# VARIATION AND OSCILLATION FOR SINGULAR INTEGRALS WITH ODD KERNEL ON LIPSCHITZ GRAPHS 

ALBERT MAS AND XAVIER TOLSA


#### Abstract

We prove that, for $\rho>2$, the $\rho$-variation and oscillation for the smooth truncations of the Cauchy transform on Lipschitz graphs are bounded in $L^{p}$ for $1<p<\infty$. The analogous result holds for the $n$-dimensional Riesz transform on $n$-dimensional Lipschitz graphs, as well as for other singular integral operators with odd kernel. In particular, our results strengthen the classical theorem on the $L^{2}$ boundedness of the Cauchy transform on Lipschitz graphs by Coifman, McIntosh, and Meyer.


## 1. Introduction

The $\rho$-variation and oscillation for martingales and some families of operators have been studied in many recent papers on probability, ergodic theory, and harmonic analysis (see [Lé], [Bo], [JKRW], [CJRW1], and [JSW], for example). The purpose of this paper is to establish some new results concerning the $\rho$-variation and oscillation for families of singular integral operators defined on Lipschitz graphs. In particular, our results include the $L^{p}$ boundedness of the $\rho$-variation and the oscillation for the smooth truncations of the Cauchy transform and the $n$-dimensional Riesz transform on Lipschitz graphs, for $1<p<\infty$ and $\rho>2$.

Given a Borel measure $\mu$ in $\mathbb{R}^{d}$, one defines the $n$-dimensional Riesz transform of a function $f \in L^{1}(\mu)$ by $R^{\mu} f(x)=\lim _{\epsilon \searrow 0} R_{\epsilon}^{\mu} f(x)$ (whenever the limit exists), where

$$
R_{\epsilon}^{\mu} f(x)=\int_{|x-y|>\epsilon} \frac{x-y}{|x-y|^{n+1}} f(y) d \mu(y), \quad x \in \mathbb{R}^{d}
$$

When $d=2$ (i.e., $\mu$ is a Borel measure in $\mathbb{C}$ ), one defines the Cauchy transform of $f \in L^{1}(\mu)$ by $C^{\mu} f(x)=\lim _{\epsilon} \searrow_{0} C_{\epsilon}^{\mu} f(x)$ (whenever the limit exists), where

$$
C_{\epsilon}^{\mu} f(x)=\int_{|x-y|>\epsilon} \frac{f(y)}{x-y} d \mu(y), \quad x \in \mathbb{C}
$$

To avoid the problem of existence of the preceding limits, it is useful to consider the maximal operators $R_{*}^{\mu} f(x)=\sup _{\epsilon>0}\left|R_{\epsilon}^{\mu} f(x)\right|$ and $C_{*}^{\mu} f(x)=\sup _{\epsilon>0}\left|C_{\epsilon}^{\mu} f(x)\right|$.

The Cauchy and Riesz transforms are two very important examples of singular integral operators with a Calderón-Zygmund kernel. The kernels $K: \mathbb{R}^{d} \backslash\{0\} \rightarrow \mathbb{R}$ that we consider in this paper satisfy

$$
\begin{equation*}
|K(x)| \leq \frac{C}{|x|^{n}}, \quad\left|\partial_{x^{i}} K(x)\right| \leq \frac{C}{|x|^{n+1}} \quad \text { and } \quad\left|\partial_{x^{i}} \partial_{x^{j}} K(x)\right| \leq \frac{C}{|x|^{n+2}}, \tag{1.1}
\end{equation*}
$$

[^0]for all $1 \leq i, j \leq d$ and $x=\left(x^{1}, \ldots, x^{d}\right) \in \mathbb{R}^{d} \backslash\{0\}$, where $0<n<d$ is some integer and $C>0$ is some constant; and moreover $K(-x)=-K(x)$ for all $x \neq 0$ (i.e. $K$ is odd). Notice that the $n$-dimensional Riesz transform corresponds to the vector kernel $\left(x^{1}, \ldots, x^{d}\right) /|x|^{n+1}$, and the Cauchy transform to $\left(x^{1},-x^{2}\right) /|x|^{2}$ (so, we may consider $K$ to be any scalar component of these vector kernels).

Given an odd kernel $K$ satisfying (1.1) and a finite Borel measure $\mu$ in $\mathbb{R}^{d}$, for each $\epsilon>0$, we consider the $\epsilon$-truncated operator

$$
T_{\epsilon} \mu(x)=\int_{|x-y|>\epsilon} K(x-y) d \mu(y), \quad x \in \mathbb{R}^{d},
$$

and then we set $T \mu(x)=\lim _{\epsilon} \searrow_{0} T_{\epsilon} \mu(x)$ whenever the limit makes sense, and $T_{*} \mu(x)=$ $\sup _{\epsilon>0}\left|T_{\epsilon} \mu(x)\right|$. Finally, given $f \in L^{1}(\mu)$, we define $T_{\epsilon}^{\mu} f(x):=T_{\epsilon}(f \mu)(x), T^{\mu} f(x):=$ $T(f \mu)(x)$ and $T_{*}^{\mu} f(x):=T_{*}(f \mu)(x)$. Thus, for a suitable choice of $K$, the operator $T^{\mu}$ coincides with the Cauchy or Riesz transforms.

Besides the operator $T_{\epsilon}$ defined above, one can consider another $\epsilon$-truncated variant that we proceed to define. First we need some additional notation. Given $x=\left(x^{1}, \ldots, x^{d}\right) \in \mathbb{R}^{d}$, we use the notation $\widetilde{x}:=\left(x^{1}, \ldots, x^{n}\right) \in \mathbb{R}^{n}$. Let $\varphi_{\mathbb{R}}:[0, \infty) \rightarrow[0, \infty)$ be a non decreasing $\mathcal{C}^{2}$ function such that $\chi_{[3 \sqrt{n}, \infty)} \leq \varphi_{\mathbb{R}} \leq \chi_{[2.1 \sqrt{n}, \infty)}$ (the numbers $3 \sqrt{n}$ and $2.1 \sqrt{n}$ are chosen just for definiteness and they are not important). Given $\epsilon>0$ and $x \in \mathbb{R}^{d}$, we denote

$$
\varphi_{\epsilon}(x):=\varphi_{\mathbb{R}}(|\widetilde{x}| / \epsilon) \quad \text { and } \quad \varphi:=\left\{\varphi_{\epsilon}\right\}_{\epsilon>0}
$$

Given $K$ as above, $x \in \mathbb{R}^{d}, 0<\epsilon$, and a finite Borel measure $\mu$, we set

$$
\left(K \varphi_{\epsilon} * \mu\right)(x):=\int \varphi_{\epsilon}(x-y) K(x-y) d \mu(y)
$$

We also denote $(K \varphi * \mu)(x):=\left\{\left(K \varphi_{\epsilon} * \mu\right)(x)\right\}_{\epsilon>0}$. Finally, given $f \in L^{1}(\mu)$, we define $T_{\varphi_{\epsilon}}^{\mu} f(x):=\left(K \varphi_{\epsilon} *(f \mu)\right)(x), T_{\varphi}^{\mu} f(x):=\lim _{\epsilon \rightarrow 0} T_{\varphi_{\epsilon}}^{\mu} f(x)$ (whenever the limit makes sense), $T_{\varphi_{*}}^{\mu} f(x):=\sup _{\epsilon>0}\left|T_{\varphi_{\epsilon}}^{\mu} f(x)\right|$, and $\mathcal{T}_{\varphi}^{\mu} f(x):=\left\{T_{\varphi_{\epsilon}}^{\mu} f(x)\right\}_{\epsilon>0}$.

Let $\mathcal{I}$ be a subset of $\mathbb{R}$ (in this paper, we will always have $\mathcal{I}=(0, \infty)$ or $\mathcal{I}=\mathbb{Z}$ ), and let $\mathcal{F}:=\left\{F_{\epsilon}\right\}_{\epsilon \in \mathcal{I}}$ be a family of functions defined on $\mathbb{R}^{d}$. Given $\rho>0$, the $\rho$-variation of $\mathcal{F}$ at $x \in \mathbb{R}^{d}$ is defined by

$$
\mathcal{V}_{\rho}(\mathcal{F})(x):=\sup _{\left\{\epsilon_{m}\right\}}\left(\sum_{m \in \mathbb{Z}}\left|F_{\epsilon_{m+1}}(x)-F_{\epsilon_{m}}(x)\right|^{\rho}\right)^{1 / \rho}
$$

where the pointwise supremum is taken over all decreasing sequences $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}} \subset \mathcal{I}$. Fix a decreasing sequence $\left\{r_{m}\right\}_{m \in \mathbb{Z}} \subset \mathcal{I}$. The oscillation of $\mathcal{F}$ at $x \in \mathbb{R}^{d}$ is defined by

$$
\mathcal{O}(\mathcal{F})(x):=\sup _{\left\{\epsilon_{m}\right\},\left\{\delta_{m}\right\}}\left(\sum_{m \in \mathbb{Z}}\left|F_{\epsilon_{m}}(x)-F_{\delta_{m}}(x)\right|^{2}\right)^{1 / 2}
$$

where the pointwise supremum is taken over all sequences $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}} \subset \mathcal{I}$ and $\left\{\delta_{m}\right\}_{m \in \mathbb{Z}} \subset \mathcal{I}$ such that $r_{m+1} \leq \epsilon_{m} \leq \delta_{m} \leq r_{m}$ for all $m \in \mathbb{Z}$.

In this paper we are interested in studying the $\rho$-variation and oscillation for the family $\mathcal{T}_{\varphi}^{\mu} f$. That is, we will deal with

$$
\begin{aligned}
& \left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mu}\right) f(x):=\mathcal{V}_{\rho}\left(\mathcal{T}_{\varphi}^{\mu} f\right)(x)=\mathcal{V}_{\rho}(K \varphi *(f \mu))(x) \quad \text { and } \\
& \left(\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mu}\right) f(x):=\mathcal{O}\left(\mathcal{T}_{\varphi}^{\mu} f\right)(x)=\mathcal{O}(K \varphi *(f \mu))(x),
\end{aligned}
$$

for a Borel measure $\mu$ and $f \in L^{1}(\mu)$. Although it is not clear from the definitions, these operators are $\mu$-measurable (see [CJRW1], [JSW]).

Given $E \subset \mathbb{R}^{d}$, we denote by $\mathcal{H}_{E}^{n}$ the $n$-dimensional Hausdorff measure restricted to $E$. Let $\Gamma:=\left\{x \in \mathbb{R}^{d}: x=(\widetilde{x}, A(\widetilde{x}))\right\}$ be the graph of a Lipschitz function $A: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ with Lipschitz constant $\operatorname{Lip}(A)$. Let $H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ and $B M O\left(\mathcal{H}_{\Gamma}^{n}\right)$ be the (atomic) Hardy space and the space of functions with bounded mean oscillation, respectively, with respect to the measure $\mathcal{H}_{\Gamma}^{n}$. The following is our main result.

Theorem 1.1. Let $\rho>2$, let $K$ be a kernel satisfying (1.1), and set $\mu:=\mathcal{H}_{\Gamma}^{n}$. The operators $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mu}$ and $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mu}$ are bounded

- in $L^{p}(\mu)$ for $1<p<\infty$,
- from $H^{1}(\mu)$ to $L^{1}(\mu)$,
- from $L^{1}(\mu)$ to $L^{1, \infty}(\mu)$, and
- from $L^{\infty}(\mu)$ to $B M O(\mu)$.

In all the cases above, the norm of $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mu}$ is bounded independently of the sequence that defines $\mathcal{O}$.

Let us recall that the $L^{2}\left(\mathcal{H}_{\Gamma}^{1}\right)$ boundedness of the Cauchy transform on Lipschitz graphs $\Gamma \subset \mathbb{C}$ with slope small enough was proved by A. P. Calderón in his celebrated paper [Ca]. The $L^{2}$ boundedness on Lipschitz graphs in full generality was proved later on by R. Coifman, A. McIntosh, and Y. Meyer [CMM].

Consider the Cauchy kernel $K(z)=1 / z(z \in \mathbb{C})$, and set $\mu:=\mathcal{H}_{\Gamma}^{1}$, so $C_{\epsilon}^{\mu}=T_{\epsilon}^{\mu}$. By standard Calderón-Zygmund theory (namely, Cotlar's inequality), the $L^{2}(\mu)$ boundedness of the Cauchy transform $C^{\mu}$ is equivalent to the $L^{2}(\mu)$ boundedness of the maximal operator $C_{*}^{\mu}$. Let $M^{\mu}$ denote the Hardy-Littlewood maximal operator with respect to the measure $\mu$. It is easy to check that, for $f \in L^{1}(\mu)$ with compact support, there exists some constant $C_{0}>0$ such that

$$
C_{\epsilon}^{\mu} f(x) \leq T_{\varphi_{\epsilon}}^{\mu} f(x)+C_{0} M^{\mu} f(x) \leq\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mu}\right) f(x)+C_{0} M^{\mu} f(x)
$$

for all $\epsilon>0$, thus $\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mu}\right)+C_{0} M^{\mu}$ controls the maximal operator $C_{*}^{\mu}$ and, in this sense, Theorem 1.1 (together with the known $L^{p}(\mu)$ boundedness of $\left.M^{\mu}\right)$ strengthens the results of [Ca] and [CMM]. Analogous conclusions hold for the $n$-dimensional Riesz transform and the maximal operator $R_{*}^{\mu}$.

The operator $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mu}$ is also related to an important open problem posed by G. David and S. Semmes which actually is our main motivation to prove Theorem 1.1. We need some definitions to state it.

Recall that a measure $\mu$ is said to be $n$-dimensional Ahlfors-David regular, or simply AD regular, if there exists some constant $C$ such that $C^{-1} r^{n} \leq \mu(B(x, r)) \leq C r^{n}$ for all $x \in \operatorname{supp} \mu$ and $0<r \leq \operatorname{diam}(\operatorname{supp} \mu)$. It is not difficult to see that such a measure $\mu$ must be of the form $\mu=h \mathcal{H}_{\operatorname{supp} \mu}^{n}$, where $h$ is some positive function bounded above and away from zero. A Borel set $E \subset \mathbb{R}^{d}$ is called AD regular if the measure $\mathcal{H}_{E}^{n}$ is AD regular. One says that $\mu$ is uniformly $n$-rectifiable, or simply uniformly rectifiable, if there exist $\theta, M>0$ so that, for each $x \in \operatorname{supp} \mu$ and $R>0$, there is a Lipschitz mapping $g$ from the $n$-dimensional ball $B^{n}(0, R) \subset \mathbb{R}^{n}$ into $\mathbb{R}^{d}$ such that $\operatorname{Lip}(g) \leq M$ and $\mu\left(B(x, R) \cap g\left(B^{n}(0, R)\right)\right) \geq \theta R^{n}$, where $\operatorname{Lip}(g)$ stands for the Lipschitz constant of $g$. In the language of [DS2], this means that $\operatorname{supp} \mu$ has big pieces of Lipschitz images of $\mathbb{R}^{n}$. A Borel set $E \subset \mathbb{R}^{d}$ is called uniformly $n$ rectifiable if $\mathcal{H}_{E}^{n}$ is $n$-uniformly rectifiable. Of course, the $n$-dimensional graph of a Lipschitz function is uniformly $n$-rectifiable.

David and Semmes asked the following question, which is still open (see [Pa, Chapter 7]):
Problem 1.2. Is it true that an n-dimensional $A D$ regular measure $\mu$ is $n$-uniformly rectifiable if and only if $R_{*}^{\mu}$ is bounded in $L^{2}(\mu)$ ?

It is proved in [DS1] that if $\mu$ is uniformly rectifiable, then $R_{*}^{\mu}$ is bounded in $L^{2}(\mu)$. However, the converse implication has been proved only in the case $n=1$ and $d=2$, by P. Mattila, M. Melnikov and J. Verdera [MMV], using the notion of curvature of measures (which seems to be useful only in this case).

Set $\mathcal{R}^{\mu}:=\left\{R_{\epsilon}^{\mu}\right\}_{\epsilon>0}$. By combining some techniques from [DS2] and [To], in our forthcoming paper [MT] we show that the $L^{2}(\mu)$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{R}^{\mu}$ implies that $\mu$ is uniformly $n$-rectifiable. Moreover, we also prove that $\mathcal{V}_{\rho} \circ \mathcal{R}^{\mu}$ is bounded in $L^{2}(\mu)$ for all AD regular uniformly $n$-rectifiable measures $\mu$. So we obtain the following theorem, which might be considered as a first approach to a possible solution of Problem 1.2:

Theorem 1.3. Let $\rho>2$. An n-dimensional $A D$ regular measure $\mu$ is uniformly $n$-rectifiable if and only if $\mathcal{V}_{\rho} \circ \mathcal{R}^{\mu}$ is a bounded operator in $L^{2}(\mu)$.

An essential ingredient for the proof of this result is Theorem 1.1 above. The arguments and techniques used to derive the $L^{2}$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{R}^{\mu}$ on uniformly rectifiable measures from the $L^{2}$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{R}_{\varphi}^{\mu}$ on Lipschitz graphs are quite delicate ( $\mathcal{R}_{\varphi}^{\mu}$ is defined as $\mathcal{R}^{\mu}$ but using the family $\varphi$ for the truncations). In particular, they involve the corona type decomposition introduced in [DS1]. For this reason, the proof of the preceding theorem is out of the scope of this paper and will appear in [MT].

Concerning the background on the $\rho$-variation and oscillation, a fundamental result is Lépingle's inequality [Lé], from which the $L^{p}$ boundedness of the $\rho$-variation and oscillation for martingales follows, for $\rho>2$ and $1<p<\infty$ (see Theorem 2.4 below for more details). From this result on martingales, one deduces that the $\rho$-variation and oscillation are also bounded in $L^{p}$ for the averaging operators (also called differentiation operators, see [JKRW]):

$$
\begin{equation*}
D_{\epsilon} f(x)=\frac{1}{|B(x, \epsilon)|} \int_{B(x, \epsilon)} f(y) d y, \quad x \in \mathbb{R} . \tag{1.2}
\end{equation*}
$$

As far as we know, the first work dealing with the $\rho$-variation and oscillation for singular integral operators is the one of J. Campbell, R. L. Jones, K. Reinhold and M. Wierdl [CJRW1], where the $L^{p}$ and weak $L^{1}$ boundedness of the $\rho$-variation (for $\rho>2$ ) and oscillation for the Hilbert transform was proved. Recall that, for $f \in L^{p}(\mathbb{R})$ and $x \in \mathbb{R}$,

$$
H_{\epsilon} f(x)=\frac{1}{\pi} \int_{|x-y|>\epsilon} \frac{1}{x-y} f(y) d y
$$

and then the Hilbert transform of $f$ is defined by $H f(x)=\lim _{\epsilon \rightarrow 0} H_{\epsilon} f(x)$, whenever the limit exists. Later on, there appeared other papers showing the $L^{p}$ boundedness of the $\rho$ variation and oscillation for singular integrals in $\mathbb{R}^{d}$ ([CJRW2]), with weights ([GT]), or for other operators such as the spherical averaging operator or singular integral operators on parabolas ([JSW]). Finally, we remark that, very recently, the case of the Carleson operator has been considered too ([LT], [OSTTW]).

Notice that the Hilbert transform is one of the simplest examples where Theorem 1.1 applies (one sets $\Gamma=\mathbb{R}$, i.e. $A \equiv 0$ ), and so one obtains a new proof of the $L^{p}$ boundedness of the $\rho$-variation and oscillation for the Hilbert transform. In the original proof in [CJRW1], a key ingredient was the following classical identity, which follows via the Fourier transform:

$$
\begin{equation*}
Q_{\epsilon}=P_{\epsilon} * H \tag{1.3}
\end{equation*}
$$

where $P_{\epsilon}$ is the Poisson kernel and $Q_{\epsilon}$ is the conjugated Poisson kernel. Using this identity and the close relationship between the operators $Q_{\epsilon}$ and $H_{\epsilon}$, Campbell et al. derived the $L^{p}$ boundedness of the $\rho$-variation and oscillation for the Hilbert transform from the one of
the family $\left\{D_{\epsilon}(H f)\right\}_{\epsilon>0}$, where $D_{\epsilon}$ is the averaging operator in (1.2) (notice that $P_{\epsilon}$ can be written as a convex combination of operators $D_{\delta}, \delta>0$ ).

In most of the previous results concerning $\rho$-variation and oscillation of families of operators from harmonic analysis, the Fourier transform is a fundamental tool. However, this is not useful in order to prove Theorem 1.1, since the graph $\Gamma$ is not invariant under translations in general. Moreover, even for the Cauchy transform, there is no formula like (1.3), which relates the truncations of a singular integral operator with an averaging operator applied to a singular integral operator, when $\Gamma$ is a general Lipschitz graph.

The main ingredients of our proof of Theorem 1.1 are the known results on the $\rho$-variation and oscillation for martingales (Lépingle's inequality [Lé]) and a multiscale analysis which stems from the geometric proof of the $L^{2}$ boundedness of the Cauchy transform on Lipschitz graphs by P. W. Jones [Jn1] and his celebrated work [Jn2] on quantitative rectifiability in the plane, using the so called $\beta$ coefficients. Some of the techniques in these papers were further developed in higher dimensions by David and Semmes [DS1] for Ahlfors-David regular sets. More recently, in $[\mathrm{To}]$ some coefficients denoted by $\alpha$, in the spirit of the Jones' $\beta$ 's, were introduced, and they were shown to be useful for the study of the $L^{p}$-boundedness of Calderón-Zygmund operators on Lipschitz graphs and on uniformly rectifiable sets (see the definition below Theorem 1.4). In our paper, the $\alpha$ and $\beta$ coefficients play a fundamental role.

Let us remark that Lépingle's inequality, which asserts the $L^{p}$ boundedness of the $\rho$ variation of martingales, fails if one assumes $\rho \leq 2$ (see [Qi] and [JW], for example). Moreover, this fact can be brought to the $\rho$-variation of averaging operators and singular integral operators, thus it is essential to assume $\rho>2$ in Theorem 1.1. Analogous conclusions hold if one replaces the $\ell^{2}$-norm by and $\ell^{\rho}$-norm with $\rho<2$ in the definition of $\mathcal{O}$. See [CJRW1], or [AJS] for the case of martingales.

Concerning the direct applications of Theorem 1.1, it is easily seen that the $L^{p}$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mu}$ yields a new proof of the existence of the principal values $T_{\varphi}^{\mu} f(x):=$ $\lim _{\epsilon \rightarrow 0} T_{\varphi_{\epsilon}}^{\mu} f(x)$ for all $f \in L^{p}(\mu)$ and almost all $x \in \Gamma$, without using a dense class of functions in $L^{p}(\mu)$ as in the classical proof. Moreover, from Theorem 1.1 one also gets some information on the speed of convergence. In fact, a classical result derived from variational inequalities is the boundedness of the $\lambda$-jump operator $N_{\lambda} \circ \mathcal{T}_{\varphi}^{\mu}$ and the ( $a, b$ )-upcrossings operator $N_{a}^{b} \circ \mathcal{T}_{\varphi}^{\mu}$. Given $\lambda>0, f \in L_{\text {loc }}^{1}(\mu)$ and $x \in \mathbb{R}^{d}$, one defines $\left(N_{\lambda} \circ \mathcal{T}_{\varphi}^{\mu}\right) f(x)$ as the supremmum of all integers $N$ for which there exists $0<\epsilon_{1}<\delta_{1} \leq \epsilon_{2}<\delta_{2} \leq \cdots \leq \epsilon_{N}<\delta_{N}$ so that

$$
\left|T_{\varphi_{\epsilon_{i}}}^{\mu} f(x)-T_{\varphi \delta_{i}}^{\mu} f(x)\right|>\lambda
$$

for each $i=1, \ldots, N$. Similarly, given $a<b$, one defines $\left(N_{a}^{b} \circ \mathcal{T}_{\varphi}^{\mu}\right) f(x)$ to be the supremmum of all integers $N$ for which there exists $0<\epsilon_{1}<\delta_{1} \leq \epsilon_{2}<\delta_{2} \leq \cdots \leq \epsilon_{N}<\delta_{N}$ so that $T_{\varphi_{i}}^{\mu} f(x)<a$ and $T_{\varphi_{\delta_{i}}}^{\mu} f(x)>b$ for each $i=1, \ldots, N$. Using Theorem 1.1 one obtains (by the same arguments as in [CJRW1, Theorem 1.3 and Corollary 7.1]) the following:

Theorem 1.4. Let $\rho>2, \lambda>0$, and let $K$, and $\mu$ be as in Theorem 1.1. For $1<p<\infty$, there exist constants $C_{1}$ and $C_{2}$ depending on $\rho, n, d, K$, and $\operatorname{Lip}(A)$ (and on $p$ for the case of $C_{1}$ ) such that

$$
\begin{gathered}
\left\|\left(\left(N_{\lambda} \circ \mathcal{T}_{\varphi}^{\mu}\right) f\right)^{1 / \rho}\right\|_{L^{p}(\mu)} \leq \frac{C_{1}}{\lambda}\|f\|_{L^{p}(\mu)} \text { and } \\
\mu\left(\left\{x \in \Gamma:\left(N_{\lambda} \circ \mathcal{T}_{\varphi}^{\mu}\right) f(x)>m\right\}\right) \leq \frac{C_{2}}{\lambda m^{1 / \rho}}\|f\|_{L^{1}(\mu)} .
\end{gathered}
$$

Trivially, $\left(N_{a}^{b} \circ \mathcal{T}_{\varphi}^{\mu}\right) f \leq\left(N_{b-a} \circ \mathcal{T}_{\varphi}^{\mu}\right) f$, thus Theorem 1.4 also holds replacing $\lambda$ by $b-a$ and $N_{\lambda}$ by $N_{a}^{b}$. In [JSW] it is shown that the results of Theorem 1.4 still hold when $\rho=2$ for the particular case of the Hilbert transform. In our paper we do not pursue this endpoint result.

