J. OPERATOR THEORY
44(2000), 63–90

BILINEAR OPERATORS ON HOMOGENEOUS GROUPS

LOUKAS GRAFAKOS and XINWEI LI

Communicated by Norberto Salinas

Abstract. Let H^p denote the Lebesgue space L^p for $p > 1$ and the Hardy space H^p for $p \leqslant 1$. For $0 < p, q, r < \infty$, we study $H^p \times H^q \to H^r$ mapping properties of bilinear operators given by finite sums of products of Calderón-Zygmund operators on stratified homogeneous Lie groups. When $r \leqslant 1$, we show that such mapping properties hold when a number of moments of the operator vanish. This hypothesis is natural and the conditions imposed are the minimal required for any operator of this type to map into the space H^r. Our proofs employ both the maximal function and atomic characterization of H^p. We also discuss some applications.

Keywords: *Biliniar operators, homogeneous groups, Hardy spaces.*
MSC (2000): 42B30.

0. INTRODUCTION

The study of multilinear operators is not motivated by a mere quest to generalize the theory of linear operators, but by their wide applicability and usability in analysis. The relation between the Cauchy integral along Lipschitz curves and the Calderón commutators is an example of this situation: the boundedness of the (bi)linear commutators is clearly connected to that of the Cauchy integral. Multilinear operators have also proved to be very useful in other fields of mathematics such as partial differential equations. The need to invert some linear partial differential operators occasionally leads to the study of multilinear singular integrals. The remarkable solution of the Korteweg-de Vries equation by the method of inverse scattering is a dramatic corroboration of this point of view.

In this article, we systematically study boundedness of bilinear operators given by sums of products of Calderón-Zygmund operators on stratified homogeneous groups. We are interested in mapping properties of these operators from $X \times Y$ to Z, where X, Y, and Z are Lebesgue spaces or Hardy spaces. We concentrate our attention on the case where Z is a Hardy space, otherwise the result is a trivial consequence of Hölder's inequality. We prove that boundedness into a

Hardy space holds exactly when a necessary number of moments of the operator vanishes. To avoid cumbersome notation we state our results for bilinear operators only; however, we note that our methods work for general multilinear operators of the same type.

1. NOTATION AND STATEMENT OF RESULTS

Let G be a stratified homogeneous Lie group with ambient space $\mathbb{R}^n$ and group dilations $\{\delta_r\}_{r>0}$. Then for some $1 = d_1 \leqslant d_2 \leqslant \cdots \leqslant d_n$,

$$\delta_r x = (r^{d_1}x_1, \ldots, r^{d_n}x_n) \quad \text{for all} \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n \text{ and } r > 0.$$

The number $D = d_1 + \cdots + d_n$ is called the homogeneous dimension of G. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $d(\alpha) = d_1\alpha_1 + \cdots + d_n\alpha_n$. Let $P(x) = \sum\limits_{\alpha} a_\alpha x^\alpha = \sum\limits_{\alpha} a_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ be a polynomial on G. The largest $d(\alpha)$ with nonzero coefficient a_α is called the homogeneous degree of $P(x)$. Let $\rho(x)$ be a C_0^∞ (away from the origin) homogeneous norm on G. Then, there exists a positive constant c such that for all $x, y \in G$,

$$\rho(xy) \leqslant c(\rho(x) + \rho(y)) \quad \text{and} \quad \rho(x^{-1}) = \rho(x). \tag{1.1}$$

Note that by [8], we can always choose an equivalent norm which has the subadditivity property with constant $c = 1$. In the rest of this paper we will assume that the constant c in (1.1) is equal to 1. All balls below will be left balls; that is, sets $Q = \{x : \rho(c_Q^{-1}x) < R\}$, where R is the radius of Q and c_Q is its center. We denote by $\mathrm{d}x$ Haar measure on G, normalized so that the measure of the unit ball $\{x : \rho(x) < 1\}$ is 1. Under this normalization, the Haar measure $|Q|$ of the ball Q is R^D and therefore $Q = \{x : \rho(c_Q^{-1}x) < |Q|^{\frac{1}{D}}\}$. For $a > 1$, aQ will be the set $Q = \{x : \rho(c_Q^{-1}x) < a|Q|^{\frac{1}{D}}\}$.

Let $\{X_1, \ldots, X_n\}$ be a basis for the space of left-invariant vector fields on G. For $\alpha = (\alpha_1, \ldots, \alpha_n)$, write $X^\alpha = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. Similarly, let $\{Y_1, \ldots, Y_n\}$ be a basis for the space of right-invariant vector fields on G, and define Y^α likewise. Unless stated otherwise, all our Taylor expansions will be based on left-invariant vector fields as in [6].

For $0 < r \leqslant 1$, the Hardy space $H^r(G)$ is defined to be the set of all distributions f on G for which the maximal function $\sup\limits_{t>0} |f * \varphi_t(x)|$ is in $L^r(G)$. Here $*$ is the group convolution on G, $\varphi_t = \frac{1}{t^D}(\varphi \circ \delta_{\frac{1}{t}})$, and φ is a Schwartz function which is also a commutative approximate identity on G; that is, $\int\limits_G \varphi \, \mathrm{d}x = 1$ and $\varphi_t * \varphi_s = \varphi_s * \varphi_t$ for all $s, t > 0$. Note that the definition of $H^r(G)$ is independent of the function φ. It was communicated to us by J. Dziubański that every stratified homogeneous group admits a compactly supported commutative approximate identity. See the appendix at the end of this paper for a proof of this fact. In the sequel, we will work with such an approximate identity.

