Sharp Weighted Non-tangential Maximal Estimates via Carleson-Sparse Domination

The Journal of Geometric Analysis (2024) 34:360
https://doi.org/10.1007/s12220-024-01814-3

Sharp Weighted Non-tangential Maximal Estimates via Carleson-Sparse Domination

Andreas Rosén[1]

Received: 19 June 2024 / Accepted: 13 September 2024

Abstract

We prove sharp weighted estimates for the non-tangential maximal function of singular integrals mapping functions from $\mathbf{R}^n$ to the half-space in $\mathbf{R}^{1+n}$ above $\mathbf{R}^n$. The proof is based on pointwise sparse domination of the adjoint singular integrals that map functions from the half-space back to the boundary. It is proved that these map L_1 functions in the half-space to weak L_1 functions on the boundary. From this a non-standard sparse domination of the singular integrals is established, where averages have been replaced by Carleson averages.

Keywords Sparse domination · Non-tangential maximal functional · Carleson functional · Sharp weighted estimates

Mathematics Subject Classification 42B35 · 42B20 · 42B37

1 Introduction

Let us recall the sparse domination paradigm for estimating singular integral operators, which has been very successful in proving sharp weighted estimates for various singular operators for more than a decade. Given a singular integral operator T on $\mathbf{R}^n$, with Calderón–Zygmund kernel $k(x, y)$, the procedure for obtaining estimates is as follows.

(a) Boundedness of T on $L_2(\mathbf{R}^n)$ is proved. For classical convolution singular integral operators the Fourier transform is used, and for nonconvolution singular integral operators Tb theorems are used.
(b) Using the Calderón–Zygmund decomposition, L_2 boundedness and estimates of $k(x, y)$, weak $L_1(\mathbf{R}^n)$ boundedness of T is proved.

Andreas Rosén—Formerly Andreas Axelsson. This study was supported by the Swedish Research Council (Grant 2022-03996).

✉ Andreas Rosén
andreas.rosen@chalmers.se

[1] Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Gothenburg, Sweden

Published online: 01 October 2024

(c) Using weak $L_1(\mathbf{R}^n)$ boundedness of T and its grand maximal truncation operator M_T, a pointwise domination

$$|Tf(x)| \lesssim \sum_{Q\in\mathcal{D}_f,\, Q\ni x} \fint_{3Q} |f(y)|dy, \qquad x\in\mathbf{R}^n,$$

by a sum of averages of f, over a sparse collection $\mathcal{D}_f$ of dyadic cubes Q, is proved. Sparse roughly means that the cubes in $\mathcal{D}_f$ are essentially disjoint. See Sect. 4 for the definition.

(d) From the sparse domination, one can prove the boundedness of T on any Banach function space on which, together with its dual space, the maximal function is bounded. For example on weighted $L_p(w)$-spaces, $1 < p < \infty$, $w \in A_p(\mathbf{R}^n)$.

See for example Lerner [12] and Lerner and Nazarov [13]. We use the analyst's inequality $X \lesssim Y$, which means that $X \le CY$ for some constant $C < \infty$ independent of relevant variables but possibly depending on some parameters which should be clear from the context. For example, the C in (c) above is independent of f and x, but may depend on the sparseness parameter η. $X \gtrsim Y$ means $Y \lesssim X$, and $X \approx Y$ means $X \lesssim Y$ and $X \gtrsim Y$.

Somewhat hand in hand with singular integrals T as above, goes the theory of singular operator families $\{\Theta_t\}_{t>0}$, where Θ_t are integral operators

$$\Theta_t f(x) = \int_{\mathbf{R}^n} k(t,x;y) f(y)dy, \qquad x\in\mathbf{R}^n, t>0. \tag{1}$$

Equivalently, the operator family defines a mapping from functions on $\mathbf{R}^n$ to functions on the upper half-space $\mathbf{R}^{1+n}_+ = \{(t,x) : t>0, x\in\mathbf{R}^n\}$, and a basic L_2 estimate is the square function estimate

$$\int_0^\infty \|\Theta_t f\|_2^2 \frac{dt}{t} \lesssim \|f\|_2^2. \tag{2}$$

For families of classical convolution operators Θ_t, the square function estimate (2) is Calderón's reproducing formula and the function $(t,x) \mapsto \Theta_t f(x)$ is a continuous wavelet transform of f. See Daubechies [5]. For families of nonconvolution operators Θ_t, there are Tb theorems for proving (2). See for example Semmes [15] and Hofmann and Grau De La Herrán [6]. Sparse domination has been extended to this framework to prove sharp weighted estimates of the *square function*

$$x \mapsto \left(\int_0^\infty |\Theta_t f(x)|^2 \frac{dt}{t}\right)^{1/2}.$$

See Lerner [11] and Bailey et al. [2].

In the above works, the kernel of Θ_t is at least integrable on $\mathbf{R}^n$ and typically satisfies an estimate

$$|k(t, x; y)| \lesssim t^{-n} \frac{1}{(1 + |x - y|/t)^{n+\delta}}, \qquad x, y \in \mathbf{R}^n, t > 0, \tag{3}$$

for some $\delta > 0$, together with suitable off-diagonal Hölder estimates. In the present work, we consider operators Θ_t with kernels that are only locally integrable and typically satisfy (3) only for $\delta = 0$. The simplest example are the (non-singular) Riesz transforms

$$R_t^j f(x) = \int_{\mathbf{R}^n} \frac{x_j - y_j}{(t^2 + |x - y|^2)^{(1+n)/2}} f(y) dy, \qquad x \in \mathbf{R}^n, t > 0,$$

where $(t, x) \mapsto R_t^j f(x)$, $j = 1, \ldots, n$, are Stein–Weiss harmonic conjugate functions to the Poisson extension

$$P_t f(x) = R_t^0 f(x) = \int_{\mathbf{R}^n} \frac{t}{(t^2 + |x - y|^2)^{(1+n)/2}} f(y) dy, \qquad x \in \mathbf{R}^n, t > 0,$$

of f.

If $\delta = 0$, the square function estimate (2) is no longer suitable as an L_2 estimate to feed into a sparse domination scheme. This is because, in contrast to the case where the kernels are integrable, we generally do not have strong convergence $\Theta_t f \to 0$ for $\delta = 0$ when $t \to 0$, so the left hand side in (2) is typically infinite. Instead, a natural object to estimate when $\delta = 0$ is the non-tangential maximal function of $\Theta_t f(x)$. We therefore replace (2) with an estimate

$$\int_{\mathbf{R}^n} \left(\sup_{(t,x):|x-z|<\alpha t} |\Theta_t f(x)| \right)^2 dz \lesssim \|f\|_2^2, \tag{4}$$

which serves as our starting L_2 estimate for step (a) in a sparse domination scheme. Given an L_2 non-tangential estimate (4), we proceed by duality and consider the map

$$Sf(y) = \iint_{\mathbf{R}_+^{1+n}} k(t, x; y) f(t, x) dt dx, \qquad y \in \mathbf{R}^n, \tag{5}$$

which maps functions $f(t, x)$ defined in the upper half-space $\mathbf{R}_+^{1+n}$ to functions $(Sf)(y)$ on its boundary $\mathbf{R}^n$. By duality, (4) corresponds to the Carleson estimate

$$\|Sf\|_2^2 \lesssim \int_{\mathbf{R}^n} \left(\sup_{Q \ni z} \frac{1}{|Q|} \iint_{\widehat{Q}} |f(t, x)| dt dx \right)^2 dz \tag{6}$$

for S, where the sup is over all cubes $Q \subset \mathbf{R}^n$ containing z, and $\widehat{Q} \subset \mathbf{R}_+^{1+n}$ is the Carleson box, the cube with Q as its base. The equivalence of (6) and (4) follows from [7, Thm. 3.2]. Our main result, Theorem 4.1, shows that (6) implies the pointwise Carleson-sparse domination

$$|Sf(y)| \lesssim \sum_{Q\in\mathcal{D}_f, Q\ni y} \frac{1}{|Q|} \iint_{\widehat{3Q}} |f(t,x)|dtdx, \qquad \text{a.e. } y \in \mathbf{R}^n, \tag{7}$$

of S. Here $3Q$ denotes the cube with the same center as Q but three times the side length. Note that the Carleson averages in this sparse domination normalize by the $\mathbf{R}^n$ measure $|Q|$, and not by the $\mathbf{R}^{1+n}$ measure $|\widehat{3Q}| \approx |\widehat{Q}|$. This is in contrast to usual sparse domination techniques.