## 2. Preliminaries

Throughout all the paper, $n$ and $d$ are two fixed integers such that $0<n<d$. Given a point $x=\left(x^{1}, \ldots, x^{d}\right) \in \mathbb{R}^{d}$, we use the notation $\widetilde{x}:=\left(x^{1}, \ldots, x^{n}\right) \in \mathbb{R}^{n}$. Given a function $f: \mathbb{R}^{m} \rightarrow \mathbb{R}$, we denote by $\nabla f$ its gradient (when it makes sense), and by $\nabla^{2} f$ the matrix of second derivatives of $f$. If $f$ depends on different points $x_{1}, x_{2}, \ldots \in \mathbb{R}^{m}$, then $\nabla_{x_{i}} f$ denotes the gradient of $f$ with respect to the $x_{i}$ variable, and analogously for $\nabla_{x_{i}}^{2} f$.

For two sets $F_{1}, F_{2} \subset \mathbb{R}^{d}$, we denote by $\operatorname{dist}_{\mathcal{H}}\left(F_{1}, F_{2}\right)$ the Hausdorff distance between $F_{1}$ and $F_{2}$. We denote by $\mathcal{L}^{n}$ the Lebesgue measure on $\mathbb{R}^{n}$, and for the sake of simplicity, we set $\|\cdot\|_{p}:=\|\cdot\|_{L^{p}\left(\mathcal{L}^{n}\right)}$ for $1 \leq p \leq \infty$, and $d y:=d \mathcal{L}^{n}(y)$ for $y \in \mathbb{R}^{n}$.

In the paper, when we refer to the angle between two affine $n$-planes in $\mathbb{R}^{d}$, we mean the angle between the $n$-dimensional subspaces associated to the $n$-planes. As usual, the letter ' $C$ ' stands for some constant which may change its value at different occurrences, and which quite often only depends on $n$ and $d$. The notation $A \lesssim B(A \gtrsim B)$ means that there is some fixed constant $C$ such that $A \leq C B(A \geq C B)$, with $C$ as above. Also, $A \approx B$ is equivalent to $A \lesssim B \lesssim A$.
2.1. More about the family $\varphi$. Given $x \in \mathbb{R}^{d}, 0<\epsilon \leq \delta$, and a finite Borel measure $\mu$, we set $\varphi_{\epsilon}^{\delta}(x):=\varphi_{\epsilon}(x)-\varphi_{\delta}(x)$ and we define

$$
\left(K \varphi_{\epsilon}^{\delta} * \mu\right)(x):=\int \varphi_{\epsilon}^{\delta}(x-y) K(x-y) d \mu(y)
$$

thus $\left(K \varphi_{\epsilon}^{\delta} * \mu\right)(x)=\left(K \varphi_{\epsilon} * \mu\right)(x)-\left(K \varphi_{\delta} * \mu\right)(x)$.
For $m \in \mathbb{N}, x \in \mathbb{R}^{m}$, and $R \geq r>0$, we denote by $B^{m}(x, r)$ the closed ball of $\mathbb{R}^{m}$ with center $x$ and radius $r$, and by $A^{m}(x, r, R)$ the closed annulus of $\mathbb{R}^{m}$ centered at $x$ with inner radius $r$ and outer radius $R$. We also use the notation $B(x, r)$ and $A(x, r, R)$ when there is no possible confusion about $m$.

Each function $\varphi_{\epsilon}^{\delta}$ is non negative, and $\operatorname{supp} \varphi_{\epsilon}^{\delta} \subset A^{n}(0,2.1 \epsilon \sqrt{n}, 3 \delta \sqrt{n}) \times \mathbb{R}^{d-n} \subset \mathbb{R}^{d}$. Moreover, $\sum_{j \in \mathbb{Z}} \varphi_{2^{-j-1}}^{2-j}(x)=1$ for $\widetilde{x} \neq 0$, and there are at most two terms that do not vanish in the previous sum for a given $x \in \mathbb{R}^{d}$.
2.2. The $\alpha$ and $\beta$ coefficients. Special dyadic lattice. Given $m \in \mathbb{N}, \lambda>0$, and a cube $Q \subset \mathbb{R}^{m}$ (i.e. $Q:=[0, b)^{m}+a$ with $a \in \mathbb{R}^{m}$ and $b>0$ ), $\ell(Q)$ denotes the side length of $Q$, $z_{Q}$ denotes the center of $Q$ and $\lambda Q$ denotes the cube with center $z_{Q}$ and side length $\lambda \ell(Q)$. Throughout the paper, we will only use cubes with sides parallel to the axes.

Let $\mu$ be a locally finite Borel measure on $\mathbb{R}^{d}$. Given $1 \leq p<\infty$ and a cube $Q \subset \mathbb{R}^{d}$, one sets (see [DS2])

$$
\begin{equation*}
\beta_{p, \mu}(Q)=\inf _{L}\left\{\frac{1}{\ell(Q)^{n}} \int_{2 Q}\left(\frac{\operatorname{dist}(y, L)}{\ell(Q)}\right)^{p} d \mu(y)\right\}^{1 / p}, \tag{2.1}
\end{equation*}
$$

where the infimum is taken over all $n$-planes $L$ in $\mathbb{R}^{d}$. For $p=\infty$ one replaces the $L^{p}$ norm by the supremum norm:

$$
\begin{equation*}
\beta_{\infty, \mu}(Q)=\inf _{L}\left\{\sup _{y \in \operatorname{supp} \mu \cap 2 Q} \frac{\operatorname{dist}(y, L)}{\ell(Q)}\right\}, \tag{2.2}
\end{equation*}
$$

where the infimum is taken over all $n$-planes $L$ in $\mathbb{R}^{d}$ again. These coefficients were introduced by P. W. Jones in [Jn1] for $p=\infty$ and by G. David and S. Semmes in [DS1] for $1 \leq p<\infty$.

Let $F \subset \mathbb{R}^{d}$ be the closure of an open set. Given two finite Borel measures $\sigma, \nu$ on $\mathbb{R}^{d}$, one sets

$$
\begin{equation*}
\operatorname{dist}_{F}(\sigma, \nu):=\sup \left\{\left|\int f d \sigma-\int f d \nu\right|: \operatorname{Lip}(f) \leq 1, \operatorname{supp} f \subset F\right\} . \tag{2.3}
\end{equation*}
$$

It is easy to check that this is a distance in the space of finite Borel measures $\sigma$ such that $\operatorname{supp} \sigma \subset F$ and $\sigma(\partial F)=0$. Moreover, it turns out that this distance is a variant of the well known Wasserstein distance $W_{1}$ from optimal transportation (see [Vi, Chapter 1]). See [Ma, Chapter 14] for other properties of $\operatorname{dist}_{F}$.

Given a cube $Q$ which intersects supp $\mu$, consider the closed ball $B_{Q}:=B\left(z_{Q}, 6 \ell(Q)\right)$. Then one defines (see [To])

$$
\begin{equation*}
\alpha_{\mu}^{n}(Q):=\frac{1}{\ell(Q)^{n+1}} \inf _{c \geq 0, L} \operatorname{dist}_{B_{Q}}\left(\mu, c \mathcal{H}_{L}^{n}\right), \tag{2.4}
\end{equation*}
$$

where the infimum is taken over all constants $c \geq 0$ and all $n$-planes $L$ in $\mathbb{R}^{d}$. For convenience, if $Q$ does not intersect $\operatorname{supp} \mu$, we set $\alpha_{\mu}^{n}(Q)=0$. To simplify notation, sometimes we will write $\alpha_{\mu}(Q)$ or $\alpha(Q)$ instead of $\alpha_{\mu}^{n}(Q)$ (and analogously for the $\beta$ 's).

The following result characterizes uniform rectifiability in terms of the $\alpha$ and $\beta$ coefficients.
Theorem 2.1. Let $\mu$ be an $n$-dimensional $A D$ regular measure on $\mathbb{R}^{d}$, and consider any $p \in[1,2]$. Then, the following are equivalent:
(a) $\mu$ is uniformly $n$-rectifiable.
(b) For any cube $R \subset \mathbb{R}^{d}$,

$$
\begin{equation*}
\sum_{Q \in \mathcal{D}_{\mathbb{R}^{d}}(R)} \beta_{p, \mu}(Q)^{2} \ell(Q)^{n} \leq C \ell(R)^{n} \tag{2.5}
\end{equation*}
$$

with $C$ independent of $R$, where $\mathcal{D}_{\mathbb{R}^{d}}(R)$ stands for the collection of cubes of $\mathbb{R}^{d}$ contained in $R$ which are obtained by splitting $R$ dyadically.
(c) There exists $C>0$ such that, for any cube $R \subset \mathbb{R}^{d}$,

$$
\begin{equation*}
\sum_{Q \in \mathcal{D}_{\mathbb{R}^{d}}(R)} \alpha_{\mu}(Q)^{2} \ell(Q)^{n} \leq C \ell(R)^{n} \tag{2.6}
\end{equation*}
$$

The equivalence $(\mathrm{a}) \Longleftrightarrow(\mathrm{b})$ in Theorem 2.1 was proved by G. David and S. Semmes in [DS1], and the equivalence $(\mathrm{a}) \Longleftrightarrow$ (c) was proved by X. Tolsa in [To].
In this paper we will use a slightly different definition of the $\alpha$ and $\beta$ coefficients adapted to the $n$-uniformly rectifiable measure $\mu=f \mathcal{H}_{\Gamma}^{n}$, where $\Gamma:=\left\{x \in \mathbb{R}^{d}: x=(\widetilde{x}, A(\widetilde{x}))\right\}$ is the $n$-dimensional graph of a given Lipschitz function $A: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ and $f \in L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$ satisfies $f(x) \approx 1$ for almost all $x \in \Gamma$. To this end, we need to introduce a special dyadic lattice of sets related to $\Gamma$. Given a cube $\widetilde{Q} \subset \mathbb{R}^{n}$ (i.e. $\widetilde{Q}:=[0, b)^{n}+a$ with $a \in \mathbb{R}^{n}$ and $b>0$ ), we define $Q:=\widetilde{Q} \times \mathbb{R}^{d-n}$. This type of set will be called $v$-cube ("vertical" cube). We denote by $\ell(Q)$ and $\widetilde{z}_{Q}$ the side length and center of $\widetilde{Q}$, respectively, and given $\lambda>0$ we set $\lambda Q:=\lambda \widetilde{Q} \times \mathbb{R}^{d-n}$. Let $\widetilde{\mathcal{D}}$ denote the standard dyadic lattice of $\mathbb{R}^{n}$, and set $\mathcal{D}:=\{Q: \widetilde{Q} \in \widetilde{\mathcal{D}}\}$. It is easy to check that the v-cubes of $\mathcal{D}$ intersected with $\Gamma$ provide a dyadic lattice associated to the graph $\Gamma$ in the sense of [Da, Appendix 1]. Finally, for $m \in \mathbb{Z}$, set $\mathcal{D}_{m}:=\left\{Q \in \mathcal{D}: \ell(Q)=2^{-m}\right\}$.

Fix a constant $C_{\Gamma}>10 \sqrt{n}(1+\operatorname{Lip}(A))$ (the precise value of $C_{\Gamma}$ will not be relevant in the proofs given in the paper). Given $1 \leq p \leq \infty$ and a v-cube $Q \subset \mathbb{R}^{d}$, we define the coefficient
$\beta_{p, \mu}(Q)$ as in (2.1) and (2.2) but replacing $2 Q$ by $C_{\Gamma} Q$. We also define $\alpha_{\mu}(Q)$ as in (2.4) but taking $B_{Q}:=B\left(\widetilde{z}_{Q}, C_{\Gamma} \ell(Q)\right) \times \mathbb{R}^{d-n} \subset \mathbb{R}^{d}$. This new definition of the $\alpha$ and $\beta$ coefficients (adapted to the graph $\Gamma$ ) is the one that we will use in the whole paper.

Remark 2.2. It is an exercise to check that, with this new definition of the $\alpha$ 's and $\beta$ 's, inequalities (2.5) and (2.6) of Theorem 2.1 still hold. Moreover, the following is an easy consequence of (2.5) and (2.6): Let $\Gamma$ be an $n$-dimensional Lipschitz graph, $f \in L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$ such that $f(x) \approx 1$ for almost all $x \in \Gamma$, and $\mu=f \mathcal{H}_{\Gamma}^{n}$. Let $1 \leq p \leq 2$. Given $C_{1}, C_{2}, C_{3} \geq 1$, there exists a constant $C_{4}>0$ such that, for any $R \in \mathcal{D}$,

$$
\sum_{Q \in \mathcal{D}: Q \subset C_{1} R}\left(\beta_{p, \mu}\left(C_{2} Q\right)^{2}+\alpha_{\mu}\left(C_{3} Q\right)^{2}\right) \mu(Q) \leq C_{4} \mu(R),
$$

and the dependence of $C_{4}$ with respect to $\Gamma$ is only on $\operatorname{Lip}(A)$.
Remark 2.3. It is shown in [To, Lemma 3.2], that $\beta_{1, \mu}(Q) \lesssim \alpha_{\mu}(Q)$ for all $Q \in \mathcal{D}$. Given $Q \in \mathcal{D}$, let $L_{Q}$ be a minimizing $n$-plane for $\alpha_{\mu}(Q)$. In general, $\beta_{\infty, \mu}(Q)$ can not be controlled by $\beta_{1, \mu}(Q)$, so given $x \in \operatorname{supp} \mu \cap C_{\Gamma} Q$, we can not control dist $\left(x, L_{Q}\right)$ by means of $\alpha_{\mu}(Q)$. But it is shown in [To, Lemma 5.2] that $\operatorname{dist}\left(x, L_{Q}\right) \lesssim \sum_{R \in \mathcal{D}: x \in R \subset Q} \alpha_{\mu}(R) \ell(R)$, and in particular, if $P \in \mathcal{D}$ is such that $P \subset Q$ and $x \in \operatorname{supp} \mu \cap C_{\Gamma} P$, and $L_{P}$ denotes a minimizing $n$-plane for $\alpha_{\mu}(P)$, one has (see [To, Remark 5.3])

$$
\begin{equation*}
\operatorname{dist}\left(x, L_{Q}\right) \lesssim \operatorname{dist}\left(x, L_{P}\right)+\sum_{R \in \mathcal{D}: P \subset R \subset Q} \alpha_{\mu}(R) \ell(R) . \tag{2.7}
\end{equation*}
$$

2.3. Martingales. First of all, let us recall a particular case of Lépingle's inequality (see [JSW], or [Lé] and [JKRW, Theorem 6.4] for martingales in a probability space):
Theorem 2.4. Let $(X, \Sigma, \lambda)$ be a $\sigma$-finite measure space and $\rho>2$. Then, there exist constants $C_{1}, C_{2}>0$ such that, for every martingale $\mathcal{G}:=\left\{G_{m}\right\}_{m \in \mathbb{Z}} \in L^{2}(\lambda)$,

$$
\left\|\mathcal{V}_{\rho}(\mathcal{G})\right\|_{L^{2}(\lambda)} \leq C_{1}\|\mathcal{G}\|_{L^{2}(\lambda)} \quad \text { and } \quad\|\mathcal{O}(\mathcal{G})\|_{L^{2}(\lambda)} \leq C_{2}\|\mathcal{G}\|_{L^{2}(\lambda)}
$$

where $\|\mathcal{G}\|_{L^{2}(\lambda)}:=\sup _{m \in \mathbb{Z}}\left\|G_{m}\right\|_{L^{2}(\lambda)}$. The constants $C_{1}$ and $C_{2}$ do not depend on the measure $\lambda$, and $C_{2}$ neither depends on the fixed sequence that defines $\mathcal{O}$.

To prove Theorem 1.1, we need to introduce a particular martingale, and to review some known results.
Lemma 2.5. Fix a cube $\widetilde{P} \subset \mathbb{R}^{n}$ (not necessarily dyadic) and a Lipschitz graph $\Gamma:=\{x \in$ $\left.\mathbb{R}^{d}: x=(\widetilde{x}, A(\widetilde{x}))\right\}$ such that $\operatorname{supp} A \subset \widetilde{P}$. Consider the measure $\mu:=f \mathcal{H}_{\Gamma}^{n}$, where $f(x)=1$ for all $\widetilde{x} \in \widetilde{P}^{c}$ and $C_{0}^{-1} \leq f(x) \leq C_{0}$ for all $\widetilde{x} \in \widetilde{P}$, for some fixed constant $C_{0}>0$. Also set $P:=\widetilde{P} \times \mathbb{R}^{d-n}$. Then, the following hold:

$$
\begin{gather*}
T_{*} \mu \in L_{l o c}^{1}(\mu), \quad T_{*}\left(\chi_{E} \mu\right) \in L_{l o c}^{1}(\mu) \text { for every compact set } E \subset \mathbb{R}^{d}, \text { and }  \tag{2.8}\\
\|T \mu\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2} . \tag{2.9}
\end{gather*}
$$

Remark 2.6. To avoid the problem of non-integrability near infinity, for this type of measures $\mu$ we redefine $T_{\epsilon} \mu(x):=\lim _{M \rightarrow \infty} \int \chi_{(\epsilon, M)}(|x-y|) K(x-y) d \mu(y)$, which exists because $\mu$ is flat outside a compact set and $K$ is odd. All the results in this paper remain valid with this new definition and the adjustments that have to be done in the proofs are minimal.

In this paper, we will deal with other integrals which concern the kernel $K$ and the measure $\mu$ near infinity. The non-integrability problem can be avoided in the same manner.

Proof of Lemma 2.5. It is known that the operator $T_{*}^{\mu}$ is bounded in $L^{2}(\mu)$, because $T_{*}^{\mu}$ is the maximal operator associated to a Calderón-Zygmund singular integral and $\mu$ is a uniformly rectifiable measure (see [DS1]). Thus, $T_{*}\left(\chi_{E} \mu\right)=T_{*}^{\mu}\left(\chi_{E}\right) \in L_{l o c}^{1}(\mu)$ for every compact set $E \subset \mathbb{R}^{d}$.

We are going to check that $\left\|T_{*} \mu\right\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2}$. This will imply that $T_{*} \mu \in L_{l o c}^{1}(\mu)$ and, since $T \mu$ exists (because $\mu$ is uniformly rectifiable) and $|T \mu| \leq T_{*} \mu$, we will also obtain $\|T \mu\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2}$; so the lemma will be proved.

Using that $T_{*}^{\mu}$ is bounded in $L^{2}(\mu)$, we have

$$
\begin{align*}
\left\|T_{*} \mu\right\|_{L^{2}(\mu)} & \leq\left\|T_{*}\left(\chi_{3 P} \mu\right)\right\|_{L^{2}(\mu)}+\left\|T_{*}\left(\chi_{(3 P)^{c}} \mu\right)\right\|_{L^{2}(\mu)} \\
& \lesssim \mu(P)^{1 / 2}+\left\|T_{*}\left(\chi_{(3 P)^{c}} \mu\right)\right\|_{L^{2}(\mu)} . \tag{2.10}
\end{align*}
$$

Set $L:=\mathbb{R}^{n} \times\{0\}^{d-n} \subset \mathbb{R}^{d}$; obviously $\chi_{P^{c}} \mu=\mathcal{H}_{L \backslash P}^{n}$. Since $L$ is an $n$-plane and $K$ is odd, $T_{*} \mathcal{H}_{L}^{n}(x)=0$ for all $x \in L$. Thus,

$$
\begin{equation*}
\left\|T_{*} \mathcal{H}_{L \backslash 3 P}^{n}\right\|_{L^{2}\left(\mathcal{H}_{L}^{n}\right)} \leq\left\|T_{*} \mathcal{H}_{L}^{n}\right\|_{L^{2}\left(\mathcal{H}_{L}^{n}\right)}+\left\|T_{*} \mathcal{H}_{L \cap 3 P}^{n}\right\|_{L^{2}\left(\mathcal{H}_{L}^{n}\right)} \lesssim \mu(P)^{1 / 2} \tag{2.11}
\end{equation*}
$$

Set $z_{P}:=\left(\widetilde{z}_{P}, 0, \ldots, 0\right) \in L$ (recall that $\widetilde{z}_{P}$ denotes the center of $\left.\widetilde{P}\right)$ and $\chi_{\epsilon}(x):=$ $\chi_{(\epsilon, \infty)}(|x|)$. It is obvious that $\int \chi_{\epsilon}\left(z_{P}-y\right) K\left(z_{P}-y\right) d \mathcal{H}_{L \backslash 3 P}^{n}(y)=0$ for all $\epsilon>0$. Thus, given $x \in \operatorname{supp} \mu \cap P$,

$$
\begin{aligned}
\left|\left(K \chi_{\epsilon} * \mathcal{H}_{L \backslash 3 P}^{n}\right)(x)\right| \leq & \int \chi_{\epsilon}(x-y)\left|K(x-y)-K\left(z_{P}-y\right)\right| d \mathcal{H}_{L \backslash 3 P}^{n}(y) \\
& +\int\left|\chi_{\epsilon}(x-y)-\chi_{\epsilon}\left(z_{P}-y\right)\right|\left|K\left(z_{P}-y\right)\right| d \mathcal{H}_{L \backslash 3 P}^{n}(y)
\end{aligned}
$$

Since $\Gamma$ is a Lipschitz graph, $\left|x-z_{P}\right| \lesssim \ell(P)$. So, the first term on right hand side of the previous inequality is easily bounded by an absolute constant independent of $\epsilon$, by standard arguments. For the second term, notice that $\operatorname{supp}\left(\chi_{\epsilon}(x-\cdot)-\chi_{\epsilon}\left(z_{P}-\cdot\right)\right) \cap(L \backslash 3 P)=\emptyset$ for all $\epsilon<\ell(P)$, and $\mathcal{H}_{L}^{n}\left(\left\{y \in \mathbb{R}^{n}: \chi_{\epsilon}(x-y)-\chi_{\epsilon}\left(z_{P}-y\right) \neq 0\right\}\right) \lesssim \ell(P) \epsilon^{n-1}$ for all $\epsilon \geq \ell(P)$. Therefore, since $\left|z_{P}-y\right| \approx \epsilon$ for all $y \in \operatorname{supp}\left(\chi_{\epsilon}(x-\cdot)-\chi_{\epsilon}\left(z_{P}-\cdot\right)\right) \cap(L \backslash 3 P)$, the second term can also be estimated by an absolute constant. Thus, we conclude $T_{*} \mathcal{H}_{L \backslash 3 P}^{n}(x)=$ $\sup _{\epsilon>0}\left|\left(K \chi_{\epsilon} * \mathcal{H}_{L \backslash 3 P}^{n}\right)(x)\right| \lesssim 1$ for all $x \in \operatorname{supp} \mu \cap P$.

Using the previous observations and (2.11), we have

$$
\begin{aligned}
\left\|T_{*}\left(\chi_{(3 P)^{c}} \mu\right)\right\|_{L^{2}(\mu)}^{2} & =\left\|T_{*} \mathcal{H}_{L \backslash 3 P}^{n}\right\|_{L^{2}\left(\chi_{P} \mu\right)}^{2}+\left\|T_{*} \mathcal{H}_{L \backslash 3 P}^{n}\right\|_{L^{2}\left(\chi_{P^{c}} \mu\right)}^{2} \\
& \leq\left\|T_{*} \mathcal{H}_{L \backslash 3 P}^{n}\right\|_{L^{2}\left(\chi_{P} \mu\right)}^{2}+\left\|T_{*} \mathcal{H}_{L \backslash 3 P}^{n}\right\|_{L^{2}\left(\mathcal{H}_{L}^{n}\right)}^{2} \lesssim \mu(P)
\end{aligned}
$$

which, combined with (2.10), gives $\left\|T_{*} \mu\right\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2}$, as desired.
We are ready to define the martingale. Let $P$ and $\mu$ be as in Lemma 2.5. Given $m \in \mathbb{Z}$ and $a \in \mathbb{R}^{n}$, we set

$$
\widetilde{D}_{m}^{a}:=a+\left[0,2^{-m}\right)^{n} \subset \mathbb{R}^{n} \quad \text { and } \quad D_{m}^{a}:=\widetilde{D}_{m}^{a} \times \mathbb{R}^{d-n} \subset \mathbb{R}^{d}
$$

Set $\mathcal{D}_{m}^{a}:=\left\{D_{m}^{a+2^{-m} k} \subset \mathbb{R}^{d}: k \in \mathbb{Z}^{n}\right\}$ (notice that $\mathcal{D}_{m}^{a}$ coincides with $\mathcal{D}_{m}$ translated by a parameter $a \in \mathbb{R}^{n}$ and, for a fixed $a, \bigcup_{m \in \mathbb{Z}} \mathcal{D}_{m}^{a}$ is a translation of the standard dyadic lattice). Notice that $\mu\left(D_{m}^{a}\right) \approx 2^{-m n}$ for all $m \in \mathbb{Z}, a \in \mathbb{R}^{n}$. For $D \in \mathcal{D}_{m}^{a}$ and $x \in D$, we set

$$
E_{D} \mu(x):=\frac{1}{\mu(D)} \int_{D} \int_{D^{c}} K(z-y) d \mu(y) d \mu(z)
$$

(take into account Remark 2.6 for the meaning of $\int_{D^{c}} K(z-y) d \mu(y)$ ). Finally, for $x \in \mathbb{R}^{d}$, we define the martingale $E_{m}^{a} \mu(x):=\sum_{D \in \mathcal{D}_{m}^{a}} \chi_{D}(x) E_{D} \mu(x), m \in \mathbb{Z}$.