The Hardy space $H^r(G)$ can also be characterized by its atomic decomposition. Every element f in $H^r(G)$ can be written as

$$f = \sum_Q \lambda_Q a_Q, \tag{1.2}$$

where Q is a ball, $\lambda_Q > 0$, and a_Q is an atom, i.e., a compactly supported bounded function with support in Q which satisfies:

(i) $|a_Q(x)| \leqslant |Q|^{-\frac{1}{r}}$ and
(ii) $\int a_Q(x)P(x)\,\mathrm{d}x = 0$

for all polynomials $P(x)$ of homogeneous degree not exceeding a fixed integer N_1 with $N_1 \geqslant \left[D\left(\frac{1}{r}-1\right)\right]$. The number N_1 can be taken arbitrarily large.

Consider a doubly indexed family of Calderón-Zygmund singular integral operators $\left\{T_i^j\right\}_{\substack{j=1,2\\ i=1,2,\ldots,N}}$ on G. The T_i^j's are given by $T_i^j f = f * K_i^j$, where $*$ is the convolution on G, and K_i^j are standard Calderón-Zygmund distribution kernels. We assume that there exists a large enough positive integer M and a constant A such that for all i, j the following hold:

(i) the T_i^j's are L^2-bounded, that is

$$\left\|T_i^j f\right\|_{L^2} \leqslant A\|f\|_{L^2},$$

for all f in a suitable dense subset of $L^2(G)$;

(ii) for all multi-indices α with $d(\alpha) \leqslant M$,

$$\left|\left(X^\alpha K_i^j\right)(x)\right| \leqslant A\rho(x)^{-D-d(\alpha)}.$$

REMARK 1.1. Condition (ii) on G is equivalent to condition (ii)$'$ below:
(ii)$'$ for all multi-indices α with $d(\alpha) \leqslant M$,

$$\left|\left(Y^\alpha K_i^j\right)(x)\right| \leqslant A\rho(x)^{-D-d(\alpha)}.$$

This is a consequence of Proposition 1.29 in [6], which states that Y^α can be written as a sum of homogeneous polynomials of degree $d(\beta) - d(\alpha)$ times X^β.

REMARK 1.2. Direct consequences of (ii) and (ii)$'$ are:
(iii) for all α with $d(\alpha) \leqslant M-1$,

$$\left|\left(X^\alpha K_i^j\right)(xy) - \left(X^\alpha K_i^j\right)(x)\right| \leqslant A\frac{\rho(y)}{\rho(x)^{D+d(\alpha)+1}} \quad \text{whenever } \rho(x) \geqslant 2\rho(y);$$

(iii)$'$ for all α with $d(\alpha) \leqslant M-1$,

$$\left|\left(Y^\alpha K_i^j\right)(yx) - \left(Y^\alpha K_i^j\right)(x)\right| \leqslant A\frac{\rho(y)}{\rho(x)^{D+d(\alpha)+1}} \quad \text{whenever } \rho(x) \geqslant 2\rho(y).$$

REMARK 1.3. We do not need to assume that K_i^j is homogeneous of degree $-D$. This property is essentially contained in (ii). However, in most interesting examples, we have this homogeneity.

REMARK 1.4. Singular integral operators with kernels satisfying (i) and (ii) can be extended to bounded operators from $L^p(G)$ to $L^p(G)$ for $1 < p < \infty$, and those with kernels satisfying (ii) and (iii) can be extended to bounded operators from $H^r(G)$ to $H^r(G)$ for $0 < r \leqslant 1$. See [6] and [12] for details.

Let $H^p(G) = L^p(G)$ for $p > 1$. We would like to study $H^p(G) \times H^q(G) \to H^r(G)$ mapping properties of bilinear operators of the form

$$B(f,g)(x) = \sum_{i=1}^{N} (T_i^1 f)(x)(T_i^2 g)(x),$$

where $0 < p, q, r < \infty$. Such boundedness properties can hold only when $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Since the case $r > 1$ is trivial, we consider below the following three situations:

(a) $p, q > 1$, $r \leqslant 1$,
(b) $p > 1$, $q \leqslant 1$, $r \leqslant 1$, and
(c) $p, q \leqslant 1$, $r \leqslant 1$.

Below we shall write $\sum\limits_i$ for the $\sum\limits_{i=1}^{N}$. In the sequel, all constants will depend on N.

THEOREM 1.5 *Let $p, q > 1$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Assume that there exists a nonnegative integer $k \leqslant D - 1$ such that*

$$\int x^\alpha B(f,g)(x)\, \mathrm{d}x = 0, \tag{1.3}$$

for all multi-indices α with $d(\alpha) \leqslant k$ and all $f, g \in L^2(G)$ with compact support. Then, for $\frac{D}{D+k+1} < r \leqslant 1$, B can be extended to a bounded operator from $L^p(G) \times L^q(G)$ into $H^r(G)$.

THEOREM 1.6. *Let $0 < p \leqslant 1$, $q > 1$, and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Assume that there exists a nonnegative integer k such that*

$$\int x^\alpha B(f,g)(x)\, \mathrm{d}x = 0, \tag{1.4}$$

for all multi-indices α with $d(\alpha) \leqslant k$, all f being p-atoms and $g \in L^2(G)$ with compact support. Then, for $\frac{D}{D+k+1} < r \leqslant 1$, B can be extended to a bounded operator from $H^p(G) \times L^q(G)$ into $H^r(G)$.