The four sections of this paper simply implement steps (a)-(d) of the sparse domination scheme described above. Starting from an L_2 non-tangential maximal estimate (4), or equivalently an L_2 Carleson estimate (6), we prove weighted L_p estimates. For Muckenhoupt weights $w \in A_p(\mathbf{R}^n)$, we estimate the $L_p(\mathbf{R}^n, w)$ operator norms by $[w]_{A_p}^{\max(p,q)/p}$, $1/p + 1/q = 1$. See Theorem 5.2 for the Carleson estimate and Theorem 5.4 for the dual non-tangential maximal estimate. The power of $[w]_{A_p}$ is optimal because it is known to be optimal for the $\mathbf{R}^n$ singular integral $\Theta_0 = \lim_{t\to 0}\Theta_t$, see Hytönen [10], which the non-tangential maximal function of $(\Theta_t)_{t>0}$ majorizes. Note that for $\delta > 0$ the non-tangential maximal function of $\Theta_t f$ is pointwise bounded by the maximal function of f. In this case the sharper weighted estimate by $[w]_{A_p}^{q/p}$ follows directly from Buckley [3, Thm. 2.5].

The sparse domination and estimates considered in this paper are related to those by Hytönen and Rosén [9]. There, L_p non-tangential maximal estimates for causal $\mathbf{R}^{1+n}_+$ Calderón–Zygmund operators

$$T^+ f(t,x) = \text{p.v.} \iint_{\mathbf{R}^{1+n}_+} K(t,x;s,y)g(s,y)dsdy, \qquad (s,t) \in \mathbf{R}^{1+n}_+$$

were proved. Here causal means the kernel condition $K(t,x;s,y) = 0$ for $s > t$, that is, T^+ is upward mapping. Both those results and the results in the present paper are proved by a sparse domination of the adjoint anti-causal, or downward mapping, operator. A natural attempt to prove an estimate of (5), for example in $L_p(\mathbf{R}^n)$-norm, is to apply the trace estimate from [8, Thm. 1.1] to obtain

$$\|Sf\|_{L_p(\mathbf{R}^n)} \lesssim \|C(\nabla\widetilde{S}f)\|_{L_p(\mathbf{R}^n)},$$

where C denotes the Carleson functional, which is defined within the square in (6) and $\widetilde{S}f$ denotes an extension of Sf to $\mathbf{R}^{1+n}_+$, defined by an auxiliary weakly singular integral operator $\widetilde{S}$ on $\mathbf{R}^{1+n}_+$. The idea would then be to apply [9, Thm. 5.1] to the singular integral $T = \nabla\widetilde{S}$ on $\mathbf{R}^{1+n}_+$. However this seems to be impossible. Technically, the boundedness of T would require a Whitney averaging in the Carleson norm. More seriously, according to [9, Ex. 2.1] it is necessary that $T = T^-$ is anti-causal in order for boundedness to be possible in the Carleson norm. To achieve anti-causality, a natural way is to truncate the kernel of $\widetilde{S}$ so that it becomes anti-causal. (This is possible without changing the boundary values Sf.) But at $t = s$, this adds an $\mathbf{R}^n$ singular integral operator to $T = \nabla\widetilde{S}$, which acts on $f(t, x)$ in the x variable for each $t > 0$. According to [9, Ex. 2.3], such a horizontal mapping singular integral will in general not be bounded in a Carleson norm. To summarize, the estimates considered in this paper are related to, but not implied by, those in [9].

2 Setup and L_2 Estimates

This section is about step (a) in the sparse domination scheme. First some notation. We fix a system of dyadic cubes $\mathcal{D} = \bigcup_{j\in\mathbf{Z}} \mathcal{D}^j$ in $\mathbf{R}^n$, where $\mathcal{D}^j$ are the cubes of side length $\ell(Q) = 2^{-j}$, such that the dyadic cubes in $\mathcal{D}$ form a connected tree under inclusion. Given $Q \in \mathcal{D}$, its parent is the minimal dyadic cube strictly containing Q, its grandparent is the parent of the parent, its children are the maximal dyadic cubes stricly contained in Q and its siblings are the other children of the parent of Q. For a given cube $Q \subset \mathbf{R}^n$, dyadic or not, we denote by cQ, $c > 0$, the cube with the same center as Q but with side length $\ell(cQ) = c\ell(Q)$. The Carleson box above Q is the $\mathbf{R}^{1+n}_+$ cube $\widehat{Q} = (0, \ell(Q)) \times Q$, and the Whitney region described by Q is the upper half $Q^w = (\ell(Q)/2, \ell(Q)) \times Q$ of $\widehat{Q}$.

With 1_E and $|E|$ we denote the indicator function and the measure of a set E. The maximal function of a function $f(x)$ on $\mathbf{R}^n$ is

$$Mf(x) = \sup_{r>0} \fint_{|y-x|<r} |f(y)|dy, \qquad x \in \mathbf{R}^n.$$

We denote the non-tangential maximal functional of a function $f(t, x)$ on $\mathbf{R}^{1+n}_+$ by

$$Nf(z) = \operatorname{esssup}_{(t,x):|x-z|<\alpha t}|f(t, x)|, \qquad z \in \mathbf{R}^n, \tag{8}$$

where the aperture $\alpha > 0$ of the cones is a fixed constant. We denote the Carleson functional of a function $f(t, x)$ on $\mathbf{R}^{1+n}_+$ by

$$Cf(z) = \sup_{Q\ni z} \frac{1}{|Q|} \iint_{\widehat{Q}} |f(t, x)|dtdx, \qquad z \in \mathbf{R}^n,$$

where the sup is over all (non-dyadic) cubes $Q \subset \mathbf{R}^n$ containing z. Versions of M, N and C, dyadic and non-dyadic, will appear below where they are needed.

Throughout this paper, we fix a kernel function $k : \mathbf{R}^{1+2n} \to \mathbf{R}$ which morally has the size estimates

$$|k(t, x; y)| \lesssim \frac{1}{|(t, x) - (0, y)|^n}.$$

However, we will not use this estimate, but only the following three. We assume, for some $\delta > 0$, off-diagonal Hölder estimates

$$|k(t+s, x+z; y) - k(t, x; y)| \lesssim \frac{|(s,z)|^\delta}{|(t,x)-(0,y)|^{n+\delta}} \tag{9}$$

for $|(s,z)| \le |(t,x)-(0,y)|/2$, and

$$|k(t, x; y+z) - k(t, x; y)| \lesssim \frac{|z|^\delta}{|(t,x)-(0,y)|^{n+\delta}} \tag{10}$$

for $|z| \le |(t,x)-(0,y)|/2$, where $s, t > 0$ and $x, y, z \in \mathbf{R}^n$.

