CORONA DECOMPOSITIONS FOR PARABOLIC UNIFORMLY RECTIFIABLE SETS

S. BORTZ, J. HOFFMAN, S. HOFMANN, J.L. LUNA-GARCIA, AND K. NYSTRÖM

ABSTRACT. We prove that parabolic uniformly rectifiable sets admit (bilateral) corona decompositions with respect to regular Lip(1,1/2) graphs. Together with our previous work, this allows us to conclude that if $\Sigma \subset \mathbb{R}^{n+1}$ is parabolic Ahlfors-David regular, then the following statements are equivalent.
   (1) $\Sigma$ is parabolic uniformly rectifiable.
   (2) $\Sigma$ admits a corona decomposition with respect to regular Lip(1,1/2) graphs.
   (3) $\Sigma$ admits a bilateral corona decomposition with respect to regular Lip(1,1/2) graphs.
   (4) $\Sigma$ is big pieces squared of regular Lip(1,1/2) graphs.

CONTENTS

1. Introduction 2
2. Preliminaries 6
3. Statement of the main results 11
   3.1. The proof of Theorem 3.1 12
4. Construction of Lip(1,1/2) graphs for stopping time regimes 14
5. Pushing the geometric square function to the graph 23
   5.1. Auxiliary lemmas 24
   5.2. Proof of Lemma 5.6 29
   5.3. Proof of Lemma 5.10 31
6. Packing the maximal cubes 38
7. The proof of Theorem 3.3 48
Appendix A. Two lemmas concerning approximating planes 52
Appendix B. Slice-wise vs. Parabolic Hausdorff Measure 53
References 56

Date: February 8, 2023.
2010 Mathematics Subject Classification. 28A75, 30L99, 43A85.
The authors J. H., S. H., and J.L. L.-G. were partially supported by NSF grants DMS-1664047 and DMS-2000048. K.N. was partially supported by grant 2017-03805 from the Swedish research council (VR).
1. Introduction

The results proved in this paper are part of a program devoted to providing parabolic analogues of the influential work of G. David and S. Semmes [DS1, DS2] concerning uniformly rectifiable sets. At the time [DS1] was written, tools from the Calderón-Zygmund school of harmonic analysis were being rapidly adapted to less classical and rougher geometrical settings. A particularly important development was the proof of the $L^2$ boundedness of the Cauchy integral on Lipschitz curves [CMM], and subsequent developments concerning the boundedness of singular integral operators of Calderón type on Lipschitz graphs and curves [CDM] (see also [CCFJR, D, M, LMcS, CS]). In addition, the Calderón-Zygmund theory was extended beyond the Euclidean setting (see e.g. [Chr]).

When restricting to the setting of $d$-dimensional subsets of $n$-dimensional Euclidean space, David and Semmes [DS1, DS2] found necessary and sufficient conditions on these sets for the validity of the $L^2$ boundedness of $d$-dimensional singular integral operators. David and Semmes coined the sets having this property uniformly rectifiable, and through a remarkable series of works they provided numerous characterizations of these sets. Of particular relevance to this paper is the following theorem stating a subset of the equivalent characterizations of uniformly rectifiable sets. We refer to [DS1] for exact definitions of the different notions appearing in Theorem 1.1.

**Theorem 1.1 ([DS1]).** Let $E \subset \mathbb{R}^{n+1}$ be $n$-dimensional Ahlfors(-David) regular. Then the following statements are equivalent.

(a) $E$ is uniformly rectifiable.

(b) $E$ admits a corona decomposition with respect to Lipschitz graphs.

(c) All sufficiently regular convolution type Calderon-Zygmund operators with odd kernels are $L^2$ bounded on $E$.

The purpose of this paper is to prove that the implication $(a) \implies (b)$ holds in the context of parabolic uniformly rectifiable sets (see Definition 2.19). To prove that the implication $(c) \implies (b)$ holds in the parabolic setting is, as of writing, an interesting and open problem. Combining the main results established in this paper, see Theorem 3.1 and Theorem 3.3 below, with the results of our previous work [BHHLN1], we obtain the following.

**Theorem 1.2.** Let $\Sigma \subset \mathbb{R}^{n+1}$ be parabolic Ahlfors-David regular (see Definition 2.14). Then the following statements are equivalent.

(a) $\Sigma$ is parabolic uniformly rectifiable.

(b) $\Sigma$ admits a corona decomposition with respect to regular Lip$(1,1/2)$ graphs.

(c) $\Sigma$ admits a bilateral corona decomposition with respect to regular Lip$(1,1/2)$ graphs.

(d) $\Sigma$ is big pieces squared of regular Lip$(1,1/2)$ graphs.
All implications stated in Theorem 1.2, but the implications (a) \(\implies\) (b) and (a) \(\implies\) (c), are contained\(^1\) in [BHHLN1]. We will prove\(^2\) directly that (a) and (b) imply (c) by essentially re-running the machine that we use to prove that (a) \(\implies\) (b), and by using additional information gleaned from the fact that (b) \(\implies\) (d). This means that the bulk of the paper is mainly devoted to the proof of the implication (a) \(\implies\) (b) and our proof will involve a very careful adaptation of the proof in [DS1], along with some additional arguments utilizing ideas from [HLN1]. The implication (a) \(\implies\) (b) is the content of Theorem 3.1 stated in precise form below. The precise statement of the result that (a) \(\implies\) (c) is given in Theorem 3.3.

The study of parabolic uniformly rectifiable sets \(\Sigma \subset \mathbb{R}^{n+1}\) emanates in [HLN1], [HLN2] where the third and fifth author, together with J. Lewis, introduced the notion of parabolic uniformly rectifiable sets. In [HLN1] the existence of big pieces of regular Lip(1,1/2) graphs, under the additional assumption that \(\Sigma\) is Reifenberg flat in the parabolic sense, is established. The results in [HLN1], [HLN2] were the first of their kind in the context of parabolic problems and the studies [HLN1], [HLN2] were motivated by the study of parabolic/caloric measure in rough time-varying domains.

By definition, a Lip(1,1/2) graph is the graph of a function which is Lipschitz with respect to the parabolic metric. A Lip(1,1/2) graph is a regular Lip(1,1/2) graph if, in addition, the 1/2-order time derivative of the function defining the graph is in (parabolic) BMO (see Definition 2.11). Combining the results of Section 2 in [HLN2] with the developments in [LM, H2, HL1], it follows that a Lip(1,1/2) graph is parabolic uniformly rectifiable in the sense of [HLN1], [HLN2] if and only if it is a regular Lip(1,1/2) graph. This explains the importance of regular Lip(1,1/2) graphs in the statement of Theorem 1.2, and in the statements of Theorem 3.1 and Theorem 3.3. As proved in [LS], [KW], the class of regular Lip(1,1/2) graphs is strictly smaller than the class of Lip(1,1/2) graphs.

The notions of regular Lip(1,1/2) graphs and parabolic uniform rectifiability are deeply rooted in the study of the Dirichlet problem for the heat equation in time-varying (graph) domains, and the solvability of the \(L^p\)-Dirichlet problem for the heat equation is intimately connected to quantitative mutual absolute continuity of the caloric measure and the surface measure. In particular, in [KW] it was proved that there are Lip(1,1/2) graph domains for which caloric measure and surface measure are mutually singular. Later, in their pioneering work, Lewis, Silver and Murray [LS, LM] proved that for regular Lip(1,1/2) graph domains, the caloric measure and the surface measure are quantitatively related in the sense that they are mutually absolutely continuous, and the associated parabolic Poisson kernel satisfies a scale-invariant reverse Hölder inequality in \(L^p\) for some \(p \in (1, \infty)\). The importance and relevance of regular Lip(1,1/2) graph domains, from the perspective of parabolic singular integrals, layer potentials, boundary value problems and inverse problems, is emphasized through the works in [LS, LM, H1, H2, HL1, HL2, LN, N1].

---

\(^1\)The implication (c) \(\implies\) (b) is trivial.

\(^2\)We could also follow the more indirect method in [HMM1], but we present an alternative approach here.
particular, in \([\text{HL}1]\) the solvability of the \(L^2\)-Dirichlet problem (and of the \(L^2\)-Neumann problem) for the heat equation, using layer potentials, was obtained in the region above a regular Lip\((1,1/2)\) graph under the restriction that the \(1/2\)-order time derivative (measured in BMO) of the function defining the graph is small\(^3\). The highly influential work \([\text{HL}3]\) (see also \([\text{DPP}]\)), devoted to parabolic operators with singular terms, should also be mentioned in this context. In particular, in \([\text{HL}3]\) the method of extrapolation of Carleson measure estimates was introduced, a method which was crucial in the resolution of the Kato conjecture, see \([\text{AHL}e\text{Mc}T], [\text{AHL}McT]\).

In \([\text{HL}N1], [\text{HL}N2]\) parts of the analysis on regular Lip\((1,1/2)\) graph domains mentioned above was extended beyond the setting of graphs. However, other than \([\text{HL}N1], [\text{HL}N2]\) only a few notable works in rougher parabolic settings have appeared \([\text{E}, \text{GH}1, \text{GH}2, \text{N}, \text{NS}, \text{MP}, \text{MPT}]\). This is in contrast to the corresponding elliptic contexts, see, for example, \([\text{DJ}, \text{HMM}1, \text{GMT}, \text{AGMT}, \text{HLMN, MT, AHMMT}]\), and it is a long term goal of the authors to develop a corresponding parabolic theory. The present paper is one of the starting points of that effort. In particular, in forthcoming work we will show, using Theorem 1.2, that the main theorem in \([\text{HMM}1]\) can be extended to the parabolic setting. In addition, in \([\text{BHHLN}2]\) we have recently obtained, by expanding on \([\text{NS}]\), a flexible parabolic analogue of the work of David and Jerison \([\text{DJ}]\). Still, to establish parabolic analogues of, for example, \([\text{GMT, AHM}M\text{MT, HLM}N]\) remain open problems which may require new methods and tools.

As mentioned, in the elliptic setting, some of the most significant recent progress in this area relies on methods which are currently unavailable in the parabolic setting. Thus, to prove parabolic versions of those results, one will require significantly new ideas. However, to prove the results of this paper, we are able to develop parabolic versions of the methods in \([\text{DS}1, \text{HMM}1]\). We prove two main results, whose statements may be found below. First, Theorem 3.1, whose elliptic analogue appears in \([\text{DS}1]\), and we give the proof of Theorem 3.1 in Section 4, Section 5 and Section 6. Second, Theorem 3.3, whose elliptic analogue appears in \([\text{HMM}1]\), and we give the proof of Theorem 3.3 in Section 7. Concerning Theorem 3.1, it turns out that we can follow quite closely the corresponding arguments in \([\text{DS}1]\). The reader who is familiar with that work will certainly recognize substantial parts of the proof. The novelties in Section 4, Section 5 and Section 6 are in many respects mainly of technical nature: they arise here to treat the extension to the parabolic setting, and they rely in part on some ideas from \([\text{HL}N1]\). On the other hand, the corona decomposition is of central importance and utility in the theory of quantitative rectifiability, and we believe that this fact justifies a complete and careful treatment in the parabolic case. Finally, we note that the converse to Theorem 3.1, which we have established (in particular) in \([\text{BHHLN}1]\), does indeed rely on some new methods that are significantly different to those previously appearing in the literature.

\(^3\)This smallness is sharp in the sense that there are regular Lip\((1,1/2)\) graph domains for which the \(L^2\)-Dirichlet problem is not solvable. On the other hand, the \(L^p\)-Dirichlet problem is solvable for some \(p < \infty\) for all regular Lip\((1,1/2)\) graph domains \([\text{LM}]\).
As we have previously mentioned in [BHHLN1, BHHLN2], it is true that in [RN1, RN2, RN3], the author took on the ambitious challenge to develop the theory of parabolic uniformly rectifiable sets. Unfortunately though, at least in the papers in [RN1, RN2], the author either gives no proofs of his claims or he supplies proofs which have fundamental errors. In the present work we give a correct proof of the main result claimed in [RN1], and we shall explain momentarily two principal errors (as well as a third, more technical error) in the latter paper. The errors in [RN2] are discussed in some detail in our paper [BHHLN2], which gives a correct proof of those results.

In [RN1] the author claims to prove that parabolic uniform rectifiability implies corona decompositions with respect to regular Lip(1,1/2) graphs. In [RN1], the author, as we do here, follows rather closely the outline of the proof of the corresponding elliptic result in [DS1], while also, as we do, employing some ideas from [HLN1] in order to adapt the arguments of [DS1] to the parabolic case. It is in the latter context that the proof in [RN1] breaks down, for the following reason: in Lemma 2.7 in [RN1], the constant \( \hat{c} \) cannot necessarily be taken small, given only the assumption of parabolic uniform rectifiability. In [HLN1], the constant is small by virtue of the assumption of Reifenberg flatness used in that paper. Without smallness, the sets \( \{E_i\} \), introduced in subsection 2.3.1 in [RN1], need not, as claimed, have bounded overlaps and therefore the estimate in display (20) in [RN1] is not valid. This is a critical gap in the proof: it means that in [RN1] there is no proof of the “regularity” (see Definition 2.11 below for the definition of a regular Lip(1,1/2) graph) of the approximating graphs, which is the essential feature required of the graphs in the parabolic setting (and which has no analogue in the elliptic case); thus, the proof in [RN1] breaks down precisely at the point in the argument where one needs to do something new in the parabolic setting that has no counterpart in the corresponding elliptic proof. Below we deal with this problem by using a counting function introduced in [DS1], see Lemma 5.12 below.

A second critical error in [RN1] appears in the claimed proof of [RN1, Theorem 1.5], which in turn relies on the claimed result [RN1, Theorem 3.1]. In fact, the latter is false. Indeed, in contrast to the situation in the elliptic setting, parabolic uniform rectifiability is not characterized by Carleson set conditions such as the “Bilateral Weak Geometric Lemma”, even in the case of a Lip(1,1/2) graph; see [BHHLN1, Observation 4.19] for a discussion of this issue and a counter-example. Thus the proof of [RN1, Theorem 1.5] collapses.

In addition, we mention a technical (but still serious) error, in the crucial estimate in the display between displays (32) and (33) in [RN1]. Here, the author claims a Poincaré inequality that need not hold if the function \( \psi_2 \) depends on time, which it certainly does since the problem is parabolic; thus, again the error arises because the author in [RN1] did not address the difference between the elliptic and

---

4Even the reduction of [RN1, Theorem 1.5] to [RN1, Theorem 3.1] is not clear to the authors. Indeed, the claimed bilateral approximation does not follow as in the elliptic setting as ‘nice’ parabolic Littlewood-Paley kernels are always odd in \( x \), but not necessarily in \( (x,t) \) jointly. As a consequence, they are unable to detect the difference between a \( t \)-independent plane and the same plane omitting the points with \( t \in (a,b) \). In particular, they do not detect all ‘holes’; this would be essential for [RN1, Claim 1, Section 3] to “be easily adapted from [DS1, page 24]”, as claimed therein.
parabolic settings. In this context, the correct estimates, based on a more refined inequality of Poincaré type appropriate to the parabolic setting, are given in our Lemma 6.31 below.

The rest of the paper is organized as follows. Section 2 is mainly of preliminary nature and we here provide notation and definitions. In Section 3 we state our main result: Theorem 3.1 and Theorem 3.3. In Subsection 3.1 we briefly discuss the outline of the proof of Theorem 3.1 to be presented in subsequent sections. The proof is divided into several parts and our argument adapts the corresponding arguments of [DS1], also utilizing some ideas from [HLN1], to deal with the extension to the parabolic setting. The proof of Theorem 3.1 is given in Section 4, Section 5 and Section 6 and progresses from the construction of a Lip(1,1) graph with small constant for each stopping time regime, to the pushing of the geometric square function estimate (the Carleson measure estimate) to the graph, also verifying that the graph is actually a regular Lip(1,1/2) graph, and finally to the packing of the stopping time regimes. In Section 7 we then prove Theorem 3.3 by improving Theorem 3.1(4) to give a bilateral approximation as in Theorem 3.3(4). The paper ends with two appendix containing a few technical estimates and observations used in the paper.

2. Preliminaries

In this section we provide notation and definitions. We will for simplicity assume $n \geq 2$. Our results also remain valid in the case $n = 1$, but in this case some notations used in the paper have to be adjusted. Points in Euclidean $(n + 1)$-space $\mathbb{R}^{n+1}$ are denoted by $(X, t) = (x_1, \ldots, x_n, t)$, where $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ are the spatial coordinates and $t$ represents the time-coordinate. We set
d := n + 1.

In general, $c$ will denote a positive constant satisfying $1 \leq c < \infty$. We write $c_1 \lesssim c_2$, if $c_1/c_2$ is bounded from above by a positive constant depending on the structural constants of a lemma or theorem (e.g. $n, d$, the constant defining the Ahlfors-David regularity or the constants defining parabolic uniform rectifiability). We write $c_1 \sim c_2$, if $c_1 \lesssim c_2$ and $c_2 \lesssim c_1$.

We let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{R}^n$ and we let $|X| = \langle X, X \rangle^{1/2}$ be the Euclidean norm of $X$. We let $||X, t|| := |X| + |t|^{1/2}$. Given $(X, t), (Y, s) \in \mathbb{R}^{n+1}$ we let
\[ d_p(X, t, Y, s) := d_p((X, t), (Y, s)) = |X - Y| + |t - s|^{1/2}, \]
and we define $d_p(X, t, E)$ to equal the parabolic distance, defined with respect to $d_p(\cdot, \cdot)$, from $(X, t) \in \mathbb{R}^{n+1}$ to $E \subset \mathbb{R}^{n+1}$. We let $\text{diam}(E)$ denote the parabolic diameter of $E$, that is,
\[ \text{diam}(E) := \sup \{d_p(X, t, Y, s) : (X, t), (Y, s) \in E\}. \]

We let
\[ (2.1) \quad C_r(X, t) := \{(Y, s) \in \mathbb{R}^n \times \mathbb{R} : |y_i - x_i| < r, |t - s| < r^2\}, \]
whenever \((X, t) \in \mathbb{R}^{n+1}, r > 0\), and we will refer to \(C_r(X, t)\) as a parabolic cube of “length” \(r\) (in \(\mathbb{R}^{n+1}\)). For \((x, t) \in \mathbb{R}^n\), we will use the notation
\[
C_r(x, t) := \{(y, s) \in \mathbb{R}^{n-1} \times \mathbb{R} : |y_i - x_i| < r, |t - s| < r^2\},
\]
eto denote a parabolic cube in \(\mathbb{R}^n\).

We let \(dX, dx\) denote Lebesgue measure on \(\mathbb{R}^n, \mathbb{R}^{n-1}\), respectively, and, given \(\eta > 0\), we let \(H^\eta\) denote the (standard Euclidean) \(\eta\)-dimensional Hausdorff measure. To continue we need to define Hausdorff measure adapted to the parabolic setting. This is done in a straightforward way replacing the standard Euclidean diameter with its parabolic counterpart.

**Definition 2.3.** For \(\eta, \delta > 0, E \subseteq \mathbb{R}^{n+1}\), we define
\[
H_{\eta, \delta}^\eta(E) := \inf \left\{ \sum \text{diam}(E_j)^\eta : E \subseteq \bigcup_j E_j, \text{diam}(E_j) \leq \delta \right\}.
\]
We define the parabolic Hausdorff measure, of homogeneous dimension \(\eta\), as
\[
H^\eta_p(E) := \lim_{\delta \to 0} H_{\eta, \delta}^\eta(E).
\]

Note that the limit in (2.4) exists (possibly as \(+\infty\)) since \(H_{\eta, \delta}^\eta(E)\) is a decreasing function on \((0, \infty)\).

In the sequel, \(\Sigma \subset \mathbb{R}^{n+1}\) will denote a closed set which will be clear from the context. Below we will, for \((X, t) \in \Sigma\) and \(r > 0\), use the notation \(\Delta(X, t, r) = \Delta(X, t) := \Sigma \cap C_r(X, t)\) to denote a (parabolic) “surface ball”. We define
\[
\sigma := \sigma_\Sigma := H^{n+1}_p|_\Sigma,
\]
to denote the parabolic surface measure on \(\Sigma\).

Also other notions of surface measure can be considered in the parabolic context. Indeed, we let
\[
\mu(E) := \int_{\mathbb{R}} \int_{\mathbb{R} \times \{t\}} 1_E(X, t) d\mathcal{H}^{n-1}(X) d\mathcal{H}^1(t),
\]
where \(1_E\) is the indicator function for the set \(E\), denote a slice-wise or product-like measure of the set \(E \subset \mathbb{R}^{n+1}\). This measure has been used in previous works concerning parabolic uniform rectifiability as in [KW, LM, H2, HL1, HL2, HLN1, HLN2], the parabolic surface measure was defined as
\[
\sigma^\delta := \sigma|_\Sigma.
\]

In this paper we do not use the measure in (2.7), instead we use the measure defined in (2.5). We here collect a number of remarks concerning this choice. The non-trivial statements in Remark 2.8, (i), (v), and (vi), are proved in Appendix B below.

**Remark 2.8.**
(i) If \(\sigma^\delta\) (or for that matter any measure \(m\) defined on \(\Sigma\)) satisfies the parabolic Ahlfors-David Regularity (p-ADR) condition (see Definition 2.14 below), then so does \(\sigma\), and in that case the two measures are of course equivalent. This follows easily from the definition of the \(H^{n+1}_p\) measure, and it is really
just the same phenomenon that occurs in the classical (elliptic) case, see [DS1]. We give the proof for $\sigma^a$ in Appendix B.

(ii) Consequently, if $\Sigma$ is a Lip($1,1/2$) graph, then $\sigma \approx \sigma^a$. In particular, on a hyperplane $\mathcal{P} \subset \mathbb{R}^{n+1}$ parallel to the $t$-axis, which we may identify with Euclidean space $\mathbb{R}^n$, we have that $\mathcal{H}^a|_{\mathcal{P}} \approx \mathcal{H}^{a+1}|_{\mathcal{P}}$, since the former is just $n$-dimensional Lebesgue measure on $\mathcal{P}$, which is parabolic ADR on $\Sigma = \mathcal{P}$.

(iii) If $\mathcal{P}$ is a hyperplane parallel to the $t$-axis, and if $\pi$ is the orthogonal projection operator onto $\mathcal{P}$, then $\mathcal{H}^{a+1}$ measure does not increase under the action of $\pi$. In particular, by virtue of (ii), we have, for any Borel set $A$, that $\mathcal{H}^a(\pi(A)) = \mathcal{H}^{a+1}(\pi(A)) \leq \mathcal{H}^{a+1}(A)$.

(iv) If $\Sigma$ is parabolic uniformly rectifiable (p-UR set; see Definition 2.19 below), defined with respect to $\sigma^a$, then using (i) one deduces that $\Sigma$ is also p-UR defined with respect to $\sigma$. Thus, a priori p-UR with respect to $\sigma^a$ is a stronger notion, that is, it implies p-UR with respect to $\sigma$.

(v) If $\Sigma$ is parabolic uniformly rectifiable (p-UR; see Definition 2.19 below), with respect to $\sigma$, then the two measures $\sigma^a$ and $\sigma$ are equivalent and hence one deduces (in the same way as (iv)) that $\Sigma$ is parabolic uniformly rectifiable, defined with respect to $\sigma^a$. Thus, a posteriori the notions of p-UR are equivalent.

(vi) The measures are not equivalent in general, even in the p-ADR setting. In fact, $\sigma^a \leq c(n)\sigma$, but the other direction does not need to hold.

Definition 2.9. We say that a $n$-dimensional hyperplane $P \subset \mathbb{R}^{n+1}$ is a $t$-independent plane if $P$ contains a line in the $t$-direction. Equivalently, if $\hat{N}_P$ is the normal to $P$, then $\hat{N}_P \cdot (0,1) = 0$. We let $\mathcal{P}$ denote the collection of all $t$-independent planes. If $P \in \mathcal{P}$ we let $P_\times \subset \mathbb{R}^n$ be the $(n-1)$-dimensional plane that defines $P = P_\times \times \mathbb{R}$.

Given a function $\psi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ we let $D_{1/2}^t\psi(x,t)$ denote the $1/2$ derivative in $t$ of $\psi(x,\cdot)$, $x$ fixed. This half derivative in time can be defined by way of the Fourier transform or by

$$D_{1/2}^t\psi(x,t) \equiv \hat{\psi}(x,s) - \psi(x,t) \overline{\hat{\psi}(x,s)} \| s - \hat{t} \|^{3/2} ds,$$

for properly chosen $\hat{\psi}$. We let $\| \cdot \|$ denote the norm in parabolic $BMO(\mathbb{R}^n)$ (replace standard cubes by parabolic cubes in the definition of $BMO$).

Definition 2.11. A function $\psi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ is called Lip($1,1/2$) with constant $b_1$, if

$$|\psi(x,t) - \psi(y,s)| \leq b_1(|x-y| + |t-s|^{1/2})$$

whenever $(x,t) \in \mathbb{R}^{n-1}, (y,s) \in \mathbb{R}^n$. If $\Sigma = \{(x,\psi(x,t),t) : (x,t) \in \mathbb{R}^{n-1} \times \mathbb{R}\}$ in the coordinates $P_\times \times (P_\times)^\perp \times \mathbb{R}$, for some $t$-independent plane $P \in \mathcal{P}$ and Lip($1,1/2$) function $\psi$, then we say that $\Sigma$ is a Lip($1,1/2$) graph. We say that $\psi = \psi(x,t)$ :
$\mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ is a \textit{regular} $\text{Lip}(1,1/2)$ function with parameters $b_1$ and $b_2$, if $\psi$ satisfies (2.12) and if
\begin{equation}
D_{1/2}^t \psi \in BMO(\mathbb{R}^n), \quad \|D_{1/2}^t \psi\|_* \leq b_2 < \infty. \tag{2.13}
\end{equation}
If $\Sigma = \{(x, \psi(x, t), t) : (x, t) \in \mathbb{R}^{n-1} \times \mathbb{R}\}$ in the coordinates $P_s \times (P_s)^2 \times \mathbb{R}$, for some $t$-independent plane $P \in \mathcal{P}$ and \textit{regular} $\text{Lip}(1,1/2)$ function $\psi$, then we say that $\Sigma$ is a \textit{regular} $\text{Lip}(1,1/2)$ graph.

\textbf{Definition 2.14.} Let $\Sigma \subset \mathbb{R}^{n+1}$ be a closed set. We say that a measure $\mu$, defined on $\Sigma$, is \textit{parabolic Ahlfors-David regular} (parabolic ADR) for short (or simply p-ADR) with constant $M \geq 1$, if
\begin{equation}
M^{-1} r^{d} \leq \mu(\Delta(X, t, r)) \leq M r^{d}, \tag{2.15}
\end{equation}
whenever $0 < r < \text{diam} \Sigma$, $(X, t) \in \Sigma$, $T_0 < t < T_1$ and where $\text{diam} \Sigma$ is the (parabolic) diameter of $\Sigma$ (which may be infinite). If
\begin{equation}
M^{-1} r^{d} \leq \sigma(\Delta(X, t, r)) \leq M r^{d}, \tag{2.16}
\end{equation}
then we simply say that $\Sigma$ is \textit{parabolic ADR} (p-ADR, or just ADR).