Let us make some comments to understand better the nature of $E_{m}^{a} \mu$. First of all notice that, since $\mu(\partial D)=0$, for any $D \in \mathcal{D}_{m}^{a}$ and $\mu$-almost all $z \in D$ we have

$$
\begin{equation*}
\int_{D^{c}} K(z-y) d \mu(y)=\lim _{\epsilon \rightarrow 0} \int_{D^{c}} \chi_{\epsilon}(z-y) K(z-y) d \mu(y), \tag{2.12}
\end{equation*}
$$

and for any $\epsilon>0$, we have

$$
\begin{equation*}
\int_{D} \int_{D} \chi_{\epsilon}(z-y) K(z-y) d \mu(y) d \mu(z)=0 \tag{2.13}
\end{equation*}
$$

because of the antisymmetry of $K$. Therefore, by (2.12), (2.13), (2.8), and the dominated convergence theorem, $\int_{D}\left|\int_{D^{c}} K(z-y) d \mu(y)\right| d \mu(z)<\infty$ (in particular, we have seen that $E_{m}^{a} \mu$ is well defined) and $\int_{D} T\left(\chi_{D} \mu\right) d \mu=0$. Using this and (2.12), we finally have that

$$
\begin{equation*}
E_{m}^{a} \mu(x)=\frac{1}{\mu(D)} \int_{D} T\left(\chi_{D^{c}} \mu\right) d \mu=\frac{1}{\mu(D)} \int_{D} T \mu d \mu \tag{2.14}
\end{equation*}
$$

for $x \in D \in \mathcal{D}_{m}^{a}$, thus $E_{m}^{a} \mu(x)$ is the average of the function $T \mu$ on the v-cube $D \in \mathcal{D}_{m}^{a}$ which contains $x$. So, it is completely clear that, for a fixed $a \in \mathbb{R}^{n},\left\{E_{m}^{a} \mu\right\}_{m \in \mathbb{Z}}$ is a martingale. In [MV] it is shown that $\left\{E_{m}^{a} \mu\right\}_{m \in \mathbb{Z}}$ is well defined and it is a martingale without the assumption of the existence of $T \mu$ (i.e., for more general measures $\mu$ ).

Now, we can use (2.14), the $L^{2}$ boundedness of the dyadic maximal operator and (2.9) to deduce that

$$
\begin{equation*}
\left\|E_{m}^{a} \mu\right\|_{L^{2}(\mu)} \lesssim\|T \mu\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2} \tag{2.15}
\end{equation*}
$$

for all $a \in \mathbb{R}^{n}$ and $m \in \mathbb{Z}$, where the constants that appear in the previous inequalities only depend on $C_{0}, n, d$ and $\operatorname{Lip}(A)$.

Set $E^{a} \mu:=\left\{E_{m}^{a} \mu\right\}_{m \in \mathbb{Z}}$. Then, the martingale $E^{a} \mu$ belongs to $L^{2}(\mu)$ by (2.15); thus by Theorem 2.4, for all $a \in \mathbb{R}^{n}$,

$$
\begin{align*}
\left\|\mathcal{V}_{\rho}\left(E^{a} \mu\right)\right\|_{L^{2}(\mu)} & \lesssim\left\|E^{a} \mu\right\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2} \quad \text { for } \rho>2, \\
\left\|\mathcal{O}\left(E^{a} \mu\right)\right\|_{L^{2}(\mu)} & \lesssim\left\|E^{a} \mu\right\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2} \tag{2.16}
\end{align*}
$$

where the constants in the previous inequalities only depend on $C_{0}, n, d$, and $\operatorname{Lip}(A)$ (and on $\rho$, in the case of $\mathcal{V}_{\rho}$ ).

Finally, for $x \in \mathbb{R}^{d}$, we define

$$
E_{m} \mu(x):=2^{m n} \int_{\left\{a: x \in D_{m}^{a}\right\}} E_{m}^{a} \mu(x) d a
$$

(notice that $\left.\mathcal{L}^{n}\left(\left\{a: x \in D_{m}^{a}\right\}\right)=2^{-m n}\right)$. Thus, $E_{m} \mu$ is an average (of the $m^{\prime}$ 'th term) of some martingales depending on a parameter $a \in \mathbb{R}^{n}$.

Set $E \mu:=\left\{E_{m} \mu\right\}_{m \in \mathbb{Z}}$. We want to obtain estimates like (2.16) for $\mathcal{V}_{\rho}(E \mu)$ and $\mathcal{O}(E \mu)$. We will only show the details for $\mathcal{V}_{\rho}(E \mu)$, because the case of $\mathcal{O}(E \mu)$ follows by similar arguments.

One can easily check that $E_{m} \mu(x)=2^{M n} \int_{\left[0,2^{-M}\right]^{n}} E_{m}^{a} \mu(x) d a$ for all $m, M \in \mathbb{Z}$ with $M \leq m$. Therefore, for all $M, r, s \in \mathbb{Z}$ with $M \leq r \leq s$, we have

$$
\begin{equation*}
E_{r} \mu(x)-E_{s} \mu(x)=2^{M n} \int_{\left[0,2^{-M}\right]^{n}}\left(E_{r}^{a} \mu(x)-E_{s}^{a} \mu(x)\right) d a . \tag{2.17}
\end{equation*}
$$

Given $M \in \mathbb{Z}$, we consider the auxiliary transformation

$$
\mathcal{V}_{\rho, M}(E \mu)(x):=\sup _{\left\{r_{m}\right\}}\left(\sum_{m \in \mathbb{Z}}\left|E_{r_{m+1}} \mu(x)-E_{r_{m}} \mu(x)\right|^{\rho}\right)^{1 / \rho},
$$

where the pointwise supremum is taken over all decreasing sequences of integers $\left\{r_{m}\right\}_{m \in \mathbb{Z}}$ such that $r_{m} \geq M$ for all $m \in \mathbb{Z}$. With this definition it is obvious that the sequence $\left\{\mathcal{V}_{\rho, M}(E \mu)(x)\right\}_{M \in \mathbb{Z}}$ is non increasing and $\mathcal{V}_{\rho}(E \mu)(x)=\lim _{M \rightarrow-\infty} \mathcal{V}_{\rho, M}(E \mu)(x)$ for all $x \in$ $\mathbb{R}^{d}$. Minkowski's integral inequality and (2.17) yield the pointwise estimate

$$
\begin{aligned}
\mathcal{V}_{\rho, M}(E \mu)(x) & =\sup _{\left\{r_{m}\right\}: r_{m} \geq M}\left(\sum_{m \in \mathbb{Z}}\left|E_{r_{m+1}} \mu(x)-E_{r_{m}} \mu(x)\right|^{\rho}\right)^{1 / \rho} \\
& \leq 2^{M n} \int_{\left[0,2^{-M}\right]^{n}} \sup _{\left\{r_{m}\right\}}\left(\sum_{m \in \mathbb{Z}}\left|E_{r_{m+1}}^{a} \mu(x)-E_{r_{m}}^{a} \mu(x)\right|^{\rho}\right)^{1 / \rho} d a \\
& =2^{M n} \int_{\left[0,2^{-M]^{n}}\right.} \mathcal{V}_{\rho}\left(E^{a} \mu\right)(x) d a .
\end{aligned}
$$

Therefore, by the previous estimate, Minkowski's integral inequality and (2.16),

$$
\left\|\mathcal{V}_{\rho, M}(E \mu)\right\|_{L^{2}(\mu)} \leq 2^{M n} \int_{\left[0,2^{-M]^{n}}\right.}\left\|\mathcal{V}_{\rho}\left(E^{a} \mu\right)\right\|_{L^{2}(\mu)} d a \leq C \mu(P)^{1 / 2}
$$

where $C>0$ only depends on $C_{0}, n, d, \operatorname{Lip}(A)$, and $\rho$. By the monotone convergence theorem, we conclude that $\left\|\mathcal{V}_{\rho}(E \mu)\right\|_{L^{2}(\mu)} \lesssim \mu(P)^{1 / 2}$. Thus we have proved the following theorem (which can be considered the starting point to prove Theorem 1.1):
Theorem 2.7. Fix a cube $\widetilde{P} \subset \mathbb{R}^{n}$. Set $\Gamma:=\left\{x \in \mathbb{R}^{d}: x=(\widetilde{x}, A(\widetilde{x}))\right\}$, where $A: \mathbb{R}^{n} \rightarrow$ $\mathbb{R}^{d-n}$ is a Lipschitz function supported in $\widetilde{P}$, and set $P:=\widetilde{P} \times \mathbb{R}^{d-n}$. Set $\mu:=f \mathcal{H}_{\Gamma}^{n}$, where $f(x)=1$ for all $\widetilde{x} \in \widetilde{P}^{c}$ and $C_{0}^{-1} \leq f(x) \leq C_{0}$ for all $\widetilde{x} \in \widetilde{P}$, for some constant $C_{0}>0$.

Let $\rho>2$. Then, there exist constants $C_{1}, C_{2}>0$ such that $\left\|\mathcal{V}_{\rho}(E \mu)\right\|_{L^{2}(\mu)} \leq C_{1} \mu(P)^{1 / 2}$ and $\|\mathcal{O}(E \mu)\|_{L^{2}(\mu)} \leq C_{2} \mu(P)^{1 / 2}$, where $C_{1}$ and $C_{2}$ only depend on $C_{0}$, $n$, $d$, and $\operatorname{Lip}(A)$ (and on $\rho$ in the case of $C_{1}$ ).

We need to introduce additional notation in order to express $E_{m} \mu$ in a more convenient way for our purposes. Let $\mu_{1}, \ldots, \mu_{k}$ be a finite collection of positive Borel measures such that $\mu_{l}\left(D_{m}^{a}\right)>0$ for all $a \in \mathbb{R}^{n}, m \in \mathbb{Z}$ and $l=1, \ldots, k$. Given $m \in \mathbb{Z}$ and $x_{1}, \ldots, x_{i}, y_{1}, \ldots, y_{j} \in$ $\mathbb{R}^{d}$, we define

$$
\Lambda_{m}^{\mu_{1}, \ldots, \mu_{k}}\left(x_{1}, \ldots, x_{i} ; y_{1}, \ldots, y_{j}\right):=2^{n m} \int_{\left\{a: x_{1}, \ldots, x_{i} \in D_{m}^{a}, y_{1}, \ldots, y_{j} \notin D_{m}^{a}\right\}} \frac{d a}{\prod_{l=1}^{k} \mu_{l}\left(D_{m}^{a}\right)} .
$$

Then, by Fubini's theorem,

$$
\begin{align*}
E_{m} \mu(x) & =\int_{\left\{a: x \in D_{m}^{a}\right\}} \frac{2^{m n}}{\mu\left(D_{m}^{a}\right)} \int_{D_{m}^{a}} \int_{\left(D_{m}^{a}\right)^{c}} K(z-y) d \mu(y) d \mu(z) d a \\
& =\iint\left(2^{m n} \int_{\left\{a: x, z \in D_{m}^{a}, y \notin D_{m}^{a}\right\}} \frac{d a}{\mu\left(D_{m}^{a}\right)}\right) K(z-y) d \mu(z) d \mu(y)  \tag{2.18}\\
& =\iint \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y) .
\end{align*}
$$

## 3. Sketch of the proof of Theorem 1.1

The proof relies on two basic facts: the known $L^{2}$ boundedness of the $\rho$-variation and oscillation of martingales explained in the previous section and the good geometric properties of Lipschitz graphs from a measure-theoretic point of view.

As we said above, the starting point of the proof is Theorem 2.7 , where the $L^{2}$ boundedness of the $\rho$-variation and oscillation (of a convex combination) of some particular martingales is stated. So, the first step consists in relating the results on martingales in Theorem 2.7 with the $\rho$-variation and oscillation of singular integrals on Lipschitz graphs, and this is the aim of the following two theorems:

Theorem 3.1. Let $\Gamma$ and $\mu$ be as in Theorem 2.7. For each $x \in \Gamma$, define

$$
\begin{equation*}
W \mu(x)^{2}:=\sum_{m \in \mathbb{Z}}\left|\left(K \varphi_{2^{-m}} * \mu\right)(x)-E_{m} \mu(x)\right|^{2} \tag{3.1}
\end{equation*}
$$

Then, $\|W \mu\|_{L^{2}(\mu)}^{2} \leq C_{1} \sum_{Q \in \mathcal{D}}\left(\alpha_{\mu}\left(C_{2} Q\right)^{2}+\beta_{2, \mu}(Q)^{2}\right) \mu(Q)$, where $C_{1}, C_{2}>0$ depend only on $C_{0}, n, d, K$, and $\operatorname{Lip}(A)$.

Theorem 3.2. Let $\Gamma$ and $\mu$ be as in Theorem 2.7. For each $x \in \Gamma$, define

$$
\begin{equation*}
S \mu(x)^{2}:=\sup _{\left\{\epsilon_{m}\right\}} \sum_{j \in \mathbb{Z}} \sum_{m \in \mathbb{Z}: \epsilon_{m}, \epsilon_{m+1} \in I_{j}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{2}, \tag{3.2}
\end{equation*}
$$

where $I_{j}=\left[2^{-j-1}, 2^{-j}\right)$ and the supremum is taken over all decreasing sequences of positive numbers $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$. Then, $\|S \mu\|_{L^{2}(\mu)}^{2} \leq C \sum_{Q \in \mathcal{D}}\left(\alpha_{\mu}(Q)^{2}+\beta_{2, \mu}(Q)^{2}\right) \mu(Q)$, where $C>0$ only depends on $C_{0}, n, d, K$, and $\operatorname{Lip}(A)$.

Two fundamental tools to study $W \mu$ and $S \mu$ are the $\alpha$ and $\beta$ coefficients, which will be used to measure the flatness of $\Gamma$ at different scales, in order to estimate the terms which appear in the sums in (3.1) and (3.2). This will be done in sections 4 and 5 . To use the $\alpha$ coefficients to relate the $\rho$-variation of martingales with the $\rho$-variation of singular integrals, it is a key fact that we are considering a "smooth" family like $\varphi$, because the $\alpha$ 's are defined in terms of Lipschitz functions but $T_{\epsilon}$ is defined by means of a rough truncation. Moreover, we are taking a truncation only on the first $n$-coordinates because the average of martingales that we are using is taken over the parameter $a \in \mathbb{R}^{n}$, using the v-cubes $D_{M}^{a}$ (see subsection 2.3).

Combining Theorem 3.1 and Theorem 3.2 with the $L^{2}$ estimates of the $\rho$-variation and oscillation on the average of martingales $E \mu$ in Theorem 2.7, we are able to obtain local $L^{2}$ estimates of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}}{ }_{\Gamma}^{n}$ and $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ when $\Gamma$ is any Lipschitz graph. More precisely, we separate the sum in the definition of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ into two parts, which are classically called short and long variation (and analogously for $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ ). The short variation corresponds to the sum $S \mu$ in Theorem 3.2 (here $\mu$ is a suitable modification of $\mathcal{H}_{\Gamma}^{n}$ ), where the indices run over $m \in \mathbb{Z}$ such that both $\epsilon_{m}$ and $\epsilon_{m+1}$ lie in the same dyadic interval, and can be handled using the $\alpha$ 's and $\beta$ 's. The long variation corresponds to the sum over the indices $m \in \mathbb{Z}$ such that $\epsilon_{m}$ and $\epsilon_{m+1}$ lie in different dyadic intervals, so one may assume that the $\epsilon_{m}$ 's are dyadic numbers. It is handled by comparing $K \varphi_{2^{-m}} * \mu$ with $E_{m} \mu$, and then using Theorem 3.1 and the fact the $\rho$-variation and oscillation of $E \mu$ are bounded in $L^{2}(\mu)$, by Theorem 2.7. This will be done in section 6 (see Theorem 6.1).

Using the local $L^{2}$ estimates of Theorem 6.1, combined with rather standard techniques in Calderón-Zygmund theory, in section 7 we obtain the $H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ and $L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow$ $\operatorname{BMO}\left(\mathcal{H}_{\Gamma}^{n}\right)$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ and $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$. Then, by interpolation, we obtain the $L^{p}$ boundedness of these operators in the whole range $1<p<\infty$, and in particular the $L^{2}$ boundedness (see Theorem 7.1). Moreover, [CJRW2, Theorem B] can be adapted to prove that the $L^{2}\left(\mathcal{H}_{\Gamma}^{n}\right)$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ and $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ also yields the boundedness of these operators from $L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ to $L^{1, \infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$.

Let us stress that almost all the estimates in the proof of Theorem 1.1 (in particular, the constants involved in the relationships $\lesssim, \gtrsim$ and $\approx)$ depend either on $n, d, K$ or $\operatorname{Lip}(A)$, and possibly on other variables such as $\rho$ or $p$.

## 4. Proof of Theorem 3.1

In order to study the difference $\left(K \varphi_{2-m} * \mu\right)(x)-E_{m} \mu(x)$, we are going to split $E_{m} \mu(x)$ into two parts, the one we will compare with $\left(K \varphi_{2^{-m}} * \mu\right)(x)$ (which corresponds to integrate, in the definition of $E_{m} \mu(x)$, over the points $y \in \mathbb{R}^{d}$ such that $\left.2^{-m} \lesssim|\widetilde{x}-\widetilde{y}|\right)$, and the remaining part. Then, we will estimate each part of $\left(K \varphi_{2-m} * \mu\right)(x)-E_{m} \mu(x)$ separately, using the cancelation properties of the kernel $K$ and the uniform rectifiability of $\mu$.

Recall from (2.18) that $E_{m} \mu(x)=\iint \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y)$. Given $\epsilon>0$, we set $\gamma_{\epsilon}:=1-\varphi_{\epsilon}$. Then,

$$
\begin{aligned}
E_{m} \mu(x)= & \iint \varphi_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y) \\
& +\iint \gamma_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y)
\end{aligned}
$$

The first term in the previous sum is the one that we will compare with $\left(K \varphi_{2-m} * \mu\right)(x)$. For all $a \in \mathbb{R}^{n}$ such that $x \in D_{m}^{a}$, we have $\operatorname{supp} \varphi_{2^{-m}}(x-\cdot) \cap D_{m}^{a}=\emptyset$, and thus $\left(K \varphi_{2-m} * \mu\right)(x)=$ $\left(K \varphi_{2^{-m}} *\left(\chi_{\left(D_{m}^{a}\right)^{c}} \mu\right)\right)(x)$. Hence, using Fubini's theorem and the definition of $\Lambda_{m}^{\mu}(x, z ; y)$,

$$
\begin{aligned}
\left(K \varphi_{2^{-m}} * \mu\right)(x) & =2^{m n} \int_{\left\{a: x \in D_{m}^{a}\right\}}\left(K \varphi_{2^{-m}} *\left(\chi_{\left(D_{m}^{a}\right)^{c}} \mu\right)\right)(x) d a \\
& =2^{m n} \int_{\left\{a: x \in D_{m}^{a}\right\}} \mu\left(D_{m}^{a}\right)^{-1} \int_{D_{m}^{a}}\left(K \varphi_{2^{-m}} *\left(\chi_{\left(D_{m}^{a}\right)^{c}} \mu\right)\right)(x) d \mu(z) d a \\
& =\iint \varphi_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) K(x-y) d \mu(z) d \mu(y) .
\end{aligned}
$$

We can decompose $\left(K \varphi_{2^{-m}} * \mu\right)(x)-E_{m} \mu(x)$ as

$$
\begin{align*}
\left(K \varphi_{2^{-m}} *\right. & \mu)(x)-E_{m} \mu(x) \\
= & \iint \varphi_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y)(K(x-y)-K(z-y)) d \mu(z) d \mu(y) \\
& -\iint \gamma_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y)  \tag{4.1}\\
= & \sum_{j<m} F_{j}^{m}(x)-\sum_{j \in \mathbb{Z}} G_{j}^{m}(x),
\end{align*}
$$

where

$$
\begin{align*}
F_{j}^{m}(x) & :=\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\mu}(x, z ; y)(K(x-y)-K(z-y)) d \mu(z) d \mu(y)  \tag{4.2}\\
G_{j}^{m}(x) & :=\iint \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y) \tag{4.3}
\end{align*}
$$

Fix a v-cube $D \in \mathcal{D}_{m}$, for some $m \in \mathbb{Z}$. In subsection 4.1 (see (4.18)) we will prove that

$$
\begin{equation*}
\sum_{j<m}\left|F_{j}^{m}(x)\right| \lesssim \frac{\operatorname{dist}\left(x, L_{D}\right)}{\ell(D)}+\sum_{Q \in \mathcal{D}: D \subset Q} \frac{\ell(D)}{\ell(Q)} \alpha(Q) \tag{4.4}
\end{equation*}
$$

for all $x \in D \cap \Gamma$, where $L_{D}$ denotes an $n$-plane that minimizes $\alpha(D)$, and in subsection 4.2 (see (4.37)) we will prove that there exists a constant $C_{b}>1$ such that

$$
\begin{equation*}
\sum_{j \in \mathbb{Z}}\left|G_{j}^{m}(x)\right| \lesssim \alpha\left(C_{b} D\right)+\sum_{Q \in \mathcal{D}: Q \subset C_{b} D} \frac{\ell(Q)^{n+1}}{\ell(D)^{n+1}} \alpha(Q) \tag{4.5}
\end{equation*}
$$

for all $x \in D \cap \Gamma$. Assuming that these estimates hold, by (4.1),

$$
\begin{align*}
\|W \mu\|_{L^{2}(\mu)}^{2}= & \sum_{m \in \mathbb{Z}} \sum_{D \in \mathcal{D}_{m}} \int_{D}\left|\left(K \varphi_{2^{-m}} * \mu\right)(x)-E_{m} \mu(x)\right|^{2} d \mu(x) \\
\lesssim & \sum_{D \in \mathcal{D}} \int_{D}\left(\frac{\operatorname{dist}\left(x, L_{D}\right)}{\ell(D)}\right)^{2} d \mu(x)+\sum_{D \in \mathcal{D}}\left(\sum_{\substack{Q \in \mathcal{D}: \\
D \subset Q}} \frac{\ell(D)}{\ell(Q)} \alpha(Q)\right)^{2} \mu(D)  \tag{4.6}\\
& +\sum_{D \in \mathcal{D}} \alpha\left(C_{b} D\right)^{2} \mu(D)+\sum_{D \in \mathcal{D}}\left(\sum_{\substack{Q \in \mathcal{D}: \\
Q \subset C_{b} D}} \frac{\ell(Q)^{n+1}}{\ell(D)^{n+1}} \alpha(Q)\right)^{2} \mu(D) \\
= & : W_{1} \mu+W_{2} \mu+W_{3} \mu+W_{4} \mu .
\end{align*}
$$

If $L_{D}^{1}$ and $L_{D}^{2}$ denote a minimizing $n$-plane for $\beta_{1}(D)$ and $\beta_{2}(D)$, respectively, one can show that dist $\left.\mathcal{H}^{( } L_{D} \cap C_{\Gamma} D, L_{D}^{1} \cap C_{\Gamma} D\right) \lesssim \alpha(D) \ell(D)$ and $\operatorname{dist}_{\mathcal{H}}\left(L_{D}^{1} \cap C_{\Gamma} D, L_{D}^{2} \cap C_{\Gamma} D\right) \lesssim \beta_{2}(D) \ell(D)$. This easily implies that, for $x \in D \cap \Gamma, \operatorname{dist}\left(x, L_{D}\right) \lesssim \operatorname{dist}\left(x, L_{D}^{2}\right)+\beta_{2}(D) \ell(D)+\alpha(D) \ell(D)$, so $W_{1} \mu \lesssim \sum_{D \in \mathcal{D}}\left(\alpha(D)^{2}+\beta_{2}(D)^{2}\right) \mu(D)$.

By Cauchy-Schwarz inequality,

$$
\begin{aligned}
W_{2} \mu & \leq \sum_{D \in \mathcal{D}} \mu(D)\left(\sum_{Q \in \mathcal{D}: D \subset Q} \frac{\ell(D)}{\ell(Q)} \alpha(Q)^{2}\right)\left(\sum_{Q \in \mathcal{D}: D \subset Q} \frac{\ell(D)}{\ell(Q)}\right) \\
& \approx \sum_{D \in \mathcal{D}} \sum_{Q \in \mathcal{D}: D \subset Q} \frac{\ell(D)^{n+1}}{\ell(Q)} \alpha(Q)^{2} \approx \sum_{Q \in \mathcal{D}} \alpha(Q)^{2} \mu(Q),
\end{aligned}
$$

and also

$$
\begin{aligned}
W_{4} \mu & \leq \sum_{D \in \mathcal{D}} \mu(D)\left(\sum_{Q \in \mathcal{D}: Q \subset C_{b} D} \frac{\ell(Q)^{n+1}}{\ell(D)^{n+1}} \alpha(Q)^{2}\right)\left(\sum_{Q \in \mathcal{D}: Q \subset C_{b} D} \frac{\ell(Q)^{n+1}}{\ell(D)^{n+1}}\right) \\
& \approx \sum_{D \in \mathcal{D}} \sum_{Q \in \mathcal{D}: Q \subset C_{b} D} \ell(Q)^{n} \frac{\ell(Q)}{\ell(D)} \alpha(Q)^{2} \lesssim \sum_{Q \in \mathcal{D}} \alpha(Q)^{2} \mu(Q) .
\end{aligned}
$$

Therefore, using (4.6) and that $\alpha(Q) \lesssim \alpha\left(C_{b} Q\right)$, we conclude that

$$
\|W \mu\|_{L^{2}(\mu)}^{2} \lesssim \sum_{Q \in \mathcal{D}}\left(\alpha\left(C_{b} Q\right)^{2}+\beta_{2}(Q)^{2}\right) \mu(Q)
$$

and the theorem follows. It only remains to prove (4.4) and (4.5).
4.1. Estimate of $\sum_{j<m} F_{j}^{m}(x)$ when $x \in D \cap \Gamma$ for some $D \in \mathcal{D}_{m}$. Assume that $x \in$ $D \cap \Gamma$ for some $D \in \mathcal{D}_{m}$ and $j<m$. Let $L_{D}$ be an $n$-plane that minimizes $\alpha(D)$ and let $\sigma_{D}:=c_{D} \mathcal{H}_{L_{D}}^{n}$ be a minimizing measure of $\alpha(D)$. Let $L_{D}^{x}$ be the $n$-plane parallel to $L_{D}$ that contains $x$ and set $\sigma_{D}^{x}:=c_{D} \mathcal{H}_{L_{D}^{x}}^{n}$.

