

Tuomas Hytönen · Andreas Rosén<sup>1</sup>

## Bounded variation approximation of $L_p$ dyadic martingales and solutions to elliptic equations

Received July 1, 2015 and in revised form June 16, 2016

**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. We also prove  $L_p$  Carleson approximability results for solutions to elliptic non-smooth divergence form equations, which generalize results in the case  $p = \infty$  by Hofmann, Kenig, Mayboroda and Pipher.

**Keywords.** Extension map, Carleson functional, approximability, stopping time argument, Corona Theorem, elliptic equation, bounded variation

### 1. Introduction

Estimates of traces  $u|_{\partial D}$  of functions  $u : D \rightarrow \mathbb{R}$  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 such 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

---

<sup>1</sup>Formerly Andreas Axelsson.

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

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

*Mathematics Subject Classification (2010):* Primary 42B37; Secondary 35J15, 42B25

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 best known trace result is the Sobolev trace theorem. It states that the trace map

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

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  $B_{p,q}^s$ , where the trace map

$$\gamma : B_{p,q}^s(\mathbb{R}_+^{1+n}) \rightarrow B_{p,q}^{s-1/p}(\mathbb{R}^n)$$

is bounded and surjective when  $s > 1/p$ . Here also  $\gamma : B_{p,1}^{1/p}(\mathbb{R}_+^{1+n}) \rightarrow L_p(\mathbb{R}^n)$  is bounded and surjective whenever  $1 \leq p < \infty$ , whereas the  $L_2$  Sobolev scale of spaces is  $H^s = B_{2,2}^s$ .

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

$$C\mu(x) := \sup_{Q \ni x} \frac{1}{|Q|} \iint_{\widehat{Q}} d|\mu|(t, y)$$

for locally finite measures  $\mu$  on  $\mathbb{R}_+^{1+n}$ . Here the supremum is over all cubes  $Q \subset \mathbb{R}^n$  containing  $x$ , and  $\widehat{Q} := (0, \ell(Q)) \times Q$  denotes the Carleson box above  $Q$ , with side length  $\ell(Q)$ . Note that  $C(\nabla u)$  is well defined for any  $u \in \text{BV}_{\text{loc}}(\mathbb{R}_+^{1+n})$  of locally bounded variation.

In addition to the quantitative condition involving the norm  $\|C(\nabla u)\|_{L_p(\mathbb{R}^n)}$ , we also need some decay at infinity, which can be assumed in various forms. The weakest condition suitable for our needs is

$$\langle |u| \rangle_{W(t,x)} := \frac{1}{|W(t,x)|} \iint_{W(t,x)} |u(s,y)| \, ds \, dy \xrightarrow{t \rightarrow \infty} 0$$

for all  $x \in \mathbb{R}^n$ , where 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  $c_1 > 0$ . The above convergence is in particular implied by the stronger quantitative bound

$$Nu \in L_p(\mathbb{R}^n),$$

where

$$Nu(x) := \text{ess sup}_{|y-x|<t} |u(t,y)|, \quad x \in \mathbb{R}^n,$$

denotes the non-tangential maximal function. Indeed, it is easy to check that  $\langle |u| \rangle_{W(t,x)} \leq \inf_{y \in B(x,ct)} Nu(y) \rightarrow 0$  as  $t \rightarrow \infty$  if  $Nu \in L_p(\mathbb{R}^n)$ . Thus we have the nested interior function spaces

$$\begin{aligned} \mathcal{V}_p^0 &:= \left\{ u \in \text{BV}_{\text{loc}}(\mathbb{R}_+^{1+n}) : C(\nabla u) \in L_p(\mathbb{R}^n), \langle |u| \rangle_{W(t,x)} \xrightarrow{t \rightarrow \infty} 0 \text{ for all } x \in \mathbb{R}^n \right\} \supset \\ \mathcal{V}_p^N &:= \{ u \in \text{BV}_{\text{loc}}(\mathbb{R}_+^{1+n}) : C(\nabla u) \in L_p(\mathbb{R}^n), Nu \in L_p(\mathbb{R}^n) \} \supset \\ \tilde{\mathcal{V}}_p^N &:= \{ u \in C^1(\mathbb{R}_+^{1+n}) : C(\nabla u) \in L_p(\mathbb{R}^n), Nu \in L_p(\mathbb{R}^n) \}. \end{aligned}$$

With the help of these spaces, we formulate our first  $L_p$  extension result:

**Theorem 1.1.** *Let  $1 < p < \infty$ . Consider the normed linear function space  $\mathcal{V}_p^0$  with norm  $\|C(\nabla(\cdot))\|_p$ . Then the trace  $\gamma u$  of any  $u \in \mathcal{V}_p^0$ , is well defined almost everywhere in the sense of convergence of Whitney averages*

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

The trace map  $\gamma : \mathcal{V}_p^0 \rightarrow L_p(\mathbb{R}^n)$  is well defined, and there exists  $c_p < \infty$  such that

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

for all  $u \in \mathcal{V}_p^0$ . Moreover, the trace map  $\gamma$  is surjective, and given any  $g \in L_p(\mathbb{R}^n)$  there exists an extension  $u \in \mathcal{V}_p^0$  such that  $\gamma u = g$ , with estimates

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

In fact, this extension may be chosen so that  $u \in \tilde{\mathcal{V}}_p^N$ , with the additional estimate

$$\|Nu\|_{L_p(\mathbb{R}^n)} \leq c_p \|g\|_{L_p(\mathbb{R}^n)},$$

and the pointwise non-tangential limits  $\lim_{(t,y) \rightarrow (0,x), |y-x| < \alpha t} u(t, y) = g(x)$  exist at each Lebesgue point of  $g$ , for any fixed  $\alpha < \infty$ .

We remark that the extension operator  $g \mapsto u$  is non-linear, even though  $\gamma$  itself is of course linear.

The corresponding trace result in the case  $p = \infty$ , proved by Varopoulos [16, 17], is that there is a bounded and surjective trace map

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

and a corresponding non-linear bounded extension operator, where  $\text{BMO}(\mathbb{R}^n)$  stands for the John–Nirenberg space of functions of bounded mean oscillation. Following Varopoulos [17], we obtain the extensions in Theorem 1.1 from a result on approximate extensions of Lebesgue functions on  $\mathbb{R}^n$ . This main component of Theorem 1.1, contained in our Theorem 1.2, generalizes well known techniques in the endpoint case  $p = \infty$  related

to the Corona Theorem, first proved by Carleson [3]. Our proof of Theorem 1.2, though, is more in the spirit of Garnett [9, Ch. VIII, Thm. 6.1].

The statement below refers to dyadic versions of the non-tangential maximal functional, the Carleson functional and the Hardy–Littlewood maximal functional, defined respectively by

$$\begin{aligned}
 N_{\mathcal{D}}f(x) &:= \sup_{Q: x \in Q \in \mathcal{D}} \sup_{(t,y) \in W_Q} |f(t,y)|, & W_Q &:= (\ell(Q)/2, \ell(Q)) \times Q, \\
 C_{\mathcal{D}}f(x) &:= \sup_{Q: x \in Q \in \mathcal{D}} \frac{1}{|Q|} \int_{\widehat{Q}} |f(t,y)| dt dy, & \widehat{Q} &:= (0, \ell(Q)) \times Q, \\
 M_{\mathcal{D}}g(x) &:= \sup_{Q: x \in Q \in \mathcal{D}} \frac{1}{|Q|} \int_Q |g(y)| dy,
 \end{aligned}$$

where  $\mathcal{D}$  is a system of dyadic cubes in  $\mathbb{R}^n$ .

**Theorem 1.2.** *Fix  $1 < p < \infty$ . Consider  $g \in L_p(\mathbb{R}^n)$  and define the dyadic average extension*

$$u(t,x) := \int_Q g(y) dy, \quad (t,x) \in W_Q,$$

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  $f : \mathbb{R}_+^{1+n} \rightarrow \mathbb{R}$  which is constant on each dyadic Whitney region, with pointwise estimates

$$N_{\mathcal{D}}(f - u) \leq \epsilon M_{\mathcal{D}}g, \quad C_{\mathcal{D}}(\nabla f) \lesssim \epsilon^{-1} M_{\mathcal{D}}(M_{\mathcal{D}}g),$$

and implied norm estimates

$$\|N(f - u)\|_{L_p(\mathbb{R}^n)} \leq \epsilon \|g\|_{L_p(\mathbb{R}^n)}, \quad \|C(\nabla f)\|_{L_p(\mathbb{R}^n)} \lesssim \epsilon^{-1} \|g\|_{L_p(\mathbb{R}^n)}.$$

Moreover, for any fixed  $\alpha < \infty$ , the non-tangential limits  $\lim_{(t,y) \rightarrow (0,x), |y-x| < \alpha t} f(t,y) =: f(0,x)$  exist almost everywhere, with  $\|f(0, \cdot) - g\|_{L_p(\mathbb{R}^n)} \leq \epsilon \|g\|_{L_p(\mathbb{R}^n)}$ .

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

$$u(x) := \sum_{Q \subset (0,1), |Q| \geq 2^{-k}} (\chi_{Q^l}(x) - \chi_{Q^r}(x)), \quad x \in \mathbb{R},$$

where the sum is over dyadic subintervals of  $(0, 1)$  of length at least  $2^{-k}$ , and the summand involves the characteristic functions of the left and right dyadic children of  $Q$ . Then one checks that  $\|u\|_p \lesssim \sqrt{k}$ , whereas the dyadic average extension  $u(t,x)$  is seen to satisfy

$$C(\nabla u)(x) \gtrsim k \quad \text{for all } x \in (0, 1).$$

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 the endpoint BMO case.

Our second main result is an approximation result analogous to Theorem 1.2, but with the dyadic martingale  $u$  replaced by a solution  $u$  to an elliptic divergence form equation  $\operatorname{div} A \nabla u = 0$  in  $\mathbb{R}_+^{1+n}$ . This generalizes results in the endpoint case  $p = \infty$  by Garnett [9] for the Laplace equation in  $\mathbb{R}_+^2$ , by Dahlberg [5] for the Laplace equation on Lipschitz domains in  $\mathbb{R}^n$ , and by Kenig, Koch, Pipher and Toro [13] and Hofmann, Kenig, Mayboroda and Pipher [11] for divergence form equations on Lipschitz domains.

**Theorem 1.3.** *Fix  $1 < p < \infty$  and coefficients  $A \in L_\infty(\mathbb{R}^n; \mathcal{L}(\mathbb{R}^{1+n}))$  which are accretive in the sense that there exists  $\lambda_A > 0$  such that*

$$(A(x)v, v) \geq \lambda_A |v|^2$$

*for almost all  $x \in \mathbb{R}^n$  and all  $0 \neq v \in \mathbb{R}^{1+n}$ . Then for any  $0 < \epsilon < 1$ , there exists  $c_\epsilon < \infty$  such that the following holds. Given any weak solution  $u : \mathbb{R}_+^{1+n} \rightarrow \mathbb{R}$  to the  $t$ -independent real scalar, but possibly non-symmetric, divergence form elliptic equation*

$$\operatorname{div}_{t,x} A(x) \nabla_{t,x} u(t, x) = 0,$$

*with  $Nu \in L_p(\mathbb{R}^n)$ , there exists a function  $f$  in  $\mathbb{R}_+^{1+n}$  of locally bounded variation with estimates*

$$\|N(f - u)\|_{L_p(\mathbb{R}^n)} \leq \epsilon \|Nu\|_{L_p(\mathbb{R}^n)}, \quad \|C(\nabla f)\|_{L_p(\mathbb{R}^n)} \leq c_\epsilon \|Nu\|_{L_p(\mathbb{R}^n)}.$$

We shall informally refer to such an  $f$  as an *approximant* of  $u$ . Note that the functions  $f$  and  $u$  share the same domain of definition,  $\mathbb{R}_+^{1+n}$ , and no extensions are involved here, in contrast to Theorem 1.2, which established an *approximate extension*  $f$  on  $\mathbb{R}_+^{1+n}$  of an initial function  $g$  on  $\mathbb{R}^n$ . While Theorem 1.2 also featured a function  $u$  on  $\mathbb{R}_+^{1+n}$  in a seemingly similar role to Theorem 1.3, the actual role of  $u$  in Theorem 1.2 was mainly auxiliary, as an intermediate object in the construction of  $f$ , while in Theorem 1.3 we regard the solution  $u$  itself, rather than its boundary limit, as the primary object of interest that we wish to approximate.

