Approximate and Exact Extensions of Lebesgue Boundary Functions

Tuomas Hytönen and Andreas Rosén

Abstract. We prove continuity and surjectivity of the trace map onto $L^p(\mathbb{R}^n)$, from a space of functions of locally bounded variation, defined by the Carleson functional. The extension map is constructed through a stopping time argument. This extends earlier work by Varopoulos in the BMO case, related to the Corona theorem.

1. Introduction

Estimates of traces $u|_{\partial D}$ of functions $u : D \to \mathbb{R}$ in some given domain $D$, say in the Euclidean space, are important in analysis, for example in boundary value problems for partial differential equations. By local parametrization, it often suffices to consider the case where $D$ is the half-space

$$\mathbb{R}^{1+n}_+ := \{(t, x) : t > 0, x \in \mathbb{R}^n\}$$

and the traces are defined on the boundary $\partial \mathbb{R}^{1+n}_+ = \mathbb{R}^n = \{(0, x) : x \in \mathbb{R}^n\}$. We shall concentrate on this case here. A first problem is to show boundedness of the trace map

$$\gamma : u(t, x) \mapsto g(x) = (\gamma u)(x) := u(0, x).$$

This amounts to identifying norms $\| \cdot \|_D$ and $\| \cdot \|_{\partial D}$ on the function spaces for $u$ and $g$ respectively, so that an estimate $\|g\|_{\partial D} \lesssim \|u\|_D$ holds. A second problem is to determine whether $\gamma$, as a map between the corresponding function spaces, is surjective. One wants that any $g$ can be extended to some $u$ in $D$ such that $\gamma(u) = g$, with estimates $\|u\|_D \lesssim \|g\|_{\partial D}$.

The most well-known trace result is the Sobolev trace theorem. This states that the trace map

$$\gamma : H^s(\mathbb{R}^{1+n}_+) \to H^{s-1/2}(\mathbb{R}^n)$$

is bounded and surjective when $s > 1/2$. It is important to note that the Sobolev trace theorem breaks down in the limit case of regularity $s = 1/2$, and does not yield a bounded trace map onto the Lebesgue boundary space $L_2(\mathbb{R}^n)$. One way to solve this problem is to consider instead the scale of Besov spaces $B^{s}_{p,q}$, where the trace map

$$\gamma : B^s_{p,q}(\mathbb{R}^{1+n}_+) \to B^{s-1/2}_{p,q}(\mathbb{R}^n),$$

is bounded and surjective when $s > 1/p$. Here also $\gamma : B^{1/p}_{p,1}(\mathbb{R}^{1+n}_+) \to L^p(\mathbb{R}^n)$ is bounded and surjective whenever $1 \leq p < \infty$.

1 Formerly Andreas Axelsson.

T.H. is supported by the European Research Council through the ERC Starting Grant “Analytic–probabilistic methods for borderline singular integrals”. A.R. is supported by Grant 621-2011-3744 from the Swedish research council, VR.
Our first main result provides a new bounded and surjective trace map onto $L_p(\mathbb{R}^n)$, from a space of functions of locally bounded variation in the half-space, with norm

$$\|C(\nabla u)\|_{L_p(\mathbb{R}^n)},$$

using the Carleson functional

$$Cf(x) := \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(t,y)| dt dy, \quad x \in \mathbb{R}^n,$$

for functions in $\mathbb{R}^{1+n}_+$. Here the supremum is over all cubes $Q \subset \mathbb{R}^n$ containing $x$, and $\hat{Q} := (0, \ell(Q)) \times Q$ denotes the Carleson box above $Q$, with side length $\ell(Q)$. Note that $C(\nabla u)$ is well defined in the natural way for any $u \in BV_{\text{loc}}(\mathbb{R}^{1+n}_+)$ of locally bounded variation. However, for technical reasons, to avoid issues with interior regularity and constants, we choose to formulate the result using the interior of locally bounded variation. However, for technical reasons, to avoid issues with interior regularity and constants, we choose to formulate the result using the interior function space

$$V_p := \{u \in C^1(\mathbb{R}^{1+n}_+) : C(\nabla u) \in L_p(\mathbb{R}^n), Nu \in L_p(\mathbb{R}^n)\},$$

where $Nu(x) := \text{ess sup}_{|y-x| < c_1 t} |u(t,y)|, x \in \mathbb{R}^n$, denotes the non-tangential maximal function. For the definition of the trace map below, we use averages over Whitney regions

$$W(t,x) := \{ (s,y) : c_0^{-1} < s/t < c_0, |y-x| < c_1 t \},$$

with some fixed parameters $c_0 > 1$ and $c_1 > 0$.

**Theorem 1.1.** Let $1 < p < \infty$. Consider the normed linear function space $V_p$ with norm $\|C(\nabla(\cdot))\|_p$. Then the trace $\gamma u$ of any $u \in V_p$, is well defined almost everywhere in the sense of convergence of Whitney averages

$$(\gamma u)(x) := \lim_{t \to 0^+} |W(t,x)|^{-1} \int_{W(t,x)} u(s,y) ds dy, \quad x \in \mathbb{R}^n.$$ 

The trace map $\gamma : V_p \to L_p(\mathbb{R}^n)$ is well defined, and there exists $c_p < \infty$ so that estimates

$$\|\gamma u\|_{L_p(\mathbb{R}^n)} \leq c_p \|C(\nabla u)\|_{L_p(\mathbb{R}^n)}$$

hold for all $u \in V_p$. Moreover, the trace map $\gamma$ is surjective, and given any $g \in L_p(\mathbb{R}^n)$ there exists an extension $u \in V_p$ such that $\gamma u = g$, with estimates

$$\|C(\nabla u)\|_{L_p(\mathbb{R}^n)} \leq c_p \|g\|_{L_p(\mathbb{R}^n)}.$$ 

We remark that the extension operator $g \mapsto u$ is in general non-linear, even though $\gamma$ itself of course is linear. The extension operator constructed is also such that non-tangential estimates $\|Nu\|_{L_p(\mathbb{R}^n)} \leq c_p \|g\|_{L_p(\mathbb{R}^n)}$ hold and pointwise non-tangential limits $\lim_{t,y \to (0,x), |y-x| < \alpha t} u(t,y) = g(x)$ exist at each Lebesgue point of $g$, for any fixed $\alpha < \infty$.