Using the kernel $k(t, x; y)$ we define integral operators Θ_t as in (1) and an operator S as in (5), and we assume that L_2 Carleson estimates

$$\|Sf\|_{L_2(\mathbf{R}^n)} \lesssim \|Cf\|_{L_2(\mathbf{R}^n)} \tag{11}$$

hold, that is, (6). By duality, (11) is equivalent to L_2 non-tangential maximal estimates

$$\|N(S^* f)\|_{L_2(\mathbf{R}^n)} \lesssim \|f\|_{L_2(\mathbf{R}^n)}, \tag{12}$$

that is, (4), for the adjoint operator

$$S^* f(t,x) = \int_{\mathbf{R}^n} k(t,x;y) f(y) dy, \qquad (t,x) \in \mathbf{R}^{1+n}_+,$$

that is, (1). A standard method for verifying the non-tangential estimate (12) is to derive it from a gradient square function estimate

$$\int_0^\infty \|t\nabla\Theta_t f\|_2^2 \frac{dt}{t} \lesssim \|f\|_2^2,$$

which in turn can be proved by a Tb theorem applied to the kernel $t\nabla_{t,x}k(t,x;y)$. Indeed, $(t,x) \mapsto \Theta_t f(x)$ often solves an elliptic equation in applications, where L_2 estimates between non-tangential maximal functionals and square functionals are known. See [1, Sec. 10.1] for a large class of elliptic equations.

For the Riesz transforms mentioned in the introduction, there is the following alternative algebraic argument, also for Lipschitz domains.

Example 2.1 To simplify the notation, we only consider dimension $1+n = 2$. The following argument can be performed in higher dimensions, replacing the complex algebra with the Clifford algebra. See [14, Sec. 8.3] for the higher dimensional Cauchy integral that would be used for this argument.

Let $t = \phi(x)$ be the graph of a Lipschitz function $\phi : \mathbf{R} \to \mathbf{R}$, and consider the family of operators

$$\Theta_t f(x) = \frac{1}{2\pi i} \int_{\mathbf{R}} \frac{f(y)\,(1 + i\phi'(y))dy}{y + i\phi(y) - x - i(t + \phi(x))}, \qquad x \in \mathbf{R}, t > 0,$$

which represent the Cauchy integral that acts on functions on the graph. If $\phi = 0$ and f is real-valued, then $\operatorname{Re}\Theta_t f$ is the Poisson integral of f and $\operatorname{Im}\Theta_t f$ is the Riesz transform of f, but for general ϕ, both real and imaginary parts have only kernels with the decay $\delta = 0$ in (3).

We now realise that by the Cauchy integral formula, we have the algebraic identity

$$\Theta_t f = \Theta_t(\Theta_0 f),$$

since $g = \Theta_0 f = \lim_{t\to 0}\Theta_t f$ represents the trace on the graph of the analytic function $\Theta_t f$ above this graph. The function g belongs to the upper Hardy subspace, and therefore its Cauchy extension $\Theta_{-t} g = 0$, $t > 0$, vanishes below the graph. Thus

$$\Theta_t f = (\Theta_t - \Theta_{-t})(\Theta_0 f), \qquad t > 0,$$

where one proves that the kernel of $\Theta_t - \Theta_{-t}$ has Poisson kernel estimates $\lesssim t/(t^2 + (x-y)^2)$. From this follows the pointwise maximal estimate

$$N(S^* f)(z) \lesssim M(\Theta_0 f)(z), \qquad z \in \mathbf{R}.$$

Since $M : L_p(\mathbf{R}^n) \to L_p(\mathbf{R}^n)$ is bounded for $1 < p \le \infty$ and the singular Cauchy integral $\Theta_0 : L_p(\mathbf{R}^n) \to L_p(\mathbf{R}^n)$ is bounded for $1 < p < \infty$, (12) follows. However, the sharp weighted estimates in Theorem 5.4 below do not follow.

3 Weak L_1 Estimates

This section is about step (b) in the sparse domination scheme.

Proposition 3.1 *Assume that the operator S has L_2 Carleson estimates* (11) *and that the kernel k has x-regularity* (9)*. Then S has weak L_1-estimates*

$$|\{y \in \mathbf{R}^n : |Sf(y)| > \lambda\}| \lesssim \lambda^{-1}\|f\|_{L_1(\mathbf{R}^{1+n}_+)}, \qquad \lambda > 0.$$

For the proof, we need a twin of the Carleson functional, the area functional

$$Af(z) = \iint_{|x-z|<\alpha t} |f(t,x)|\,\frac{dt dx}{t^n}, \qquad z \in \mathbf{R}^n,$$

where the aperture $\alpha > 0$ of the cones is a fixed constant. Recall that for $1 \le p < \infty$, different apertures $\alpha > 0$ give equivalent norms $\|Af\|_{L_p(\mathbf{R}^n)}$. See [8, Prop. 2.2].

Lemma 3.2 *Let $g : \mathbf{R}^{1+n}_+ \to \mathbf{R}$ be a function such that $\|Cg\|_{L_\infty(\mathbf{R}^n)} \le \lambda$. Then*

$$\int_{\mathbf{R}^n} |Ag(z)|^2 dz \lesssim \lambda \int_{\mathbf{R}^n} |Ag(z)| dz.$$

Recall that

$$\|Af\|_{L_p(\mathbf{R}^n)} \eqsim \|Cf\|_{L_p(\mathbf{R}^n)} \tag{13}$$

for $1 < p < \infty$. See [8, Prop. 2.4]. However, for $p = \infty$ the area functional can be unbounded, even if the Carleson functional is bounded. We take the opportunity to correct a typo in the counterexample given in the first display after [8, Prop. 2.4]: It should read $f(t, x) = (t + |x|)^{-1}$ in each dimension n. Lemma 3.2 means that Ag is close enough to be bounded for the stated estimate to hold.

Proof The following is inspired by the good lambda inequality from [4, Thm. 3(a)]. Denote by $A^{(\alpha)}g$ and $A^{(\beta)}g$ the area functionals with the apertures α and β respectively. Assuming that $\|Cg\|_{L_\infty(\mathbf{R}^n)} \le \lambda$, we claim that the estimate

$$|\{z \,:\, A^{(\alpha)}g(z) > 2s\}| \lesssim (\lambda/s)|\{z \,:\, A^{(\beta)}g(z) > s\}|$$

holds for all $s > 0$, provided that $\alpha < \beta$. To see this, let $U = \bigcup_j Q_j$ be a Whitney decomposition of the open set $U = \{z \,:\, A^{(\beta)}g(z) > s\}$ into disjoint cubes $Q_j \subset \mathbf{R}^n$ such that $\ell(Q_j) \eqsim \operatorname{dist}(Q_j, \mathbf{R}^n \setminus U)$. It suffices to show

$$|\{z \in Q_j \,:\, A^{(\alpha)}g(z) > 2s\}| \lesssim (\lambda/s)|Q_j| \tag{14}$$

for each Q_j, since summing over j then gives the desired estimate. Let $M_j = \{z \in Q_j \,:\, A^{(\alpha)}g(z) > 2s\}$ and choose a point $z_j \in \mathbf{R}^n \setminus U$ such that $\operatorname{dist}(z_j, Q_j) \eqsim \ell(Q_j)$, so that $A^{(\beta)}g(z_j) \le s$. Now note that

$$\begin{aligned}\{(t, x) \,:\, |x - z| < \alpha t\} \subset \{&(t, x) \,:\, |x - z_j| < \beta t\} \\ &\cup \{(t, x) \,:\, |x - z| < \alpha t, t < c\ell(Q_j)\}\end{aligned}$$

for all $z \in Q_j$, for some constant $c > 1$ only depending on α, β. Indeed, if $|x - z| < \alpha t$ and $t \ge c\ell(Q_j)$, then $|x - z_j| \le \alpha t + |z - z_j|$ where $|z - z_j| \eqsim \ell(Q_j) \le t/c$. This gives the estimate

$$\begin{aligned}|M_j| &\le \frac{1}{2s}\int_{M_j} A^{(\alpha)}g(z)dz \\ &\le \frac{1}{2s}\int_{M_j} A^{(\beta)}g(z_j)dz + \frac{1}{2s}\int_{Q_j}\left(\iint_{|x-z|<\alpha t, t<c\ell(Q_j)} |g(t,x)|t^{-n}dtdx\right)dz \\ &\le |M_j|/2 + c's^{-1}\iint_{(1+2c\alpha)Q_j\times(0,c\ell(Q_j))} |g(t,x)|dtdx,\end{aligned}$$

for some $c' < \infty$. Since $\|Cg\|_{L_\infty(\mathbf{R}^n)} < \lambda$, this yields $|M_j| \lesssim s^{-1}|Q_j|\lambda$ as claimed.