Again, our proof in fact gives the pointwise bounds

$$N_{\mathcal{D}^\delta}(f - u) \leq \epsilon M_{\mathcal{D}^\delta}(Nu) \quad \text{and} \quad C_{\mathcal{D}^\delta}(\nabla f) \leq c_\epsilon M_{\mathcal{D}^\delta}(Nu),$$

where  $\mathcal{D}^\delta$  denotes modified dyadic versions of the maximal and Carleson functionals (see Section 6), from which the asserted norm bounds are immediate by the maximal inequality. The dependence of  $c_\epsilon$  on  $\epsilon$  given by our proof, is certainly worse than  $1/\epsilon$ . We also note that when  $p$  is large enough, it follows from [11] that the  $L_p$  Dirichlet problem is well posed since the  $L$ -harmonic measure, for  $L = -\operatorname{div}_{t,x} A(x) \nabla_{t,x}$ , is in  $A_\infty(dx)$ , and therefore the norm  $\|Nu\|_{L_p(\mathbb{R}^n)}$  can be replaced by  $\|u|_{\mathbb{R}^n}\|_{L_p(\mathbb{R}^n)}$  in Theorem 1.3 in this case.

The previously studied endpoint case  $p = \infty$  of Theorem 1.3 is a key tool in the study of boundary value problems for elliptic equations  $Lu = 0$  as above in [13, 11]. The existence of such an approximant  $f \approx u$  in the  $L_\infty$  norm is referred to as  $\epsilon$ -approximability (of solutions), and it is proved in [13] that it implies the  $A_\infty$  property of the harmonic measure. However, the latter is well known by Dahlberg, Jerison and Kenig [6] to imply the comparability of non-tangential maximal functions and square functions of solutions  $u$ , which in turn is a key tool in [13, 11] in the construction of approximants  $f \approx u$ . Therefore bounded variation approximability  $f \approx u$ ,  $A_\infty$  control of harmonic measure and  $N \approx S$  comparability turn out to be equivalent. The importance of the approximability property, through this circle of arguments in the  $p = \infty$  case, motivates our generalization in Theorem 1.3 to the case  $p < \infty$ . We show that these three properties are also equivalent when  $p = \infty$  in the approximability property is replaced by  $n/(n-1) \leq p < \infty$ . See Section 5 for the detailed statements. In particular, we extend [13, Thm. 2.3] to  $n/(n-1) \leq p < \infty$ .

The outline of the paper is as follows. In Section 2, we survey the basic estimates for the functionals defining our spaces. In Section 4 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 3. Finally in Section 6 we prove Theorem 1.3, using local  $N \approx S$  estimates which we borrow from [11], as a replacement of the stopped square function estimates which we used in the martingale case of Theorem 1.2.

## 2. The basic functionals

In this section, we collect well known facts concerning the functionals that we use to define several 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) are denoted by  $Q, R, S, \dots$ , and we assume that they are open. The Carleson box above a cube  $Q \subset \mathbb{R}^n$  is denoted

$$\widehat{Q} := (0, \ell(Q)) \times Q \subset \mathbb{R}_+^{1+n},$$

where  $\ell(Q)$  denotes the sidelength of  $Q$ . We write  $cQ$  to denote the cube with the same centre as  $Q$  but with  $\ell(cQ) = c\ell(Q)$ .

Let  $\mathcal{D} = \bigcup_{j \in \mathbb{Z}} \mathcal{D}_j$  denote a system of dyadic cubes in  $\mathbb{R}^n$ , with  $\mathcal{D}_j$  being the cubes of sidelength  $\ell(Q) = 2^{-j}$ , such 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  $(t, x) \in \mathbb{R}_+^{1+n}$  is defined to be

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

where  $c_0, c_1 > 0$  are fixed parameters.

The Hardy–Littlewood maximal function of  $f \in L_1^{\text{loc}}(\mathbb{R}^n)$  that we use is

$$Mf(x) := \sup_{Q \ni x} \int_Q |f(y)| dy,$$

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  $M_{\mathcal{D}}f(x)$ . We also require the following truncated (to large cubes) version of the dyadic Hardy–Littlewood maximal function:

$$M_{\mathcal{D}}f(Q) = \sup_{R \supset Q, R \in \mathcal{D}} \int_R |f(y)| dy, \quad Q \in \mathcal{D}. \tag{1}$$

**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*  $Af$  as

$$\begin{aligned} Nf(x) &:= \operatorname{ess\,sup}_{|y-x| < \alpha t} |f(t, y)|, \\ Cf(x) &:= \sup_{Q \ni x} |Q|^{-1} \iint_{\widehat{Q}} |f(t, y)| dt dy, \\ Af(x) &:= \iint_{|y-x| < \alpha t} |f(t, y)| t^{-n} dt dy. \end{aligned}$$

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 on 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

$$\begin{aligned} N_{\mathcal{D}}f(x) &:= \sup_{x \in Q \in \mathcal{D}} |f_Q|, \\ C_{\mathcal{D}}f(x) &:= \sup_{x \in Q \in \mathcal{D}} |Q|^{-1} \sum_{R \in \mathcal{D}, R \subset Q} |f_R| |W_R|, \\ A_{\mathcal{D}}f(x) &:= \sum_{x \in Q \in \mathcal{D}} |f_Q| \ell(Q). \end{aligned}$$

The reader is invited to check that these definitions agree with those given in the Introduction without assuming the constant values over the regions  $W_Q$ .

We want to point out that, throughout this paper, we are using the measure  $dt dx$  and not the measure  $t^{-1} dt dx$ , although the latter is quite 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  $\|N^{(\alpha)} f\|_p \approx \|N^{(\beta)} f\|_p$  for any  $1 \leq p \leq \infty$ , and  $\|A^{(\alpha)} f\|_p \approx \|A^{(\beta)} f\|_p$  for any  $1 \leq p < \infty$ .*

*Proof.* The estimates for  $Nf$  are proved in Fefferman and Stein [8, Lem. 1]. To prove the estimate for the  $A$ -functional, we follow Coifman, Meyer and Stein [4, Prop. 4, case  $2 \leq p < \infty$ ] and consider  $0 < \alpha < \beta < \infty$ : Dualize against  $\|h\|_{p'} = 1$  to get

$$\begin{aligned} \|A^{(\beta)} f\|_p &= \int_{\mathbb{R}^n} \left( \iint_{|y-x| < \beta t} |f(t, y)| t^{-n} dt dy \right) h(x) dx \\ &= \iint_{\mathbb{R}_+^{1+n}} |f(t, y)| \left( t^{-n} \int_{|x-y| < \beta t} h(x) dx \right) dt dy \\ &\lesssim \iint_{\mathbb{R}_+^{1+n}} |f(t, y)| \left( t^{-n} \int_{|x-y| < \alpha t} Mh(x) dx \right) dt dy \\ &= \int_{\mathbb{R}^n} A^{(\alpha)} f(x) Mh(x) dx \lesssim \|A^{(\alpha)} f\|_p. \quad \square \end{aligned}$$

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

**Proposition 2.3.** *We have*

$$\begin{aligned} \|Nf\|_p &\approx \|N_{\mathcal{D}}f\|_p, & 1 \leq p \leq \infty, \\ \|Cf\|_p &\approx \|C_{\mathcal{D}}f\|_p, & 1 < p \leq \infty, \\ \|Af\|_p &\approx \|A_{\mathcal{D}}f\|_p, & 1 \leq p < \infty, \end{aligned}$$

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

*Proof.* For proofs for  $N$  and  $C$ , we refer to [12]. Consider now the area functional  $A$ . As in the proof of Proposition 2.2, the proof is an adaption of [4, Prop. 4, case  $2 \leq p < \infty$ ]. Dualize against  $\|h\|_{p'} = 1$  to get

$$\begin{aligned} \|Af\|_p &= \int_{\mathbb{R}^n} \left( \iint_{|y-x| < \alpha t} |f(t, y)| t^{-n} dt dy \right) h(x) dx \\ &= \iint_{\mathbb{R}_+^{1+n}} |f(t, y)| \left( t^{-n} \int_{|x-y| < \alpha t} h(x) dx \right) dt dy \\ &\lesssim \sum_{Q \in \mathcal{D}} |f_Q| |W_Q| \left( \int_Q (Mh)(x) dx \right) = \int_{\mathbb{R}^n} (A_{\mathcal{D}}f)(Mh) dx \lesssim \|A_{\mathcal{D}}f\|_p. \quad \square \end{aligned}$$

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 depends 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 simply defines the function space  $L_1(\mathbb{R}_+^{1+n})$ .

Proposition 2.4 is a reformulation of [4, Thm. 3]. The proof below contains some novelties in the estimate  $A \lesssim C$ , by using a duality argument rather than a good lambda estimate.

*Proof of Proposition 2.4.* For  $C \lesssim A$  we have

$$\begin{aligned} M(Af)(x) &= \sup_{Q \ni x} \int_Q \left( \iint_{|y-x| < \alpha t} |f(t, y)| t^{-n} dt dy \right) dx \\ &= \sup_{Q \ni x} |Q|^{-1} \iint \left( t^{-n} \int_{|x-y| < \alpha t, x \in Q} dx \right) |f(t, y)| dt dy \gtrsim Cf(x). \end{aligned}$$

For  $A \lesssim C$ , we argue by duality with a suitable  $\|h\|_{p'} = 1$ :

$$\begin{aligned} \|Af\|_p &= \int_{\mathbb{R}^n} \left( \iint_{|y-x| < \alpha t} |f(t, y)| t^{-n} dt dy \right) h(x) dx \\ &= \iint_{\mathbb{R}_+^{1+n}} |f(t, y)| \left( t^{-n} \int_{|x-y| < \alpha t} h(x) dx \right) dt dy \\ &=: \iint_{\mathbb{R}_+^{1+n}} |f(t, y)| H(t, y) dt dy = \int_0^\infty \left( \iint_{\substack{(t,y) \in \mathbb{R}_+^{1+n} \\ H(t,y) > \lambda}} |f(t, y)| dt dy \right) d\lambda, \end{aligned}$$

where  $H(t, y)$  is defined by the penultimate equality to be the parenthetical quantity on the line above. If  $H(t, y) > \lambda$ , there is a cube  $Q$  such that  $\int_Q h(x) dx > c\lambda$  and  $(t, y) \in \widehat{Q}$ . By the Whitney covering lemma, there is a collection  $\mathcal{Q}_\lambda$  of such cubes such that the  $\widehat{Q}$  are pairwise disjoint, and the  $5\widehat{Q}$ ,  $Q \in \mathcal{Q}_\lambda$ , cover all the points  $(t, y)$  with  $H(t, y) > \lambda$ . Thus

$$\begin{aligned} \iint_{\substack{(t,y) \in \mathbb{R}_+^{1+n} \\ H(t,y) > \lambda}} |f(t, y)| dt dy &\leq \sum_{Q \in \mathcal{Q}_\lambda} \iint_{5\widehat{Q}} |f(t, y)| dt dy \leq \sum_{Q \in \mathcal{Q}_\lambda} |5Q| \inf_{x \in Q} Cf(x) \\ &\lesssim \sum_{Q \in \mathcal{Q}_\lambda} \int_Q Cf(x) dx \leq \int_{\{Mh > c\lambda\}} Cf(x) dx. \end{aligned}$$

Substituting back shows that

$$\|Af\|_p \lesssim \int_0^\infty \int_{\{Mh > c\lambda\}} Cf(x) dx d\lambda \lesssim \int_{\mathbb{R}^n} Cf(x) Mh(x) dx \lesssim \|Cf\|_p. \quad \square$$

### 3. A dyadic weighted stopped square function estimate

In this section we prove an auxiliary weighted norm inequality, which will be used in the construction of extensions in the subsequent section.

Let  $\omega_* \subset \mathcal{D}$  be any collection of dyadic cubes. Given any  $Q \in \mathcal{D}$ , define its stopping parent  $Q_*$  to be the minimal  $Q_* \in \omega_*$  such that  $Q_* \supseteq Q$ . If no such  $Q_*$  exists, we let  $Q_* := Q$ . Define the stopped square function

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

**Lemma 3.1.** *The stopped square function  $S_{\omega_*}$  has estimates*

$$\begin{aligned} |\{S_{\omega_*} u > \lambda\}| &\lesssim \lambda^{-1} \|u\|_{L_1(\mathbb{R}^n)}, \quad \lambda > 0, \\ \|S_{\omega_*} u\|_{L_2(\mathbb{R}^n)} &\lesssim \|u\|_{L_2(\mathbb{R}^n)}, \end{aligned}$$

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

A standard Calderón–Zygmund decomposition argument yields the weak  $L_1$  estimate, given the  $L_2$  estimate. The  $L_2$  estimate is in turn proved by using a well known martingale square functions estimate (see for example Garnett [9, Ch. VIII, Lem. 6.4]). For completeness, we include the details of the proof.