The corresponding trace result in the case $p = \infty$, proved by Varopoulos [7, 8] is that there is a bounded and surjective trace map

$$\|u\|_{BV(\mathbb{R}^n)} \lesssim \|C(\nabla u)\|_{L_{\infty}(\mathbb{R}^n)},$$

and a corresponding non-linear bounded extension operator, where $BV(\mathbb{R}^n)$ stands for the John–Nirenberg space of functions of bounded mean oscillation. Following Varopoulos [5], we obtain the extensions in Theorem 1.1 from a result on approximate extensions of Lebesgue functions on $\mathbb{R}^n$. This main result in Theorem 1.1 contained in our Theorem 1.2 generalizes well known techniques in the end point
case $p = \infty$ related to the Corona Theorem, first proved by Carleson [11]. Our proof of Theorem 1.2 though, is more in the spirit of Garnett [4, Ch. VIII, Thm. 6.1].

**Theorem 1.2.** Fix $1 < p < \infty$. Consider $g \in L_p(\mathbb{R}^n)$ and define the dyadic average extension
\[
u(t, x) := |Q|^{-1} \int_Q g(y) \, dy, \quad (t, x) \in W_Q,
\]
where $W_Q := (\ell(Q)/2, \ell(Q)) \times Q$ denotes the dyadic Whitney region above a dyadic cube $Q \subset \mathbb{R}^n$ of side length $\ell(Q)$. Then, for any $0 < \epsilon < 1$, there exists $f : \mathbb{R}^{1+n} \to \mathbb{R}$ which is constant on each dyadic Whitney region, with estimates
\[
\begin{cases}
\|N(f - u)\|_{L_p(\mathbb{R}^n)} \leq \epsilon \|g\|_{L_p(\mathbb{R}^n)}, \\
\|C(\nabla f)\|_{L_p(\mathbb{R}^n)} \leq \epsilon^{-1} \|g\|_{L_p(\mathbb{R}^n)}.
\end{cases}
\]
Moreover, for any fixed $\alpha < \infty$, non-tangential limits $\lim_{(t,y) \to (0,x),|y-x|<\alpha} f(t,y) =: f(0,x)$ exist almost everywhere, so that $\|f(0,\cdot) - g\|_{L_p(\mathbb{R}^n)} \leq \epsilon \|g\|_{L_p(\mathbb{R}^n)}$.

We remark that our proof in fact yields the pointwise estimates $N_p(f - u) \leq \epsilon M_p g$ and $C_p(\nabla u) \leq \epsilon^{-1} M_p(M_p g)$, using the dyadic non-tangential maximal functional, the dyadic Carleson functional and the dyadic Hardy–Littlewood maximal functional. See Section 2.

That the construction of approximate extensions $f$ as above with control of $C(\nabla f)$ is indeed non-trivial, can be seen as follows. Consider a “lacunary” function, which in a standard Haar basis would mean something like
\[
u(x) := \sum_{Q \subset [0,1],|Q| \geq 2^{-k}} (\chi_{Q^+}(x) - \chi_{Q^-}(x)), \quad x \in \mathbb{R},
\]
where the sum is over dyadic subintervals of $(0,1)$ of length at least $2^{-k}$, and the integrand uses the characteristic functions of the left and right dyadic children of $Q$. Then one checks that $\|\nu\|_p \lesssim \sqrt{k}$, whereas the dyadic average extension $\nu(t,x)$ is seen to satisfy
\[C(\nabla \nu)(x) \gtrsim k, \quad \text{for all } x \in (0,1).
\]
Therefore the dyadic average extension, or the closely related Poisson extension, will not satisfy the required estimates. Instead Theorem 1.2 is proved using a stopping time construction, where we modify the stopping condition used in endpoint BMO case.

The outline of the paper is as follows. In Section 2 we survey the basic estimates for the functionals defining our spaces. In Section 3 we deduce Theorem 1.1 from Theorem 1.2 and prove the latter using a weighted stopped square function estimate, the proof of which is in Section 4.

2. The basic functionals

In this section, we collect well known facts concerning the functionals we use to define norms of functions in the half space $\mathbb{R}^{1+n}_+$. First we fix notation. We write the $L_p(\mathbb{R}^n)$ norm as $\|\cdot\|_p$. Cubes in $\mathbb{R}^n$ (dyadic or not) we denote by $Q, R, S, \ldots$, and we assume that these are open. The Carleson box above a cube $Q \subset \mathbb{R}^n$ is denoted
\[
\hat{Q} := (0,\ell(Q)) \times Q \subset \mathbb{R}^{1+n}_+.
\]
where \( \ell(Q) \) denotes the sidelength of \( Q \). We write \( cQ \) to denote the cube with same center as \( Q \) but with \( \ell(cQ) = c\ell(Q) \).

