}^d \setminus \Omega$.

The Calderón-Zygmund decomposition is based on a Whitney type decomposition. According to [16, p.15] there are constants $K_1 > 1$, $K_2 > 2$, $K_3 > 1$ (depending only on the distance function ρ), and a sequence of balls $B_1, \dots, B_j, \dots$, with $B_j = B(x_j, \rho_j)$, and a sequence $\mathfrak{W}$ of measurable sets ('generalized Whitney cubes') $w_1, \dots, w_j, \dots$, so that the following properties are satisfied:

(a) The B_j are pairwise disjoint.
(b) If $B_j^* = B(x_j, K_1\rho_j)$ then the numbers $K_1\rho_j$ belong to $\{2^j : j \in \mathbb{Z}\}$ and $\bigcup_j B_j^* = \Omega$. Moreover each $x \in \Omega$ is contained in no more than K_3 of the balls B_j^*.
(c) $B_j \subset w_j \subset B_j^*$
(d) The w_j are pairwise disjoint, and we have $\bigcup w_j = \Omega$.
(e) If $B_j^{**} = B(x_j, K_2\rho_j)$ then $B_j^{**} \cap (\mathbb{R}^d \setminus \Omega) \ne \emptyset$.
(f) Each B_j^{**} is contained in $\Omega^* = \{x : M_{HL}(\chi_\Omega) > (10K_2)^{-\tau}\}$ and thus

$$\text{meas}(\Omega^*) \lesssim \alpha^{-1}\|f\|_1 \lesssim \int \Phi(|f|/\alpha)dx. \tag{3.2}$$

We thus get a decomposition $f = g + \sum_{w \in \mathfrak{W}} f_w$ where $f_w(x) = f(x)$ if $x \in w$ and $|f(x)| > \alpha$ and $f_w(x) = 0$ otherwise; moreover $|g(x)| \lesssim \alpha$ and $|w|^{-1} \int |f_w|dx \lesssim \alpha$ for each w. The sets w play the role of the usual Whitney cubes. For each $w \in \mathfrak{W}$ we assign a point x_w and an integer $r(w)$ by setting $x_{w_j} = x_j$ and $r(w_j) = \log_2(K_1\rho_j)$.

In what follows we choose $c > 0$ small, specifically the choice

$$c < \frac{1}{2}\min\{1, \gamma\} \tag{3.3}$$

works. We then further decompose f_w by setting

$$f_w^n(x) = f_w(x) \quad \text{if} \quad 2^{c(n-1)}\alpha < |f_w(x)| \le 2^{cn}\alpha.$$

Observe that $f_w = \sum_{n=1}^\infty f_w^n$ and

$$\sum_{n=1}^\infty \frac{1}{|w|} \int |f_w^n(x)|dx \lesssim \alpha.$$

We also let

$$g_w^n(x) = \chi_w(x)\frac{1}{|w|}\int_w f_w^n(y)dy,$$
$$b_w^n(x) = f_w^n(x) - g_w^n(x),$$

and

$$g^n(x) = \sum_w g_w^n(x), \qquad b^n(x) = \sum_w b_w^n(x).$$

Now

$$\sum_{n=1}^{\infty} |g_w^n(x)| \le \frac{1}{|w|}\int_w \sum_{n=1}^{\infty} |f_w^n(y)| dy\, \chi_w(x) \le \frac{1}{|w|}\int_w |f_w(y)| dy\, \chi_w(x) \lesssim \alpha; \tag{3.4}$$

moreover

$$\sum_{n=1}^{\infty} |g^n(x)| \lesssim \alpha \tag{3.5}$$

and

$$\sum_{n=1}^{\infty} [\|g_w^n\|_1 + \|b_w^n\|_1] \lesssim \int_w |f(x)| dx \lesssim \alpha|w|. \tag{3.6}$$

It will also be necessary to decompose the measure μ further. Let μ^n be the regularization defined in (2.22) and let

$$\mu_k^n(x) = 2^{-k\tau} \mu^n(2^{-kP}x).$$

For our basic decomposition of the singular Radon transform we set $f^n = \sum_w f_w^n$ and using $f = g + \sum_n f^n = g + \sum_n g^n + \sum_n b^n$ we split

$$\sum_{k\in\mathbb{Z}} \mu_k * f = H_{I,1} + H_{I,2} + H_{I,3} + H_b$$

where

$$\begin{aligned} H_{I,1} &= \sum_{k\in\mathbb{Z}} \mu_k * g \\ H_{I,2} &= \sum_{k\in\mathbb{Z}} \sum_{n\ge 1} (\mu_k - \mu_k^n) * f^n \\ H_{I,3} &= \sum_{k\in\mathbb{Z}} \sum_{n\ge 1} \mu_k^n * g^n \\ H_b &= \sum_{k\in\mathbb{Z}} \sum_{n\ge 1} \mu_k^n * b^n. \end{aligned} \tag{3.7}$$

A further decomposition is necessary for H_b. For given $n \ge 1$, $l \in \mathbb{Z}$ we define

$$\begin{aligned} I_l^n &= [ln, (l+1)n) \\ (I_l^n)^* &= [(l-1)n, (l+1+\frac{2}{a})n] \end{aligned} \tag{3.8}$$

and set

$$B_l^n = \sum_{w: r(w)\in I_l^n} b_w^n.$$

We split $H_b = H_{II} + H_{III}$ where

$$
(3.9)\qquad
\begin{aligned}
H_{II} &= \sum_{n\geq 1}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}\setminus(I_l^n)^*} \mu_k^n * B_l^n\\
H_{III} &= \sum_{n\geq 1}\sum_{l\in\mathbb{Z}}\sum_{k\in(I_l^n)^*} \mu_k^n * B_l^n.
\end{aligned}
$$

Note that H_{II} is the portion of H_b where the scaling of the measures μ_k^n is very different from the scaling of the balls w, which enables us to use standard L^1 arguments in the complement of the set Ω^*. The difficult term to estimate is H_{III}.

We shall show that

$$
(3.10.1)\qquad \sum_{i=1}^{3}\|H_{I,i}\|_2 \lesssim \alpha^{1/2}\|f\|_1^{1/2}
$$

$$
(3.10.2)\qquad \|H_{II}\|_{L^1(\mathbb{R}^d\setminus\Omega^*)} \lesssim \|f\|_1
$$

¿From (3.10.1/2) we get by Chebyshev's inequality

$$
(3.11)\qquad
\begin{aligned}
\operatorname{meas}\Big(\Big\{x : \sum_{i=1}^{3}|H_{I,i}(x)| > \alpha/10\Big\}\Big) &\lesssim \alpha^{-2}\Big\|\sum_{i=1}^{3}|H_{I,i}|\Big\|_2^2\\
&\lesssim \alpha^{-2}\Big[\sum_{i=1}^{3}\|H_{I,i}\|_2\Big]^2 \lesssim \alpha^{-1}\|f\|_1
\end{aligned}
$$

and

$$
(3.12)\qquad \operatorname{meas}\big(\{x\in\mathbb{R}^d\setminus\Omega^* : |H_{II}(x)| > \alpha/10\}\big) \lesssim \alpha^{-1}\|f\|_1.
$$

We now prove the L^2 bounds (3.10.1) using standard arguments. The cancellation of $\mu = \mu^0$ implies that $\widehat{\mu^0}(\xi) = O(|\xi|)$ and since μ_0 is smooth we get

$$
(3.13)\qquad |\widehat{\mu^0}(\xi)| \lesssim \min\{|\xi|, |\xi|^{-N}\}
$$

for large N.

Even without such a cancellation assumption the difference $\mu^n - \mu^{n-1}$ does have cancellation and using the decay assumption (3.1) on the Fourier transform of μ it is straightforward to check that for $m \geq 1$

$$
(3.14)\qquad |\widehat{\mu^m}(\xi) - \widehat{\mu^{m+1}}(\xi)| \lesssim 2^{-m\gamma}\min\{2^{-m}|\xi|, (2^{-m}|\xi|)^{-N}\}.
$$

Indeed the left hand side of (3.14) is $\lesssim (1+|\xi|)^{-\gamma}|\widehat{\phi}(2^{-m}\xi) - \widehat{\phi}(2^{-m-1}\xi)|$ and since $\widehat{\phi}(\eta) = 1 + O(|\eta|^d)$ we obtain the bound $2^{-m\gamma}(2^{-m}|\xi|)^{d-\gamma}$ which yields the claim for $|\xi| \leq 2^{m+1}$ since also $d-\gamma > 1$. For $|\xi| \geq 2^{m+1}$ we use that $|\widehat{\mu}^m(\xi)| \leq C_N|\xi|^{-\gamma}(1+2^{-m}|\xi|)^{-N}$.