Finally, multiplying (14) by s and integrating over $s \in (0, \infty)$ gives

$$\begin{aligned}\int_{\mathbf{R}^n} |A^{(\alpha)} g(z)|^2 dz &\eqsim \int_0^\infty s|\{z : A^{(\alpha)} g(z) > 2s\}|ds \\ &\lesssim \int_0^\infty \lambda|\{z : A^{(\beta)} g(z) > s\}|ds \\ &= \lambda \int_{\mathbf{R}^n} |A^{(\beta)} g(z)|dz.\end{aligned}$$

Since different apertures give equivalent L_2 norms and equivalent L_1 norms, this proves the lemma. □

Proof of Proposition 3.1 (i) We estimate f using a Calderón–Zygmund argument based on Carleson averages. Given $\lambda > 0$, let $Q_j \in \mathcal{D}$ denote the maximal dyadic cubes for which

$$\iint_{\widehat{Q_j}} |f(t,x)|dtdx > \lambda|Q_j|.$$

Define

$$b_j(t,x) = \begin{cases} f(t,x) - \frac{1}{|\widehat{Q_j}|} \iint_{\widehat{Q_j}} f(s,y)dsdy, & (t,x) \in \widehat{Q_j}, \\ 0, & \text{else,} \end{cases}$$

and $g = f - \sum_j b_j$. Note that in the definition of b_j, we do normalize the integral by $|\widehat{Q_j}|$ and not $|Q_j|$.

(ii) Estimating first Sg, we note that $\|Cg\|_{L_\infty(\mathbf{R}^n)} \lesssim \lambda$. Indeed, since g is constant equal to the $\widehat{Q}_j$-average of f on each $\widehat{Q}_j$, it follows that $|Q|^{-1} \iint_{\widehat{Q}} |g|dtdx \lesssim |Q_j|^{-1} \iint_{\widehat{Q_j}} |f|dtdx$ if $Q \subset Q_j$. Note that $\iint_{\widehat{Q_j}} |f|dtdx \le \iint_{\widehat{R}} |f|dtdx \le \lambda|R| \eqsim \lambda|Q_j|$, where R is the dyadic parent of Q_j. For Q not contained in any Q_j, we have $|Q|^{-1} \iint_{\widehat{Q}} |f|dtdx \le \lambda$. We can now apply (11), (13) and Lemma 3.2 to get

$$\begin{aligned}&|\{y \in \mathbf{R}^n : |Sg(y)| > \lambda\}| \le \lambda^{-2} \int_{\mathbf{R}^n} |Sg(y)|^2 dy \\ &\lesssim \lambda^{-2} \int_{\mathbf{R}^n} |Cg(y)|^2 dy \eqsim \lambda^{-2} \int_{\mathbf{R}^n} |Ag(y)|^2 dy \\ &\lesssim \lambda^{-1} \int_{\mathbf{R}^n} |Ag(y)|dy \eqsim \lambda^{-1} \|g\|_{L_1(\mathbf{R}^{1+n}_+)} \le \lambda^{-1} \|f\|_{L_1(\mathbf{R}^{1+n}_+)}.\end{aligned}$$

(iii) To estimate Sb_j, we do the standard estimate

$$|\{y : \left|S\Big(\sum_j b_j\Big)\right| > \lambda\}| \le \sum_j |3Q_j| + \lambda^{-1} \sum_j \int_{\mathbf{R}^n \setminus (3Q_j)} |Sb_j(y)|dy,$$

where

$$\sum_j |3Q_j| \lesssim \sum_j \lambda^{-1} \iint_{\widehat{Q_j}} |f(t,x)| dt dx \leq \lambda^{-1} \|f\|_{L_1(\mathbf{R}^{1+n}_+)}.$$

For the second term, we use that $\iint_{\widehat{Q_j}} b_j(t,x) dt dx = 0$ and (9) to estimate

$$\begin{aligned}
&\int_{\mathbf{R}^n \setminus (3Q_j)} |Sb_j(y)| dy \\
&= \int_{\mathbf{R}^n \setminus (3Q_j)} \left| \iint_{\widehat{Q_j}} (k(t,x;y) - k(t_Q, x_Q; y)) b_j(t,x) dt dx \right| dy \\
&\lesssim \iint_{\widehat{Q_j}} \left(\int_{\mathbf{R}^n \setminus (3Q_j)} \frac{\ell(Q)^\delta}{|(t_Q, x_Q) - (0,y)|^{n+\delta}} dy \right) |b_j(t,x)| dt dx \\
&\lesssim \iint_{\widehat{Q_j}} |b_j(t,x)| dt dx \lesssim \iint_{\widehat{Q_j}} |f(t,x)| dt dx,
\end{aligned}$$

where (t_Q, x_Q) denotes the center of $\widehat{Q}_j$. This yields

$$\sum_j \int_{\mathbf{R}^n \setminus (3Q_j)} |Sb_j(y)| dy \lesssim \sum_j \iint_{\widehat{Q_j}} |f(t,x)| dt dx \lesssim \|f\|_{L_1(\mathbf{R}^{1+n}_+)},$$

which completes the proof.

□

Remark 3.3 In the proof of Proposition 3.1, we can easily reduce to the case when f is constant on each dyadic Whitney region Q^w. Indeed, given $f \in L_1(\mathbf{R}^{1+n}_+)$, we write $f = f_0 + f_1$, where $\iint_{Q^w} f_0(t,x) dt dx = 0$ and f_1 is constant on each dyadic Whitney region Q^w. For f_0, using (9), we have strong L_1 estimates

$$\begin{aligned}
\int_{\mathbf{R}^n} |Sf_0(y)| dy &= \int_{\mathbf{R}^n} \left| \sum_{Q \in \mathcal{D}} \iint_{Q^w} (k(t,x;y) - k(t_Q, x_Q; y)) f_0(t,x) dt dx \right| dy \\
&\lesssim \sum_{Q \in \mathcal{D}} \iint_{Q^w} \left(\int_{\mathbf{R}^n} \frac{\ell(Q)^\delta}{|(t_Q, x_Q) - (0,y)|^{n+\delta}} dy \right) |f_0(t,x)| dt dx \\
&\lesssim \sum_{Q \in \mathcal{D}} \iint_{Q^w} |f_0(t,x)| dt dx = \|f_0\|_{L_1(\mathbf{R}^{1+n}_+)},
\end{aligned}$$

where (t_Q, x_Q) now denotes the center of Q^w. Thus it remains to estimate Sf_1.

For the sparse domination of S, we require the maximal operator

$$M_S f(y) = \sup_{Q \ni y} \|S(1_{\mathbf{R}^{1+n}_+ \setminus \widehat{3Q}} f)\|_{L_\infty(Q)}, \qquad y \in \mathbf{R}^n,$$

where the sup is taken over all dyadic cubes $Q \subset \mathbf{R}^n$ containing y. This is a version of Lerner's grand maximal truncation operator from [12], which is enough for our purposes.