Let \( \mathcal{D} = \bigcup_{j \in \mathbb{Z}} \mathcal{D}_j \) denote a fixed system of dyadic cubes in \( \mathbb{R}^n \), with \( \mathcal{D}_j \) being the cubes of sidelength \( \ell(Q) = 2^{-j} \). We assume that the dyadic cubes in \( \mathcal{D} \) form a connected tree under inclusion. Let \( W_Q := (\ell(Q)/2,\ell(Q)) \times Q \) denote dyadic Whitney regions. The corresponding non-dyadic Whitney region around a point \((t,x) \in \mathbb{R}^{1+n}_+\) we define to be

\[
W(t,x) := \{(s,y) : c_0^{-1} < s/t < c_0, |y - x| < c_1 t\},
\]

where \( c_0 > 1 \) and \( c_1 > 0 \) are fixed parameters.

The Hardy-Littlewood maximal function of \( f \in L^{loc}_1(\mathbb{R}^n) \) that we use is

\[
Mf(x) := \sup_{Q \ni x} |Q|^{-1} \int_Q |f(y)|dy,
\]

where the supremum in \( Mf \) is over all cubes \( Q \subset \mathbb{R}^n \) containing \( x \). Restricting the cubes to the dyadic ones in the supremum yields the dyadic Hardy-Littlewood maximal function \( M_D f(x) \). We also require the following truncated (to large cubes) version of the dyadic Hardy-Littlewood maximal function:

\[
(1) \quad M_D f(Q) = \sup_{R \supset Q, R \in \mathcal{D}} |R|^{-1} \int_R |f(y)|dy, \quad Q \in \mathcal{D}.
\]

**Definition 2.1.** For a locally integrable function \( f(t,x) \) in \( \mathbb{R}^{1+n}_+ \) we define, for \( x \in \mathbb{R}^n \), the non-tangential maximal functional \( Nf \), the Carleson functional \( Cf \) and the area functional \( Af \)

\[
Nf(x) := \text{ess sup}_{|y-x|<\alpha t} |f(t,y)|,
\]

\[
Cf(x) := \sup_{Q \ni x} |Q|^{-1} \int_Q |f(t,y)|dtdy,
\]

\[
Af(x) := \int_{|y-x|<\alpha t} |f(t,y)|t^{-n}dtdy.
\]

Here the supremum in \( Cf \) is over all cubes \( Q \subset \mathbb{R}^n \) containing \( x \). In the definition of \( Nf \) and \( Af \), the parameter \( \alpha > 0 \) denotes some fixed aperture of the cones. To emphasize the exact dependence of the aperture, we sometimes write \( N^{(\alpha)} \) and \( A^{(\alpha)} \).

For a function \( f(t,x) \) in \( \mathbb{R}^{1+n}_+ \), having constant value \( f_Q \) on each dyadic Whitney region \( W_Q \), we also define dyadic versions of these functionals by

\[
N_D f(x) := \sup_{x \in Q \in \mathcal{D}} |f_Q|,
\]

\[
C_D f(x) := \sup_{x \in Q \in \mathcal{D}} |Q|^{-1} \sum_{R \in \mathcal{D}, R \subset Q} |f_R||W_R|,
\]

\[
A_D f(x) := \sum_{x \in Q \in \mathcal{D}} |f_Q|\ell(Q).
\]

We want to point out that we are using the measure \( dtdx \) and not the measure \( t^{-1}dtdx \) throughout this paper, which is also common in the literature. Note that the functionals \( A \) and \( C \) extend in a natural way to the case when \( f \) is a signed measure on \( \mathbb{R}^{1+n}_+ \), and in particular to the case of gradients of functions of locally bounded variation.
We record the following norm equivalences between different choices for the aperture of the cones.

**Proposition 2.2.** Fix \(0 < \alpha, \beta < \infty\). Then for any \(1 \leq p \leq \infty\), we have \(\|N^{(\alpha)} f\|_p \approx \|N^{(\beta)} f\|_p\), and for any \(1 \leq p < \infty\), we have \(\|A^{(\alpha)} f\|_p \approx \|A^{(\beta)} f\|_p\).

**Proof.** The estimates for \(Nf\) are proved in Fefferman and Stein [3, Lem. 1]. To prove the estimate for the \(A\)-functional, we follow Coifman, Meyer and Stein [2, Prop. 4, case \(2 \leq p < \infty\)] and consider \(0 < \alpha < \beta < \infty\): Dualize against \(\|h\|_{p'} = 1\) to get

\[
\|A^{(\beta)} f\|_p = \int_{\mathbb{R}^n} \left( \int_{|y-x|<\beta t} |f(t,y)| t^{-n} dt dy \right) h(x) dx \\
= \int_{\mathbb{R}_+^{1+n}} |f(t,y)| \left( t^{-n} \int_{|x-y|<\beta t} h(x) dx \right) dt dy \\
\lesssim \int_{\mathbb{R}_+^{1+n}} |f(t,y)| \left( t^{-n} \int_{|x-y|<\alpha t} M h(x) dx \right) dt dy \\
= \int_{\mathbb{R}^n} A^{(\alpha)} f(x) M h(x) dx \lesssim \|A^{(\alpha)} f\|_p.
\]

□

We also record the following equivalence of norms between the corresponding dyadic and non-dyadic functionals.

**Proposition 2.3.** We have

\[
\|N f\|_p \approx \|N_D f\|_p, \quad 1 \leq p \leq \infty,
\]

\[
\|C f\|_p \approx \|C_D f\|_p, \quad 1 < p \leq \infty,
\]

\[
\|A f\|_p \approx \|A_D f\|_p, \quad 1 < p < \infty,
\]

uniformly for all functions \(f(t,x)\) in \(\mathbb{R}_+^{1+n}\) that are constant on each dyadic Whitney region.