*Proof of Lemma 3.1.* (a) For the  $L_2$  estimate, we write  $\omega_* = \bigcup_{k=-\infty}^{\infty} \omega_k$ , where the cubes in  $\omega_k$  are disjoint and  $\omega_{k-1} = \{Q_* : Q \in \omega_k\}$ . We define the martingale  $\{u_k\}_{k=-\infty}^{\infty}$ , where

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

This yields

$$\begin{aligned} \|S_{\omega_*} u\|_2^2 &= \sum_k \sum_{Q \in \omega_k} |u_Q - u_{Q_*}|^2 |R| \leq \sum_k \int_{\mathbb{R}^n} |u_{k+1} - u_k|^2 dx \\ &= \sum_k \int_{\mathbb{R}^n} (u_{k+1}^2 + u_k^2 - 2u_{k+1}u_k) dx = \sum_k \int_{\mathbb{R}^n} (u_{k+1}^2 - u_k^2) dx \leq \int_{\mathbb{R}^n} u^2 dx, \end{aligned}$$

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

(b) Let  $Q_k$  denote the maximal dyadic cubes contained in  $\{M_{\mathcal{D}} u > \lambda\}$ . 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

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

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

$$\left| \left\{ S_{\omega_*} \left( \sum_k b_k \right) > \lambda/2 \right\} \right| \lesssim \sum_k |Q_k| = |\{M_{\mathcal{D}} u > \lambda\}| \lesssim \lambda^{-1} \int |u| dx,$$

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

The main result of this section is the following weighted estimate for  $S_{\omega_*}$ , inspired by the work of Gundy and Wheeden [10, Thm. 2] for the non-stopped square function (i.e., the case  $\omega_* = \mathcal{D}$ ).

**Proposition 3.2.** *Fix a Muckenhoupt weight  $w \in A_\infty(dx)$  and an exponent  $1 \leq p < \infty$ . Then we have the stopped square function estimate*

$$\|S_{\omega_*}u\|_{L_p(\mathbb{R}^n;w)} \lesssim \|M_{\mathcal{D}}u\|_{L_p(\mathbb{R}^n;w)},$$

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

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

$$w(\{S_{\omega_*}u > 2\lambda, M_{\mathcal{D}}u < \gamma\lambda\}) \lesssim \gamma^\delta w(\{S_{\omega_*}u > \lambda\})$$

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

$$|\{S_{\omega_*}u > 2\lambda, M_{\mathcal{D}}u < \gamma\lambda\} \cap Q| \lesssim \gamma|Q|$$

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

$$\begin{aligned} 4\lambda^2 &< \sum_{R \in \omega_*, R_* \subset Q} |u_R - u_{R_*}|^2 1_R(x) + \sum_{R \in \omega_*, R \subset Q \subsetneq R_*} |u_R - u_{R_*}|^2 1_R(x) \\ &+ \sum_{R \in \omega_*, R \supseteq Q} |u_R - u_{R_*}|^2 1_R(x) \leq S_{\omega_*}(u1_Q)(x) + 4(\gamma\lambda)^2 + \lambda^2, \end{aligned}$$

by using  $M_{\mathcal{D}}u(x) < \gamma\lambda$  for the second term and the maximality of  $Q$  for the last term. Therefore, assuming  $\gamma < 1/2$ , we have  $S_{\omega_*}(u1_Q)(x) > \lambda$ , so

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

From Lemma 3.1, we get the estimate

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

We may assume that  $\{S_{\omega_*}u > 2\lambda, M_{\mathcal{D}}u < \gamma\lambda\} \cap Q \neq \emptyset$ , and in particular  $\int_Q |u| dx \leq \gamma\lambda|Q|$ . Put together, this proves that  $|\{S_{\omega_*}u > 2\lambda, M_{\mathcal{D}}u < \gamma\lambda\} \cap Q| \lesssim \gamma|Q|$ .  $\square$

#### 4. Construction of extensions

In this section we prove Theorems 1.1 and 1.2, assuming a square function estimate which we prove in Section 3. We first prove Theorem 1.1, where we use Theorem 1.2 in the construction of extensions.

*Proof of Theorem 1.1, part I: existence and bound of the trace of  $u \in \mathcal{V}_p^0$ .* We fix  $x \in \mathbb{R}^n$  and consider two Whitney regions  $W(t_1, x)$  and  $W(t_2, x)$ , with  $t_1 < t_2$ . We estimate

$$\begin{aligned}
 & \left| |W(t_2, x)|^{-1} \iint_{W(t_2, x)} u(s, y) \, ds \, dy - |W(t_1, x)|^{-1} \iint_{W(t_1, x)} u(s, y) \, ds \, dy \right| \\
 & \quad \approx \left| \iint_{W(1,0)} (u(t_2 s, x + t_2 y) - u(t_1 s, x + t_1 y)) \, ds \, dy \right| \\
 & \quad \lesssim \iint_{W(1,0)} \int_{t_1}^{t_2} |\nabla u(t s, x + t y)| \, dt \, ds \, dy \\
 & \quad = \int_{t_1}^{t_2} \int_{t/c_0}^{c_0 t} \int_{|y'-x| < c_1 t} |\nabla u(s', y')| t^{-1-n} \, dy' \, ds' \, dt \\
 & \quad \lesssim \int_{t_1/c_0}^{c_0 t_2} \int_{|y'-x| < c_1 s'} |\nabla u(s', y')| (s')^{-n} \, dy' \, ds'. \tag{2}
 \end{aligned}$$

Since  $A(\nabla u) \in L_p(\mathbb{R}^n)$  by Proposition 2.4, we have  $A(\nabla u)(x) < \infty$  for almost all  $x \in \mathbb{R}^n$ . For such  $x$ , it follows from the above estimate that Whitney averages converge as  $t \rightarrow 0$ . Thus, in this sense we have a well defined trace almost everywhere on  $\mathbb{R}^n$ .

On the other hand, the definition of  $\mathcal{V}_p^0$  yields  $|W(t_2, x)|^{-1} \iint_{W(t_2, x)} u(s, y) \, ds \, dy \rightarrow 0$  as  $t_2 \rightarrow \infty$ . The estimate

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

follows. Indeed, if the right hand side is finite, then at almost every  $x \in \mathbb{R}^n$ , we see from (2) that the trace  $\gamma u(x)$  exists in the sense of convergence of Whitney averages, since the right hand side in (2) has zero limit as  $t_2 \rightarrow 0$ . Then letting  $t_2 \rightarrow \infty$  shows the pointwise estimate  $|\gamma u(x)| \lesssim |A(\nabla u)(x)|$ .

This completes the first part of the proof of Theorem 1.1. □

*Proof of Theorem 1.1, part II: construction of the extension assuming Theorem 1.2.* We construct the extension  $u$  of  $g \in L_p(\mathbb{R}^n)$  as follows. Define functions  $g_k, u_k$  and  $f_k, k = 0, 1, 2, \dots$ , 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

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

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

$$f := \sum_{k=0}^{\infty} f_k.$$

This is an exact extension of  $g$  since  $0 = \lim_{k \rightarrow \infty} \|g_{k+1}\|_p = \|g - \sum_{j=0}^k f_j|_{\mathbb{R}^n}\|_p$ .  
 Moreover,

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

for some fixed  $0 < \epsilon < 1$ , and similarly

$$\|Nf\|_p \leq \sum_{k=0}^{\infty} \|Nf_k\|_p \leq \sum_{k=0}^{\infty} (\|Nu_k\|_p + \epsilon \|g_k\|_p) \lesssim \sum_{k=0}^{\infty} \|g_k\|_p \lesssim \|g\|_p.$$

This shows that we have an extension  $f \in \mathcal{V}_p^N$ .

It remains to mollify  $f$  to obtain another extension

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

where  $\eta \in C_0^\infty(W(1, 0))$  has  $\iint \eta = 1$ . Then it is straightforward to verify that  $u \in \tilde{\mathcal{V}}_p^N$ , with the stated estimate of

$$\nabla u(t, x) = \iint_{\mathbb{R}_+^{1+n}} \begin{bmatrix} s & y^i \\ 0 & I \end{bmatrix} \nabla f(ts, x + ty)\eta(s, y) ds dy. \quad \square$$

For the proof of Theorem 1.2, we require the following lemma for the truncated dyadic maximal function from (1).

**Lemma 4.1.** *For any  $g \in L_1^{\text{loc}}(\mathbb{R}^n)$  and  $Q \in \mathcal{D}$ , we have*

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

*Proof.* Define

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

The weak  $L_1$  estimate for  $M_{\mathcal{D}}$  yields

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

Thus

$$\frac{|Q|}{M_{\mathcal{D}}g(Q)} \leq 2 \frac{|Q \setminus E_Q|}{M_{\mathcal{D}}g(Q)} \leq 4 \int_{Q \setminus E_Q} \frac{dx}{M_{\mathcal{D}}g(x)} \leq 4 \int_Q \frac{dx}{M_{\mathcal{D}}g(x)}. \quad \square$$

The following lemma shows that the horizontal derivatives are essentially controlled by the vertical ones, reducing the problem to controlling the latter.

**Lemma 4.2.** *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*

$$\iint_{\widehat{Q}} |\nabla_x u| dt dx \lesssim \iint_{\widehat{Q}} |\partial_t u| dt dx + |Q| \sum_{Q'} |u_{Q'}|,$$

where the last sum is over  $Q' \in \mathcal{D}$  with  $\ell(Q') = \ell(Q)$  and  $\partial Q' \cap \partial Q \neq \emptyset$ .

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

*Proof.* Fix a dyadic cube  $Q$ . Consider a contribution to  $\nabla_x u$  from the jump across  $\partial W_R \cap \partial W_S \subset \widehat{Q}$ , where  $\ell(R) = \ell(S)$ . Go up through ancestors to a common dyadic ancestor  $R_N = S_N$ , and write

$$R = R_0 \subset R_1 \subset \dots \subset R_N = S_N \supset \dots \supset S_1 \supset S_0 = S.$$

If  $R_N = S_N \subset Q$ , then

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

For some fixed subcube  $R_k \subsetneq Q$ , there arises in this way one such term  $|u_{R_k} - u_{R_{k-1}}| |R_k|$  from each subcube  $R$  of  $R_k$  such that  $\partial R \cap \partial R_k \neq \emptyset$ . There are at most  $C2^{(n-1)k}$  such subcubes with  $\ell(R) = 2^{-k}\ell(R_k)$ .

If  $R_N = S_N \not\subset Q$ , then we estimate as above, but stop at  $|R_K| = |S_K| = |Q|$ , 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

$$\iint_{\widehat{Q}} |\nabla_x u| dt dx \lesssim \iint_{\widehat{Q}} |\partial_t u| dt dx + |Q| \sum_{Q'} |u_{Q'}|. \quad \square$$

*Proof of Theorem 1.2.* (1) We first localize the problem to a large top cube  $Q_0$ . Choose  $Q_0 \in \mathcal{D}$  large enough that

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

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

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

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

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

$$f_2(t, x) := \begin{cases} \int_{Q_0} g \, dy, & (t, x) \in \widehat{Q}_0, \\ 0, & (t, x) \notin \widehat{Q}_0. \end{cases}$$

Then  $N_{\mathcal{D}}(f_2 - u_2)(x) \leq M_{\mathcal{D}}g(x)$  if  $x \notin Q_0$ , and  $N_{\mathcal{D}}(f_2 - u_2)(x) \leq \inf_{Q_1} M_{\mathcal{D}}g$  if  $x \in Q_0$ , where  $Q_1$  is the sibling of  $Q_0$ . Thus

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

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

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

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

We then define generations of stopping cubes under  $Q_0$  inductively as follows:

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

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

Furthermore, for  $Q \in \omega_*$  we define the ‘‘dyadic sawtooth’’ region

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

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  $Q \in \omega_*$ , i.e.,

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

and  $f = 0$  on  $\mathbb{R}_+^{1+n} \setminus \widehat{Q}_0$ . 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

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

from which the estimate follows.