THEOREM 1.7. *Let $0 < p, q, r \leqslant 1$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Assume that for some integer $k \geqslant \left[D\left(\frac{1}{p} - 1\right)\right] + \left[D\left(\frac{1}{q} - 1\right)\right] + D + 2$ we have*

$$\int x^\alpha B(f,g)(x)\, \mathrm{d}x = 0 \tag{1.5}$$

for all multi-indices α with $d(\alpha) \leqslant k$ and all f being p-atoms and g being q-atoms. Then B can be extended to a bounded operator from $H^p(G) \times H^q(G)$ into $H^r(G)$.

REMARK 1.8. Note that the hypotheses of Theorem 1.7 imply that $\frac{D}{D+k+1} < r \leqslant 1$.

Remark 1.9. Having k vanishing moments is a necessary requirement for $B(f,g)$ to belong to H^r for $r > \frac{D}{D+k+1}$.

Remark 1.10. It is easy to see that the integrals in conditions (1.3), (1.4) and (1.5) are well-defined. Moreover, by the Campbell-Hausdorff formula, (1.5) is equivalent to either one of the following two conditions:

$$(1.5)' \qquad \int (yx)^\alpha B(f,g)(x)\,\mathrm{d}x = 0 \quad \forall\, y \in G,$$

$$(1.5)'' \qquad \int (xy)^\alpha B(f,g)(x)\,\mathrm{d}x = 0 \quad \forall\, y \in G.$$

Remark 1.11. Theorems 1.5 and 1.7 generalize the results in [1] and [7] to the context of stratified homogeneous groups; however, the approach taken in the proof of Theorem 1.7 is different. Theorem 1.6 fills in the missing link between Theorems 1.5 and 1.7 and does not seem to have appeared in the literature before.

Although we discuss bilinear operators only, our methods can be adapted to multilinear operators of this kind as well.

We would like to thank J. Dziubański for communicating to us the existence of a smooth compactly supported commutative approximate identity on stratified homogeneous groups. We are also grateful to Professors B. Blank, M. Christ, G. Folland, Y. Han, A. Hulanicki, M. Taibleson, and G. Weiss for helpful discussion and suggestions.

2. THE PROOF OF THEOREM 1.5

Let φ be a smooth compactly supported commutative approximate identity on G as in Section 1. Without loss of generality, we may assume that $\operatorname{supp}\varphi \subseteq \{x : \rho(x) \leqslant 1\}$. For $x_0 \in G$, let $\varphi_{t,x_0}(x) = t^{-D}\varphi(\delta_{\frac{1}{t}}(x^{-1}x_0)), x \in G$. We need to show that the function

$$x_0 \to \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) B(f,g)(x)\,\mathrm{d}x \Big|$$

is in $L^r(G)$. Fix a function $\eta(x)$ in $C^\infty(G)$ satisfying $\eta \equiv 1$ on $\{x : \rho(x) < 2\}$, and $\operatorname{supp}\eta \subseteq \{x : \rho(x) < 4\}$. Let $\eta_0(x) = \eta\big(\delta_{\frac{1}{t}}(x^{-1}x_0)\big)$ and $\eta_1(x) = 1 - \eta_0(x)$. Also fix $f, g \in L^2(G)$ with compact support. Split the operator B as follows:

$$(2.1) \qquad B(f,g) = B(\eta_0 f, \eta_0 g) + B(f, \eta_1 g) + B(\eta_1 f, g) - B(\eta_1 f, \eta_1 g).$$

Let us consider the term $B(f, \eta_1 g)$ first. For $\rho(x^{-1}x_0) \leqslant t$, we have

$$\begin{aligned}
&\sup_{t>0} |T_i^2(\eta_1 g)(x) - T_i^2(\eta_1 g)(x_0)| \\
(2.2) \qquad &= \sup_{t>0} \Big| \int\limits_{\rho(y^{-1}x_0)\geqslant 2t} \big(K_i^2(y^{-1}x) - K_i^2(y^{-1}x_0)\big)\eta_1(y)g(y)\,\mathrm{d}y \Big| \\
&\leqslant c \sup_{t>0} \int\limits_{\rho(y^{-1}x_0)\geqslant 2t} \rho(x_0^{-1}x) \frac{|\eta_1(y)g(y)|}{\rho(y^{-1}x_0)^{D+1}}\,\mathrm{d}y \qquad \text{by condition (iii)} \\
&\leqslant cM(\eta_1 g)(x_0),
\end{aligned}$$

where M is the Hardy-Littlewood maximal operator on G, and c is an absolute positive constant. Throughout this article, $c > 0$ will denote a constant whose value may vary. Thus

$$\begin{aligned}
&\sup_{t>0}\Big|\int \varphi_{t,x_0}(x)B(f,\eta_1 g)(x)\,\mathrm{d}x\Big| \\
&\qquad\leqslant \sum_i \sup_{t>0}\int |\varphi_{t,x_0}(x)|\,\big|T_i^1 f(x)\big|\,\big|T_i^2(\eta_1 g)(x)-T_i^2(\eta_1 g)(x_0)\big|\,\mathrm{d}x \\
&\qquad\quad+\sum_i \sup_{t>0}\int |\varphi_{t,x_0}(x)|\,\big|T_i^1 f(x)\big|\,\big|T_i^2(\eta_1 g)(x_0)\big|\,\mathrm{d}x \\
&\qquad\leqslant c\sum_i M\big(T_i^1 f\big)(x_0)M(\eta_1 g)(x_0)+c\sum_i M\big(T_i^1 f\big)(x_0)\big|T_i^2(\eta_1 g)(x_0)\big|.
\end{aligned}\tag{2.3}$$