**Proof.** For proofs of the results for \(N\) and \(C\), we refer to [6]. Consider now the area functional \(A\). As in the proof of Proposition 2.2, the proof is an adaption of [2, Prop. 4, case \(2 \leq p < \infty\)]. Dualize against \(\|h\|_{p'} = 1\) to get

\[
\|A f\|_p = \int_{\mathbb{R}^n} \left( \int_{|y-x|<\alpha t} |f(t,y)| t^{-n} dt dy \right) h(x) dx \\
= \int_{\mathbb{R}_+^{1+n}} |f(t,y)| \left( t^{-n} \int_{|x-y|<\alpha t} h(x) dx \right) dt dy \\
\lesssim \sum_{Q \in \mathcal{D}} |f_Q| |W_Q| \left( |Q|^{-1} \int_Q (M h)(x) dx \right) \\
= \int_{\mathbb{R}^n} (A_D f)(M h) dx \lesssim \|A_D f\|_p.
\]

□

Less obvious is the following important \(L_p\) equivalence of the \(A\) and \(C\) functionals.
Proposition 2.4. For $1 \leq p < \infty$, we have
\[ \|Af\|_p \lesssim \|Cf\|_p. \]

For $1 < p \leq \infty$, we have
\[ \|Cf\|_p \lesssim \|Af\|_p \]
(for any fixed aperture $\alpha$ in the case $p = \infty$).

At the endpoint $p = \infty$, the $A$-functional is dependent on the choice of aperture, and should be replaced by the Carleson functional, which is strictly smaller, as seen from the example
\[ f(t, x) = (t + |x|)^{-n}. \]

At the endpoint $p = 1$, we have $\|Cf\|_1 < \infty$ only if $f = 0$, so the Carleson functional should be replaced by the area functional, which in this case defines simply the function space $L^1(\mathbb{R}^{1+n})$.

Proposition 2.4 is a reformulation of Coifman, Meyer and Stein [2, Thm. 3]. The proof below contains some novelties in the estimate $A \lesssim C$.

Proof of Proposition 2.4. For $C \lesssim A$ we have
\[
M(Af)(x) = \sup_{Q \ni x} |Q|^{-1} \int_Q \left( \int_{|y-x| < \alpha t} |f(t, y)| t^{-n} dy \right) dx
= \sup_{Q \ni x} |Q|^{-1} \int \left( \int_{|y-x| < \alpha t} h(x) dx \right) |f(t, y)| dy \gtrsim Cf(x).
\]

For $A \lesssim C$, we argue by duality with a suitable $\|h\|_p = 1$:
\[
\|Af\|_p = \int_{\mathbb{R}^n} \left( \int_{|y-x| < \alpha t} |f(t, y)| t^{-n} dy \right) h(x) dx
= \int_{\mathbb{R}^{1+n}} |f(t, y)| H(t, y) dtdy = \int_0^\infty \left( \int_{H(t, y) > \lambda} \int_{|y-x| < \alpha t} h(x) dx \right) d\lambda.
\]

If $H(t, y) > \lambda$, there is a cube $Q$ such that $|Q|^{-1} \int_Q h(x) dx > c\lambda$ and $(t, y) \in \hat{Q}$. By the Whitney covering lemma, there is a collection $Q_\Lambda$ of these cubes such that the $\hat{Q}$ are pairwise disjoint, and the $5\hat{Q}$, $Q \in Q_\Lambda$, cover all the points $(t, y)$ with $H(t, y) > \lambda$. Thus
\[
\int_{H(t, y) > \lambda} \int_{|y-x| < \alpha t} |f(t, y)| dy dx \lesssim \sum_{Q \in Q_\Lambda} \int_{\hat{Q}} |f(t, y)| dy dx \lesssim \sum_{Q \in Q_\Lambda} |5\hat{Q}| \inf_{x \in Q} Cf(x)
\lesssim \sum_{Q \in Q_\Lambda} \int_Q Cf(x) dx \leq \int_{\{Mh > \lambda\}} Cf(x) dx.
\]

Substituting back, this shows that
\[
\|Af\|_p \lesssim \int_0^\infty \int_{\{Mh > \lambda\}} Cf(x) dx d\lambda \lesssim \int_{\mathbb{R}^n} Cf(x) Mh(x) dx \lesssim \|Cf\|_p.
\]
3. CONSTRUCTION OF EXTENSIONS

In this section we prove Theorems 1.1 and 1.2, assuming a square function estimate which we prove in Section 4. We first deduce Theorems 1.1 from Theorem 1.2.

Proof of Theorem 1.2 assuming Theorem 1.1

(1) First we consider the existence and bound of the trace of $u \in \mathcal{V}_p$. Fix $x \in \mathbb{R}^n$ and consider two Whitney regions $W(t_1, x)$ and $W(t_2, x)$, with $t_1 < t_2$. Estimate

$$
\left| W(t_2, x) \right|^{-1} \int_{W(t_2, x)} u(s, y) ds dy - \left| W(t_1, x) \right|^{-1} \int_{W(t_1, x)} u(s, y) ds dy
\approx \left| \int_{W(1, 0)} (u(t_2 s + t_2 y) - u(t_1 s + t_1 y)) ds dy \right|
\lesssim \int_{W(1, 0)} \int_{t_1}^{t_2} |\nabla u(t s + t y)| dt ds dy
= \int_{t_1}^{t_2} \int_{t/c_0}^{c_0 t} \int_{|y' - x| < c_1 t} |\nabla u(s', y')| t^{-1-n} dy' ds' dt
\lesssim \int_{t_1/c_0}^{c_0 t_2} \int_{|y' - x| < c_1 s'} |\nabla u(s', y')| (s')^{-n} dy' ds'
$$

Since $A(\nabla u) \in L_p(\mathbb{R}^n)$ by Proposition 2.4, we have for almost all $x \in \mathbb{R}^n$ that $A(\nabla u)(x) < \infty$. For such $x$, it follows from the above estimate that Whitney averages converges as $t \to 0$. Thus, in this sense we have a well defined trace almost everywhere on $\mathbb{R}^n$. Moreover, from $N u \in L_p(\mathbb{R}^n)$ we have the existence in particular of $x_0 \in \mathbb{R}^n$ where $N(\alpha) u(x_0) < \epsilon$ for any given $\epsilon > 0$ and $\alpha > c_0 c_1$. This shows that $|W(t_2, x)|^{-1} \int_{W(t_2, x)} u(s, y) ds dy \to 0$ as $t_2 \to \infty$. The estimate

$$
\|\gamma u\|_p \lesssim \|A(\nabla u)\|_p \approx \|C(\nabla u)\|_p
$$

follows.