Since $\widehat{\mu_k^n}(\xi) = \widehat{\mu^n}(\delta_k^*\xi)$ we obtain using (3.13), (3.14) that

$$
(3.15)\qquad
\begin{aligned}
&\sum_{k\in\mathbb{Z}}|\widehat{\mu_k^0}(\xi)| \lesssim 1\\
&\sum_{k\in\mathbb{Z}}|\widehat{\mu_k^m}(\xi) - \widehat{\mu_k^{m-1}}(\xi)| \lesssim 2^{-m\gamma}.
\end{aligned}
$$

We recover the well-known result that T is L^2 bounded, and as a consequence of the last displayed inequality we also get

$$\Big\|\sum_{k\in\mathbb{Z}}(\mu_k-\mu_k^n)*f\Big\|_2\lesssim\sum_{m=n}^{\infty}\Big\|\sum_{k\in\mathbb{Z}}(\mu_k^{m+1}-\mu_k^m)*f\Big\|_2\lesssim 2^{-n\gamma}\|f\|_2.$$

Now clearly

$$\|H_{I,1}\|_2^2=\Big\|\sum_{k\in\mathbb{Z}}\mu_k*g\Big\|_2^2\lesssim\|g\|_2^2\lesssim\alpha\|f\|_1$$

and using (3.13) and (3.14) we also obtain

$$\begin{aligned}\|H_{I,2}\|_2^2&\le\Big(\sum_{n\ge1}\Big\|\sum_{k\in\mathbb{Z}}(\mu_k-\mu_k^n)*f^n\Big\|_2\Big)^2\lesssim\Big(\sum_{n\ge1}2^{-n\gamma}\|f^n\|_2\Big)^2\\&\lesssim\sum_{n\ge1}2^{-n\gamma}\|f^n\|_2^2\lesssim\sum_{n\ge1}2^{-n\gamma}\|f^n\|_1 2^{c(n+1)}\alpha\lesssim\alpha\|f\|_1\end{aligned}$$

by our choice of c in (3.3). Moreover

$$\begin{aligned}\|H_{I,3}\|_2^2&=\Big\|\sum_{k\in\mathbb{Z}}\sum_{n\ge1}\big(\mu_k^0+\sum_{m=0}^{n-1}(\mu_k^{m+1}-\mu_k^m)*g^n\big)\Big\|_2^2\\&\le\Big(\Big\|\sum_{k\in\mathbb{Z}}\mu_k^0*\sum_{n\ge1}g^n\Big\|_2+\sum_{m=0}^{\infty}\Big\|\sum_{k\in\mathbb{Z}}(\mu_k^{m+1}-\mu_k^m)*\sum_{n>m}g^n\Big\|_2\Big)^2\\&\lesssim\Big(\sum_{m=0}^{\infty}2^{-m\gamma}\Big\|\sum_{n>m}g^n\Big\|_2\Big)^2\lesssim\alpha\|f\|_1.\end{aligned}$$

Finally we prove the L^1 bound (3.10.2). Suppose that $r(w)\in I_l^n$. For $k\ge\max(I_n^l)^*$ (thus $k-r(w)\ge 2n/a$) we use the cancellation of b_w^n and obtain with $y_w\in w$

$$\begin{aligned}\mu_k^n*b_w^n(x)&=\int 2^{-k\tau}\big[\mu^n(\delta_{-k}(x-y))-\mu^n(\delta_{-k}(x-y_w))\big]b_w^n(y)dy\\&=2^{-k\tau}\int\langle\delta_{-k}(y-y_w),\nabla\mu^n(\delta_{-k}(x-y_w+s(y-y_w)))\rangle b_w^n(y)dy\end{aligned}$$

and since $|\delta_{-k}(y-y_w)|\lesssim 2^{(r(w)-k)a}$ for $y\in w$ and $\|\nabla\mu^n\|_1=O(2^n)$ we get

$$\int|\mu_k^n*b_w^n(x)|dx\lesssim 2^n 2^{(r(w)-k)a}\|b_w^n\|_1.$$

Moreover notice that by our assumption that μ is supported in the unit ball we have that $\mu_k^n*b_w^n$ is supported in Ω^* if $k<\min(I_l^n)^*$.

Thus

$$\begin{aligned}\|H_{II}\|_{L^1(\mathbb{R}^d\setminus\Omega^*)}&\le\sum_{n\ge1}\sum_{l\in\mathbb{Z}}\sum_{k\ge\max(I_l^n)^*}\|\mu_k^n*B_l^n\|_1\\&\lesssim\sum_{n\ge1}\sum_{l\in\mathbb{Z}}\sum_{k\ge\max(I_l^n)^*}\sum_{r(w)\in I_l^n}2^n2^{(r(w)-k)a}\|b_w^n\|_1\\&\lesssim\sum_{n\ge1}\sum_{l\in\mathbb{Z}}2^{-n}\sum_{r(w)\in I_l^n}\|b_w^n\|_1\lesssim\|f\|_1,\end{aligned}$$

by the definition of $(I_l^n)^*$. Thus (3.10.2) is proved.

A decomposition similar to (3.7), (3.9) applies to the maximal operator where no cancellation on μ is assumed. We have

$$\sup_k |\mu_k * f| \le M_{I,1} + M_{I,2} + M_{I,3} + M_{II} + M_{III}$$

where

$$\begin{aligned} M_{I,1} &= \sup_{k\in\mathbb{Z}} |\mu_k * g| \\ M_{I,2} &= \sum_{n\ge 1} \sup_{k\in\mathbb{Z}} |(\mu_k - \mu_k^n) * f^n| \\ M_{I,3} &= \sum_{n\ge 1} \sup_{k\in\mathbb{Z}} |\mu_k^n * g^n| \\ M_{II} &= \sum_{n\ge 1}\sum_{l\in\mathbb{Z}} \sup_{k\in\mathbb{Z}\setminus(I_l^n)^*} |\mu_k^n * B_l^n| \\ M_{III} &= \sum_{n\ge 1}\sum_{l\in\mathbb{Z}} \sup_{k\in(I_l^n)^*} |\mu_k^n * B_l^n| \end{aligned} \tag{3.16}$$

Concerning the L^2 boundedness we observe that $\sup_k |\mu_k^0 * f|$ is pointwise controlled by the Hardy-Littlewood maximal function $M_{HL}f$, associated to the given dilation group. Therefore

$$\big\| \sup_k |\mu_k^0 * f| \big\|_2 \lesssim \|f\|_2. \tag{3.17}$$

Again by Fourier transform arguments as above

$$\begin{aligned} &\big\| \sup_k |(\mu_k^m - \mu_k^{m-1}) * f| \big\|_2 \lesssim \Big\| \Big(\sum_k |(\mu_k^m - \mu_k^{m-1}) * f|^2 \Big)^{1/2} \Big\|_2 \\ &\lesssim 2^{-m\gamma} \Big(\int \sum_k |\widehat{\mu_k^m}(\xi) - \widehat{\mu_k^{m-1}}(\xi)|^2 |\widehat{f}(\xi)|^2 d\xi \Big)^{1/2} \lesssim 2^{-m\gamma} \|\widehat{f}\|_2 \lesssim 2^{-m\gamma} \|f\|_2. \end{aligned}$$

This shows that we can repeat the arguments for H_I above and get

$$\sum_{i=1}^{3} \|M_{I,i}\|_2 \lesssim \alpha^{1/2} \|f\|_1^{1/2}. \tag{3.18}$$

In the definition of M_{II} we may replace the sup over $k \notin (I_l^n)^*$ by the sum and the estimation is exactly the same as for H_{II} above. This yields

$$\|M_{II}\|_{L^1(\mathbb{R}^d\setminus\Omega^*)} \lesssim \|f\|_1. \tag{3.19}$$

We combine these estimates with (3.2) and we see that in order to prove Theorems 1.1 and 1.2 we are left to prove the inequalities

$$\operatorname{meas}\{x : |M_{III}| > \tfrac{4}{5}\alpha\} \lesssim \int \frac{|f(x)|}{\alpha} \log\log(e^2 + \frac{|f(x)|}{\alpha}) dx \tag{3.20}$$

$$\operatorname{meas}\{x : |H_{III}| > \tfrac{4}{5}\alpha\} \lesssim \int \frac{|f(x)|}{\alpha} \log\log(e^2 + \frac{|f(x)|}{\alpha}) dx \tag{3.21}$$

This will be done in §5 and §6 below.

Weak type $L\log L$ estimates. We note that weak type $L\log L$ inequalities for T and $\mathcal{M}$ can be already obtained from trivial L^1 estimates for H_{III} and M_{III}. Here we are essentially reproving the result in [4].

Corollary 3.1. *Let μ be a compactly supported Borel measure satisfying*

$$|\widehat{\mu}(\xi)| \le C(1+|\xi|)^{-\gamma}.$$

Then M is of weak type $L\log L$. If in addition the cancellation condition $\int d\mu(x) = 0$ holds, then T is of weak type $L\log L$.

Proof. Given our previous estimates we just have to estimate the measure of the sets where $M_{III} > \alpha$ or $|H_{III}| > \alpha$. We simply use Chebyshev's inequality and are left with estimating $\alpha^{-1}\|M_{III}\|_1$ and $\alpha^{-1}\|H_{III}\|_1$, respectively. Using that the L^1 norm of μ_k^n is uniformly bounded in k, n we get

$$\begin{aligned}\|H_{III}\|_1 &\le \sum_{n\ge1}\sum_{l\in\mathbb{Z}}\sum_{k\in(I_l^n)^*}\|\mu_k^n * B_l^n\|_1 \lesssim \sum_{n\ge1}\sum_{l\in\mathbb{Z}}\sum_{k\in(I_l^n)^*}\sum_{r(w)\in I_l^n}\|b_w^n\|_1\\ &\lesssim \sum_{n\ge1}\sum_{l\in\mathbb{Z}}\sum_{r(w)\in I_l^n} n\|b_w^n\|_1 \lesssim \sum_{n\ge1} n\|f^n\|_1 \lesssim \int |f(x)|\log(e+\frac{|f(x)|}{\alpha})dx\end{aligned}$$

and the same argument applies to M_{III}. □

4. A stopping time argument

In order to refine the previous estimates for M_{III} and H_{III} we need a further decomposition of b_w^n. Here we use a stopping time argument based on length Λ_n (and thickness Θ_n). The reader will note some similarities with Christ's stopping time argument in [2].

In what follows $\mathfrak{Q}_0$ will denote the set of dyadic unit cubes of the form $(n_1, \dots, n_d) + [0,1)^d$, $n_i \in \mathbb{Z}$.

Proposition 4.1. *For every n and every w with $r(w) \in I_l^n$ there is a decomposition*

$$b_w^n = \sum_{\kappa\in(I_l^n)^*} f_w^{n,\kappa} \tag{4.1}$$

so that the following properties are satisfied.