As noted above (see Remark 2.8 (i)), if (2.15) holds for a measure $\mu$ on $\Sigma$, then it holds for $\sigma$ as in (2.5), i.e. (2.16) holds for a possibly different but still universal choice of $M$.

Following [HLN1], [HLN2] we next introduce parabolic $\beta_2$ numbers, which we will denote by $\gamma$. These are defined in a way analogous to the usual $\beta$ numbers in [DS1, DS2] except that they incorporate parabolic scalings, parabolic surface measure and they only allow for approximation by $t$-independent planes. Indeed, assume that $\Sigma \subset \mathbb{R}^{n+1}$ is parabolic ADR in the sense of Definition 2.14. We then let
\begin{equation}
\gamma(Z, \tau, r) := \gamma_2(Z, \tau, r) := \inf_{P \in \mathcal{P}} \left( \iint_{\Delta(Z, \tau, r)} \left( \frac{d_P(Y, s, P)}{r} \right)^2 \, d\sigma(Y, s) \right)^{1/2}, \tag{2.17}
\end{equation}
whenever $(Z, \tau, r) \in \Sigma$, $r > 0$, and where $\mathcal{P}$ is the set of $n$-dimensional hyperplanes $P$ containing a line parallel to the $t$ axis. We also introduce
\begin{equation}
dv(Z, \tau, r) := \frac{d\gamma_2(Z, \tau, r)}{r} = (\gamma_2(Z, \tau, r))^2 \, d\sigma(Z, \tau) r^{-1} \, dr. \tag{2.18}
\end{equation}
Recall that $\nu$ is defined to be a Carleson measure on $\Delta(Y, s, R) \times (0, R)$, if there exists $\Gamma < \infty$ such that
\begin{equation}
\nu(\Delta(X, t, \rho) \times (0, \rho)) \leq \Gamma \rho^d, \tag{2.19}
\end{equation}
whenever $(X, t) \in \Sigma$ and $C_\rho(X, t) \subset C_R(Y, s)$. The least such $\Gamma$ in (2.19) is called the Carleson norm of $\nu$ on $\Delta(Y, s, R) \times (0, R)$.

We are now ready to give the definition of parabolic uniform rectifiability.

\textbf{Definition 2.19.} Assume that $\Sigma \subset \mathbb{R}^{n+1}$ is parabolic ADR in the sense of Definition 2.14 with constant $M$. Let $\nu = \gamma_2$ be defined as in (2.17). Then $\Sigma$ is parabolic uniformly rectifiable, parabolic UR for short, with UR constants $(M, \Gamma)$ if
\begin{equation}
\|\nu\| := \sup_{(X, t) \in \Sigma, \rho > 0} \rho^{-d} \nu(\Delta(X, t, \rho) \times (0, \rho)) \leq \Gamma. \tag{2.20}
\end{equation}
We will also need a parabolic dyadic decomposition of $\Sigma$. This is a special case of a more general result in [Chr].

**Lemma 2.21.** Assume that $\Sigma \subset \mathbb{R}^{n+1}$ is parabolic ADR in the sense of Definition 2.14 with constant $M$. Then $\Sigma$ admits a parabolic dyadic decomposition in the sense that there exist constants $\alpha > 0$, $\beta > 0$ and $c_\ast < \infty$, such that for each $k \in \mathbb{Z}$ there exists a collection of Borel sets, $\mathcal{D}_k$, which we will call (dyadic) cubes, such that

$$\mathcal{D}_k := \{ Q_k^j \subset \Sigma : j \in \mathcal{S}_k \},$$

where $\mathcal{S}_k$ denotes some index set depending on $k$, satisfying

(i) $\Sigma = \bigcup_j Q_k^j$ for each $k \in \mathbb{Z}$.

(ii) If $m \geq k$ then either $Q_m^m \subset Q_k^j$ or $Q_m^m \cap Q_k^j = \emptyset$.

(iii) For each $(j, k)$ and each $m < k$, there is a unique $i$ such that $Q_k^i \subset Q_m^m$.

(iv) $\operatorname{diam}(Q_k^j) \leq c_* 2^{-k}$.

(v) Each $Q_k^j$ contains $\Sigma \cap C_{\alpha_2^2}(Z_k^j, t_k^j)$ for some $(Z_k^j, t_k^j) \in \Sigma$.

(vi) $\sigma((Z, t) \in Q_k^j : d_p(Z, t, \Sigma \cap Q_k^j) \leq \rho 2^{-k}) \leq c_* \rho^\beta \sigma(Q_k^j)$, for all $k, j$ and for all $\rho \in (0, \alpha)$.

**Remark 2.22.** We denote by $\mathcal{D} = \mathcal{D}(\Sigma)$ the collection of all $Q_k^j$, i.e.,

$$\mathcal{D} := \bigcup_k \mathcal{D}_k.$$

For a dyadic cube $Q \in \mathcal{D}_k$, we let $\ell(Q) := 2^{-k}$, and we will refer to this quantity as the size of $Q$. Evidently, $\ell(Q) \sim \operatorname{diam}(Q)$ with constants of comparison depending at most on $n$ and $M$. Note that (iv) and (v) of Lemma 2.21 imply, for each cube $Q \in \mathcal{D}_k$, that there is a point $(X_Q, t_Q) = (X_Q, t_Q) \in \Sigma$, and a cube $C_r(X_Q, t_Q)$, such that

$$r \approx 2^{-k} \approx \operatorname{diam}(Q), \quad \text{and} \quad \Sigma \cap C_r(X_Q, t_Q) \subset Q \subset \Sigma \cap C_{c r}(X_Q, t_Q),$$

for some uniform constant $c$. We will denote the associated surface ball by

$$\Delta_Q(r) := \Sigma \cap C_r(X_Q, t_Q),$$

and we shall refer to the point $(X_Q, t_Q)$ as the center of $Q$. Given a dyadic cube $Q \subset \Sigma$ and $\lambda > 1$, we define

$$\lambda Q := \Delta_Q(\lambda \operatorname{diam}(Q)).$$

**Definition 2.26.** [DS1] Let $S \subset \mathcal{D}(\Sigma)$. We say that $S$ is coherent if the following conditions hold.

(a) $S$ contains a unique maximal element $Q(S)$ which contains all other elements of $S$ as subsets.

(b) If $Q$ belongs to $S$, and if $Q \subset \tilde{Q} \subset Q(S)$, then $\tilde{Q} \in S$.

(c) Given a cube $Q \in S$, either all of its children belong to $S$, or none of them do.

We say that $S$ is semi-coherent if only conditions (a) and (b) hold.
3. **Statement of the main results**

As stated in the introduction, the purpose of this work is to prove (a) \( \implies \) (b) and (a) \( \implies \) (c) in Theorem 1.2. To prove (a) \( \implies \) (b) we show the following.

**Theorem 3.1.** Suppose that \( \Sigma \subset \mathbb{R}^{n+1} \) is parabolic uniformly rectifiable with constants \( (M, \Gamma) \). Let \( \delta \ll 1 \) and \( \kappa \gg 1 \) be two given positive constants. Then there exists a disjoint decomposition \( \mathbb{D}(\Sigma) = \mathbb{B} \cup \mathbb{G} \), satisfying the following properties.

1. The collection \( \mathcal{B} \) satisfies the Carleson packing condition:
   \[
   \sum_{Q' \subseteq Q, Q' \in \mathcal{B}} \sigma(Q') \leq c(n, M, \Gamma, \delta, \kappa) \sigma(Q), \quad \forall Q \in \mathbb{D}(\Sigma).
   \]

2. The collection \( \mathcal{G} \) is further subdivided into disjoint stopping time regimes \( \{S\}_{S \in \mathcal{F}} \), such that each such regime \( S \) is coherent.

3. The maximal cubes \( Q(S) \) for the stopping time regimes satisfy the Carleson packing condition:
   \[
   \sum_{S : Q(S) \subseteq Q} \sigma(Q(S)) \leq c(n, M, \Gamma, \delta, \kappa) \sigma(Q), \quad \forall Q \in \mathbb{D}(\Sigma).
   \]

4. For each \( S \), there exists a coordinate system \(^7\) and a regular Lip(1,1/2) function \( \psi_S := \psi = \psi(x, t) : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) with parameters \( b_1 = c(n, M) \cdot \delta \) and \( b_2 = b_2(n, M, \Gamma) \), such if we define \( \Sigma_S := \Sigma_\psi := \{(x, \psi(x, t), t) : (x, t) \in \mathbb{R}^{n-1} \times \mathbb{R}\} \), then
   \[
   \sup_{(x, t) \in Q} d_\psi(X, t, \Sigma_S) \leq \delta \operatorname{diam}(Q),
   \]
   for all \( Q \in \mathcal{S} \).

The second theorem we prove is a strengthening of Theorem 3.1 in the sense that it gives instead a bilateral approximation in (4) and proves (a) \( \implies \) (c) in Theorem 1.2.

**Theorem 3.3.** Suppose that \( \Sigma \subset \mathbb{R}^{n+1} \) is parabolic uniformly rectifiable with constants \( (M, \Gamma) \). Let \( \delta \ll 1 \) and \( \kappa \gg 1 \) be two given positive constants. Then there exists a disjoint decomposition \( \mathbb{D}(\Sigma) = \mathcal{B}^* \cup \mathcal{G}^* \), satisfying the following properties.

1. The collection \( \mathcal{B}^* \) satisfies the Carleson packing condition:
   \[
   \sum_{Q' \subseteq Q, Q' \in \mathcal{B}^*} \sigma(Q') \leq c(n, M, \Gamma, \delta, \kappa) \sigma(Q), \quad \forall Q \in \mathbb{D}(\Sigma).
   \]

2. The collection \( \mathcal{G}^* \) is further subdivided into disjoint stopping time regimes \( \{S\}_{S \in \mathcal{F}^*}, \) such that each such regime \( S \) is coherent.

3. The maximal cubes \( Q(S^*) \) for the stopping time regimes satisfy the Carleson packing condition:
   \[
   \sum_{S' : Q(S^*) \subseteq Q} \sigma(Q(S^*)) \leq c(n, M, \Gamma, \delta, \kappa) \sigma(Q), \quad \forall Q \in \mathbb{D}(\Sigma).
   \]

---

\(^7\)This means that we identify \( P_i \times P_i^* \times \mathbb{R} \) with \( \mathbb{R}^n \times \mathbb{R} \) for some \( i \)-independent plane \( P \in \mathcal{P} \), see Definition 2.9.
3.1. The proof of Theorem 3.1. We here briefly discuss the outline of the proof of Theorem 3.1 to be presented in subsequent sections.

In the sequel, \( \Sigma \subset \mathbb{R}^{n+1} \) is parabolic uniformly rectifiable with constants \((M, \Gamma)\). Throughout the proof, four parameters will consistently appear: \(K, \epsilon, \delta\) and \(\kappa\). The parameters \(\delta\) and \(\kappa\) are exactly as in the statement of Theorem 3.1. \(K\) and \(\epsilon\) are auxiliary parameters. Both \(\epsilon\) and \(\delta\) will be small, \(\epsilon \ll \delta \ll 1\), and \(\epsilon\) is chosen/defined in (6.46). This implies that \(\epsilon = \epsilon(n, M, \Gamma, K, \delta)\). \(K\) and \(\kappa\) will be large, \(K \gg \kappa \gg 1\), and to achieve that the constants in Theorem 3.1 only depend on \(n, M, \Gamma, \delta\) and \(\kappa\), we will enforce the relation \(K/\kappa = c(n, M) \gg 1\).

For \(K \gg \kappa \gg 1\) fixed, we define, for each \(Q \in \mathcal{D}(\Sigma)\),

\[
\beta_{\infty}(Q) := \beta_{\infty}(Q, K) \quad \text{and} \quad \beta_2(Q) := \beta_2(Q, K),
\]

where

\[
\beta_{\infty}(Q, K) := \inf_{P \in \mathcal{P}} \left( \operatorname{diam}(Q)^{-1} \sup_{|Y, s| \in 8KQ} d_\rho(Y, s, P) \right),
\]

and

\[
\beta_2(Q, K) := \inf_{P \in \mathcal{P}} \left( \operatorname{diam}(Q)^{-d} \int_{8KQ} \left( \frac{d_\rho(Y, s, P)}{\operatorname{diam}(Q)} \right)^2 d\sigma(Y, s) \right)^{1/2}.
\]

In the following we drop \(K\) to ease the notation. Note that

\[
\beta_{\infty}(Q)^{n+3} \leq \beta_2^2(Q^*),
\]

where \(Q^*\) is an appropriately chosen ancestor to \(Q\) such that

\[100c \operatorname{diam}(Q) \leq \operatorname{diam}(Q^*) \leq \operatorname{diam}(Q)\]

for \(c = c(n, M) \gg 2\) to be chosen. Indeed, let \(\tilde{P}\) be a plane realizing the infimum in the definition of \(\beta_2(Q^*)\). Clearly there exists \((Z_1, \tau_1) \in 8K\tilde{Q}\) with

\[
\rho := \operatorname{diam}(Q) \beta_{\infty}(Q) \leq d_\rho(Z_1, \tau_1, \tilde{P}).
\]

We have \(\sigma(\Delta(Z_1, \tau_1, \rho/8)) \geq \rho^d\) as \(\Sigma\) is parabolic ADR. Moreover every point in \(\Delta(Z_1, \tau_1, \rho/8)\) is contained in \(8cKQ\), for some \(c = c(n, M) \gg 2\), and lies at least \(\rho/2\) from \(\tilde{P}\). Thus, (3.6) is valid.

Using that \(v\) is a Carleson measure, a standard argument (see e.g. [R, Lemma 4.8]) shows that

\[
\sum_{Q < Q^*} \beta_2^2(Q) \sigma(Q) \leq \sigma(Q^*), \quad Q^* \in \mathcal{D}(\Sigma),
\]

with implicit constant depending on \(K, n, M\) and \(\Gamma\).
To this end, we define
\[ B := B(\epsilon, K) := \{ Q \in \mathcal{D}(\Sigma) : \beta_\infty(Q) \geq \epsilon \}, \]
\[ G := G(\epsilon, K) = \mathcal{D}(\Sigma) \setminus B(\epsilon, K). \]

Then (3.6) and (3.7) imply that
\[
\sum_{Q \in \tilde{Q}, Q \in B} \sigma(Q) \leq \sigma(\tilde{Q}), \quad \tilde{Q} \in \mathcal{D}(\Sigma),
\]
where the implicit constant now depends on \( \epsilon, K, n, M \) and \( \Gamma \).

For \( Q \in \mathcal{D}(\Sigma) \) we let \( P_Q \in \mathcal{P} \) be a \( t \)-independent plane which achieves the infimum in the definition of \( \beta_\infty(Q) \). Then, by the definition of \( G \),
\[
d_p(Y, s, P_Q) \leq \epsilon \operatorname{diam}(Q) \quad \forall (Y, s) \in 8KQ, \quad Q \in G.
\]

Following [DS1, Section 7] (or equivalently, [DS2, Lemma IV.2.35]) we may construct an augmented collection \( B' \supset B \), and a refined collection \( G' \subset G \), for which we have a disjoint decomposition of the dyadic cubes \( \mathcal{D}(\Sigma) = B' \cup G' \). Moreover, \( B' \) still satisfies the packing condition (3.8), and \( G' \) may be partitioned into a collection of coherent stopping time regimes \( \{ S \}_{S \in \mathcal{F}} \), with maximal cubes \( \{ Q(S) \}_{S \in \mathcal{F}} \), with the following properties:

(A) If \( Q \in S \), then \( \angle(P_Q, P_{Q(S)}) \leq \delta \).

(B) If \( Q \) is a minimal cube of \( S \), then either a child of \( Q \) is in \( B' \)
\[
\text{or } \angle(P_Q, P_{Q(S)}) \geq \delta/2.
\]

Given a stopping time regime \( S \), we define \( m(S) \) to be the collection of minimal cubes in \( S \).

Following [DS1], we will construct a \( \text{Lip}(1,1/2) \) graph associated to each stopping time regime. After doing so, proving Theorem 3.1 is a matter of demonstrating that the graph is, in fact, a \( \text{regular} \) \( \text{Lip}(1,1/2) \) graph and that the maximal cubes \( \{ Q(S) \}_{S \in \mathcal{F}} \) pack. These objectives will be achieved in a somewhat simultaneous manner in a way similar to [DS1]. The insight (and courage) of David and Semmes in providing such a proof is quite remarkable.

The proof of Theorem 3.1 is divided up into three sections: in Section 4, we construct a \( \text{Lip}(1,1/2) \) graph with small constant for each stopping time regime. In Section 5, we push the geometric square function estimate (the Carleson measure estimate for \( \nu \)) to the graph. The control of the geometric square function on the graph will verify, by [HLN1, HLN2], that the graph is indeed regular. In Section 6, we use the fact that we pushed the geometric square function in a sharp manner to pack the stopping time regimes.

Our argument will also rely on two simple lemmas concerning the approximation of surface cubes by planes, these lemmas can be found in Appendix A. The construction of the graph relies on condition (3.10) (A) and that the geometric lemma holds for \( \Sigma \). It is also worth making the remark that we only use condition (3.10) (B) above when packing the stopping time regimes in Section 6.
4. Construction of $\text{Lip}(1,1/2)$ graphs for stopping time regimes

In this section we shall follow very closely the arguments in [DS1, Chapter 8] to construct, for $S \in \mathcal{F}$ fixed, a $\text{Lip}(1,1/2)$ graph which approximates $Q \in S$ as in Theorem 3.1(4). We can without loss of generality assume that

$$P_{Q(S)} := \{(x, n, t) \in \mathbb{R}^{n+1} \times \mathbb{R} \times \mathbb{R} : x_n = 0\}.$$

We will often identify $P_{Q(S)}$ with $\mathbb{R}^n$ and employ the notation $C'_\delta(x, t)$ introduced in (2.2) for a $n$-dimensional parabolic cube in $P_{Q(S)}$. We let $\pi$ denote the projection onto $P_{Q(S)}$, that is, $\pi(x, n, t) = (x, t) \in \mathbb{R}^n$. We let $\pi^+$ denote the projection onto the normal to $P_{Q(S)}$, that is, $\pi^+(x, n, t) = n$. Given $Q \in \mathcal{D}(\Sigma)$ we let $(x_Q, t_Q) := \pi(X_Q, t_Q)$ denote the projected center of $Q$.

Given $S \in \mathcal{F}$, and following [DS1], we introduce distance functions adapted to the stopping regime. Indeed, if $(x, t) \in \mathbb{R}^{n+1}$ then we define

$$d(x, t) = d_S(x, t) := \inf_{Q \in S} \left[ d_p(X, t, Q) + \text{diam}(Q) \right],$$

and we let $D = D_S: \mathbb{R}^n \to [0, \infty)$ be defined as

$$D(x, t) := \inf_{(x, t) \in \pi^+(x, t)} d(x, t) = \inf_{Q \in S} \left[ d_p(x, t, \pi(Q)) + \text{diam}(Q) \right].$$

In the sequel, we set $F := \{(x, t) \in \Sigma : d(x, t) = 0\}$.

Lemma 4.2. Let $S \in \mathcal{F}$ and let $P_{Q(S)} = \mathbb{R}^n$. Then $\pi$ is one-to-one on $F$ and we can define a function $\hat{\psi}$ on $\pi(F) \subset \mathbb{R}^n$ as follows. Given $(x, t) = (x, n, t) \in F$, we set $\hat{\psi}(x, t) := \pi^+(x, t) = x_n$. Then $\hat{\psi}$ is a well-defined $\text{Lip}(1,1/2)$ function with constant bounded by $2\delta$ on $\pi(F)$, i.e.,

$$|\hat{\psi}(x, t) - \hat{\psi}(y, s)| \leq 2\delta d_p(x, t, y, s),$$

whenever $(x, t), (y, s) \in \pi(F)$. More generally, suppose that $(x, t), (y, s) \in 16\kappa Q(S)$, and that for some $N \geq 1$, we have $\min\{d(x, t), d(y, s)\} \leq N d_p(X, t, Y, s)$; then

$$|\pi^+(x, t) - \pi^+(y, s)| \leq 2\delta d_p(\pi(x, t), \pi(y, s)),$$

provided that $\epsilon \ll \delta$ is small enough, depending on $N$, and $K \gg 1$ large enough, depending on $N$ and $\kappa$. (Remark: in the sequel, we apply the lemma with $N$ equal to an absolute constant.)

Proof. Clearly, it suffices to prove (4.4). The proof follows that of [DS1, Lemma 8.4]. Let $(x, t), (y, s) \in 16\kappa Q(S)$ and without loss of generality, suppose that $d(x, t) \leq N d_p(X, t, Y, s)$. Let $Q \in S$ be such that

$$d_p(X, t, Q) + \text{diam}(Q) \leq 2N d_p(X, t, Y, s).$$

We can, if necessary, replace $Q$ by one of its ancestors (which is still in $S$) to obtain $\kappa^{-1}d_p(X, t, Y, s) \leq \text{diam} Q \leq N d_p(X, t, Y, s)$. By construction, $Q \in \mathcal{G}$ and hence $\beta_{\infty}(Q) \leq \epsilon$, so for $K = K(N, \kappa)$ large enough, there is a plane $P_Q$ such that

$$d_p(X, t, P_Q) + d_p(Y, s, P_Q) \leq \epsilon d_Q \leq \epsilon N d_p(X, t, Y, s) \ll \delta d_p(X, t, Y, s),$$

as $\epsilon \ll \delta$. The conclusion in (4.4) now follows since, by construction of the stopping time, the angle between $P_{Q(S)}$ and $P_Q$ is at most $\delta$. In particular, the reader
can verify this claim using elementary geometry and the fact that $\sin^2(\delta) \approx \delta^2$ for $\delta$ small. 

With Lemma 4.2 as the starting point we next construct a function $\psi$ on $\mathbb{R}^n = P_{Q(S)}$ which coincides with $\hat{\psi}$ on $\pi(F)$. We will extend $\hat{\psi}$ off $\pi(F)$ using an appropriate Whitney type extension following the corresponding construction in [DS1].

For $(x, t) \in \mathbb{R}^n$ and $(x, t)$ not on the boundary of any dyadic cube, let $I_{(x,t)}$ be the largest (closed) dyadic cube in $\mathbb{R}^n$ containing $(x, t)$ and satisfying

$$\text{diam}(I_{(x,t)}) \leq \frac{1}{20} \inf_{(z, \tau) \in I_{(x,t)}} D(z, \tau).$$

Let $\{I_i\}$ be a labelling of the set of all these cubes $I_{(x,t)}$ without repetition. By construction the collection $\{I_i\}$ consists of pairwise non-overlapping closed dyadic cubes in $\mathbb{R}^n$. These cubes cover $\mathbb{R}^n \setminus \pi(F)$ and they do not intersect $\pi(F)$. For the construction we need the following, whose proof follows that of [DS1, Lemma 8.7], essentially verbatim.

**Lemma 4.6.** Given $I_i$ we have

$$10 \text{ diam } I_i \leq D(y, s) \leq 60 \text{ diam } I_i \quad \forall (y, s) \in 10I_i. \tag{4.7}$$

In particular, if $10I_i \cap 10I_j \neq \emptyset$, then

$$\text{diam } I_i \approx \text{ diam } I_j. \tag{4.8}$$

**Proof.** Fix $i$ and let $(y, s) \in I_i$. As $D(\cdot, \cdot)$ is $\text{Lip}(1, 1/2)$ with norm 1, we have

$$D(y, s) \geq \inf_{(x, t) \in I_i} D(x, t) - 10 \text{ diam } I_i \geq 10 \text{ diam } I_i,$$

where we used that $\text{diam } I_i \leq 20^{-1} \inf_{(x, t) \in I_i} D(x, t)$. On the other hand, if $I$ is the dyadic parent of $I_i$, then there exists $(z, \tau) \in I$ such that $D(z, \tau) < 20 \text{ diam } I = 40 \text{ diam } I_i$. Thus,

$$D(y, s) \leq D(z, \tau) + 20 \text{ diam } I_i \leq 60 \text{ diam } I_i.$$

This proves (4.7), and hence also (4.8). \qed

We continue to follow the constructions in [DS1, Chapter 8]. Set

$$R = R_S := \text{diam}(Q(S)). \tag{4.9}$$

Given the center $(X_{Q(S)}, t_{Q(S)})$ of $Q(S)$, we recall that

$$(x_{Q(S)}, t_{Q(S)}) = \pi(X_{Q(S)}, t_{Q(S)}),$$

and we introduce the index set

$$\Lambda = \{i : I_i \cap C_{2R}^+(x_{Q(S)}, t_{Q(S)}) \neq \emptyset\}. \tag{4.10}$$

We claim, for each $i \in \Lambda$, that there exists $Q(i) \in S$ such that

$$\kappa^{-1} \text{ diam } I_i \leq \text{ diam } Q(i) \leq 120 \text{ diam } I_i, \quad d_p(\pi(Q(i)), I_i) \leq 120 \text{ diam } I_i.\tag{4.11}$$

To see this, note that given $I_i$ and $(x, t) \in I_i$, there exists a cube $Q \in S$ such that $d_p(x, t, \pi(Q)) + \text{diam}(Q) \leq 2D(x, t) - \text{diam}(I_i)$. Moreover, $D(x, t) \leq c \text{ diam}(Q(S))$, for a constant $c \geq 1$ depending only on $n, M$, and

$$d_p(x, t, \pi(Q^n)) \leq 2D(x, t) \leq 120 \text{ diam } I_i,$$
for all \( Q \subseteq Q^* \subseteq Q(S) \). Therefore we can choose \( Q(i) \) so that \( Q \subseteq Q(i) \subseteq Q(S) \) and

\[
(ck)^{-1} D(x, t) \leq \text{diam}(Q(i)) \leq 2D(x, t) \leq 120 \text{diam } I_i,
\]

provided \( \kappa \) is chosen large enough.

In the sequel, we will, at instances, use the notation

\[
r_i := \text{diam } I_i, \quad \text{so that } r_i/\kappa \leq \text{diam } Q(i) \leq r_i,
\]

by (4.11).