Proposition 3.4 *Assume that the operator S has L_2 Carleson estimates* (11) *and that the kernel k has the x-regularity* (9) *and the y-regularity* (10). *Then M_S has the weak L_1 bound*

$$|\{y \in \mathbf{R}^n : |M_S f(y)| > \lambda\}| \lesssim \lambda^{-1} \|f\|_{L_1(\mathbf{R}^{1+n}_+)}, \qquad \lambda > 0.$$

Proof This result follows by tweaking the standard proof of Cotlar's lemma. Fix a cube $Q \subset \mathbf{R}^n$. Let

$$f_1 = 1_{\widehat{3Q}} f \quad \text{and} \quad f_2 = 1_{\mathbf{R}^{1+n}_+ \setminus \widehat{3Q}} f.$$

At a given point $y_1 \in Q$, we estimate $Sf_2(y_1)$ by averaging over a variable point $y \in Q$. Write

$$Sf_2(y_1) = (Sf_2(y_1) - Sf_2(y)) + Sf(y) - Sf_1(y) = I + II + III.$$

Fixing $p > 1$, we estimate term II as

$$\fint_Q |Sf(y)|^{1/p} dy \lesssim \inf_Q M(|Sf|^{1/p})$$

and term III as

$$\fint_Q |Sf_1(y)|^{1/p} dy \lesssim |Q|^{-1/p} \|f_1\|^{1/p}_{L_1(\mathbf{R}^{1+n}_+)} \lesssim (\inf_Q Cf)^{1/p},$$

using Kolmogorov's inequality and Proposition 3.1. To estimate term I, we use (10) to get

$$|Sf_2(y_1) - Sf_2(y)| \lesssim \iint_{\mathbf{R}^{1+n}_+ \setminus \widehat{3Q}} \frac{\ell(Q)^\delta}{|(t,x) - (0,y_0)|^{n+\delta}} |f(t,x)| dt dx,$$

for any $y_0 \in Q$. Writing $r = |(t,x) - (0,y_0)|$ and $r^{-n-\delta} \eqsim \int_r^\infty s^{-n-1-\delta} ds$, we get

$$\begin{aligned} |Sf_2(y_1) - Sf_2(y)| &\lesssim \ell(Q)^\delta \iint_{\mathbf{R}^{1+n}_+ \setminus \widehat{3Q}} \left(\int_r^\infty \frac{ds}{s^{n+1+\delta}} \right) |f(t,x)| dt dx \\ &= \ell(Q)^\delta \int_{\ell(Q)}^\infty \frac{ds}{s^{n+1+\delta}} \left(\iint_{\{|(t,x-y_0)|<s\} \setminus \widehat{3Q}} |f(t,x)| dt dx \right) \\ &\lesssim \ell(Q)^\delta \int_{\ell(Q)}^\infty \frac{ds}{s^{1+\delta}} Cf(y_0) \eqsim Cf(y_0). \end{aligned}$$

Collecting the estimates, we have prove the pointwise estimate

$$M_S f \lesssim Cf + M(|Sf|^{1/p})^p + Cf.$$

Here M is bounded on $L_{p,\infty}(\mathbf{R}^n)$, which combined with Proposition 3.1 gives the estimate for the second term. A standard Vitali covering argument finally shows that

$$|\{y \in \mathbf{R}^n : |Cf(y)| > \lambda\}| \lesssim \lambda^{-1}\|f\|_{L_1(\mathbf{R}^{1+n}_+)}, \qquad \lambda > 0,$$

which completes the proof. □

4 Carleson-Sparse Domination

This section is about step (c) in the sparse domination scheme.

We recall that a collection of dyadic cubes $\widetilde{\mathcal{D}} \subset \mathcal{D}$ is called η-sparse, $\eta > 0$, if each $Q \in \widetilde{\mathcal{D}}$ contains a subset $E_Q \subset Q$ such that $|E_Q| \geq \eta|Q|$ and $E_R \cap E_Q = \emptyset$ whever $R \neq Q$, $Q, R \in \widetilde{\mathcal{D}}$.

Our main result in this paper is the following Carleson-sparse domination. The proof given below is an adaption of the estimate in [12] to singular integrals mapping from $\mathbf{R}^{1+n}_+$ to $\mathbf{R}^n$.

Theorem 4.1 *Assume that the operator S has L_2 Carleson estimates (11) and that the kernel k has the x-regularity (9) and the y-regularity (10). Fix $0 < \eta < 1$. Then for any $f \in L_1(\mathbf{R}^{1+n}_+)$ with bounded support, there exists an η-sparse family $\mathcal{D}_f$ of dyadic cubes such that*

$$|Sf(y)| \lesssim \sum_{Q \in \mathcal{D}_f, Q \ni y} \frac{1}{|Q|} \iint_{\widehat{3Q}} |f(t,x)| dt dx, \qquad \textit{for a.e. } y \in \mathbf{R}^n.$$

Proof (i) Let $c, \alpha > 0$ be constants to be chosen below. Let $Q \in \mathcal{D}$ be such that $\operatorname{supp} f \subset \widehat{3Q}$. Define

$$E = \Big\{ y \in \mathbf{R}^n : \max\big(|Sf(y)|, M_S f(y)\big) > c|Q|^{-1} \iint_{\widehat{3Q}} |f| dt dx \Big\}, \qquad (15)$$

and let $R_j \in \mathcal{D}$ be the maximal subcubes $R_j \subset Q$ such that

$$|R_j \cap E| > \alpha |R_j|.$$

Using maximality, we can also obtain a converse estimate by noting that

$$|R_j \cap E| \leq |R_j^p \cap E| \leq \alpha |R_j^p| = \alpha 2^n |R_j|,$$

where R_j^p denotes the dyadic parent of R_j. We choose $\alpha = 1/(2^n + 1)$, so that each R_j contains a substantial part of both E and $\mathbf{R}^n \setminus E$ in the sense that

$$\min(|R_j \cap E|, |R_j \setminus E|) \geq |R_j|/(2^n + 1). \tag{16}$$

Lebesgue's differentiation theorem shows that

$$E \cap Q \subset \bigcup_j R_j, \tag{17}$$

modulo a set of measure zero. From Propositions 3.1 and 3.4, we obtain

$$|E| \lesssim \left(c|Q|^{-1} \iint_{\widehat{3Q}} |f| dt dx\right)^{-1} \iint_{\widehat{3Q}} |f| dt dx = |Q|/c.$$

Since $\bigcup R_j = \{M(1_E) > \alpha\}$, by the weak L_1 boundedness of M we have

$$\sum_j |R_j| \lesssim \alpha^{-1}|E| \lesssim |Q|/(\alpha c).$$

Having fixed $\alpha = 1/(2^n + 1)$, we choose c large enough so that

$$\sum_j |R_j| \leq (1 - \eta)|Q|. \tag{18}$$

(ii) Recall that supp $f \subset \widehat{3Q}$, so that $1_{\widehat{3Q}} f = f$. Now write

$$\begin{aligned} 1_Q S(1_{\widehat{3Q}} f) &= 1_{Q \setminus \bigcup R_j} S(1_{\widehat{3Q}} f) + \sum_j 1_{R_j} S(1_{\mathbf{R}^{1+n}_+ \setminus \widehat{3R_j}} f) + \sum_j 1_{R_j} S(1_{\widehat{3R_j}} f) \\ &= I + II + \sum_j 1_{R_j} S(1_{\widehat{3R_j}} f). \end{aligned}$$

- For term I, by (17) we have $y \notin E$ for a.e. $y \in Q \setminus \bigcup R_j$, and therefore (15) gives

$$|S(1_{\widehat{3Q}} f)(y)| \lesssim |Q|^{-1} \iint_{\widehat{3Q}} |f| dt dx.$$

- For subcube R_j in term II, by (16) there exists $y' \in R_j \setminus E$. For all $y \in R_j$, we therefore have

$$|S(1_{\mathbf{R}^{1+n}_+ \setminus \widehat{3R_j}}) f(y)| \lesssim M_S f(y') \lesssim |Q|^{-1} \iint_{\widehat{3Q}} |f| dt dx,$$

using (15).

To summarize, we have shown that

$$1_Q S(1_{\widehat{3Q}} f) \leq C_1 |Q|^{-1} \iint_{\widehat{3Q}} |f| dt dx + \sum_j 1_{R_j} S(1_{\widehat{3R_j}} f), \tag{19}$$

for some constant $C_1 < \infty$, where the disjoint subcubes satisfy (18).