(i)

$$\sum_{\kappa\in(I_l^n)^*} |f_w^{n,\kappa}| = |b_w^n|. \tag{4.2}$$

(ii) For every $q \in \mathfrak{Q}_0$, $\kappa \in (I_l^n)^*$

$$\Lambda_n\big[\sum_{r(w)<\kappa} f_w^{n,\kappa}(\delta_\kappa\cdot)\chi_q\big] \le \alpha^{-1}\sum_{r(w)<\kappa}\int_q |f_w^{n,\kappa}(\delta_\kappa y)|dy. \tag{4.3}$$

(iii) For every $q \in \mathfrak{Q}_0$, and for every $\kappa \in (I_l^n)^*$ and $s \ge 1$ with $\kappa + s \in (I_l^n)^*$,

$$\Theta_n\big[\sum_{r(w)\le\kappa} f_w^{n,\kappa}(\delta_{\kappa+s}\cdot)\chi_q\big] \le 16(n+1)\alpha. \tag{4.4}$$

Proof. This is proved by an inductive construction.

We shall give a decomposition of

$$\mathcal{G}^0 = \sum_{w:r(w)\in I_l^n} b_w^n;$$

since the w are disjoint this will yield a decomposition of each b_w^n. Set $\kappa_{n,l}^{\max} = \max(I_l^n)^*$ and $\kappa_j = \kappa_{n,l}^{\max} - j$. We shall establish the following

Claim. *For $N = 0, 1, \dots$ we can decompose*

$$\mathcal{G}^0 = \sum_{j=0}^{N} [H^j + S^j] + G^N$$

so that

(i) $G^{j-1} = H^j + S^j + G^j$ *if* $j \geq 1$

(ii) $G^j = \sum_{q \in \mathfrak{Q}_0} \sum_{\nu=1}^{L(j,Q)} G_\nu^{j,q}$, *where* $G_\nu^{j,q}$ *vanishes in the complement of* $\delta_{\kappa_j} q$ *and*

$$\Theta_n[G_\nu^{j,q}(\delta_{\kappa_j} \cdot)] \leq 8\alpha.$$

Moreover

$$L(j, Q) \leq n + 1.$$

(iii)

$$H^j(x) = 0 \qquad \text{if } x \notin \bigcup_{r(w)<\kappa_j} w$$

$$S^j(x) = 0 \qquad \text{if } x \notin \bigcup_{r(w)=\kappa_j} w$$

$$G^N(x) = 0 \qquad \text{if } x \notin \bigcup_{r(w)<\kappa_N} w.$$

(iv) For each $q \in \mathfrak{Q}_0$,

$$\Lambda_n\big[H^j(\delta_{\kappa_j} \cdot)\chi_q\big] \leq \alpha^{-1} \int_q |H^j(\delta_{\kappa_j} y)| dy.$$

(v) For $\kappa > \kappa_j$, $\kappa \in (I_l^n)^*$ *and each* $q \in \mathfrak{Q}_0$,

$$\Theta_n\big[H^j(\delta_\kappa \cdot)\chi_q\big] + \Theta_n\big[S^j(\delta_\kappa \cdot)\chi_q\big] \leq 16(n+1)\alpha.$$

(vi) The functions G^j, $G_\nu^{j,q}$, H^j, S^j *are nonnegative at* x *(nonpositive) if and only if* $f(x)$ *is nonnegative (nonpositive).*

If we accept the claim then in order to complete the proof of the proposition we observe that in the above statement $\kappa = \kappa_j = \kappa_{n,l}^{\max} - j$ and thus we merely have to define

$$f_w^{n,\kappa}(x) = \begin{cases} H^{\kappa_{n,l}^{\max} - \kappa}(x) & \text{if } x \in w, r(w) < \kappa \leq \kappa_{n,l}^{\max}, \\ S^{\kappa_{n,l}^{\max} - \kappa}(x) & \text{if } x \in w, r(w) = \kappa \\ 0 & \text{if } x \notin w \text{ or if } \kappa < r(w). \end{cases}$$

Then (4.1) follows from (iii) and (4.2) from (4.1) and (vi). (4.3) is a consequence of (iv) and (4.4) follows from (v).

Proof of the Claim. We argue by induction and assume that either $N = 0$ or that $N > 0$ and statements (i)-(vi) hold for all $j \leq N - 1$.

If $N = 0$ we set $S^0 = H^0 = 0$ and $G^0 = \mathcal{G}^0$. If $N \geq 1$ we begin by defining functions S^N, $\mathcal{G}^N$ where $S^N(x) = G^{N-1}(x)$ if $x \in \bigcup_{r(w)=\kappa_N} w$ and $S^N(x) = 0$ otherwise, and $\mathcal{G}^N(x) = G^{N-1}(x) - S^N(x)$. Thus $\mathcal{G}^N$ is supported on $\bigcup_{r(w)<\kappa_N} w$ and coincides with G^{N-1} there. Note that $\mathcal{G}^N$ vanishes if $\kappa_j < \min I_l^n$ and the construction stops then.

We now use Proposition 2.3 to decompose for $q \in \mathfrak{Q}_0$

$$\mathcal{G}^N(\delta_{\kappa_N} x)\chi_q(x) = \sum_{\nu=1}^{L} g_\nu^{N,q} + h_L^{N,q}$$

so that $\Theta_n[g_\nu^{N,q}]\Lambda_n[h_{\nu-1}^{N,q}] \le 8\int |g_\nu^{N,q}|dx$ and $h_L^{N,q}$ vanishes for $L \ge n+1$. Also the signs of the functions $g_\nu^{N,q}$, $h_L^{N,q}$ coincides with the sign of $\mathcal{G}^N(\delta_{\kappa_N}(x))\chi_q(x)$ and we have $h_{\nu-1}^{N,q} = g_\nu^{N,q} + h_\nu^{N,q}$ for $\nu \ge 1$ with $\Lambda_n[h_\nu^{N,q}] \le \Lambda_n[h_{\nu-1}^{N,q}]/2$.

Let $L(N,q)$ be the *minimal* integer L so that

$$\Lambda_n[h_L^{N,q}] \le \alpha^{-1} \int |h_L^{N,q}(y)|dy. \tag{4.5}$$

Then $L(N,q) \le n+1$ (since $h_L^{N,q}$ vanishes for $L \ge n+1$).

Now, $\Lambda_n[h_{\nu-1}^{N,q}] \ge \alpha^{-1}\int |h_{\nu-1}^{N,q}(y)|dy$ for $\nu \le L(N,q)$, by the minimality of $L(N,q)$, and since $|g_\nu^{N,q}| \le |h_{\nu-1}^{N,q}|$ we get

$$\Theta_n[g_\nu^{N,q}] \le 8\frac{\int |g_\nu^{N,q}(y)|dy}{\Lambda_n[h_{\nu-1}^{N,q}]} \le 8\alpha. \tag{4.6}$$

Now define $G_\nu^{N,q}(x) = g_\nu^{N,q}(\delta_{-\kappa_N} x)$, for $\nu \le L(N,q)$, and $G^N(x) = \sum_{q\in\mathfrak{Q}_0}\sum_{\nu=1}^{L(N,q)} G_\nu^{N,q}(x)$. Moreover $H^{N,q}(x) = h_{L(N,q)}^{N,q}(\delta_{-\kappa_N}x)$ and $H^N(x) = \sum_{q\in\mathfrak{Q}_0} H^{N,q}(x)$. Then the statement (vi) about the sign of $G_\nu^{N,q}$, G^N and H^N holds. (iv) follows from (4.5). Statements (i) and (iii) hold by construction, and the inequality for the thickness in (ii) holds by (4.6) by (4.6).

In view of (i), (vi) we also have $|H^N| + |S^N| \le |G^{N-1}| \le |G^{N-s}|$ for $s \ge 1$ so that by statement (ii) for $j \le N-1$ we get

$$\begin{aligned}
&\Theta_n[H^N(\delta_{\kappa_{N+s}}\cdot)\chi_q] + \Theta_n[S^N(\delta_{\kappa_{N+s}}\cdot)\chi_q] = \Theta_n[H^N(\delta_{\kappa_{N-s}}\cdot)\chi_q] + \Theta_n[S^N(\delta_{\kappa_{N-s}}\cdot)\chi_q]\\
&\le 2\Theta_n[G^{N-s}(\delta_{\kappa_{N-s}}\cdot)\chi_q] \le 16\sum_{\nu=1}^{L(N-s,q)} \Theta_n[G_\nu^{N-s,q}(\delta_{\kappa_{N-s}}\cdot)] \le 16L(N-s,q)\alpha \le 16(n+1)\alpha
\end{aligned}$$

This implies (v) for $j = N$ and the Claim is proved. $\square$

5. The main estimate for the maximal function

We shall prove the nontrivial estimate (3.20) for the maximal function, assuming again that the curvature assumption in the introduction is satisfied, and prove the inequality

$$\operatorname{meas}\Big(\Big\{x : \sup_k \Big|\sum_{\substack{n,l\\ k\in(I_l^n)^*}} \mu_k^n * B_l^n\Big| > \alpha\Big\}\Big) \le \int \Phi(|f|/\alpha)dx \tag{5.1}$$

with $\Phi(t) = t\log\log(e^2+t)$.

We use the decomposition in Proposition 4.1 and form an additional exceptional set $\mathcal{O}_1$. To define it we set for $q \in \mathfrak{Q}_0$, $\kappa \in (I_l^n)^*$,

$$F_q^{n,l,\kappa}(x) = \sum_{\substack{r(w)\in I_l^n\\ r(w)<\kappa}} f_w^{n,\kappa}(x)\chi_q(\delta_{-\kappa}x). \tag{5.2}$$

and define

$$\mathcal{O}_1 = \bigcup_{n=1}^{\infty} \bigcup_{l\in\mathbb{Z}} \bigcup_{\kappa\in(I_l^n)^*} \bigcup_{q\in\mathfrak{Q}_0} \bigcup_{\substack{k\in(I_l^n)^* \\ k\le\kappa}} \text{supp } \big(\mu_k^n * F_q^{n,l,\kappa}\big); \tag{5.3}$$

moreover we define

$$\mathcal{O} = \mathcal{O}_1 \cup \Omega^* \tag{5.4}$$

where Ω^* is as in (3.2).