Let \( \{\tilde{\nu}_i\} \) be a class of infinitely differentiable functions on \( \mathbb{R}^n \) such that

\[(i) \quad \tilde{\nu}_i \equiv 1 \text{ on } 2I_i \text{ and } \tilde{\nu}_i \equiv 0 \text{ in } \mathbb{R}^n \setminus 3I_i \text{ for all } i,
\]

\[(ii) \quad r_i^l \left| \frac{\partial^j}{\partial x^l} \tilde{\nu}_i \right| + r_i^j \left| \frac{\partial^j}{\partial t^l} \tilde{\nu}_i \right| \leq c(l, n) \text{ for } l = 1, 2, \ldots .
\]

In (ii), \( \frac{\partial^j}{\partial x^l} \) denotes an arbitrary partial derivative with respect to the space variable \( x \) and of order \( l \). We also introduce \( \{\tilde{\nu}_i\} \) defined as a partition of unity adapted to \( \{2I_i\} \), i.e.,

\[
\nu_i(x, t) = \tilde{\nu}_i(x, t) (\sum_j \tilde{\nu}_j(x, t))^{-1}.
\]

Then \( \{\nu_i\} \) is also a class of infinitely differentiable functions on \( \mathbb{R}^n \) and

\[
(4.14) \quad r_i^l \left| \frac{\partial^j}{\partial x^l} \nu_i \right| + r_i^j \left| \frac{\partial^j}{\partial t^l} \nu_i \right| \leq c(l, n) \text{ for } l = 1, 2, \ldots .
\]

We now ready to construct the extension of \( \hat{\psi} \) off \( \pi(F) \). If \( i \in \Lambda \), then for \( I_i \) we have an associated dyadic cube \( Q(i) \in S \) as in (4.11), we first express the associated hyperplane \( P_{Q(i)} \) as

\[
(4.15) \quad P_{Q(i)} = \{(x, B_i(x, t), t) : (x, t) \in \mathbb{R}^n\},
\]

i.e., \( B_i : \mathbb{R}^n \to \mathbb{R} \) is the affine function whose graph is \( P_{Q(i)} \). If \( i \notin \Lambda \), set \( B_i = 0 \) and we note, in this case, \( B_i \) is the affine function whose graph is \( P_{Q(S)} \). Using this notation we let

\[
(4.16) \quad \psi(x, t) = \hat{\psi}(x, t) \text{ when } (x, t) \in \pi(F),
\]

\[
\psi(x, t) = \sum_i B_i(x, t) \nu_i(x, t) \text{ when } (x, t) \in \mathbb{R}^n \setminus \pi(F).
\]

This defines \( \psi \) on \( \mathbb{R}^n \), with \( \psi = \hat{\psi} \) on \( \pi(F) \) and by construction

\[
(4.17) \quad \psi \equiv 0 \text{ on } \mathbb{R}^n \setminus C_{4\kappa}(x_{Q(S)}, t_{Q(S)}).
\]

Based on the construction we can prove the following lemma.

**Lemma 4.18.** For \( (y, s), (z, \tau) \in \mathbb{R}^n \) we have

(i) \( |B_i(y, s) - B_j(z, \tau)| \leq \delta d_p(y, s, z, \tau) \) for all \( i \),

(ii) \( |B_i(y, s) - B_j(y, s)| \leq \varepsilon \min(r_i, r_j), \text{ if } (y, s) \in 100I_i \cup 100I_j, \text{ if } 10I_i \cap 10I_j \neq \emptyset. \)
Proof. By construction of the stopping regime, the Lip(\(1/2\)) norm of \(B_t\) is bounded by \(2\delta\), so \((i)\) is trivial. To prove \((ii)\), we follow the proof of [DS1, Lemma 8.17]. Let \(I_i\) and \(I_j\) be two cubes such that \(10I_i \cap 10I_j \neq \emptyset\), thus \(\diam I_i = r_i \approx r_j = \diam I_j\) by Lemma 4.6. We claim that if \(i \in \Lambda\) or \(j \in \Lambda\), then
\begin{equation}
(4.19) \quad d_p(Q(i), Q(j)) \leq r_i \approx r_j.
\end{equation}
Let us prove the claim. Let \(i \in \Lambda\). Let \((X, t) \in Q(i)\) and \((Y, s) \in Q(j)\). We may assume that \(d_p(X, t, Y, s) \geq \diam Q(i)\), the other case being trivial. Since \(d(X, t) \leq \diam Q(i)\), by (4.4) we see that
\begin{equation}
(4.20) \quad |\pi^+(X, t) - \pi^+(Y, s)| \leq d_p(\pi(X, t), \pi(Y, s)).
\end{equation}
By (4.11), \(d_p(\pi(X, t), \pi(Y, s)) \leq r_i \approx r_j\), so the claim (4.19) holds.

Suppose that \(i \in \Lambda\). Then by (4.19) and (4.11), there is a “surface ball” \(\Delta_r\), centered in \(Q(i)\), such that \(r \approx r_i \approx r_j\), and
\[Q(i) \cup Q(j) \subset \Delta_r \subset \Delta_{2r} \subset KQ(i) \cap KQ(j),\]
for \(K\) large enough. Using the second part of Lemma A.2 (with \(\Delta = \Delta_r\), \(P_1 = P_{Q(i)}\), \(P_2 = P_{Q(j)}\)), we see that if \((y, s) \in 100I_i \cup 100I_j\), with \(10I_i \cap 10I_j \neq \emptyset\), then
\begin{equation}
(4.21) \quad |B_t(y, s) - B_j(y, s)| \leq \epsilon r.
\end{equation}
The case \(j \in \Lambda\) is exactly the same, so this completes the proof in the case that at least one of \(i\) or \(j\) belongs to \(\Lambda\).

If, on the other hand, \(i \notin \Lambda\) and \(j \notin \Lambda\), then \(Q(i) = Q(j) = Q(S)\), and hence \(B_i \equiv B_j\), so the estimate holds trivially. This completes the proof of Lemma 4.18. □

Before proving that \(\psi\) is Lip(\(1/2\)), we present one more preliminary lemma, which is the parabolic version of [DS1, Lemma 8.21].

Lemma 4.22. Let
\begin{equation}
(4.23) \quad (x, t) \in C'_4(-\kappa, t_{Q(S)}), \quad \kappa \geq 0 \text{ with } D(x, t) \leq r < 2\kappa R,
\end{equation}
and let \(Q \in S\) be such that \(d_p(x, t, \pi(Q)) \leq \lambda r\) and \(\lambda^{-1}r \leq \diam Q \leq \lambda r\) for some \(\lambda \geq 1\). Then
\begin{equation}
(4.24) \quad \pi^{-1}\left(C'_4(x, t)\right) \cap 4\kappa Q(S) \text{ is contained in } \lambda Q,
\end{equation}
where \(\lambda\) depends on \(\lambda\) and \(\kappa\).

Furthermore, for some \(\hat{\lambda}\) depending on \(\kappa\), we have
\begin{equation}
(4.25) \quad \hat{\lambda}^{-1}d(X, t) \leq D(\pi(X, t)) \leq d(X, t) \text{ for all } (X, t) \in 2\kappa Q(S).
\end{equation}

Proof. The proof follows that of [DS1, Lemma 8.21] essentially verbatim. To prove (4.24), we consider \((Y, s) \in Q, (Z, \tau) \in \pi^{-1}\left(C'_4(x, t)\right) \cap 4\kappa Q(S)\). First, if
\[d_p(Y, s, (Z, \tau)) \leq \diam Q,\]
then \((Z, \tau) \in \lambda Q\), for, say, \(\lambda \geq 5\). Second, if
\[d_p(Y, s, (Z, \tau)) > \diam Q,\]
then
\[d(Y, s) \leq \diam Q \leq d_p(Y, s, (Z, \tau)).\]
Hence, by Lemma 4.2

\[ |\pi^+(Y, s) - \pi^+(Z, \tau)| \leq d_p(\pi(Y, s), \pi(Z, \tau)) \leq \lambda \text{ diam } Q. \]

(The last inequality follows from our assumptions that \(d_p(X, t, \pi(Q)) \leq r \approx \text{ diam } Q\), and that \((Y, s) \in Q\), and \(\pi(Z, \tau) \in C'(x, t)\). This proves (4.24).

To deduce (4.25), we note that the second inequality is trivial by definition. To prove the first inequality, we consider first the case that \(\pi(X, t) \in \pi(F)\), thus, \(\pi(X, t) = \pi(X', t)\) for some \((X', t) \in F\), so that \(d(X', t) = 0\) by definition. But then by Lemma 4.2, \((X, t) = (X', t)\), so \(d(X, t) = 0\), and (4.25) holds in this case. Otherwise, if \(\pi(X, t) \notin \pi(F)\), we then set \((x, t) := \pi(X, t)\), and \(r := D(x, t) > 0\), and note that by definition of \(D(x, t)\), there is a \(Q \in S\) such that \(d_p(x, t, \pi(Q)) \leq r \approx \text{ diam } Q\) (here, there can be possible implicit dependence on \(\kappa\) in the case that \(R \ll r \ll \kappa R\)). Hence, we may apply (4.24) to deduce that \((X, t) \in \overline{\lambda Q}\), and therefore,

\[ d(X, t) \leq d_p(X, t, Q) + \text{ diam } Q \leq \text{ diam } Q \approx r = D(\pi(X, t)). \]

\[ \square \]

Lemma 4.26. Let \(\psi : \mathbb{R}^n \to \mathbb{R}\) be defined as in (4.16). Then \(\psi : \mathbb{R}^n \to \mathbb{R}\) is \(\text{Lip}(1, 1/2)\) with constant on the order of \(\delta\), i.e.

\[ |\psi(y, s) - \psi(z, \tau)| \leq \delta d_p(y, s, z, \tau), \quad \forall (y, s), (z, \tau) \in \mathbb{R}^n. \]

Proof. Again, we shall follow the corresponding arguments in [DS1, Chapter 8]. However, we point out that the proof in [DS1] is not fully explicated, and in particular, does not address the case that \(F = \emptyset\) (which, in contrast to the case of the classical Whitney extension theorem, can actually happen in the present setting). Fortunately, the methods of [DS1] require only a modest refinement to address this situation.

Recall that we use the notation \(r_i := \text{diam } I_i\). There are four cases to consider:

1. \((y, s), (z, \tau) \in \pi(F)\).
2. \((y, s) \in \mathbb{R}^n \setminus \pi(F)\), with \((y, s) \in 2I_j\), and \((z, \tau) \in \pi(F)\).
3. \((y, s), (z, \tau) \in \mathbb{R}^n \setminus \pi(F)\), with \((y, s) \in 2I_j\), \((z, \tau) \in 2I_k\), and 
   \[ \text{(WLOG) } r_k \leq r_j, \quad (z, \tau) \notin 10I_j. \]
4. \((y, s), (z, \tau) \in \mathbb{R}^n \setminus \pi(F)\), with \((y, s) \in 2I_j\), \((z, \tau) \in 2I_k\), and 
   \[ \text{(WLOG) } r_k \leq r_j, \quad (z, \tau) \notin 10I_j. \]

We make one more preliminary observation: by (4.17), we may suppose that \((y, s)\) and \((z, \tau)\) both lie in the closure of \(C' := C'_{\delta R}(x, Q(S), t, Q(S))\). Indeed, if both points lie outside of this closed cube, then (4.27) is trivial, and if one (say \((y, s)\)) lies outside, and the other inside, then we may replace \((y, s)\) by another point \((y', s') \in \partial C'\) such that \(d_p(z, \tau, y', s') \leq d_p(z, \tau, y, s)\), since for such points \((y', s')\) and \((y, s)\), we have \(\psi(y, s) = 0 = \psi(y', s')\).

Let us now discuss the various cases (1)-(4).

Case (1) follows immediately from Lemma 4.2, and the definition of \(\psi\) (see (4.16)).
Consider next case (3). In this case, \((y, s), (z, \tau) \in 10I_j\). By Lemma 4.6, (4.14) and Lemma 4.18, and the fact that \(\sum_i (v_i(y, s) - v_i(z, \tau)) = 0\),

\[
|\psi(y, s) - \psi(z, \tau)| \leq \left| \sum_i (B_i(y, s) - B_i(z, \tau))v_i(y, s) \right|
\]

\[
+ \left| \sum_i (B_i(z, \tau) - B_j(z, \tau))(v_i(y, s) - v_i(z, \tau)) \right|
\]

\[
\leq \delta d_p(y, s, z, \tau) + r_j^{-1} ed_p(y, s, z, \tau) r_j \leq \delta d_p(y, s, z, \tau),
\]

since each \(B_i\) has Lip\((1,1/2)\) norm at most \(2\delta\), and \(\epsilon \ll \delta\).

It remains to treat cases (2) and (4), which we shall do more or less simultaneously. We remark that if \(F\) is non-empty, one can reduce case (4) to case (2), as in the proof of the classical Whitney extension theorem. However, in the present setting, it may be that \(F\) is empty, so we shall treat case (4) directly. For the sake of specificity, let us do this first, as the proof in case (2) will be similar, but a bit simpler.

In case (4), we decompose

\[
|\psi(y, s) - \psi(z, \tau)| \leq a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7,
\]

where for some \((U, t) \in Q(j), (U', t') \in Q(k)\) to be specified more precisely below, setting \((u, t) := \pi(U, t), (u', t') := \pi(U', t')\), we define

\[
a_1 = |\psi(y, s) - B_j(y, s)|, \quad a_2 = |B_j(y, s) - B_j(u, t)|,
\]

\[
a_3 = |B_j(u, t) - \pi^+(U, t)|, \quad a_4 = |\pi^+(U, t) - \pi^+(U', t')|,
\]

\[
a_5 = |\pi^+(U', t') - B_k(u', t')|, \quad a_6 = |B_k(u', t') - B_k(z, \tau)|,
\]

\[
a_7 = |B_k(z, \tau) - \psi(z, \tau)|.
\]

We estimate these terms as follows. For future reference, we observe that since \((z, \tau) \notin 10I_j\), and \((y, s) \in 2I_j\), therefore, for a purely dimensional implicit constant,

\[
r_k \leq r_j \leq d_p(y, s, z, \tau).
\]

Consider for now any \((U, t) \in Q(j), \) and any \((U', t') \in Q(k)\). By definition of \(\psi\), since \(\sum_i v_i(y, s) = 1\), we have

\[
a_1 = \left| \sum_i v_i(y, s) (B_i(y, s) - B_j(y, s)) \right| \leq \epsilon r_j \leq \epsilon d_p(y, s, z, \tau),
\]

by Lemma 4.18 (ii) and (4.29). Similarly,

\[
a_7 \leq \epsilon d_p(y, s, z, \tau).
\]

Since each \(B_i\) has Lip\((1,1/2)\) norm at most \(2\delta\), by the second inequality in (4.11),

\[
a_2 \leq \delta d_p(y, s, u, t) \leq \delta r_j \leq \delta d_p(y, s, z, \tau),
\]

again using (4.29). Similarly,

\[
a_6 \leq \delta d_p(y, s, z, \tau).
\]
Since $P_{Q(j)}$ is the graph of $B_j$, which has small slope, and since $\beta_{ac}(Q(j)) \leq \epsilon$, we have
\begin{equation}
4.31 \quad a_2 \approx \text{dist}((U, t), P_{Q(j)}) \leq \epsilon \text{ diam } Q(j) \leq 120er_j \leq \epsilon d_p(y, s, z, \tau),
\end{equation}
where we have used (4.11) and then (4.29) in the last two steps. Similarly,
\begin{equation}
4.32 \quad a_3 \approx \epsilon d_p(y, s, z, \tau).
\end{equation}

It remains to deal with $a_4$. Since $(U, t) \in Q(j)$ and $(y, s) \in 2I_j$, we have
\begin{equation}
4.33 \quad d(U, t) \leq \text{diam } Q(j) \leq 120r_j, \quad d_p(u, t, y, s) \leq 300r_j,
\end{equation}
by (4.11). Similarly, since $(U', t') \in Q(k)$, and $(z, \tau) \in 2I_k$,
\begin{equation}
4.34 \quad d(U', t') \leq \text{diam } Q(k) \leq 120r_k \leq 120r_j, \quad d_p(u', t', z, \tau) \leq 300r_k \leq 300r_j.
\end{equation}

Consequently, by (4.32) and (4.29),
\begin{equation}
4.35 \quad d(U, t) \leq 120r_j \leq d_p(y, s, z, \tau).
\end{equation}

To treat $a_4$, we now consider two sub-cases:
- (Sub-case A) $d_p(y, s, z, \tau) \geq 1000r_j$.
- (Sub-case B) $d_p(y, s, z, \tau) < 1000r_j$.

We first consider the former. In this case, since $d_p(y, s, z, \tau) \geq 1000r_j$, the right hand inequalities in (4.32) and (4.33) imply that
\begin{equation}
4.36 \quad d_p(y, s, z, \tau) \approx d_p(u, t, u', t') \leq d_p(U, t, U', t'),
\end{equation}
Combining the latter estimate with (4.34), we obtain
\begin{equation}
4.37 \quad d(U, t) \leq d_p(U, t, U', t').
\end{equation}
Thus, we may apply Lemma 4.2, with $N$ equal to an absolute constant, and with $\epsilon \ll \delta$, to obtain
\begin{equation}
4.38 \quad a_4 \leq 2\delta d_p(u, t, u', t') \approx \delta d_p(y, s, z, \tau).
\end{equation}

On the other hand, consider now sub-case B, i.e., $d_p(y, s, z, \tau) < 1000r_j$. Thus
\begin{equation}
4.39 \quad d_p(y, s, z, \tau) \approx r_j,
\end{equation}
by (4.29). In the current scenario, we consider two further sub-cases.

(Sub-case B1): $Q(j)$ meets $Q(k)$. In this case we simply let $(U, t) = (U', t')$ be any point in $Q(j) \cap Q(k)$, so that, trivially, $a_4 = 0$.

(Sub-case B2): $Q(j)$ and $Q(k)$ are disjoint. In this case, we let $(U', t')$ be any point in $Q(k)$, and we select $(U, t)$ as follows. Let $C(j) := C_r(X_{Q(j), t_{Q(j)}})$, with $r \approx \text{diam } Q(j)$, be the cube defined in (2.23), with respect to $Q = Q(j)$. Since $\Sigma \cap C(j) \subset Q(j)$ by construction (see (2.23)), it follows that $(U', t') \notin C(j)$. We then set $(U, t) := (X_{Q(j), t_{Q(j)}})$ (the center of $C(j)$), and observe that in this case
\begin{equation}
4.40 \quad d_p(U, t, t', t') \geq \text{diam } Q(j),
\end{equation}
where the implicit constants here depend on those in (2.23).

Note that (4.32) and (4.33) continue to hold, since we still have $(U, t) \in Q(j)$, $(y, s) \in 2I_j$, $(U', t') \in Q(k)$, and $(z, \tau) \in 2I_k$. In particular,
\begin{equation}
4.41 \quad d(U, t) \leq \text{diam } Q(j), \quad d_p(u, t, y, s) + d_p(u', t', z, \tau) \leq r_j.
\end{equation}
By (4.36) and the second inequality in (4.38), we see that
\begin{equation}
\tag{4.39}
d_p(u, u', t') \leq r_j \approx d_p(y, s, z, \tau).
\end{equation}
Combining (4.37) and the first inequality in (4.38), we find that
\[ d(U, t) \leq \text{diam } Q(j) \leq d_p(U, t, U', t'). \]
Consequently, we may as above apply Lemma 4.2 to obtain
\[ a_4 \leq 2\delta d_p(u, t, u', t') \leq \delta d_p(y, s, z, \tau), \]
where in the last step we have used (4.39).

Summing our estimates for the terms $a_1$ through $a_7$, and using that $\epsilon \ll \delta$, we obtain the desired bound (4.27) in case (4), in general.

In case (2), we proceed in a similar fashion to case (4), but now matters are a bit simpler: we decompose
\[ |\psi(y, s) - \psi(z, \tau)| \leq a_1 + a_2 + a_3 + a_4, \]
with $a_1, a_2$ and $a_3$ exactly as above (and enjoying the very same bounds), and with
\[ a_4 = |\pi^+(U, t) - \psi(z, t)|, \]
where $(U, t)$ is an arbitrary point in $Q(j)$. Thus, again, (4.32) holds, and combining the latter with (4.29), we see as before that
\begin{equation}
\tag{4.40}
d_p(u, t, y, s) \leq d_p(y, s, z, \tau).
\end{equation}
Recall that in case (2), we assume $(z, \tau) \in \pi(F)$. Consequently, by (the first part of) Lemma 4.2, and the fact that $\psi = \hat{\psi}$ on $\pi(F)$, we have $d(z, \tau, \psi(z, \tau)) = 0$, and $\pi^+(z, \tau, \psi(z, \tau)) = \psi(z, \tau)$. Hence, by the second part of Lemma 4.2, and (4.40), we find that
\[ a_4 \leq 2\delta d_p(u, t, z, \tau) \leq \delta d_p(y, s, z, \tau) \]
We now sum the bounds for terms $a_1$ through $a_4$ to complete the proof of the lemma. \hfill \Box

**Lemma 4.41.** Let $(X, t) \in 2\kappa Q(S)$. Then
\begin{equation}
\tag{4.42}
d_p(X, t, (x, \psi(x, t), t)) \leq \epsilon d(X, t), \quad (x, t) := \pi(X, t).
\end{equation}
In particular, (3.2) holds provided $\epsilon \ll \delta / \kappa$.

**Proof.** Note that if $Q \in S$, then $d(X, t) \leq \kappa \text{diam}(Q)$ for all $(X, t) \in \kappa Q$. Using this observation, we see that (4.42) implies (3.2) provided $\epsilon \ll \delta / \kappa$.

We turn now to (4.42), following the proof of [DS1, Proposition 8.2]. If $(X, t) \in F$, there is nothing to prove, so we consider $(X, t) \in 2\kappa Q(S) \setminus F$. In particular, $d(X, t) > 0$. Set $(x, t) := \pi(X, t)$. Estimate (4.25) implies that $D(x, t) > 0$ and hence $(x, t) \in I_i$ for some $i$. Using (4.24) with $Q := Q(i)$ and $r := D(x, t) \approx \text{diam } I_i$, we see that $(X, t) \in \lambda Q(i)$, where by (4.11), the constant $\lambda$ in Lemma 4.22 depends on $\kappa$, and therefore $\lambda$ depends on $\kappa$. Since we assume that $K$ is large, depending on $\kappa$, we may in particular take $K$ much larger than $\lambda$, so that
\[ |\pi^+(X, t) - B_i(x, t)| \leq 2\epsilon \text{diam}(Q(i)) \leq \epsilon \text{diam } I_i \approx \epsilon D(x, t) \leq \epsilon d(X, t), \]
where we have used (3.9) with $Q = Q(i)$, along with the definition of $B_i$. Estimate (4.42) now follows, by Lemma 4.18 (ii) and the definition of $\psi$. \hfill \Box
Lemma 4.43.  

(4.44) \[ |\nabla^2_y \psi(y, s)| + |\partial_y \psi(y, s)| \leq \epsilon r_j^{-1} \text{ if } (y, s) \in 2I \]

Proof. We first prove the bound for \(|\nabla^2_y \psi(y, s)|\) by following [DS1]. By the definition of \(\Lambda\) (see (4.10)), and the adapted partition of unity \(\{\nu_i\}\), we have

\[
\sum_{i \in \Lambda} \nu_i(y, s) = 1, \quad \text{and} \quad \nabla \nu_i(y, s) = \nabla_x \left( \sum_{i \in \Lambda} \nu_i(y, s) \right) = 0,
\]

whenever \((y, s) \in 2I_j\). Let \(\alpha, \beta\) be two spatial indices and let \(\partial_\alpha, \partial_\beta, \text{ and } \partial_{\alpha \beta}\) denote the corresponding partial differential operators of order one and two. Using (4.45) we see that if \((y, s) \in 2I_j\), then

\[
\partial_{\alpha \beta} \psi(y, s) = \partial_{\alpha \beta} \left( \sum_i \nu_i(y, s) B_i(y, s) \right)
\]

\[
= \sum_i (\partial_{\alpha \beta} \nu_i(y, s)) B_i(y, s) + \sum_i (\partial_\beta \nu_i(y, s)) (\partial_\alpha B_i(y, s)) + \sum_i (\partial_\alpha \nu_i(y, s)) (\partial_\beta B_i(y, s))
\]

\[
= \sum_i (\partial_{\alpha \beta} \nu_i(y, s)) (B_i(y, s) - B_j(y, s)) + \sum_i (\partial_\alpha \nu_i(y, s)) (\partial_\beta B_i(y, s) - \partial_\beta B_j(y, s)) + \sum_i (\partial_\alpha \nu_i(y, s)) (\partial_\beta B_i(y, s) - \partial_\beta B_j(y, s)) =: I + II + III.
\]

The desired bound for term \(I\) follows immediately from Lemma 4.18 (ii) and (4.14), and the fact that the cubes \(3I_j\) (which contain the support of \(\nu_i\)) have bounded overlaps, by virtue of Lemma 4.6 (specifically (4.8)). A similar argument yields the desired bound for terms \(II\) and \(III\), once we observe that, if \(10I_j \cap 10I_j \neq \emptyset\), then \(|\nabla_i B_i - \nabla_j B_j| \leq \epsilon\). To see the latter, we translate and (spatially) rotate coordinates (only for purposes of the present argument) so that \(P_{(i, j)}\) is the hyperplane \(\{(x, x_n, t) : \ x_n = 0\}\), hence in these new coordinates \(B_j \equiv 0\). By Lemma A.2, the angle between \(P_{(i, j)}\) and \(P_{(i, j)}\) is at most \(C \epsilon\), i.e., in the new coordinates \(B_i\) is affine with \(\nabla_i B_i \sim \nabla_i B_j \leq C \epsilon\).

The bound for \(|\partial_y \psi(y, s)|\) can be produced similarly. In this case we first note that \(\partial_i B_i = 0\) (since \(P_{(i, j)} \in \mathcal{P}\)) and that

\[
0 = \partial_i \left( \sum_i \nu_i(y, s) \right) = \sum_i \partial_i \nu_i(y, s)
\]

for \((y, s) \in 2I_j\). Using this, Lemma 4.18 (ii), and (4.14), we deduce, for \((y, s) \in 2I_j\),

\[
|\partial_i \psi(y, s)| = |\partial_i \sum_i \nu_i(y, s) B_i(y, s)| = |\sum_i \partial_i \nu_i(y, s) B_i(y, s)|
\]

\[
= \sum_i |\partial_i \nu_i(y, s) (B_i(y, s) - B_j(y, s))| \leq \sum_i |\partial_i \nu_i| \epsilon r_j \leq \epsilon r_j^{-1}.
\]

Here we have also used \(\text{diam}(I_i) \sim \text{diam}(I_j)\) whenever \(10I_i \cap 10I_j \neq \emptyset\). \(\square\)
5. Pushing the geometric square function to the graph

In the previous section we proved that there exists, for arbitrary but fixed \( S \in \mathcal{F} \), a coordinate system and a \( \text{Lip}(1,1/2) \) function \( \psi_S := \psi = \phi(x,t) : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) with parameter \( b_1 \leq \delta \), such that if we define
\[
\Sigma_S := \Sigma_\phi := \{(x,\psi(x,t),t) : (x,t) \in \mathbb{R}^{n-1} \times \mathbb{R}\},
\]
then
\[
(5.1) \quad \sup_{(X,t) \in 2\sigma Q} d_p(X,t,\Sigma_S) \leq C_\ell(Q), \quad \forall Q \in \mathcal{S}.
\]
(The latter bound \((5.1)\) is proved in Lemma 4.41.) Recall that we have already fixed some of the relations amongst the parameters \( K, \epsilon, \delta \) and \( \kappa \): e.g., we have specified that \( \epsilon \ll \delta \), depending on \( \kappa \) (see Lemma 4.41), and that \( K \) is very large, also depending on \( \kappa \) (see Lemma 4.2 and the proof of Lemma 4.41). A little more precisely, we have \( 0 < C(\kappa) \epsilon \ll \delta \ll 1 \), for some large enough constant \( C(\kappa) \), and \( K = c(\kappa,n,M) \gg 1 \). In addition to these existing relations, in the final part of the argument, see \((6.46)\), \( \epsilon \) will be chosen to depend further upon \( n, K \), the ADR constant \( M \), and the parabolic UR constant \( \Gamma \) (see Definition 2.19), as well as on \( \kappa \) and \( \delta \).