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

$$\sum_{Q \in \omega_*, Q \subset Q_1} |u_Q - u_{Q_*}| |Q| \leq \frac{1}{\epsilon} \sum_{Q \in \omega_*, Q \subset Q_1} |u_Q - u_{Q_*}|^2 \frac{|Q|}{M_{\mathcal{D}}g(Q)},$$

where we write  $Q_*$  for the stopping parent of  $Q$ , that is, the smallest  $Q_* \in \omega_*$  such that  $Q_* \supsetneq Q$ , and exceptionally  $(Q_0)_* := Q_0$ .

Define the square function

$$Sg(x) := \left( \sum_{Q \in \omega_*, Q \subset Q_1} |\langle g \rangle_Q - \langle g \rangle_{Q_*}|^2 1_Q(x) \right)^{1/2}, \quad \langle g \rangle_Q := \int_Q g(y) dy.$$

Recall that  $u_Q = \langle g \rangle_Q$  for  $Q \in \omega_*$ . Lemma 4.1 gives

$$\begin{aligned} & \sum_{Q \in \omega_*, Q \subset Q_1} |u_Q - u_{Q_*}|^2 \frac{|Q|}{M_{\mathcal{D}}g(Q)} \\ & \lesssim \sum_{Q \in \omega_*, Q \subset Q_1} |\langle g \rangle_Q - \langle g \rangle_{Q_*}|^2 \int_{Q_1} 1_Q(x) \frac{dx}{M_{\mathcal{D}}g(x)} = \int_{Q_1} |Sg(x)|^2 \frac{dx}{M_{\mathcal{D}}g(x)}. \end{aligned}$$

We now use some properties of (dyadic versions of) the Muckenhoupt weight classes  $A_p$ ; these are easy variants of well known results for the usual  $A_p$  classes, found e.g. in [7, Ch. 7]. Write

$$(M_{\mathcal{D}}g)^{-1} = 1 \cdot ((M_{\mathcal{D}}g)^\gamma)^{1-q},$$

for some  $\gamma \in (0, 1)$  and  $q = 1 + 1/\gamma \in (2, \infty)$ . It follows (cf. [7, Thm. 7.7(1) and Prop. 7.2(3)]) that  $(M_{\mathcal{D}}g)^\gamma \in A_1(dx)$  and  $(M_{\mathcal{D}}g)^{-1} \in A_q(dx) \subset A_\infty(dx)$ , with  $A_q$  constants independent of  $g$ .

We now apply Proposition 3.2, with the collection of cubes  $\tilde{\omega}_* := \{Q \in \omega_* : Q \subset Q_1\}$ , the function

$$\tilde{g} := \begin{cases} g(x) - f_{Q_1} g dx, & x \in Q_1, \\ 0, & x \notin Q_1, \end{cases}$$

the weight  $w := (M_{\mathcal{D}}g)^{-1}$ , and  $p = 2$ . This gives

$$\int_{\mathbb{R}^n} |S_{\tilde{\omega}_*} \tilde{g}|^2 dw \lesssim \int_{\mathbb{R}^n} |M_{\mathcal{D}} \tilde{g}|^2 dw = \int_{Q_1} |M_{\mathcal{D}} \tilde{g}|^2 dw \lesssim \int_{Q_1} |M_{\mathcal{D}}g|^2 dw.$$

Thus

$$\begin{aligned} \int_{Q_1} |Sg(x)|^2 dw & \lesssim \int_{Q_1} |S_{\tilde{\omega}_*} g(x)|^2 dw + \int_{Q_1} |M_{\mathcal{D}}g|^2 dw \\ & \leq \int_{\mathbb{R}^n} |S_{\tilde{\omega}_*} \tilde{g}(x)|^2 dw + \int_{Q_1} |M_{\mathcal{D}}g|^2 dw \lesssim \int_{Q_1} |M_{\mathcal{D}}g|^2 dw \leq |Q_1| \inf_{Q_1} M_{\mathcal{D}}(M_{\mathcal{D}}g), \end{aligned}$$

and so

$$\|C(\partial_t f)\|_p \lesssim \epsilon^{-1} \|M_{\mathcal{D}}(M_{\mathcal{D}}g)\|_p \lesssim \epsilon^{-1} \|g\|_p.$$



(4) To complete the proof, we use Lemma 4.2 and obtain the Carleson estimate

$$\|C(\nabla f)\|_p \approx \|C(\partial_t f)\|_p + \|C(\nabla_x f)\|_p \lesssim \|C(\partial_t f)\|_p + \|M_{\mathcal{D}}g\|_p \lesssim \|g\|_p. \quad \square$$

## 5. Application to harmonic measure

Before proving Theorem 1.3, we discuss in this section an important application, to the solvability of the Dirichlet boundary value problem, of such  $L_p$  approximability for solutions to an elliptic equation. As noted in the introduction, in the endpoint case  $p = \infty$ , there is a well known equivalence between

- (1) comparability of non-tangential maximal functions and square functions for solutions,
- (2) approximability of solutions by functions of bounded variation, and
- (3)  $A_\infty$  control of harmonic measure,

for a given real elliptic divergence form equation. It is important to note that this equivalence holds for all equations with real and possibly non-symmetric coefficients, including those which depend on the transversal direction  $t$ . There are known examples by Caffarelli, Fabes and Kenig [2] of symmetric coefficients for which harmonic measure is not  $A_\infty$ , and therefore the approximability and comparability properties may fail as well for  $t$ -dependent coefficients.

In this section, the goal is to demonstrate that in the above equivalences, we may replace (2) by the following local version of the conclusion in Theorem 1.3.

- (2p) For each  $0 < \epsilon < 1$ , there exists  $c'_\epsilon < \infty$  such that for every weak solution  $u : \mathbb{R}_+^{1+n} \rightarrow \mathbb{R}$  to  $\operatorname{div}_{t,x} A(t, x) \nabla_{t,x} u(t, x) = 0$  with  $\|u\|_{L_\infty(\mathbb{R}_+^{1+n})} \leq 1$ , and every cube  $Q \subset \mathbb{R}^n$ , there exists a function  $f_Q$  in  $\mathbb{R}_+^{1+n}$  of locally bounded variation with estimates

$$\|N_\ell(f_Q - u)\|_{L_p(Q)} \leq \epsilon |Q|^{1/p}, \quad \|A_\ell(\nabla f_Q)\|_{L_p(Q)} \leq c_\epsilon |Q|^{1/p}.$$

Here  $N_\ell$  and  $A_\ell$  denote versions of the non-tangential maximal and area functionals from Definition 2.1 using cones  $\{(y, t) : |y - x| \leq \alpha t, t < \ell\}$  truncated at height  $\ell = \ell(Q)$ .

Note that we have used the truncated area functional  $A_\ell$  in the second estimate in (2p), in contrast to the Carleson functional  $C$  that we used in the global version in Theorem 1.3. This, however, is inessential by Proposition 2.4, which also easily extends to the truncated situation by routine modifications. The chosen formulation of property (2p) is motivated by its application to the harmonic measure below, where the area functional leads to the most immediate connection.

Assuming  $A_\infty$  control of the harmonic measure, comparability of non-tangential maximal functions and square functions for solutions follows by [6]. Given such  $N \approx S$  comparability, approximability follows, both in the case  $p = \infty$  as in [13, 11], and for  $1 < p < \infty$ , as shown in Section 6 of this paper. Also the local approximability (2p) follows from  $N \approx S$ , since our estimates are derived from pointwise estimates. In general, without assuming such comparability, we note the following.

**Proposition 5.1.** *Consider a possibly  $t$ -dependent real equation*

$$\operatorname{div}_{t,x} A(t, x) \nabla_{t,x} u(t, x) = 0.$$

*If the approximability property for solutions in Theorem 1.3 holds and if*

$$n/(n - 1) \leq p < \infty,$$

*then the local approximability property (2p) also holds.*

*Proof.* Let  $u$  be a solution with properties as in (2p). Given a cube  $Q \subset \mathbb{R}^n$ , write  $u = u_0 + u_1$ , where  $u_0$  is the solution to the Dirichlet problem with boundary data  $\eta_Q u|_{\mathbb{R}^n}$ , where  $\eta_Q = 1$  on  $5Q$  and supported on  $6Q$ . From the maximum principle and [11, Lem. 4.9], we have the estimate

$$|u_0(t, x)| \lesssim \min\left(1, \left[\frac{\ell(Q)}{|(t, x) - (0, x_Q)|}\right]^{n-1+\nu}\right), \quad (t, x) \in \mathbb{R}_+^{1+n},$$

for some  $\nu > 0$ , where  $x_Q$  denotes the centre of  $Q$ . In particular, we have  $Nu_0(x) \lesssim \min(1, (\ell(Q)/|x - x_Q|)^{n-1+\nu})$  for  $x \in \mathbb{R}^n$ , and thus  $\|Nu_0\|_{L_p(\mathbb{R}^n)} \lesssim |Q|^{1/p}$  if  $p \geq n/(n - 1)$ .

Let  $f_Q := f_0 + u_1$ , where  $f_0$  is the approximant to  $u_0$  given by the assumed global approximability, so that  $f_Q - u = f_0 - u_0$ . We will show that this  $f_Q$  qualifies for (2p). For the first estimate, this is immediate from the assumed properties of the global approximant and the observations just made, namely

$$\|N_\ell(f_Q - u)\|_{L_p(Q)} \leq \|N(f_0 - u_0)\|_{L_p(\mathbb{R}^n)} \leq \epsilon \|Nu_0\|_{L_p(\mathbb{R}^n)} \lesssim \epsilon |Q|^{1/p}.$$

For the second estimate, we separately consider the two terms  $f_0$  and  $u_1$ . First,

$$\|A_\ell(\nabla f_0)\|_{L_p(Q)} \lesssim \|C(\nabla f_0)\|_{L_p(\mathbb{R}^n)} \lesssim c_\epsilon \|Nu_0\|_{L_p(\mathbb{R}^n)} \lesssim c_\epsilon |Q|^{1/p}.$$

Finally, we turn to  $u_1$ . Let  $R \subset \mathbb{R}^n$  be a cube with  $\max(\ell(R), \operatorname{dist}(R, Q)) \leq \ell := \ell(Q)$ . Since  $u_1$  is a bounded solution with vanishing boundary values on  $5Q \supset 2R$ , we may apply the boundary Caccioppoli estimate (see for example [13, (1.3B)]) to deduce that

$$\int_{\widehat{R}} |\nabla u_1| \lesssim \left(\int_{\widehat{R}} |\nabla u_1|^2\right)^{1/2} \lesssim \frac{1}{\ell(R)} \left(\int_{2\widehat{R}} |u_1|^2\right)^{1/2} \lesssim \frac{1}{\ell(R)},$$

which shows that

$$C_\ell(\nabla u_1)(x) = \sup_{\substack{R \ni x \\ \ell(R) \leq \ell}} \frac{1}{|R|} \int_{\widehat{R}} |\nabla u_1| \lesssim 1$$

in a neighbourhood of  $Q$ . By the  $L_p$ -comparability of  $C_\ell$  and  $A_\ell$  (a routine modification of Proposition 2.4), this gives

$$\|A_\ell(\nabla u_1)\|_{L_p(Q)} \lesssim |Q|^{1/p}$$

and completes the verification of (2p). □

We now consider the main result in this section, namely that (2p) implies (3).