To estimate the measure of $\mathcal{O}_1$ observe that supp $(\mu_k^n * F_q^{n,l,\kappa}) = \delta_k\text{supp } (\mu_0^n * [F_q^{n,l,\kappa}(\delta_k\cdot)])$ and since for $k\le\kappa$ the function $F_q^{n,l,\kappa}(\delta_k\cdot)$ is supported in a set of bounded diameter we get by (2.29) and (4.3)

$$\begin{aligned}
&\text{meas}\big(\text{supp } (\mu_k^n * F_q^{n,l,\kappa})\big) = 2^{k\tau}\text{meas}\big(\text{supp } (\mu_0^n * [F_q^{n,l,\kappa}(\delta_k\cdot)])\big) \\
&\lesssim 2^{k\tau}\Lambda_n[F_q^{n,l,\kappa}(\delta_k\cdot)] \lesssim 2^{k\tau}2^{(\kappa-k)A}\Lambda_n[F_q^{n,l,\kappa}(\delta_\kappa\cdot)] \\
&\lesssim 2^{k\tau}2^{(\kappa-k)A}\alpha^{-1}\int |F_q^{n,l,\kappa}(\delta_\kappa y)|dy \lesssim 2^{(k-\kappa)(\tau-A)}\alpha^{-1}\int |F_q^{n,l,\kappa}(y)|dy.
\end{aligned}$$

Thus, we can sum a geometric series in $k\le\kappa$ and obtain

$$\begin{aligned}
\text{meas}(\mathcal{O}_1) &\lesssim \sum_{n=1}^{\infty}\sum_{l\in\mathbb{Z}}\sum_{\kappa\in(I_l^n)^*}\sum_{q\in\mathfrak{Q}_0}\alpha^{-1}\int |F_q^{n,l,\kappa}(y)|dy \lesssim \sum_{n=1}^{\infty}\sum_{w}\alpha^{-1}\int |b_w^n(y)|dy \\
&\lesssim \alpha^{-1}\sum_{n=1}^{\infty}\sum_{w}\int |f_w^n(y)|dy \lesssim \alpha^{-1}\int |f(y)|dy
\end{aligned} \tag{5.5}$$

and the measure of $\mathcal{O} = \mathcal{O}_1 \cup \Omega^*$ satisfies the same estimate. Note that the contributions for $k\le\kappa$, $r(w)=\kappa$ are also supported in $\mathcal{O}$ since μ is assumed to be supported in the unit ball and thus

$$\bigcup_{n=1}^{\infty}\bigcup_{l\in\mathbb{Z}}\bigcup_{w:r(w)\in I_l^n}\bigcup_{k\le r(w)} \text{supp } \big(\mu_k^n * f_w^{n,r(w)}\big) \subset \Omega^*.$$

It now remains to handle the contribution in the complement of $\mathcal{O}$ which only involves the scales $k>\kappa$ and contributions for $r(w)\in I_l^n$ with $r(w)\le\kappa$; to simplify the notation below we set

$$I_l^{n,\kappa} = \{r\in I_l^n : r\le\kappa\}.$$

We shall first cut out a contribution from 'flat' parts of Σ. We recall that the curvature does not vanish to infinite order on Σ and therefore there is a number $\eta>0$ such that

$$\int_\Sigma |K(x)|^{-\eta}d\sigma(x) < \infty. \tag{5.6}$$

This is well known (for example, one may use an argument in [16, p.343] to reduce to an inequality in one dimension where one can use Hölder's inequality and compactness).

By Chebyshev's inequality (5.6) implies that

$$|\{x\in\Sigma : |K(x)|\le n^{-3/\eta}\}| \lesssim n^{-3}. \tag{5.7}$$

Now we use a partition of unity to write

$$\mu = \sum_{i\in\mathcal{J}^n} \nu^{i,n}$$

where each $\nu^{i,n}$ is supported on a cube R_i of diameter $\varepsilon_1 n^{-3/\eta}$ (here ε_1 will be as in Lemma 2.6) and the supports of the $\nu^{i,n}$ have bounded overlap, independent of n. Note that then

$$\operatorname{card}(\mathcal{J}^n) \lesssim n^{3(d-1)/\eta}. \tag{5.8}$$

We split the index set into disjoint subsets as $\mathcal{J}^n = \mathcal{J}_1^n \cup \mathcal{J}_2^n$ where $\mathcal{J}_2^n$ consists of all $i \in \mathcal{J}$ with the property that $|K(x')| \le n^{-3/\eta}$ for all $x' \in \operatorname{supp} R_i$.

Then by (5.7) we have that the sum of the total variations of the $\nu^{i,n}$, for which $i \in \mathcal{J}_2^n$, satisfies the bound

$$\sum_{i\in\mathcal{J}_2^n} \|\nu^{i,n}\| \lesssim n^{-3}.$$

Let

$$\mu^{i,n} = \nu^{i,n} * \phi_n$$

and $\mu_k^{i,n} = 2^{-k\tau}\mu^{i,n}(2^{-kP}\cdot)$.

Since the cardinality of $(I_l^n)^*$ is $O(n)$ and $\sum_{i\in\mathcal{J}_2^n} \|\mu_k^{i,n}\|_1 = O(n^{-3})$ the contribution of the measures $\sum_{i\in\mathcal{J}_2^n} \mu_k^{i,n}$, $k \in (I_l^n)^*$ can be handled by a straightforward L^1 estimate:

$$\begin{aligned}
&\operatorname{meas}\Big(\Big\{x : \sup_k \Big| \sum_{\substack{n,l\\ k\in(I_l^n)^*}} \sum_{\substack{\kappa\in(I_l^n)^*\\ \kappa\le k}} \sum_{i\in\mathcal{J}_2^n} |\mu_k^{i,n} * \sum_{r(w)\in I_l^{n,\kappa}} f_w^{n,\kappa}\Big| > \alpha/10\Big\}\Big)\\
&\lesssim \alpha^{-1}\Big\|\sum_{n,l}\sum_{k\in(I_l^n)^*}\sum_{\substack{\kappa\in(I_l^n)^*\\ \kappa\le k}}\sum_{i\in\mathcal{J}_2^n} |\mu_k^{i,n} * \sum_{r(w)\in I_l^{n,\kappa}} f_w^{n,\kappa}|\Big\|_1\\
&\lesssim \alpha^{-1}\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{k\in(I_l^n)^*}\sum_{i\in\mathcal{J}_2^n} \big\|\mu_k^{i,n}\big\|_1 \sum_{r(w)\in I_l^{n,\kappa}} \|f_w^{n,\kappa}\|_1\\
&\lesssim \alpha^{-1}\sum_{n,l}\sum_{\kappa\in(I_l^n)^*} n^{-2} \sum_{r(w)\in I_l^{n,\kappa}} \|f_w^{n,\kappa}\|_1 \lesssim \alpha^{-1}\|f\|_1.
\end{aligned} \tag{5.9}$$

Next choose a large constant C_0; specifically the choice

$$C_0 \ge \frac{100}{a}(1+\frac{d}{\eta})\max\{1, \frac{A}{\tau - A}\} + 10 + \log_2\big(\frac{C_1}{c_1}\big) \tag{5.10}$$

will work where $c_1 \le C_1$ are as in (2.27). Then the contribution for the scales $\kappa \le k \le \kappa + C_0 \log n$ is also handled by an L^1 estimate:

$$\begin{aligned}
&\operatorname{meas}\Big(\Big\{x:\sup_k\Big|\sum_{\substack{n,l\\ k\in(I_l^n)^*}}\sum_{\substack{\kappa\in(I_l^n)^*\\ \kappa\le k\le\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\mu_k^{i,n}*\sum_{r(w)\in I_l^{n,\kappa}}f_w^{n,\kappa}(x)\Big|>\alpha/10\Big\}\Big)\\
&\lesssim\alpha^{-1}\Big\|\sum_{n,l}\sum_{k\in(I_l^n)^*}\sum_{\substack{\kappa\in(I_l^n)^*\\ \kappa\le k\le\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}|\mu_k^{i,n}*\sum_{r(w)\in I_l^{n,\kappa}}f_w^{n,\kappa}|\Big\|_1\\
&\lesssim\alpha^{-1}\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{w:r(w)\in I_l^{n,\kappa}}\sum_{\kappa\le k\le\kappa+C_0\log n}\|f_w^{n,\kappa}\|_1\\
&\lesssim\alpha^{-1}\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{r(w)\in I_l^{n,\kappa}}\log n\|f_w^{n,\kappa}\|_1\\
&\lesssim\alpha^{-1}\sum_n\log n\|f^n\|_1\lesssim\int\frac{|f(x)|}{\alpha}\log\log\big(e^2+\frac{|f(x)|}{\alpha}\big)dx
\end{aligned}\tag{5.11}$$

It remains to show

$$\operatorname{meas}\Big(\Big\{x:\sup_k\Big|\sum_{\substack{n,l\\ k\in(I_l^n)^*}}\sum_{\substack{\kappa\in(I_l^n)^*\\ k>\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\mu_k^{i,n}*\sum_{r(w)\in I_l^{n,\kappa}}f_w^{n,\kappa}(x)\Big|>\alpha/10\Big\}\Big)\lesssim\alpha^{-1}\|f\|_1\tag{5.12}$$

and this will be accomplished by proving L^2 estimates.