(2) We construct the extension $u$ of $g \in L_p(\mathbb{R}^n)$ as follows. Define functions $g_k$, $u_k$ and $f_k$, $k = 0, 1, 2, \ldots$ inductively: Let $g_0 := g$. Given $g_k \in L_p(\mathbb{R}^n)$, $k \geq 0$, we apply Theorem 1.2 to define the dyadic extension $u_k$ and its approximation $f_k$, with estimates

$$
\|N(f_k - u_k)\|_p \leq \epsilon \|g_k\|_p,
\|C(\nabla f_k)\|_p \leq C \epsilon^{-1} \|g_k\|_p.
$$

Then let $g_{k+1} := g_k - f_k |\mathbb{R}^n$. We have

$$
\|g_{k+1}\|_p \leq \|N(u_k - f_k)\|_p \leq \epsilon \|g_k\|_p
$$

and therefore $\|g_k\|_p \leq \epsilon^k \|g\|_p$. Define

$$
f := \sum_{k=0}^{\infty} f_k.
$$
which is an exact extension of $g$ since

$$0 = \lim_{k \to \infty} \|g_{k+1}\|_p = \left\| g - \sum_{j=0}^{k} f_j \right\|_{p}. $$

Moreover, we have estimates

$$\|C(\nabla f)\|_p \leq \sum_{k=0}^{\infty} C\epsilon^{-1} \epsilon \|g\|_p \lesssim \|g\|_p,$$

fixing some $0 < \epsilon < 1$.

It remains to mollify $f$ to obtain

$$u(t, x) := \int_{\mathbb{R}^{1+n}} f(ts, x + ty) \eta(s, y) ds dy, \quad (t, x) \in \mathbb{R}^{1+n},$$

where $\eta \in C^\infty_0(W(1,0))$ has $\int \eta = 1$. Then it is straightforward to verify that $u \in V_p$, with the stated estimate of

$$\nabla u(t, x) = \int_{\mathbb{R}^{1+n}} \left[ \begin{array}{c} s \\ 0 \\ 1 \end{array} \right] \nabla f(ts, x + ty) \eta(s, y) ds dy. \quad \square$$

For the proof of Theorem 1.2, we require the following lemma for the truncated dyadic maximal function from [11].

**Lemma 3.1.** For any $g \in L^1_\text{loc}(\mathbb{R}^n)$ and $Q \in D$, we have

$$\frac{|Q|}{M_D g(Q)} \leq 4 \int_Q \frac{dx}{M_D g(x)}.$$

**Proof.** Define

$$E_Q := \{ x \in Q : M_D g(x) > 2M_D g(Q) \} = \{ x \in Q : M_D (g1_Q)(x) > 2M_D g(Q) \}.$$

The weak $L_1$ estimate for $M_D$ yields

$$|E_Q| \leq \frac{1}{2M_D g(Q)} \int_Q |g| dx \leq \frac{1}{2} |Q|.$$

Thus

$$\frac{|Q|}{M_D g(Q)} \leq 2 \frac{|Q \setminus E_Q|}{M_D g(Q)} \leq 4 \int_{Q \setminus E_Q} \frac{dx}{M_D g(x)} \leq 4 \int_Q \frac{dx}{M_D g(x)}.$$

**Proof of Theorem 1.2** modulo Lemma 3.2 and Proposition 4.1 (1) We first localize the problem to a large top cube $Q_0$. Choose $Q_0 \in D$ large enough so that

$$\int_{\mathbb{R}^n \setminus Q_0} |M_D g|^p dx \leq \delta,$$

where $\delta > 0$ is to be chosen below. Define

$$g_2(x) := \begin{cases} |Q_0|^{-1} \int_{Q_0} g dy, & x \in Q_0, \\ g(x), & x \notin Q_0, \end{cases}$$

and let $g_1 := g - g_2$. Let $u_1$ and $u_2$ be the dyadic extensions of $g_1$ and $g_2$ respectively, so that $u_1$ is non-zero only on $\hat{Q}_0$.

Define the approximation $f_2$ of $u_2$ to be

$$f_2(t, x) := \begin{cases} |Q_0|^{-1} \int_{Q_0} g dy, & (t, x) \in \hat{Q}_0, \\ 0, & (t, x) \notin \hat{Q}_0. \end{cases}$$
Then $N_D(f_2 - u_2)(x) \leq M_D g(x)$ if $x \notin Q_0$ and $N_D(f_2 - u_2)(x) \leq \inf_{Q_1} M_D g$ if $x \in Q_0$, where $Q_1$ is the sibling of $Q_0$. Thus

$$
\|N_D(f_2 - u_2)\|_p^p \leq 2 \int_{\mathbb{R}^n \setminus Q_0} |M_D g|^p dx \leq (\epsilon/2)^p \|g\|_p^p,
$$

provided $2\delta \leq (\epsilon/2)^p \int_{\mathbb{R}^n} |g|^p dx$. Furthermore $\|C(\nabla f_2)\|_p \leq |Q_0|^{1/n}(|Q_0|^{-1} \int_{Q_0} g dy) \leq \|g\|_p$. Thus we have reduced to the problem of approximating $u_1 \approx f_1$.