Since $\frac{1}{r}=\frac{1}{p}+\frac{1}{q}$, Hölder's inequality gives

$$\begin{aligned}
&\int \sup_{t>0}\Big|\int \varphi_{t,x_0}(x)B(f,\eta_1 g)(x)\,\mathrm{d}x\Big|^r\,\mathrm{d}x_0 \\
&\qquad\leqslant c\sum_i \big\|M\big(T_i^1 f\big)\big\|_p^r\big(\|M(\eta_1 g)\|_q^r+\big\|T_i^2(\eta_1 g)\big\|_q^r\big)\leqslant c\|f\|_p^r\,\|g\|_q^r.
\end{aligned}\tag{2.4}$$

The estimate for term $B(\eta_1 f, g)$ is similar and is omitted. We now consider the last term in (2.1). Write $B(\eta_1 f,\eta_1 g)=\Sigma_1+\Sigma_2+\Sigma_3+\Sigma_4$, where

$$\begin{aligned}
\Sigma_1(x)&=\sum_i\big(T_i^1(\eta_1 f)(x)-T_i^1(\eta_1 f)(x_0)\big)\big(T_i^2(\eta_1 g)(x)-T_i^2(\eta_1 g)(x_0)\big),\\
\Sigma_2(x)&=\sum_i T_i^1(\eta_1 f)(x)T_i^2(\eta_1 g)(x_0),\\
\Sigma_3(x)&=\sum_i T_i^1(\eta_1 f)(x_0)T_i^2(\eta_1 g)(x),\\
\Sigma_4(x)&=-\sum_i T_i^1(\eta_1 f)(x_0)T_i^2(\eta_1 g)(x_0).
\end{aligned}\tag{2.5}$$

By (2.2), $\Sigma_1(x)\leqslant cM(\eta_1 f)(x_0)M(\eta_1 g)(x_0)$ for all x in the support of φ_{t,x_0}. Hölder's inequality now gives the desired estimate for Σ_1.

Terms Σ_2 and Σ_3 are similar to the second sum of (2.3) and are estimated likewise. The estimate for Σ_4 is easier and is omitted. We conclude that

$$\int \sup_{t>0}\Big|\int \varphi_{t,x_0}(x)B(\eta_1 f,\eta_1 g)(x)\,\mathrm{d}x\Big|^r\,\mathrm{d}x_0\leqslant c\|f\|_p^r\,\|g\|_q^r.\tag{2.6}$$

Now we consider the main term $B(\eta_0 f,\eta_0 g)$. Hypothesis (1.3) implies that

$$\sum_i\Big(T_i^1\Big)^*\big(P_k T_i^2(\eta_0 g)\big)=0\quad\text{a.e. on } G,\tag{2.7}$$

for all polynomials P_k of homogeneous degree less than or equal to k, where $(T_i^1)^*$ is the adjoint operator of T_i^1. Since $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}<\frac{D+k+1}{D}$ and $k+1\leqslant D$, we can choose $1<p_1<p$ and $1<q_1<q$, such that

$$\frac{1}{p_1}+\frac{1}{q_1}=\frac{D+k+\varepsilon}{D},$$

for some $0 < \varepsilon < 1$.

As in [5], let $\Gamma_{k+\varepsilon}$ be the homogeneous Lipschitz space of G, defined as the set of all continuous and bounded functions on G which satisfy

$$\|h\|_{\Gamma_{k+\varepsilon}} \equiv \sup_{u,v\in G} \sup_{d(\alpha)\leqslant k} \frac{|(X^\alpha h)(uv) - (X^\alpha h)(u)|}{\rho(v)^\varepsilon} < \infty.$$

Let $P_x^k(z) = \sum_{d(\alpha)\leqslant k} C_{\alpha,x_0}(x)(x^{-1}z)^\alpha$ be the Taylor polynomial of homogeneous degree k of the function $\varphi_{t,x_0}(\,\cdot\,)$ at the point x. Using Taylor's theorem and the easy fact that the C^∞ compactly supported function φ_{t,x_0} is in $\Gamma_{k+\varepsilon}$ with $\|\varphi_{t,x_0}\|_{\Gamma_{k+\varepsilon}} \leqslant ct^{-D-k-\varepsilon}$, we deduce that

$$\big|\varphi_{t,x_0}(y) - P_x^k(x^{-1}y)\big| \leqslant ct^{-D-k-\varepsilon}\rho(x^{-1}y)^{k+\varepsilon} \tag{2.8}$$

for some constant $c > 0$ depending only on φ, k and ε.

Using the Campbell-Hausdorff formula, we can write the polynomial $y \to P_x^k(x^{-1}y)$ as a sum of powers of y with coefficients in x. Applying (2.7) to this polynomial we obtain

$$\sum_i \big(T_i^1\big)^*\big[P_x^k(x^{-1}\,\cdot\,)T_i^2(\eta_0 g)(\,\cdot\,)\big] = 0. \tag{2.9}$$