(iii) We can now iterate (19) to get the stated sparse estimate as follows. Choose $Q_1 \in \mathcal{D}$ such that $\text{supp}\, f \subset \widehat{Q_1}$. Set $P_1 = Q_1$ and recursively define P_{j+1} to be the dyadic parent of P_j, $j = 1, 2 \ldots$. Let $Q_2, Q_3, \ldots$ be an ordering of all the siblings of all the cubes P_j, $j = 1, 2, \ldots$. We obtain a disjoint union $\mathbf{R}^n = \bigcup_{j=1}^{\infty} Q_j$ modulo a zero set. Since Q_1 is contained in a sibling of Q_j, it follows that $3Q_j \supset Q_1$, and therefore $\widehat{3Q_j} \supset \text{supp}\, f$, $j = 1, 2, \ldots$. Hence

$$Sf = \sum_j 1_{Q_j} Sf = \sum_j 1_{Q_j} S(1_{\widehat{3Q_j}} f). \tag{20}$$

We now apply the estimate in step (ii) above to each $Q = Q_j$, which produces first generations of subcubes $R_k \subset Q_j$. Then we apply the estimate in step (ii) above to each such first generation subcube $Q = R_k$, which produces second generations of subcubes. Continuing recursively in this way, we define the family of dyadic cubes $\mathcal{D}_f$ as the union of all Q_j along with all generations of subcubes R_k. For $Q \in \mathcal{D}_f$, we define

$$E_Q = Q \setminus \bigcup_k R_k,$$

where R_k are all the subcubes of Q, constucted from Q as in step (i) above. It follows from (18) that $\mathcal{D}_f$ is η-sparse. Combining (20) and recursively (19), the stated sparse domination of Sf follows. This completes the proof.

□

5 Sharp Weighted Estimates

This section is about step (d) in the sparse domination scheme. We first derive weighted L_p Carleson estimates of S from Theorem 4.1, and then use duality between the Carleson and non-tangential maximal functionals to obtain weighted L_q non-tangential maximal estimates of S^*.

We fix $1 < p < \infty$ and $1/p + 1/q = 1$, a let $w(x) > 0$, $x \in \mathbf{R}^n$, be an A_p weight, that is,

$$[w]_{A_p}^{1/p} = \sup_Q \left(\fint_Q w dx \right)^{1/p} \left(\fint_Q \nu dx \right)^{1/q} < \infty,$$

where the sup is over all cubes Q, and $\nu = w^{-q/p}$ is the dual weight. It is readily checked that ν is an A_q weight with $[\nu]_{A_q}^{1/q} = [w]_{A_p}^{1/p}$. Write $w(Q) = \int_Q dw$ for the w-measure of Q, where $dw = wdx$, and similarly for ν.

Some additional technicalities arise since we aim to prove weighted estimates which are sharp with the respect to the dependence on the A_p characteristic $[w]_{A_p}$ of the weight. In particular we need to avoid using the doubling property of the measures quantitatively. To this end, dyadic operators are preferred. We also need to handle dilations $3Q$ of dyadic cubes $Q \in \mathcal{D}$ coming from Theorem 4.1.

Lemma 5.1 *Let $1 < p \leq \infty$. Then we have maximal function estimates*

$$\|M_w^{3D} f\|_{L_p(w)} \lesssim \|f\|_{L_p(w)}$$

for the centered dyadic weighted maximal operator

$$M_w^{3D} f(z) = \sup_{Q \in \mathcal{D}, Q \ni z} \frac{1}{w(3Q)} \int_{3Q} |f| dw.$$

The same estimate holds for the standard dyadic weighted maximal operator

$$M_w^{D} f(z) = \sup_{Q \in \mathcal{D}, Q \ni z} \frac{1}{w(Q)} \int_{Q} |f| dw.$$

The implicit constant is independent of w, but depends on p.

Proof It suffices to consider M_w^{3D} as the proof for M_w^D is similar but simpler. For $p = \infty$, clearly $\|M_w^{3D} f\|_{L_\infty(w)} \leq \|f\|_{L_\infty(w)}$. By the Marcinkiewicz interpolation theorem, it suffices to show the weak L_1 estimate

$$w(\{z : M^{3D} f(z) > \lambda\}) \lesssim \lambda^{-1} \|f\|_{L_1(w)}.$$

To see this, for given $\lambda > 0$ we consider the dyadic cubes $Q \in \mathcal{D}$ such that $\iint_{3Q} |f| dw > \lambda w(3Q)$. Denote by $\{Q_j\}$ the maximal such cubes. Then Q_j are disjoint and $\{z : M^{3D} f(z) > \lambda\} = \bigcup_j Q_j$. This gives

$$\begin{aligned} w(\{z : M^{3D} f(z) > \lambda\}) &= \sum_j w(Q_j) \leq \sum_j w(3Q_j) \\ &\leq \lambda^{-1} \sum_j \int_{3Q_j} |f| dw = \lambda^{-1} \int_{\mathbf{R}^n} \sum_j 1_{3Q_j} |f| dw \\ &\leq \lambda^{-1} 3^n \|f\|_{L_1(w)}. \end{aligned}$$

□

Theorem 5.2 *Let $1 < p < \infty$. Assume that the operator S has estimates (11) and that the kernel k has estimates (9) and (10). Then we have estimates*

$$\|Sf\|_{L_p(w)} \lesssim [w]_{A_p}^{\max(1,q/p)} \|C_\nu^D(f\nu^{-1})\|_{L_p(\nu)},$$

where the dyadic weighted Carleson functional is

$$C_\nu^D f(z) = \sup_{Q\in\mathcal{D}, Q\ni z} \frac{1}{\nu(Q)} \iint_{\widehat{Q}} |f(t,x)| dt d\nu.$$

Here the implicit constant in the estimate is independent of w and ν.

Proof (i) We first estimate Sf by the auxiliary Carleson functional

$$C_\nu^{3D} f(z) = \sup_{Q\in\mathcal{D}, Q\ni z} \frac{1}{\nu(cQ)} \iint_{\widehat{3Q}} |f(t,x)| dt d\nu,$$

where $c \geq 1$ is a fixed constant. Let $g : \mathbf{R}^n \to \mathbf{R}$ be a dual function for which equality holds in Hölder's inequality $\int_{\mathbf{R}^n} (Sf) g dx \leq \|Sf\|_{L_p(w)} \|g\|_{L_q(\nu)}$. Using Theorem 4.1 and adapting the argument [12, Sec. 5], we estimate

$$\begin{aligned}
\int_{\mathbf{R}^n} (Sf) g dx &\lesssim \sum_{Q\in\mathcal{D}_f} \left(\frac{1}{|Q|} \iint_{\widehat{3Q}} |f| dt dx\right)\left(\int_Q |g| dx\right) \\
&= \sum_{Q\in\mathcal{D}_f} B_Q \left(\frac{\nu(E_Q)^{1/p}}{\nu(cQ)} \iint_{\widehat{3Q}} |f| dt dx\right)\left(\frac{w(E_Q)^{1/q}}{w(cQ)} \int_Q |g| dx\right) \\
&\leq \sup_{Q\in\mathcal{D}} B_Q \left(\sum_{Q\in\mathcal{D}_f} \left(\frac{1}{\nu(cQ)} \iint_{\widehat{3Q}} |f| dt dx\right)^p \nu(E_Q)\right)^{1/p} \\
&\quad \times \left(\sum_{Q\in\mathcal{D}_f} \left(\frac{1}{w(cQ)} \int_Q |g| dx\right)^q w(E_Q)\right)^{1/q} \\
&\leq \left(\sup_{Q\in\mathcal{D}} B_Q\right) \|C_\nu^{3D}(f\nu^{-1})\|_{L_p(\nu)} \|M_w^D(gw^{-1})\|_{L_q(w)},
\end{aligned}$$

using that $w(cQ) \geq w(Q)$ in the last inequality, where $E_Q \subset Q$ are the disjoint ample subsets of $Q \in \mathcal{D}_f$ and

$$\begin{aligned}
B_Q &= \frac{\nu(cQ) w(cQ)}{|Q| \nu(E_Q)^{1/p} w(E_Q)^{1/q}} \\
&= \left(\frac{w(cQ)}{w(E_Q)}\right)^{1/q} \left(\frac{\nu(cQ)}{\nu(E_Q)}\right)^{1/p} \left(\left(\frac{w(cQ)}{|Q|}\right)^{1/p} \left(\frac{\nu(cQ)}{|Q|}\right)^{1/q}\right).
\end{aligned}$$