(2) Replacing $g$ by $g_1$, it follows from step (1) that we may assume that $\text{supp } g \subset Q_0 \subset \mathcal{D}$ and $f_{Q_0} = 0$. Denote by $u$ the dyadic average extension of $g$, and write $u_Q := |Q|^{-1} \int_Q g(y) dy$. We construct the approximant $f$ using the following stopping time argument. Given any cube $Q \in \mathcal{D}$, define stopping cubes

$$
\omega(Q) := \{ \text{maximal } R \in \mathcal{D} \text{ such that } R \subset Q \text{ and } |u_R - u_Q| \geq \epsilon M_D g(R) \}.
$$

Define generations of stopping cubes under $Q_0$ inductively as follows.

$$
\omega_0 := \{Q_0\}, \quad \omega_1 := \omega(Q_0), \quad \omega_{k+1} := \bigcup_{Q \in \omega_k} \omega(Q), \quad k = 1, 2, \ldots
$$

$$
\omega_s := \bigcup_{k=0}^\infty \omega_k.
$$

Furthermore, for $Q \in \omega_s$ we define the region

$$
\Omega(Q) := \hat{Q} \setminus \bigcup_{R \in \omega(Q)} \overline{R} \subset \hat{Q}.
$$

We define $f$ to be the locally constant function in $\mathbb{R}^{1+n}$ which takes the value $u_Q$ on $\Omega(Q)$ for each $Q \in \omega_s$, i.e.,

$$
f_R := u_Q, \quad \text{when } W_R \subset \Omega(Q), Q \in \omega_s,
$$

and $f = 0$ on $\mathbb{R}^{1+n} \setminus \hat{Q}_0$. From this construction it is clear that $f$ has non-tangential limits almost everywhere. To verify that $\|N(f - u)\|_p \leq \epsilon \|g\|_p$, we note directly from the stopping condition that

$$
N_D(f - u)(x) = \sup_{Q \ni x, Q \subset \mathcal{D}} |f_Q - u_Q| \leq \epsilon \sup_{x \in Q \subset \mathcal{D}} M_D g(Q) = \epsilon M_D g(x),
$$

from which the estimate follows.

(3) We next establish the main estimate, namely that of $C(\partial_1 f)$. We fix $Q_1 \subset \mathcal{D}$ with $Q_1 \subset Q_0$ and estimate

$$
\sum_{Q \in \omega_s, Q \subset Q_1} |u_Q - u_{Q_1}| \leq \frac{1}{\epsilon} \sum_{Q \in \omega_s, Q \subset Q_1} |u_Q - u_{Q_1}|^2 \frac{|Q|}{M_D g(Q)},
$$

where we write $Q_s$ for the stopping parent of $Q$, that is the smallest $Q_s \in \omega_s$ such that $Q_s \supset Q$, and exceptionally $(Q_0)_s := Q_0$.

Define the square function

$$
Sg(x) := \left( \sum_{Q \in \omega_s, Q \subset Q_1} |\langle g \rangle_Q - \langle g \rangle_{Q_s}|^2 1_Q(x) \right)^{1/2}, \quad \langle u \rangle_Q := |Q|^{-1} \int_Q g(y) dy.
$$
Recalling that \( u_Q = \langle g \rangle_Q \) for \( Q \in \omega_* \), Lemma 3.1 gives
\[
\sum_{Q \in \omega_*, Q \subset Q_1} |u_Q - u_{Q_*}|^2 \frac{|Q|}{M_D g(Q)} \lesssim \sum_{Q \in \omega_*, Q \subset Q_1} |\langle g \rangle_Q - \langle g \rangle_Q|^2 \int_{Q_1} 1_Q(x) \frac{dx}{M_D g(x)} = \int_{Q_1} |Sg(x)|^2 \frac{dx}{M_D g(x)}.
\]

Writing
\[
(M_D g)^{-1} = 1 \cdot ((M_D g)^\gamma)^{1-g},
\]
for some \( \gamma \in (0, 1) \) and \( q = 1 + 1/\gamma \in (2, \infty) \), it follows that \( (M_D g)^\gamma \in A_1(dx) \) and \( (M_D g)^{-1} \in A_q(dx) \subset A_\infty(dx) \), with \( A_q \) constants independent of \( g \), by well known properties of Muckenhoupt weights. We now apply Proposition 4.1 below, with the collection of cubes \( \omega_* := \{ Q \in \omega_* : Q \subset Q_1 \} \), the function
\[
\tilde{g} := \begin{cases} g(x) - |Q_1|^{-1} \int_{Q_1} g dx, & x \in Q_1, \\ 0, & x \notin Q_1, 
\end{cases}
\]
the weight \( w := (M_D g)^{-1} \) and \( p = 2 \). This gives
\[
\int_{\mathbb{R}^n} |S_{\omega_*} \tilde{g}|^2 dw \lesssim \int_{\mathbb{R}^n} |M_D \tilde{g}|^2 dw = \int_{Q_1} |M_D \tilde{g}|^2 dw \lesssim \int_{Q_1} |M_D g|^2 dw.
\]
Thus
\[
\int_{Q_1} |Sg(x)|^2 dw \lesssim \int_{Q_1} |S_{\omega_*} g(x)|^2 dw + \int_{Q_1} |M_D g|^2 dw
\leq \int_{\mathbb{R}^n} |S_{\omega_*} \tilde{g}(x)|^2 dw + \int_{Q_1} |M_D g|^2 dw \lesssim \int_{Q_1} |M_D g|^2 dw \leq |Q_1| \inf_{Q_1} M_D (M_D g),
\]
and so
\[
\|C(\partial_t f)\|_p \lesssim \epsilon^{-1} \|M_D (M_D g)\|_p \lesssim \epsilon^{-1} \|g\|_p.
\]