Then we have

$$\begin{aligned}
&\sup_{t>0}\Big|\int\varphi_{t,x_0}B(\eta_0 f,\eta_0 g)(x)\,\mathrm{d}x\Big| = \sup_{t>0}\Big|\int\varphi_{t,x_0}\sum_i T_i^1(\eta_0 f)(x)T_i^2(\eta_0 g)(x)\,\mathrm{d}x\Big| \\
&= \sup_{t>0}\Big|\int\sum_i(\eta_0 f)(x)\big(T_i^1\big)^*\big(\varphi_{t,x_0}T_i^2(\eta_0 g)\big)(x)\,\mathrm{d}x\Big| \\
&= \sup_{t>0}\Big|\int\sum_i(\eta_0 f)(x)\big[\big(T_i^1\big)^*\big[\big(\varphi_{t,x_0}(\,\cdot\,)-P_x^k(x^{-1}\,\cdot\,)\big)T_i^2(\eta_0 g)(\,\cdot\,)\big](x)\big]\,\mathrm{d}x\Big|,
\end{aligned} \tag{2.10}$$

where we used (2.9) in the last inequality above. Let $\frac{1}{p_1} + \frac{1}{p_1'} = 1$. We claim that

$$\big\|\big(T_i^1\big)^*\big[(\varphi_{t,x_0}(\,\cdot\,) - P_x^k(x^{-1}\,\cdot\,))T_i^2(\eta_0 g)(\,\cdot\,)\big]\big\|_{p_1'} \leqslant ct^{-D-k-\varepsilon}\|\eta_0 g\|_{q_1},$$

where $q_1 > 1$ and $\frac{1}{p_1'} = \frac{1}{q_1} - \frac{k+\varepsilon}{D}$.

Assuming the claim, we control (2.10) by

$$\begin{aligned}
&\sum_i \sup_{t>0}\|\eta_0 f\|_{p_1}\big\|(T_i^1)^*\big[\big(\varphi_{t,x_0}(\,\cdot\,) - P_x^k(x^{-1}\,\cdot\,)\big)T_i^2(\eta_0 g)(\,\cdot\,)\big]\big\|_{p_1'} \\
&\qquad\leqslant c\sup_{t>0}\|\eta_0 f\|_{p_1}t^{-D-k-\varepsilon}\|\eta_0 g\|_{q_1} \\
&\qquad\leqslant c\sup_{t>0}t^{-D-k-\varepsilon}M(|f|^{p_1})^{\frac{1}{p_1}}(x_0)t^{\frac{D}{p_1}}M(|g|^{q_1})^{\frac{1}{q_1}}(x_0)t^{\frac{D}{q_1}} \\
&\qquad= cM(|f|^{p_1})^{\frac{1}{p_1}}(x_0)M(|g|^{q_1})^{\frac{1}{q_1}}(x_0),
\end{aligned}$$

since $\frac{D}{p_1} + \frac{D}{q_1} = D + k + \varepsilon$. Thus

$$\begin{aligned}\int \sup_{t>0} \Big| \int \varphi_{t,x_0} B(\eta_0 f, \eta_0 g)\,\mathrm{d}x \Big|^r \mathrm{d}x_0 &\leqslant c \int M(|f|^{p_1})^{\frac{r}{p_1}}(x_0) M(|g|^{q_1})^{\frac{r}{q_1}}(x_0)\,\mathrm{d}x_0 \\ &\leqslant c\Big(\int M(|f|^{p_1})^{\frac{p}{p_1}}(x_0)\,\mathrm{d}x_0\Big)^{\frac{r}{p}} \Big(\int M(|g|^{q_1})^{\frac{q}{q_1}}(x_0)\,\mathrm{d}x_0\Big)^{\frac{r}{q}} \leqslant c\|f\|_p^r\,\|g\|_q^r\end{aligned}$$

since $p > p_1$ and $q > q_1$.

To complete our proof we need to prove the claim. Let $(K_i^1)^*$ be the kernel of $(T_i^1)^*$. Using (2.8) we obtain

$$\begin{aligned}&\big|(T_i^1)^*\big[\big(\varphi_{t,x_0}(\,\cdot\,) - P_x^k(x^{-1}\,\cdot\,)\big)T_i^2(\eta_0 g)(\,\cdot\,)\big]\big| \\ &\qquad\leqslant c\int \big|\big(K_i^1\big)^*(y^{-1}x)\big| t^{-D-k-\varepsilon}\rho(x^{-1}y)^{k+\varepsilon}\big|T_i^2(\eta_0 g)(y)\big|\,\mathrm{d}y \\ &\qquad\leqslant ct^{-D-k-\varepsilon}\int \frac{\big|T_i^2(\eta_0 g)(y)\big|}{\rho(y^{-1}x)^{D-(k+\varepsilon)}}\,\mathrm{d}y \qquad\qquad \text{by condition (ii).}\end{aligned}$$

This last estimate, together with the fractional integral theorem on homogeneous groups ([6]), and the boundedness of Calderón-Zygmund operators give

$$\begin{aligned}\big|(T_i^1)^*\big[(\varphi_{t,x_0}(\,\cdot\,) - P_x^k(x^{-1}\,\cdot\,)\big)T_i^2(\eta_0 g)(\,\cdot\,)\big]\big\|_{p_1'} &\leqslant ct^{-D-k-\varepsilon}\big\|T_i^2(\eta_0 g)\big\|_{q_1} \\ &\leqslant ct^{-D-k-\varepsilon}\|\eta_0 g\|_{q_1},\end{aligned}$$

where $q_1 > 1$ and $\frac{1}{p_1'} = \frac{1}{q_1} - \frac{k+\varepsilon}{D}$.

This finishes the proof of the claim and thus of Theorem 1.5 ∎

3. THE PROOF OF THEOREM 1.6

Fix $0 < p \leqslant 1$, $f \in H^p(G)$, and $g \in L^2(G)$ with compact support. Let f have a decomposition as in (1.2). We will show that for any atom a_Q that appears in the decomposition of f, the following holds:

$$\int \sup_{t>0} \Big| \int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)(T_i^2 g)\,\mathrm{d}x \Big|^r \mathrm{d}x_0 \leqslant c\|g\|_q^r. \tag{3.1}$$

Then, summing (3.1) over all such atoms will give the conclusion. Let

$$I(x_0) = \sup_{t>0} \Big| \int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 g\big)\,\mathrm{d}x \Big|^r.$$

We will use the notation $c_Q' = c_Q^{-1}x_0$. Pick a fixed l with $1 < l < q$.