**Theorem 5.2.** *Let  $1 < p < \infty$  and consider a real equation  $\operatorname{div}_{t,x} A(t, x) \nabla_{t,x} u(t, x) = 0$ . If the local approximability property (2p) holds, then harmonic measure belongs to  $A_\infty$ .*

*Proof.* Our proof is an adaption of the proof of [13, Thm. 2.3, the case  $p = \infty$ ], and we only point out the changes needed for  $p < \infty$ . Fix a solution  $u$  to  $\operatorname{div}_{t,x} A(t, x) \nabla_{t,x} u(t, x) = 0$  with  $\|u\|_{L^\infty(\mathbb{R}_+^{1+n})} \leq 1$ . Following [13], but not their notation, we consider the counting function

$$K_r(x) := \max\{k : \exists k \text{ points } z_i = (x_i, t_i) \in \Gamma_r(x) \text{ such that } t_i < \theta t_{i-1} \text{ and } |u(z_i) - u(z_{i-1})| \geq \epsilon_0\},$$

where  $\epsilon_0, \theta \in (0, 1)$  are parameters, and

$$\Gamma_r(x) := \{(y, t) : |y - x| \leq \alpha_0 t, t < r\}$$

is a truncated cone based at  $x$ . In the proof of [13, Thm. 2.3], the classical  $\epsilon$ -approximability property is only used through the following consequence established in [13, Lem. 2.9]:

$$\int_Q K_{\ell(Q)}(x) dx \leq c(\epsilon_0, \theta).$$

Thus, it suffices to establish the same conclusion under our approximation property (2p). We will in fact show that

$$\int_Q K_{\ell(Q)}(x)^p dx \leq C(\epsilon_0, \theta), \tag{3}$$

from which the earlier estimate follows by Hölder’s inequality. To prove (3), we fix a cube  $Q$ , pick an approximant  $f_Q$  given by the hypothesis (2p), and note the estimate

$$|Q_b| := |\{y \in Q : N_\ell(f_Q - u) > C_1 \epsilon\}| \leq (C_1 \epsilon)^{-p} \|N_\ell(f_Q - u)\|_{L^p(Q)}^p \leq |Q|/C_1^p.$$

Let  $\mathcal{W}$  denote a Whitney covering of  $Q_b$  by cubes  $R \subset Q$ . For  $x \in Q$ , we note that the pointwise estimate

$$|u - f_Q| \leq C_1 \epsilon$$

holds in  $\Gamma_{\ell(Q)}(x) \setminus \bigcup_{R \in \mathcal{W}} \widehat{C_2 R}$  for some  $C_2 < \infty$ , provided that the  $\alpha$  appearing in (2p) is chosen large enough depending on  $\alpha_0$ . Let  $R_x$  be a largest cube  $R \in \mathcal{W}$  such that  $\widehat{C_2 R}$  intersects  $\Gamma_{\ell(Q)}(x)$ . Then  $x \in C'_2 R_x$  for some  $C'_2 \geq C_2$ , and the pointwise estimate above holds throughout  $\Gamma_{\ell(Q)}(x) \setminus \Gamma_{\ell(C_2 R_x)}(x)$ .

Now, let  $z_i = (t_i, y_i) \in \Gamma_{\ell(Q)}(x)$ , with  $i = 1, \dots, k$ , be points as in the definition of  $K_r(x)$ , and let  $h$  be the largest index such that  $t_h > C_2 \ell(R_x)$ . It follows from interior Hölder regularity that a jump estimate  $|u(w_i) - u(w_{i-1})| \geq \frac{3}{4} \epsilon_0$  persists for all  $w_j \in B(y_j, \eta t_j) \times \{t_j\}$  and a suitably small  $\eta$ . If  $C_1 \epsilon \leq \epsilon_0/4$ , it follows that also  $|f_Q(w_i) - f_Q(w_{i-1})| \geq \frac{1}{4} \epsilon_0$  for  $w_j \in (B(y_j, \eta t_j) \times \{t_j\}) \cap \Gamma_r(x)$  and  $i \leq h$ . Estimating

the difference  $|f_Q(w_i) - f_Q(w_{i-1})|$  by an integral of  $\nabla f_Q$  over the connecting line, and averaging over  $w_j \in (B(y_j, \eta t_j) \times \{t_j\}) \cap \Gamma_r(x)$  for  $j = i, i - 1$ , implies that

$$\frac{1}{4}\epsilon_0 \lesssim \int_{\Gamma_{t_{i-1}}(x) \setminus \Gamma_{t_i}(x)} |\nabla f_Q(t, y)| t^{-n} dt dy,$$

and summing over  $i = 2, \dots, h$  yields

$$h - 1 \lesssim A_{\ell(Q)}(\nabla f_Q)(x).$$

On the other hand, the remaining points  $z_i$ , with  $i = h+1, \dots, k$ , all belong to  $\Gamma_{\ell(C_2 R_x)}(x)$ , so that  $k - h \leq K_{\ell(C_2 R_x)}(x)$ , by definition. So altogether we have

$$k = 1 + (h - 1) + (k - h) \leq 1 + C_3 A_{\ell(Q)}(\nabla f_Q)(x) + K_{\ell(C_2 R_x)}(x).$$

Recalling that  $K_{\ell(Q)}(x)$  is the maximal value of such numbers  $k$ , and that  $C'_2 R_x \ni x$ , we arrive at

$$K_{\ell(Q)} \leq 1 + C_3 A_{\ell(Q)}(\nabla f_Q) + \sup_{R \in \mathcal{W}} 1_{C'_2 R} K_{\ell(C'_2 R)},$$

where we have also estimated  $C_2 \leq C'_2$ .

To prove (3), we set

$$D := \sup_Q \int_Q K_{\ell(Q)}(x)^p dx$$

and integrate

$$\begin{aligned} K_{\ell(Q)}(x)^p &\leq 3^{p-1} \left( 1 + C_3^p A_{\ell(Q)}(\nabla f_Q)^p + \sup_{R \in \mathcal{W}} 1_{C'_2 R}(x) K_{\ell(C'_2 R)}(x)^p \right) \\ &\leq 3^{p-1} \left( 1 + C_3^p A_{\ell(Q)}(\nabla f_Q)^p + \sum_{R \in \mathcal{W}} 1_{C'_2 R}(x) K_{\ell(C'_2 R)}(x)^p \right) \end{aligned}$$

to get

$$\int_Q K_{\ell(Q)}(x)^p dx \leq 3^{p-1} (1 + C_3^p c_\epsilon^p + |Q|^{-1} (C'_2)^n |Q_b| D).$$

Choosing  $C_1$  large (and then  $\epsilon$  small) and taking supremum over  $Q$ , we can hide the second term of the right hand side on the left hand side. (To guarantee the finiteness of  $D$  in the first place, one may initially replace  $K_r(x)$  by  $\min(K_r(x), M)$  and pass to the monotone limit  $M \rightarrow \infty$  in the end.) This proves (3), which concludes the proof.  $\square$

### 6. Approximation of solutions to elliptic equations

In this section, we prove Theorem 1.3, but first fix notation. We only consider dyadic cubes of the  $N$ th generation for some fixed  $N$ , and refer to these simply as “dyadic cubes”, written

$$\mathcal{D}^\delta := \{Q \in \mathcal{D} : \ell(Q) = 2^{-kN}, k \in \mathbb{Z}\},$$

where we write  $\delta := 2^{-N}$  for the change of scale between consecutive generations of these dyadic cubes. The main advantage of this is to gain convenient control of boundary effects: For positive integers  $m$ , we have the dichotomy that each dyadic subcube  $Q' \subsetneq Q$  is either a ‘‘centred cube’’, meaning

$$Q' \subset (1 - 2m\delta)Q,$$

or a ‘‘boundary cube’’, meaning

$$Q' \subset Q \setminus (1 - 2m\delta)Q.$$

As long as  $2m\delta < 1$ , both these consist of a positive fraction of the total volume of  $Q$ , but this could never be achieved with  $\delta = 2^{-1}$  when every dyadic child is a boundary cube.

The  $\mathcal{D}^\delta$ -dyadic versions of the functionals  $M$ ,  $N$  and  $C$  are denoted by  $M_{\mathcal{D}^\delta}$ ,  $N_{\mathcal{D}^\delta}$  and  $C_{\mathcal{D}^\delta}$ . One verifies that the estimates analogous to Proposition 2.3 hold. We also use the following notation for  $Q \in \mathcal{D}^\delta$ , where the parameter  $\eta > 0$  will be eventually chosen small relative to the given  $\epsilon$  appearing in the statement of Theorem 1.3. (The parameter  $\delta = 2^{-N}$  will also be chosen small, but independent of  $\epsilon$ .)

- $\widehat{Q} := (0, \ell(Q)) \times Q$  is the Carleson box, as before.
- $p_Q := ((1 - \eta)\ell(Q), c_Q)$  is the ‘‘corkscrew point’’ of  $\widehat{Q}$ , where  $c_Q$  denotes the centre of  $Q$ .
- $\widetilde{Q} := \{\ell(Q)\} \times \eta Q$  is a small hypersurface on the top boundary of  $\widehat{Q}$ , around the centre.
- $W_Q := \widehat{Q} \setminus \bigcup_{\mathcal{D}^\delta \ni Q' \subsetneq Q} \widehat{Q}' = [\delta\ell(Q), \ell(Q)) \times Q$  is a Whitney-type rectangle.
- $\Gamma_Q := \{(t, x) : t > \delta\ell(Q) + \text{dist}(x, Q)\}$  is an epigraph domain containing  $W_Q$ .

We now outline the proof of Theorem 1.3. We construct the approximant  $f$  as follows. We define a family of ‘‘stopping cubes’’  $\mathcal{S}$  in (8), and the corresponding sawtooth regions  $\Omega_{\mathcal{S}}(S) := \bigcup_{Q: \pi_{\mathcal{S}} Q = S} W_Q$ , where  $\pi_{\mathcal{S}} Q = S$  means that  $S$  is the smallest stopping cube such that  $S \supset Q$ ; the family  $\mathcal{S}$  is built in such a way that the value of  $u(p_Q)$  varies relatively little among all  $Q$  with  $\pi_{\mathcal{S}} Q = S$ . The first approximation to  $u$  is then given by  $\varphi_1 := \sum_{S \in \mathcal{S}} u(p_S) \cdot 1_{\Omega_{\mathcal{S}}(S)}$ .

However, this approximation fails to be good on the Whitney regions  $W_R$ , where the oscillation of  $u$  is relatively large, more precisely, when it happens that

$$\text{osc}_{W_R} u := \sup_{z, w \in W_R} |u(z) - u(w)| > \epsilon M_{\mathcal{D}^\delta}(Nu)(R) := \epsilon \sup_{Q \supset R} \int_Q Nu(x) dx. \quad (4)$$

Note that here the defining condition is simpler than the stopping conditions considered above, in that it can be directly checked for any cube, without reference to the previously chosen members of the stopping family. We label by  $\mathcal{R}$  the family of cubes  $R$  in (4), and introduce the additional correction  $\varphi_2 := \sum_{R \in \mathcal{R}} (u - \varphi_1) \cdot 1_{W_R}$ . The final approximation is then given by  $f := \varphi_1 + \varphi_2$ . The verification of  $N_{\mathcal{D}^\delta}(f - u) \lesssim \epsilon Nu$  will then be straightforward from the construction of the collections  $\mathcal{S}$  and  $\mathcal{R}$ . The pointwise estimate for the Carleson functional,  $C_{\mathcal{D}^\delta}(\nabla f) \lesssim_\epsilon M_{\mathcal{D}^\delta}(Nu)$ , is verified separately for  $\varphi_1$  and  $\varphi_2$  in place of  $f$ ; these bounds depend in particular on the Carleson property of both  $\mathcal{S}$  and  $\mathcal{R}$ , established in Lemmas 6.4 and 6.5, and the estimates are completed in Lemmas 6.6 and 6.8.

Lemmas 6.4 and 6.5 ultimately build on two estimates which we borrow from [11]:

$$\int_{\theta Q} |u(\psi(x), x) - u(p_Q)|^2 dx \lesssim \iint_{(t,x) \in \widehat{Q}, t > \psi(x)} |\nabla u(t, x)|^2 (t - \psi(x)) dt dx, \quad (5)$$

$$\frac{1}{|Q|} \iint_{(t,x) \in \widehat{Q}, t > \psi(x)} |\nabla u(t, x)|^2 (t - \psi(x)) dt dx \lesssim \sup_{\substack{(t,x) \in \mathbb{R}_+^{1+n} \\ t > \psi(x)}} |u(t, x)|^2, \quad (6)$$

for weak solutions  $u$  to an elliptic equation  $Lu = 0$  as in Theorem 1.3. Here  $\theta \in (0, 1)$ ,  $p_Q$  is a ‘‘corkscrew point’’ above  $Q$  in the Carleson box  $\widehat{Q}$ , and  $\psi \geq 0$  is a Lipschitz function. Note that the implicit constants in the two estimates depend on the ellipticity constants  $\lambda_A$  and  $\|A\|_\infty$  from Theorem 1.3 and on  $\|\nabla \psi\|_\infty$  and dimension, but not otherwise on  $A$ ,  $\psi$ ,  $u$  or  $Q$ . The first estimate (5) follows from [11, Cor. 1.17] upon replacing  $u$  by  $u - u(p_Q)$  and using interior regularity and Poincaré’s inequality to remove the error term. The second estimate (6) follows from [11, Cor. 1.10] upon pulling back that result from the half-space to the epigraph domain  $t > \psi(x)$ .