Notice that, because of $x \in L_{D}^{x}$, the antisymmetry of $\varphi_{2^{-j-1}}^{2-j} K$, and since $j<m$ (so, if $x \in D_{m}^{a}$ and $y \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(x-\cdot)$, then $\left.y \notin D_{m}^{a}\right)$, we have

$$
\begin{align*}
0 & =\int \varphi_{2^{-j-1}}^{2^{-j}}(x-y) K(x-y) d \sigma_{D}^{x}(y) \\
& =\int_{\left\{a: x \in D_{m}^{a}\right\}} \frac{2^{m n}}{\sigma_{D}^{x}\left(D_{m}^{a}\right)} \int_{D_{m}^{a}} \int_{\left(D_{m}^{a}\right)^{c}} \varphi_{2^{-j-1}}^{2^{-j}}(x-y) K(x-y) d \sigma_{D}^{x}(y) d \sigma_{D}^{x}(z) d a  \tag{4.7}\\
& =\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y) K(x-y) d \sigma_{D}^{x}(z) d \sigma_{D}^{x}(y)
\end{align*}
$$

Given $a \in \mathbb{R}^{n}$, let $b:=a+\left\{2^{-m-1}\right\}^{n} \in \mathbb{R}^{n}$ be the center of $\widetilde{D}_{m}^{a}$. For $u \in \mathbb{R}^{n}$ we denote $\|u\|_{\infty}:=\max _{i=1, \ldots, n}\left|u^{i}\right|$. Then, given $t \in \mathbb{R}^{d}$, it is clear that $t \in \overline{D_{m}^{a}}$ if and only if $\|\tilde{t}-b\|_{\infty} \leq$ $2^{-m}$. Using that $\sigma_{D}^{x}$ is a Hausdorff measure on an $n$-plane, that $K$ is antisymmetric and that $\varphi_{2^{-j-1}}^{2^{-j}}$ is symmetric, one can show that

$$
0=\int_{\|\widetilde{x}-b\|_{\infty} \leq 2^{-m}} \int_{\|\tilde{z}-b\|_{\infty} \leq 2^{-m}} \int_{\|\widetilde{y}-b\|_{\infty}>2^{-m}} \varphi_{2^{-j-1}}^{2^{-j}}(x-y) K(z-y) d \sigma_{D}^{x}(y) d \sigma_{D}^{x}(z) d b
$$

By the change of variable $b=a+\left\{2^{-m-1}\right\}^{n}$, it is easy to see that this triple integral is equal to $\int_{\left\{a: x \in D_{m}^{a}\right\}} \int_{D_{m}^{a}} \int_{\left(D_{m}^{a}\right)^{c}} \varphi_{2^{-j-1}}^{2^{-j}}(x-y) K(z-y) d \sigma_{D}^{x}(y) d \sigma_{D}^{x}(z) d a$. Thus, since $\sigma_{D}^{x}\left(D_{m}^{a}\right)$ does not depend on $a \in \mathbb{R}^{n}$ because $\sigma_{D}^{x}$ is flat,

$$
\begin{align*}
0 & =\int_{\left\{a: x \in D_{m}^{a}\right\}} \frac{2^{m n}}{\sigma_{D}^{x}\left(D_{m}^{a}\right)} \int_{D_{m}^{a}} \int_{\left(D_{m}^{a} c^{c}\right.} \varphi_{2^{-j-1}}^{2^{-j}}(x-y) K(z-y) d \sigma_{D}^{x}(y) d \sigma_{D}^{x}(z) d a  \tag{4.8}\\
& =\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y) K(z-y) d \sigma_{D}^{x}(z) d \sigma_{D}^{x}(y) .
\end{align*}
$$

By (4.7) and (4.8), we conclude that

$$
\begin{equation*}
0=\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y)(K(x-y)-K(z-y)) d \sigma_{D}^{x}(z) d \sigma_{D}^{x}(y) \tag{4.9}
\end{equation*}
$$

By definition, it is clear that $\Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y)=\Lambda_{m}^{\sigma_{D}}(x, z ; y)$. Therefore, using (4.9), we can decompose

$$
\begin{equation*}
F_{j}^{m}(x)=F 1_{j}^{m}(x)+F 2_{j}^{m}(x)+F 3_{j}^{m}(x)+F 4_{j}^{m}(x), \tag{4.10}
\end{equation*}
$$

where

$$
\begin{align*}
& F 1_{j}^{m}(x):=\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\mu}(x, z ; y)  \tag{4.11}\\
& \quad(K(x-y)-K(z-y)) d\left(\mu-\sigma_{D}\right)(z) d \mu(y), \\
& F 2_{j}^{m}(x):=\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\mu}(x, z ; y)  \tag{4.12}\\
& \quad(K(x-y)-K(z-y)) d \sigma_{D}(z) d\left(\mu-\sigma_{D}\right)(y), \\
& F 3_{j}^{m}(x):=\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y)\left(\Lambda_{m}^{\mu}(x, z ; y)-\Lambda_{m}^{\left.\sigma_{D}(x, z ; y)\right)} \begin{array}{r}
(K(x-y)-K(z-y)) d \sigma_{D}(z) d \sigma_{D}(y), \\
F 4_{j}^{m}(x):=\iint \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y) \\
\quad(K(x-y)-K(z-y)) d\left(\sigma_{D} \times \sigma_{D}-\sigma_{D}^{x} \times \sigma_{D}^{x}\right)(z, y) .
\end{array}\right. \tag{4.13}
\end{align*}
$$

In the next subsections we will prove the following estimates:

$$
\begin{gather*}
\left|F 1_{j}^{m}(x)\right|+\left|F 3_{j}^{m}(x)\right| \lesssim 2^{j-m} \alpha(D),  \tag{4.15}\\
\left|F 2_{j}^{m}(x)\right| \lesssim 2^{j-m} \sum_{Q \in \mathcal{D}: D \subset Q, \ell(Q) \leq 2^{-j}} \alpha(Q),  \tag{4.16}\\
\left|F 4_{j}^{m}(x)\right| \lesssim 2^{j-m} \frac{\operatorname{dist}\left(x, L_{D}\right)}{\ell(D)} . \tag{4.17}
\end{gather*}
$$

Then, using (4.10), we will finally get that, for all $D \in \mathcal{D}_{m}$ and $x \in D \cap \Gamma$,

$$
\begin{align*}
\sum_{j<m}\left|F_{j}^{m}(x)\right| & \lesssim \frac{\operatorname{dist}\left(x, L_{D}\right)}{\ell(D)}+\sum_{j \leq m} 2^{j-m} \sum_{Q \in \mathcal{D}: D \subset Q, \ell(Q) \leq 2^{-j}} \alpha(Q) \\
& \lesssim \frac{\operatorname{dist}\left(x, L_{D}\right)}{\ell(D)}+\sum_{Q \in \mathcal{D}: D \subset Q} \frac{\ell(D)}{\ell(Q)} \alpha(Q), \tag{4.18}
\end{align*}
$$

which gives (4.4).
4.1.1. Estimate of $F 1_{j}^{m}(x)$. Notice that, if $|\widetilde{x}-\widetilde{z}|>2^{-m} \sqrt{n}$, there is no $a \in \mathbb{R}^{n}$ such that $x, z \in D_{m}^{a}$, and this means that $\Lambda_{m}^{\mu}(x, z ; y)=0$. Thus, we can assume that $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$. Therefore, if the constant $C_{\Gamma}$ (see the definition of the $\alpha$ 's in subsection 2.2) is big enough, $\operatorname{supp} \Lambda_{m}^{\mu}(x, \cdot ; y) \subset B_{D}$.

For $y, z \in \Gamma$ such that $y \in \operatorname{supp} \varphi_{2-j-1}^{2^{-j}}(x-\cdot), j<m$ and $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$ (so, in particular, $|x-z| \lesssim|x-y|)$, we have the following estimates:

$$
\begin{gathered}
|K(x-y)-K(z-y)| \lesssim|x-z||x-y|^{-n-1} \lesssim 2^{j(n+1)-m} \\
\left|\nabla_{z}(K(x-y)-K(z-y))\right|=\left|\nabla_{z} K(z-y)\right| \lesssim 2^{j(n+1)}
\end{gathered}
$$

Claim 4.1. We have $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m n}$ and $\left|\nabla_{z} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$ for all $x, y, z \in \mathbb{R}^{d}$.
Claim 4.1 and the subsequent ones $4.2, \ldots, 4.7$ will be proved in subsection 4.3 below.
Putting all these estimates together we obtain that

$$
\left|\nabla_{z}\left(\Lambda_{m}^{\mu}(x, z ; y)(K(x-y)-K(z-y))\right)\right| \lesssim 2^{j(n+1)+m n}
$$

and, since $\operatorname{supp} \Lambda_{m}^{\mu}(x, \cdot ; y) \subset B_{D}$, recalling the definition of $\operatorname{dist}_{B_{D}}$ in (2.3),

$$
\left|\int \Lambda_{m}^{\mu}(x, z ; y)(K(x-y)-K(z-y)) d\left(\mu-\sigma_{D}\right)(z)\right| \lesssim 2^{j(n+1)+m n} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) .
$$

We can use this last estimate in (4.11) to obtain

$$
\begin{aligned}
\left|F 1_{j}^{m}(x)\right| & \lesssim 2^{j(n+1)+m n} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \int \varphi_{2^{-j-1}}^{2^{-j}}(x-y) d \mu(y) \\
& \lesssim 2^{j+m n} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \approx 2^{j-m} \ell(D)^{-n-1} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \lesssim 2^{j-m} \alpha(D)
\end{aligned}
$$

which, together with the estimate of $\left|F 3_{j}^{m}(x)\right|$ in subsection 4.1.3, gives (4.15).
4.1.2. Estimate of $F 2_{j}^{m}(x)$. Arguing as in subsection 4.1.1, we can obtain the following estimates for $x, y, z$ as above:

$$
\begin{gather*}
\left|\varphi_{2^{-j-1}}^{2^{-j}}(x-y)\right| \leq 1 \quad \text { and } \quad\left|\nabla_{y} \varphi_{2^{-j-1}}^{2^{-j}}(x-y)\right| \lesssim 2^{j},  \tag{4.19}\\
|K(x-y)-K(z-y)| \lesssim|x-z||x-y|^{-n-1} \lesssim 2^{j(n+1)-m},  \tag{4.20}\\
\left|\nabla_{y}(K(x-y)-K(z-y))\right| \lesssim|x-z||x-y|^{-n-2} \lesssim 2^{j(n+2)-m} . \tag{4.21}
\end{gather*}
$$

Claim 4.2. For $j<m, y \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(x-\cdot)$, and $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$, the following hold: $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m n}$ and $\nabla_{y} \Lambda_{m}^{\mu}(x, z ; y)=0$.

Notice that the first estimate in Claim 4.2 is the same as the first one in Claim 4.1.
Let $D_{j} \in \mathcal{D}_{j}$ be the unique dyadic v-cube with $\ell\left(D_{j}\right)=2^{-j}$ which contains $D$. Then, $\operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(x-\cdot) \subset B_{D_{j}}$ for $C_{\Gamma}$ big enough. Therefore, we can use the previous estimates to see that the gradient of the term inside the integral with respect to $y$ in (4.12) is bounded by $2^{j(n+2)+m(n-1)}$ and is supported in $B_{D_{j}}$, and then by (2.3) we derive that

$$
\begin{align*}
\left|F 2_{j}^{m}(x)\right| & \leq \int \mid \int \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) \\
& (K(x-y)-K(z-y)) d\left(\mu-\sigma_{D}\right)(y) \mid d \sigma_{D}(z)  \tag{4.22}\\
& \lesssim \int_{|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}} 2^{j(n+2)+m(n-1)} \operatorname{dist}_{B_{D_{j}}}\left(\mu, \sigma_{D}\right) d \sigma_{D}(z) \\
& \lesssim 2^{j(n+2)-m} \operatorname{dist}_{B_{D_{j}}}\left(\mu, \sigma_{D}\right)
\end{align*}
$$

We shall estimate $\operatorname{dist}_{B_{D_{j}}}\left(\mu, \sigma_{D}\right)$ in terms of the $\alpha$ coeficients. Consider the unique sequence of dyadic v-cubes $D=: D_{m} \subset \ldots \subset D_{i+1} \subset D_{i} \subset \ldots \subset D_{j}$ such that each $D_{i}$ belongs to $\mathcal{D}_{i}$, for $i=j, \ldots, m$. Let $L_{D_{i}}$ be an $n$-plane that gives the minimum in the definition of $\alpha\left(D_{i}\right)$ and let $\sigma_{D_{i}}:=c_{D_{i}} d \mathcal{H}_{L_{D_{i}}}^{n}$ be a minimizing measure. We will prove that

$$
\begin{equation*}
\operatorname{dist}_{B_{D_{j}}}\left(\mu, \sigma_{D}\right) \lesssim 2^{-j(n+1)} \sum_{i=j}^{m-1} \alpha\left(D_{i}\right) \tag{4.23}
\end{equation*}
$$

Combining (4.23) with (4.22), we will finally obtain that $\left|F 2_{j}^{m}(x)\right| \lesssim 2^{j-m} \sum_{i=j}^{m-1} \alpha\left(D_{i}\right)$, which gives (4.16).

Let us prove (4.23). By the triangle inequality,

$$
\begin{aligned}
\operatorname{dist}_{B_{D_{j}}}\left(\mu, \sigma_{D}\right) & \leq \operatorname{dist}_{B_{D_{j}}}\left(\mu, \sigma_{D_{j}}\right)+\sum_{i=j}^{m-1} \operatorname{dist}_{B_{D_{j}}}\left(\sigma_{D_{i}}, \sigma_{D_{i+1}}\right) \\
& \lesssim 2^{-j(n+1)} \alpha\left(D_{j}\right)+\sum_{i=j}^{m-1} \operatorname{dist}_{B_{D_{j}}}\left(\sigma_{D_{i}}, \sigma_{D_{i+1}}\right)
\end{aligned}
$$

so we are reduced to prove that, for all $i=j, \ldots, m-1$,

$$
\begin{equation*}
\operatorname{dist}_{B_{D_{j}}}\left(\sigma_{D_{i}}, \sigma_{D_{i+1}}\right) \lesssim 2^{-j(n+1)} \alpha\left(D_{i}\right) \tag{4.24}
\end{equation*}
$$

By definition, $\operatorname{dist}_{B_{D_{j}}}\left(\sigma_{D_{i}}, \sigma_{D_{i+1}}\right)=\sup \left|\int g d\left(c_{D_{i}} \mathcal{H}_{L_{D_{i}}}^{n}-c_{D_{i+1}} \mathcal{H}_{L_{D_{i+1}}}^{n}\right)\right|$, where the supremum is taken over all Lipschitz functions $g$ supported in $B_{D_{j}}$ such that $\operatorname{Lip}(g) \leq 1$. Fix one of such Lipschitz functions $g$. Then,

$$
\begin{align*}
\int g d\left(c_{D_{i}} \mathcal{H}_{L_{D_{i}}}^{n}-c_{D_{i+1}} \mathcal{H}_{L_{D_{i+1}}}^{n}\right)= & \left(c_{D_{i}}-c_{D_{i+1}}\right) \int g d \mathcal{H}_{L_{D_{i}}}^{n}  \tag{4.25}\\
& +c_{D_{i+1}} \int g d\left(\mathcal{H}_{L_{D_{i}}}^{n}-\mathcal{H}_{L_{D_{i+1}}}^{n}\right)
\end{align*}
$$

It is shown in [To, Lemma 3.4] that $\left|c_{D_{i}}-c_{D_{i+1}}\right| \lesssim \alpha\left(D_{i}\right)$, so the first term on the right hand side of (4.25) is bounded in absolute value by $C 2^{-j(n+1)} \alpha\left(D_{i}\right)$.

In order to estimate the second term of the right hand side of (4.25), set $L_{D_{i+1}}=$ $\left\{(\widetilde{t}, a(\widetilde{t})) \in \mathbb{R}^{d}: \widetilde{t} \in \mathbb{R}^{n}\right\}$ (where $a: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ is an appropriate affine map), and let $p: L_{D_{i}} \rightarrow L_{D_{i+1}}$ be the projection defined by $p(t):=(\widetilde{t}, a(\widetilde{t}))$. Since $\Gamma$ is a Lipschitz graph,
$a$ is well defined and $p$ is a homeomorphism. Let $p_{\sharp} \mathcal{H}_{L_{D_{i}}}^{n}$ be the image measure of $\mathcal{H}_{L_{D_{i}}}^{n}$ by $p$. It is easy to check that $\mathcal{H}_{L_{D_{i+1}}}^{n}=\tau p_{\sharp} \mathcal{H}_{L_{D_{i}}}^{n}$, where $\tau$ is some positive constant such that $|\tau-1| \lesssim \alpha\left(D_{i}\right)$ and $\tau \lesssim 1$. Therefore,

$$
\begin{align*}
\mid \int g d\left(\mathcal{H}_{L_{D_{i}}}^{n}-\right. & \left.\mathcal{H}_{L_{D_{i+1}}}^{n}\right)\left|=\left|\int(g(t)-\tau g(p(t))) d \mathcal{H}_{L_{D_{i}}}^{n}(t)\right|\right. \\
& \leq\left|\int(1-\tau) g(t) d \mathcal{H}_{L_{D_{i}}}^{n}(t)\right|+\left|\int \tau(g(t)-g(p(t))) d \mathcal{H}_{L_{D_{i}}}^{n}(t)\right|  \tag{4.26}\\
& \lesssim 2^{-j(n+1)} \alpha\left(D_{i}\right)+\int \mid\left(g(t)-g(p(t)) \mid d \mathcal{H}_{L_{D_{i}}}^{n}(t) .\right.
\end{align*}
$$

Since $g$ and $g \circ p$ are supported in $B_{D_{j}}$ and $g$ is 1-Lipschitz, by [To, Lemma 3.4],

$$
\begin{aligned}
\int|(g-g \circ p)| d \mathcal{H}_{L_{D_{i}}}^{n} & \lesssim \int_{B_{D_{j}}} \operatorname{dist}_{\mathcal{H}}\left(L_{D_{i}} \cap B_{D_{j}}, L_{D_{i+1}} \cap B_{D_{j}}\right) d \mathcal{H}_{L_{D_{i}}}^{n} \\
& \lesssim 2^{-j n} \operatorname{dist}_{\mathcal{H}}\left(L_{D_{i}} \cap B_{D_{j}}, L_{D_{i+1}} \cap B_{D_{j}}\right) \\
& \lesssim 2^{-j n} 2^{i-j} \operatorname{dist}_{\mathcal{H}}\left(L_{D_{i}} \cap B_{D_{i}}, L_{D_{i+1}} \cap B_{D_{i}}\right) \lesssim 2^{-j(n+1)} \alpha\left(D_{i}\right) .
\end{aligned}
$$

This last estimate together with (4.26) and the fact that $\left|c_{D_{i+1}}\right| \lesssim 1$ implies that the second term on the right hand side of (4.25) is also bounded in absolute value by $C 2^{-j(n+1)} \alpha\left(D_{i}\right)$. Therefore, to obtain (4.24) we only have to take the supremum in (4.25) over all admissible functions $g$.
4.1.3. Estimate of $F 3_{j}^{m}(x)$. Notice that, by Fubini's theorem,

$$
\begin{aligned}
\Lambda_{m}^{\mu}(x, z ; y) & -\Lambda_{m}^{\sigma_{D}}(x, z ; y)=2^{m n} \int_{\left\{a: x, z \in D_{m}^{a}, y \notin D_{m}^{a}\right\}}\left(\frac{1}{\mu\left(D_{m}^{a}\right)}-\frac{1}{\sigma_{D}\left(D_{m}^{a}\right)}\right) d a \\
& =2^{m n} \int_{\left\{a: x, z \in D_{m}^{a}, y \notin D_{m}^{a}\right\}} \frac{\sigma_{D}\left(D_{m}^{a}\right)-\mu\left(D_{m}^{a}\right)}{\mu\left(D_{m}^{a}\right) \sigma_{D}\left(D_{m}^{a}\right)} d a \\
& =2^{m n} \int_{\left\{a: x, z \in D_{m}^{a}, y \notin D_{m}^{a}\right\}}\left(\int_{t \in D_{m}^{a}} d\left(\sigma_{D}-\mu\right)(t)\right) \mu\left(D_{m}^{a}\right)^{-1} \sigma_{D}\left(D_{m}^{a}\right)^{-1} d a \\
& =\int \Lambda_{m}^{\mu, \sigma_{D}(x, z, t ; y) d\left(\sigma_{D}-\mu\right)(t) .}
\end{aligned}
$$

Since $\Lambda_{m}^{\mu, \sigma_{D}}(x, z, t ; y)=0$ if $|\widetilde{x}-\widetilde{t}|>2^{-m} \sqrt{n}$, we may assume that $\operatorname{supp} \Lambda_{m}^{\mu, \sigma_{D}}(x, z, \cdot ; y) \subset$ $B_{D}$ (by taking $C_{\Gamma}$ big enough).
Claim 4.3. We have $\left|\Lambda_{m}^{\mu, \sigma_{D}}(x, z, t ; y)\right| \lesssim 2^{2 m n}$ and $\left|\nabla_{t} \Lambda_{m}^{\mu, \sigma_{D}}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)}$ for all $x, y, z, t \in \mathbb{R}^{d}$.

Using Claim 4.3, we deduce that $\left|\Lambda_{m}^{\mu}(x, z ; y)-\Lambda_{m}^{\sigma_{D}}(x, z ; y)\right| \lesssim 2^{m(2 n+1)} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right)$, and then,

$$
\begin{aligned}
\left|F 3_{j}^{m}(x)\right| & \lesssim \int \varphi_{2^{-j-1}}^{2^{-j}}(x-y) \int_{|\tilde{x}-\tilde{z}| \leq 2^{-m} \sqrt{n}} 2^{m(2 n+1)} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \\
& \lesssim x-z| | x-\left.y\right|^{-n-1} d \sigma_{D}(z) d \sigma_{D}(y) \\
& \lesssim 2^{2 m n+j(n+1)} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \iint_{\substack{\tilde{\tilde{x}-\tilde{y} \mid \leq 2^{-j}}|\sqrt{n} \sqrt{n}\\
| x-\tilde{z} \mid \leq 2^{-m} \sqrt{n}}} d \sigma_{D}(z) d \sigma_{D}(y) \\
& \lesssim 2^{m n+j} \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \lesssim 2^{j-m} \alpha(D),
\end{aligned}
$$

which, together with the estimate of $\left|F 1_{j}^{m}(x)\right|$ (see subsection 4.1.1), gives (4.15).
4.1.4. Estimate of $F 4_{j}^{m}(x)$. Set $L_{D}=\left\{(\widetilde{y}, a(\widetilde{y})) \in \mathbb{R}^{d}: \widetilde{y} \in \mathbb{R}^{n}\right\}$, where $a: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ is an appropriate affine map, and let $p: L_{D}^{x} \rightarrow L_{D}$ be the projection defined by $p(y):=(\widetilde{y}, a(\widetilde{y}))$. Since $\Gamma$ is a Lipschitz graph, $a$ is well defined and $p$ is a homeomorphism. If $p_{\sharp} \sigma_{D}^{x}$ is the image measure of $\sigma_{D}^{x}$ under $p$, we obviously have $\sigma_{D}=p_{\sharp} \sigma_{D}^{x}$ because $L_{D}$ and $L_{D}^{x}$ differ by a translation. Therefore, since $\widetilde{p(y)}=\widetilde{y}$, (4.14) becomes

$$
\begin{aligned}
F 4_{j}^{m}(x)=\iint(K(x-p(y))-K(p(z)-p(y))-(K(x-y)-K(z-y))) \\
\varphi_{2-j-1}^{2^{-j}}(x-y) \Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y) d \sigma_{D}^{x}(z) d \sigma_{D}^{x}(y) .
\end{aligned}
$$

For $y, z \in L_{D}^{x}$ such that $\varphi_{2^{-j-1}}^{2-j}(x-y) \Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y) \neq 0$, we have $K(p(z)-p(y))-K(z-y)=0$, so we can estimate

$$
\begin{aligned}
\mid K(x-p(y))-K(p(z)-p(y)) & -(K(x-y)-K(z-y))|=|K(x-p(y))-K(x-y)| \\
& \lesssim \frac{|y-p(y)|}{|x-y|^{n+1}} \lesssim 2^{j(n+1)}|y-p(y)| \approx 2^{j(n+1)} \operatorname{dist}\left(x, L_{D}\right) .
\end{aligned}
$$