Reintroducing cancellation. The decomposition in (4.1) was needed to exploit the geometry of the exceptional set; however we paid the price of destroying the cancellation properties of the b_w^n. As the information on the support of the $f_w^{n,\kappa}$ has been used and is not needed anymore for the scales $k>\kappa+C_0\log n$ we shall now modify the functions $f_w^{n,\kappa}$ to reintroduce some cancellation. Namely let $\{P_i\}_{i=1}^{M_d}$ be an orthonormal basis of the space of polynomials of degree $\le d$ on the unit ball $\{x:|x|\le1\}$ and for given w define the projection operator Π_w by

$$\Pi_w[h](x)=\chi_w(x)\sum_{i=1}^{M_d}P_i(\delta_{-r(w)}(x-x_w))\int_w h(y)P_i(\delta_{-r(w)}(y-x_w))2^{-r(w)\tau}\,dy.$$

Note that

$$\big|\Pi_w[h](x)\big|\le C\frac{1}{|w|}\int_w|h(y)|dy\tag{5.13}$$

where C is independent of h and w.

Let

$$\begin{aligned}
g_w^{n,\kappa}(x)&=\Pi_w[f_w^{n,\kappa}](x),\\
b_w^{n,\kappa}(x)&=f_w^{n,\kappa}(x)-g_w^{n,\kappa}(x),
\end{aligned}$$

so that $b_w^{n,\kappa}$ vanishes off w and for polynomials p

$$\int_w b_w^{n,\kappa}(x)p(x)dx=0\quad\text{ if }\deg(p)\le d.\tag{5.14}$$

We observe that since the w's are generalized Whitney cubes for Ω (see §3), we have

$$\sum_{n,\kappa}\big|\Pi_w[f_w^{n,\kappa}](x)\big|\lesssim\chi_w(x)\frac{1}{|w|}\int_w|f(x)|dx\lesssim\alpha;\tag{5.15}$$

moreover by (5.13)

$$\sum_{n,\kappa}\left[\|b_w^{n,\kappa}\|_1+\|g_w^{n,\kappa}\|_1\right]\lesssim\sum_{n,\kappa}\|f_w^{n,\kappa}\|_1\lesssim\int_w|f(x)|dx. \tag{5.16}$$

Now (5.12) will follow from

$$\Big\|\sup_k\Big|\sum_{\substack{n,l\\ k\in(I_l^n)^*}}\sum_{\substack{\kappa\in(I_l^n)^*\\ k>\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\mu_k^{i,n}*\sum_{r(w)\in I_l^{n,\kappa}}g_w^{n,\kappa}\Big|\Big\|_2^2\lesssim\alpha\|f\|_1 \tag{5.17}$$

$$\Big\|\sup_k\Big|\sum_{\substack{n,l\\ k\in(I_l^n)^*}}\sum_{\substack{\kappa\in(I_l^n)^*\\ k>\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\mu_k^{i,n}*\sum_{r(w)\in I_l^{n,\kappa}}b_w^{n,\kappa}\Big|\Big\|_2^2\lesssim\alpha\|f\|_1. \tag{5.18}$$

The estimation (5.17) is straightforward. If $d\sigma$ denotes surface measure on Σ and $d\sigma_k$ the dilate $2^{-k\tau}d\sigma(\delta_{-k}\cdot)$ then the maximal function

$$Mf(x)=\sup_{k\in\mathbb{Z}}|d\sigma_k*f|$$

defines a bounded operator on L^2. By the positivity of this maximal operator the left side of (5.17) is bounded by a constant times

$$\Big\|M_{HL}M\Big[\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_w|g_w^{n,\kappa}|\Big]\Big\|_2^2\lesssim\alpha\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_w\|g_w^{n,\kappa}\|_1\lesssim\alpha\|f\|_1;$$

here we used (5.15/16).

For the remainder of this section we prove (5.18).

We first replace the sup in k by an ℓ^2 sum and then, for fixed k, we apply Schwarz' inequality in the form $[\sum_n|a_n|]^2\lesssim\sum|na_n|^2$. Next we observe that for fixed n the number k is contained in at most $3+2/a$ of the intervals $(I_l^n)^*$. Then we apply Schwarz' inequality for the sim in κ yielding a factor of $O(n)$ and for the sum in i yielding a factor of $O(n^{3(d-1)/\eta})$. Finally we group the sum over w into groups for which $r(w)=r$, $r\in I_l^n$ and apply Schwarz' inequality in r which yields one more factor of $O(n)$. Thus we see that the left side of (5.18) is dominated by a constant times

$$\sum_{\substack{k,n,l\\ k\in(I_l^n)^*}}\sum_{\substack{\kappa:\kappa<\\ k-C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\sum_{r\in I_l^{n,\kappa}}n^{(4+\frac{3(d-1)}{\eta})}\Big\|\mu_k^{i,n}*\sum_{r(w)=r}b_w^{n,\kappa}\Big\|_2^2 \tag{5.20}$$

We note that the some of the applications of Schwarz' inequality above are not really necessary but it turns out that the polynomial factors in n are irrelevant in the range $\kappa<k-C_0\log n$.

Now, for fixed κ,k, define

$$M(\kappa,k)=\left[k-(k-\kappa)\frac{a}{2A}+\log_2\frac{C_1}{c_1}+2\right] \tag{5.21}$$

where $[v]$ denotes the largest integer $\leq v$. Note that for $\kappa<k-C_0\log n$ we have $M(\kappa,k)<k$. Let $\mathfrak{R}(\kappa,k)$ be the collection of dilates $\delta_{M(\kappa,k)}q$, where $q\in\mathfrak{Q}_0$. For each w with $r(w)=r\leq\kappa$ we assign $R\in\mathfrak{R}(\kappa,k)$ so that $w\cap R\neq\emptyset$. We write $R=R_{\kappa,k}(w)$ or simply $R=R(w)$ if the dependence on k,κ is clear.

Let $\widetilde{\mathfrak{R}}(\kappa,k)$ be a subcollection of $\mathfrak{R}(\kappa,k)$ with the property that if $R,R'\in\mathfrak{R}(\kappa,k)$, $R\neq R'$ and $R=\delta_{M(\kappa,k)}q$, $R'=\delta_{M(\kappa,k)}q'$ then $\operatorname{dist}(q,q')>10$.

We shall show for fixed n, l, $k\in(I_l^n)^*$, $\kappa\in(I_l^n)^*$, $r\in I_l^{n,\kappa}$ that

$$\Big\|\sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)}\mu_k^{i,n}*\sum_{\substack{r(w)=r\\ R_{\kappa,k}(w)=R}}b_w^{n,\kappa}\Big\|_2^2\lesssim n^{2+3(d+3)/\eta}2^{-(k-\kappa)c_0}\alpha\sum_{r(w)=r}\|b_w^{n,\kappa}\|_1 \tag{5.22}$$

where

$$c_0=\frac{a}{2}\min\{1,\frac{\tau-A}{A}\} \tag{5.23}$$

Given (5.22), the proof of (5.18) is a quick consequence. First note that $\mathfrak{R}(\kappa,k)$ can be split into $O(10^d)$ families of type $\widetilde{\mathfrak{R}}(\kappa,k)$. Thus Minkowski's inequality and (5.22) imply that (5.22) holds also with $\widetilde{\mathfrak{R}}(\kappa,k)$ replaced by $\mathfrak{R}(\kappa,k)$. Then we obtain from (5.20) and the modified (5.22) that the left side of (5.18) is controlled by

$$\sum_{n,l}\sum_{k\in(I_l^n)^*}\sum_{\substack{\kappa\in(I_l^n)^*:\\ \kappa<k-C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\sum_{r\in I_l^{n,\kappa}}n^{6(1+\frac{d-1}{\eta})}2^{-(k-\kappa)c_0}\alpha\sum_{r(w)=r}\|b_w^{n,\kappa}\|_1$$
$$\lesssim\sum_{n,l,\kappa}n^{6+\frac{9(d-1)}{\eta}}\sum_{k\geq\kappa+C_0\log n}2^{-(k-\kappa)c_0}\alpha\sum_{\substack{r\in I_l^n\\ r<\kappa}}\sum_{r(w)=r}\|b_w^{n,\kappa}\|_1.$$

Now we sum the geometric series

$$\sum_{k\geq\kappa+C_0\log n}2^{-(k-\kappa)c_0}\lesssim n^{-c_0C_0}$$

and using (5.23) and our choice of C_0 in (5.10) we observe that $n^{-c_0C_0}\leq n^{-50(1+d/\eta)}$; this yields that the left side of (5.18) is controlled by

$$\alpha\sum_{n,l,\kappa}\sum_{r\in I_l^{n,\kappa}}\sum_{r(w)=r}\|b_w^{n,\kappa}\|_1\lesssim\alpha\|f\|_1.$$

Thus the proof will be finished when inequality (5.22) is verified.

Proof of (5.22).