The construction in this section builds on that in the case  $p = \infty$  from [9, 5, 13, 11], but with non-trivial modifications. The construction (4) of the family  $\mathcal{R}$  of Whitney regions with large oscillation of  $u$  goes back to [9], as does the stopping construction (8). The main novelty here is that for  $L_p$ ,  $p < \infty$ , we require a variable threshold in these constructions, expressed in terms of the maximal function of  $u$ . This requires a second parallel stopping construction (7), which has the effect of freezing this threshold. Such multiple stopping time constructions have appeared earlier in [1, 14, 15].

Finally, to pass from dyadic sawtooths to Lipschitz sawtooths to be able to use the above estimates (5) and (6), we follow the construction of [11] in Lemma 6.1. We now turn to the details.

### 6.1. Construction of stopping cubes

We start with some generalities. Let  $\mathcal{C}(Q', Q) \in \{\text{true}, \text{false}\}$  be some ‘‘criterion’’ that assigns a truth value to every pair of (dyadic) cubes  $Q' \subset Q$ . We specifically agree that  $\mathcal{C}(Q, Q) = \text{false}$  for every cube  $Q$ . By the ‘‘stopping family’’ with initial collection  $\mathcal{I} \subset \mathcal{D}^\delta$  and stopping criterion  $\mathcal{C}$  we understand the family  $\mathcal{F} = \mathcal{F}(\mathcal{I}, \mathcal{C})$  of dyadic cubes constructed as follows: We initialize  $\mathcal{F} := \mathcal{I}$ . Then we add to  $\mathcal{F}$  all  $F' \in \mathcal{D}^\delta \setminus \mathcal{F}$  such that

- (a)  $\mathcal{C}(F', F)$  is true for some  $F \in \mathcal{F}$  with  $F' \subsetneq F$ , and
- (b)  $F'$  is not contained in any  $F'' \subsetneq F$  with either  $F'' \in \mathcal{F}$  or  $\mathcal{C}(F'', F)$  true.

We repeat this addition indefinitely. This is seen to yield a well defined family  $\mathcal{F} \subset \mathcal{D}^\delta$ .

For every  $Q \in \mathcal{D}^\delta$ , let  $\pi_{\mathcal{F}} Q$  denote the minimal  $F \in \mathcal{F}$  such that  $Q \subset F$ , where the possibility that  $F = Q$  is not excluded. The stopping family with initial collection  $\mathcal{I}$  and stopping criterion  $\mathcal{C}$  has the property that  $\mathcal{C}(Q, F)$  is false whenever  $\pi_{\mathcal{F}} Q = F$ ; namely, the latter means by definition that there *does not exist* any intermediate stopping cube  $F'$  with  $Q \subset F' \subsetneq F$ , thus intermediate cubes  $Q'$  with  $Q \subset Q' \subsetneq F$  do *not* satisfy the stopping condition  $\mathcal{C}(Q', F)$ , and in particular, taking  $Q' = Q$ , the condition  $\mathcal{C}(Q, F)$  is false. For  $F \in \mathcal{F}$ , we denote by  $\text{ch}_{\mathcal{F}}(F) := \{F' \in \mathcal{F} \text{ maximal} : F' \subsetneq F\}$  the family of its  $\mathcal{F}$ -children.

A stopping criterion  $\mathcal{C}$  is *sparse* if

$$\sum_{\substack{Q' \subset Q \text{ maximal:} \\ \mathcal{C}(Q', Q) \text{ is true}}} |Q'| \leq \tau |Q|$$

for some fixed  $\tau < 1$ . It is straightforward to check that if the initial collection  $\mathcal{I}$  is *Carleson*, that is,

$$\sup_{Q \in \mathcal{I}} \frac{1}{|Q|} \sum_{\mathcal{I} \ni R \subset Q} |R| < \infty,$$

and if the stopping criterion  $\mathcal{C}$  is sparse, then the stopping family  $\mathcal{F} = \mathcal{F}(\mathcal{I}, \mathcal{C})$  is *Carleson*.

A priori, the stopping collections produced by a dyadic algorithm may not be so well behaved geometrically. This is to some extent remedied by the following lemma, which builds on ideas from [11, Sec. 5].

**Lemma 6.1.** *Consider a cube  $Q$  and a disjoint collection  $\mathcal{Q}$  of its dyadic subcubes. We say that  $Q' \in \mathcal{Q}$  is*

- *centred (in  $Q$ ) if  $Q' \subset (1 - 2\delta)Q$ ;*
- *uncovered (by  $Q$ ) if  $(\ell(Q'') - \ell(Q'))/\text{dist}(Q', Q'') \leq \delta^{-1}$  for all  $Q'' \in \mathcal{Q}$ .*

(See Figure 1 below for an illustration of covered and uncovered cubes.) Let

$$\mathcal{Q}^* := \{Q' \in \mathcal{Q} : Q' \text{ is centred and uncovered}\}.$$

For some constants  $\tau \in (0, 1)$  and  $C$ , we then have

$$\sum_{Q' \in \mathcal{Q}} |Q'| \leq \tau |Q| + C \sum_{Q'' \in \mathcal{Q}^*} |Q''|.$$

*Proof.* If  $Q'$  is not uncovered, then  $\ell(Q'') - \delta^{-1} \text{dist}(Q', Q'') > \ell(Q')$  for some  $Q'' \in \mathcal{Q}$ , and we say that this  $Q''$  covers  $Q'$ . Then in particular  $\ell(Q') < \ell(Q'')$ , and hence  $\ell(Q') \leq \delta \ell(Q'')$ , but also  $\text{dist}(Q', Q'') < \delta \ell(Q'')$ .

Further, if  $Q'$  is not uncovered, it is covered by some  $Q_1$  which, if not uncovered, is covered by some  $Q_2$ , and so on. Since  $\ell(Q_k)$  increases geometrically and is bounded by  $\ell(Q)$ , the chain must terminate after finitely many steps with some uncovered  $Q_k$ . The  $\ell^\infty$ -distance of the furthest point of  $Q' =: Q_0$  from the centre of  $Q_k$  can be at most

$$\begin{aligned} \sum_{j=0}^{k-1} [\ell(Q_j) + \text{dist}(Q_j, Q_{j+1})] + \frac{1}{2} \ell(Q_k) &\leq \sum_{j=0}^{k-1} 2\delta^{k-j} \ell(Q_k) + \frac{1}{2} \ell(Q_k) \\ &< \left( \frac{2\delta}{1-\delta} + \frac{1}{2} \right) \ell(Q_k) \leq \frac{5}{2} \ell(Q_k), \end{aligned}$$

since  $\delta \leq 1/2$ , and hence  $Q' \subset 5Q_k$ . We have

$$\sum_{Q' \in \mathcal{Q}} |Q'| = \sum_{\substack{Q' \in \mathcal{Q} \\ Q' \subset (1-2m\delta)Q}} |Q'| + \sum_{\substack{Q' \in \mathcal{Q} \\ Q' \subset Q \setminus (1-2m\delta)Q}} |Q'| =: I + II,$$

where  $I \leq (1 - (1 - 2m\delta)^n)|Q|$  by disjointness. On the other hand, every  $Q'$  appearing in  $I$  is contained in  $5Q''$  for some uncovered  $Q'' \in \mathcal{Q}$ . In particular,  $5Q''$  intersects  $Q' \subset (1 - 2m\delta)Q$ . Thus, the  $\ell^\infty$ -distance of the furthest point of  $5Q''$  from the centre of  $Q$  is at most

$$(1 - 2m\delta)\ell(Q)/2 + 5\ell(Q'') \leq (5\delta + 1/2 - m\delta)\ell(Q),$$

and hence  $Q'' \subset 5Q'' \subset (1 - 2(m - 5)\delta)Q$ . Since all  $Q'$  in  $I$  are covered by such  $5Q''$ , we have

$$I \leq \sum_{\substack{Q'' \text{ uncovered} \\ Q'' \subset (1-2(m-5)\delta)Q}} |5Q''| \leq 5^n \sum_{Q'' \in \mathcal{Q}^*} |Q''|$$

provided that we take  $m \geq 6$ , and also  $\delta < 1/(2m)$  for term  $I$ . □

Let us now fix as our initial collection  $\mathcal{I}$  some increasing chain of cubes  $I_0 \subsetneq I_1 \subsetneq \dots$  that exhaust  $\mathbb{R}^n$ . Clearly this is Carleson.

We define the ‘‘principal cubes’’  $\mathcal{P}$  as the stopping family with initial collection  $\mathcal{I}$  and the stopping criterion  $\mathcal{C}(Q', Q)$  given by

$$M_{\mathcal{D}^\delta}(Nu)(Q') = \sup_{R \supseteq Q'} \int_R Nu(x) dx > A \cdot M_{\mathcal{D}^\delta}(Nu)(Q), \tag{7}$$

for some fixed  $A > 1$  to be chosen. To verify that this criterion is sparse, and therefore  $\mathcal{P}$  is Carleson, select a disjoint family of subcubes  $Q' \subset Q$  that satisfy (7). Then

$$\begin{aligned} \sum |Q'| &\leq |\{M_{\mathcal{D}^\delta}(1_Q Nu) > A \cdot M_{\mathcal{D}^\delta}(Nu)(Q)\}| \\ &\leq \frac{1}{A \cdot M_{\mathcal{D}^\delta}(Nu)(Q)} \int_Q Nu(x) dx = \frac{|Q|}{A} \end{aligned}$$

by the weak-type (1, 1) estimate for the maximal operator  $M_{\mathcal{D}^\delta}$ .

The usefulness of the numbers  $M_{\mathcal{D}^\delta}(Nu)(Q)$  lies in the fact that they control the values of  $u$  in the entire graph-domain  $\Gamma_Q$ .

**Lemma 6.2.** *We have the estimate*

$$\sup_{\Gamma_Q} |u| \leq M_{\mathcal{D}^\delta}(Nu)(Q).$$

*Proof.* It is enough to observe that  $\Gamma_Q \subset \Gamma_{1/\delta}(x) := \{(t, y) : |y - x| < \delta^{-1}t\}$  for any  $x \in Q$ , and therefore

$$\sup_{\Gamma_Q} |u| \leq \inf_{x \in Q} \sup_{\Gamma_{1/\delta}(x)} |u| \leq \inf_{x \in Q} Nu(x) \leq \int_Q Nu dx \leq M_{\mathcal{D}^\delta}(Nu)(Q)$$

provided that the aperture defining  $N$  is at least  $\delta^{-1}$ . □

Finally, we define the ‘‘stopping cubes’’  $\mathcal{S}$  as the stopping family with initial collection  $\mathcal{P}$  and the stopping criterion  $\mathcal{C}(Q', Q)$  given by

$$|u(p_{Q'}) - u(p_Q)| > \epsilon M_{\mathcal{D}^\delta}(Nu)(Q'). \tag{8}$$

We observe the following self-improvement of this criterion.



**Lemma 6.3.** *Under the condition (8), we have*

$$|u(z) - u(p_Q)| \gtrsim \epsilon M_{\mathcal{D}^\delta}(Nu)(Q') \quad \text{for all } z \in \tilde{Q}' := \{\ell(Q')\} \times \eta Q'$$

provided that  $\eta$  satisfies  $\eta^\alpha \ll \epsilon$ , where  $\alpha > 0$  is the Hölder exponent from interior regularity estimates for  $u$ .

*Proof.* If  $z \in \tilde{Q}'$ , the interior regularity of solutions  $u$  to  $\operatorname{div} A \nabla u = 0$  shows that

$$|u(z) - u(p_{Q'})| \lesssim \left( \frac{|z - p_{Q'}|}{\ell(Q')} \right)^\alpha \langle |u|^2 \rangle_{\tilde{W}_{Q'}}^{1/2} \lesssim \eta^\alpha \inf_{Q'} Nu \ll \epsilon M_{\mathcal{D}^\delta}(Nu)(Q'),$$

where  $\tilde{W}_{Q'}$  is a slight expansion of  $W_{Q'}$ . Thus, we have

$$|u(z) - u(p_Q)| \gtrsim \epsilon M_{\mathcal{D}^\delta}(Nu)(Q') \quad \text{for all } z \in \tilde{Q}'. \quad \square$$

The main estimate here is that both the stopping family  $\mathcal{S}$ , and the collection  $\mathcal{R}$  of large oscillation cubes introduced in (4), satisfy the Carleson condition.

**Lemma 6.4.** *For the stopping cubes  $\mathcal{S}$ , we have the Carleson measure estimate*

$$\sum_{S \in \mathcal{S}, S \subset Q_0} |S| \lesssim |Q_0|$$

for all dyadic cubes  $Q_0$ .