Case 1. $x_0 \in 3Q$. In this case

$$\begin{aligned}
\int\limits_{3Q} I(x_0)\,\mathrm{d}x_0 &\leqslant \sum_i \int\limits_{3Q} \sup_{t>0} \Big(\int |\varphi_{t,x_0}| \big|T_i^1 a_Q T_i^2 g\big|\,\mathrm{d}x\Big)^r \mathrm{d}x_0 \\
&\leqslant c \sum_i \int\limits_{3Q} M\big(\big|T_i^1 a_Q T_i^2 g\big|\big)^r \mathrm{d}x_0 \\
&\leqslant c \sum_i \Big(\int\limits_{G} M\big(\big|T_i^1 a_Q T_i^2 g\big|\big)^l \mathrm{d}x_0\Big)^{\frac{r}{l}} \Big(\int\limits_{3Q} 1\,\mathrm{d}x_0\Big)^{\frac{l-r}{l}} \\
&\leqslant c|Q|^{\frac{l-r}{l}} \sum_i \Big(\int \big|T_i^1 a_Q T_i^2 g\big|^l \mathrm{d}x\Big)^{\frac{r}{l}} \\
&\leqslant c|Q|^{\frac{l-r}{l}} \sum_i \Big(\int |T_i^1 a_Q|^{\frac{lq}{q-l}} \mathrm{d}x\Big)^{\frac{q-l}{q}\frac{r}{l}} \Big(\int \big|T_i^2 g\big|^q \mathrm{d}x\Big)^{\frac{r}{q}} \\
&\leqslant c|Q|^{\frac{l-r}{l}} \Big(\int |a_Q|^{\frac{lq}{q-l}} \mathrm{d}x\Big)^{\frac{q-l}{q}\frac{r}{l}} \|g\|_q^r \qquad\qquad \text{since } \frac{q}{l} > 1 \\
&\leqslant c\|g\|_q^r.
\end{aligned}$$

Case 2. $x_0 \notin 3Q$. We have

$$\begin{aligned}
I(x_0) \leqslant \sup_{0<t\leqslant \frac12 \rho(c_Q')} &\Big| \int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 g\big)\,\mathrm{d}x\Big|^r \\
&+ \sup_{t>\frac12 \rho(c_Q')} \Big| \int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 g\big)\,\mathrm{d}x\Big|^r = I_1(x_0) + I_2(x_0).
\end{aligned}$$

We now have $I_1(x_0) \leqslant I_{11}(x_0) + I_{12}(x_0)$, where

$$\begin{aligned}
I_{11}(x_0) &= \sup_{0<t\leqslant \frac12 \rho(c_Q')} \Big| \int\limits_{2Q} \varphi_{t,x_0}(x) \sum_i \big(T_i^1 a_Q\big)(x)\big(T_i^2 g\big)(x)\,\mathrm{d}x\Big|^r \\
I_{12}(x_0) &= \sup_{0<t\leqslant \frac12 \rho(c_Q')} \Big| \int\limits_{(2Q)^c} \varphi_{t,x_0}(x) \sum_i \big(T_i^1 a_Q\big)(x)\big(T_i^2 g\big)(x)\,\mathrm{d}x\Big|^r.
\end{aligned}$$

Consider $I_{11}(x_0)$ first. As before, let $P^k_{c_Q}(y)$ be the Taylor polynomial of homogeneous degree k of $\varphi_{t,x_0}(\,\cdot\,)$ at the point c_Q. We claim that in this case, $P^k_{c_Q}(\,\cdot\,) \equiv 0$, since for $t \leqslant \frac12 \rho(c_Q')$, φ_{t,x_0} is identically equal to zero near c_Q. Therefore $P^k_{c_Q}(c_Q^{-1}x) = 0$ for $x \in 2Q$. By Taylor's Theorem we have

$$\begin{aligned}
\big|\varphi_{t,x_0}(x) - P^k_{c_Q}(c_Q^{-1}x)\big| &\leqslant c \frac{1}{t^{D+k+1}} \rho(c_Q^{-1}x)^{k+1} \chi_{\rho(x^{-1}x_0)\leqslant t} \\
&\leqslant c \frac{|Q|^{\frac{k+1}{D}}}{\rho(x^{-1}x_0)^{D+k+1}} \qquad\qquad \text{since } x \in 2Q \\
&\leqslant c \frac{|Q|^{\frac{k+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+k+1}} \qquad\qquad \text{by (1.1) since } x_0 \notin 3Q.
\end{aligned}$$

Replacing $\varphi_{t,x_0}(x)$ by $\varphi_{t,x_0}(x) - P^k_{c_Q}(c_Q^{-1}x)$ in the definition of $I_{11}(x)$ and using the estimate above, we obtain:

$$\begin{aligned} I_{11}(x_0) &\leqslant c\frac{|Q|^{\frac{r(k+1)}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}\sum_i\Big(\int\big|T_i^1 a_Q(x)\big|\,\big|T_i^2 g(x)\big|\,\mathrm{d}x\Big)^r \\ &\leqslant c\frac{|Q|^{\frac{r(k+1)}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}\sum_i\big\|T_i^1 a_Q\big\|^r_{q'}\|T_i^2 g\|^r_q, \quad \text{where } \frac{1}{q'}+\frac{1}{q}=1 \qquad (3.2)\\ &\leqslant c\frac{|Q|^{\frac{r(k+1)}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}\|a_Q\|^r_{q'}\|g\|^r_q \leqslant c\|g\|^r_q\frac{|Q|^{-1+\frac{D+k+1}{D}r}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}. \end{aligned}$$