We split for fixed $n,l,\ k,\kappa\in(I_l^n)^*$, $i\in\mathcal{J}_1^n$ and $r\in I_l^{n,\kappa}$,

$$\Big\|\sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)}\mu_k^{i,n}*\sum_{\substack{r(w)=r\\ R(w)=R}}b_w^{n,\kappa}\Big\|_2^2=I+II$$

where

$$I=\sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)}\int\widetilde{\mu_k^{i,n}}*\mu_k^{i,n}*\sum_{\substack{r(w)=r\\ R(w)=R}}b_w^{n,\kappa}(x)\sum_{\substack{r(w')=r\\ R(w')=R}}\overline{b_{w'}^{n,\kappa}(x)}\,dx \tag{5.24}$$

$$II=\sum_{\substack{R,R'\in\widetilde{\mathfrak{R}}(\kappa,k)\\ R\neq R'}}\int\widetilde{\mu_k^{i,n}}*\mu_k^{i,n}*\sum_{\substack{r(w)=r\\ R(w)=R}}b_w^{n,\kappa}(x)\sum_{\substack{r(w')=r\\ R(w')=R'}}\overline{b_{w'}^{n,\kappa}(x)}dx. \tag{5.25}$$

We shall first estimate II. Fix w, w' occuring in the expression (5.25). Then using the cancellation of the $b_w^{n,\kappa}$ we get

$$
\begin{aligned}
&|\widetilde{\mu_k^{i,n}} * \mu_k^{i,n} * b_w^{n,\kappa}(x)| \\
&= \Big| \int 2^{-k\tau} \Big[\widetilde{\mu_0^{i,n}} * \mu_0^{i,n}(\delta_{-k}(x-y)) - \sum_{j=0}^{d-1} \frac{1}{j!} \langle \delta_{-k}(y-x_w), \nabla\rangle^j \widetilde{\mu_0^{i,n}} * \mu_0^{i,n}(\delta_{-k}(x-x_w))\Big] b_w^{n,\kappa}(y)\, dy \\
&= \Big| \int_0^1 \frac{(1-s)^{d-1}}{(d-1)!} \int 2^{-k\tau} \langle \delta_{-k}(y-x_w), \nabla\rangle^d \widetilde{\mu_0^{i,n}} * \mu_0^{i,n}(\delta_{-k}(x-x_w+sx_w-sy))\Big] b_w^{n,\kappa}(y)\, dyds \Big|
\end{aligned}
$$

(5.26)

$$
\lesssim n^{\frac{3}{\eta}(2d-(d-3))} \int_0^1 \int_w \frac{2^{-k\tau}|\delta_{-k}(y-x_w)|^d}{|\delta_{-k}(x-x_w+sx_w-sy)|^{d+1}} |b_w^{n,\kappa}(y)| dyds
$$

by Lemma 2.5 applied to the measure $\mu_0^{i,n}$, with $\beta = n^{-3/\eta}$.

Now if $x \in w'$, $y \in w$ with $w' \cap R' \neq \emptyset$, $w \cap R \neq \emptyset$, and if $R \neq R'$ then by the separation property of the sets in $\widetilde{\mathfrak{R}}(\kappa, k)$

$$
|\delta_{-k}(x-x_w)| \geq c_1 2^{(M_{\kappa,k}-k)A} |\delta_{-M_{\kappa,k}}(x-x_w)| \geq 10 c_1 2^{(M_{\kappa,k}-k)A} \geq 5 C_1 2^{-(k-\kappa)a/2} \tag{5.27}
$$

while

$$
|\delta_{-k}(y-x_w)| + |\delta_{-k}(x-x_{w'})| \leq 2C_1 2^{-a(k-r)} \leq 2C_1 2^{-a(k-\kappa)}.
$$

Thus for $x \in w'$ we may replace $|\delta_{-k}(x-x_w+sx_w-sy)|$ in the denominator of (5.26) by $|\delta_{-k}(x'_w - x_w)|$. We also take into account that $\|b_{w'}^{n,\kappa}\|_1 \lesssim \alpha|w'|$ and thus obtain the bound

$$
II \lesssim n^{3(d+3)/\eta} \sum_{R \in \widetilde{\mathfrak{R}}(\kappa,k)} \sum_{\substack{R(w)=R \\ r(w)=r}} \|b_w^{n,\kappa}\|_1 \sum_{\substack{R' \in \widetilde{\mathfrak{R}}(\kappa,k) \\ R \neq R'}} \sum_{\substack{R(w')=R' \\ r(w')=r}} \alpha |w'| \frac{2^{-k\tau} 2^{-(k-r)ad}}{|\delta_{-k}(x_{w'}-x_w)|^{d+1}}. \tag{5.28}
$$

Now we calculate using (5.27)

$$
\begin{aligned}
&\sum_{\substack{R' \in \widetilde{\mathfrak{R}}(\kappa,k) \\ R \neq R'}} \sum_{\substack{R(w')=R' \\ r(w')=r}} |w'| \frac{2^{-k\tau} 2^{-(k-r)ad}}{|\delta_{-k}(x_{w'}-x_w)|^{d+1}} \lesssim 2^{-(k-r)ad} \sum_{\substack{R' \in \widetilde{\mathfrak{R}}(\kappa,k) \\ R \neq R'}} \sum_{\substack{R(w')=R' \\ r(w')=r}} \int_{\delta_{-k}(-x_w+w')} |u|^{-d-1} du \\
&\lesssim 2^{-(k-\kappa)ad} \int_{|u| \geq 2^{-(k-\kappa)a/2}} |u|^{-d-1} du \lesssim 2^{-\frac{a}{2}(k-\kappa)(2d-1)}.
\end{aligned}
$$

Combining this with (5.28) yields the bound

$$
II \lesssim n^{3\frac{d+3}{\eta}} 2^{-\frac{a}{2}(k-\kappa)(2d-1)} \sum_{R \in \widetilde{\mathfrak{R}}(\kappa,k)} \sum_{\substack{R(w)=R \\ r(w)=r}} \|b_w^{n,\kappa}\|_1 \tag{5.29}
$$

which is controlled by the right hand side of (5.22).

We now estimate the contribution I. Unfortunately, in introducing the cancellation and passing from $f_w^{n,\kappa}$ to $b_w^{n,\kappa}$ we have obscured the geometrical information on the thickness of $f_w^{n,\kappa}$. As the cancellation is not needed anymore for I we (partially) undo it and estimate

$$
I \leq I_1 + I_2
$$

where

$$I_1 = \sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)} \int \widetilde{\mu_k^{i,n}} * \mu_k^{i,n} * \sum_{\substack{r(w)=r\\ R(w)=R}} f_w^{n,\kappa}(x) \sum_{\substack{r(w')=r\\ R(w')=R}} \overline{b_{w'}^{n,\kappa}(x)}\, dx \tag{5.30}$$

$$I_2 = \sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)} \int \widetilde{\mu_k^{i,n}} * \mu_k^{i,n} * \sum_{\substack{r(w)=r\\ R(w)=R}} g_w^{n,\kappa}(x) \sum_{\substack{r(w')=r\\ R(w')=R}} \overline{b_{w'}^{n,\kappa}(x)}\, dx. \tag{5.31}$$

Since $|g_w^{n,\kappa}(x)| \lesssim \alpha\chi_w(x)$ we get

$$|I_2| \lesssim n^{-3\frac{d-3}{\eta}}\alpha \sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)} \sum_{\substack{r(w)=r\\ R(w)=R}} \sum_{\substack{r(w')=r\\ R(w')=R}} \int_{w'}\int_{w} \frac{2^{-k\tau}}{|\delta_{-k}(x-y)|} dy |b_{w'}^{n,\kappa}(x)|\, dx \tag{5.32}$$

and

$$\sum_{\substack{r(w)=r\\ R(w)=R}} \int_w \frac{2^{-k\tau}}{|\delta_{-k}(x-y)|} dy \lesssim 2^{-(k-M(\kappa,k))(\tau-A)} \int_R \frac{2^{-M(\kappa,k)\tau}}{|\delta_{-M(\kappa,k)}(x-y)|} dy$$
$$\lesssim 2^{-(k-\kappa)(\tau-A)\frac{a}{2A}} \int_{|u|\lesssim 1} |u|^{-1} du \lesssim 2^{-(k-\kappa)(\tau-A)\frac{a}{2A}}.$$

Thus

$$|I_2| \lesssim n^{-3\frac{d-3}{\eta}} \alpha 2^{-(k-\kappa)(\tau-A)\frac{a}{2A}} \sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)} \sum_{\substack{r(w')=r\\ R(w')=R}} \|b_{w'}^{n,\kappa}\|_1. \tag{5.33}$$

Finally for the main term I_1 we use Lemma 2.6, then (2.28) and then part (iii) of Proposition 4.1 to bound

$$\Big|\widetilde{\mu_k^{i,n}} * \mu_k^{i,n} * \sum_{\substack{r(w)=r\\ R(w)=R}} f_w^{n,\kappa}(x)\Big|$$
$$= \Big| \int \widetilde{\mu_0^{i,n}} * \mu_0^{i,n}(\delta_{-k}x - y) \sum_{\substack{r(w)=r\\ R(w)=R}} f_w^{n,\kappa}(\delta_k y) dy\Big|$$
$$\lesssim n^{1-3(d-3)/\eta}\Theta_n\big[\sum_{\substack{r(w)=r\\ R(w)=R}} f_w^{n,\kappa}(\delta_k\cdot)\big]$$
$$\lesssim n^{1-3(d-3)/\eta} 2^{-(k-M(\kappa,k))(\tau-A)}\Theta_n\big[\sum_{\substack{r(w)=r\\ R(w)=R}} f_w^{n,\kappa}(\delta_{M(\kappa,k)}\cdot)\big]$$
$$\lesssim n^{2-3(d-3)/\eta} 2^{-(k-M(\kappa,k))(\tau-A)}\alpha.$$

Since $k - M(\kappa,k) \geq (k-\kappa)a/2A$ we obtain

$$|I_1| \lesssim \alpha 2^{-(k-\kappa)(\tau-A)\frac{a}{2A}} n^{2-3\frac{d-3}{\eta}} \sum_{R\in\widetilde{\mathfrak{R}}(\kappa,k)} \sum_{\substack{r(w')=r\\ R(w')=R}} \|b_{w'}^{n,\kappa}\|_1. \tag{5.34}$$

(5.33/34) and (5.29) certainly imply (5.22). This concludes the proof of Theorem 1.1. □

Remark. The above argument also applies to maximal functions associated to certain surfaces with low codimension, for example if we assume that for every normal vector the Gaussian curvature is bounded away from zero. In this case we have to work with the notions $\Lambda_{n,\beta}$, $\Theta_{n,\beta}$ in the remark following the proof of Proposition 2.1; here β is the codimension. The condition about nonvanishing Gaussian curvature is never satisfied for manifolds with high codimension such as curves in three or more dimensions. In those cases it is presently open whether the weak type $L\log L$ inequality of Corollary 3.1 above can be improved.