To complete the proof of Theorem 3.1 we have to prove two things. First, we have to prove that the constructed function \( \psi_S := \psi = \phi(x,t) : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R} \) in fact is a regular \( \text{Lip}(1,1/2) \) function with parameters \( b_1 \leq \delta \) and \( b_2 = b_2(n,M,\Gamma) \). Second, we have to prove that the maximal cubes \( Q(S) \) for the stopping time regimes satisfy the Carleson packing condition
\[
(5.2) \quad \sum_{S:Q(S) \preceq Q} \sigma(Q(S)) \leq C(n,M,\Gamma,\delta,\kappa) \sigma(Q), \quad \forall Q \in \mathcal{D}(\Sigma).
\]
The key to both of these arguments is to make use of the geometric square function which is part of the definition of parabolic uniform rectifiability. In this section we prove how this information can be pushed to the constructed graph.

Recall that \( \Sigma \) is equipped with the geometric square function \( \gamma \) and the Carleson measure \( \nu \). We let \( \bar{\Sigma} \) be the graph of \( \psi \) and we denote the corresponding geometric square function and Carleson measure by \( \bar{\gamma} \) and \( \bar{\nu} \), respectively. In addition, for \( r > 0, (z,\tau) \in \mathbb{R}^n \) we introduce
\[
(5.3) \quad \bar{\gamma}(z,\tau,r) := \left( \inf_{L} \frac{1}{|C_{(z,\tau)}|} \left( \frac{|\psi(y,s) - L(y)|}{r} \right)^{2} \right)^{1/2},
\]
where the infimum is with respect to all linear functions of \( y \) (only). Using the regularity of the \( \psi \) constructed, we see that\(^8\)
\[
(5.4) \quad d\sigma_{\bar{\Sigma}} \approx d(\mu_{\bar{\Gamma}})(Y,s) = \sqrt{1 + |\nabla_{y}\psi(y,s)|^{2}} \, dy \, ds \sim dH^{n-1}(y) \, ds,
\]
for all \( (Y,s) \in \bar{\Sigma} \), and we recall that \( \mu \) is the slice-wise measure (see \((2.6)\)). The reader can easily verify that the slice-wise measure on the graph is parabolic ADR, and hence after an application of Proposition B.2 that the implicit constants in the

\(^8\)See Remark 2.8.
first comparability in display 5.4 above depend only on constants which depend on dimension and . Combining (5.4) with (4.27) and (4.44) it follows that
\begin{equation}
\hat{\gamma}(z, \tau, r) \sim \tilde{\gamma}(Z, \tau, r) \text{ for all } (Z, \tau) = (z, \psi(z, \tau), \tau) \in \widetilde{\Sigma}, \ r > 0.
\end{equation}

The essence of the section is to prove that we can control integrals of \(\hat{\gamma}(z, \tau, r)\), and hence integrals of \(\gamma(Z, \tau, r)\), with corresponding quantities involving \(\gamma\), with an error controlled by \(\epsilon^2\). In particular, the most important lemma in this section is the following.

**Lemma 5.6.** Fix \(S \in \mathcal{F}\). Consider \((\hat{z}, \hat{\tau}) \in \mathbb{R}^n, \rho > 0\), and let
\begin{equation}
\hat{\nu}(\hat{z}, \hat{\tau}, \rho) := \int_0^\rho \int \int_{C(\hat{z}, \hat{\tau})} (\hat{\gamma}(z, \tau, r))^2 \frac{dz \, d\tau \, dr}{r}.
\end{equation}
Assume further that
\begin{equation}
C_{2\rho}(\hat{z}, \hat{\tau}) \subseteq C_{2\rho}(\chi_{Q(S)}, \tau_{Q(S)}), \quad \rho \leq \kappa R/10.
\end{equation}
Then
\begin{equation}
\hat{\nu}(\hat{z}, \hat{\tau}, \rho) \leq \epsilon^2 \rho^d + \int \int_{E(\hat{z}, \hat{\tau}, \rho)} (\gamma(Z, \tau, Kr))^2 \frac{d\sigma(Z, \tau) \, dr}{r}
\end{equation}
where \(E(\hat{z}, \hat{\tau}, \rho)\) is the set of all \((Z, \tau, r) \in [\Sigma \cap \pi^{-1}(C_{K\rho}(\hat{z}, \hat{\tau}))] \times (0, \infty)\) such that \(K^{-1} d(Z, \tau) \leq r \leq Kp\).

From the previous lemma we shall be able to deduce the following lemma which will yield the desired regularity of the graph (see Lemma 5.40).

**Lemma 5.10.** Let \((\hat{z}, \hat{\tau}) \in \mathbb{R}^n, \rho > 0\), and let \(\hat{\nu}(\hat{z}, \hat{\tau}, \rho)\) be defined as in Lemma 5.6. Then
\begin{equation}
\hat{\nu}(\hat{z}, \hat{\tau}, \rho) \leq (\epsilon^2 + \|\nu\|)\rho^d.
\end{equation}

The rest of this section is devoted to the proofs of Lemma 5.6 and Lemma 5.10. To achieve this we will prove a number of auxiliary lemmas. Recall that \(R = \text{diam}(Q(S))\).

### 5.1. Auxiliary lemmas.

**Lemma 5.12.** Let \((z, \tau) \in C_{2\rho}(\chi_{Q(S)}, \tau_{Q(S)})\) and \(D(z, \tau)/60 \leq r \leq \kappa R/10\). Let \(I(z, \tau, r)\) be the set of all \(i \in \Lambda\) such that \(C_i(z, \tau) \cap I_i \neq \emptyset\). Let, for \(i \in \Lambda\), \(J(i)\) be the subset of \(I(z, \tau, r)\) which consists of those \(j\) which satisfy \(\text{diam}(Q(j)) \leq \text{diam}(Q(i))\) and \(2Q(j) \cap 2Q(i) \neq \emptyset\). Let
\[
N_i(Y, s) := \sum_{j \in J(i)} \chi_{2Q(j)}(Y, s).
\]
Then
\begin{equation}
\int \int N_i(Y, s) \, d\sigma(Y, s) \leq \sigma(Q(i)).
\end{equation}
Furthermore, there exists, for any \(\beta > 1\), a constant \(c = c(n, M, \beta) > 1\), such that
\begin{equation}
\sum_{i \in I(z, \tau, r)} (N_i(Y, s))^\beta \chi_{2Q(i)}(Y, s) \leq c,
\end{equation}
where \(\chi_{2Q(i)}\) is the characteristic function of \(2Q(i)\).
for all \((Y, s) \in \bigcup_{i \in I(z, r)} (2Q(i))\). Here we interpret\(^9\) \((N_i(Y, s))^{-\beta} = 0\) if \(N_i(Y, s) = \infty\).

**Proof.** To prove (5.13) of the lemma it suffices to note that

\[
\iint N_i(Y, s) \, d\sigma(Y, s) \leq \sum_{j \in J(i)} \sigma(Q(j)) \leq \sum_{j \in J(i)} \sigma(I_j) \leq \sigma(Q(i))
\]

since \(\{I_j\}\) is a disjoint collection, with \(d_p(I_j, \pi(Q(i))) \leq r_i\) for each \(j \in J(i)\).

To prove (5.14), we consider \((Y, s) \in \bigcup_{i \in I(z, r)} 2Q(i)\) fixed. Note that in the sum in (5.14) we only consider indices \(i \in I(z, r)\) such that \((Y, s) \in 2Q(i)\) and in the following we, for simplicity, denote this collection \(\tilde{I}\). An important observation is that there exists \(a' = a'(n, M) \in \mathbb{N}\), such that if \(i, i' \in \tilde{I}, Q(i) \in D_k\) and \(Q(i') \in D_{k'}\), for some \(k\) and \(k'\) such that \(k \leq k' - a'\), then \(i' \in J(i)\). This is a consequence of the facts that \(\text{diam}(Q) \approx 2^{-k}\) for all \(Q \in D_k\), and that \((Y, s) \in 2Q(i) \cap 2Q(i')\).

We divide the proof of (5.14) into two cases. First, assume that \(#\tilde{I} = \infty\). In this case, since each \(Q(i) \subseteq Q(S)\), there must exist a sequence of cubes \(\{Q(i_m)\}_m\) with \(i_m \in \tilde{I}\), such that \(\text{diam}(Q(i_m)) \to 0\) as \(m \to \infty\). Using the observation made above, this implies that \(N_i(Y, s) = \infty\) for all \(i \in \tilde{I}\). In particular,

\[
\sum_{i \in I(z, r)} (N_i(Y, s))^{-\beta} \mathbb{1}_{2Q(i)}(Y, s) = 0.
\]

Second, assume that \(#\tilde{I} < \infty\). In this case, the sets \(\{Q(i)\}\), where \(i \in \tilde{I}\), belong to a finite number of generations of \(D\). With this in mind, we let \(\{k_m\}_{m=1}^{m_0}, k_1 > k_2 > \cdots > k_{m_0}\), be all the integers for which there exists \(Q(i) \in D_{k_m}\) with \(i \in \tilde{I}\). For \(m \in \{1, 2, \ldots, m_0\}\) we introduce

\[
G_m := \{Q(i) : i \in \tilde{I}, Q(i) \in D_{k_m}\},
\]

and we note that \(G_m\) is non-empty. Again by the observation made above, we can conclude that if \(Q(i) \in G_m\) and \(Q(i') \in G_{m'}\), with \(m \geq m' + a'\), then \(i' \in J(i)\). Thus,

\[
N_i(Y, s) \geq \max\{1, m - a'\}, \quad \forall Q(i) \in G_m.
\]

For each dyadic generation \(k\), every \(Q \in D_k\) such that \((Y, s) \in 2Q\), must lie within a distance \(\leq \text{diam} Q \approx 2^{-k}\) of \((Y, s)\), hence, by the dyadic cube construction, there are at most a uniformly bounded number of such cubes; in particular, \(#G_m \leq c(n, M)\).

Consequently,

\[
\sum_{i \in I(z, r)} (N_i(Y, s))^{-\beta} \mathbb{1}_{2Q(i)}(Y, s) = \sum_{m=1}^{m_0} \sum_{Q(i) \in G_m} (N_i(Y, s))^{-\beta} \mathbb{1}_{2Q(i)}(Y, s)
\]

\[
\leq \sum_{m=1}^{m_0} #G_m (\max\{1, m - a'\})^{-\beta}
\]

\[
\leq c'(n, M, \beta),
\]

where we have used that \(a' = a'(n, M)\).

\(^9\)This interpretation will be valid when we use the functions \(N_i\), since by (5.13) the set of such points has measure zero.
The following lemma will be an important ingredient in the proof of Lemma 5.6.

**Lemma 5.16.** Let \((z, \tau) \in C^*_C(x_{Q(S)}, t_{Q(S)})\) and \(D(z, \tau)/60 \leq r \leq kR/10\). Let \(I(z, \tau, r)\) be the set of all \(i \in I\) such that \(C_i(z, \tau) \cap I_i \neq \emptyset\). Let \((X, t) \in \Sigma, (X, t) = (X(z, \tau, r), t(z, \tau, r))\), be such that \((X, t) \in Q(S)\) and \(d_p(z, \tau, \pi(X, t)) \leq r\). Then

\[
(\hat{\gamma}(z, \tau, r))^2 \leq r^{-d} \int_{\Sigma} (\gamma(Y, s, Kr/10))^2 \, d\sigma(Y, s) + \sum_{i \in I(z, \tau, r)} e^2(r_i/r)^{d+2},
\]

**Proof.** Let \((z, \tau) \in C^*_C(x_{Q(S)}, t_{Q(S)}), I(z, \tau, r)\) and

\((X, t) = (X(z, \tau, r), t(z, \tau, r)) \in \Sigma, \)

be as stated in the lemma. We emphasize that by construction the point \((X, t)\) depends on \((z, \tau)\) and \(r\). Let \(P := P(X, t, r) \in \mathcal{P}\) be the time independent plane for which the infimum in the definition of \(\gamma(X, t, Kr/10)\) is realized. Recall the notation

\[
\tilde{\Sigma} = \{(y, \psi(y, s), s) : (y, s) \in \mathbb{R}^n\}.
\]

To denote points on the approximating graph, we will in the following also use the notation \(\tilde{\Sigma}(y, s) := (y, \psi(y, s), s)\).

To start the proof we first note that

\[
(\hat{\gamma}(z, \tau, r))^2 \leq r^{-d} \int_{C(z, r)} \left(\frac{d_p(\tilde{\Sigma}(y, s), P)}{r}\right)^2 \, dy \, ds,
\]

and we introduce

\[
T := r^{-d} \int_{C(z, r) \cap \pi(F)} \left(\frac{d_p(\tilde{\Sigma}(y, s), P)}{r}\right)^2 \, dy \, ds,
\]

\[
T_i := r^{-d} \int_{C(z, r) \cap \pi_i} \left(\frac{d_p(\tilde{\Sigma}(y, s), P)}{r}\right)^2 \, dy \, ds.
\]

Using this notation we can continue the estimate in (5.18) and write

\[
(\hat{\gamma}(z, \tau, r))^2 \leq T + \sum_{i \in I(z, \tau, r)} T_i.
\]

It is straightforward to lift the integral in the definition of \(T\) to the graph, and to produce the estimate

\[
T \leq r^{-d} \int_{\pi^{-1}(C(z, r)) \cap F} \left(\frac{d_p((Y, s), P)}{r}\right)^2 \, d\sigma(Y, s) \leq (\gamma(X, t, Kr/10))^2.
\]

Note that in this estimate we have also used (5.4).

Fix \(i \in I(z, \tau, r)\). Then

\[
T_i \leq \hat{T}_i + \tilde{T}_i,
\]

where

\[
\hat{T}_i := r^{-d} \int_{C(z, r) \cap \pi_i} \left(\frac{d_p(\tilde{\Sigma}(y, s), P_{Q(i)})}{r}\right)^2 \, dy \, ds,
\]

and

\[
\tilde{T}_i := r^{-d} \int_{C(z, r) \cap \pi_i} \left(\frac{d_p(\tilde{\Sigma}(y, s), P_{Q(i)})}{r}\right)^2 \, dy \, ds.
\]
and

\[ \tilde{T}_i := r^{-d} r_i^d \sup \left\{ \left( \frac{d_p(Y, s, P)}{r} \right)^2 : (Y, s) \in P_{Q(i)}, d_p(Y, s, Q(i)) \leq c r_i \right\}. \]

Recall (4.15) and (4.16). Using that \( P_{Q(i)} \) is the graph of \( B_i \),

\[ \tilde{T}_i = r^{-d} \int_{C'(\xi, r) \cap I_i} \left( \frac{\|\psi(y, s) - B_i(y, s)\|}{r} \right)^2 \, dy \, ds \leq \varepsilon^2 r^{-d} r_i^d (r_i/r)^2, \]

where we have used that \( \|\psi(y, s) - B_i(y, s)\| \leq \varepsilon r_i \) on \( I_i \), see Lemma 4.18.

To estimate \( \tilde{T}_i \), we use Lemma A.1 applied to \( Q = Q(i) \). Indeed, let \((Z_j, \tau_j) \in Q(i)\) for \( j = 0, 1, \ldots, n \), be as in the statement of the lemma and let \( L_{Q(j)} = L_{n-1} \) be the spatial \((n-1)\)-dimensional plane which passes through \( Z_0, Z_1, \ldots, Z_{n-1} \). Then

\[ \tilde{T}_i \leq r^{-d} r_i^d \sup_{j \in \{0, \ldots, n-1\}} \left\{ \left( \frac{d_p(Z_j, \tau_j, P)}{r} \right)^2 + \left( \frac{d_p(Z_j, \tau_j, P_{Q(j)})}{r} \right)^2 \right\}. \]

To see this, we note that by Lemma A.1 we have that the points \((Z_j, \tau_j)\) stay at an ample distance from each other, and the point \((Y_j, s_j)\), realizing the infimum

\[ \inf_{(Y, s) \in P_{Q(i)}} d_p((Y, s), (Z_j, \tau_j)), \]

is within \( \varepsilon \) diam \( Q(i) \) of \((Z_j, \tau_j)\). Hence, for sufficiently small \( \varepsilon \), we can ensure that

\[ \| (Y_j, s_j) - (Y_{j'}, s_{j'}) \| \geq \text{diam} Q(i) \text{ whenever } j' \neq j. \]

In particular, the points \((Y_j, s_j)\) generate the plane \( P_{Q(j)} \), and for each \((Y, s) \in P_{Q(i)}\) such that \( d_p(Y, s, Q(i)) \leq c r_i \), we have

\[ d_p(Y, s, L_{n-1}) \leq \max_{j \in \{0, \ldots, n-1\}} d_p(Y_j, s_j, Z_j, \tau_j) = \max_{j \in \{0, \ldots, n-1\}} d_p(Z_j, \tau_j, P_{Q(j)}). \]

An even simpler argument shows that \( d_p(L_{n-1}, P) \) is controlled by

\[ \max_{j \in \{0, \ldots, n-1\}} d_p(Z_j, \tau_j, P). \]

In this case we do not need to find a generating set for \( P \). Put together we can conclude that (5.23) holds.

Using that \( Q(i) \in S \) we see that

\[ \sup_{j \in \{0, \ldots, n\}} \left\{ \left( \frac{d_p(Z_j, \tau_j, P_{Q(i)})}{r} \right)^2 \right\} \leq \varepsilon^2 (r_i/r)^2, \]

and hence, based on (5.23),

\[ \tilde{T}_i \leq r^{-d} r_i^d \sup_{j \in \{0, \ldots, n\}} \left\{ \left( \frac{d_p(Z_j, \tau_j, P)}{r} \right)^2 \right\} + r^{-d} r_i^d \varepsilon^2 (r_i/r)^2. \]

We now note that the statement in (5.24) remains true with \((Z_j, \tau_j)\) replaced by any \((\tilde{Z}_j, \tilde{\tau}_j) \in \Sigma\) satisfying \( d_p(Z_j, \tau_j, \tilde{Z}_j, \tilde{\tau}_j) \leq \eta r_i \) for \( \eta = \eta(n, M) \) sufficiently small. This can be seen from the proof of Lemma A.1 because we are free, at each iteration of the argument, to replace the points \((Z_j, \tau_j)\) with sufficiently close points \((\tilde{Z}_j, \tilde{\tau}_j)\), as this will preserve the desired properties of the (new) approximating planes and collection of points. Combining this observation, (5.24), (5.20), (5.21), (5.22) and averaging over such \((\tilde{Z}_j, \tilde{\tau}_j)\), we conclude that
(5.25) \((\hat{\gamma}(z, \tau)) \leq (\gamma(X, t, Kr/10))^2 + \sum_{i \in I(\tau, r)} e^2 r^{-d} r_i^d (r_i/r)^2 + \sum_{i \in I(\tau, r)} r^{-d} r_i^d \beta_{Q(i)},\)

where

\[
\beta_{Q(i)} := \left( \iiint_{2Q(i)} \left( \frac{d_p(Y, s, P)}{r} \right)^{2/3} d\sigma(Y, s) \right)^3.
\]

We now let \(J(i)\) be the subset of \(I(\tau, r)\) which consists of those \(j\) which are such that \(\text{diam}(Q(j)) \leq \text{diam}(Q(i))\) and \(2Q(j) \cap 2Q(i) \neq \emptyset\). We let

\[
N_i(Y, s) := \sum_{j \in J(i)} \chi_{2Q(j)}(Y, s).
\]

Then using Lemma 5.12 we have, for \(\beta > 1\) given, that there exists a constant \(c = c(n, M, \beta), 1 \leq c < \infty,\) such that

(5.26) \[\sum_i (N_i(Y, s))^{-\beta} \chi_{2Q(i)}(Y, s) \leq c,\]

for all \((Y, s)\) and

(5.27) \[\iiint_{2Q(i)} N_i(Y, s) d\sigma(Y, s) \leq 1.\]

To estimate \(\beta_{Q(i)}\), we apply Hölder’s inequality and use (5.27) to deduce that

(5.28) \[\beta_{Q(i)} = \left( \iiint_{2Q(i)} \left( \frac{d_p((Y, s, P))}{r} \right)^{2/3} (N_i(Y, s))^{-2/3} (N_i(Y, s))^{2/3} d\sigma(Y, s) \right)^3 \leq \left( \iiint_{2Q(i)} \left( \frac{d_p((Y, s, P))}{r} \right)^2 (N_i(Y, s))^{-2} d\sigma(Y, s) \right).\]

In particular we can rewrite (5.25) as

(5.29) \[\hat{\gamma}(z, \tau) \leq (\gamma(X, t, Kr/10))^2 + \sum_{i \in I(\tau, r)} e^2 r^{-d} r_i^d (r_i/r)^2 \]

\[+ r^{-d} \sum_{i \in I(\tau, r)} \iiint_{2Q(i)} \left( \frac{d_p((Y, s, P))}{r} \right)^2 (N_i(Y, s))^{-2} d\sigma(Y, s).\]

Then using (5.26) with \(\beta = 2\), we conclude that

(5.30) \[\hat{\gamma}(z, \tau) \leq (\gamma(X, t, Kr/10))^2 + \sum_{i \in I(\tau, r)} e^2 r^{-d} r_i^d (r_i/r)^2 \]

\[+ r^{-d} \iint_{\cup i \in I(\tau, r) 2Q(i)} \left( \frac{d_p((Y, s, P))}{r} \right)^2 d\sigma(Y, s).\]

Thus, to complete the proof of the lemma it now suffices to prove the estimate

(5.31) \[r^{-d} \iint_{\cup i \in I(\tau, r) 2Q(i)} \left( \frac{d_p((Y, s, P))}{r} \right)^2 d\sigma(Y, s) \leq (\gamma(X, t, Kr/10))^2.\]
To this end, we note that in turn, it is enough to prove that, for sufficiently large \( \kappa \),
\[
2Q(i) \subseteq C_{\kappa r/10}(X, t) \text{ for all } i \in I(z, \tau, r).
\]

To establish this inclusion, we observe that if \( i \in I(z, \tau, r) \) and \( (\hat{z}, \hat{\tau}) \in I_i \), then
\[
D(\hat{z}, \hat{\tau}) \leq D(z, \tau) + r \leq r,
\]
and hence diam \( I_i \leq r \). Because \( d_p(\pi(Q(i)), I_i) \leq \text{diam} \, I_i \), it follows that
\[
d_p(\pi(Q(i)), \pi(X, t)) \leq d_p(\pi(Q(i)), I_i) + d_p(I_i, (z, \tau)) + d_p(\pi(X, t), (z, \tau)) \leq r.
\]

Hence
\[
d_p(\pi(Q(i)), \pi(X, t)) \leq r.
\]

To see that \( 2Q(i) \subseteq C_{\kappa r}(X, t) \), choose \( (Y, s) \in 2Q(i) \). Since \( Q(i) \in S \), by (4.41) and our previous observation that diam \( I_i \leq r \), we obtain that
\[
d(Y, s) \leq \text{diam}(Q(i)) \leq \text{diam} \, I_i \leq r.
\]

Set \( \pi(Y, s) = (y, s), \pi(X, t) = (x, t) \). If \( d_p(Y, s, X, t) \leq r \), then there is nothing to prove. Otherwise, if \( r \leq d_p(Y, s, X, t) \), then using (5.34), we may apply Lemma 4.2 to conclude that
\[
|\pi^+(Y, s) - \pi^+(X, t)| \leq d_p(\pi(Y, s), \pi(X, t)) \leq r,
\]
where in the last step we have also used (5.33). Consequently, \( (Y, s) \in C_{\kappa r}(X, t) \), so that (5.32) holds for \( \kappa \) sufficiently large.

Combining the estimates deduced, we see that
\[
(\hat{y}(z, \tau, r))^2 \leq (y(X, t, Kr/10))^2 + \sum_{i \in I(z, \tau, r)} \epsilon^2 r^{-d} t_i^d (r_i/r)^2.
\]

Furthermore, the argument can be repeated with any \( (Y, s) \in \Sigma \cap C_r(X, t) \) in place of \( (X, t) \). Hence, averaging gives
\[
(\hat{y}(z, \tau, r))^2 \leq r^{-d} \int_{\Sigma \cap C_r(X, t)} (y(Y, s, Kr/10))^2 \, d\sigma(Y, s) + \sum_{i \in I(z, \tau, r)} \epsilon^2 r^{-d} t_i^d (r_i/r)^2,
\]
and we have arrived at the conclusion of the lemma.

\[\square\]

5.2. Proof of Lemma 5.6. Recall that we are assuming that \( \rho \leq \kappa R/10 \) (see (5.8)). Let \( \Lambda' \) be the set of all \( i \in \Lambda \) such that \( I_i \cap C'_{\rho}(\hat{z}, \hat{\tau}) \neq \emptyset \). Using the construction, (4.44) and Taylor’s theorem we immediately deduce that
\[
\sum_{i \in \Lambda} \int_{r_i}^{r_i + \delta} \left( \hat{y}(z, \tau, r) \right)^2 \frac{dr}{r} \leq \epsilon^2 \rho^d.
\]

It therefore remains to estimate \( \hat{y}(z, \tau, r) \) when either \( (z, \tau) \in \pi(F) \cap C'_{\rho}(\hat{z}, \hat{\tau}) \), or \( (z, \tau) \in I_i \cap C'_{\rho}(\hat{z}, \hat{\tau}), i \in \Lambda', \) but \( r_i < r \leq \rho \). By (4.47), both of these cases are covered if we assume \( (z, \tau) \in C'_{\rho}(\hat{z}, \hat{\tau}) \) and that \( D(z, \tau)/60 \leq r \leq \rho \). With this set up we are in a position to apply Lemma 5.16. To do so we let \( I(z, \tau, r) \) be the set of all \( i \in \Lambda \) such that \( C'_{\rho}(z, \tau) \cap I_i \neq \emptyset \). Let \( (X, t) = (X(z, \tau, r), t(z, \tau, r)) \in \Sigma \) be such that \( (X, t) \in Q(S) \) and \( d_p(\pi(X, t), \pi(X, t)) \leq r \). Then, by Lemma 5.16, we have estimate (5.17), which we reproduce here for the reader’s convenience:
\[
(\hat{y}(z, \tau))^2
\]
\[
\leq r^{-d} \int_{\Sigma \cap C_i(x,t)} (\gamma(Y, s, Kr/10))^2 \, d\sigma(Y, s) + \sum_{i \in I(z, r)} e^2 (r_i/r)^{d+2}.
\]

The estimate in (5.37) is valid for all \((z, \tau) \in C_i'(\hat{z}, \hat{\tau})\) satisfying \(D(z, \tau)/60 \leq r \leq \rho\). Hence,

\[
(5.38) \quad \int_{C_i'(\hat{z}, \hat{\tau})} \int_{D(z, \tau)/60}^{\rho} (\hat{\gamma}(z, \tau, r))^2 \, \frac{d\tau \, dr \, dz}{r} \leq A + e^2 B,
\]

where we define \(A\) to equal

\[
\int_{C_i'(\hat{z}, \hat{\tau})} \int_{D(z, \tau)/60}^{\rho} \left( \sum_{i \in I(z, r)} (r_i/r)^{d+2} \right) \, d\tau \, dr \, dz.
\]

and

\[
B := \int_{C_i'(\hat{z}, \hat{\tau})} \int_{D(z, \tau)/60}^{\rho} (\gamma(Y, s, Kr/10))^2 \, d\sigma(Y, s) \, \frac{dr \, dz \, d\tau}{r}.
\]

Note that in \(A\) we have written \((X(z, \tau, r), t(z, \tau, r))\) to highlight the dependence on \((z, \tau, r)\).