To estimate B_Q, write $a = w(cQ)/w(E_Q), b = \nu(cQ)/\nu(E_Q)$ and $\gamma = [w]_{A_p}^{1/p}$, so that

$$B_Q \lesssim a^{1/q} b^{1/p} \gamma, \tag{21}$$

since $|Q| \eqsim |cQ|$. Using that $|Q| \lesssim |E_Q| \le w(E_Q)^{1/p} \nu(E_Q)^{1/q}$ by Hölder's inequality, we have

$$a^{1/p} b^{1/q} \lesssim \gamma. \tag{22}$$

Combining (21) and (22) and noting $a, b \ge 1$, we have $B_Q \le (\gamma b^{-1/q})^{p/q} b^{1/p} \gamma \le \gamma^p$ if $p \ge 2$ and $B_Q \le a^{1/q} (\gamma a^{-1/p})^{q/p} \gamma \le \gamma^q$ if $p \le 2$. This proves that $\sup_{Q \in \mathcal{D}} B_Q \lesssim [w]_{A_p}^{\max(1,q/p)}$ and yields the stated estimate of $\|Sf\|_{L_p(w)}$, since

$$\|M_w^D(gw^{-1})\|_{L_q(w)} \lesssim \|gw^{-1}\|_{L_q(w)} = \|g\|_{L_q(\nu)}.$$

(ii) Next we prove the stated estimate by C_ν^D. By (i), it suffices to show that

$$\|C_\nu^{3D} f\|_{L_p(\nu)} \lesssim \|C_\nu^D f\|_{L_p(\nu)}, \tag{23}$$

for some $c \ge 1$ in the definition of $C_\nu^{3D} f$. Assume that $C_\nu^{3D} f(z) > \lambda$. Then there exists $Q \in \mathcal{D}$ such that $Q \ni z$ and

$$\iint_{\widehat{3Q}} |f| dt d\nu > \lambda \nu(cQ).$$

Let $P_j \in \mathcal{D}$ denote the at most 2^n dyadic cubes of side length $4\ell(Q)$ such that $P_j \cap (3Q) \ne \emptyset$, where $P_1 \supset Q$ is the grandparent of Q. By the pigeonhole principle,

$$\iint_{\widehat{P_j}} |f| dt d\nu \ge 2^{-n} \iint_{\widehat{3Q}} |f| dt d\nu$$

holds for at least one of the dyadic Carleson regions $\widehat{P_j}$, since $\bigcup_j \widehat{P_j} \supset \widehat{3Q}$. We get

$$\frac{1}{\nu(P_j)} \iint_{\widehat{P_j}} |f| dt d\nu \ge 2^{-n} \lambda \nu(cQ)/\nu(P_j),$$

and so $C_\nu^D f \ge 2^{-n} \lambda \nu(cQ)/\nu(P_j)$ on P_j. Since $z \in P_1 \in \mathcal{D}$ and $3P_1 \supset P_j$, we have

$$M_\nu^{3D}(C_\nu^D f)(z) \ge \frac{1}{\nu(3P_1)} \int_{P_j} 2^{-n} \lambda \nu(cQ)/\nu(P_j) d\nu = \lambda 2^{-n} \nu(cQ)/\nu(3P_1).$$

Choose $c = 15$. Then $cQ \supset 3P_1$ and therefore $M_\nu^{3D}(C_\nu^D f)(z) \ge \lambda 2^{-n}$. Letting $\lambda \to C_\nu^{3D} f(z)$, we have shown that

$$C_\nu^{3D} f \le 2^n M_\nu^{3D}(C_\nu^D f).$$

The estimate (23) now follows from Lemma 5.1, which completes the proof. □

To obtain sharp weighted estimates of $N^D(S^*)$ via duality, where N^D is the dyadic non-tangential maximal functional

$$N_D f(z) = \sup_{Q\in\mathcal{D}, Q\ni z} \|f\|_{L_\infty(Q^w)}, \qquad z \in \mathbf{R}^n,$$

we need the following dyadic weighted duality estimates.

Proposition 5.3 *Let $1 < p < \infty$. Then*

$$\left| \iint_{\mathbf{R}^{1+n}_+} fg dt d\nu \right| \lesssim \|N^D f\|_{L_q(\nu)} \|C^D_\nu g\|_{L_p(\nu)}. \tag{24}$$

Moreover, given any $f : \mathbf{R}^{1+n}_+ \to \mathbf{R}$ with $\|N^D f\|_{L_q(\nu)} < \infty$, there exists a non-vanishing $g : \mathbf{R}^{1+n}_+ \to \mathbf{R}$ such that

$$\|N^D f\|_{L_q(\nu)} \|C^D_\nu g\|_{L_p(\nu)} \lesssim \iint_{\mathbf{R}^{1+n}_+} fg dt d\nu. \tag{25}$$

The implicit constants in the estimates are independent of the weight ν.

Proof The result to be proved is a weighted extension of [7, Thm. 2.2], and the proof is a straightforward modification of the unweighted proof. Since we shall not use (24), we only give the proof of (25) to convince the reader that the constants there are independent of ν.

Given f and $\lambda > 0$, we let $\mathcal{D}_\lambda \subset \mathcal{D}$ be the set of maximal cubes $Q \in \mathcal{D}$ such that $\|f\|_{L_\infty(Q^w)} > \lambda$. Then the cubes in $\mathcal{D}_\lambda$ are disjoint and we have $\{z : N^D f(z) > \lambda\} = \bigcup \mathcal{D}_\lambda$. We compute

$$\begin{aligned}
\|N^D f\|^q_{L_q(\nu)} &= \int_0^\infty \lambda^q \, \nu(\{N^D f(z) > \lambda\}) \frac{d\lambda}{\lambda} \\
&= \int_0^\infty \lambda^q \sum_{Q\in\mathcal{D}_\lambda} \nu(Q) \frac{d\lambda}{\lambda} = \sum_{Q\in\mathcal{D}} \nu(Q) \int_{\lambda:\mathcal{D}_\lambda \ni Q} \lambda^q \frac{d\lambda}{\lambda} \\
&\le \sum_{Q\in\mathcal{D}} \|f\|_{L_\infty(Q^w)} \left(\nu(Q) \int_{\lambda:\mathcal{D}_\lambda\ni Q} \lambda^{q-1} \frac{d\lambda}{\lambda} \right) = \sum_{Q\in\mathcal{D}} \|f\|_{L_\infty(Q^w)} g_Q,
\end{aligned}$$

where we have set $g_Q = \nu(Q) \int_{\lambda:\mathcal{D}_\lambda\ni Q} \lambda^{q-1} \frac{d\lambda}{\lambda}$. For the inequality, we used that $\lambda < \|f\|_{L_\infty(Q^w)}$ when $Q \in \mathcal{D}_\lambda$. Now define g to be a function on $\mathbf{R}^{1+n}_+$, whose restriction to Q^w satisfies

$$\iint_{Q^w} fg dt d\nu \eqsim \|f\|_{L_\infty(Q^w)} \|g\|_{L_1(Q^w, dt d\nu)} = \|f\|_{L_\infty(Q^w)} g_Q.$$