By the same arguments as in the proof of Claim 4.1, one can easily see that $\left|\Lambda_{m}^{\sigma_{D}^{x}}(x, z ; y)\right| \lesssim$ $2^{m n}$. Therefore,

$$
\begin{aligned}
\left|F 4_{j}^{m}(x)\right| & \lesssim 2^{j(n+1)} \operatorname{dist}\left(x, L_{D}\right) 2^{m n} \iint_{\substack{|\tilde{x}-\tilde{y}| \leq 2^{-j} 3_{\sqrt{n}} \\
\mid x-z} 2^{-m} \sqrt{n}} d \sigma_{D}^{x}(z) d \sigma_{D}^{x}(y) \\
& \lesssim 2^{j} \operatorname{dist}\left(x, L_{D}\right) \approx 2^{j-m} \operatorname{dist}\left(x, L_{D}\right) / \ell(D),
\end{aligned}
$$

which gives (4.17).
4.2. Estimate of $\sum_{j \in \mathbb{Z}} G_{j}^{m}(x)$ when $x \in D \cap \Gamma$ for some $D \in \mathcal{D}_{m}$. Assume that $x \in D$ for some $D \in \mathcal{D}_{m}$. Recall from (4.3) that

$$
G_{j}^{m}(x)=\iint \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) \Lambda_{m}^{\mu}(x, z ; y) K(z-y) d \mu(z) d \mu(y),
$$

where $0 \leq \gamma_{2^{-m}}(x-y) \leq 1,\left|\nabla_{y} \gamma_{2^{-m}}(x-y)\right| \lesssim 2^{m}$ for all $x, y \in \mathbb{R}^{d}$, and $\gamma_{2^{-m}}(x-y)=0$ whenever $|\widetilde{x}-\widetilde{y}|>2^{-m} 3 \sqrt{n}$. Notice that $\Lambda_{m}^{\mu}(x, z ; y)=0$ if $|\widetilde{x}-\widetilde{z}|>2^{-m} \sqrt{n}$, thus we can assume that $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$ and $|\widetilde{z}-\widetilde{y}| \leq 2^{-m+2} \sqrt{n}$ in the integral that defines $G_{j}^{m}(x)$. Hence, if $j \leq m-2, \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \Lambda_{m}^{\mu}(x, z ; y)=0$ for all $z, y \in \mathbb{R}^{d}$, because $\varphi_{2^{-j-1}}^{2^{-j}}(z-y)=0$ if $|\widetilde{z}-\widetilde{y}| \leq 2^{-j-1} 2.1 \sqrt{n}$, and $2^{-j-1} 2.1 \sqrt{n} \geq 2^{-m+2} \sqrt{n}$ when $j \leq m-2$. Therefore, $G_{j}^{m}(x)=0$ for $j \leq m-2$, and then

$$
\begin{equation*}
\sum_{j \in \mathbb{Z}} G_{j}^{m}(x)=\sum_{j \geq m-1} G_{j}^{m}(x) ; \tag{4.27}
\end{equation*}
$$

so, from now on, we assume that $j \geq m-1$.
Let $L_{D}$ be an $n$-plane that minimizes $\alpha(D)$ and let $\sigma_{D}:=c_{D} \mathcal{H}_{L_{D}}^{n}$ be a minimizing measure of $\alpha(D)$. As we did in the beginning of subsection 4.1, given $a \in \mathbb{R}^{n}$, let $b:=a+\left\{2^{-m-1}\right\}^{n} \in$ $\mathbb{R}^{n}$ be the center of $\widetilde{D}_{m}^{a}$. Recall that, for $t \in \mathbb{R}^{d}, t \in \overline{D_{m}^{a}}$ if and only if $\|\tilde{t}-b\|_{\infty} \leq 2^{-m}$. Using that $\sigma_{D}$ is a Hausdorff measure on an $n$-plane, that $K$ is antisymmetric and that $\varphi_{2^{-j-1}}^{2^{-j}}$ and
$\gamma_{2^{-m}}$ are symmetric, one can show that

$$
0=\int_{\|\tilde{x}-b\|_{\infty} \leq 2^{-m}} \int_{\|\tilde{z}-b\|_{\infty} \leq 2^{-m}} \int_{\|\tilde{y}-b\|_{\infty}>2^{-m}} \varphi_{2^{-j-1}}^{2^{-j}}(z-y) .
$$

By the change of variable $b=a+\left\{2^{-m-1}\right\}^{n}$, it is easy to see that this triple integral is equal to $\int_{\left\{a: x \in D_{m}^{a}\right\}} \int_{D_{m}^{a}} \int_{\left(D_{m}^{a}\right)^{c}} \varphi_{2^{-j-1}}^{2-j}(z-y) \gamma_{2^{-m}}(x-y) K(z-y) d \sigma_{D}(y) d \sigma_{D}(z) d a$. Thus, since $\sigma_{D}\left(D_{m}^{a}\right)$ does not depend on $a \in \mathbb{R}^{n}$ because $\sigma_{D}$ is flat,

$$
\begin{align*}
0 & =\int_{\left\{a: x \in D_{m}^{a}\right\}} \frac{2^{m n}}{\sigma_{D}\left(D_{m}^{a}\right)} \int_{D_{m}^{a}} \int_{\left(D_{m}^{a}\right)^{c}} \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \\
& =\iint \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) \Lambda_{m}^{\sigma_{D}}(x, z ; y) K(z-y) K(z-y) d \sigma_{D}(y) d \sigma_{D}(z) d a  \tag{4.28}\\
& (z) d \sigma_{D}(y) .
\end{align*}
$$

Let $\left\{\eta_{Q}\right\}_{Q \in \mathcal{D}_{j}}$ be a partition of the unity with respect to the v-cubes $Q \in \mathcal{D}_{j}$, i.e. $\eta_{Q}: \mathbb{R}^{d} \rightarrow$ $\mathbb{R}$ are $\mathcal{C}^{\infty}$ functions such that: $\chi_{0.9 Q} \leq \eta_{Q} \leq \chi_{1.1 Q},\left|\nabla \eta_{Q}\right| \lesssim \ell(Q)^{-1}=2^{j}, \sum_{Q \in \mathcal{D}_{j}} \eta_{Q}=1$ and $\eta_{Q}(y)=\eta_{Q}(\widetilde{y}, 0)$ for all $y \in \mathbb{R}^{d}$. It is easy to check that, if $j \geq m-1, Q \in \mathcal{D}_{j}$, and $\operatorname{supp} \eta_{Q} \cap \operatorname{supp} \gamma_{2-m}(x-\cdot) \neq \emptyset$, then $Q \subset C_{e} D$ for some absolute constant $C_{e}>1$.

Given $Q \in \mathcal{D}_{j}$, let $L_{Q}$ and $\sigma_{Q}:=c_{Q} \mathcal{H}_{L_{Q}}^{n}$ be a minimizing $n$-plane and measure for $\alpha(Q)$, respectively, and consider the measure

$$
\lambda:=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \eta_{Q} \sigma_{Q}
$$

By (4.28) and the properties of the partition of the unity $\left\{\eta_{Q}\right\}_{Q \in \mathcal{D}_{j}}$, for $j \geq m-1$ we can decompose $G_{j}^{m}(x)$ as

$$
\begin{equation*}
G_{j}^{m}(x)=G 1_{j}^{m}(x)+G 2_{j}^{m}(x)+G 3_{j}^{m}(x)+G 4_{j}^{m}(x)+G 5_{j}^{m}(x), \tag{4.29}
\end{equation*}
$$

where

$$
\begin{align*}
G 1_{j}^{m}(x) & :=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \iint \ldots d\left(\mu-\sigma_{Q}\right)(z) d \mu(y),  \tag{4.30}\\
G 2_{j}^{m}(x) & :=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \iint \ldots d \sigma_{Q}(z) d\left(\mu-\sigma_{Q}\right)(y),  \tag{4.31}\\
G 3_{j}^{m}(x) & :=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \iint \ldots d\left(\sigma_{Q} \times \sigma_{Q}-\sigma_{D} \times \sigma_{D}\right)(z, y), \tag{4.32}
\end{align*}
$$

where " $\ldots$ " stands for $\varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) \eta_{Q}(y) K(z-y) \Lambda_{m}^{\mu}(x, z ; y)$, and

$$
\begin{align*}
& G 4_{j}^{m}(x):=\iint \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) K(z-y)  \tag{4.33}\\
& \\
& \quad\left(\Lambda_{m}^{\mu}(x, z ; y)-\Lambda_{m}^{\lambda}(x, z ; y)\right) d \sigma_{D}(z) d \sigma_{D}(y)
\end{align*}
$$

$$
\begin{align*}
& G 5_{j}^{m}(x):=\iint \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) K(z-y)  \tag{4.34}\\
& \\
& \qquad\left(\Lambda_{m}^{\lambda}(x, z ; y)-\Lambda_{m}^{\sigma_{D}}(x, z ; y)\right) d \sigma_{D}(z) d \sigma_{D}(y) .
\end{align*}
$$

In the next subsections we will prove the following estimates:

$$
\begin{align*}
\quad\left|G 1_{j}^{m}(x)\right|+\left|G 2_{j}^{m}(x)\right|+\left|G 4_{j}^{m}(x)\right| & \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)} \alpha(Q),  \tag{4.35}\\
\left|G 3_{j}^{m}(x)\right|+\left|G 5_{j}^{m}(x)\right| & \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right), \tag{4.36}
\end{align*}
$$

where $C_{b}$ is some absolute constant bigger than $C_{e}$. Then, using (4.27) and (4.29), we will finally obtain that, for all $D \in \mathcal{D}_{m}$ and $x \in D \cap \Gamma$,

$$
\begin{align*}
\sum_{j \in \mathbb{Z}}\left|G_{j}^{m}(x)\right| & \lesssim \sum_{j \geq m-1} \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right) \\
& \lesssim \sum_{Q \in \mathcal{D}: Q \subset C_{e} D} \frac{\ell(Q)^{n+1}}{\ell(D)^{n+1}}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right)  \tag{4.3}\\
& \lesssim \alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: R \subset C_{b} D} \frac{\ell(R)^{n+1}}{\ell(D)^{n+1}} \alpha(R),
\end{align*}
$$

which gives (4.5).
4.2.1. Estimate of $G 1_{j}^{m}(x)$. If $\varphi_{2^{-j-1}}^{2^{-j}}(z-y) \neq 0$ then $2^{-j-1} 2.1 \sqrt{n} \leq|\widetilde{z}-\widetilde{y}| \leq 2^{-j} 3 \sqrt{n}$, so if we also have that $y \in \operatorname{supp} \eta_{Q}$, then $z \in 8 \sqrt{n} Q$ because $Q \in \mathcal{D}_{j}$. Therefore, we can assume that $\operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y) \eta_{Q}(y) \subset B_{Q}$ if $C_{\Gamma}$ is big enough.

Claim 4.4. For $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2-j}(\cdot-y)$, the following hold: $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)-j}$, $\left|\nabla_{z} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$, and $\left|\nabla_{y} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$.

We have that $|K(z-y)| \lesssim 2^{j n}$ and $\left|\nabla_{z} K(z-y)\right| \lesssim 2^{j(n+1)}$ for all $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$. Using (4.19) and the first two estimates in Claim 4.4, we get

$$
\left|\nabla_{z}\left(\varphi_{2^{-j-1}}^{2^{-j}}(z-y) K(z-y) \Lambda_{m}^{\mu}(x, z ; y)\right)\right| \lesssim 2^{m(n+1)+j n} .
$$

Therefore, for $y \in \operatorname{supp} \eta_{Q}$,

$$
\begin{aligned}
\left|\int \varphi_{2^{-j-1}}^{2^{-j}}(z-y) K(z-y) \Lambda_{m}^{\mu}(x, z ; y) d\left(\mu-\sigma_{Q}\right)(z)\right| & \lesssim 2^{m(n+1)+j n} \operatorname{dist}_{B_{Q}}\left(\mu, \sigma_{Q}\right) \\
& \lesssim 2^{m(n+1)-j} \alpha(Q),
\end{aligned}
$$

and then,

$$
\left|G 1_{j}^{m}(x)\right| \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \int_{\text {supp } \eta_{Q}} 2^{m(n+1)-j} \alpha(Q) d \mu(y) \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)} \alpha(Q) .
$$

4.2.2. Estimate of $G 2_{j}^{m}(x)$. It can be estimated using the arguments of subsection 4.2.1, but now we also have to take into account that $\left|\nabla_{y} \gamma_{2^{-m}}(x-y)\right| \lesssim 2^{m} \lesssim 2^{j}$, because we are assuming $j \geq m-1$, and we have to use the last estimate in Claim 4.4.
4.2.3. Estimate of $G 3_{j}^{m}(x)$. Given $x \in D \cap \Gamma$ and $Q \in \mathcal{D}_{j}$, denote

$$
H_{Q}(y, z):=\varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y) \eta_{Q}(y) K(z-y) \Lambda_{m}^{\mu}(x, z ; y) .
$$

Then, (4.32) becomes

$$
\begin{align*}
G 3_{j}^{m}(x)= & \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \iint H_{Q}(y, z) d\left(\sigma_{Q} \times \sigma_{Q}-\sigma_{D} \times \sigma_{D}\right)(z, y) \\
= & \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \iint H_{Q}(y, z) d\left(c_{Q}^{2} \mathcal{H}_{L_{Q}}^{n} \times \mathcal{H}_{L_{Q}}^{n}-c_{D}^{2} \mathcal{H}_{L_{D}}^{n} \times \mathcal{H}_{L_{D}}^{n}\right)(z, y) \\
= & \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \iint H_{Q}(y, z)\left(c_{Q}^{2}-c_{D}^{2}\right) d \mathcal{H}_{L_{Q}}^{n}(z) d \mathcal{H}_{L_{Q}}^{n}(y)  \tag{4.38}\\
& +\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} c_{D}^{2} \iint H_{Q}(y, z) d\left(\mathcal{H}_{L_{Q}}^{n} \times \mathcal{H}_{L_{Q}}^{n}-\mathcal{H}_{L_{D}}^{n} \times \mathcal{H}_{L_{D}}^{n}\right)(z, y) \\
= & G 3 A_{j}^{m}(x)+G 3 B_{j}^{m}(x) .
\end{align*}
$$

We are going to estimate the terms $G 3 A_{j}^{m}(x)$ and $G 3 B_{j}^{m}(x)$ separately. Recall that $\ell(D)=2^{-m}$. Given a v-cube $Q \in \mathcal{D}_{j}$ such that $Q \subset C_{e} D$, let $Q=: Q_{j} \subset \ldots \subset Q_{i+1} \subset$ $Q_{i} \subset \ldots \subset Q_{m-1}$ be the sequence of v-cubes such that $Q_{i}$ belongs to $\mathcal{D}_{i}$ for $i=m-1, \ldots, j$. Evidently, $Q_{m-1} \subset C_{b} D$ for some constant $C_{b}$ big enough, because $\ell\left(Q_{m-1}\right)=2 \ell(D)$ and $Q \subset Q_{m-1} \cap C_{e} D$. Let $L_{Q_{i}}$ be an $n$-plane that minimizes $\alpha\left(Q_{i}\right)$ and let $\sigma_{Q_{i}}:=c_{Q_{i}} \mathcal{H}_{L_{Q_{i}}}^{n}$ be a minimizing measure of $\alpha\left(Q_{i}\right)$. Also, let $L_{C_{b} D}$ and $\sigma_{C_{b} D}:=c_{C_{b} D} \mathcal{H}_{L_{C_{b} D}}^{n}$ be a minimizing $n$-plane and measure of $\alpha\left(C_{b} D\right)$, respectively.

In order to estimate $G 3 A_{j}^{m}(x)$, notice that, by [To, Lemma 3.4] and the triangle inequality, $\left|c_{Q_{i}}\right| \lesssim 1$ for all $i=m-1, \ldots, j$, and

$$
\begin{align*}
\left|c_{Q}^{2}-c_{D}^{2}\right| & =\left|c_{Q}+c_{D}\right|\left|c_{Q}-c_{D}\right| \lesssim\left|c_{Q_{j}}-c_{D}\right| \\
& \lesssim\left|c_{Q_{m-1}}-c_{C_{b} D}\right|+\left|c_{C_{b} D}-c_{D}\right|+\sum_{i=m-1}^{j-1}\left|c_{Q_{i+1}}-c_{Q_{i}}\right|  \tag{4.39}\\
& \lesssim \alpha\left(C_{b} D\right)+\sum_{i=m-1}^{j-1} \alpha\left(Q_{i}\right) \lesssim \alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)
\end{align*}
$$

(in the case that $j=m-1$, there are no intermediate scales between $j$ and $m-1$, so one just obtains $\left|c_{Q}^{2}-c_{D}^{2}\right| \lesssim \alpha\left(C_{b} D\right)$ ).

Claim 4.5. For $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$, we have $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)-j}$.
Notice that this last estimate is the same as the first one in Claim 4.4. Using Claim 4.5 and that $|K(z-y)| \lesssim 2^{j n}$ for all $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2-j}(\cdot-y)$, we easily obtain $\left|H_{Q}(y, z)\right| \lesssim$ $2^{m(n+1)+j(n-1)}$. Therefore, using (4.39),

$$
\begin{align*}
\left|G 3 A_{j}^{m}(x)\right| & \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D}\left|c_{Q}^{2}-c_{D}^{2}\right| \iint\left|H_{Q}(y, z)\right| d \mathcal{H}_{L_{Q}}^{n}(z) d \mathcal{H}_{L_{Q}}^{n}(y)  \tag{4.40}\\
& \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right) .
\end{align*}
$$

To estimate $G 3 B_{j}^{m}(x)$ in (4.38), we set

$$
\begin{equation*}
G 3 B_{j}^{m}(x)=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} c_{D}^{2} G 3 B(Q)_{j}^{m}(x), \tag{4.41}
\end{equation*}
$$

where $G 3 B(Q)_{j}^{m}(x):=\iint H_{Q} d\left(\mathcal{H}_{L_{Q}}^{n} \times \mathcal{H}_{L_{Q}}^{n}-\mathcal{H}_{L_{D}}^{n} \times \mathcal{H}_{L_{D}}^{n}\right)$. Given $Q \in \mathcal{D}_{j}$ such that $Q \subset C_{e} D$, we split $G 3 B(Q)_{j}^{m}(x)$ as follows:

$$
\begin{align*}
G 3 B(Q)_{j}^{m}(x)= & \sum_{i=m-1}^{j-1} \iint H_{Q} d\left(\mathcal{H}_{L_{Q_{i+1}}}^{n} \times \mathcal{H}_{L_{Q_{i+1}}}^{n}-\mathcal{H}_{L_{Q_{i}}}^{n} \times \mathcal{H}_{L_{Q_{i}}}^{n}\right) \\
& +\iint H_{Q} d\left(\mathcal{H}_{L_{Q_{m-1}}}^{n} \times \mathcal{H}_{L_{Q_{m-1}}}^{n}-\mathcal{H}_{L_{C_{b} D}}^{n} \times \mathcal{H}_{L_{C_{b} D}}^{n}\right)  \tag{4.42}\\
& +\iint H_{Q} d\left(\mathcal{H}_{L_{C_{b} D}}^{n} \times \mathcal{H}_{L_{C_{b} D}}^{n}-\mathcal{H}_{L_{D}}^{n} \times \mathcal{H}_{L_{D}}^{n}\right)
\end{align*}
$$

(as before, in the case $j=m-1$ the first term on the right hand side of (4.42) does not exist).

Fix $i \in \mathbb{Z}$ such that $m-1 \leq i<j$. Set $L_{Q_{i+1}}=\left\{(\widetilde{y}, a(\widetilde{y})) \in \mathbb{R}^{d}: \widetilde{y} \in \mathbb{R}^{n}\right\}$, where $a: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ is an appropriate affine map, and let $p: L_{Q_{i}} \rightarrow L_{Q_{i+1}}$ be the map defined by $p(y):=(\widetilde{y}, a(\widetilde{y}))$. Let $p_{\sharp} \mathcal{H}_{L_{Q_{i}}}^{n}$ be the image measure of $\mathcal{H}_{L_{Q_{i}}}^{n}$ by $p$. It is easy to check that $\mathcal{H}_{L_{Q_{i+1}}}^{n}=\tau p_{\sharp} \mathcal{H}_{L_{Q_{i}}}^{n}$, where $\tau$ is some positive constant such that $|\tau-1| \lesssim \alpha\left(Q_{i}\right)$ and $\tau \lesssim 1$. Therefore,

$$
\begin{align*}
\iint H_{Q}(y, z) & d\left(\mathcal{H}_{L_{Q_{i+1}}}^{n} \times \mathcal{H}_{L_{Q_{i+1}}}^{n}-\mathcal{H}_{L_{Q_{i}}}^{n} \times \mathcal{H}_{L_{Q_{i}}}^{n}\right)(z, y) \\
= & \iint\left(\tau^{2} H_{Q}(p(y), p(z))-H_{Q}(y, z)\right) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y) \\
= & \iint \tau^{2}\left(H_{Q}(p(y), p(z))-H_{Q}(y, z)\right) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y)  \tag{4.43}\\
& +\iint\left(\tau^{2}-1\right) H_{Q}(y, z) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y) .
\end{align*}
$$

Since $\left|\tau^{2}-1\right| \lesssim \alpha\left(Q_{j}\right)$ and we have seen that $\left|H_{Q}(y, z)\right| \lesssim 2^{m(n+1)+j(n-1)}$ after Claim 4.5, the second term on the right side of the last equality is bounded by $C 2^{(m-j)(n+1)} \alpha\left(Q_{i}\right)$.

In order to estimate the first term on the right hand side of (4.43), notice that $\varphi_{2^{-j-1}}^{2-j}(z-$ $y), \gamma_{2^{-m}}(x-y), \eta_{Q}(y)$ and $\Lambda_{m}^{\mu}(x, z ; y)$ only depend on the first $n$ coordinates of $y$ and $z$ (i.e., on $\widetilde{y}$ and $\widetilde{z}$, thus their values coincide on $(y, z)$ and $(p(y), p(z))$. Then,

$$
\begin{aligned}
& \iint \tau^{2}\left(H_{Q}(p(y), p(z))-H_{Q}(y, z)\right) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y) \\
& =\tau^{2} \iint \varphi_{2^{-j-1}}^{2-j}(z-y) \gamma_{2-m}(x-y) \eta_{Q}(y) \Lambda_{m}^{\mu}(x, z ; y) \\
& \quad(K(p(z)-p(y))-K(z-y)) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y) .
\end{aligned}
$$

Let $\theta_{i}$ be the angle between $L_{Q_{i}}$ and $L_{Q_{i+1}}$. One can easily see that, for $y, z \in L_{Q_{i}}, \mid(p(z)-$ $p(y))-(z-y)\left|\lesssim \sin \left(\theta_{i}\right)\right| z-y\left|\lesssim \alpha\left(Q_{i}\right)\right| z-y \mid$. Thus, if also $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$,

$$
|K(p(z)-p(y))-K(z-y)| \lesssim 2^{j(n+1)}|(p(z)-p(y))-(z-y)| \lesssim 2^{j n} \alpha\left(Q_{i}\right) .
$$

Together with Claim 4.5 and the fact that $\tau^{2} \lesssim 1$, this gives

$$
\begin{aligned}
\mid \iint \tau^{2}\left(H_{Q}\right. & \left.(p(y), p(z))-H_{Q}(y, z)\right) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y) \mid \\
& \lesssim \iint \varphi_{2^{-j-1}}^{2-j}(z-y) \eta_{Q}(y) 2^{m(n+1)+j(n-1)} \alpha\left(Q_{i}\right) d \mathcal{H}_{L_{Q_{i}}}^{n}(z) d \mathcal{H}_{L_{Q_{i}}}^{n}(y) \\
& \lesssim 2^{(m-j)(n+1)} \alpha\left(Q_{i}\right)
\end{aligned}
$$

From the last estimates and (4.43), we get

$$
\left|\iint H_{Q} d\left(\mathcal{H}_{L_{Q_{i+1}}}^{n} \times \mathcal{H}_{L_{Q_{i+1}}}^{n}-\mathcal{H}_{L_{Q_{i}}}^{n} \times \mathcal{H}_{L_{Q_{i}}}^{n}\right)\right| \lesssim 2^{(m-j)(n+1)} \alpha\left(Q_{i}\right)
$$

for $i=m-1, \ldots, j-1$. With similar arguments, one also obtains

$$
\begin{aligned}
\left|\iint H_{Q} d\left(\mathcal{H}_{L_{Q_{m-1}}}^{n} \times \mathcal{H}_{L_{Q_{m-1}}}^{n}-\mathcal{H}_{L_{C_{b} D}}^{n} \times \mathcal{H}_{L_{C_{b} D}}^{n}\right)\right| & \lesssim 2^{(m-j)(n+1)} \alpha\left(C_{b} D\right) \\
\left|\iint H_{Q} d\left(\mathcal{H}_{L_{C_{b} D}}^{n} \times \mathcal{H}_{L_{C_{b} D}}^{n}-\mathcal{H}_{L_{D}}^{n} \times \mathcal{H}_{L_{D}}^{n}\right)\right| & \lesssim 2^{(m-j)(n+1)} \alpha\left(C_{b} D\right)
\end{aligned}
$$

These last three inequalities together with (4.42), (4.41) and the fact that $\left|c_{D}\right| \lesssim 1$ yield

$$
\begin{align*}
\left|G 3 B_{j}^{m}(x)\right| & \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{i=m-1}^{j-1} \alpha\left(Q_{i}\right)\right)  \tag{4.44}\\
& \leq \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right)
\end{align*}
$$

Finally, (4.40) and (4.44) applied to (4.38) give half of (4.36).
4.2.4. Estimate of $G 4_{j}^{m}(x)$. By Fubini's theorem and the definitions of $\lambda, \Lambda_{m}^{\mu}$ and $\Lambda_{m}^{\lambda}$,

$$
\begin{aligned}
\Lambda_{m}^{\mu} & (x, z ; y)-\Lambda_{m}^{\lambda}(x, z ; y)=2^{m n} \int_{\left\{a: x, z \in D_{m}^{a}, y \notin D_{m}^{a}\right\}} \frac{\lambda\left(D_{m}^{a}\right)-\mu\left(D_{m}^{a}\right)}{\mu\left(D_{m}^{a}\right) \lambda\left(D_{m}^{a}\right)} d a \\
& =2^{m n} \int_{\left\{a: x, z \in D_{m}^{a}, y \notin D_{m}^{a}\right\}}\left(\int_{t \in D_{m}^{a}} \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \eta_{Q}(t) d\left(\sigma_{Q}-\mu\right)(t)\right) \frac{d a}{\mu\left(D_{m}^{a}\right) \lambda\left(D_{m}^{a}\right)} \\
& =\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \int \eta_{Q}(t) \Lambda_{m}^{\mu, \lambda}(x, z, t ; y) d\left(\sigma_{Q}-\mu\right)(t) .
\end{aligned}
$$

We also used in the second equality that $1=\sum_{Q \in \mathcal{D}_{j}} \eta_{Q}(t)=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \eta_{Q}(t)$ for all $t \in D_{m}^{a}$ if $C_{e}$ is big enough, and this is because $j \geq m-1$ and $|\widetilde{x}-\widetilde{t}| \lesssim 2^{-m}$ for all $t \in D_{m}^{a}$.

Claim 4.6. For $x \in D, j \geq m-1,|x-y| \lesssim 2^{-m}$, and $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$, the following hold: $\left|\Lambda_{m}^{\mu, \lambda}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)-j}$ and $\left|\nabla_{t} \Lambda_{m}^{\mu, \lambda}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)}$.