(4) To complete the proof, we use Lemma 3.2 below and obtain the Carleson estimate
\[
\|C(\nabla f)\|_p \approx \|C(\partial_t f)\|_p + \|C(\nabla_x f)\|_p \lesssim \|C(\partial_t f)\|_p + \|M_D g\|_p \lesssim \|g\|_p.
\]

\[\square\]

**Lemma 3.2.** Let \( u \) be a function in \( \mathbb{R}^{1+n}_+ \) which is constant on dyadic Whitney regions, and let \( Q \in \mathcal{D} \). Uniformly for such \( u \) and \( Q \), we have the estimate
\[
\iint_{Q} |\nabla_x u| dxdt \lesssim \iint_{Q} |\partial_t u| dxdt + |Q| \sum_{Q'} |u_{Q'}|,
\]
where the last sum is over \( Q' \in \mathcal{D} \) with \( \ell(Q') = \ell(Q) \), \( \partial Q' \cap \partial Q \neq \emptyset \).

Note the obvious meaning of \( \iint_{Q} \). The contribution from \( \partial Q \cap \mathbb{R}^{1+n}_+ \) is to be counted.

**Proof.** Fix a dyadic cube \( Q \). Consider a contribution to \( \nabla_x u \) from the jump across \( \partial W_R \cap \partial W_S \subset Q \), where \( \ell(R) = \ell(S) \). Go up through ancestors to a common dyadic ancestor \( R_N = S_N \), and write
\[
R = R_0 \subset R_1 \subset \ldots \subset R_N = S_N \supset \ldots \supset S_1 \supset S_0 = S.
\]
If $R_N = S_N \subset Q$, then we estimate

$$|u_R - u_S| R \lesssim \sum_{k=1}^{N} 2^{-nk}(|u_{R_k} - u_{R_{k-1}}| R_k) + \sum_{k=1}^{N} 2^{-nk}(|u_{S_k} - u_{S_{k-1}}| S_k).$$

For some fixed sub-cube $R_k \not\subset Q$, there arise in this way one such term $|u_{R_k} - u_{R_{k-1}}| R_k$ from each sub-cube $R$ of $R_k$ such that $(\partial R) \cap (\partial R_k) \neq \emptyset$. There are at most $C2^{2(n-1)k}$ such sub-cubes with $\ell(R) = 2^{-k} \ell(R_k)$.

If $R_N = S_N \not\subset Q$, then we estimate as above, but stop at $|R_k| = |Q_k| = |Q|$, and we obtain two extra terms

$$2^{-nk}|u_{R_k}| Q + 2^{-nk}|u_{S_k}| Q.$$

Summing up, using $\sum_{0}^{\infty} 2^{(n-1)k}2^{-nk} = 2$, we get

$$\int \int_Q |\nabla u| dt dx \lesssim \int \int_Q |\partial_t u| dt dx + |Q| \sum_{Q'} |u_{Q'}|.$$

\[\square\]

4. A DYADIC WEIGHTED STOPPED SQUARE FUNCTION ESTIMATE

Let $\omega_\ast \subset D$ be any collection of dyadic cubes. Given any $Q \in D$, define its stopping parent $Q_\ast$ to be the minimal $Q_\ast \in \omega_\ast$ such that $Q_\ast \not\subset Q$. If no such $Q_\ast$ exists, we let $Q_\ast := Q$. Define the stopped square function

$$S_{\omega_\ast} u(x) := \left( \sum_{Q \in \omega_\ast} |u_Q - u_{Q_\ast}|^2 1_Q(x) \right)^{1/2}, \quad x \in \mathbb{R}^n.$$

We prove the following weighted estimates for $S_{\omega_\ast}$, inspired by the work of Gundy and Wheeden [5, Thm. 2] for the non-stopped square function (i.e., case $\omega_\ast = D$).

**Proposition 4.1.** Fix a Muckenhoupt weight $w \in A_\infty(dx)$ and an exponent $1 \leq p < \infty$. Then we have the stopped square function estimate

$$\|S_{\omega_\ast} u\|_{L_p(\mathbb{R}^n; w)} \lesssim \|M_D u\|_{L_p(\mathbb{R}^n; w)},$$

uniformly for any collection of dyadic cubes $\omega_\ast$.

**Proof.** It suffices to prove a good lambda inequality

$$w(\{S_{\omega_\ast} u > 2\lambda, M_D u < \gamma \lambda\}) \lesssim \gamma^\delta w(\{S_{\omega_\ast} u > \lambda\}),$$

for some $\delta > 0$. By the $A_\infty$ assumption, this will follow from a Lebesgue measure estimate

$$|\{S_{\omega_\ast} u > 2\lambda, M_D u < \gamma \lambda\} \cap Q| \lesssim \gamma |Q|,$$

for any maximal dyadic cube $Q \subset \{S_{\omega_\ast} u > \lambda\}$. To this end, assume that $x \in \{S_{\omega_\ast} u > 2\lambda, M_D u < \gamma \lambda\} \cap Q$. Then

$$4\lambda^2 < \sum_{R \in \omega_\ast, R \subset Q} |u_R - u_{R_\ast}|^2 1_R(x) + \sum_{R \in \omega_\ast, R \subset Q \not\subset R_\ast} |u_R - u_{R_\ast}|^2 1_R(x)$$

$$+ \sum_{R \in \omega_\ast, R \supseteq Q} |u_R - u_{R_\ast}|^2 1_R(x) \leq S_{\omega_\ast}(u 1_Q)(x) + 4(\gamma \lambda)^2 + \lambda^2,$$
using that $M_Du(x) < \gamma \lambda$ for the second term and the maximality of $Q$ for the last term. Therefore, assuming $\gamma < 1/2$, we have $S_{\omega_\gamma}(u1_Q)(x) > \lambda$, so

$$\{S_{\omega_\gamma}u > 2\lambda, M_Du < \gamma \lambda\} \cap Q \subset \{S_{\omega_\gamma}(u1_Q) > \lambda\}.$$ 