To estimate \(A\) we note that for any \((Y, s, \tau, \rho)\) arising in the integral we have

\[
d_{\rho}(z, \tau, \pi(Y, s)) \leq r,
\]

which also implies that \(D(Y, s) \leq D(z, \tau) + r \leq r\). Using this we see that \(A\) is bounded from above by

\[
\int_{\Sigma \cap C_i'(\hat{z}, \hat{\tau})} \int_{C_i(Y, \pi(Y, s))}^{\rho} r^{-d} \left( \int_{C_i(Y, \pi(Y, s))} dz \, d\tau \right) \left( \gamma(Y, s, Kr/10) \right)^2 \, dr \, d\sigma(Y, s) \, \frac{dz 
\]

\[
\leq \int_{\Sigma \cap C_i'(\hat{z}, \hat{\tau})} \int_{C_i(Y, \pi(Y, s))}^{\rho} \left( \gamma(Y, s, Kr/10) \right)^2 \, dr \, d\sigma(Y, s) \, \frac{dz 
\]

\[
\leq \int_{\Sigma \cap C_i'(\hat{z}, \hat{\tau})} \int_{C_i(Y, \pi(Y, s))}^{\rho} (\gamma(Z, \tau, Kr))^2 \, dr \, d\sigma(Z, \tau) \, d\tau.
\]

To estimate \(B\) we first note that if \(i \in I(z, \tau, r)\), then \(r_i \leq r\). Hence

\[
\sum_{i \in I(z, \tau, r)} (r_i/r)^{d+2} \leq \frac{1}{r^{d+1}} \sum_{i \in I(z, \tau, r)} r_i^{d+1},
\]

and we see that

\[
B \leq \int_{C_i'(\hat{z}, \hat{\tau})} \int_{D(z, \tau)/60}^{\rho} \left( \sum_{i \in I(z, \tau, r)} r_i^{d+1} \right) \, d\tau \, dr \, dz \, d\tau \leq \frac{1}{r^{d+2}} \sum_{i \in I(z, \tau, r)} r_i^{d+1} \, dr \, d\sigma(Z, \tau) \, d\tau \leq \frac{1}{r^{d+2}} \sum_{i \in I(z, \tau, r)} r_i^{d+1} \, dr \, d\sigma(Z, \tau) \, d\tau.
\]

and

\[
B \leq \frac{1}{r^{d+2}} \sum_{i \in I(z, \tau, r)} r_i^{d+1} \, dr \, d\sigma(Z, \tau) \, d\tau \leq \frac{1}{r^{d+2}} \sum_{i \in I(z, \tau, r)} r_i^{d+1} \, dr \, d\sigma(Z, \tau) \, d\tau \leq \frac{1}{r^{d+2}} \sum_{i \in I(z, \tau, r)} r_i^{d+1} \, dr \, d\sigma(Z, \tau) \, d\tau.
\]
We can justify the second inequality in this deduction as follows. The bounds of integration for the integral in $r$ hold because $D(z, \tau) \leq r = r_1 + r$ when $i \in I(z, \tau, r)$, and the indicator appears because when $i \in I(z, \tau, r)$, $d_p(I_i, (z, \tau)) \leq r$ must hold. This completes the proof of Lemma 5.6.

5.3. Proof of Lemma 5.10. To prove Lemma 5.10 we fix $(\hat{\tau}, \hat{\tau}) \in \mathbb{R}^n$ and $\rho > 0$, and we divide the proof into the following cases:

(i) $C_{2p}^r(\hat{\tau}, \hat{\tau}) \subset C_{2p}^r(x, t, Q_S((x, t), \rho R))$, $\rho \leq \kappa R/10$,

(ii) $C_{2p}^r(\hat{\tau}, \hat{\tau}) \subset C_{2p}^r(x, t, Q_S((x, t), \rho R))$, $\rho \geq \kappa R/10$,

(iii) $C_{2p}^r(\hat{\tau}, \hat{\tau}) \cap (\mathbb{R}^n \setminus C_{2p}^r(x, t, Q_S((x, t), \rho R))) = \emptyset$.

If $(\hat{\tau}, \hat{\tau}, \rho)$ is as in (i) then the estimate of Lemma 5.10 follows immediately from Lemma 5.6 and the fact that $\Sigma$ is parabolic UR. Assume that $(\hat{\tau}, \hat{\tau}, \rho)$ is as in (ii) and consider $(z, \tau) \in C_{2p}^r(\hat{\tau}, \hat{\tau})$, $\rho \geq \kappa R/10$. In this case

$$\hat{\nu}(z, \tau, \eta) \leq \hat{\nu}(x, t, Q_S((x, t), \rho R)) \leq (\delta^2 + \|\nu\|\|\kappa R\|d) \leq (\delta^2 + \|\nu\|\|\kappa R\|d)$$

by a covering argument and (i). Assume that $(\hat{\tau}, \hat{\tau}, \rho)$ is as in (iii). Recall that $\psi \equiv 0$ on $\mathbb{R}^n \setminus C_{4p}^r(x, t, Q_S((x, t), \rho R))$ and hence we can without loss of generality assume that

$$C_{2p}^r(\hat{\tau}, \hat{\tau}) \cap C_{4p}^r(x, t, Q_S((x, t), \rho R)) = \emptyset.$$

This case can be handled as in (5.36) by using the construction, (4.44) and Taylor’s theorem. We omit the remaining ‘routine’ details and claim that the proof of Lemma 5.10 is complete.

Lemma 5.10 also proves that $\psi$ is a regular Lip(1,1/2) graph.

Lemma 5.40.

$$|D_1^{1/2}\psi|_s \leq (\delta^2 + \|\nu\|^{1/2}).$$

In particular, $\psi_S := \psi = \psi(x, t) : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ is a regular Lip(1,1/2) function with parameters $b_1 \leq \delta$ and $b_2 = b_2(n, M, \Gamma)$.

To prove Lemma 5.40, we observe that $\hat{\nu}(\hat{\tau}, \hat{\tau}, \rho)$ is a (parabolic) Carleson measure on $\mathbb{R}^n \times \mathbb{R}_+$, by Lemma 5.10. After making this observation, one may use the arguments in Section 2 of [HLN2] or the arguments in Section 2 and at the beginning of Section 3 in [HLN1], to produce the desired estimate on $|D_1^{1/2}\psi|_s$. Note that the measure used in [HLN1] is $\sigma^s$, a measure which is equivalent to $\sigma$ in the setting of Lip(1,1/2) graphs and, consequently, this change does not change the validity of the arguments (see Remark 2.8).

Proof. We follow the argument in [HLN2]; for the reader’s convenience we include a detailed account of their result and make the appropriate changes for our purposes.

Throughout the proof we will use simple variants of the following elementary estimate:

$$\sup_{C_r(z, t)} \inf_{C_r(x, t)} \psi \leq \delta r, \quad (x, t) \in \mathbb{R}^n,$$

where the implicit constant is the same as for the Lip(1,1/2) constant for $\psi$ obtained in Lemma 4.26.
For the rest of the proof we will fix a point \((z, \tau) \in \mathbb{R}^n\) and a scale \(\rho > 0\). Associated to these parameters we choose a cut-off function \(\beta = \beta_{z, \tau, \rho} \in C^\infty_0(C'_\rho(z, \tau)),\)

\(0 \leq \beta \leq 1\), and \(\beta \equiv 1\) in \(C_{3\rho/2}(z, \tau)\). To simplify notation we will also set

\[ C := C'_\rho(z, \tau). \]

We define

\[ \psi_1 + \psi_2 := [(\psi - \psi_C)\beta] + [\beta \psi_C + (1 - \beta)\psi], \]

where as usual \(\psi_C\) denotes the average (with respect to Lebesgue measure) of \(\psi\) over \(C\).

Recall that we aim to control the \(BMO\) norm of \(D_1^{1/2}/\psi\), or in other words, we wish to prove

\[
\iint_C |D_1^{1/2}/\psi - (D_1^{1/2}/\psi)_C| \, dy \, ds \leq (\delta^2 + ||\nu||)^\rho d, \tag{5.42}
\]

with appropriate dependence of the implicit constants on allowable parameters\(^{10}\).

To prove (5.42), add and subtract a constant \(\alpha\) (to be determined) and use the decomposition \(\psi = \psi_1 + \psi_2\) from before to get

\[
\iint_C |D_1^{1/2}/\psi - (D_1^{1/2}/\psi)_C| \, dy \, ds \leq \iint_C |D_1^{1/2}/\psi_2 - \alpha| \, dy \, ds + |C||\alpha - (D_1^{1/2}/\psi)_C| + \iint_C |D_1^{1/2}/\psi_1| \, dy \, ds.
\]

Notice that for the middle term we have

\[ |\alpha - (D_1^{1/2}/\psi)_C| \leq |C|^{-1} \int_C |\alpha - D_1^{1/2}/\psi_2 - D_1^{1/2}/\psi_1| \, dy \, ds, \]

and so

\[
\iint_C |D_1^{1/2}/\psi - (D_1^{1/2}/\psi)_C| \, dy \, ds \leq 2 \iint_C |D_1^{1/2}/\psi_2 - \alpha| \, dy \, ds
\]

\[
+ 2 \iint_C |D_1^{1/2}/\psi_1| \, dy \, ds =: I + II.
\]

To handle \(II\) we need some preparation. Introduce first an approximate identity \(P_\lambda\) and abuse notation to denote both the operator and it’s kernel in this fashion. In particular, we require \(P \in C^\infty_0(C'_\lambda(0, 0))\), even and nonnegative with \(\iint_{\mathbb{R}^n} P \, dy \, ds = 1\), and as usual set \(P_\lambda(y, s) = \lambda^{-d}P(y/\lambda, s/\lambda^2)\).

In what follows \(\phi_\lambda(y, s) = \lambda^{-d}\phi(y/\lambda, s/\lambda^2)\) may denote either of the following functions:

\[ \lambda \frac{\partial P}{\partial \lambda}, \quad \lambda^2 \frac{\partial P}{\partial s}, \quad \lambda^2 \frac{\partial^2 P}{\partial y^2} \quad \text{for} \quad 1 \leq i \leq n - 1. \]

The main features of these functions for us are:

\[ \int_{\mathbb{R}^n} |\phi_\lambda| \, dy \, ds = \int_{\mathbb{R}^n} |\phi| \, dy \, ds \leq 1, \quad \sup_{\mathbb{R}^n} |\phi_\lambda| = \lambda^{-d} \sup_{\mathbb{R}^n} |\phi| \leq \lambda^{-d}. \]

\(^{10}\)As will be seen from the proof, the implicit constants will depend on the same parameters as the \(Lip(1, 1/2)\) estimate from Lemma 4.26.
For each \( 1 \leq i \leq n \) we have
\[
\iint_{\mathbb{R}^n} \phi \ dy \ ds = 0 = \iint_{\mathbb{R}^n} y_i \phi \ dy \ ds.
\]
By the second property of \( \phi \) above we have, for any linear function \( L \) of \( y \) only,
\[
(\phi + \psi)^2(y, s) = (\phi + (\psi - L))^2(y, s)
\]
where \( H \) denotes the Hilbert transform and we have used the following facts:
\[
\text{(5.45)}
\]
\[
\text{(5.47)}
\]
By the second property of \( \phi \) above we have, for any linear function \( L \) of \( y \) only,
\[
(\phi + \psi)^2(y, s) = (\phi + (\psi - L))^2(y, s)
\]
where \( H \) denotes the Hilbert transform and we have used the following facts:
\[
\text{(5.45)}
\]
\[
\text{(5.47)}
\]

Multiplying by \( \lambda^{-3} \) and integrating the above equation over \( C_p(z, \tau) \times (0, \rho) \) we arrive at
\[
\int_0^\rho \iint_{C_p(z, \tau)} \lambda^{-3}(\phi + \psi)^2 \ dy \ ds \ d\lambda \leq_n \int_0^\rho \iint_{C_p(z, \tau)} \lambda^{-1} \tilde{\psi}(y, s) \ dy \ ds \ d\lambda,
\]
or, in other words,
\[
\text{(5.44)}
\]
We are now ready to begin estimating \( II \) in (5.42). Fix \( 0 < \varepsilon < \rho \), then using the fundamental theorem of calculus
\[
\text{(5.45)}
\]
where \( \text{H} \) denotes the Hilbert transform and we have used the following facts:
\[
D_s^{1/2} D_s^{1/2} = \text{H} \partial_s, \quad \text{H}^* = -\text{H}.
\]
To handle \( II \) we first claim that it’s enough to show
\[
|D_s^{1/2} P_s \psi_1(y, s)| \leq n \ \delta \ min \ \left\{ 1, \frac{\rho^3}{|s - \tau|^{1/2}} \right\} \ 1_{|y - \beta| \leq 4 \rho}.
\]
Assuming this for the moment, we integrate in \( y, s \) to get
\[
\text{(5.46)}
\]
\[
\text{(5.47)}
\]
which is the right type of bound (see (5.42)). Therefore, it remains only to prove (5.46). For this purpose recall that for nice functions\(^{11}\) \(f\) we defined
\[
D^1f(y, s) := c \int_{\mathbb{R}} f(y, t) - f(y, s) \, dt / |t-s|^{3/2}
\]
In particular, since \(\psi_1\) is supported on \(C^2_{\rho}(\bar{z}, \tau)\), if \(|z - \tau| > 4\rho\) then \(D^1 P_{\rho}\psi_1(y, s) = 0\). We have thus reduced (5.46) to proving
\[
|D^1 P_{\rho}\psi_1(y, s)| \lesssim \delta \min\left\{1, \frac{\rho^3}{|s-\tau|^{3/2}}\right\}, \quad |y-z| \leq 4\rho.
\]
First we note that
\[
|\partial_s P_{\rho}\psi_1(y, s)| \leq \frac{\delta}{\rho}, \quad (y, s) \in \mathbb{R}^n.
\]
This follows from properties of the convolution:
\[
|\partial_s P_{\rho}\psi_1(y, s)| = \rho^{-2} |(\partial_s P)_{\rho}\psi_1(y, s)| \leq \rho^{-2} \|\partial_s P\|_{L^1(\mathbb{R}^n)} \|\psi_1\|_{L^n(\mathbb{R}^n)} \lesssim \frac{\delta}{\rho},
\]
where we used (5.41) and the definition of \(\psi_1\) in the last inequality.

We first get the uniform bound; for this we break up the integral in the definition of the half-order time derivative and the above estimate, together with (5.41), to get
\[
|D^1 P_{\rho}\psi_1(y, s)| \leq \int_{|t-s|<\rho^2} |P_{\rho}\psi_1(y, t) - P_{\rho}\psi_1(y, s)| / |t-s|^{3/2} \, dt + \int_{|t-s|>\rho^2} |P_{\rho}\psi_1(y, t) - P_{\rho}\psi_1(y, s)| / |t-s|^{3/2} \, dt
\]
\[
\leq \frac{\delta}{\rho} \int_{|t-s|<\rho^2} |t-s|^{-1/2} \, dt + \rho \int_{|t-s|>\rho^2} |t-s|^{-3/2} \, dt
\]
\[
\leq \delta.
\]
Therefore it remains to get the decay in estimate (5.48), and, for this, we need only consider\(^{12}\) \(|s-\tau| > 100\rho^2\). For such \(s\) we have \(P_{\rho}\psi_1(y, s) = 0\) and so
\[
|D^1 P_{\rho}\psi_1(y, s)| \leq \int_{\mathbb{R}} |P_{\rho}\psi_1(y, t)| / |t-s|^{3/2} \, dt = \int_{|t-s|<16\rho^2} |P_{\rho}\psi_1(y, t)| / |t-s|^{3/2} \, dt
\]
\[
\leq \rho \delta \int_{|t-s|<16\rho^2} |t-s|^{-3/2} \, dt \approx \rho \delta \rho^2 |t-\tau|^{-3/2},
\]
where we again used (5.41) in the second line and the fact that \(|t-s| \approx |\tau-s|\) with the given restrictions on \(s, \tau\). This is exactly the desired decay in (5.48), and thus we’ve finished controlling \(II_1\).

---

\(^{11}\)Such as \(P_{\rho}\psi_1 \in C^2_{\rho}(\mathbb{R}^n)\).

\(^{12}\)Perhaps worsening our implicit constant by an universal factor.
$I_2$ is a bit more elaborate. We first apply Cauchy-Schwarz to get
\[
|I_2|^2 \leq n \left( \int \int_{\mathbb{R}^n} \lambda^{-1} |H(\partial_t(P_A \psi_1))[\nu] \right)^2 \ dy \ ds \ d\lambda
\times \left( \int \int_{\mathbb{R}^n} |\partial_t(P_A \psi_1)|^2 \ dy \ ds \ d\lambda \right)
=: II_1 \cdot II_2.
\]

From the $L^2$ boundedness of the Hilbert transform on $\mathbb{R}$ we see that
\[
(5.50) \quad II_1 \leq \int \int_{\mathbb{R}^n} \lambda^{-1} |\partial_t(P_A \psi_1)|^2 \ dy \ ds \ d\lambda.
\]

With this in hand, the estimates for $II_1$ and $II_2$ will follow the same lines, the different powers of $\lambda$ accounting for the difference in scaling of $\partial_t P_A$ and $\partial_t P_A$. In particular one should keep in mind the fact that both of these functions “annihilate” first order (spatial) polynomials (see the properties of $\phi$ above).

With the above in hand, we focus on $II_1$.

For fixed $(y, s) \in C$ consider the linear polynomial
\[
L(x) := \beta(y, s) + \sum_{i=1}^{n-1} \partial_s \beta(y, s)(y_i - x_i), \quad x \in \mathbb{R}^{n-1}.
\]

Using the alluded properties of $\partial_t P_A$ and the triangle inequality we see
\[
(5.51) \quad |\partial_t(P_A \psi_1)(y, s)| \leq \int \int_{\mathbb{R}^n} |\partial_t P_A (y - x, s - t)|[\psi(x, t) - \psi_C][\beta(x, t) - L(x)] \ dx \ dr
\]
\[+ \sum_{i=1}^{n-1} |\partial_s \beta(y, s) \int \int_{\mathbb{R}^n} (y_i - x_i)\partial_t P_A (y - x, s - t)[\psi(x, t) - \psi(y, s)] \ dx \ dr
\]
\[+ |\beta(y, s)\partial_t P_A \psi(y, s)| =: A + B + C.
\]

To handle $A$ recall that $\|\partial_t P_A\|_{L^1} \leq \lambda^{-1}$, and (5.41) (together with the fact that the integration is taking place over $C(y, s)$ and $\lambda < \rho$) to get
\[
A \leq \lambda^{-1} \delta \rho \sup_{(x, t) \in C(y, s)} |\beta(x, t) - L(x)|.
\]

To control the last term we appeal to Taylor’s theorem:
\[
|\beta(x, t) - L(x)| \leq |\beta(x, t) - \beta(y, t) - \sum_{i=1}^{n-1} \partial_s \beta(y, t)(y_i - x_i)|
\]
\[+ |\beta(y, t) - \beta(y, s)| + \sum_{i=1}^{n-1} |\partial_s \beta(y, t) - \partial_s \beta(y, s)| |y_i - x_i|
\]
\[\leq |D^2_\beta(\cdot, s)|_{L^\infty(\mathbb{R}^{n-1})} |x - y|^2 + |\partial_s \beta(\cdot, \cdot)|_{L^\infty(\mathbb{R}^n)} |t - s|
\]
\[+ |\partial_s D_\beta(\cdot, \cdot)|_{L^\infty(\mathbb{R}^n)} |x - y|^2 |t - s| \leq \rho^{-2} \lambda^2 + \rho^{-2} \lambda^2 + \rho^{-3} \lambda^3 \leq \rho^{-2} \lambda^2,
\]
where we used the derivative bounds for \( \beta \) and in the second to last line the fact that \( \lambda < \rho \). Plugging this estimate, we arrive at

\[
A \lesssim \frac{\delta \lambda}{\rho}.
\]

Using again the derivative estimates for \( \beta \) and (5.41) we see

\[
B \leq \rho^{-1} \lambda \lambda^{-1} \delta \lambda = \frac{\delta \lambda}{\rho}.
\]

Trivially, since \( \beta \leq 1 \) we have

\[
C \leq |\partial_1 P_1 \psi(y, s)|,
\]

and so, combining all these estimates into (5.50) we arrive at

\[
I_{21} \leq \int_0^\rho \int_{C_{\rho}(z, t)} \lambda^{-1} \left( \frac{\delta \lambda}{\rho} + |\partial_1 P_1 \psi(y, s)| \right)^2 \ dy \ ds \ dl
\]

\[
\lesssim \int_0^\rho \int_{C_{\rho}(z, t)} \left( \frac{\delta^2 \lambda}{\rho^2} + \lambda^{-1} |\partial_1 P_1 \psi(y, s)|^2 \right) \ dy \ ds \ dl
\]

\[
\lesssim \delta^2 \rho^{d} + ||\check{\nu}|| \rho^{d},
\]

where we have used (5.44) in the last line with \( \phi = \lambda \partial_1 P_1 \).

To handle \( I_{22} \) the only change is that \( |\partial_1 P_1|_{L^1(\mathbb{R}^n)} \leq \lambda^{-2} \), which gives

\[
|\partial_1 P_1 \psi_1(y, s)| \leq \frac{\delta}{\rho} + |\partial_1 P_1 \psi(y, s)|,
\]

and thus

\[
I_{22} \leq \delta^2 \rho^d + ||\check{\nu}|| \rho^d.
\]

These last two estimates, in turn, give

\[
I_2 \lesssim (\delta^2 + ||\check{\nu}||) \rho^d.
\]

Combining these estimates we have

\[
\int_{\mathbb{R}^n} |D_1^{1/2} P_{\epsilon} \psi_1|^2 \ dy \ ds \lesssim (\delta^2 + ||\check{\nu}||) \rho^d,
\]

with bounds uniform in \( \epsilon \). In fact the implicit constants only depend on the parameters upon which the \( L^p(1, 1/2) \) bound in Lemma 4.26 depends.

By weak compactness, there is a subsequence (which we denote the same) and an \( h \in L^2(\mathbb{R}^n) \) such that \( D_1^{1/2} P_{\epsilon} \psi_1 \to h \), as \( \epsilon \to 0^+ \). Since \( P_{\epsilon} \psi_1 \in C_c^\infty(\mathbb{R}^n) \) and \( \psi_1 \) has compact support we see that \( I_{1/2} \ast (D_1^{1/2} P_{\epsilon} \psi_1) = P_{\epsilon} \psi_1 \) converges pointwise, uniformly, everywhere to \( \psi_1 \) and to \( I_{1/2} \ast h \) (here \( I_{1/2} \) is the fractional integral of order 1/2 in the time variable, appropriately normalized). This shows that \( D_1^{1/2} \psi_1 \) exists, and satisfies the estimate

\[
\int_{\mathbb{R}^n} |D_1^{1/2} \psi_1|^2 \ dy \ ds \lesssim (\delta^2 + ||\check{\nu}||) \rho^d.
\]

Therefore

\[
II = \int_{\mathbb{C}} |D_1^{1/2} \psi_1| \ dy \ ds \lesssim \rho^{d/2} \left( \int_{\mathbb{C}} |D_1^{1/2} \psi_1|^2 \ dy \ ds \right)^{1/2} = (\delta^2 + ||\check{\nu}||)^{1/2} \rho^d.
\]
It thus remains to control $I$ in (5.42). First, notice that for $(y, s) \in C'_2(z, \tau)$ we have $\psi_2(y, s) = \psi_C$, by definition of $\psi_2$, so that in particular the integral defining $D^{1/2}_s \psi_2$ is always convergent for such points. By the decomposition $\psi = \psi_1 + \psi_2$ we conclude that $D^{1/2}_s \psi$ exists in $C$.

Write

$$D^{1/2}_s \psi_2(y, s) = c \int_{|t-\tau|>9\rho^2/4} \frac{\psi_2(y, t) - \psi_2(y, s)}{|t-s|^{3/2}} \, dt, \quad (y, s) \in C,$$

and set $\alpha$ to be

$$\alpha := c \int_{|t-\tau|>9\rho^2/4} \frac{\psi(t, t) - \psi(z, \tau)}{|t-\tau|^{3/2}} \, dt.$$

If we can show, for this value of $\alpha$,

$$(5.52) \quad |D^{1/2}_s \psi_2(y, s) - \alpha| \leq \delta, \quad (y, s) \in C,$$

then integrating would give

$$I = \int_C |D^{1/2}_s \psi_2 - \alpha| \, dy \, ds \leq \delta \rho^d,$$

which combined with the estimate for $II$, gives the desired BMO bound in (5.42).

The rest of the proof is then devoted to showing (5.52).

First notice that for $(y, s) \in C$ and $|t-\tau| > 9\rho^2/4$, we have $|t-\tau| \approx |t-s|$; we will exploit this repeatedly below without mention.

From the above formulas we have

$$|D^{1/2}_s \psi_2(y, s) - \alpha| = c \int_{|t-\tau|>9\rho^2/4} \left| \frac{\psi_2(y, t) - \psi_2(y, s)}{|t-s|^{3/2}} - \frac{\psi(z, t) - \psi(z, \tau)}{|t-\tau|^{3/2}} \right| \, dt.$$

Recall $\psi_2 = (\psi_C - \psi)\beta + \psi$, so that we can rewrite the quantity in parentheses as

$$\left( \frac{(\psi_C - \psi(y, t))\beta(y, t)}{|t-s|^{3/2}} - \frac{(\psi_C - \psi(y, s))\beta(y, s)}{|t-s|^{3/2}} + \frac{(\psi(y, t) - \psi(y, s))\beta(y, s)}{|t-s|^{3/2}} \right).$$

Call these terms $A_1, A_2, A_3$ respectively.