**Lemma 6.5.** *For the large oscillation cubes  $\mathcal{R}$ , we have the Carleson measure estimate*

$$\sum_{R \in \mathcal{R}, R \subset Q_0} |R| \lesssim |Q_0|$$

for all dyadic cubes  $Q_0$ .

*Proof of Lemma 6.4.* (A) First we make a preliminary simplification of the estimate based on Lemma 6.1. By considering the maximal  $\mathcal{S}$ -cubes contained in  $Q_0$ , we may assume without loss of generality that  $Q_0 \in \mathcal{S}$ . We then write

$$\sum_{S \in \mathcal{S}, S \subset Q_0} |S| = \sum_{\substack{S \in \mathcal{S}, S \subset Q_0 \\ \pi_{\mathcal{P}} S = \pi_{\mathcal{P}} Q_0}} |S| + \sum_{P \in \mathcal{P}, P \subsetneq Q_0} \sum_{\substack{S \in \mathcal{S} \\ \pi_{\mathcal{P}} S = P}} |S|,$$

and we claim that it suffices to prove the required bound  $|Q_0|$  for the first term. Namely, if this is done, we simply apply this result, with  $Q_0 = \pi_{\mathcal{P}} Q_0 = P$ , to the inner sum in the second term, which shows that this inner sum is bounded by  $|P|$ . Then the Carleson property of the collection  $\mathcal{P}$  completes the estimate.

So we concentrate on the first term, and abbreviate  $\pi_{\mathcal{P}} Q_0 =: P$  for convenience. We also drop the summation condition “ $S \in \mathcal{S}$ ”, with the implicit understanding that this is always in force.

With Lemma 6.1 applied to  $Q = S$  and  $\mathcal{Q} = \operatorname{ch}_{\mathcal{S}}(S)$  for each relevant  $S$ , by indexing the cubes by their parents we obtain (note that each  $S$ , except for the maximal

ones, is a child of another  $S$ , and the sum over the maximal ones is bounded by  $|Q_0|$ , by disjointness)

$$\begin{aligned} \sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} |S| &\leq |Q_0| + \sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} \sum_{S' \in \text{ch}_S(S)} |S'| \\ &\leq |Q_0| + \sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} \left( \tau |S| + C \sum_{S' \in \text{ch}_S^*(S)} |S'| \right), \end{aligned}$$

where

$$\text{ch}_S^*(S) := \{S' \in \text{ch}_S(S) : S' \text{ centred in } S \text{ and uncovered by } \text{ch}_S(S)\}.$$

The second term can then be absorbed into the left side, since  $\tau < 1$ .

We further observe the following. If  $S' \in \mathcal{P}$  for some  $S'$  appearing in the inner sum on the right, this together with  $S' \in \text{ch}_S(S)$  and  $\pi_{\mathcal{P}} S = P$  implies that  $S' \in \text{ch}_{\mathcal{P}} P$ . But these cubes are pairwise disjoint. Since  $S' \subset S \subset Q_0$ , they are also contained in  $Q_0$ , hence their total volume adds up to at most  $|Q_0|$ , which may be absorbed into the first term on the right. So altogether we find that

$$\sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} |S| \lesssim |Q_0| + \sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} \sum_{S' \in \text{ch}_S^*(S) \setminus \mathcal{P}} |S'|,$$

and it remains to bound the last double sum by  $|Q_0|$ .

(B) We now aim to use the local  $N \lesssim S$  estimate (5). We first treat one of the inner sums over  $S' \in \text{ch}_S^*(S) \setminus \mathcal{P}$  for a fixed  $S$ . The significance of the restriction  $S' \notin \mathcal{P}$  comes from the fact that we then know that  $S'$  was chosen as a stopping cube by the criterion (8). By Lemma 6.3, this gives

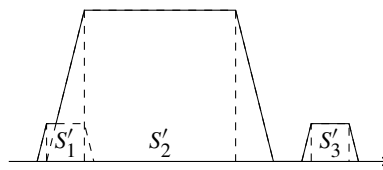
$$M_{\mathcal{D}^\delta}(Nu)(P)^2 |S'| \leq M_{\mathcal{D}^\delta}(Nu)(S')^2 |S'| \lesssim M_{\mathcal{D}^\delta}(Nu)(S')^2 |\tilde{S}'| \lesssim \int_{\tilde{S}'} |u - u(p_S)|^2 dx,$$

where we allow the dependence on  $\epsilon$  in the implicit constants.

We then consider the Lipschitz function

$$\psi_S^1(x) := \sup_{S' \in \text{ch}_S(S)} \psi_{S'}(x), \quad \psi_{S'}(x) := \max(\ell(S') - \delta^{-1} \text{dist}(x, S'), 0).$$

This is closely related to the notion of coveredness, and illustrated in Figure 1.



**Fig. 1.** A possible configuration of cubes  $S'_i \in \text{ch}_S(S)$ :  $S'_2$  and  $S'_3$  are uncovered, but  $S'_1$  is covered by  $S'_2$ . Dashed lines show the Carleson boxes  $\widehat{S}'_i$  and the graphs of  $\psi_{S'_i}$ , except where overwritten by the continuous line, which shows the graph of  $\psi_S^1$ .

The function  $\psi_S^1$  has two important features:

- $\widehat{S}' \cap \Omega_{\psi_S^1} = \emptyset$ , where  $\Omega_{\psi_S^1} := \{(t, x) \in \mathbb{R}_+^{1+n} : \psi_S^1(x) < t\}$ , for every  $S' \in \text{ch}_S(S)$ .
- $\psi_S^1(x) = \ell(S')$  for every  $x \in S'$  if  $S' \in \text{ch}_S(S)$  is uncovered.

This allows us to write, using Lemma 6.3 in the first step,

$$\begin{aligned} M_{\mathcal{D}^\delta}(Nu)(P)^2 \sum_{S' \in \text{ch}_S^*(S) \setminus \mathcal{P}} |S'| &\lesssim \sum_{S' \in \text{ch}_S^*(S) \setminus \mathcal{P}} \int_{\widehat{S}'} |u - u(p_S)|^2 dx \\ &\leq \int_{(1-\delta)S} |u(\psi_S^1(x), x) - u(p_S)|^2 dx \lesssim \iint_{\Omega_{\psi_S^1} \cap \widehat{S}} |\nabla u|^2 (t - \psi_S^1(x)) dt dx, \end{aligned} \tag{9}$$

where in the last step we have used the local  $N \lesssim S$  estimate (5).

(C) We now aim to use the local  $S \lesssim N$  estimate (6). In order to sum over all relevant  $S$ , set

$$\psi_P^2(x) := \inf_{Q: \pi_{\mathcal{P}} Q = P} [\delta \ell(Q) + \text{dist}(x, Q)].$$

Then  $\psi_P^2$  is a Lipschitz function, and

$$\Omega_{\psi_P^2} := \{(t, x) : \psi_P^2(x) < t\} = \bigcup_{Q: \pi_{\mathcal{P}} Q = P} \Gamma_Q \supset \bigcup_{Q: \pi_{\mathcal{P}} Q = P} W_Q,$$

and hence  $|u| \lesssim M_{\mathcal{D}^\delta}(Nu)(P)$  on this set by Lemma 6.2 and the stopping condition (7).

Returning to (9), we observe that

$$\Omega_{\psi_S^1} \cap \widehat{S} \subset \widehat{S} \setminus \bigcup_{S' \in \text{ch}_S(S)} \widehat{S}' = \bigcup_{Q: \pi_S Q = S} W_Q \subset \widehat{S} \cap \Omega_{\psi_P^2}$$

for all  $S$  such that  $\pi_{\mathcal{P}} S = P$ , where the first inclusion shows that these sets are pairwise disjoint in  $S$ . Since both ends of the inclusion involve graph-domains, we also see that  $\psi_S^1(x) \geq \psi_P^2(x)$ . This allows us to estimate and sum over  $S$  in (9) as follows:

$$\begin{aligned} \sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} \iint_{\Omega_{\psi_S^1} \cap \widehat{S}} |\nabla u|^2 (t - \psi_S^1(x)) dt dx &\leq \iint_{\Omega_{\psi_P^2} \cap \widehat{Q}_0} |\nabla u|^2 (t - \psi_P^2(x)) dt dx \\ &\lesssim |Q_0| \cdot \|u\|_{L^\infty(\Omega_{\psi_P^2})}^2 \lesssim |Q_0| \cdot M_{\mathcal{D}^\delta}(Nu)(P)^2, \end{aligned} \tag{10}$$

where the penultimate step is by the local  $S \lesssim N$  estimate (6).

A combination of (9) and (10) shows that the factor  $M_{\mathcal{D}^\delta}(Nu)(P)^2$  cancels from both sides, and we are left with

$$\sum_{S \subset Q_0, \pi_{\mathcal{P}} S = P} \sum_{S' \in \text{ch}_S^*(S) \setminus \mathcal{P}} |S'| \lesssim |Q_0|,$$

which completes the proof of the lemma. □

*Proof of Lemma 6.5.* By arguing as at the beginning of the proof of Lemma 6.4, we find that it is enough to prove that

$$\sum_{\substack{R \in \mathcal{R}, R \subset Q_0 \\ \pi_{\mathcal{P}} R = P}} |R| \lesssim |Q_0|,$$

where  $P := \pi_{\mathcal{P}} Q_0$ .

By interior regularity of solutions to  $u$  to  $\operatorname{div} A \nabla u = 0$ , for all  $z, w \in W_R$ ,

$$\begin{aligned} |u(z) - u(w)| &\lesssim \left( \frac{|z - w|}{\ell(R)} \right)^\alpha \ell(R) \left( \frac{1}{|W_R|} \iint_{\tilde{W}_R} |\nabla u|^2 dt dx \right)^{1/2} \\ &\lesssim \left( \frac{1}{|R|} \iint_{\tilde{W}_R} |\nabla u|^2 t dt dx \right)^{1/2}, \end{aligned}$$

where  $\tilde{W}_R$  is a slight expansion of the Whitney rectangle  $W_R$ . Hence for  $R \in \mathcal{R}$ , we have

$$M_{\mathcal{D}^\delta}(Nu)(R)^2 |R| \lesssim \left( \operatorname{osc}_{W_R} u \right)^2 |R| \lesssim \iint_{\tilde{W}_R} |\nabla u|^2 t dt dx. \tag{11}$$

Let  $W_R^* = [\delta' \ell(R), \kappa' \ell(R)] \times R^*$  be a slightly bigger expansion and  $R^*$  its projection onto  $\mathbb{R}^n$ , where  $\delta' \in (0, \delta)$ . Let

$$\psi_R^*(x) := \delta' \ell(R^*) + \operatorname{dist}(x, R^*), \quad \Gamma_R^* := \{(t, x) : t > \psi_R^*(x)\}.$$

Then  $\tilde{W}_R \subset \Gamma_R^*$  and

$$t \lesssim t - \psi_R^*(x) \quad \text{for all } (t, x) \in \tilde{W}_R. \tag{12}$$

Let further

$$\psi_P^{**}(x) := \inf_{Q: \pi_{\mathcal{P}} Q = P} \psi_R^*(x),$$

so that

$$\bigcup_{Q: \pi_{\mathcal{P}} Q = P} W_Q \subset \bigcup_{Q: \pi_{\mathcal{P}} Q = P} \Gamma_Q^* = \{(t, x) : t > \psi_P^{**}(x)\} =: \Omega_{\psi_P^{**}}.$$

It follows that

$$\begin{aligned} M_{\mathcal{D}^\delta}(Nu)(P)^2 \sum_{\substack{R \in \mathcal{R}, R \subset Q_0 \\ \pi_{\mathcal{P}} R = P}} |R| &\leq \sum_{\substack{R \in \mathcal{R}, R \subset Q_0 \\ \pi_{\mathcal{P}} R = P}} M_{\mathcal{D}^\delta}(Nu)(R)^2 |R| \\ &\lesssim \sum_{\substack{R \in \mathcal{R}, R \subset Q_0 \\ \pi_{\mathcal{P}} R = P}} \iint_{\tilde{W}_R} |\nabla u|^2 (t - \psi_P^{**}) dt dx \quad \text{by (11) and (12)} \\ &\stackrel{(*)}{\lesssim} \iint_{\Omega_{\psi_P^{**}} \cap \widehat{Q}_0} |\nabla u|^2 (t - \psi_P^{**}) dt dx \\ &\lesssim |Q_0| \cdot \|u\|_{L^\infty(\Omega_{\psi_P^{**}})}^2 \quad \text{by the local } S \lesssim N \text{ bound (6)} \\ &\stackrel{(**)}{\lesssim} |Q_0| \cdot M_{\mathcal{D}^\delta}(Nu)(P)^2. \end{aligned} \tag{13}$$

In (\*), we have used the bounded overlap of the regions  $\widetilde{W}_R$ , which is an easy consequence of the geometry of the Whitney regions, and their containment in  $\widehat{Q}_0$ , a slight expansion of the Carleson box  $\widehat{Q}_0$ . In the last step (\*\*), we have used the fact that  $\Gamma_R^* \subset \Gamma_\gamma(x) := \{(t, y) : |y - x| < \gamma t\}$  for all  $x \in R$  provided that  $\gamma$  is large enough, and therefore

$$\sup_{\Gamma_R^*} |u| \leq \inf_{x \in R} Nu(x) \leq M_{\mathcal{D}^\delta}(Nu)(R) \lesssim M_{\mathcal{D}^\delta}(Nu)(P)$$

whenever  $\pi_{\mathcal{P}} R = P$ , provided that the aperture defining  $Nu$  is large enough.