From Lemma 4.2 below, we get the estimate

$$\|S_{\omega_\gamma}(u1_Q) > \lambda\| \lesssim \lambda^{-1} \int_Q |u| dx.$$ 

We may assume that $\{S_{\omega_\gamma}u > 2\lambda, M_Du < \gamma \lambda\} \cap Q \neq \emptyset$, and in particular that $\int_Q |u| dx \leq \gamma \lambda|Q|$. Put together, this proves that $\|S_{\omega_\gamma}(u1_Q) > \lambda\| \lesssim \gamma|Q|$. □

**Lemma 4.2.** The stopped square function $S_{\omega_\gamma}$ defined above, has estimates

$$\|S_{\omega_\gamma}u > \lambda\| \lesssim \lambda^{-1} \|u\|_{L^1(\mathbf{R}^n)}, \quad \lambda > 0,$$

$$\|S_{\omega_\gamma}u\|_{L^2(\mathbf{R}^n)} \lesssim \|u\|_{L^2(\mathbf{R}^n)};$$

uniformly for any collection of dyadic cubes $\omega_\gamma$.

A standard Calderón–Zygmund decomposition argument yields the weak $L_1$ estimate, given the $L_2$ estimate. This $L_2$ estimate is in turn proved by a well known martingale square functions estimate, see for example Garnett [4, Ch. VIII, Lem. 6.4]. For completeness, we include the details of the proof.

**Proof.** (a) For the $L_2$ estimate, we write $\omega_\gamma = \bigcup_{k=\infty}^{\infty} \omega_k$, where the cubes in $\omega_k$ are disjoint and $\omega_{k-1} = \{Q_\gamma : Q \in \omega_k\}$. We define the martingale $\{u_k\}_{k=\infty}^{\infty}$, where

$$u_k(x) := \begin{cases} |Q|^{-1} \int_Q u(y) dy, & x \in Q \in \omega_k, \\ u(x), & x \notin \bigcup_{Q \in \omega_k} Q. \end{cases}$$

This yields

$$\|S_{\omega_\gamma}u\|_2^2 = \sum_k \sum_{Q \in \omega_k} |u_Q - u_{Q_\gamma}|^2 |R| \leq \sum_k \int_{\mathbf{R}^n} |u_{k+1} - u_k|^2 dx$$

$$= \sum_k \int_{\mathbf{R}^n} (u_{k+1}^2 + u_k^2 - 2u_{k+1}u_k) dx = \sum_k \int_{\mathbf{R}^n} (u_{k+1}^2 - u_k^2) dx \leq \int_{\mathbf{R}^n} u_k^2 dx,$$

where we have used that $\int u_k u_{k+1} dx = \int u_k^2 dx$.

(b) Let $Q_k$ denote the maximal dyadic cubes contained in $\{M_Du > \lambda\}$. Write $u = g + \sum_k b_k$, where $|g| \leq \lambda$ and $\text{supp} \ b_k \subset Q_k$ with $\int_{Q_k} b_k = 0$. The stated estimate follows from the two estimates

$$\|S_{\omega_\gamma}g > \lambda/2\| \lesssim \lambda^{-2} \int |S_{\omega_\gamma}g|^2 dx \lesssim \lambda^{-2} \int |g|^2 dx \lesssim \lambda^{-1} \int |g| dx \leq \lambda^{-1} \int |u| dx,$$

using (a) and that $\int_{Q_k} |u| dx \approx \lambda|Q_k|$, and

$$\|\{\sum_{k} b_k > \lambda/2\}\| \lesssim \sum_k |Q_k| = \|M_Du > \lambda\| \lesssim \lambda^{-1} \int |u| dx,$$

using that $\text{supp} \ S_{\omega_\gamma}b_k \subset Q_k$ and the weak $L_1$ bound of the Hardy–Littlewood maximal function. □
REFERENCES

[1] CARLESON, L. Interpolations by bounded analytic functions and the corona problem. Ann. of Math. (2) 76 (1962), 547–559.
[2] COIFMAN, R. R., MEYER, Y., AND STEIN, E. M. Some new function spaces and their applications to harmonic analysis. J. Funct. Anal. 62, 2 (1985), 304–335.
[3] FEFFERMAN, C., AND STEIN, E. M. Hp spaces of several variables. Acta Math. 129, 3–4 (1972), 137–193.
[4] GARNETT, J. Bounded analytic functions. No. 96 in Pure and Applied Mathematics. Academic Press, New York-London, 1981.
[5] GUNDY, R. F., AND WHEEDEN, R. L. Weighted integral inequalities for the nontangential maximal function, Lusin area integral, and Walsh-Paley series. Studia Math. 49 (1973/74), 107–124.
[6] HYTÖNEN, T., AND ROSÉN, A. On the Carleson duality. Ark. Mat. 51 (2013), 293–313.
[7] VAROPOULOS, N. BMO functions and the Ïø-equation. Pacific J. Math. 71, 1 (1977), 221–273.
[8] VAROPOULOS, N. A remark on functions of bounded mean oscillation and bounded harmonic functions. Pacific J. Math. 74, 1 (1978), 257–259.

Tuomas Hytönen, Department of Mathematics and Statistics, P.O. Box 68 (Gustaf Hällströmin katu 2b), FI-00014 University of Helsinki, Finland
E-mail address: tuomas.hytonen@helsinki.fi

Andreas Rosén, Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Göteborg, Sweden
E-mail address: andreas.rosen@chalmers.se