Using Claim 4.6 and the properties of $\eta_{Q}$, we obtain

$$
\begin{aligned}
\left|\Lambda_{m}^{\mu}(x, z ; y)-\Lambda_{m}^{\lambda}(x, z ; y)\right| & \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{m(2 n+1)} \operatorname{dist}_{B_{Q}}\left(\mu, \sigma_{Q}\right) \\
& \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{m(2 n+1)-j(n+1)} \alpha(Q)
\end{aligned}
$$

Plugging this estimate into the definition of $G 4_{j}^{m}(x)$ in (4.33), we get

$$
\begin{aligned}
&\left|G 4_{j}^{m}(x)\right| \lesssim \iint \varphi_{2^{-j-1}}^{2^{-j}}(z-y) \gamma_{2^{-m}}(x-y)|K(z-y)| \\
& \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{m(2 n+1)-j(n+1)} \alpha(Q) d \sigma_{D}(z) d \sigma_{D}(y) \\
& \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)} \alpha(Q),
\end{aligned}
$$

which, together with the estimates of $\left|G 1_{j}^{m}(x)\right|$ and $\left|G 2_{j}^{m}(x)\right|$ in subsections 4.2.1 and 4.2.2, gives (4.35).
4.2.5. Estimate of $G 5_{j}^{m}(x)$. Arguing as in subsection 4.2.4, we have

$$
\begin{align*}
\Lambda_{m}^{\lambda}(x, z ; y)-\Lambda_{m}^{\sigma_{D}}(x, z ; y) & =\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \int \eta_{Q}(t) \Lambda_{m}^{\lambda, \sigma_{D}}(x, z, t ; y) d\left(\sigma_{D}-\sigma_{Q}\right)(t)  \tag{4.45}\\
& =\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \int H_{Q}(t) d\left(\sigma_{D}-\sigma_{Q}\right)(t)
\end{align*}
$$

where we have set $H_{Q}(t):=\eta_{Q}(t) \Lambda_{m}^{\lambda, \sigma_{D}}(x, z, t ; y)$.
We are going to estimate the right hand side of (4.45) using the techniques of subsection 4.2.3. We have

$$
\begin{equation*}
\int H_{Q} d\left(\sigma_{D}-\sigma_{Q}\right)=\left(c_{D}-c_{Q}\right) \int H_{Q} d \mathcal{H}_{L_{D}}^{n}+c_{Q} \int H_{Q} d\left(\mathcal{H}_{L_{D}}^{n}-\mathcal{H}_{L_{Q}}^{n}\right) \tag{4.46}
\end{equation*}
$$

We introduce the intermediate v-cubes between $Q \in \mathcal{D}_{j}$ and $D \in \mathcal{D}_{m}$ to obtain

$$
\begin{equation*}
\left|c_{D}-c_{Q}\right| \lesssim \alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R) . \tag{4.47}
\end{equation*}
$$

Claim 4.7. For $x \in D, j \geq m-1,|x-y| \lesssim 2^{-m}$, and $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$, the following holds: $\left|\Lambda_{m}^{\lambda, \sigma_{D}}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)-j}$.

Combining Claim 4.7 with (4.47), we derive that

$$
\begin{equation*}
\left|c_{D}-c_{Q}\right| \int\left|H_{Q}\right| d \mathcal{H}_{L_{D}}^{n} \lesssim 2^{m(2 n+1)-j(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right) \tag{4.48}
\end{equation*}
$$

For the second term on the right side of (4.46), one can also use the arguments in subsection 4.2.3 (see (4.42) and following) to show that

$$
\begin{equation*}
\left|\int H_{Q} d\left(\mathcal{H}_{L_{D}}^{n}-\mathcal{H}_{L_{Q}}^{n}\right)\right| \lesssim 2^{m(2 n+1)-j(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right) \tag{4.49}
\end{equation*}
$$

(now it is easier because the function $H_{Q}(t)$ only depends on the first $n$ coordinates of the points involved, i.e., it depends only on $\widetilde{x}, \widetilde{y}, \widetilde{z}$ and $\widetilde{t}$, so when we project vertically to deal with the image measure, the function $H_{Q}$ is not affected). Therefore, by (4.48), (4.49), (4.46), and (4.45), we obtain

$$
\left|\Lambda_{m}^{\lambda}(x, z ; y)-\Lambda_{m}^{\sigma_{D}}(x, z ; y)\right| \lesssim \sum_{\substack{Q \in \mathcal{D}_{j}: \\ Q \subset C_{e} D}} 2^{m(2 n+1)-j(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{\substack{R \in \mathcal{D}: \\ Q \subset R \subset C_{b} D}} \alpha(R)\right)
$$

From the definition of $G 5_{j}^{m}(x)$ in (4.34), we conclude that

$$
\begin{aligned}
\left|G 5_{j}^{m}(x)\right| & \lesssim \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{m(2 n+1)-j(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right) \\
& \iint \varphi_{2^{-j-1}(z-y) \gamma_{2-m}^{2-j}(x-y)|K(z-y)| d \sigma_{D}(z) d \sigma_{D}(y)} \alpha \sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} 2^{(m-j)(n+1)}\left(\alpha\left(C_{b} D\right)+\sum_{R \in \mathcal{D}: Q \subset R \subset C_{b} D} \alpha(R)\right)
\end{aligned}
$$

which, together with the estimate of $\left|G 3_{j}^{m}(x)\right|$ in subsection 4.2.3, gives (4.36).
4.3. Proof of Claims $4.1, \ldots, 4.7$. We have to prove:

- Claim 4.1: We have $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m n}$ and $\left|\nabla_{z} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$ for all $x, y, z \in \mathbb{R}^{d}$.
- Claim 4.2: For $j<m, y \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(x-\cdot)$, and $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$, the following hold: $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m n}$ and $\nabla_{y} \Lambda_{m}^{\mu}(x, z ; y)=0$.
- Claim 4.3: We have $\left|\Lambda_{m}^{\mu, \sigma_{D}}(x, z, t ; y)\right| \lesssim 2^{2 m n}$ and $\left|\nabla_{t} \Lambda_{m}^{\mu, \sigma_{D}}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)}$ for all $x, y, z, t \in \mathbb{R}^{d}$.
- Claim 4.4: For $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$, the following hold: $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)-j}$, $\left|\nabla_{z} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$, and $\left|\nabla_{y} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$.
- Claim 4.5: For $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y),\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)-j}$.
- Claim 4.6: For $x \in D, j \geq m-1,|x-y| \lesssim 2^{-m}$, and $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$, the following hold: $\left|\Lambda_{m}^{\mu, \lambda}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)-j}$ and $\left|\nabla_{t} \Lambda_{m}^{\mu, \lambda}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)}$.
- Claim 4.7: For $x \in D, j \geq m-1,|x-y| \lesssim 2^{-m}$, and $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$, the following holds: $\left|\Lambda_{m}^{\lambda, \sigma_{D}}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)-j}$.
To prove the claims, we need to express the function $\Lambda$ at the end of subsection 2.3 in a more convenient way. Notice that we can replace $D_{m}^{a}$ by $\overline{D_{m}^{a}}$ in the definition of $\Lambda$ because $\mu$ and the $n$-dimensional Hausdorff measure vanish on $\partial D_{m}^{a}$.

For $u \in \mathbb{R}^{n}$ and $r>0$, we denote $|u|_{\infty}:=\max _{i=1, \ldots, n}\left|u^{i}\right|, B_{\infty}(u, r):=\left\{v \in \mathbb{R}^{n}\right.$ : $\left.|u-v|_{\infty} \leq r\right\}$, and $B_{\infty}^{m}(u):=B_{\infty}\left(u, 2^{-m-1}\right)$. Given $a \in \mathbb{R}^{n}$, let $b:=a+\left\{2^{-m-1}\right\}^{n} \in \mathbb{R}^{n}$ be the center of $\widetilde{D}_{m}^{a}$. Then, given $q \in \mathbb{R}^{d}$,

$$
q \in \overline{D_{m}^{a}} \Longleftrightarrow|\tilde{q}-b|_{\infty} \leq 2^{-m} \Longleftrightarrow b \in B_{\infty}^{m}(\widetilde{q})
$$

Let $\mu_{1}, \ldots, \mu_{k}$ be positive Borel measures such that $\mu_{l}\left(D_{m}^{a}\right)>0$ and $\mu_{l}\left(\partial D_{m}^{a}\right)=0$ for all $a \in \mathbb{R}^{n}, m \in \mathbb{Z}$ and $l=1, \ldots, k$. Given $m \in \mathbb{Z}$ and $x_{1}, \ldots, x_{i}, y_{1}, \ldots, y_{j} \in \mathbb{R}^{d}$ we have, by the change of variable $b=a+\left\{2^{-m-1}\right\}^{n} \in \mathbb{R}^{n}$,

$$
\begin{align*}
\Lambda_{m}^{\mu_{1}, \ldots, \mu_{k}}\left(x_{1}, \ldots, x_{i} ; y_{1}, \ldots, y_{j}\right) & =\int_{\left\{a \in \mathbb{R}^{n}: x_{1}, \ldots, x_{i} \in \overline{D_{m}^{a}}, y_{1}, \ldots, y_{j} \notin \overline{D_{m}^{a}}\right\}} \frac{2^{n m} d a}{\prod_{l=1}^{k} \mu_{l}\left(D_{m}^{a}\right)} \\
& =2^{n m} \int \frac{\chi_{B_{\infty}^{m}\left(\widetilde{x_{1}}\right) \cap \ldots \cap B_{\infty}^{m}\left(\widetilde{x_{i}}\right) \cap B_{\infty}^{m}\left(\widetilde{y_{1}}\right)^{c} \cap \ldots \cap B_{\infty}^{m}\left(\widetilde{y_{j}}\right)^{c}(b)}^{\prod_{l=1}^{k} \mu_{l}\left(D_{m}^{b-\left\{2^{-m-1}\right\}^{n}}\right)} d b .}{} \tag{4.50}
\end{align*}
$$

Proof of Claim 4.1. By (4.50), we have

$$
\begin{equation*}
\Lambda_{m}^{\mu}(x, z ; y)=2^{n m} \int \mu\left(D_{m}^{b-\left\{2^{-m-1}\right\}^{n}}\right)^{-1} \chi_{B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}}(b) d b \tag{4.51}
\end{equation*}
$$

Since $\mu\left(D_{m}^{b}\right) \gtrsim 2^{-m n}$ for all $b \in \mathbb{R}^{n}$,

$$
\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{2 m n} \mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \leq 2^{2 m n} \mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{x})\right) \leq 2^{m n}
$$

To deal with the second inequality in Claim 4.1, we will estimate

$$
\left|\Lambda_{m}^{\mu}\left(x, z_{1} ; y\right)-\Lambda_{m}^{\mu}\left(x, z_{2} ; y\right)\right| /\left|z_{1}-z_{2}\right|
$$

for $z_{1}$ close enough to $z_{2}$. Recall that, given two sets $F_{1}, F_{2} \subset \mathbb{R}^{n}, F_{1} \Delta F_{2}:=\left(F_{1} \backslash F_{2}\right) \cup\left(F_{2} \backslash\right.$ $F_{1}$ ) denotes their symmetric difference. Using (4.51), we get

$$
\begin{align*}
& \mid \Lambda_{m}^{\mu}(x,\left.z_{1} ; y\right)-\Lambda_{m}^{\mu}\left(x, z_{2} ; y\right) \mid \\
& \lesssim 2^{2 n m} \int \mid \chi_{B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}\left(\widetilde{z_{1}}\right) \cap B_{\infty}^{m}(\widetilde{y})^{c}(b)-\chi_{B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}\left(\widetilde{z_{2}}\right) \cap B_{\infty}^{m}\left(\widetilde{y^{c}}\right.}(b) \mid d b}  \tag{4.52}\\
& \quad=2^{2 n m} \mathcal{L}^{n}\left(\left(B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}\left(\widetilde{z_{1}}\right) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \Delta\left(B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}\left(\widetilde{z_{2}}\right) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right)\right) \\
& \quad \leq 2^{2 n m} \mathcal{L}^{n}\left(B_{\infty}^{m}\left(\widetilde{z_{1}}\right) \Delta B_{\infty}^{m}\left(\widetilde{z_{2}}\right)\right) \lesssim 2^{2 n m}\left|\widetilde{z_{1}}-\widetilde{z_{2}}\right| 2^{-m(n-1)} \leq 2^{m(n+1)}\left|z_{1}-z_{2}\right|,
\end{align*}
$$

and the claim follows.
Proof of Claim 4.2. The first estimate has been already proved in Claim 4.1. Let us deal with the second one. Notice that if $y \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(x-\cdot)$ then $|\widetilde{x}-\widetilde{y}| \geq 2^{-j-1} 2.1 \sqrt{n}$. Thus, if also $j<m$ and $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$, then $|\widetilde{x}-\widetilde{y}|>2^{-m} \sqrt{n}$ and $|\widetilde{z}-\widetilde{y}|>2^{-m} \sqrt{n}$. Therefore, $B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}=B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\widetilde{z})$ for all $y \in \operatorname{supp} \varphi_{2^{-j-1}}^{2-j}(x-\cdot)$, if $|\widetilde{x}-\widetilde{z}| \leq 2^{-m} \sqrt{n}$. This means, using (4.51), that $\Lambda_{m}^{\mu}(x, z ; y)$ does not depend on $y$, so $\nabla_{y} \Lambda_{m}^{\mu}(x, z ; y)=0$ for all $y \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(x-\cdot)$, and the claim is proved.

Proof of Claim 4.3. This claim follows by arguments very similar to the ones in the proof of Claim 4.1. Just notice that $\mu\left(D_{m}^{b}\right) \sigma_{D}\left(D_{m}^{b}\right) \gtrsim 2^{-2 m n}$ for all $b \in \mathbb{R}^{n}$.

Proof of Claim 4.4. Using (4.51), we have that

$$
\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{2 m n} \mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \leq 2^{2 m n} \mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right)
$$

Notice that $\mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \lesssim 2^{-m(n-1)}|\widetilde{y}-\widetilde{z}|$. Since $z \in \operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y),|\widetilde{y}-\widetilde{z}| \leq$ $2^{-j} 3 \sqrt{n}$. Thus, $\mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \lesssim 2^{-m(n-1)-j}$, and then $\left|\Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)-j}$.

In Claim 4.1 we already proved that $\left|\nabla_{z} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$. Finally, to prove that $\left|\nabla_{y} \Lambda_{m}^{\mu}(x, z ; y)\right| \lesssim 2^{m(n+1)}$, one can repeat the computations done in (4.52) but applied to the $y$ coordinate and use that $B_{\infty}^{m}\left(\widetilde{y_{1}}\right)^{c} \Delta B_{\infty}^{m}\left(\widetilde{y_{2}}\right)^{c}=B_{\infty}^{m}\left(\widetilde{y_{1}}\right) \Delta B_{\infty}^{m}\left(\widetilde{y_{2}}\right)$.

Proof of Claim 4.5. This claim is included in the previous one.
Proof of Claim 4.6. Recall that $\lambda=\sum_{Q \in \mathcal{D}_{j}: Q \subset C_{e} D} \eta_{Q} \sigma_{Q}$, where $C_{e}$ is some fixed constant big enough (see the beginning of subsection 4.2). Using the properties of $\eta_{Q}$ and that $C_{e}$ is big enough, it is not difficult to show that $\lambda\left(D_{m}^{b-\left\{2^{-m-1}\right\}^{n}}\right) \gtrsim 2^{-m n}$ for all $b \in \mathbb{R}^{n}$ such that $b \in B_{\infty}^{m}(\widetilde{x})$ (recall that $x \in D$ and $j \geq m-1$ ). Therefore, by (4.50),

$$
\begin{aligned}
\left|\Lambda_{m}^{\mu, \lambda}(x, z, t ; y)\right| & \lesssim 2^{3 n m} \mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{t}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \\
& \leq 2^{3 n m} \mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) .
\end{aligned}
$$

As in the proof of Claim 4.4, we have $\mathcal{L}^{n}\left(B_{\infty}^{m}(\widetilde{z}) \cap B_{\infty}^{m}(\widetilde{y})^{c}\right) \lesssim 2^{-m(n-1)-j}$ for all $z \in$ $\operatorname{supp} \varphi_{2^{-j-1}}^{2^{-j}}(\cdot-y)$. Thus, $\left|\Lambda_{m}^{\mu, \lambda}(x, z, t ; y)\right| \lesssim 2^{m(2 n+1)-j}$, as wished.

For the second estimate in Claim 4.6, we argue as in (4.52). For $t_{1}$ and $t_{2}$ close enough,

$$
\begin{aligned}
& \left|\Lambda_{m}^{\mu, \lambda}\left(x, z, t_{1} ; y\right)-\Lambda_{m}^{\mu, \lambda}\left(x, z, t_{2} ; y\right)\right| \\
& \lesssim 2^{3 n m} \int \mid \chi_{B_{\infty}^{m}(\widetilde{x}) \cap B_{\infty}^{m}(\tilde{z}) \cap B_{\infty}^{m}\left(\tilde{t_{1}}\right) \cap B_{\infty}^{m}(\widetilde{y})^{c}}(b)-\chi_{B_{\infty}^{m}(\tilde{x}) \cap B_{\infty}^{m}(\tilde{z}) \cap B_{\infty}^{m}\left(\tilde{t_{2}}\right) \cap B_{\infty}^{m}(\widetilde{y})^{c}(b) \mid d b} \\
& \leq 2^{3 n m} \mathcal{L}^{n}\left(B_{\infty}^{m}\left(\widetilde{t_{1}}\right) \Delta B_{\infty}^{m}\left(\widetilde{t_{2}}\right)\right) \lesssim 2^{3 n m}\left|\widetilde{t_{1}}-\widetilde{t_{2}}\right| 2^{-m(n-1)} \leq 2^{m(2 n+1)}\left|t_{1}-t_{2}\right|,
\end{aligned}
$$

and the claim follows by letting $t_{1} \rightarrow t_{2}$.
Proof of Claim 4.7. This claim is proved as the first estimate in Claim 4.6, replacing $\mu$ by $\sigma_{D}$ (we only used that $\mu\left(D_{m}^{b}\right) \gtrsim 2^{-m n}$ for all $b \in \mathbb{R}^{n}$, which also holds for $\sigma_{D}$ ).

## 5. Proof of Theorem 3.2

Given $x \in \Gamma$, let $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ be a decreasing sequence of positive numbers such that

$$
\begin{equation*}
S \mu(x)^{2} \leq 2 \sum_{j \in \mathbb{Z}} \sum_{m \in \mathbb{Z}: \epsilon_{m}, \epsilon_{m+1} \in I_{j}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{2}, \tag{5.1}
\end{equation*}
$$

so $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ depends on $x$.
Fix $j \in \mathbb{Z}$ and assume that $x \in D$, for some $D \in \mathcal{D}_{j}$. Let $L_{D}$ be an $n$-plane that minimizes $\alpha(D)$ and let $\sigma_{D}:=c_{D} \mathcal{H}_{L_{D}}^{n}$ be a minimizing measure for $\alpha(D)$. Let $L_{D}^{x}$ be the $n$-plane parallel to $L_{D}$ which contains $x$, and set $\sigma_{D}^{x}:=c_{D} \mathcal{H}_{L_{D}^{x}}^{n}$.

By the antisymmetry of the function $\varphi_{\epsilon_{m+1}}^{\epsilon_{m}} K$, and since $\sigma_{D}^{x}$ is a Hausdorff measure on the $n$-plane $L_{D}^{x}$ and $x \in L_{D}^{x}$, we have

$$
\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \sigma_{D}^{x}\right)(x)=\int \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y) K(x-y) d \sigma_{D}^{x}(y)=0
$$

for all $m \in \mathbb{Z}$. Therefore, we can decompose

$$
\begin{equation*}
\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)=\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\mu-\sigma_{D}\right)\right)(x)+\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\sigma_{D}-\sigma_{D}^{x}\right)\right)(x) \tag{5.2}
\end{equation*}
$$

For every $m \in \mathbb{Z}$ such that $\epsilon_{m}, \epsilon_{m+1} \in I_{j}$ we will prove the following inequalities:

$$
\begin{gather*}
\left|\left(K \varphi_{\epsilon_{m+1} \epsilon_{m}} *\left(\mu-\sigma_{D}\right)\right)(x)\right| \lesssim 2^{j}\left|\epsilon_{m}-\epsilon_{m+1}\right| \alpha(D)  \tag{5.3}\\
\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\sigma_{D}-\sigma_{D}^{x}\right)\right)(x)\right| \lesssim 2^{2 j}\left|\epsilon_{m}-\epsilon_{m+1}\right| \operatorname{dist}\left(x, L_{D}\right) . \tag{5.4}
\end{gather*}
$$

Assume for a moment that these estimates hold. Then, by (5.2),

$$
\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right| \lesssim 2^{j}\left|\epsilon_{m}-\epsilon_{m+1}\right|\left(\alpha(D)+2^{j} \operatorname{dist}\left(x, L_{D}\right)\right)
$$

Then, using (5.1), we conclude that

$$
\begin{aligned}
\|S \mu\|_{L^{2}(\mu)}^{2} & \leq 2 \sum_{j \in \mathbb{Z}} \sum_{D \in \mathcal{D}_{j}} \int_{D} \sum_{m \in \mathbb{Z}: \epsilon_{m}, \epsilon_{m+1} \in I_{j}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{2} d \mu(x) \\
& \lesssim \sum_{j \in \mathbb{Z}} \sum_{D \in \mathcal{D}_{j}} \int_{D}\left(\alpha(D)+\frac{\operatorname{dist}\left(x, L_{D}\right)}{2^{-j}}\right)^{2} \sum_{\substack{m \in \mathbb{Z}: \\
\epsilon_{m}, \epsilon_{m+1} \in I_{j}}}\left(\frac{\left|\epsilon_{m}-\epsilon_{m+1}\right|}{2^{-j}}\right)^{2} d \mu(x) \\
& \lesssim \sum_{D \in \mathcal{D}} \alpha(D)^{2} \mu(D)+\sum_{D \in \mathcal{D}} \int_{D}\left(\frac{\operatorname{dist}\left(x, L_{D}\right)}{\ell(D)}\right)^{2} d \mu(x) .
\end{aligned}
$$

Notice that, under the integral sign on the right hand side of the first inequality above, $\epsilon_{m}$ and $\epsilon_{m+1}$ depend on $x$. It is not obvious that the resulting function is measurable. To deal with this issue more carefully, we might first ask $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ to lie in some finite set, prove
the variational norm inequality with constants independent of the set, and then enlarge the set. Then, by monotone convergence we would obtain the result with $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ restricted to a countable set dense in $(0, \infty)$. The final theorem would follow then from the continuity properties of the operators involved. This applies to other similar situations in the paper. However, for the sake of conciseness, we will skip further details.

The second term on the right hand side of the last inequality coincides with $W_{1} \mu$ (see (4.6)), thus it is bounded (modulo constants) by $\sum_{D \in \mathcal{D}}\left(\alpha(D)^{2}+\beta_{2}(D)^{2}\right) \mu(D)$, and Theorem 3.2 is proved.

It only remains to verify (5.3) and (5.4) for $x \in D \in \mathcal{D}_{j}$ and $m \in \mathbb{Z}$ such that $\epsilon_{m}, \epsilon_{m+1} \in I_{j}$. First of all, notice that $\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}$ satisfies

$$
\begin{align*}
\left|\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\right| & =\left|\varphi_{\mathbb{R}}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m+1}}\right)-\varphi_{\mathbb{R}}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m}}\right)\right| \leq\left\|\varphi_{\mathbb{R}}^{\prime}\right\|_{L^{\infty}(\mathbb{R})}\left|\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m+1}}-\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m}}\right|  \tag{5.5}\\
& =\left\|\varphi_{\mathbb{R}}^{\prime}\right\|_{\infty}|\widetilde{x}-\widetilde{y}| \frac{\epsilon_{m}-\epsilon_{m+1}}{\epsilon_{m} \epsilon_{m+1}} \lesssim 2^{j}\left|\epsilon_{m}-\epsilon_{m+1}\right|
\end{align*}
$$

for all $y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot)$. For $i=1, \ldots, d$,

$$
\partial_{y^{i}}\left(\varphi_{\epsilon_{m}}(x-y)\right)=\varphi_{\mathbb{R}}^{\prime}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m}}\right) \frac{y^{i}-x^{i}}{\epsilon_{m}|\widetilde{x}-\widetilde{y}|} \chi_{[1, n]}(i)
$$

Hence,

$$
\begin{aligned}
\mid \partial_{y^{i}}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}\right. & (x-y))\left|\leq\left|\varphi_{\mathbb{R}}^{\prime}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m}}\right) \frac{1}{\epsilon_{m}}-\varphi_{\mathbb{R}}^{\prime}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m+1}}\right) \frac{1}{\epsilon_{m+1}}\right|\right. \\
& \leq\left|\varphi_{\mathbb{R}}^{\prime}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m}}\right)\right|\left|\frac{1}{\epsilon_{m}}-\frac{1}{\epsilon_{m+1}}\right|+\left|\varphi_{\mathbb{R}}^{\prime}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m}}\right)-\varphi_{\mathbb{R}}^{\prime}\left(\frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m+1}}\right)\right| \frac{1}{\epsilon_{m+1}} \\
& \leq\left(\left\|\varphi_{\mathbb{R}}^{\prime}\right\|_{\infty}+\left\|\varphi_{\mathbb{R}}^{\prime \prime}\right\|_{\infty} \frac{|\widetilde{x}-\widetilde{y}|}{\epsilon_{m+1}}\right) \frac{\epsilon_{m}-\epsilon_{m+1}}{\epsilon_{m} \epsilon_{m+1}}
\end{aligned}
$$

Since $\epsilon_{m}, \epsilon_{m+1} \in I_{j}$, we deduce from the previous estimate that, for $x-y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}$,

$$
\begin{equation*}
\left|\nabla_{y}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\right)\right| \lesssim \frac{\epsilon_{m}-\epsilon_{m+1}}{\epsilon_{m} \epsilon_{m+1}} \approx 2^{2 j}\left|\epsilon_{m}-\epsilon_{m+1}\right| \tag{5.6}
\end{equation*}
$$

We are going to use (5.5) and (5.6) to prove (5.3) and (5.4). Let us start with (5.3). Since $\epsilon_{m}, \epsilon_{m+1} \in I_{j}$, we can assume that $\operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot) \subset B_{D}$, by taking $C_{\Gamma}$ big enough.