6. Estimates for the singular integral operators

The proof of the weak type $L\log\log L$ estimate for the singular Radon transforms relies to a large extent on the same arguments as for the maximal operator. We shall just indicate the necessary modifications.

We need to prove inequality (3.21). The definition of the exceptional set $\mathcal{O}$ and estimate (5.5) remains the same. Thus we are left to show (again with $\Phi(s) = s\log\log(e^2+s)$)

$$\text{meas}\Big(\Big\{x : \Big|\sum_{n,l}\sum_{k\in(I_l^n)^*}\mu_k^n * \sum_{\substack{\kappa\in(I_l^n)^*\\ \kappa\le k}}\sum_{r(w)\in I_l^{n,\kappa}} f_w^{n,\kappa}\Big| > \frac{4}{5}\alpha\Big\}\Big) \lesssim \int \Phi\Big(\frac{|f(x)|}{\alpha}\Big)dx. \tag{6.1}$$

Now, as in §5, we wish to decompose the measure into a part with curvature and a part with flatness (with the splitting depending on n). Some care is needed now since we need to preserve the cancellation of the measure when acting on the α-bounded contributions. Before doing this decomposition we shall reverse the order of the steps (5.9), (5.11) and first get an analogue of (5.11) for the functions μ_k^n. Indeed since $\|\mu_k^n\|_1 = O(1)$ the argument for (5.11) yields

$$\text{meas}\Big(\Big\{x : \Big|\sum_{n,l,\kappa}\sum_{\substack{k\in(I_l^n)^*\\ \kappa\le k\le\kappa+C_0\log n}}\mu_k^n * \sum_{r(w)\in I_l^{n,\kappa}} f_w^{n,\kappa}\Big| > \frac{\alpha}{10}\Big\}\Big) \lesssim \int \Phi\Big(\frac{|f(x)|}{\alpha}\Big)dx \tag{6.2}$$

and therefore we have to bound

$$\text{meas}\Big(\Big\{x : \Big|\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\ge\kappa+C_0\log n}}\mu_k^n * \sum_{r(w)\in I_l^{n,\kappa}} f_w^{n,\kappa}\Big| > \frac{7}{10}\alpha\Big\}\Big). \tag{6.3}$$

As before we split $f_w^{n,\kappa} = g_w^{n,\kappa} + b_w^{n,\kappa}$ and we first show that

$$\text{meas}\Big(\Big\{x : \Big|\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\ge\kappa+C_0\log n}}\mu_k^n * \sum_{r(w)\in I_l^{n,\kappa}} g_w^{n,\kappa}\Big| > \frac{\alpha}{10}\Big\}\Big) \lesssim \frac{\|f\|_1}{\alpha}. \tag{6.4}$$

We use the nonisotropic version of an inequality in [6, p. 548] for the maximal version of the singular integral, namely we have

$$\Big\|\sup_{K_1,K_2}\Big|\sum_{k=K_1}^{K_2}\widetilde{\mu}_k * u\Big|\Big\|_2 \lesssim \|u\|_2. \tag{6.5}$$

Here $\widetilde{\mu}_k$ is the reflection of μ_k. Indeed for (6.5) one just needs $|\widehat{\mu}(\xi)| \le \min\{|\xi|, |\xi|^{-\gamma}\}$ for some $\gamma > 0$ (*cf.* (3.15)). In order to use (6.5) we have to split $\mu_k^n = \mu_k - (\mu_k - \mu_k^n)$.

¿From (6.5) and (5.17) we get

$$\begin{aligned}
&\Big\| \sum_{n,l} \sum_{\kappa \in (I_l^n)^*} \sum_{\substack{k \in (I_l^n)^* \\ k \geq \kappa + C_0 \log n}} \mu_k * \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa} \Big\|_2^2 \\
&= \sup_{\|u\|_2 \leq 1} \Big| \int \sum_{n,l} \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa}(y) \sum_{\substack{k \in (I_l^n)^* \\ k \geq \kappa + C_0 \log n}} \widetilde{\mu}_k * u(y) dy \Big|^2 \\
&\leq \sup_{\|u\|_2 \leq 1} \Big\| \sup_{K_1, K_2} \Big| \sum_{k=K_1}^{K_2} \widetilde{\mu}_k * u \Big| \Big\|_2^2 \Big\| \sum_{n,l} \sum_{\kappa \in (I_l^n)^*} \Big| \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa} \Big| \Big\|_2^2 \\
&\lesssim \Big\| \sum_{n,l} \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} |g_w^{n,\kappa}| \Big\|_2^2 . \qquad (6.6)
\end{aligned}$$

For each w and $x \in w$ we have

$$\sum_{n,l} \sum_{\substack{\kappa \in (I_l^n)^* \\ r(w) \in I_l^{n,\kappa}}} |g_w^{n,\kappa}(x)| \lesssim \sum_{n,l} \sum_{\substack{\kappa \in (I_l^n)^* \\ r(w) \in I_l^{n,\kappa}}} \frac{1}{|w|} \int_w |f_w^{n,\kappa}(x)| dx \lesssim \frac{1}{|w|} \int_w |f(x)| dx \lesssim \alpha$$

and in view of the disjointness of the sets w the expression (6.6) is controlled by

$$\sum_w \Big\| \sum_{n,l} \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} |g_w^{n,\kappa}| \Big\|_2^2 \lesssim \alpha \|f\|_1. \qquad (6.7)$$

Moreover for fixed n, and $m \geq n$ we get using (3.15)

$$\begin{aligned}
&\Big\| \sum_l \sum_{k \in (I_l^n)^*} (\mu_k^m - \mu_k^{m+1}) * \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa} \Big\|_2 \\
&\lesssim 2^{-m\gamma} \sum_l \sum_{k \in (I_l^n)^*} \Big\| \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa} \Big\|_2^2
\end{aligned}$$

and thus using the telescoping sum $\mu_k^n - \mu_k = \sum_{m=n}^{\infty} (\mu_k^m - \mu_k^{m+1})$ we obtain

$$\begin{aligned}
&\Big\| \sum_{n,l} \sum_{k \in (I_l^n)^*} (\mu_k^n - \mu_k) * \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa} \Big\|_2 \\
&\lesssim \Big[\sum_n \sum_{m=n}^{\infty} 2^{-n\gamma} \Big(\sum_l \sum_{k \in (I_l^n)^*} \Big\| \sum_{\kappa \in (I_l^n)^*} \sum_{r(w) \in I_l^{n,\kappa}} g_w^{n,\kappa} \Big\|_2^2 \Big)^{1/2} \Big] \\
&\lesssim \Big[\sum_n 2^{-n\gamma} n^2 \Big(\sum_l \sum_{\substack{w : r(w) \\ \in I_l^n}} \Big\| \sum_{\kappa \in (I_l^n)^*} g_w^{n,\kappa} \Big\|_2^2 \Big)^{1/2} \Big]
\end{aligned}$$

which by the argument above is dominated by a constant times $\alpha \|f\|_1$. Combining these estimates with Chebyshev's inequality we see that (6.4) holds.

We are left to prove

$$\operatorname{meas}\Big(\Big\{ x : \Big| \sum_{n,l} \sum_{\kappa \in (I_l^n)^*} \sum_{\substack{k \in (I_l^n)^* \\ k \geq \kappa + C_0 \log n}} \mu_k^n * \sum_{r(w) \in I_l^{n,\kappa}} b_w^{n,\kappa} \Big| > \frac{\alpha}{10} \Big\} \Big) \lesssim \frac{\|f\|_1}{\alpha}. \qquad (6.8)$$

We now let $\mu^{i,n,m} = \nu^{i,n} * \phi_m$ (which was previously considered only for the case $m = n$) and define the L^1 dilate $\mu_k^{i,n,m} = 2^{-k\tau}\mu^{i,n,m}(\delta_{-k}\cdot)$. Split (with $\mathcal{J}_1^n$ and $\mathcal{J}_2^n$ as in §5)

$$\begin{aligned}\mu_k^n &= \mu_k^0 + (\mu_k^n - \mu_k^0)\\ &= \mu_k^0 + \sum_{i\in\mathcal{J}_1^n}\sum_{m=1}^{n}(\mu_k^{i,n,m} - \mu_k^{i,n,m-1}) + \sum_{i\in\mathcal{J}_2^n}(\mu_k^{i,n,n} - \mu_k^{i,n,0}).\end{aligned} \tag{6.9}$$

Let $h_k^{i,n,m} = \mu_k^{i,n,m} - \mu_k^{i,n,m-1}$ and $h_k^{n,m} = \sum_{i\in\mathcal{J}_1^n} h_k^{i,n,m}$. Using (6.9) we split

$$\sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\geq\kappa+C_0\log n}}\mu_k^n * \sum_{r(w)\in I_l^{n,\kappa}} b_w^{n,\kappa} = I + II + \sum_{m=1}^{\infty} III_m$$

where

$$\begin{aligned}I &= \sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\geq\kappa+C_0\log n}}\mu_k^0 * \sum_{r(w)\in I_l^{n,\kappa}} b_w^{n,\kappa}\\ II &= \sum_{n,l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\geq\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_2^n}(\mu_k^{i,n,n} - \mu_k^{i,n,0}) * \sum_{r(w)\in I_l^{n,\kappa}} b_w^{n,\kappa}\\ III_m &= \sum_{n\geq m}\sum_{l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\geq\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}(\mu_k^{i,n,m} - \mu_k^{i,n,m-1}) * \sum_{r(w)\in I_l^{n,\kappa}} b_w^{n,\kappa}.\end{aligned}$$

We show that

$$\|I\|_{L^1(\mathbb{R}^d\setminus\Omega^*)} + \|II\|_1 \lesssim \|f\|_1 \tag{6.10}$$

and

$$\|III_m\|_2^2 \lesssim (1+m)^{-2}\alpha\|f\|_1. \tag{6.11}$$

(6.10/11) imply that the sets where $|I| > \alpha/10$, $|II| > \alpha/10$, and $\sum_{m=1}^{\infty}|III_m| > \alpha/10$ all have measure $\lesssim \alpha^{-1}\|f\|_1$. Combining this with the estimate (3.2) for the measure of Ω^* yields (6.8).