For the first one, notice that $\beta(y, t)$, as a function of $t$, is supported on $|t-\tau| < 4\rho^2$ by definition, so

$$\int_{|t-\tau|>9\rho^2/4} |A_1| \, dt = \int_{|t-\tau|=4\rho^2} \frac{|\psi_C - \psi(y, t)|}{|t-s|^{3/2}} \, dt \approx \int_{|t-\tau|=4\rho^2} \frac{|\psi_C - \psi(y, t)|}{|t-\tau|^{3/2}} \, dt,$$

and since then $(y, t) \in C'_{4\rho}(z, \tau)$, we use (5.42) to bound the numerator by $\delta \rho$, which gives

$$\int_{|t-\tau|>9\rho^2/4} |A_1| \, dt \leq \delta \int_{|t|=\rho^2} r^{-3/2} \, dt \leq \delta.$$

For $A_2$ we use that $|\beta| \leq 1$ and $(y, s) \in C$ to again bound the numerator by $\rho \delta$ and so

$$\int_{|t-\tau|>9\rho^2/4} |A_2| \, dt \leq \rho \delta \int_{|t|=\rho^2} r^{-3/2} \, dt \leq \delta.$$
For $A_3$ we first put a common denominator on the quantity in parentheses and rewrite the result as

\[
\frac{|(\psi(y, t) - \psi(z, t)) - (\psi(y, s) - \psi(z, \tau))| |t - \tau|^{3/2} + (\psi(z, t) - \psi(z, \tau))(|t - \tau|^{3/2} - |s - t|^{3/2})}{|t - s|^{3/2}}.
\]

Now notice that for any given $t$, the points $(y, t), (z, t)$ lie in a (parabolic) cylinder of size $< \rho$, and so each term inside the brackets can be controlled, using (5.41), by $\delta \rho$. This an the $Lip(1, 1/2)$ estimate for $\psi$ directly then give

\[
|A_3| \leq \frac{\delta \rho}{|t - s|^{3/2}} + \delta \frac{|t - \tau|^{1/2} |t - \tau|^{3/2} - |s - t|^{3/2}}{|t - \tau|^{3/2} |t - s|^{3/2}}
\]

\[
\leq \frac{\delta \rho}{|t - \tau|^{3/2}} + \delta \frac{|t - \tau|^{1/2} |t - \tau|^{3/2} + |s - t|^{1/2} |s - \tau|}{|t - \tau|^{3/2} |t - s|^{3/2}}
\]

\[
\leq \frac{\delta \rho}{|t - \tau|^{3/2}} + \frac{\delta \rho^2}{|t - \tau|^2},
\]

where we used the elementary inequality $|a^p - b^p| \leq (a^{p-1} + b^{p-1})|a - b|$ in the second line. Integrating this gives

\[
\int_{|t - \tau| > \delta \rho^2 / 4} |A_3| \, dt \leq \delta,
\]

which finishes the proof. \hfill \Box

6. Packing the maximal cubes

In this section we conclude the proof of Theorem 3.1, by proving the packing condition for the maximal cubes $\{Q(S)\}_{S \in \mathcal{F}}$. In order to reduce matters, we need to introduce some additional notation. Recall that $m(S)$ is the collection of minimal cubes in a stopping time regime $S \in \mathcal{F}$. We let $m_0(S)$ denote the cubes in $m(S)$ which are such that $Q$ has a child in $B'$ and we let $m_1(S)$ denote the cubes in $m(S)$ for which the angle between $P_{Q(S)}$ and $P_Q$ is greater or equal to $\delta / 2$. Hence, by construction

\[m(S) = m_0(S) \cup m_1(S).\]

Following [DS1], we now let

\[
\mathcal{F}_0 := \{S \in \mathcal{F} : \sigma(\cup_{Q \in m_0(S)} Q) \geq \sigma(Q(S))/4\},
\]

\[\mathcal{F}_1 := \{S \in \mathcal{F} : \sigma(\cup_{Q \in m_1(S)} Q) \geq \sigma(Q(S))/2\},
\]

\[\mathcal{F}_2 := \{S \in \mathcal{F} : \sigma(Q(S) \setminus \cup_{Q \in m(S)} Q) \geq \sigma(Q(S))/4\}.
\]

Then $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2$. It follows (see [DS1, Lemma 7.4]) that

\[
\sum_{S \in \mathcal{F}_0 \cup \mathcal{F}_2 : Q(S) \subset Q} \sigma(Q(S)) \leq c(n, M, \Gamma, \epsilon, \delta, K) \sigma(Q), \quad \forall Q \in \mathcal{D}(\Sigma).
\]

The ideas behind this estimate are that to obtain the bound for $\mathcal{F}_2$, one uses that the sets $Q(S) \setminus (\cup_{Q \in m(S)} Q)$ are pairwise disjoint, while the bound for $\mathcal{F}_0$ follows from the packing condition which by construction holds for $B'$. We include the details.
of these arguments below. In the case of $F_2$ we have for each $Q(S)$ a set $F_S \subseteq Q(S)$ with each $F_S$ disjoint and $\sigma(F_S) \geq (1/4)\sigma(Q(S))$. Thus, for $\bar{Q}$ any cube
\[
\sum_{\bar{Q} \subseteq Q} \sigma(Q(S)) \leq 4 \sum_{\bar{Q} \subseteq Q} \sigma(F_S) \leq 4\sigma(\bar{Q}),
\]
where we used that $F_S \subseteq Q(S) \subseteq \bar{Q}$ are disjoint, so their measures cant add up to more than $\sigma(\bar{Q})$.

The $F_0$ stopping times have a different packing argument. If $S$ is in $F_0$ we have for each $Q(S)$ and associated collection of subcubes $Q_j$ of $Q(S)$, call them $B_S$, which have a child $Q_j$ in $B'$. In particular, we take $B_S$ to be the $m_0(S)$ cubes and they satisfy
\[
\sum_{Q \in B_S} \sigma(Q) = \sigma \left( \bigcup_{Q \in B_S} Q \right) \geq (1/4)\sigma(Q(S)).
\]
One notes that if $S, S'$ is in $F_0$ are distinct then $B_S \cap B_{S'}$, which can be seen by the disjointness of the stopping times. Thus, if $\bar{Q}$ is a cube
\[
\sum_{\bar{Q} \subseteq Q} \sigma(Q(S)) \leq 4 \sum_{\bar{Q} \subseteq Q} \sum_{Q \in B_S} \sigma(Q) \leq 4C \sum_{\bar{Q}' \subseteq \bar{Q}} \sigma(Q') \leq 4C \sigma(\bar{Q}),
\]
where the $C$ is the packing constant for the cubes $B'$ and $C_*$ is a constant so that if $Q'$ is a child of $Q$ then $\sigma(Q) \leq C_* \sigma(Q')$. Thus, to pack the maximal cubes (and hence prove Theorem 3.1) it suffices to prove that
\[
\sum_{S \in F_1 \cap Q(S) \subseteq Q} \sigma(Q(S)) \leq c(n, M, \Gamma, \epsilon, \delta, K) \sigma(Q), \quad \forall Q \in \mathbb{D}(\Sigma).
\]

Consider $S \in F$ and assume that
\[
\int_{E(S)} (\gamma(Z, \tau, Kr))^2 \frac{\text{d}\sigma(Z, \tau)}{r} dr \leq \varepsilon^2 \sigma(Q(S)),
\]
where $E(S) := E(x_{Q(S)}, t_{Q(S)}, R)$ is the region appearing in Lemma 5.6, and $R := \text{diam}(Q(S))$, see (4.9). I.e. $E(S)$ is the set of all $(Z, \tau, r) \in [\Sigma \cap \pi^{-1}(C_{KR}(x_{Q(S)}, t_{Q(S)}))] \times (0, \infty)$ such that $K^{-1}d(Z, \tau) \leq r \leq KR$.

We are the going to prove that there exists $\varepsilon$ small, independent of $S$, such that if (6.3) holds then, $S \not\in F_1$. Hence, if $S \in F_1$ then the opposite inequality in (6.3) must hold and therefore, if $Q \in \mathbb{D}(\Sigma)$ then
\[
\sum_{S \in F_1 \cap Q(S) \subseteq Q} \sigma(Q(S)) \leq \varepsilon^{-2} \sum_{S \in F_1 \cap Q(S) \subseteq Q} \int_{E(S)} (\gamma(Z, \tau, Kr))^2 \frac{\text{d}\sigma(Z, \tau)}{r} dr \leq c(n, M, \Gamma, \epsilon, \delta, K) \sigma(Q),
\]
by the facts that $\Sigma$ is parabolic UR, that $\sigma$ is a doubling measure, and that the sets $\{E(S)\}_{S \in F}$ have bounded overlap in $\Sigma \times (0, \infty)$. Verification of the latter fact follows
by first noting, for each \((Z, \tau, r) \in E(S)\), that there is some cube \(Q \in S\) such that
\[
d_p(Z, \tau, Q) + \operatorname{diam} Q \leq 2Kr.
\]
Hence, upon replacing \(Q\) with a suitable ancestor in \(S\), we have
\[
d_p(Z, \tau, Q) \leq r \quad \text{and} \quad \operatorname{diam} Q \approx r.
\]
Clearly there are only a bounded number of cubes satisfying the inequality in this display for each \((Z, \tau, r)\), hence the sets \(\{E(S)\}_{S \in \mathcal{F}}\) have bounded overlap in \(\Sigma \times (0, \infty)\).

Based on this, and for the sake of obtaining a contradiction, we will from now on assume that \((6.3)\) holds and that \(S \in \mathcal{F}_1\). In the following we will at instances use the short notation \(C_p := C_p(x_{Q(S)}, t_{Q(S)}), \rho > 0\). Using this notation, the assumption in \((6.3)\) and Lemma 5.6, we can conclude the validity of the important estimate
\[
\int_0^{kR/10} \int_{C_{p, \rho}} \frac{(\hat{y}(z, t, r))^2}{r} \, \frac{dz \, dr}{r} \leq \varepsilon^2 \sigma(Q(S)).
\]

We are now going to study the implications of \((6.5)\). To do this, let \(\phi\) be an infinitely differentiable real valued function on \(\mathbb{R}^d\) with compact support in \(C'_1(0, 0)\) and set
\[
\phi_\lambda(z, \tau) \equiv \lambda^{-d} \phi(z/\lambda, \tau/\lambda^2).
\]
For \(f : \mathbb{R}^n \to \mathbb{R}\), we introduce the usual convolution of \(\phi\) with \(f\) by
\[
(\phi * f)(z, \tau) = \int_{\mathbb{R}^n} \phi(z - y, \tau - s)f(y, s) \, dy \, ds,
\]
whenever this convolution makes sense. Furthermore, we suppose that
\[
\int_{\mathbb{R}^n} \phi(y, s) \, dy \, ds = 0 \quad \text{and} \quad \int_{\mathbb{R}^n} y_i \phi(y, s) \, dy \, ds = 0 \quad \text{for} \quad 1 \leq i \leq n - 1.
\]
If \(L\) is a linear function of \(y\) only, then
\[
(\phi_\lambda * \psi)^2(z, \tau) = (\phi_\lambda * (\psi - L))^2(z, \tau)
\]
\[
\leq \lambda^{-d} \int_{C'_1(z, \tau)} |\psi - L|^2 \, dy \, ds.
\]
In particular,
\[
(\phi_\lambda * \psi)^2(z, \tau) \leq \lambda^2 (\hat{y}(z, \tau, \lambda))^2.
\]
From this observation and \((6.5)\) we see that
\[
\int_0^\rho \int_{C_{p, \rho}^*(\hat{z}, \hat{\tau})} (\phi_\lambda * \psi)^2(\hat{z}, \hat{\tau}) \, \frac{dz \, dr \, d\lambda}{\lambda^3} \leq \int_0^\rho \int_{C_{p, \rho}^*(\hat{z}, \hat{\tau})} (\hat{y}(\hat{z}, \hat{\tau}, \lambda))^2 \, \frac{dz \, dr \, d\lambda}{\lambda}
\]
\[
\leq \varepsilon^2 \sigma(Q(S))
\]
whenever \(C_{p, \rho}^*(\hat{z}, \hat{\tau}) \subset C_{kR}^*(x_{Q(S)}, t_{Q(S)})\) and \(\rho \leq kR/10\).

We will now use a parabolic Calderón reproducing formula. We here follow the construction in [H2, pp 227-228] and we let \(P_1f := k_1 * f\), where
\[
k_1(x, t) := \lambda^{-d} k(x^{-1} x, \lambda^{-2} t),
\]
and where \( k \in C_0^\infty(C_1(0,0)) \) is a non-negative even function, with
\[
\int_{\mathbb{R}^n} k \, dy \, ds = 1.
\]
Thus, \( P_\lambda \) is a nice standard parabolic approximation of the identity. Consequently,
\begin{equation}
I = - \int_0^\infty \frac{\partial}{\partial \lambda} P_\lambda^2 \, d\lambda = \int_0^\infty P_\lambda Q_\lambda \, \frac{d\lambda}{\lambda},
\end{equation}
in the strong operator topology on \( \mathcal{B}(L^2(\mathbb{R}^n)) \), where \( Q_\lambda := -2\lambda \frac{\partial P_\lambda}{\partial \lambda} \). In particular, if we let \( \phi_\lambda := -2\lambda \frac{\partial k_\lambda}{\partial \lambda} \), then, as \( k \) is even, we see that (6.7) holds for \( \phi_\lambda \), and by (6.11)
\begin{equation}
\psi(z, \tau) = \int_0^\infty \left( k_\lambda \ast \phi_\lambda \ast \psi \right)(z, \tau) \, \frac{d\lambda}{\lambda},
\end{equation}
in the sense of \( L^2 \) and with pointwise almost everywhere.

Using the representation in (6.12) we decompose \( \psi \) as
\[
\psi = \psi_1 + \psi_2, \quad \psi_1 = \psi_{11} + \psi_{12},
\]
where
\[
\psi_{11}(z, \tau) := \int_R \left( k_\lambda \ast \phi_\lambda \ast \psi \right)(z, \tau) \, \frac{d\lambda}{\lambda} = - \int_R \frac{\partial}{\partial \lambda} P_\lambda^2 \psi(z, \tau) \, d\lambda,
\]
\[
\psi_{12}(z, \tau) := \int_0^R \int_{\mathbb{R}^n \setminus C_{200R}^0} k_\lambda(z - \hat{z}, \tau - \hat{\tau}) (\phi_\lambda \ast \psi)(\hat{z}, \hat{\tau}) \, \frac{d\hat{z} \, d\hat{\tau}}{\lambda},
\]
\[
\psi_{2}(z, \tau) := \int_0^R \int_{C_{200R}^0} k_\lambda(z - \hat{z}, \tau - \hat{\tau}) (\phi_\lambda \ast \psi)(\hat{z}, \hat{\tau}) \, \frac{d\hat{z} \, d\hat{\tau}}{\lambda},
\]
for \((z, \tau) \in \mathbb{R}^n \) and we again recall that \( R = \text{diam}(Q(S)) \).

**Lemma 6.13.** The function \( \psi_1 : \mathbb{R}^n \to \mathbb{R} \) satisfies
\begin{equation}
|\psi_1(x, t) - \psi_1(y, s)| \leq \delta d_\rho(x, t, y, s),
\end{equation}
whenever \((x, t), (y, s) \in C_{100R}^r(x_Q(S), t_Q(S)) \). Furthermore,
\begin{equation}
|\nabla_x^2 \psi_1(x, t)| + |\partial_x \psi_1(x, t)| \leq \delta R^{-1} \text{ on } C_{100R}^r(x_Q(S), t_Q(S)).
\end{equation}

**Proof.** Note that if \((x, t) \in C_{100R}^r(x_Q(S), t_Q(S)) \), then \( \psi_{12}(x, t) = 0 \) by construction and hence it suffices to treat \( \psi_{11} \). Recall that by Lemma 4.26,
\begin{equation}
|\psi(x, t) - \psi(y, s)| \leq \delta d_\rho(x, t, y, s)
\end{equation}
for all \((x, t), (y, s) \in \mathbb{R}^n \). Observe that by construction,
\[
\psi_{11} = \tilde{P}_R \psi = \tilde{k}_R \ast \psi,
\]
where \( \tilde{k}_R := k_R \ast k_R \), and \( \tilde{P}_R := P_\lambda^2 \). Since \( \tilde{k}_R \) is non-negative with integral 1, we see that \( |\psi_{11}(x, t) - \psi_{11}(y, s)| \) can be bounded by
\[
\leq \int_{\mathbb{R}^n} \tilde{k}_R(z, \tau) |\psi(x - z, t - \tau) - \psi(y - z, s - \tau)| \, dz \, d\tau \leq \delta d_\rho(x, t, y, s),
\]
by (6.16) and the translation invariance of the parabolic distance $d_p$. This proves (6.14).

To prove (6.15), we first observe that

$$\left| \nabla_x^2 \tilde{k}_R(x, t) \right| + \left| \partial_t \tilde{k}_R(x, t) \right| \leq R^{-d-2} \mathbb{1}_{C^{2}_{2000}(0,0)}(x, t).$$

Moreover,

$$\int_{\mathbb{R}^n} \nabla^2 \tilde{k}_R(x, t) \, dx \, dt = 0 = \int_{\mathbb{R}^n} \partial_t \tilde{k}_R(x, t) \, dx \, dt.$$ 

Therefore $|\nabla^2 \psi_{11}(x, t)| + |\partial_t \psi_{11}(x, t)|$ can be bounded by

$$\leq R^{-d-2} \int_{\mathbb{R}^n} \mathbb{1}_{C^{2}_{2000}(0,0)}(x-y, t-s) |\psi(y, s) - \psi(x, t)| \, dx \, dt$$

$$\leq \delta R^{-d-2} \int_{\mathbb{R}^n} \mathbb{1}_{C^{2}_{2000}(0,0)}(x-y, t-s) \, d_p(x,t,y,s) \, dx \, dt \leq \delta R^{-1},$$

where in the next-to-last step we have again used (6.16). \qed

**Lemma 6.17.** The function $\psi_2: \mathbb{R}^n \to \mathbb{R}$ satisfies

$$|\psi_2(x, t) - \psi_2(y, s)| \leq \delta d_p(x, t, y, s),$$

whenever $(x, t), (y, s) \in C'_{100R}(x_{Q(s)}, t_{Q(s)})$. Furthermore,

$$\int_{\mathbb{R}^n} |\nabla_x \psi_2|^2 + |D_{1/2}^i \psi_2|^2 \, dz \, dt \leq \epsilon^2 \sigma(Q(S)).$$

**Proof.** (6.18) follows immediately from (6.14) and (6.16). To prove (6.19), let $f \in L^2(\mathbb{R}^n)$ and note that

$$\int_{\mathbb{R}^n} \partial_{\lambda_i} \psi_2(z, \tau) f(z, \tau) \, dz \, dt = - \int_0^R \int_{C^{2000}_R} ((\partial_{\lambda_i} k)_\lambda \ast f)(\phi_\lambda \ast \psi) \frac{d\xi \, d\tau \, d\lambda}{\lambda^2}.$$ 

Hence,

$$\int_{\mathbb{R}^n} |\nabla_x \psi_2|^2 \, dz \, dt \leq \int_0^R \int_{C^{2000}_R} |\phi_\lambda \ast \psi|^2 \frac{d\xi \, d\tau \, d\lambda}{\lambda^3} \leq \epsilon^2 \sigma(Q(S)),$$

by Littlewood-Paley theory and (6.10) (provided we ensure $\kappa \geq 2000$ so that $\kappa R/10 \geq 200R$). Similarly,

$$\int_{\mathbb{R}^n} D_{1/2}^i \psi_2(z, \tau) f(z, \tau) \, dz \, dt = \int_0^R \int_{C^{2000}_R} ((D_{1/2}^i k)_\lambda \ast f)(\phi_\lambda \ast \psi) \frac{d\xi \, d\tau \, d\lambda}{\lambda^2},$$

and again

$$\int_{\mathbb{R}^n} |D_{1/2}^i \psi_2|^2 \, dz \, dt \leq \int_0^R \int_{C^{2000}_R} |\phi_\lambda \ast \psi|^2 \frac{d\xi \, d\tau \, d\lambda}{\lambda^3} \leq \epsilon^2 \sigma(Q(S)),$$

where we have used that

$$\int_0^R \int_{\mathbb{R}^n} |(D_{1/2}^i k)_\lambda \ast f(\hat{\xi}, \hat{\tau})|^2 \frac{d\xi \, d\tau \, d\lambda}{\lambda} \leq \int_\mathbb{R} \int |f(\hat{z}, \hat{\tau})|^2 \, d\hat{\xi} \, d\hat{\tau}.$$ 

In turn, the latter estimate follows readily via Plancherel’s theorem in $\mathbb{R}^n$, using that the half time derivative corresponds to the Fourier multiplier $|\tau|^{1/2}$ as follows.
First, the equivalence between our notion of half-time derivative and the Fourier multiplier one can be found, for instance, in [DPV] (see in particular the beginning of Section 3 and Proposition 3.3 there).

Recall that \( k \in C^\infty_0(C_1'(0,0)) \). Taking the Fourier transform in \((\hat{\xi}, \hat{\tau})\) and using Plancherel’s formula we have, using \((\xi, \zeta) \in \mathbb{R}^{n-1} \times \mathbb{R}\) for the dual variables,

\[
\left\langle (D_{1/2}^1k)_\lambda \ast f(\hat{\xi}, \hat{\tau}) \right\rangle^2 \leq \int_{\mathbb{R}^n} |\lambda^2 \xi|^2 |\hat{k}(\lambda \xi, \lambda^2 \zeta)\hat{f}(\xi, \zeta)|^2 \, d\xi \, d\zeta
\]

\[
\leq \left\langle \int_{\mathbb{R}^n} \min(\lambda^2|\xi|, (\lambda^2|\zeta|)^{-2}) |\hat{f}(\xi, \zeta)|^2 \, d\xi \, d\zeta \right\rangle^2
\]

where we used the fact that \( \hat{k} \) is a Schwartz function in the last inequality; more specifically we used the bounds

\[
|\hat{k}(\xi, \zeta)| \leq \min\{1, |\xi|^{-3/2}\}, \quad (\xi, \zeta) \in \mathbb{R}^n.
\]

Using the previous estimate and integrating in \( \lambda \) we arrive at

\[
\int_0^\infty \int_{\mathbb{R}^n} |(D_{1/2}^1k)_\lambda \ast f(\hat{\xi}, \hat{\tau})|^2 \frac{d\xi \, d\zeta}{\lambda} \leq \int_{\mathbb{R}^n} |\hat{f}(\xi, \zeta)|^2 \left[ \int_0^\infty \min(\lambda^2|\xi|, (\lambda^2|\zeta|)^{-2}) \frac{d\lambda}{\lambda} \right] \, d\xi \, d\zeta
\]

\[
= \int_{\mathbb{R}^n} |\hat{f}(\xi, \zeta)|^2 \, d\xi \, d\zeta.
\]

\( \Box \)

We next prove two lemmas concerning the oscillations of \( \psi_2 \) and \( \psi \).

**Lemma 6.20.** Consider \( C_\lambda'(z, \tau) \subset \mathbb{R}^n \), let

\[
m_{C_\lambda'(z, \tau)} \psi_2 := \int_{C_\lambda'(z, \tau)} \psi_2 \, d\hat{\xi} \, d\hat{\tau},
\]

and

\[
osc_{C_\lambda'(z, \tau)} \psi_2 := \sup_{(\hat{\xi}, \hat{\tau}) \in \partial C_\lambda'(z, \tau)} \left| \psi_2(\hat{\xi}, \hat{\tau}) - m_{C_\lambda'(z, \tau)} \psi_2 \right|.
\]

Assume that \( C_\lambda'(z, \tau) \subset C_{100u}(x_Q(S), t_Q(S)) \). Then

\[
osc_{C_\lambda'(z, \tau)} \psi_2 \leq r \left\{ \frac{1}{r} \int_{C_\lambda'(z, \tau)} \left| \psi_2(\hat{\xi}, \hat{\tau}) - m_{C_\lambda'(z, \tau)} \psi_2 \right| \, d\hat{\xi} \, d\hat{\tau} \right\}^{1/(d+1)} \delta^{-d/(d+1)}.
\]

**Proof.** Let \((\hat{\xi}, \hat{\tau}) \in \partial C_\lambda'(z, \tau)\) be such that

\[
\left| \psi_2(\hat{\xi}, \hat{\tau}) - m_{C_\lambda'(z, \tau)} \psi_2 \right| = osc_{C_\lambda'(z, \tau)} \psi_2 =: \mu
\]

Using Lemma 6.13 and Lemma 6.17 we note that

\[
(6.21) \quad |\psi_2(\hat{\xi}, \hat{\tau}) - \psi_2(\hat{\xi}, \hat{\tau})| \leq \beta_0 d_p(\hat{\xi}, \hat{\tau}, \hat{\xi}, \hat{\tau}),
\]
whenever \((\hat{z}, \hat{\tau}), (\tilde{z}, \tilde{\tau}) \in C^{' \prime}_{(z, \tau)}\), and for every function \(\tilde{\psi}\) in the set \(\{\psi, \psi_1, \psi_2\}\). Here \(\beta\) is the implicit constant in (6.18). If \((\tilde{z}, \tilde{\tau}) \in C^{' \prime}_{z, \tau}\) is such that \(d_{\beta}(\hat{z}, \hat{\tau}, \tilde{z}, \tilde{\tau}) \leq \mu/(2\beta\delta)\), then

\[
|\psi_2(\hat{z}, \hat{\tau}) - \psi_2(\tilde{z}, \tilde{\tau})| \leq \mu/2,
\]

by (6.21) and hence

\[
|\psi_2(\hat{z}, \hat{\tau}) - m_{C^{' \prime}_{z, \tau}}\psi_2| \geq \mu/2,
\]

by the definition of \(\mu\) and \((\hat{z}, \hat{\tau})\).

Assume that \(\mu/(2\beta\delta) \leq r\). Then, using (6.23) we see that

\[
\int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\hat{z}, \hat{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\hat{z} \, d\hat{\tau} \geq c\mu(\mu/(\beta\delta))^d = c\mu^{d+1}\beta^{-d}\delta^{-d},
\]

for a constant \(c = c(n)\). Hence, in this case, with implicit constants depending only on \(n\), we have

\[
\mu^{d+1} \leq \beta^d\delta^d \int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\hat{z}, \hat{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\hat{z} \, d\hat{\tau}.
\]

This completes the proof of the lemma in this case.