Integrating the above over $(3Q)^c$ and using that

$$\int\limits_{(3Q)^c}\frac{|Q|^{-1+\frac{D+k+1}{D}r}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}\,\mathrm{d}x_0 \leqslant c, \qquad (3.3)$$

we obtain that for $r > \frac{D}{D+k+1}$

$$\int\limits_{(3Q)^c} I_{11}(x_0)\,\mathrm{d}x_0 \leqslant c\|g\|^r_q.$$

Now consider $I_{12}(x_0)$. Let N_1 be as in Section 1. For any nonnegative integer $m \leqslant N_1$, let $P_x(\,\cdot\,)$ be the right Taylor polynomial of $K_i^1(\,\cdot\,x)$ at x of homogeneous degree m. By (iii)$'$,

$$\big|K_i^1(y^{-1}x) - P_{c_Q^{-1}x}(y^{-1}c_Q)\big| \leqslant c\frac{\rho(y^{-1}c_Q)^{m+1}}{\rho\big(c_Q^{-1}x\big)^{m+1+D}}, \qquad (3.4)$$

whenever $\rho\big(c_Q^{-1}x\big) \geqslant 2\rho(y^{-1}c_Q)$. Now

$$\big|T_i^1 a_Q(x)\big| = \Big|\int\big(K_i^1(y^{-1}x) - P_{c_Q^{-1}x}(y^{-1}c_Q)\big)a_Q(y)\,\mathrm{d}y\Big|, \qquad (3.5)$$

by the cancellation property of a_Q and the Campbell-Hausdorff formula. Thus, using (3.4) and (3.5), we obtain the following pointwise estimate,

$$\big|T_i^1 a_Q(x)\big| \leqslant c\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho(c_Q^{-1}x)^{m+1+D}}, \quad x \notin 2Q, \qquad (3.6)$$

for all nonnegative integers $m \leqslant N_1$. We now have that $I_{12}(x_0)$ is bounded above by

$$\begin{aligned} &c\sum_i\sup_{0<t\leqslant\frac{1}{2}\rho(c'_Q)}\Big(\int|\varphi_{t,x_0}(x)|\big|T_i^1 a_Q(x)\big|^{l'}\chi_{(2Q)^c}\,\mathrm{d}x\Big)^{\frac{r}{l'}}\Big(\int|\varphi_{t,x_0}(x)|\,\big|T_i^2 g(x)\big|^l\,\mathrm{d}x\Big)^{\frac{r}{l}} \\ &\qquad\leqslant c\sum_i\sup_{0<t\leqslant\frac{1}{2}\rho(c'_Q)}\Big(\int|\varphi_{t,x_0}(x)|\,\big|T_i^1 a_Q(x)\big|^{l'}\chi_{(2Q)^c}\,\mathrm{d}x\Big)^{\frac{r}{l'}}M\big(\big|T_i^2 g(x)\big|^l\big)^{\frac{r}{l}}(x_0). \end{aligned}$$

Integrating the above on $x_0 \notin 3Q$ and applying Hölder's inequality, we obtain the following estimate for $\int\limits_{(3Q)^c} I_{12}(x_0)\,\mathrm{d}x_0$:

$$
\begin{aligned}
(3.7)\qquad & c\sum_i \Big(\int\limits_{(3Q)^c} \sup_{0<t\leqslant\frac{1}{2}\rho(c'_Q)} \Big(\int |\varphi_{t,x_0}(x)|\big|T_i^1 a_Q(x)\big|^{l'}\chi_{(2Q)^c}\,\mathrm{d}x\Big)^{\frac{p}{l'}}\,\mathrm{d}x_0\Big)^{\frac{r}{p}} \\
&\qquad\cdot\Big(\int\limits_{(3Q)^c} M\big(\big|T_i^2 g(x)\big|^l\big)^{\frac{q}{l}}(x_0)\,\mathrm{d}x_0\Big)^{\frac{r}{q}}.
\end{aligned}
$$

By the Hardy-Littlewood maximal theorem, the second factor in (3.7) is bounded by $c\|g\|_q^r$. By (3.6), the first factor in (3.7) satisfies

$$
\begin{aligned}
&\Big(\int\limits_{(3Q)^c} \sup_{0<t\leqslant\frac{1}{2}\rho(c'_Q)} \Big(\int |\varphi_{t,x_0}(x)||T_i^1 a_Q(x)|^{l'}\chi_{(2Q)^c}\,\mathrm{d}x\Big)^{\frac{p}{l'}}\,\mathrm{d}x_0\Big)^{\frac{r}{p}} \\
&\quad\leqslant c\Big(\int\limits_{(3Q)^c} \sup_{0<t\leqslant\frac{1}{2}\rho(c'_Q)} \Big(\int\limits_{(2Q)^c} |\varphi_{t,x_0}(x)|\frac{|Q|^{(-\frac{1}{p}+1+\frac{m+1}{D})l'}}{\rho(c_Q^{-1}x)^{(m+1+D)l'}}\,\mathrm{d}x\Big)^{\frac{p}{l'}}\,\mathrm{d}x_0\Big)^{\frac{r}{p}} \\
&\quad\leqslant c\Big(\int\limits_{(3Q)^c} \frac{|Q|^{(-\frac{1}{p}+1+\frac{m+1}{D})p}}{\rho(c_Q^{-1}x_0)^{(m+1+D)p}}\,\mathrm{d}x_0\Big)^{\frac{r}{p}} \\
&\quad\leqslant c\big(|Q|^{-1+p+\frac{m+1}{D}p}\,|Q|^{\frac{1}{D}(D-(m+1+D)p)}\big)^{\frac{r}{p}} = c,
\end{aligned}
$$

where we use that $\rho(c_Q^{-1}x) = \rho\big(c_Q^{-1}x_0x_0^{-1}x\big) \geqslant \rho\big(c_Q^{-1}x_0\big) - \rho\big(x_0^{-1}x\big) \geqslant \frac{1}{2}\rho(c_Q^{-1}x_0)$. We also picked an m in (3.6) with $m > \frac{D}{p} - D - 1$ to make the last integral convergent.