The inequality

$$\|I\|_{L^1(\mathbb{R}^d\setminus\Omega^*)} \lesssim \|f\|_1$$

follows from the standard estimates for singular integrals (in view of the regularity of $\mu * \phi_0$). The bound for $\|II\|_1$ is proved exactly as in estimate (5.9). Thus we are left to check (6.11).

Concerning the terms III_m we apply Cauchy-Schwarz' inequality and estimate

$$\|III_m\|_2^2 \lesssim \sum_{n\geq m} n^2 \Big\|\sum_{l}\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\geq\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n} h_k^{i,n,m} * \sum_{r(w)\in I_l^{n,\kappa}} b_w^{n,\kappa}\Big\| \lesssim C_2 IV_m + V_m$$

where $C_2 > 10/a$,

$$IV_m = \sum_{n\geq m} n^2 \sum_{l}\Big\|\sum_{\kappa\in(I_l^n)^*}\sum_{\substack{k\in(I_l^n)^*\\ k\geq\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n} h_k^{i,n,m} * \sum_{r(w)\in I_l^{n,\kappa}} b_w^{n,\kappa}\Big\|_2^2 \tag{6.12}$$

and

$$(6.13)\quad V_m=\sum_{n\ge m}n^2\sum_{\substack{l,l'\\|l-l'|\ge C_2}}\sum_{\kappa,\kappa'\in(I_l^n)^*}\sum_{\substack{k,k'\in(I_l^n)^*\\k\ge\kappa+C_0\log n\\k'\ge\kappa'+C_0\log n}}\sum_{i,i'\in\mathcal{J}_1^n}\big\langle\widetilde{h}_{k'}^{i',n,m'}*h_k^{i,n,m}*\sum_{r(w)\in I_l^{n,\kappa}}b_w^{n,\kappa},\sum_{r(w')\in I_{l'}^{n,\kappa'}}\overline{b_{w'}^{n,\kappa'}}\big\rangle.$$

The inner product in the second term is estimated by Plancherel's theorem. By van der Corput's Lemma and cancellation there is the Fourier transform estimate

$$\big|\mathcal{F}[h_0^{i,n,m}](\xi)\big|\lesssim\min\{|\xi|,|\xi|^{-\gamma}\}$$

and thus

$$\big|\mathcal{F}\big[\widetilde{h}_{k'}^{n,m,i'}*h_k^{n,m,i}\big]\big|\lesssim 2^{-|k-k'|a\gamma}$$

which is $O(2^{-n|l-l'|\gamma a/2})$ if $k\in(I_l^n)^*$, $k'\in(I_{l'}^n)^*$, $|l-l'|\ge C_2>10/a$. Set

$$\mathcal{E}_{l,n}(x)=\sum_{\kappa\in(I_l^n)^*}\sum_{r(w)\in I_l^{n,\kappa}}b_w^{n,\kappa}(x).$$

We may apply Cauchy-Schwarz and Parseval's theorem to bound

$$\begin{aligned}V_m&\lesssim\sum_{n\ge m}n^{4+6(d-1)/\eta}\sum_{\substack{l,l'\\|l-l'|\ge C_2}}2^{-n|l-l'|\gamma}\int|\widehat{\mathcal{E}_{l,n}}(\xi)||\widehat{\mathcal{E}_{l',n}}(\xi)|d\xi\\&\lesssim\sum_{n\ge m}n^{4+6(d-1)/\eta}\sum_{\substack{l,l'\\|l-l'|\ge C_2}}2^{-n|l-l'|\gamma}\|\mathcal{E}_{l,n}\|_2\|\mathcal{E}_{l',n}\|_2\\&\lesssim\sum_{n\ge m}n^{4+6(d-1)/\eta}2^{-C_2a\gamma n}\sum_l\|\mathcal{E}_{l,n}\|_2^2\end{aligned}$$

Now

$$\|\mathcal{E}_{l,n}\|_2^2\lesssim 2^{c(n+1)}\alpha\Big\|\sum_{r(w)\in I_l^{n,\kappa}}|b_w^{n,\kappa}|\Big\|_1$$

where c is as in (3.3) and hence we obtain

$$(6.14)\qquad V_m\lesssim\alpha\sum_{n\ge m}2^{-5a\gamma n}\Big\|\sum_l\sum_{\kappa\in(I_l^n)^*}\sum_{r(w)\in I_l^{n,\kappa}}|b_w^{n,\kappa}|\Big\|_1\lesssim 2^{-5a\gamma m}\alpha\|f\|_1.$$

For the term IV_m we have by Cauchy-Schwarz for the k summation and other applications of Cauchy-Schwarz leading to (5.20)

$$\begin{aligned}IV_m&\lesssim\sum_{n\ge m}n^3\sum_l\sum_{k\in(I_l^n)^*}\Big\|\sum_{\substack{\kappa\in(I_l^n)^*\\k\ge\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}h_k^{i,n,m}*\sum_{r(w)\in I_l^{n,\kappa}}b_w^{n,\kappa}\Big\|_2^2\\&\lesssim\sum_{n\ge m}n^{5+\frac{3(d-1)}{\eta}}\sum_l\sum_{k\in(I_l^n)^*}\sum_{\substack{\kappa\in(I_l^n)^*\\k\ge\kappa+C_0\log n}}\sum_{i\in\mathcal{J}_1^n}\sum_{r\in I_\kappa^{n,l}}\Big\|\mu_k^{i,n,m}*\sum_{r(w)=r}b_w^{n,\kappa}\Big\|_2^2.\end{aligned}\qquad(6.15)$$

Now $\mu_k^{i,n,m}$ satisfies similar quantitative properties as $\mu_k^{i,n,n}\equiv\mu_k^{i,n}$ considered in §5; in particular we have $|\partial^\alpha(\mu_0^{i,n,n}*\mu_0^{i,n,m})(x)|\lesssim n^{-3(d-3-2|\alpha|)/\eta}(2^{-m}+|x|)^{-1-|\alpha|}$. Thus the estimates for expression (5.20), are applicable and we obtain the bound

$$\begin{aligned}IV_m&\lesssim\alpha\sum_{n\ge m}n^{8+9(d+1)/\eta}\sum_\kappa\sum_{k\ge\kappa+C_0\log n}2^{-(k-\kappa)c_0}\sum_r\sum_{r(w)=r}\|b_w^{n,\kappa}\|_1\\&\lesssim\sum_{n\ge m}n^{8+9(d+1)/\eta-C_0c_0}\alpha\|f\|_1\lesssim(1+m)^{-2}\alpha\|f\|_1.\end{aligned}\qquad(6.16)$$

This shows (6.11) and thus (6.8) and the proof of Theorem 1.2. is complete. □

References

1. C. P. Calderón, *Lacunary spherical means*, Illinois J. Math. **23** (1979), 476–484.
2. M. Christ, *Weak type (1,1) bounds for rough operators*, Annals of Math. **128** (1988), 19–42.
3. ______, *Examples of singular maximal functions unbounded on L^p*, Publicacions Mathemàtiques **35** (1991), 269–279.
4. M. Christ and E. M. Stein, *A remark on singular Calderón-Zygmund theory*, Proc. Amer. Math. Soc. **99** (1987), 71–75.
5. R. R. Coifman and G. Weiss, *Book review*, Bull. Amer. Math. Soc. **84** (1978), 242–250.
6. J. Duoandikoetxea and J.-L. Rubio de Francia, *Maximal and singular integral operators via Fourier transform estimates*, Invent. Math. **84** (1986), 541–561.
7. C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36.
8. L. Grafakos, *Endpoint bounds for an analytic family of Hilbert transforms*, Duke Math. J. **62** (1991), 23–59.
9. A. Nagel, N. M. Rivière and S. Wainger, *On Hilbert transforms along curves, II*, Amer. J. Math. **98** (1976), 395–403.
10. N. M. Rivière, *Singular integrals and multiplier operators*, Ark. Mat. **9** (1971), 243–278.
11. A. Seeger and T. Tao, *Sharp Lorentz space estimates for rough operators*, Math. Annalen **320** (2001), 381–415.
12. A. Seeger, T. Tao and J. Wright, *Endpoint mapping properties of spherical maximal operators*, preprint, 2001.
13. ______, *Pointwise convergence of lacunary spherical means*, expository note (to appear).
14. E. M. Stein, *On limits of sequences of maximal operators*, Annals of Math. **74** (1961), 140–170.
15. ______, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, N.J., 1971.
16. ______, *Harmonic analysis: Real variable methods, orthogonality and oscillatory integrals*, Princeton Univ. Press, 1993.
17. E. M. Stein and S. Wainger, *Problems in harmonic analysis related to curvature*, Bull. Amer. Math. Soc. **84** (1978), 1239–1295.

Department of Mathematics, University of Wisconsin, Madison, WI 53706-1388, USA

E-mail address: seeger@math.wisc.edu

Department of Mathematics, University of California, Los Angeles, CA 90095-1555, USA

E-mail address: tao@math.ucla.edu

Department of Mathematics and Statistics, University of Edinburgh, King's Buildings, Mayfield Rd., Edinburgh EH3 9JZ, U.K.

E-mail address: wright@maths.ed.ac.uk