Observing that  $M_{\mathcal{D}^\delta}(Nu)(P)^2$  cancels from both sides of (13), we have established the required bound.  $\square$

### 6.2. The $\epsilon$ -approximating functions

As a first approximation, consider the piecewise constant function

$$\varphi_1 := u(p_S) \quad \text{on } \Omega_S(S) := \bigcup_{Q: \pi_S Q = S} W_Q.$$

**Lemma 6.6.** *We have the estimate*

$$\iint |\nabla(1_{\widehat{Q}_0} \varphi_1)| \, dt \, dx \lesssim \int_{Q_0} Nu \, dx$$

for all dyadic cubes  $Q_0$ .

*Proof.* Let us abbreviate  $S_0 := \pi_S Q_0$ . Then we have

$$\begin{aligned} 1_{\widehat{Q}_0} \varphi_1 &= \sum_{S \in \mathcal{S}, S \subset Q_0} u(p_S) 1_{\Omega_S(S)} + u(p_{S_0}) 1_{\Omega_{S_0}(S_0) \cap \widehat{Q}_0}, \\ |\nabla(1_{\widehat{Q}_0} \varphi_1)| &\leq \sum_{S \in \mathcal{S}, S \subset Q_0} |u(p_S)| \cdot |\nabla 1_{\Omega_S(S)}| + |u(p_{S_0})| \cdot |\nabla 1_{\Omega_{S_0}(S_0) \cap \widehat{Q}_0}| \\ &= \sum_{S \in \mathcal{S}, S \subset Q_0} |u(p_S)| \cdot H^n \llbracket \partial \Omega_S(S) \rrbracket + |u(p_{S_0})| \cdot H^n \llbracket \Omega_{S_0}(S_0) \cap \widehat{Q}_0 \rrbracket, \end{aligned} \tag{14}$$

where  $H^n$  is the  $n$ -dimensional Hausdorff measure, or more simply, the  $n$ -dimensional Lebesgue measure on the hyperplanes to which it is restricted.

We can then compute

$$\begin{aligned} \iint |\nabla(1_{\widehat{Q}_0} \varphi_1)| \, dt \, dx &\lesssim \sum_{S \in \mathcal{S}, S \subset Q_0} |u(p_S)| \cdot |S| + |u(p_{S_0})| \cdot |Q_0| \\ &\leq \sum_{S \in \mathcal{S}, S \subset Q_0} \inf_S Nu \cdot |S| + \inf_{Q_0} Nu \cdot |Q_0| \lesssim \int_{Q_0} Nu \, dx, \end{aligned}$$

where the first estimate is based on simple geometric observations concerning the shape of the sets  $\Omega_S(S)$  and  $\Omega_{S_0}(S_0) \cap \widehat{Q}_0$ , and the last one on the Carleson inequality and the Carleson property of the collection  $\mathcal{S}$  for the first term from Lemma 6.4, and a trivial estimate for the second.  $\square$

**Remark 6.7.** It is perhaps interesting to remark that, in bounding the gradient as in (14), we make an apparently crude estimate of the jumps of  $\varphi_1$  in the interior of  $\widehat{Q}_0$ , in that we dominate a jump  $|u(p_{S'}) - u(p_S)|$  simply by  $|u(p_{S'})| + |u(p_S)|$ . However, a comparison with the stopping criterion (8) shows that this is not so crude after all: it is easy to check that  $|u(p_{S'})| + |u(p_S)| \lesssim M_{\mathcal{D}^\delta}(Nu)(S')$  for all  $S' \subset S$ , and the very stopping criterion (8) says that, for consecutive stopping cubes  $S' \subsetneq S$ , the difference is already essentially as big as this maximal quantity. This means that if we ignore the dependence on  $\epsilon$  as we do, there is no essential loss in making this apparently crude estimate. Note, however, that we argued somewhat differently in the context of dyadic martingales, where we did trace a good dependence on  $\epsilon$ .

The function  $\varphi_1$  provides a good  $\epsilon$ -approximation of  $u$  in all those  $W_Q$  where

$$\operatorname{osc}_{W_Q} u < \epsilon M_{\mathcal{D}^\delta}(Nu)(Q),$$

that is, whenever  $Q \notin \mathcal{R}$ . Likewise, it is clear that  $\varphi_1$  fails to be a good approximation in any  $W_R$  with  $R \in \mathcal{R}$ . Our final  $\epsilon$ -approximation will be  $\varphi_1 + \varphi_2$ , where

$$\varphi_2|_{W_Q} := \begin{cases} (u - \varphi_1)|_{W_Q} = u|_{W_Q} - u(p_{\pi_S Q}) & \text{if } Q \in \mathcal{R}, \\ 0 & \text{else.} \end{cases}$$

It remains to show that  $\varphi_2$  satisfies the needed Carleson measure estimate.

**Lemma 6.8.** *We have the estimate*

$$\iint |\nabla(1_{\widehat{Q}_0}\varphi_2)| \, dt \, dx \lesssim \int_{Q_0} Nu \, dx$$

for all dyadic cubes  $Q_0$ .

*Proof.* We have

$$\begin{aligned} 1_{\widehat{Q}_0}\varphi_2 &= \sum_{R \in \mathcal{R}, R \subset Q_0} (u - u(p_{\pi_S R})) \cdot 1_{W_R}, \\ \nabla(1_{Q_0}\varphi_2) &= \sum_{R \in \mathcal{R}, R \subset Q_0} [\nabla u \cdot 1_{W_R} + (u - u(p_{\pi_S R})) \cdot \nabla 1_{W_R}], \end{aligned}$$

and hence

$$\iint |\nabla(1_{Q_0}\varphi_2)| \, dt \, dx \lesssim \sum_{R \in \mathcal{R}, R \subset Q_0} \iint_{W_R} |\nabla u| \, dt \, dx + \sum_{R \in \mathcal{R}, R \subset Q_0} \inf_R Nu \cdot |R|, \quad (15)$$

using again

$$|\nabla 1_{W_R}| = H^n \lfloor \partial W_R, \quad H^n(\partial W_R) \lesssim |R|.$$

By Caccioppoli's inequality, we can estimate the first term in (15) by

$$\begin{aligned} \iint_{W_R} |\nabla u| dt dx &\leq \left( \iint_{W_R} |\nabla u|^2 dt dx \right)^{1/2} |W_R|^{1/2} \\ &\lesssim \frac{1}{\ell(R)} \left( \iint_{\tilde{W}_R} |u|^2 dt dx \right)^{1/2} |W_R|^{1/2} \\ &\lesssim \frac{1}{\ell(R)} \left( \iint_{\tilde{W}_R} \inf_R (Nu)^2 dt dx \right)^{1/2} |W_R|^{1/2} \\ &\lesssim \frac{1}{\ell(R)} \inf_R (Nu) |W_R| = \inf_R (Nu) |R|, \end{aligned}$$

which coincides with the second term in (15). So altogether

$$\iint |\nabla(1_{Q_0}\varphi_2)| dt dx \lesssim \sum_{R \in \mathcal{R}, R \subset Q_0} \inf_R Nu \cdot |R| \lesssim \int_{Q_0} Nu dx,$$

by Carleson's inequality and the Carleson property of  $\mathcal{R}$  from Lemma 6.5 in the last step.  $\square$

*Acknowledgments.* T.H. is supported by the European Research Council through the ERC Starting Grant "Analytic-probabilistic methods for borderline singular integrals". He is a member of the Finnish Centre of Excellence in Analysis and Dynamics Research. A.R. is supported by Grant 621-2011-3744 from the Swedish research council, VR.

## References

- [1] Auscher, P., Rosén, A., Rule, D.: Boundary value problems for degenerate elliptic equations and systems. *Ann. Sci. École. Norm. Sup. (4)* **48**, 951–1000 (2015) [Zbl 1328.35049](#) [MR 3377070](#)
- [2] Caffarelli, L. A., Fabes, E. B., Kenig, C. E.: Completely singular elliptic-harmonic measures. *Indiana Univ. Math. J.* **30**, 917–924 (1981) [Zbl 0482.35020](#) [MR 0632860](#)
- [3] Carleson, L.: Interpolations by bounded analytic functions and the corona problem. *Ann. of Math. (2)* **76**, 547–559 (1962) [Zbl 0112.29702](#) [MR 0141789](#)
- [4] Coifman, R. R., Meyer, Y., Stein, E. M.: Some new function spaces and their applications to harmonic analysis. *J. Funct. Anal.* **62**, 304–335 (1985) [Zbl 0569.42016](#) [MR 0791851](#)
- [5] Dahlberg, B. E. J.: Approximation of harmonic functions. *Ann. Inst. Fourier (Grenoble)* **30**, no. 2, 97–107 (1980) [Zbl 0417.31005](#) [MR 0584274](#)
- [6] Dahlberg, B. E. J., Jerison, D. S., Kenig, C. E.: Area integral estimates for elliptic differential operators with nonsmooth coefficients. *Ark. Mat.* **22**, 97–108 (1984) [Zbl 0537.35025](#) [MR 0735881](#)
- [7] Duoandikoetxea, J.: *Fourier Analysis*. Grad. Stud. Math. 29, Amer. Math. Soc., Providence, RI (2001) [Zbl 0969.42001](#) [MR 1800316](#)
- [8] Fefferman, C., Stein, E. M.:  $H^p$  spaces of several variables. *Acta Math.* **129**, 137–193 (1972) [Zbl 0257.46078](#) [MR 0447953](#)
- [9] Garnett, J.: *Bounded Analytic Functions*. Pure Appl. Math. 96, Academic Press, New York (1981) [Zbl 0469.30024](#) [MR 0628971](#)

- [10] Gundy, R. F., Wheeden, R. L.: Weighted integral inequalities for the nontangential maximal function, Lusin area integral, and Walsh–Paley series. *Studia Math.* **49**, 107–124 (1973/74) [Zbl 0271.28002](#) [MR 0352854](#)
- [11] Hofmann, S., Kenig, C., Mayboroda, S., Pipher, J.: Square function/non-tangential maximal function estimates and the Dirichlet problem for non-symmetric elliptic operators. *J. Amer. Math. Soc.* **28**, 483–529 (2015) [Zbl 1326.42028](#) [MR 3300700](#)
- [12] Hytönen, T., Rosén, A.: On the Carleson duality. *Ark. Mat.* **51**, 293–313 (2013) [Zbl 1294.42002](#) [MR 3090198](#)
- [13] Kenig, C., Koch, H., Pipher, J., Toro, T.: A new approach to absolute continuity of elliptic measure, with applications to non-symmetric equations. *Adv. Math.* **153**, 231–298 (2000) [Zbl 0958.35025](#) [MR 1770930](#)
- [14] Lerner, A., Nazarov, F.: Intuitive dyadic calculus: the basics. [arXiv:1508.05639](#) (2015)
- [15] Rosén, A.: A local  $Tb$  theorem for matrix weighted paraproducts. *Rev. Mat. Iberoamer.* **32**, 1259–1276 (2016) [Zbl 1366.42017](#) [MR 3593522](#)
- [16] Varopoulos, N.: BMO functions and the  $\bar{\partial}$ -equation. *Pacific J. Math.* **71**, 221–273 (1977) [Zbl 0371.35035](#) [MR 0508035](#)
- [17] Varopoulos, N.: A remark on functions of bounded mean oscillation and bounded harmonic functions. *Pacific J. Math.* **74**, 257–259 (1978) [Zbl 0382.31004](#) [MR 0508036](#)