By (5.5) and (5.6), for all $y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot)$,

$$
\left|\nabla_{y}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y) K(x-y)\right)\right| \lesssim 2^{j(n+2)}\left|\epsilon_{m}-\epsilon_{m+1}\right|,
$$

hence

$$
\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\mu-\sigma_{D}\right)\right)(x)\right| \lesssim 2^{j(n+2)}\left|\epsilon_{m}-\epsilon_{m+1}\right| \operatorname{dist}_{B_{D}}\left(\mu, \sigma_{D}\right) \lesssim 2^{j}\left|\epsilon_{m}-\epsilon_{m+1}\right| \alpha(D)
$$

which gives (5.3).
In order to prove $(5.4)$, set $L_{D}^{x}=\left\{(\widetilde{t}, a(\widetilde{t})) \in \mathbb{R}^{d}: \widetilde{t} \in \mathbb{R}^{n}\right\}$, where $a: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ is an appropriate affine map, and let $p: L_{D} \rightarrow L_{D}^{x}$ be the map defined by $p(t):=(\widetilde{t}, a(\widetilde{t}))$. Since $\Gamma$ is a Lipschitz graph, $a$ is well defined and $p$ is a homeomorphism. Let $p_{\sharp} \mathcal{H}_{L_{D}}^{n}$ be the image measure of $\mathcal{H}_{L_{D}}^{n}$ by $p$. It is easy to see that, $|y-p(y)| \approx \operatorname{dist}\left(x, L_{D}\right)$ for all $y \in L_{D}$. Notice also that the image measure $p_{\sharp} \mathcal{H}_{L_{D}}^{n}$ coincides with $\mathcal{H}_{L_{D}^{x}}^{n}$. Therefore, since $\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)$ only
depends on $\widetilde{x}-\widetilde{y}$,

$$
\begin{align*}
\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\sigma_{D}-\sigma_{D}^{x}\right)\right)(x) & =c_{D} \int \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y) K(x-y) d\left(\mathcal{H}_{L_{D}}^{n}-p_{\sharp} \mathcal{H}_{L_{D}}^{n}\right)(y)  \tag{5.7}\\
& =c_{D} \int \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)(K(x-y)-K(x-p(y))) d \mathcal{H}_{L_{D}}^{n}(y) .
\end{align*}
$$

For $y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot) \cap L_{D}$, we have

$$
|K(x-y)-K(x-p(y))| \lesssim 2^{j(n+1)}|y-p(y)| \approx 2^{j(n+1)} \operatorname{dist}\left(x, L_{D}\right)
$$

Plugging this estimate and (5.5) into (5.7), we conclude that

$$
\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\sigma_{D}-\sigma_{D}^{x}\right)\right)(x)\right| \lesssim 2^{2 j}\left|\epsilon_{m}-\epsilon_{m+1}\right| \operatorname{dist}\left(x, L_{D}\right)
$$

which gives (5.4); and the theorem follows.

$$
\text { 6. } L^{2} \text { LOCALIZATION OF } \mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}} \text { AND } \mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}
$$

From here till the end of the paper, $\Gamma:=\left\{x \in \mathbb{R}^{d}: x=(\widetilde{x}, A(\widetilde{x}))\right\}$ will be the graph of a Lipschitz function $A: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$, without any assumption on the support of $A$.

Theorem 6.1. Let $\rho>2$. There exist $C_{1}, C_{2}>0$ such that, for every $f \in L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$ supported in $\Gamma \cap D$ (where $D:=\widetilde{D} \times \mathbb{R}^{d-n}$ and $\widetilde{D}$ is a cube of $\mathbb{R}^{n}$ ),

$$
\begin{align*}
& \int_{D}\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f\right)^{2} d \mathcal{H}_{\Gamma}^{n} \leq C_{1}\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}^{2} \mathcal{H}_{\Gamma}^{n}(D) \quad \text { and }  \tag{6.1}\\
& \int_{D}\left(\left(\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f\right)^{2} d \mathcal{H}_{\Gamma}^{n} \leq C_{2}\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}^{2} \mathcal{H}_{\Gamma}^{n}(D) \tag{6.2}
\end{align*}
$$

The constant $C_{2}$ does not depend on the fixed sequence that defines $\mathcal{O}$.
We will only give the proof of (6.1), because the proof of (6.2) follows by very similar arguments and computations.

We claim that it is enough to prove (6.1) for all functions $f$ such that $f(x) \approx 1$ for all $x \in \Gamma \cap D$. Otherwise, we consider $g(x):=\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}^{-1} f(x)+2 \chi_{D}(x)$, which clearly satisfies $g(x) \approx 1$ for all $x \in \Gamma \cap D$. Since $f=\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}\left(g-2 \chi_{D}\right)$,

$$
\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f(x) \leq\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) g(x)+2\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) \chi_{D}(x)\right)
$$

Applying (6.1) to the functions $g$ and $\chi_{D}$, we finally get

$$
\int_{D}\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f\right)^{2} d \mathcal{H}_{\Gamma}^{n} \lesssim\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}^{2} \mathcal{H}_{\Gamma}^{n}(D)
$$

Given $f$ and $D$ as in Theorem 6.1, from now on, we assume that $f \approx 1$ in $\Gamma \cap D$. Let $\widetilde{z}_{D}$ be the center of $\widetilde{D}$ and set $z_{D}:=\left(\widetilde{z}_{D}, A\left(\widetilde{z}_{D}\right)\right)$. One can easily construct a Lipschitz function $A_{D}: \mathbb{R}^{n} \rightarrow \mathbb{R}^{d-n}$ such that $\operatorname{Lip}\left(A_{D}\right) \lesssim \operatorname{Lip}(A), A_{D}(\widetilde{x})=A\left(\widetilde{z}_{D}\right)$ for all $\widetilde{x} \in(3 \widetilde{D})^{c}$, and $A_{D}(\widetilde{x})=A(\widetilde{x})$ for all $\widetilde{x} \in \widetilde{D}$. Let $\Gamma_{D}$ be the graph of $A_{D}$ and define the measure $\mu:=\mathcal{H}_{\Gamma_{D} \backslash D}^{n}+f \mathcal{H}_{\Gamma_{D} \cap D}^{n}$. Notice that $\chi_{(3 D)^{c}} \mu$ is supported in the $n$-plane $L:=\mathbb{R}^{n} \times\left\{A\left(\widetilde{z}_{D}\right)\right\}$ and $\chi_{D} \mu=f \mathcal{H}_{\Gamma \cap D}^{n}$.

Since $f \approx 1$ in $\Gamma \cap D$ and $\chi_{D} \mu=\left(1-\chi_{(3 D)^{c}}-\chi_{3 D \backslash D}\right) \mu$, we can decompose

$$
\begin{aligned}
& \int_{D}\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f\right)^{2} d \mathcal{H}_{\Gamma}^{n} \approx \int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{D} \mu\right)\right)^{2} d \mu \\
& \lesssim \int_{D}\left(\mathcal{V}_{\rho}(K \varphi * \mu)+\mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{(3 D)^{c}} \mu\right)\right)+\mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{3 D \backslash D} \mu\right)\right)\right)^{2} d \mu
\end{aligned}
$$

In the next subsections, we will see that $\int_{D} \mathcal{V}_{\rho}(K \varphi * \mu)^{2} d \mu, \int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{(3 D)^{c}} \mu\right)\right)^{2} d \mu$, and $\int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{3 D \backslash D} \mu\right)\right)^{2} d \mu$ are bounded by $C \mu(D)$, and (6.1) will follow.
6.1. Proof of $\int_{D} \mathcal{V}_{\rho}(K \varphi * \mu)^{2} d \mu \lesssim \mu(D)$. Fix $x \in \operatorname{supp} \mu$, and let $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ be a decreasing sequence of positive numbers (which depends on $x$ ) such that

$$
\begin{equation*}
\left(\mathcal{V}_{\rho}(K \varphi * \mu)(x)\right)^{\rho} \leq 2 \sum_{m \in \mathbb{Z}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{\rho} . \tag{6.3}
\end{equation*}
$$

For $j \in \mathbb{Z}$ we set $I_{j}:=\left[2^{-j-1}, 2^{-j}\right)$. We decompose $\mathbb{Z}=\mathcal{S} \cup \mathcal{L}$, where

$$
\begin{align*}
\mathcal{S} & :=\bigcup_{j \in \mathbb{Z}} \mathcal{S}_{j}, \quad \mathcal{S}_{j}:=\left\{m \in \mathbb{Z}: \epsilon_{m}, \epsilon_{m+1} \in I_{j}\right\},  \tag{6.4}\\
\mathcal{L} & :=\left\{m \in \mathbb{Z}: \epsilon_{m} \in I_{i}, \epsilon_{m+1} \in I_{j} \text { for } i<j\right\} .
\end{align*}
$$

Then, $\sum_{m \in \mathbb{Z}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{\rho}=\sum_{m \in \mathcal{S}}\left|\left(K \varphi_{\epsilon_{m+1} \epsilon_{m}} * \mu\right)(x)\right|^{\rho}+\sum_{m \in \mathcal{L}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{\rho}$.
Notice that, since the $\ell^{\rho}(\mathbb{Z})$-norm is smaller than the $\ell^{2}(\mathbb{Z})$-norm for $\rho>2$,

$$
\begin{equation*}
\sum_{m \in \mathcal{S}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{\rho} \leq S \mu(x)^{\rho}, \tag{6.5}
\end{equation*}
$$

where $S \mu(x)$ has been defined in Theorem 3.2.
Let us now estimate the sum over the indices $m \in \mathcal{L}$. For $m \in \mathbb{Z}$ we define $j\left(\epsilon_{m}\right)$ as the integer such that $\epsilon_{m} \in I_{j\left(\epsilon_{m}\right)}$. Since $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ is decreasing, given $j \in \mathbb{Z}$, there is at most one index $m \in \mathcal{L}$ such that $\epsilon_{m} \in I_{j}$. Thus, if $k, m \in \mathcal{L}$ and $k<m$, one has $j\left(\epsilon_{k}\right)<j\left(\epsilon_{m}\right)$.

With this notation, we have

$$
\begin{align*}
& \sum_{m \in \mathcal{L}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} * \mu\right)(x)\right|^{\rho}=\sum_{m \in \mathcal{L}}\left|\left(K \varphi_{\epsilon_{m+1}} * \mu\right)(x)-\left(K \varphi_{\epsilon_{m}} * \mu\right)(x)\right|^{\rho} \\
& \quad \lesssim \sum_{m \in \mathcal{L}}\left|\left(K \varphi_{\epsilon_{m+1}} * \mu\right)(x)-\left(K \varphi_{2^{-j\left(\epsilon_{m+1}\right)-1}} * \mu\right)(x)\right|^{\rho} \\
& \quad+\sum_{m \in \mathcal{L}}\left|\left(K \varphi_{2^{-j\left(\epsilon_{m+1}\right)-1}} * \mu\right)(x)-E_{j\left(\epsilon_{m+1}\right)+1} \mu(x)\right|^{\rho} \\
& \quad+\sum_{m \in \mathcal{L}}\left|E_{j\left(\epsilon_{m+1}\right)+1} \mu(x)-E_{j\left(\epsilon_{m}\right)+1} \mu(x)\right|^{\rho}  \tag{6.6}\\
& \quad+\sum_{m \in \mathcal{L}}\left|E_{j\left(\epsilon_{m}\right)+1} \mu(x)-\left(K \varphi_{2^{-j\left(\epsilon_{m}\right)-1}} * \mu\right)(x)\right|^{\rho} \\
& \quad+\sum_{m \in \mathcal{L}}\left|\left(K \varphi_{2-j\left(\epsilon_{m}\right)-1} * \mu\right)(x)-\left(K \varphi_{\epsilon_{m}} * \mu\right)(x)\right|^{\rho} \\
& \quad \lesssim S \mu(x)^{\rho}+W \mu(x)^{\rho}+\mathcal{V}_{\rho}(E \mu)(x)^{\rho},
\end{align*}
$$

where $S \mu(x)$ and $W \mu(x)$ have been defined in Theorems 3.2 and 3.1, respectively, and $\mathcal{V}_{\rho}(E \mu)$ is the $\rho$-variation of the average of martingales $\left\{E_{m} \mu\right\}_{m \in \mathbb{Z}}$ from subsection 2.3. Therefore, by (6.5), (6.6), and (6.3), we deduce that

$$
\mathcal{V}_{\rho}(K \varphi * \mu)(x) \lesssim S \mu(x)+W \mu(x)+\mathcal{V}_{\rho}(E \mu)(x)
$$

for all $x \in \operatorname{supp} \mu$, and so

$$
\begin{equation*}
\int_{D} \mathcal{V}_{\rho}(K \varphi * \mu)^{2} d \mu \lesssim\|S \mu\|_{L^{2}(\mu)}^{2}+\|W \mu\|_{L^{2}(\mu)}^{2}+\left\|\mathcal{V}_{\rho}(E \mu)\right\|_{L^{2}(\mu)}^{2} \tag{6.7}
\end{equation*}
$$

Clearly, Theorem 2.7, Theorem 3.1, and Theorem 3.2 can be applied to the measure $\mu$, because $\operatorname{supp} \mu$ is a translation of the graph of a Lipschitz function with compact support. These theorems in combination with (6.7) yield

$$
\begin{equation*}
\int_{D} \mathcal{V}_{\rho}(K \varphi * \mu)^{2} d \mu \leq C_{1}\left(\mu(3 D)+\sum_{Q \in \mathcal{D}}\left(\alpha_{\mu}\left(C_{2} Q\right)^{2}+\beta_{2, \mu}(Q)^{2}\right) \mu(Q)\right) \tag{6.8}
\end{equation*}
$$

where $C_{1}, C_{2}>0$ only depend on $n, d, K, \operatorname{Lip}(A)$, and $\rho$ (the condition $\rho>2$ is used to ensure the $L^{2}$ boundedness of $\mathcal{V}_{\rho}(E \mu)$ ). Obviously, $\mu(3 D) \approx \mu(D)$ and, since $\chi_{(3 D)^{c}} \mu$ coincides with the $n$-dimensional Hausdorff measure on an $n$-plane, using Remark 2.2 it is easy to check that $\sum_{Q \in \mathcal{D}}\left(\alpha_{\mu}\left(C_{2} Q\right)^{2}+\beta_{2, \mu}(Q)^{2}\right) \mu(Q) \lesssim \mu(3 D)$. Hence, we conclude that $\int_{D} \mathcal{V}_{\rho}(K \varphi * \mu)^{2} d \mu \lesssim \mu(D)$ by (6.8).
6.2. Proof of $\int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{(3 D)^{c}} \mu\right)\right)^{2} d \mu \lesssim \mu(D)$. Fix $x \in \operatorname{supp} \mu \cap D$, and let $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ be a decreasing sequence of positive numbers (which depends on $x$ ) such that

$$
\begin{equation*}
\left(\mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{(3 D)^{c}} \mu\right)\right)(x)\right)^{\rho} \leq 2 \sum_{m \in \mathbb{Z}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\chi_{(3 D)^{c}} \mu\right)\right)(x)\right|^{\rho} . \tag{6.9}
\end{equation*}
$$

Recall that $\widetilde{z}_{D}$ is the center of $\widetilde{D}, z_{D}:=\left(\widetilde{z}_{D}, A\left(\widetilde{z}_{D}\right)\right)$ and $L:=\mathbb{R}^{n} \times\left\{A\left(\widetilde{z}_{D}\right)\right\}$. Since $\chi_{(3 D)^{c}} \mu=\mathcal{H}_{L \backslash 3 D}^{n}$ and $z_{D}$ is the center of $L \cap D,\left(K \varphi_{\epsilon}^{\delta} *\left(\chi_{(3 D)^{c}} \mu\right)\right)\left(z_{D}\right)=0$ for all $0<\epsilon \leq \delta$. Thus, $\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\chi_{(3 D)^{c}} \mu\right)\right)(x)\right|=\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\chi_{(3 D)^{c}} \mu\right)\right)(x)-\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(\chi_{(3 D)}{ }^{c} \mu\right)\right)\left(z_{D}\right)\right| \leq$ $\Theta 1_{m}+\Theta 2_{m}$, where

$$
\begin{align*}
& \Theta 1_{m}:=\int_{(3 D)^{c}} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\left|K(x-y)-K\left(z_{D}-y\right)\right| d \mu(y),  \tag{6.10}\\
& \Theta 2_{m}:=\int_{(3 D)^{c}}\left|\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)-\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}\left(z_{D}-y\right)\right| K\left(z_{D}-y\right) d \mu(y) .
\end{align*}
$$

Since $x \in \operatorname{supp} \mu \cap D$ and $A$ is a Lipschitz function, we have $\left|x-z_{D}\right| \lesssim \ell(D)$, and then $\left|K(x-y)-K\left(z_{D}-y\right)\right| \lesssim\left|x-z_{D}\right|\left|z_{D}-y\right|^{-n-1} \lesssim \ell(D)\left|z_{D}-y\right|^{-n-1}$ for all $y \in(3 D)^{c}$. Therefore, using that $\sum_{m \in \mathbb{Z}} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} \leq 1$ and that $\rho>1$,

$$
\begin{equation*}
\left(\sum_{m \in \mathbb{Z}} \Theta 1_{m}^{\rho}\right)^{1 / \rho} \leq \sum_{m \in \mathbb{Z}} \Theta 1_{m} \lesssim \int_{(3 D)^{c}} \ell(D)\left|z_{D}-y\right|^{-n-1} d \mu(y) \lesssim 1 . \tag{6.11}
\end{equation*}
$$

To deal with $\Theta 2_{m}$, we decompose $\mathbb{Z}=\mathcal{S} \cup \mathcal{L}$ as in (6.4). As before, given $m \in \mathbb{Z}$, let $j\left(\epsilon_{m}\right)$ be the integer such that $\epsilon_{m} \in I_{j\left(\epsilon_{m}\right)}$. Observe that

$$
\operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot) \subset A\left(\widetilde{x}, 2.1 \sqrt{n} 2^{-j\left(\epsilon_{m+1}\right)-1}, 3 \sqrt{n} 2^{-j\left(\epsilon_{m}\right)}\right) \times \mathbb{R}^{d-n}=: A_{m}(x)
$$

Notice also that the sets $A_{m}(x)$ have finite overlap for $m \in \mathcal{L}$, and the same is true for the sets $A_{j}^{\prime}(x):=A\left(\widetilde{x}, 2.1 \sqrt{n} 2^{-j-1}, 3 \sqrt{n} 2^{-j}\right) \times \mathbb{R}^{d-n}$ for $j \in \mathbb{Z}$. The same observations hold if we replace $x$ by $z_{D}$ (and $\widetilde{x}$ by $\widetilde{z}_{D}$ ). Obviously, $A_{m}(x) \subset A_{j}^{\prime}(x)$ (and $A_{m}\left(z_{D}\right) \subset A_{j}^{\prime}\left(z_{D}\right)$ ) for $m \in \mathcal{S}_{j}$.

Assume that $m \in \mathcal{S}$. With the same computations as those carried out in (5.6), one can easily prove that, for all $z-y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}$,

$$
\left|\nabla_{z}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(z-y)\right)\right| \lesssim\left(\left\|\varphi_{\mathbb{R}}^{\prime}\right\|_{L^{\infty}(\mathbb{R})}+\left\|\varphi_{\mathbb{R}}^{\prime \prime}\right\|_{L^{\infty}(\mathbb{R})} \frac{|\widetilde{z}-\widetilde{y}|}{\epsilon_{m+1}}\right) \frac{\epsilon_{m}-\epsilon_{m+1}}{\epsilon_{m} \epsilon_{m+1}} \lesssim 2^{j\left(\epsilon_{m}\right)} \frac{\left|\epsilon_{m}-\epsilon_{m+1}\right|}{|\widetilde{z}-\widetilde{y}|}
$$

because $|\widetilde{z}-\widetilde{y}| \approx \epsilon_{m} \approx \epsilon_{m+1} \approx 2^{j\left(\epsilon_{m}\right)}$ for all $z-y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}$ and $m \in \mathcal{S}$. In particular, if $z \in D$ and $y \in(3 D)^{c},\left|\nabla_{z}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(z-y)\right)\right| \lesssim 2^{j\left(\epsilon_{m}\right)}\left|\epsilon_{m}-\epsilon_{m+1}\right|\left|\widetilde{z}_{D}-\widetilde{y}\right|^{-1}$. Hence,

$$
\Theta 2_{m} \lesssim \int_{\left(A_{m}(x) \cup A_{m}\left(z_{D}\right)\right) \backslash 3 D} \ell(D) 2^{j\left(\epsilon_{m}\right)} \frac{\left|\epsilon_{m}-\epsilon_{m+1}\right|}{\left|\widetilde{z}_{D}-\widetilde{y}\right|^{n+1}} d \mu(y),
$$

and then,

$$
\begin{align*}
\left(\sum_{m \in \mathcal{S}} \Theta 2_{m}^{\rho}\right)^{1 / \rho} & \lesssim \sum_{m \in \mathcal{S}} \int_{\left(A_{m}(x) \cup A_{m}\left(z_{D}\right)\right) \backslash 3 D} \ell(D) 2^{j\left(\epsilon_{m}\right)} \frac{\left|\epsilon_{m}-\epsilon_{m+1}\right|}{\left|\widetilde{z}_{D}-\widetilde{y}\right|^{n+1}} d \mu(y) \\
& \leq \sum_{j \in \mathbb{Z}} \int_{\left(A_{j}^{\prime}(x) \cup A_{j}^{\prime}\left(z_{D}\right)\right) \backslash 3 D} \frac{\ell(D)}{\left|\widetilde{z}_{D}-\widetilde{y}\right|^{n+1}} \sum_{m \in \mathcal{S}_{j}} \frac{\left|\epsilon_{m}-\epsilon_{m+1}\right|}{2^{-j}} d \mu(y)  \tag{6.12}\\
& \lesssim \int_{(3 D)^{c}} \frac{\ell(D)}{\left|\widetilde{z}_{D}-\widetilde{y}\right|^{n+1}} d \mu(y) \lesssim 1 .
\end{align*}
$$

Assume now that $m \in \mathcal{L}$. It is easy to check that $\left|\nabla_{z}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(z-y)\right)\right| \lesssim|\widetilde{z}-\widetilde{y}|^{-1}$ for all $z, y \in \mathbb{R}^{d}$. So, if also $z \in D$ and $y \in(3 D)^{c},\left|\nabla_{z}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(z-y)\right)\right| \lesssim\left|\widetilde{z}_{D}-\widetilde{y}\right|^{-1}$. Therefore,

$$
\begin{align*}
\left(\sum_{m \in \mathcal{L}} \Theta 2_{m}^{\rho}\right)^{1 / \rho} & \lesssim \sum_{m \in \mathcal{L}} \int_{\left(A_{m}(x) \cup A_{m}\left(z_{D}\right)\right) \backslash 3 D} \frac{\ell(D)}{\left|z_{D}-y\right|^{n+1}} d \mu(y)  \tag{6.13}\\
& \lesssim \int_{(3 D)^{c}} \frac{\ell(D)}{\left|\widetilde{z}_{D}-\widetilde{y}\right|^{n+1}} d \mu(y) \lesssim 1 .
\end{align*}
$$

Finally combining (6.11), (6.12), and (6.13), with (6.9) and the fact that $\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\right.$ $\left.\left(\chi_{(3 D)^{c}} \mu\right)\right)(x) \mid \leq \Theta 1_{m}+\Theta 2_{m}$, we conclude that

$$
\mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{(3 D)^{c}} \mu\right)\right)(x) \lesssim\left(\sum_{m \in \mathbb{Z}} \Theta 1_{m}^{\rho}\right)^{1 / \rho}+\left(\sum_{m \in \mathcal{S}} \Theta 2_{m}^{\rho}\right)^{1 / \rho}+\left(\sum_{m \in \mathcal{L}} \Theta 2_{m}^{\rho}\right)^{1 / \rho} \lesssim 1
$$

for all $x \in \operatorname{supp} \mu \cap D$. Therefore, $\int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{(3 D)^{c}} \mu\right)\right)^{2} d \mu \lesssim \mu(D)$.
6.3. Proof of $\int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{3 D \backslash D} \mu\right)\right)^{2} d \mu \lesssim \mu(D)$. Fix $x \in \operatorname{supp} \mu \cap D$. Since $\rho>1$,

$$
\begin{aligned}
& \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{3 D \backslash D} \mu\right)\right)(x)=\sup _{\left\{\epsilon_{m}\right\}}\left(\sum_{m \in \mathbb{Z}}\left|\int_{3 D \backslash D} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y) K(x-y) d \mu(y)\right|^{\rho}\right)^{1 / \rho} \\
& \quad \leq \sup _{\left\{\epsilon_{m}\right\}} \sum_{m \in \mathbb{Z}} \int_{3 D \backslash D} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)|K(x-y)| d \mu(y) \leq \int_{3 D \backslash D}|K(x-y)| d \mu(y) .
\end{aligned}
$$

By a standard computation, one can show that

$$
\int_{D}\left(\int_{3 D \backslash D}|K(x-y)| d \mu(y)\right)^{2} d \mu(x) \lesssim \mu(D),
$$

hence we conclude that $\int_{D} \mathcal{V}_{\rho}\left(K \varphi *\left(\chi_{3 D \backslash D} \mu\right)\right)^{2} d \mu \lesssim \mu(D)$.