Assume that \(\mu/(2\beta\delta) \geq r\). In this case, using (6.21) and (6.22) we see that

\[
|\psi_2(\tilde{z}, \tilde{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \geq \mu/2
\]

on an ample portion of \(C^{' \prime}_{(z, \tau)}\) and hence, for \(c = c(n)\), we have

\[
\int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\tilde{z} \, d\tilde{\tau} \geq c^{-1}\mu.
\]

Moreover, using (6.21), we also have

\[
r^{-1} \int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\tilde{z} \, d\tilde{\tau} \leq \beta\delta.
\]

In particular, combining (6.25) and (6.26) we deduce

\[
\mu \leq r \left\{ r^{-1} \int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\tilde{z} \, d\tilde{\tau} \right\}^{1/(d+1)} \leq r \left\{ r^{-1} \int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\tilde{z} \, d\tilde{\tau} \right\}^{1/(d+1)},
\]

where the various implicit constants are purely dimensional, i.e. depend only on \(n\). The conclusion of the lemma follows also in this case. □

**Lemma 6.28.** Let

\[
N(\psi_2)(\tilde{z}, \tilde{\tau}) := \sup_{C^{' \prime}_{(z, \tau)}(\tilde{z}, \tilde{\tau})} \left\{ r^{-1} \int\int_{C^{' \prime}_{(z, \tau)}} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C^{' \prime}_{(z, \tau)}}\psi_2| \, d\tilde{z} \, d\tilde{\tau} \right\}.
\]
Let $\theta \in (0, 1)$ be a degree of freedom and let
\[
\tilde{F} := \{(\tilde{z}, \tilde{\tau}) \in C_{S0R}^\prime(x_{Q(S)}, t_{Q(S)}) : N(\psi_2(\tilde{z}, \tilde{\tau}) \leq \theta^{d+1} \delta)\}.
\]
Let $\eta \in (0, 1)$, consider $(z_0, \tau_0) \in \mathbb{R}^n$ and assume that $\tilde{F} \cap C'_r(z_0, \tau_0) \neq \emptyset$ for some $r \leq \eta R$. Then
\[
\sup_{(z, \tau) \in C_r'(z_0, \tau_0)} |\psi(z, \tau) - \psi(z_0, \tau_0) - \nabla_z \psi_1(z_0, \tau_0) \cdot (z - z_0)| \leq (\theta + \eta)\delta r.
\]

**Proof.** Let $(z, \tau) \in C_r'(z_0, \tau_0)$. Then
\[
|\psi(z, \tau) - \psi(z_0, \tau_0) - \nabla_z \psi_1(z_0, \tau_0) \cdot (z - z_0)|
\leq |\psi_2(z, \tau) - \psi_2(z_0, \tau_0)| + |\psi_1(z, \tau) - \psi_1(z_0, \tau_0) - \nabla_z \psi_1(z_0, \tau_0) \cdot (z - z_0)|
\leq 2 \text{osc}_{C_r'(z_0, \tau_0)} \psi_2 + |\psi_1(z, \tau) - \psi_1(z_0, \tau_0) - \nabla_z \psi_1(z_0, \tau_0) \cdot (z - z_0)|.
\]
Using Lemma 6.13, and in particular (6.15)), and Taylor’s Theorem, we see that
\[
|\psi_1(z, \tau) - \psi_1(z_0, \tau_0) - \nabla_z \psi_1(z_0, \tau_0) \cdot (z - z_0)| \leq \delta R^{-1} r^2 \leq \delta \eta r,
\]
by our assumption that $r \leq \eta R$. Furthermore, as $\tilde{F} \cap C'_r(z_0, \tau_0) \neq \emptyset$ we can apply Lemma 6.20 to conclude that
\[
\text{osc}_{C_r'(z_0, \tau_0)} \psi_2 \leq \theta \delta r.
\]
Combining the latter estimate with (6.30), we obtain (6.29). \qed

**Lemma 6.31.** Let $N(\psi_2), \theta$ and $\tilde{F}$ be defined as in Lemma 6.28. Then
\[
\mathcal{H}_p^{n+1}(C_{S0R}^\prime(x_{Q(S)}, t_{Q(S)}) \setminus \tilde{F}) \leq \delta^{-2} \theta^{-2(d+1)} e^2 \sigma(Q(S)).
\]

**Proof.** Note that
\[
\mathcal{H}_p^{n+1}(C_{S0R}^\prime(x_{Q(S)}, t_{Q(S)}) \setminus \tilde{F}) \leq \delta^{-2} \theta^{-2(d+1)} \int_{\mathbb{R}^n} |N(\psi_2)(\tilde{z}, \tilde{\tau})|^2 \, d\tilde{z} \, d\tilde{\tau}.
\]
To control $N(\psi_2)(\tilde{z}, \tilde{\tau})$ we need to use a Poincaré inequality. To this end, for locally integrable $g : \mathbb{R}^n \to \mathbb{R}$, we let $M(g)$ be the $n$-dimensional parabolic Hardy-Littlewood maximal function
\[
M(g)(x, t) := \sup_{\rho > 0} \frac{1}{C_r^\prime(x, t)} \int_{C_r^\prime(x, t)} |g| \, dy \, ds,
\]
and we let $M_x$ and $M_t$ denote (respectively) the standard Hardy-Littlewood maximal operators in the $x$ and $t$ variables only. Consider $(\tilde{z}, \tilde{\tau}) \in \mathbb{R}^n$ and assume that $(\tilde{z}, \tilde{\tau}) \in C_r^\prime(z, \tau)$. Then $C_r^\prime(z, \tau) \subset C_{2r}^\prime(\tilde{z}, \tilde{\tau})$ and
\[
\int_{C_r^\prime(z, \tau)} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C_r^\prime(z, \tau)} \psi_2| \, d\tilde{z} \, d\tilde{\tau}
\leq r^{-1} \int_{C_{2r}^\prime(\tilde{z}, \tilde{\tau})} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C_{2r}^\prime(\tilde{z}, \tilde{\tau})} \psi_2| \, d\tilde{z} \, d\tilde{\tau}.
\]
Using Lemma 2.2 in [AEN] we see that
\[
\int_{C_{2r}^\prime(\tilde{z}, \tilde{\tau})} |\psi_2(\tilde{z}, \tilde{\tau}) - m_{C_{2r}^\prime(\tilde{z}, \tilde{\tau})} \psi_2| \, d\tilde{z} \, d\tilde{\tau}
\leq \delta^{-2} \theta^{-2(d+1)} e^2 \sigma(Q(S)).
\]
Lemma 6.37.  
from which the result follows.

We have
\[ \mathcal{M}(|\nabla_x \psi_2|)(\tilde{x}, \tilde{\tau}) + \mathcal{M}_d(|H_t D_t^{1/2} \psi_2|)(\tilde{x}, \tilde{\tau}). \]

In particular,
\[ N(\psi_2)(\tilde{x}, \tilde{\tau}) \leq \mathcal{M}(|\nabla_x \psi_2|)(\tilde{x}, \tilde{\tau}) + \mathcal{M}_d(|H_t D_t^{1/2} \psi_2|)(\tilde{x}, \tilde{\tau}). \]

Hence,
\[ \mathcal{H}^{n+1}(\delta) \leq \delta^{-2} \theta^{-2(d+1)} \int_{\mathbb{R}^n} |\nabla_x \psi_2|^2 + |D_t^{1/2} \psi_2|^2 \, d\zeta \, d\tau \]
\[ \leq \delta^{-2} \theta^{-2(d+1)} \tilde{\varepsilon}^2 \sigma(Q(S)), \]
by Lemma 6.17.

Let \( N(\psi_2) \), \( \theta \) and \( \tilde{F} \) be defined as in Lemma 6.28. Let \( \eta \in (0, 1) \), consider \((z_0, \tau_0) \in \mathbb{R}^n\), and assume that \( \tilde{F} \cap C_{z_0, \tau_0} \neq \emptyset \) for some \( r \leq \eta R \). By Lemma 6.28 we have
\[ \sup_{(z, \tau) \in C_{z_0, \tau_0}} \frac{1}{diam(\tilde{F})} \leq \frac{1}{(\theta + \delta) r}. \]

Letting \( \theta := \eta \) we can conclude that
\[ \sup_{(z, \tau) \in C_{z_0, \tau_0}} |\psi(z, \tau) - \psi(z_0, \tau_0) - \nabla_x \psi_1(z_0, \tau_0) \cdot (z - z_0)| \leq (\theta + \delta) r. \]

(6.33)

Let \( \tilde{P} = \tilde{P}_{C_{z_0, \tau_0}} \) be the time-independent plane which is the graph of the affine function
\[ L(z) := \psi(z_0, \tau_0) - \nabla_x \psi_1(z_0, \tau_0) \cdot (z - z_0). \]

Then by (6.33) we have
\[ \sup_{(Z, \tau) \in \Sigma(\tilde{P})} r^{-1} d_p(Z, \tau, \tilde{P}) \leq \eta \delta. \]

Recall that the we are seeking a contradiction to the assumption that (6.3) holds for some \( S \in F_1 \). We will achieve this by proving that the conclusion in (6.34) is incompatible with \( S \in F_1 \). Next, we prove two additional auxiliary lemmas to be used in the final argument.

Lemma 6.35. Let \( \eta, 0 < \eta \ll 1 \) and \( A \gg 1 \) be given. Then there exists \( \epsilon_0 = \epsilon_0(n, M, \eta, A) > 0 \) such that if \( \epsilon \leq \epsilon_0 \), then the angle between \( P_Q \) and \( P_{Q(S)} \) is \( \leq \delta/100 \) for all \( Q \in S \) with \( diam(Q) \geq \eta R/A \).

Proof. Consider \( Q \in S \), let \( Q_0 = Q \) and let \( Q_1, \ldots, Q_M \) denote the successive ancestors of \( Q \) with \( Q_M = Q(S) \). By definition, see (3.9), we have
\[ d_p(Y, s, P_{Q_i}) \leq \epsilon \text{diam}(Q_i) \text{ for all } (Y, s) \in 8KQ_i \text{ and for all } i \in \{0, \ldots, M\}. \]
(6.36)
This implies, see Lemma A.2, that the angle between \( P_{Q_i} \) and \( P_{Q_{i+1}} \) is bounded by \( \leq \epsilon \). In particular, the angle between \( P_Q \) and \( P_{Q(S)} \) is bounded by \( \leq \epsilon M \leq \epsilon \log(2R/ \text{diam}(Q)) \), from which the result follows.

Lemma 6.37. If \( Q \in m_1(S) \), then \( d_p(\tilde{F}, \pi(Q)) > \text{diam}(Q) \).
\textbf{Proof.} Assume, for the sake of obtaining a contradiction, that there exists \( Q \in m_1(S) \) such that \( d_p(\bar{F}, \pi(Q)) \leq \text{diam}(Q) \). Let \((X_Q, t_Q)\) be the center of \( Q \), let \((z_0, \tau_0) := (x_Q, t_Q) := \pi(X_Q, t_Q), \ r = 10 \text{diam}(Q) \). Consider \( C'(z_0, \tau_0) \). By the assumption and the construction \( \bar{F} \cap C'(z_0, \tau_0) \neq \emptyset \). Recall also that \( m_1(S) \) denotes the cubes in \( m(S) \) for which the angle between \( P_Q \) and \( P_{Q(S)} \) is greater or equal to \( \delta/2 \). Hence, by Lemma 6.35 we can conclude that \( r \leq \eta \delta \) if \( \epsilon \) sufficiently small. Furthermore, \( C'(z_0, \tau_0) \leq C'_{100\delta}(x_Q(S), t_Q(S)) \). By Lemma 4.41, for \((Z, \tau) \in Q\) we have
\begin{equation}
(6.38) \quad d_p(Z, \tau, (z, \psi(z, \tau), \tau)) \leq \epsilon d(z, \tau) \leq \epsilon \text{diam}(Q), \ (z, \tau) = \pi(Z, \tau).
\end{equation}
Furthermore, by (6.34) we also know that
\begin{equation}
(6.39) \quad d_p(Z, \tau, \bar{P}) \leq \eta \delta r,
\end{equation}
for a time independent plane \( \bar{P} \in \mathcal{P} \). The inequalities in the last two displays imply that if \((Z, \tau) \in Q\), and if \( \epsilon \leq \eta \delta \), then
\begin{equation}
(6.40) \quad d_p(Z, \tau, \bar{P}) \leq \eta \delta r.
\end{equation}
Upon observing that we can produce \( n + 1 \) points of \( Q \) with ample distance from each other using Lemma A.1, the estimate (6.39) implies that the angle between \( \bar{P} \) and \( P_Q \) is \( \leq \eta \delta < \delta/100 \) if we choose \( \eta \) and \( \epsilon \) small. Let \( Q^* \) be the largest ancestor to \( Q \) such that \( 10 \text{diam}(Q^*) \leq \eta \delta R \). The same argument as above then gives that the angle between \( \bar{P} \) and \( P_Q^* \) is \( \leq \delta/100 \). In particular, the angle between \( P_Q^* \) and \( P_Q \) must be \( \leq \delta/50 \). On the other hand, we may apply Lemma 6.35 to \( Q^* \) to deduce that the angle between \( P_Q^* \) and \( P_{Q(S)} \) is at most \( \delta/100 \), and thus that the angle between \( P_Q \) and \( P_{Q(S)} \) is at most \( \delta/33 \). This is incompatible with \( Q \in m_1(S) \) and hence we have reached a contradiction. \( \square \)

We are now ready finally to complete the proof of Theorem 3.1 by proving that \( S \notin \mathcal{F}_1 \). We see that by Lemma 6.37 it suffices to prove that (6.3) implies that there is at least one \( Q \in m_1(S) \) with \( d_p(\bar{F}, \pi(Q)) \leq \text{diam}(Q) \) where \( \bar{F} \) is defined as in Lemma 6.28. Let
\begin{equation}
(6.41) \quad \mathbf{F} = \bigcup_{Q \in m_1(S)} Q.
\end{equation}
There exists a collection \( S' \subseteq m_1(S) \) such that \( \mathbf{F} \) can be covered by
\begin{equation}
\{C_{20 \text{diam}(Q)}(Z_Q, \tau_Q)\}_{Q \in S'}
\end{equation}
in such a way that
\begin{equation}
\{C_{3 \text{diam}(Q)}(Z_Q, \tau_Q)\}_{Q \in S'}
\end{equation}
is a disjoint collection. Hence,
\begin{equation}
\sigma(\mathbf{F}) \leq \sigma\left(\bigcup_{Q \in S'} C_{20 \text{diam}(Q)}(Z_Q, \tau_Q) \cap \Sigma\right) \leq \sum_{Q \in S'} (\text{diam}(Q))^d.
\end{equation}
Using this fact, and our usual notation \((z_Q, \tau_Q) = \pi(Z_Q, \tau_Q)\), we can use by proof of Lemma 4.2 (specifically, the fact that (4.4) holds whenever \((X, t), (Y, s) \in 16\kappa(Q(S)\))
and \( \min(d(X, t), d(Y, s)) \leq 1000d_p(X, t, Y, s) \), to conclude that the collection

\[
\{C_{\text{diam}}(Q)(z_Q, \tau_Q)\}_{Q \in S'}
\]

is also pairwise disjoint. Hence

\[
(6.42) \quad \mathcal{H}^{p+1}_p \left( \bigcup_{Q \in S'} C_{\text{diam}}(Q)(z_Q, \tau_Q) \right) \geq \sigma(F).
\]

Moreover, every set in \( \{C'_{\text{diam}}(Q)(z_Q, \tau_Q)\}_{Q \in S'} \) is contained in \( C'_{50\rho}(x_Q(S), t_Q(S)) \) and by Lemma 6.37, none of the sets in the collection \( \{C'_{\text{diam}}(Q)(z_Q, \tau_Q)\}_{Q \in S'} \) intersect \( F \). Hence, using this observation and (6.42), we see that

\[
(6.43) \quad \sigma(F) \leq \mathcal{H}^{p+1}_p \left( \bigcup_{Q \in F'} C_{\text{diam}}(Q)(z_Q, \tau_Q) \right) \leq \mathcal{H}^{p+1}_p (C_{50\rho}(x_Q(S), t_Q(S)) \setminus \tilde{F}).
\]

Using Lemma 6.31 with \( \theta = \eta \) we have

\[
(6.44) \quad \sigma(F) \leq \delta^{-2} \theta^{-2(d+1)} \epsilon^2 \sigma(Q(S)) \approx \delta^{-2} \eta^{-2(d+1)} \epsilon^2 \sigma(Q(S)).
\]

Hence,

\[
(6.45) \quad \sigma(F) \leq c \delta^{-2} \eta^{-2(d+1)} \epsilon^2 \sigma(Q(S)),
\]

for some \( c = c(n, M, \Gamma, \kappa, K) \geq 1 \). Given \( \delta \) and \( \eta \), we let \( \epsilon_1 := \delta \eta^{d+1}/\sqrt{4c} \). Let also \( \epsilon_0 = \epsilon_0(n, M, \eta, A) > 0 \), with \( A \gg 1 \) fixed, be as in the statement of Lemma 6.35. We now choose

\[
(6.46) \quad \epsilon := \min\{\epsilon_0, \epsilon_1\}.
\]

Then \( \epsilon = \epsilon(n, M, \Gamma, \kappa, K, \delta, \eta, A) \), and in fact, as \( \eta \) and \( A \) can be chosen as constants with no particular dependencies, \( \epsilon \) can be seen simply as a function of \( n, M, \Gamma, K, \kappa, \) and \( \delta \). By the choice for \( \epsilon \),

\[
(6.47) \quad \sigma(F) \leq \sigma(Q(S))/4.
\]

The estimate (6.47) contradicts the statement that \( S \in \mathcal{F}_1 \). As we had reduced matters to showing \( S \notin \mathcal{F}_1 \), we have proved Theorem 3.1.

7. The Proof of Theorem 3.3

In this section we prove Theorem 3.3 by improving Theorem 3.1(4) to give a bilateral approximation as in Theorem 3.3(4). The latter is a parabolic version of a result appearing in [HMM1]. The most important observation here is that in the proof of Theorem 3.1 the condition (3.10) (B), for the stopping time regimes, was only used to pack the maximal cubes of the stopping time regimes (Section 6). In particular, the construction of the graphs requires only property (3.10) (A) of the stopping time regimes (and the fact that \( \Sigma \) is p-UR). With this in mind, we are going to use the packing of the stopping time regimes in Theorem 3.1 to prove the packing of stopping time regimes where (3.10) (A) still holds, along with a stronger, bilateral approximation by a \( t \)-independent plane. Then we can run the machine that produces the regular Lip(1,1/2) graphs and never carry out the arguments of Section 6 (since the stopping time regimes will already pack) and we
will use this bilateral approximation by planes to improve the approximation by graphs to a bilateral one.

**Definition 7.1.** Let \( \Sigma \) be a parabolic Ahlfors-David Regular set with dyadic grid \( \mathcal{D}(\Sigma) \). For \( Q \in \mathcal{D}(\Sigma) \) and \( K > 1 \) we let \( b_{\beta_\infty}(Q) := b_{\beta_\infty}(Q, K) \) denote the number

\[
\inf_{P \in \mathcal{P}} \frac{\text{diam}(Q)}{\text{diam}(Q)^{-1}} \left[ \sup_{\{(Y,s) \in \mathcal{S}\}} d_P(Y,s,P) + \sup_{\{(X,t) \in \mathcal{C}\}} d_P(X,t,\Sigma) \right].
\]

Given \( \epsilon > 0, K > 1 \), we say that the \((\epsilon,K)\)-parabolic bilateral weak geometric lemma \(((\epsilon,K)\text{-PBWGL})\) holds for \( \Sigma \), if there exists a constant \( c(\epsilon,K) \) such that

\[
\sum_{\mathcal{D}(\Sigma) : b_{\beta_\infty}(Q,K) \geq \epsilon} \sigma(Q) \leq c(\epsilon,K)\sigma(Q^*), \quad \forall Q^* \in \mathcal{D}(\Sigma).
\]

Combining Theorem 3.1 with our previous work [BHHLN1] we immediately obtain the following 13.

**Theorem 7.2 ([BHHLN1, Theorem 4.15]).** Let \( \Sigma \) be a parabolic Ahlfors-David regular set. If \( \Sigma \) is parabolic UR with constants \( M \) and \( \Gamma \), then \( \Sigma \) satisfies the \((\epsilon,K)\text{-PBWGL} \) for every \( \epsilon \in (0,1] \) and \( K > 1 \).

The reader may notice that the \( \beta \) numbers used here are defined slightly differently compared to [BHHLN1], and that the (dilation) parameter \( K \) does not appear in Theorems 4.15 [BHHLN1]. However, these differences represent no problems, see [BHHLN1, Remark 2.24]. In particular, it turns out that \((\epsilon,K)\text{-PBWGL}\) for some fixed \( K \) and all \( \epsilon > 0 \) implies that the \((\epsilon,K)\text{-PBWGL}\) for all \( K \) and all \( \epsilon > 0 \), so that we may use the result of [BHHLN1] to first prove the \((\epsilon,2)\text{-PBWGL}\) for all \( \epsilon > 0 \) and obtain \((\epsilon,K)\text{-PBWGL}\) for all \( K \) and all \( \epsilon > 0 \). This improvement of \( K \) be proved with techniques in [DS2, I. 3.2] specifically using Lemma 3.27 in part I.

The following lemma will be important in the proof of Theorem 3.3.

**Lemma 7.3.** Suppose \( \Sigma \) is a parabolic UR set with parameters \( M \) and \( \Gamma \). Let \( \epsilon \ll \delta \) and \( K \gg \kappa \) be as in the proof of Theorem 3.1. If we define

\[
\mathcal{G}' := \{ Q \in \mathcal{D}(\Sigma) : b_{\beta_\infty}(Q,K) < \epsilon \},
\]

\[
\mathcal{B}' := \{ Q \in \mathcal{D}(\Sigma) : b_{\beta_\infty}(Q,K) \geq \epsilon \},
\]

then \( \mathcal{D}(\Sigma) = \mathcal{G}' \cup \mathcal{B}' \), and the following hold.

1. The cubes \( \mathcal{G}' \) can be decomposed into disjoint, stopping time regimes \( \{ S^* \}_{S^* \in \mathcal{F}^*} \), such that each \( S^* \) is coherent and

\[
\text{Angle}(P_Q, P_{Q(S^*)}) \leq 4\delta, \quad \forall Q \in S^*,
\]

where \( P_Q \in \mathcal{P} \) is the plane minimizing the quantity in the definition of \( b_{\beta_\infty}(Q,K) \).

---

13 The notation GPG in [BHHLN1] means ‘good parabolic graph’, that is, a regular Lip(1,1/2) graph.
with the decomposition in (1), the ‘bad’ cubes $B^*$ and the maximal cubes $\{Q(S^*)\}_{S \in F^*}$ satisfy a packing condition
\[
\sum_{Q \in B^*} \sigma(Q) + \sum_{Q : Q(S) \subseteq Q^*} \sigma(Q(S)) \leq c(\varepsilon, K) \sigma(Q^*), \quad \forall Q^* \in \mathcal{D}(\Sigma).
\]

Proof. The packing for the cubes $B^*$ is a consequence of Lemma 7.2. By the proof of Theorem 3.1 the decomposition above holds with $G'(\varepsilon, K)$ replaced by
\[
G = G(\varepsilon, K) = \{Q \in \mathcal{D}(\Sigma) : \beta(\varepsilon, K, Q) < \varepsilon\},
\]
and with $B^*$ replaced by the complement of $G$ in $\mathcal{D}(\Sigma)$. In particular, there exists a decomposition of $G$ into stopping time regimes $\{S\}_{S \in F}$ such that
\[
\text{Angle}(P_Q, P_{Q(S)}) \leq \delta, \quad \forall Q \in S,
\]
where $P_Q$ is the plane that minimizes $\beta(\varepsilon, K, Q)$. Now this implies that
\[
\text{Angle}(P_Q, P_R) \leq 2\delta, \quad \forall Q, R \in S.
\]
Moreover, since $\varepsilon \ll \delta$, Lemma A.2 implies that
\[
\text{Angle}(P_Q, P_R) \leq \delta, \quad \forall Q, R \in S \cap G^*.
\]
Now to prove the lemma, we follow the most natural course of action and simply check that for each $S$, we can decompose $S \cap G^*$ into stopping time regimes $\{S^*\}_{F_S}$ in a way that maintains the packing condition. This can be done exactly as in [HMM1, Lemma 2.2]. We describe the process of partitioning $S \cap G^*$, and we leave the verification of the packing condition to the interested reader. Given $S \in F$, we let $M_S$ be the set of $Q \in S \cap G^*$ for which either $Q = Q(S)$ or else the dyadic parent of $Q$ or one of its siblings belongs to $B^*$. For each $Q \in M_S$ we form a new stopping time regime $S^* = S^*(Q)$ as follows. Set $Q(S^*) = Q$ and ‘divide’ $Q(S^*)$ ‘dyadically’ until we reach a subcube $Q'$ such that $Q' \notin S$, or else $Q'$ or one of its siblings belongs to $B^*$. In any such scenario, $Q'$ and all of its siblings are omitted from $S^*$ and $Q'$ is a minimal cube of $S^*$. It then follows that $S^* \subseteq S \cap G^*$ and that $S^*$ is coherent. Moreover,
\[
S \cap G^* = \bigcup_{Q \in M_S} S^*(Q),
\]
and since $G = \cup_{S \in F} S$ and $G^* \subseteq G$, it holds that
\[
G^* = \bigcup_{F^*} S^*,
\]
where $F^* = \{S^*(Q)\}_{Q \in M_S, S \in F}$. This concludes the proof of the lemma. \hfill $\Box$

Proof of Theorem 3.3. We will prove Theorem 3.3 with $\delta$ in (4) replaced by $2\delta$. We begin, as in Theorem 3.1, by considering $\varepsilon \ll \delta$ and $K \gg \kappa$, and compared to Theorem 3.1, a difference is that we now work with the stopping time regimes $\{S^*\}_{S \in F^*}$, produced in Lemma 7.3. Here we have no intention, or need, to carry out the analysis in Section 6, and we can repeat the arguments verbatim up until that point. Since we already have proven the packing condition for the maximal cubes

\footnote{This verification is carried out in detail in [HMM1].}
in Lemma 7.3, all conclusions of Theorem 3.3 hold with the potential exception of the estimate
\[
\sup_{(Y, s) \in C_{s, \text{diam}(Q)} \cap \Sigma_S} d_p(Y, s, \Sigma) \leq 2\delta \text{diam}(Q), \quad \forall Q \in S^*.
\]
Hence, to prove Theorem 3.3 we have to prove (7.4). We continue to use the notation \(P_Q\) of Lemma 7.3. Fix \(Q \in S^*\) and \((Y, s) \in C_{s, \text{diam}(Q)} \cap \Sigma_S\). Since \(Q \in S^*\) we have that \(P_Q\) satisfies
\[
\text{diam}(Q)^{-1} \left[ \sup_{(Y, s) \in \mathbb{C} Q} d_p(Y, s, P_Q) + \sup_{(X, t) \in C_{s, \text{diam}(Q)} \cap \Sigma_S} d_p(X, t, \Sigma) \right] < \epsilon,
\]
and
\[
\text{Angle}(P_Q, P_{Q(S^*)}) \leq 4\delta.
\]
We assume, by translation and a change of coordinates, that \(P_Q = \mathbb{R}^{n-1} \times \{0\} \times \mathbb{R}\).
Recalling that \(\Sigma_S\) is the graph of a Lip(1,1/2) function with Lip(1,1/2) constant \(\leq \delta\),
with respect to the coordinates induced by \(P_Q\), it follows, if we choose \(\delta\)
to be sufficiently small, that \(\Sigma_S\) is the graph of a Lip(1,1/2) function \(\tilde{\psi} : \mathbb{R}^n \to \mathbb{R}\)
with Lip(1,1/2) constant \(\leq \delta\). In particular, if \((x, 0, t), (y, 0, t) \in P_Q\) then
\[
d_p((x, \tilde{\psi}(x, t), t), (y, \tilde{\psi}(y, s), t)) \leq c(1 + \delta)d_p(x, t, y, s).
\]
Now let \((y, 0, s)\) be the projection of \((Y, s)\) into \(P_Q\). Then, \((y, 0, s) \in C_{s, \text{diam}(Q)}(X, t) \cap P_Q\) as \(b\beta_{so}(Q, K) < \epsilon\). Again using that \(b\beta_{so}(Q, K) < \epsilon\),
we can conclude that there exists \((X, t) \in \Sigma \cap (7/8)\kappa Q \subseteq 2\kappa Q(S^*)\) such that
\[
d_p(X, t, (y, 0, s)) < \epsilon \text{diam}(Q),
\]
and
\[
d_p(X, t, P_Q) = d_p((X, t), (x, 0, t)) < \epsilon \text{diam}(Q).
\]
The inequalities (7.6) and (7.7) imply
\[
d_p((x, 0, t), (y, 0, s)) \leq 2\epsilon \text{diam}(Q),
\]
and hence, by (7.5),
\[
d_p((x, \tilde{\psi}(x, t), t), (Y, s)) = d_p((x, \tilde{\psi}(x, t), t), (y, \tilde{\psi}(y, s), t)) \leq c(1 + \delta)2\epsilon \text{diam}(Q) \leq c\epsilon \text{diam}(Q).
\]
By Lemma 4.41, which holds by construction, we have
\[
d_p((X, t), (x, \tilde{\psi}(x, t), t)) \leq 2d_p(X, t, \Sigma_S) \leq c\epsilon \text{diam}(Q),
\]
where we have used that \(\tilde{\psi}(x, t)\) is Lip(1,1/2) with small constant. Assuming that \(\epsilon \ll \delta/\kappa\),
the inequality in (7.9) yields
\[
d_p((X, t), (x, \tilde{\psi}(x, t), t)) \leq \delta \text{diam}(Q).
\]
Combining this inequality with (7.8), we find that
\[
d_p(X, t, Y, s) \leq d_p((X, t), (x, \tilde{\psi}(x, t), t)) + d_p((x, \tilde{\psi}(x, t), t), (Y, s)) \leq \delta + c\epsilon < 2\delta.
\]
This proves (7.4) and Theorem 3.3. \(\square\)
Appendix A. Two lemmas concerning approximating planes

In this appendix we prove two auxiliary and technical lemmas on approximating hyperplanes.