To estimate $C_\nu^D g$, for $Q \in \mathcal{D}$ we compute

$$\begin{aligned}\frac{1}{\nu(Q)} \iint_{\widehat{Q}} |g| dt d\nu &= \frac{1}{\nu(Q)} \sum_{R \subset Q} \nu(R) \int_{\lambda : \mathcal{D}_\lambda \ni R} \lambda^{q-1} \frac{d\lambda}{\lambda} \\ &= \frac{1}{\nu(Q)} \int_0^\infty \lambda^{q-1} \sum_{R \in \mathcal{D}_\lambda, R \subset Q} \nu(R) \frac{d\lambda}{\lambda} \\ &= \frac{1}{\nu(Q)} \int_0^\infty \lambda^{q-1} \, \nu(\{N^D f > \lambda\} \cap Q) \frac{d\lambda}{\lambda} \\ &= \frac{1}{\nu(Q)} \int_Q (N^D f)^{q-1} d\nu \le \inf_Q M_\nu^D ((N^D f)^{q-1}).\end{aligned}$$

Using Lemma 5.1, this shows that

$$\|C_\nu^D g\|_{L_p(\nu)} \le \|M_\nu^D ((N^D f)^{q-1})\|_{L_p(\nu)} \lesssim \|(N^D f)^{q-1}\|_{L_p(\nu)} = \|N^D f\|_{L_q(\nu)}^{q-1}.$$

Therefore

$$\|N^D f\|_{L_q(\nu)} \|C_\nu^D g\|_{L_p(\nu)} \lesssim \|N^D f\|_{L_q(\nu)}^q \le \sum_{Q \in \mathcal{D}} \|f\|_{L_\infty(Q^w)} g_Q \eqsim \iint_{\mathbf{R}_+^{1+n}} f g dt d\nu,$$

which completes the proof. □

With the above, we are now in position to prove sharp weighted non-tangential maximal estimates for $\{\Theta_t\}_{t>0}$.

Theorem 5.4 *Let* $1 < q < \infty$. *Assume that the operator* S^* *has estimates* (12) *and that the kernel* k *has estimates* (9) *and* (10). *Then for* $f \in L_q(\mathbf{R}^n, \nu)$, *we have estimates*

$$\|N(S^* f)\|_{L_q(\nu)} \lesssim [\nu]_{A_q}^{\max(1, p/q)} \|f\|_{L_q(\nu)},$$

for any fixed aperture $\alpha > 0$ *used in the definition* (8) *of* N. *Here the implicit constant in the estimate is independent of* ν.

Proof (i) The corresponding dyadic estimate, that is, with N replaced by N^D, follows immediately from Proposition 5.3 and Theorem 5.2 since

$$\begin{aligned}\|N^D(S^* f)\|_{L_q(\nu)} &\lesssim \iint_{\mathbf{R}_+^{1+n}} (S^* f)(t, x) g(t, x) dt d\nu \Big/ \|C_\nu^D g\|_{L_p(\nu)} \\ &= \int_{\mathbf{R}^n} f(x) \, S(g\nu)(x) dx \Big/ \|C_\nu^D g\|_{L_p(\nu)} \\ &\le \|f\|_{L_q(\nu)} \|S(g\nu)\|_{L_p(w)} \Big/ \|C_\nu^D g\|_{L_p(\nu)} \\ &\lesssim \|f\|_{L_q(\nu)} [w]_{A_p}^{\max(1, q/p)} = [\nu]_{A_q}^{\max(1, p/q)} \|f\|_{L_q(\nu)},\end{aligned}$$

where g is a function dual to $S^* f$, provided by Proposition 5.3.

(ii) For the estimate of $N(S^* f)(z)$, $z \in \mathbf{R}^n$, let (t_z, x_z) be such that $|x_z - z| < \alpha t_z$ and $N(S^* f)(z) \eqsim |S^* f(t_z, x_z)|$. Write

$$S^* f(t_z, x_z) = (S^* f(t_z, x_z) - S^* f(t_z, z)) + S^* f(t_z, z) = I + II.$$

For term II, we have $|S^* f(t_z, z)| \le N^D f(z)$. For term I, we have

$$|S^* f(t_z, x_z) - S^* f(t_z, z)| \lesssim \int_{\mathbf{R}^n} \frac{|x_z - z|^\delta}{|(t_z, z) - (0, y)|^{n+\delta}} |f(y)| dy.$$

Using $|x_z - z|^\delta \lesssim t_z^\delta$ and writing

$$\frac{1}{(1+r)^{n+\delta}} \eqsim \int_r^\infty \frac{ds}{(1+s)^{n+\delta+1}},$$

this gives

$$|S^* f(t_z, x_z) - S^* f(t_z, z)| \lesssim \int_{\mathbf{R}^n} \frac{t_z^{-n}}{(1 + |y - z|/t_z)^{n+\delta}} |f(y)| dy$$
$$\eqsim \int_0^\infty t_z^{-n} \left(\int_{|y-z|/t_z < s} |f(y)| dy \right) \frac{ds}{(1+s)^{n+\delta+1}} \lesssim Mf(z).$$

Therefore we have the pointwise estimate

$$N(S^* f) \lesssim N^D f + Mf$$

The estimate in (i) above and the estimate $\|Mf\|_{L_q(\nu)} \lesssim [\nu]_{A_q}^{p/q} \|f\|_{L_q(\nu)}$ from [3, Thm. 2.5] now completes the proof.

□

We note that in the proof of Theorem 5.4, it would have sufficed to estimate the vertical maximal function $\| \sup_{t>0} |\Theta_t f| \|_{L_q(\nu)}$ in step (i).

Funding Open access funding provided by University of Gothenburg.

Data availability Not applicable.

References

1. Auscher, P., Axelsson, A.: Weighted maximal regularity estimates and solvability of non-smooth elliptic systems. I. Invent. Math. **184**(1), 47–115 (2011)
2. Bailey, J., Brocchi, G., Reguera, M.C.: Quadratic sparse domination and weighted estimates for non-integral square functions. J. Geom. Anal. **33**(1), Paper No. 20, 49 (2023)
3. Buckley, S.M.: Estimates for operator norms on weighted spaces and reverse Jensen inequalities. Trans. Am. Math. Soc. **340**(1), 253–272 (1993)
4. Coifman, R.R., Meyer, Y., Stein, E.M.: Some new function spaces and their applications to harmonic analysis. J. Funct. Anal. **62**(2), 304–335 (1985)
5. Daubechies, I.: Ten lectures on wavelets, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1992)
6. Grau De La Herrán, A., Hofmann, S.: Generalized local Tb theorems for square functions. Mathematika **63**(1), 1–28 (2017)
7. Hytönen, T., Rosén, A.: On the Carleson duality. Ark. Mat. **51**(2), 293–313 (2013)
8. Hytönen, T., Rosén, A.: Bounded variation approximation of L_p dyadic martingales and solutions to elliptic equations. J. Eur. Math. Soc. (JEMS) **20**(8), 1819–1850 (2018)
9. Hytönen, T., Rosén, A.: Causal sparse domination of Beurling maximal regularity operators. J. Anal. Math. **150**(2), 645–672 (2023)
10. Hytönen, T.P.: The sharp weighted bound for general Calderón–Zygmund operators. Ann. Math. (2) **175**(3), 1473–1506 (2012)
11. Lerner, A.K.: Sharp weighted norm inequalities for Littlewood–Paley operators and singular integrals. Adv. Math. **226**(5), 3912–3926 (2011)
12. Lerner, A.K.: On pointwise estimates involving sparse operators. N. Y. J. Math. **22**, 341–349 (2016)
13. Lerner, A.K., Nazarov, F.: Intuitive dyadic calculus: the basics. Expos. Math. **37**(3), 225–265 (2019)
14. Rosén, A.: Geometric multivector analysis. Birkhäuser Advanced Texts: Basler Lehrbücher. Birkhäuser Advanced Texts: Basel Textbooks. Birkhäuser/Springer, Cham (2019) ©2019. From Grassmann to Dirac
15. Semmes, S.: Square function estimates and the $T(b)$ theorem. Proc. Am. Math. Soc. **110**(3), 721–726 (1990)