## 7. $L^{p}$ AND ENDPOINT ESTIMATES FOR $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ and $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$

We denote by $H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ and $B M O\left(\mathcal{H}_{\Gamma}^{n}\right)$ the (atomic) Hardy space and the space of functions with bounded mean oscillation, respectively, with respect to the measure $\mathcal{H}_{\Gamma}^{n}$. These spaces are defined as the classical $H^{1}\left(\mathbb{R}^{d}\right)$ and $B M O\left(\mathbb{R}^{d}\right)$ (see [Du, Chapter 6], for example), but by replacing the true cubes of $\mathbb{R}^{d}$ by our special v-cubes.

Theorem 7.1. Let $\rho>2$. The operators $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ and $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ are bounded

- in $L^{p}\left(\mathcal{H}_{\Gamma}^{n}\right)$ for $1<p<\infty$,
- from $H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ to $L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$,
- from $L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ to $L^{1, \infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$, and
- from $L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$ to $\operatorname{BMO}\left(\mathcal{H}_{\Gamma}^{n}\right)$,
and the norm of $\mathcal{O} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ in the cases above is bounded independently of the sequence that defines $\mathcal{O}$.

We will only give the proof of Theorem 7.1 in the case of the $\rho$-variation, because the proof for the oscillation follows by analogous arguments.
7.1. The operator $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}: H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ is bounded. Fix a cube $\widetilde{D} \subset \mathbb{R}^{n}$ and set $D:=\widetilde{D} \times \mathbb{R}^{d-n}$. Let $f$ be an atom, i.e., a function defined on $\Gamma$ and such that

$$
\begin{equation*}
\operatorname{supp} f \subset D, \quad\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \leq \frac{1}{\mathcal{H}_{\Gamma}^{n}(D)}, \quad \text { and } \quad \int f d \mathcal{H}_{\Gamma}^{n}=0 \tag{7.1}
\end{equation*}
$$

We have to prove that $\int\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f d \mathcal{H}_{\Gamma}^{n} \leq C$, for some constant $C>0$ which does not depend on $f$ or $D$. Since $\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f(x)$ is well defined and non negative for $f \in L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$, the uniform boundedness of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ on atoms yields its boundedness from $H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ to $L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ by standard arguments. We omit the details.

First of all, by Hölder's inequality, Theorem 6.1, and (7.1),

$$
\begin{aligned}
\int_{3 D}\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f d \mathcal{H}_{\Gamma}^{n} & \leq \mathcal{H}_{\Gamma}^{n}(3 D)^{1 / 2}\left(\int_{3 D}\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f\right)^{2} d \mathcal{H}_{\Gamma}^{n}\right)^{1 / 2} \\
& \lesssim \mathcal{H}_{\Gamma}^{n}(3 D)^{1 / 2}\left(\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}^{2} \mathcal{H}_{\Gamma}^{n}(3 D)\right)^{1 / 2} \lesssim 1
\end{aligned}
$$

Thus, it remains to prove that $\int_{(3 D)^{c}}\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f d \mathcal{H}_{\Gamma}^{n} \leq C$.
Given $x \in \Gamma \backslash 3 D$, let $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ be a decreasing sequence of positive numbers (which depends on $x$ ) such that

$$
\begin{equation*}
\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f(x)\right)^{\rho} \leq 2 \sum_{m \in J}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f \mathcal{H}_{\Gamma}^{n}\right)\right)(x)\right|^{\rho} \tag{7.2}
\end{equation*}
$$

where $J:=\left\{m \in \mathbb{Z}: \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot) \cap \operatorname{supp} f \neq \emptyset\right\}$.
Set $z_{D}:=\left(\widetilde{z}_{D}, A\left(\widetilde{z}_{D}\right)\right) \in D \cap \Gamma$, where $\widetilde{z}_{D}$ is the center of $\widetilde{D}$. By (7.1), we have $\int \varphi_{\epsilon}^{\delta}(x-$ $\left.z_{D}\right) K\left(x-z_{D}\right) f(y) d \mathcal{H}_{\Gamma}^{n}(y)=0$ for all $0<\epsilon \leq \delta$. Thus, given $m \in J$, we can decompose

$$
\begin{aligned}
\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f \mathcal{H}_{\Gamma}^{n}\right)\right)(x)= & \int \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\left(K(x-y)-K\left(x-z_{D}\right)\right) f(y) d \mathcal{H}_{\Gamma}^{n}(y) \\
& +\int\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)-\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}\left(x-z_{D}\right)\right) K\left(x-z_{D}\right) f(y) d \mathcal{H}_{\Gamma}^{n}(y)
\end{aligned}
$$

and we obtain $\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f \mathcal{H}_{\Gamma}^{n}\right)\right)(x)\right| \leq\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}\left(\Theta 1_{m}+\Theta 2_{m}\right)$, where

$$
\begin{aligned}
\Theta 1_{m} & :=\int_{D} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\left|K(x-y)-K\left(x-z_{D}\right)\right| d \mathcal{H}_{\Gamma}^{n}(y) \\
\Theta 2_{m} & :=\int_{D}\left|\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)-\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}\left(x-z_{D}\right)\right|\left|K\left(x-z_{D}\right)\right| d \mathcal{H}_{\Gamma}^{n}(y)
\end{aligned}
$$

The term $\Theta 1_{m}$ can be easily handled. For $x \in \Gamma \backslash 3 D$, we have

$$
\begin{equation*}
\Theta 1_{m} \lesssim \ell(D) \operatorname{dist}(x, D)^{-n-1} \int_{D} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y) d \mathcal{H}_{\Gamma}^{n}(y) . \tag{7.3}
\end{equation*}
$$

Let us estimate $\Theta 2_{m}$. Decompose $J=\mathcal{S} \cup \mathcal{L}$, where $\mathcal{S}$ and $\mathcal{L}$ are as in (6.4) but replacing $m \in \mathbb{Z}$ by $m \in J$, and as before, let $j\left(\epsilon_{m}\right)$ be the integer such that $\epsilon_{m} \in I_{j\left(\epsilon_{m}\right)}$. Using that $x \in \Gamma \backslash 3 D$ and $\operatorname{supp} f \subset D$, one can easily check that $\mathcal{L}$ contains a finite number of elements, and this number only depends on $n$ and $d$. Similarly, $\mathcal{S}_{j}=\emptyset$ for all $j \in \mathbb{Z}$ except on a finite number which only depends on $n$ and $d$.

Assume that $m \in \mathcal{S}$. With the same computations as those in (5.6), one can prove that, for all $y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot),\left|\nabla_{y}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\right)\right| \lesssim 2^{j\left(\epsilon_{m}\right)}\left|\epsilon_{m}-\epsilon_{m+1}\right||\widetilde{x}-\widetilde{y}|^{-1}$, because $|\widetilde{x}-\widetilde{y}| \approx \epsilon_{m} \approx \epsilon_{m+1} \approx 2^{-j\left(\epsilon_{m}\right)}$ for all $y \in \operatorname{supp} \varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-\cdot)$. Thus,

$$
\begin{equation*}
\Theta 2_{m} \lesssim \ell(D)^{n+1} \operatorname{dist}(x, D)^{-n-1} 2^{j\left(\epsilon_{m}\right)}\left|\epsilon_{m}-\epsilon_{m+1}\right| . \tag{7.4}
\end{equation*}
$$

Assume now that $m \in \mathcal{L}$. It is easy to verify that $\left|\nabla_{y}\left(\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)\right)\right| \lesssim|\widetilde{x}-\widetilde{y}|^{-1}$, so $\Theta 2_{m} \lesssim \ell(D)^{n+1} \operatorname{dist}(x, D)^{-n-1}$.

Combining this last estimate with (7.3), (7.4), the fact that $\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f \mathcal{H}_{\Gamma}^{n}\right)\right)(x)\right| \leq$ $\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}\left(\Theta 1_{m}+\Theta 2_{m}\right)$, the remark on $\mathcal{S}$ and $\mathcal{L}$ made just after (7.3), (7.2), and that $\rho>1$, we have that, for all $x \in \Gamma \backslash 3 D$,

$$
\begin{aligned}
& \left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f(x) \lesssim\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}\left(\sum_{m \in J} \Theta 1_{m}+\sum_{m \in \mathcal{S}} \Theta 2_{m}+\sum_{m \in \mathcal{L}} \Theta 2_{m}\right) \\
& \quad \lesssim \frac{\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \ell(D)^{n+1}}{\operatorname{dist}(x, D)^{n+1}}\left(\sum_{m \in J} \int_{D} \frac{\varphi_{\epsilon_{m+1}}^{\epsilon_{m}}(x-y)}{\ell(D)^{n}} d \mathcal{H}_{\Gamma}^{n}(y)+\sum_{m \in \mathcal{S}} \frac{\left|\epsilon_{m}-\epsilon_{m+1}\right|}{2^{-j\left(\epsilon_{m}\right)}}+\sum_{m \in \mathcal{L}} 1\right) \\
& \quad \lesssim \frac{\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \ell(D)^{n+1}}{\operatorname{dist}(x, D)^{n+1}} .
\end{aligned}
$$

Then, using (7.1) and standard estimates, we conclude that

$$
\int_{(3 D)^{c}}\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f(x) d \mathcal{H}_{\Gamma}^{n}(x) \lesssim \int_{(3 D)^{c}} \frac{\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \ell(D)^{n+1}}{\operatorname{dist}(x, D)^{n+1}} d \mathcal{H}_{\Gamma}^{n}(x) \lesssim 1 .
$$

7.2. The operator $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}}{ }_{\Gamma}^{n}: L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow B M O\left(\mathcal{H}_{\Gamma}^{n}\right)$ is bounded. We have to prove that there exists a constant $C>0$ such that, for any $f \in L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$ and any cube $\widetilde{D} \subset \mathbb{R}^{n}$, there exists some constant $c$ depending on $f$ and $\widetilde{D}$ such that

$$
\int_{D}\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f-c\right| d \mathcal{H}_{\Gamma}^{n} \leq C\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \mathcal{H}_{\Gamma}^{n}(D)
$$

Let $f$ and $D$ be as above, and set $f_{1}:=f \chi_{3 D}$ and $f_{2}:=f-f_{1}$. First of all, by Hölder's inequality and Theorem 6.1, we have

$$
\begin{aligned}
\int_{D}\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{1} d \mathcal{H}_{\Gamma}^{n} & \leq \mathcal{H}_{\Gamma}^{n}(D)^{1 / 2}\left(\int_{3 D}\left(\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{1}\right)^{2} d \mathcal{H}_{\Gamma}^{n}\right)^{1 / 2} \\
& \lesssim \mathcal{H}_{\Gamma}^{n}(D)^{1 / 2}\left(\left\|f_{1}\right\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}^{2} \mathcal{H}_{\Gamma}^{n}(3 D)\right)^{1 / 2} \lesssim\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \mathcal{H}_{\Gamma}^{n}(D) .
\end{aligned}
$$

Notice that $\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right)\left(f_{1}+f_{2}\right)-\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}\right| \leq\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{1}$, because $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ is sublinear and positive. Then, for any $c \in \mathbb{R}$,

$$
\begin{aligned}
\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right)\left(f_{1}+f_{2}\right)-c\right| & \leq\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right)\left(f_{1}+f_{2}\right)-\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}\right|+\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}-c\right| \\
& \leq\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{1}+\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}-c\right|
\end{aligned}
$$

hence we are reduced to prove that, for some constant $c \in \mathbb{R}$,

$$
\begin{equation*}
\int_{D}\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}-c\right| d \mathcal{H}_{\Gamma}^{n} \leq C\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)} \mathcal{H}_{\Gamma}^{n}(D) . \tag{7.5}
\end{equation*}
$$

Set $z_{D}:=\left(\widetilde{z}_{D}, A\left(\widetilde{z}_{D}\right)\right)$, where $z_{D}$ is the center of $\widetilde{D}$, and take $c:=\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}\left(z_{D}\right)$. We may assume that $c<\infty$ (this is the case if, for example, $f$ has compact support).

Given a family $a:=\left\{a_{\epsilon}\right\}_{\epsilon>0} \subset \mathbb{C}$, define its $\rho$-variation semi-norm to be

$$
\|a\|_{\mathcal{V}_{\rho}}=\sup _{\left\{\epsilon_{m}\right\}}\left(\sum_{m \in \mathbb{Z}}\left|a_{\epsilon_{m+1}}-a_{\epsilon_{m}}\right|^{\rho}\right)^{1 / \rho}
$$

where the supremum is taken over all decreasing sequences $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}} \subset(0, \infty)$. Since $\|\cdot\|_{\nu_{\rho}}$ satisfies the triangle inequality, we have $\left|\|a\|_{\mathcal{V}_{\rho}}-\|b\|_{\mathcal{V}_{\rho}}\right| \leq\|a-b\|_{\mathcal{V}_{\rho}}$ for all $a:=\left\{a_{\epsilon}\right\}_{\epsilon>0}$ and $b:=\left\{b_{\epsilon}\right\}_{\epsilon>0}$, where $a-b:=\left\{a_{\epsilon}-b_{\epsilon}\right\}_{\epsilon>0}$. Hence,

$$
\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}(x)-c\right|^{\rho} \leq \sup _{\left\{\epsilon_{m} \searrow 0\right\}} \sum_{m \in \mathbb{Z}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f_{2} \mathcal{H}_{\Gamma}^{n}\right)\right)(x)-\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f_{2} \mathcal{H}_{\Gamma}^{n}\right)\right)\left(z_{D}\right)\right|^{\rho}
$$

Given $x \in \Gamma \cap D$, let $\left\{\epsilon_{m}\right\}_{m \in \mathbb{Z}}$ be a decreasing sequence of positive numbers (which depends on $x$ ) such that

$$
\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}(x)-c\right|^{\rho} \leq 2 \sum_{m \in \mathbb{Z}}\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f_{2} \mathcal{H}_{\Gamma}^{n}\right)\right)(x)-\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f_{2} \mathcal{H}_{\Gamma}^{n}\right)\right)\left(z_{D}\right)\right|^{\rho}
$$

Notice that $\left|\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f_{2} \mathcal{H}_{\Gamma}^{n}\right)\right)(x)-\left(K \varphi_{\epsilon_{m+1}}^{\epsilon_{m}} *\left(f_{2} \mathcal{H}_{\Gamma}^{n}\right)\right)\left(z_{D}\right)\right| \leq\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}\left(\Theta 1_{m}+\Theta 2_{m}\right)$, where $\Theta 1_{m}$ and $\Theta 2_{m}$ are as in (6.10) but replacing $\mu$ by $\mathcal{H}_{\Gamma}^{n}$. It is straightforward to check that the arguments and computations given in subsection 6.2 to estimate the two terms in (6.10) (see (6.11), (6.12), and (6.13)) still hold if we replace $\mu$ by $\mathcal{H}_{\Gamma}^{n}$. Therefore, we have

$$
\sum_{m \in \mathbb{Z}=\mathcal{S} \cup \mathcal{L}}\left(\Theta 1_{m}+\Theta 2_{m}\right)^{\rho} \lesssim 1
$$

which impies that $\left|\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f_{2}(x)-c\right| \lesssim\|f\|_{L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right)}$ and, by integrating in $D$, gives (7.5).
7.3. The operator $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}: L^{p}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow L^{p}\left(\mathcal{H}_{\Gamma}^{n}\right)$ is bounded for all $1<p<\infty$. Since $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ is sublinear, the $L^{p}$ boundedness follows by applying the results of subsection 7.1 and subsection 7.2 , and the interpolation theorem between the pairs $\left(H^{1}\left(\mathcal{H}_{\Gamma}^{n}\right), L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)\right)$ and $\left(L^{\infty}\left(\mathcal{H}_{\Gamma}^{n}\right), B M O\left(\mathcal{H}_{\Gamma}^{n}\right)\right)$ in [Ju, page 43].

Given a v-cube $Q \subset \mathbb{R}^{d}$, set $m_{Q}(f):=\mathcal{H}_{\Gamma}^{n}(Q)^{-1} \int_{Q} f d \mathcal{H}_{\Gamma}^{n}$, and let $M$ denote the HardyLittlewood maximal operator with respect to $\Gamma$, i.e. for $x \in \Gamma, M(f)(x):=\sup m_{Q}(|f|)$, where the supremum is taken over all v-cubes $Q \subset \mathbb{R}^{d}$ containing $x \in \Gamma$. Let $M^{\sharp}$ be the sharp maximal operator defined by $M^{\sharp}(f)(x):=\sup m_{Q}\left(\left|f-m_{Q}(f)\right|\right)$, where the supremum is also taken over all v-cubes $Q \subset \mathbb{R}^{d}$ containing $x \in \Gamma$.

One comment about the interpolation theorem in [Ju, page 43] is in order. Given an operator $F$ bounded form $H^{1}$ to $L^{1}$ and from $L^{\infty}$ to $B M O$, in the proof of the interpolation theorem applied to $F$, one uses that $M^{\sharp} \circ F$ is sublinear (i.e. $\left(M^{\sharp} \circ F\right)(f+g) \leq\left(M^{\sharp} \circ F\right) f+$
$\left(M^{\sharp} \circ F\right) g$ for all functions $\left.f, g\right)$. This is the case when $F$ is linear. However, $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ is not linear, and then it is not clear if $M^{\sharp} \circ \mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ is sublinear. Nevertheless, this problem can be fixed easily using that $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}}{ }_{\Gamma}^{n}$ is sublinear and positive (that is $\left(\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}\right) f(x) \geq 0$ for all $f$ and $x$ ), as the following lemma shows.
Lemma 7.2. Let $F: L_{l o c}^{1}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow L_{l o c}^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ be a positive and sublinear operator. Then $\left(M^{\sharp} \circ F\right)(f+g) \lesssim(M \circ F) f+\left(M^{\sharp} \circ F\right) g$ for all functions $f, g$.
Proof. If $F$ is sublinear and positive, one has that $|F(f)(x)-F(g)(x)| \leq F(f-g)(x)$ for all functions $f, g \in L_{l o c}^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$. In particular, $|F(f+g)(x)-F(g)(x)| \leq F(f)(x)$. Then, for $x, y \in Q \cap \Gamma$,

$$
\begin{aligned}
\left|F(f+g)(y)-m_{Q}(F g)\right| & \leq|F(f+g)(y)-F g(y)|+\left|F g(y)-m_{Q}(F g)\right| \\
& \leq|F f(y)|+\left|F g(y)-m_{Q}(F g)\right| .
\end{aligned}
$$

Hence, $m_{Q}\left|F(f+g)-m_{Q}(F g)\right| \leq m_{Q}|F f|+m_{Q}\left|F g-m_{Q}(F g)\right| \leq(M \circ F) f(x)+\left(M^{\sharp} \circ F\right) g(x)$ and, by taking the supremum over all possible v-cubes $Q \ni x$, we conclude $\left(M^{\sharp} \circ F\right)(f+$ $g)(x) \lesssim(M \circ F) f(x)+\left(M^{\sharp} \circ F\right) g(x)$.

By using Lemma 7.2 and the fact that $\|M f\|_{L^{p}\left(\mathcal{H}_{\Gamma}^{n}\right)} \lesssim\left\|M^{\sharp} f\right\|_{L^{p}\left(\mathcal{H}_{\Gamma}^{n}\right)}$ for $f \in L^{p_{0}}\left(\mathcal{H}_{\Gamma}^{n}\right) \cap$ $L^{p}\left(\mathcal{H}_{\Gamma}^{n}\right)$ and $1 \leq p_{0} \leq p<\infty$, one can reprove Journé's interpolation theorem applied to $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}$ with minor modifications in the original proof.
7.4. The operator $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}_{\Gamma}^{n}}: L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right) \rightarrow L^{1, \infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$ is bounded. By adapting [CJRW2, Theorem B] to our setting and using the smoothnes of the family $\varphi$, one can show that the $L^{2}\left(\mathcal{H}_{\Gamma}^{n}\right)$ boundedness of $\mathcal{V}_{\rho} \circ \mathcal{T}_{\varphi}^{\mathcal{H}}{ }_{\Gamma}^{n}$ yields the boundedness of this operator from $L^{1}\left(\mathcal{H}_{\Gamma}^{n}\right)$ to $L^{1, \infty}\left(\mathcal{H}_{\Gamma}^{n}\right)$. The interested reader may see $[\mathrm{MT}]$, where a more general result is proved.

## References

[AJS] M. Akcoglu, R. L. Jones, and P. Schwartz, Variation in probability, ergodic theory and analysis, Illinois J. of Math, 42 (1998), 154-177.
[Bo] J. Bourgain, Pointwise ergodic theorems for arithmetic sets, Inst. Hautes Etudes Sci. Publ. Math. 69 (1989), 5-45.
[Ca] A. P. Calderón, Cauchy integrals on Lipschitz curves and related operators, Proc. Nat. Acad. Sci. U.S.A 74 (1977), 1324-1327.
[CJRW1] J. Campbell, R. L. Jones, K. Reinhold, and M. Wierdl, Oscillation and variation for the Hilbert transform, Duke Math. J. 105 (2000), 59-83.
[CJRW2] J. Campbell, R. L. Jones, K. Reinhold, and M. Wierdl, Oscillation and variation for singular integrals in higher dimensions, Trans. Amer. Math. Soc. 35 (2003), 2115-2137.
[CMM] R. Coifman, A. McIntosh, and Y. Meyer, L'intégrale de Cauchy définit un opérateur borné sur $L^{2}$ pour les courbes lipschitziennes, Ann. of Math. 116 (1982), 361-387.
[Da] G. David, Wavelets and singular integrals on curves and surfaces, Lecture Notes in Math. 1465, Springer-Verlag, Berlin, 1991.
[DS1] G. David and S. Semmes, Singular integrals and rectifiable sets in $R^{n}$ : au-delà des graphes lipschitziens, Astérisque No. 193 (1991).
[DS2] G. David and S. Semmes, Analysis of and on uniformly rectifiable sets, Mathematical Surveys and Monographs, 38, American Mathematical Society, Providence, RI, 1993.
[Du] J. Duoandikoetxea, Fourier Analysis, Graduate Studies in Mathematics, vol. 29 (2001). American Math. Soc., Providence, RI.
[GT] T. A. Gillespie and J. L. Torrea Dimension free estimates for the oscillation of Riesz transforms, Israel Journal of Math. 141 (2004), 125-144.
[Jn1] P. W. Jones, Square functions, Cauchy integrals, analytic capacity, and harmonic measure, Harmonic analysis and partial differential equations (El Escorial, 1987), 24-68, Lecture Notes in Math., 1384, Springer, Berlin, 1989.
[Jn2] P. W. Jones, Rectifiable sets and the travelling salesman problem, Invent. Math. 102 (1990), 1-15.
[JKRW] R. L. Jones, R. Kaufman, J. Rosenblatt, and M. Wierdl, Oscillation in ergodic theory, Ergodic Theory and Dynam. Sys. 18 (1998), 889-936.
[JRW] R. L. Jones, J. Rosenblatt, and M. Wierdl, Oscillation in ergodic theory: higher dimensional results, Israel Journal of Math. 135 (2003), 1-27.
[JSW] R. L. Jones, A. Seeger, and J. Wright, Strong variational and jump inequalities in harmonic analysis, Trans. Amer. Math. Soc. 360 (2008), 6711-6742.
[JW] R. L. Jones and G. Wang, Variation inequalities for Fejér and Poisson kernels, Trans. Amer. Math. Soc. 356 (2004), 4493-4518.
[Ju] J. Journé, Calderón-Zygmund operators, pseudo-differential operators and the Cauchy integral of Calderón, Lecture Notes in Math., 994, Springer, Berlin, 1983.
[LT] M. Lacey and E. Terwilleger, A Wiener-Wintner theorem for the Hilbert transform, Ark. Mat. 46 (2008), 2, 315-336.
[Lé] D. Lépingle, La variation d'ordre p des semi-martingales, Z. Wahrscheinlichkeitstheorie Verw. Gebiete 36 (1976), 295-316.
[MT] A. Mas and X. Tolsa, Variation for the Riesz transform and uniform rectifiability, submitted.
[Ma] P. Mattila. Geometry of sets and measures in Euclidean spaces, Cambridge Stud. Adv. Math. 44, Cambridge Univ. Press, Cambridge, 1995.
[MMV] P. Mattila, M. S. Melnikov and J. Verdera, The Cauchy integral, analytic capacity, and uniform rectifiability, Ann. of Math. (2) 144 (1996), 127-136.
[MV] P. Mattila and J. Verdera Convergence of singular integrals with general measures, J. Eur. Math. Soc. 11 (2009), 257-271.
[OSTTW] R. Oberlin, A. Seeger, T. Tao, C. Thiele, and J. Wright A variation norm Carleson theorem, accepted in J. Eur. Math. Soc.
[Pa] H. Pajot, Analytic capacity, rectifiability, Menger curvature and the Cauchy integral, Lecture Notes in Math. 1799, Springer, 2002.
[Qi] J. Qian, The p-variation of partial sum processes and the empirical process, Ann. of Prob. 26 (1998), 1370-1383.
[To] X. Tolsa, Uniform rectifiability, Calderón-Zygmund operators with odd kernel, and quasiorthogonality, Proc. London Math. Soc. 98(2) (2009), 393-426.
[Vi] C. Villani, Topics in optimal transportation, Graduate Studies in Mathematics, vol. 58 (2003). American Math. Soc., Providence, RI.

Albert Mas. Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Catalonia, Spain

E-mail address: amblesa@mat.uab.cat
Xavier Tolsa. Institució Catalana de Recerca i Estudis Avançats (ICREA) and Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Catalonia

E-mail address: xtolsa@mat.uab.cat


[^0]:    Date: March, 2011.
    2010 Mathematics Subject Classification. Primary 42B20, 42B25.
    Key words and phrases. $\rho$-variation, oscillation, Calderón-Zygmund singular integrals.
    Both authors are partially supported by grants 2009SGR-000420 (Generalitat de Catalunya) and MTM2010-16232 (Spain). Albert Mas is also supported by grant AP2006-02416 (FPU program, Spain).