**Lemma A.1.** Let $\Sigma \subset \mathbb{R}^{n+1}$ be a parabolic Ahlfors-David regular set with dyadic cubes $\mathcal{D}(\Sigma)$. Then there exists a constant $A = A(n, M) \gg 1$ such that the following holds. Let $Q \in \mathcal{D} = \mathcal{D}(\Sigma)$. Suppose that $(Z_0, \tau_0) \in Q$. Then there exist points $(Z_i, \tau_i) \in Q$ for $i = 1, \ldots, n$, such that if we let, for $j = 1, \ldots, n$, $L_{j-1}$ be the spatial $(j-1)$-dimensional plane which passes through $Z_0, Z_1, \ldots, Z_{j-1}$, then

$$d_p(Z_j, \tau_j, L_{j-1} \times \mathbb{R}) \geq A^{-1} \text{diam}(Q)$$

for $j = 1, \ldots, n$. Moreover, the same result holds if $Q$ is replaced by $\lambda Q$ for some $\lambda \geq 1$, with $A$ now depending on $\lambda$ as well.

**Proof.** We only prove the result for $Q$ as the same proof works for any dilate $\lambda Q$ of $Q$, $\lambda \geq 1$. We will prove the lemma by induction on $j$ for $j = 1, \ldots, n$. Let $L_0 := \{Z_0\}$. Then $L_0$ is a 0-dimensional plane. Assume that $d_p(Z, \tau, L_0 \times \mathbb{R}) \leq A^{-1} \text{diam}(Q)$ for all $(Z, \tau) \in Q$. Then $Q$ can be covered by roughly $A^2$ cubes of size $2A^{-1} \text{diam}(Q)$. Hence

$$(\text{diam}(Q))^{n+1} \leq \sigma(Q) \leq A^3 (A^{-1} \text{diam}(Q))^{n+1}$$

which is impossible if $A$ is large enough. Hence there exists a point $(Z_1, \tau_1) \in Q$ such that $d_p(Z_1, \tau_1, L_0 \times \mathbb{R}) \geq A^{-1} \text{diam}(Q)$ and the conclusion is true for $j = 1$ and we let $L_1$ be the linear span of $Z_0$ and $Z_1$. Let $1 \leq j < n-1$ and suppose that we have found points $(Z_i, \tau_i) \in Q$ for $i = 1, \ldots, j$ as stated and that $L_{j-1}$ is the (spatial) $(j-1)$-dimensional plane spanned by $Z_0, Z_1, \ldots, Z_{j-1}$. Assume that $d_p(Z, \tau, L_{j-1} \times \mathbb{R}) \leq A^{-1} \text{diam}(Q)$ for all $(Z, \tau) \in Q$. Then $Q$ can be covered by roughly $A^{j-1} A^2$ cubes of size $2A^{-1} \text{diam}(Q)$ and, again using that $\Sigma$ has the parabolic ADR property, we see that

$$1 \leq A^{j-1} A^2 A^{-n+1}$$

which again is impossible, as $j < n$, if $A = A(n, M) \gg 1$ is large enough. Hence there exists $(Z_j, \tau_j) \in Q$ such that $d_p(Z_j, \tau_j, L_{j-1} \times \mathbb{R}) \geq A^{-1} \text{diam}(Q)$ and we let $L_j$ be the (spatial) $j$-dimensional plane spanned by $Z_0, Z_1, \ldots, Z_j$. The lemma now follows from the induction hypothesis. $\square$

**Lemma A.2.** Let $\Sigma \subset \mathbb{R}^{n+1}$ be a parabolic Ahlfors-David regular set with dyadic cubes $\mathcal{D}(\Sigma)$. Let $Q \in \mathcal{D} = \mathcal{D}(\Sigma)$ and assume that $P_1$ and $P_2$ are two time-independent hyperplanes such that $d_p(X, t, P_i) \leq \epsilon \text{diam}(Q)$ for all $(X, t) \in Q$ and for $i \in \{1, 2\}$. Then

$$(i) \ d_p(Y, s, P_2) \leq \epsilon (\text{diam}(Q) + d_p(Y, s, Q)), \ (Y, s) \in P_1,$$

(A.3) $$(ii) \ d_p(Y, s, P_1) \leq \epsilon (\text{diam}(Q) + d_p(Y, s, Q)), \ (Y, s) \in P_2.$$ 

In particular, the angle between $P_1$ and $P_2$ is bounded by $\leq \epsilon$.

Moreover, (A.3) continues to hold (with slightly different implicit constants) with $Q$ replaced by any “surface ball” $\Delta = \Delta_0(X, t) = C_1(X, t) \cap \Sigma$, where $(X, t) \in \Sigma$, provided that $d_p(X, t, P_i) \leq \epsilon \text{diam}(\Delta)$ for all $(X, t) \in 2\Delta$ and for each $i \in \{1, 2\}$. 
Proof. It is enough to prove the result for a cube $Q$: indeed, given a surface ball $\Delta = \Delta_r$ for which $d_p(X, t, P_i) \leq \varepsilon \text{diam}(\Delta)$ for all $(X, t) \in 2\Delta$ and for each $i \in \{1, 2\}$, we may cover $\Delta$ by a bounded number of disjoint cubes $Q_j$, with $\text{diam}(Q_j) \approx r$, such that $\cup_j Q_j \subset 2\Delta$, and observe that (A.3) (with $\varepsilon$ replaced by $C\varepsilon$) may be applied to each $Q_j$.

Fix now a dyadic cube $Q \subset \Sigma$, satisfying the hypotheses of the lemma. Let $(Z_i, \tau_i) \in Q$ for $i = 0, 1, \ldots, n$, be as in the statement of Lemma A.1 and let $L_Q = L_{n-1}$ be the spatial $(n-1)$-dimensional plane which passes through $Z_0, Z_1, \ldots, Z_{n-1}$. Let $P := L_Q \times \mathbb{R}$. Consider $(Y, s) \in P_1$. Then

$$d_p(Y, s, P_2) \leq d_p(Y, s, P) + d_p(P, P_2) \leq \max_{i \in \{0, \ldots, n\}} d_p(Y, s, Z_i, \tau_i) + \max_{i \in \{0, \ldots, n\}} d_p(Z_i, \tau_i, P_2) \leq \varepsilon \text{diam}(Q) + \max_{i \in \{0, \ldots, n\}} d_p(Y, s, Z_i, \tau_i) \leq \varepsilon(\text{diam}(Q) + d_p(Y, s, Q)).$$

This proves the lemma. \qed

APPENDIX B. SLICE-WISE VS. PARABOLIC HAUSDORFF MEASURE

In this appendix we prove the (non-trivial) assertions made in Remark 2.8. Recall $\mu$ is the (global) slice-wise measure defined as

$$\mu(E) = \int_R \mathcal{H}^{n-1}(E_t) d\mathcal{H}^1(t),$$

where for $t \in \mathbb{R}$, $E_t := E \cap (\mathbb{R}^n \times \{t\})$.

The following proposition proves Remark 2.8(vi).

Proposition B.1. There exists a constant $c(n)$ such that $\mu \leq c(n)\mathcal{H}^{n+1}_p$.

Proof. In the proof, for any parabolic cube $Q$, we will write $Q = Q' \times I$, where $I$ is an interval, so that the side length of $Q'$, $\ell(Q')$, and the length of $I$, $\ell(I)$, satisfy $\ell(I)^2 = \ell(Q')$ and we let $\ell(Q) := \ell(Q')$. Suppose that $E \subset \mathbb{R}^{n+1}$, a Borel set, is such that $\mathcal{H}^{n+1}_p(E) < \infty$ (otherwise there is nothing to prove). Recall that $\mathcal{H}^{n+1}_p(E)$ is a decreasing function in $\delta$. Let $\delta > 0$ be arbitrary. By definition of $\mathcal{H}^{n+1}_p(E)$ there exists a countable collection of cubes $Q_i$ such that

$$\sum_i \ell(Q_i)^{n+1} \leq 2^{n+1} \mathcal{H}^{n+1}_p(E) + \delta, \quad E \subseteq \cup_i Q_i \quad \text{and} \quad \text{diam}(Q_i) \leq \delta.$$

To see this, we use a covering $\{E_i\}$ which nearly minimizes the quantity in the definition of $\mathcal{H}^{n+1}_p(E)$, where $(2\sqrt{n} + \sqrt{2})\delta' = \delta$, then for each $i$ we take $Q_i = C_r(X_i, t_i)$ with $\ell(Q_i)/2 = r_i = \text{diam}(E_i)$ and $(X_i, t_i)$ an arbitrary point in $E_i$.

\[15\] Here we are taking an ‘honest’ parabolic cube, different from the $C_r(X, t)$ defined above. This means $Q = (X, t) : |X' - X| < \ell / 2, |t - t'| < \ell^2 / 2$, where $(X', t')$ is the center of the cube and $\ell$ is the side length.
Let $I(t) = \{i : t \in I_i\}$. Then as $E \subset Q_t$ we have $E_i \subseteq \bigcup_{i \in I(t)} Q'_i$ and $\text{diam}(Q'_i) \leq \delta$. Therefore

$$\mathcal{H}^{n-1}_0(E_i) \leq c'(n) \sum_{i \in I(t)} \ell(Q'_i)^{n-1} = c'(n) \sum_{i \in I(t)} \ell(Q_i)^{n-1},$$

where $c'(n) := (\sqrt{n})^{n-1}$. Hence

$$\int \mathcal{H}^{n-1}_0(E_i) d\mathcal{H}^1(t) \leq c'(n) \int \sum_{i \in I(t)} \ell(Q_i)^{n-1} d\mathcal{H}^1(t)$$

$$\leq c'(n) \sum_i \int \ell(Q_i)^{n-1} d\mathcal{H}^1(t) = c'(n) \sum_i \ell(Q_i)^{n+1}$$

$$\leq c'(n) 2^{n+1} \mathcal{H}^{n+1}_p(E) + c'(n).$$

The result now follows from the monotone convergence theorem as $\mathcal{H}^{n-1}_0(E_i)$ increases as $\delta$ decreases. □

The following proposition proves Remark 2.8(i) for the measure $\sigma^\delta$.

**Proposition B.2.** Let $E \subset \mathbb{R}^{n+1}$ be closed. There exists $c(n) > 0$ such that the following holds. If there is a constant $c$ such that

$$c^{-1} r^{n+1} \leq \mu(C_r(X, t) \cap E) \leq c r^{n+1}, \quad \forall (X, t) \in E, r > 0,$$

then

$$c(n)^{-1} c^{-1} r^{n+1} \leq \mathcal{H}^{n+1}_p(C_r(X, t) \cap E) \leq c(n) c r^{n+1}, \quad \forall (X, t) \in E, r > 0.$$

**Proof.** The lower bound follows directly from the previous proposition. To prove the other inequality fix $(X, t) \in E, r > 0$, let $\delta \in (0, r)$ be arbitrary and let $r_\delta$ be such that $\text{diam}(C_{r_\delta}(0)) = \delta$. Clearly,

$$C_r(X, t) \cap E \subseteq \bigcup_{(Y, s) \in C_r(X, t) \cap E} C_{r_\delta}(Y, s) \cap E.$$

By the $5r$-covering lemma (see e.g. [Mat, Theorem 2.1]) there exists a countable collection of $(Y_i, s_i) \in E \cap C_r(X, t)$ such that the cubes $C_{r_\delta}(Y_i, s_i)$ are disjoint and

$$C_r(X, t) \cap E \subseteq \bigcup_i C_{r_\delta}(Y_i, s_i) \cap E.$$

It follows that

$$\mathcal{H}^{n+1}_{p, \delta}(C_r(X, t) \cap E) \leq \sum_i \text{diam}(C_{r_\delta}(Y_i, s_i))^{n+1} \leq c \sum_i \mu(C_{r_\delta}(Y_i, s_i) \cap E)^{n+1}$$

$$\leq \mu(C_{2r}(X, t) \cap E) \leq c r^{n+1},$$

where now $\leq$ means that constants only depend on $n$, and where we have used that $\text{diam}(C_{r_\delta}(Y_i, s_i)) < \delta < r$ in the second-to-last inequality. Letting $\delta \to 0^+$ yields the upper bound. □

On the other hand, $\mathcal{H}^{n+1}_p \not\leq \mu$, even when the set is $p$-ADR with respect to $\mathcal{H}^{n+1}_p$, as the following example shows.
Example B.3. Let $E = [C_{1-1/2n}]^{n} \times C_{3/4}$, where $C_{s}$ denotes a (generalized) $s$-dimensional Cantor-type set in $\mathbb{R}$. Then $E$ is p-ADR with respect to $\mathcal{H}^{n+1}_p$, but $\mu(E) = 0$. The fact that $\mu(E) = 0$ follows from the fact that $\mathcal{H}^{1}(C_{3/4}) = 0$. Moreover, one can verify that $E$ is p-ADR with respect to the measure $\mathcal{H}^{n+1/2} \times \mathcal{H}^{3/4}$ and hence p-ADR with respect to $\mathcal{H}^{n+1}_p$.

We will require the following definition.

Definition B.4. Let $E \subset \mathbb{R}^{n+1}$ be a p-ADR set (with respect to $\mathcal{H}^{n+1}_p$) and let $S$ be a collection of p-ADR sets, with uniform control on the p-ADR constant. We say $E$ is big pieces of $S$, written $E$ is $BP(S)$, if there exists $\theta > 0$ such that for every $(X, t) \in E$ and $r \in (0, \text{diam}(E))$ there exists $\Gamma \in S$ with

$$\mathcal{H}^{n+1}_p(E \cap C_{r}(X, t) \cap \Gamma) \geq \theta \mathcal{H}^{n+1}_p(E \cap C_{r}(X, t)).$$

We say $E$ is big pieces squared of $S$, written $E$ is $BP^2(S)$ if there exists a constant $M$ such that $E$ is $BP(BP(S, M))$, where $BP(S, M)$ is the collection of all p-ADR sets with p-ADR constant less than $M$ which are $BP(S)$.

Proposition B.5. Let $S$ be a collection of closed subsets of $\mathbb{R}^{n+1}$, which are uniformly p-ADR with respect to $\mu$, that is, there exists $M > 1$ such that for all $\Gamma \in S$ it holds

$$M^{-1} r^{n+1} \leq \mu(\Gamma \cap C_{r}(X, t)) \leq M r^{n+1}, \quad \forall (X, t) \in \Gamma, r > 0.$$ 

If $E$ is p-ADR (with respect to $\mathcal{H}^{n+1}_p$) and $E$ is $BP(S)$, then $E$ is p-ADR with respect to $\mu$ with constant depending only on $M, n$, the p-ADR constant of $E$, and the constant $\theta$ in the definition of $BP(S)$. In particular, $\mu|_{E} \approx \mathcal{H}^{n+1}_p|_{E}$, with implicit constant depending only on $M, n$, the p-ADR constant of $E$, and the constant $\theta$ in the definition of $BP(S)$.

Proof. Let $(X, t) \in E$ and $r \in (0, \text{diam}(E))$. By Proposition B.1

$$\mu(C_{r}(X, t) \cap E) \leq c(n)\mathcal{H}^{n+1}_p(E \cap C_{r}(X, t)) \lesssim r^{n+1},$$

where the implicit constant depends on the p-ADR constant of $E$ and $n$. To prove the lower bound, we note that for $\Gamma \in S$, $\mathcal{H}^{n+1}_p|_{\Gamma} \approx \mu|_{\Gamma}$, with implicit constant depending on $M$ since any set which is p-ADR with respect to $\mu$ is p-ADR with respect to $\mathcal{H}^{n+1}_p$ (by Proposition B.2). Then using that $E$ is p-ADR and $E$ is $BP(S)$ there exists $\Gamma \in S$ such that

$$r^{n+1} \approx \mathcal{H}^{n+1}(E \cap C_{r}(X, t)) \lesssim \theta^{-1} \mathcal{H}^{n+1}(E \cap C_{r}(X, t) \cap \Gamma) \approx_{M} \theta^{-1} \mu(E \cap C_{r}(X, t) \cap \Gamma) \lesssim \theta^{-1} \mu(E \cap C_{r}(X, t)).$$

This proves the proposition. □

The following corollary finishes the proof of the assertion in Remark 2.8(v).

Corollary B.6. If $E$ is parabolic UR then $\mu|_{E} \approx \mathcal{H}^{n+1}_p|_{E}$, that is, $\sigma^{+} \approx \sigma$. Here the implicit constants depend on dimension and the parabolic UR constants for $E$. 

Proof. We have shown in Theorem 3.1 that if $E$ is parabolic UR, then $E$ admits a corona decomposition with respect to regular Lip($1,1/2$) graphs, with (uniform) control on the Lip($1,1/2$) constant in terms of the parabolic UR constants for $E$. Then it follows from [BHHLN1, Theorem 1.1] that $E$ is $BP^2(S)$, where $S$ is a collection of Lip($1,1/2$) graphs with uniform control on the Lip($1,1/2$) constant. In particular, $S$ is a collection of sets which are uniformly p-ADR with respect to $\mu$. Applying Proposition B.5 twice yields the corollary. □

References

[AEN] P. Auscher, M. Egert and K. Nyström, The Dirichlet problem for second order parabolic operators in divergence form. J. École Polytech. Math. 5 (2018), 407-441.

[AHLeMcT] P. Auscher, S. Hofmann, J. Lewis, and P. Tchamitchian. Extrapolation of Carleson measures and the analyticity of Kato’s square root operators, Acta Math. 187 (2001), 161-190.

[AHLMcT] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, and P. Tchamitchian. The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$, Ann. of Math. (2) 156, 2 (2002), 633-654.

[AGMT] J. Azzam, J. Garnett, M. Mourgoglou and X. Tolsa. Uniform rectifiability, elliptic measure, square functions, and $\varepsilon$-approximability via an ACF monotonicity formula. preprint arXiv:1612.02650.

[AHMMLT] J. Azzam, S. Hofmann, J. M. Martell, M. Mourgoglou, X. Tolsa, Harmonic measure and quantitative connectivity: geometric characterization of the $L^p$-solvability of the Dirichlet problem, Invent. Math. 222 (2020), no. 3, 881-993.

[BHHLN1] S. Bortz, J. Hoffmann, S. Hofmann, J.-L. Luna Garcia, K. Nyström, Coronizations and big pieces in metric spaces. To appear in Ann. Inst. Fourier.

[BHHLN2] S. Bortz, J. Hoffmann, S. Hofmann, J.-L. Luna Garcia, K. Nyström, On big pieces approximations of parabolic hypersurfaces. Ann. Fenn. Math. 47 (2022), no. 1, 533-571.

[CCFJR] A. P. Calderón, C. P. Calderón, E. Fabes, M. Jodeit, and N. M. Riviere, Applications of the Cauchy integral on Lipschitz curves, Bull. Amer. Math. Soc. 84, Number 2 (1978), 287-290.

[Chr] M. Christ, A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral, Colloq. Math., LX/LXI (1990), 601–628.

[CMM] R. Coifman, A. McIntosh and Y. Meyer, L’intégrale de Cauchy définit un opérateur borné sur $L^2$ pour les courbes lipschitziennes. [The Cauchy integral defines a bounded operator on $L^2$ for Lipschitz curves] Ann. of Math. (2) 116 (1982), no. 2, 361-387.

[CDM] R. Coifman, G. David, G. and Y. Meyer, La solution des conjecture de Calderón. [Solution of Calderón’s conjectures] Adv. in Math. 48 (1983), no. 2, 144-148.

[CS] R. Coifman and S. Semmes, $L^2$ estimates in nonlinear Fourier analysis. Harmonic analysis (Sendai, 1990), 79–95, ICM-90 Satell. Conf. Proc., Springer, Tokyo, 1991.

[D] G. David, Wavelets and singular integrals on curves and surfaces. Lecture Notes in Mathematics, 1465. Springer-Verlag, Berlin, 1991. x+107 pp.

[DJ] G. David and D. Jerison, Lipschitz approximation to hypersurfaces, harmonic measure, and singular integrals. Indiana Univ. Math. J. 39 (1990), no. 3, 831-845.

[DS1] G. David and S. Semmes, Singular integrals and rectifiable sets in $\mathbb{R}^n$: Beyond Lipschitz graphs, Astérisque 193 (1991).

[DS2] G. David and S. Semmes, Analysis of and on Uniformly Rectifiable Sets, Mathematical Monographs and Surveys 38, AMS 1993.
[DPV] E. Di Nezza, G. Palatucci and E. Valdinoci, Hitchhiker’s guide to the fractional Sobolev spaces, Bulletin des Sciences Mathématiques 136, 2012. 43

[DPP] M. Dindoš, S. Petermichl and J. Pipher, BMO solvability and the $A_v$ condition for second order parabolic operators. Ann. Inst. H. Poincaré Anal. Non Linéaire 34 (2017), no. 5, 1155-1180. 4

[E] M. Engelstein, A free boundary problem for the parabolic Poisson kernel. Adv. Math. 314 (2017), 835–947. 4

[GMT] J. Garnett, M. Mourgoglou and X. Tolsa, Uniform rectifiability from Carleson measure estimates and $\epsilon$-approximability of bounded harmonic functions. Duke Math. J. 167 (2018), no. 8, 1473-1524. 4

[GH1] A. Genschaw, and S. Hofmann, A Weak Reverse Hölder Inequality for Caloric Measure. J. Geom. Anal. 30 (2020), no. 2, 1530-1564. 4

[GH2] A. Genschaw, and S. Hofmann, BMO Solvability and Absolute Continuity of Caloric Measure, preprint arXiv:1904.08407, to appear, Potential Analysis. 4

[H1] S. Hofmann, A characterization of commutators of parabolic singular integrals, Proceedings of conference on Harmonic Analysis and PDE, held at Miraflores de la Sierra, Spain, 1992, J. Garcia-Cuerva, E. Hernández, F.Soria, editors, CRC press, Boca Raton (1995), pp. 195-210. 3

[H2] S. Hofmann, Parabolic singular integrals of Calderon-type, rough operators, and caloric layer potentials, Duke Math. J., Vol. 90 (1997), pp 209-259. 3, 7, 40

[HLMN] S. Hofmann, P. Le, J. M. Martell, and K. Nyström. The weak-$A_v$ property of harmonic and $p$-harmonic measures implies uniform rectifiability. Anal. PDE. 10 (2017), no. 3, 513–558. 4

[HL1] S. Hofmann and J. L. Lewis, $L^2$ Solvability and representation by caloric layer potentials in time-varying domains. Annals of Math. 144 (1996), 349-420. 3, 4, 7

[HL2] S. Hofmann and J. L. Lewis, Square functions of Calderon type, and applications, Rev. Math. Ibero. 17 (2001), pp 1-20. 3, 7

[HL3] S. Hofmann and J. Lewis, The Dirichlet problem for parabolic operators with singular drift terms, Mem. Amer. Math. Soc. 151 (2001). 4

[HLN1] S. Hofmann, J. Lewis and K. Nyström, Existence of big pieces of graphs for parabolic problems, Annales Acad. Sci. Fenn. Mathematica 28 (2003), 355-384. 3, 4, 5, 6, 7, 9, 13, 31

[HLN2] S. Hofmann, J. Lewis and K. Nyström, Caloric measure in parabolic flat domains, Duke Math. J. 122 (2004), 281-345. 3, 4, 7, 9, 13, 31

[HMM1] S. Hofmann, J.M. Martell and S. Mayboroda, Uniform rectifiability, Carleson measure estimates, and approximation of harmonic functions. Duke Math. J. 165 (2016), no. 12, 2331-2389. 3, 4, 48, 50

[KW] R. Kaufman and J.M. Wu, Parabolic measure on domains of class Lip$_{1/2}$, Compositio Mathematica 65 (1988), 201-207. 3, 7

[LM] J. Lewis and M. A. M. Murray, The method of layer potentials for the heat equation in time-varying domains, Mem. Amer. Math. Soc. 545 (1995) 1 - 157. 3, 4, 7

[LN] J. Lewis and K. Nyström, On a Parabolic Symmetry Problem, Revista Matematica Iberoamericana, Volume 23/Number 2 (2007), 513-536. 3

[LS] J. Lewis and J. Silver, Parabolic measure and the Dirichlet problem for the heat equation in two dimensions, Indiana Univ. Math. J. 37 (1988), 801-839. 3

[LV] J. Lewis and A. Vogel, Uniqueness in a free boundary problem, Communications in PDE 31 (2006), 1591-1614.

[LMcS] Chun Li, A. McIntosh and S. Semmes, Convolution singular integrals on Lipschitz surfaces, J. Amer. Math. Soc. 5 (1992), 455 - 481. 2
M. Mourgoglou and C. Puliatti, Blow-ups of caloric measure in time varying domains and applications to two-phase problems. preprint arXiv:2008.06968.

M. Mourgoglou and X. Tolsa. Harmonic measure and Riesz transform in uniform and general domains. J. Reine Angew. Math. 758 (2020), 183–221.

P. Mattila, Geometry of sets and measures in Euclidean spaces, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

J. Mateu, L. Prat, and X. Tolsa, Removable singularities for Lipschitz caloric functions in time varying domains, preprint arXiv:2005.03397.

M.A.M. Murray, The Cauchy integral, Calderón commutators, and conjugations of singular integrals in $\mathbb{R}^n$. Trans. Amer. Math. Soc. 289 (2) (1985) 497-518.

K. Nyström. Caloric measure and Reifenberg flatness. Ann. Acad. Sci. Fenn. Math. 31: 405–436, 2006.

K. Nyström. On an inverse type problem for the heat equation in parabolic regular graph domains. Math. Z. 31: 197–222, 2012.

K. Nyström, and M. Strömqvist, On the parabolic Lipschitz approximation of parabolic uniform rectifiable sets. Rev. Mat. Iberoam. 33 (2017), no. 4, 1397-1422.

S. Rigot, Quantitative notions of rectifiability in the Heisenberg groups. preprint, arXiv:1904.06904.

J. Rivera-Noriega, A parabolic version of Corona decompositions. Illinois Math Journal 53 (2009), 533-559.

J. Rivera-Noriega, Two results over sets with big pieces of parabolic Lipschitz graphs. Houston Journal of Mathematics 36 (2010) 619-635.

J. Rivera-Noriega, Parabolic singular integrals and uniform rectifiable sets in the parabolic sense, J. Geometrical Analysis 23 (2013), 1140-1157.