We have now proved (3.1) for $I_1(x_0)$. Next, we discuss $I_2(x_0)$. Note that by (1.4), Hölder's inequality, and the L^q boundedness of T_i^2, we have

$$
\begin{aligned}
I_2(x_0) &= \sup_{t>\frac{1}{2}\rho(c'_Q)} \Big|\int \big(\varphi_{t,x_0}(x) - P_{c_Q}^k(c_Q^{-1}x)\big)\sum_i \big(T_i^1 a_Q\big)(x)\big(T_i^2 g\big)(x)\,\mathrm{d}x\Big|^r \\
&\leqslant c\sum_i \sup_{t>\frac{1}{2}\rho(c'_Q)} \Big(\int \frac{\rho\big(c_Q^{-1}x\big)^{k+1}}{t^{D+k+1}}\big|T_i^1 a_Q(x)\big|\,\big|T_i^2 g(x)\big|\,\mathrm{d}x\Big)^r \\
(3.8)\qquad &\leqslant c\|g\|_q^r \sum_i \sup_{t>\frac{1}{2}\rho(c'_Q)} \Big(\int \frac{\rho\big(c_Q^{-1}x\big)^{(k+1)q'}}{t^{(D+k+1)q'}}\big|T_i^1 a_Q(x)\big|^{q'}\,\mathrm{d}x\Big)^{\frac{r}{q'}}.
\end{aligned}
$$

Since $\frac{r}{q'} < 1$, we have that the supremum in (3.8) is controlled by

$$
\sup_{t>\frac{1}{2}\rho(c'_Q)} \Big(\int\limits_{x\in 2Q} \mathrm{d}x\Big)^{\frac{r}{q'}} + \sup_{t>\frac{1}{2}\rho(c'_Q)} \Big(\int\limits_{x\notin 2Q} \mathrm{d}x\Big)^{\frac{r}{q'}} \equiv I_{21}(x_0) + I_{22}(x_0).
$$

It is easy to obtain that

$$
(3.9)\qquad I_{21}(x_0) \leqslant c\frac{|Q|^{\frac{r(k+1)}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}\|T_i^1 a_Q\|_{q'}^r \leqslant c\frac{|Q|^{-1+\frac{D+k+1}{D}r}}{\rho\big(c_Q^{-1}x_0\big)^{(D+k+1)r}}.
$$

For $I_{22}(x_0)$, we use (3.6) (taking $m = k$) to obtain

$$\begin{aligned} I_{22}(x_0) &\leqslant c \sup_{t > \frac{1}{2}\rho(c'_Q)} \Big(\int \frac{\rho(c_Q^{-1}x)^{(k+1)q'}}{t^{(D+k+1)q'}} \frac{|Q|^{(-\frac{1}{p}+1+\frac{k+1}{D})q'}}{\rho\big(c_Q^{-1}x\big)^{(k+1+D)q'}} \,\mathrm{d}x \Big)^{\frac{r}{q'}} \\ &\leqslant c \frac{|Q|^{(-\frac{1}{p}+1+\frac{k+1}{D})r}}{\rho\big(c_Q^{-1}x_0\big)^{(k+1+D)r}} \Big(\int_{x \notin 2Q} \frac{1}{\rho\big(c_Q^{-1}x\big)^{Dq'}} \,\mathrm{d}x \Big)^{\frac{r}{q'}} \\ &\leqslant c \frac{|Q|^{-1+\frac{D+k+1}{D}r}}{\rho\big(c_Q^{-1}x_0\big)^{(k+1+D)r}}. \end{aligned} \tag{3.10}$$

Combining (3.8), (3.9), and (3.10), we deduce

$$I_2(x_0) \leqslant c\|g\|_q^r \frac{|Q|^{-1+\frac{D+k+1}{D}r}}{\rho(c_Q^{-1}x_0)^{(D+k+1)r}}.$$

Integrating the above over the set $x_0 \notin 3Q$, and using (3.3), we obtain

$$\int_{(3Q)^c} I_2(x_0)\,\mathrm{d}x_0 \leqslant c\|g\|_q^r.$$

We have now finished the proof of (3.1) and we derive Theorem 1.6. Let $f = \sum \lambda_Q a_Q$ be a finite sum of atoms in H^p. We have

$$\begin{aligned} &\int \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) \sum_i (T_i^1 f)(x)(T_i^2 g)(x)\,\mathrm{d}x \Big|^r \mathrm{d}x_0 \\ &\qquad \leqslant \sum_Q \lambda_Q^r \int \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) \sum_i (T_i^1 a_Q)(x)(T_i^2 g)(x)\,\mathrm{d}x \Big|^r \mathrm{d}x_0 \\ &\qquad \leqslant c \sum_Q \lambda_Q^r \|g\|_q^r \quad \text{by (3.1).} \end{aligned}$$

Since $q > 1$, we claim that $\frac{p}{r} \leqslant 2$; otherwise, $\frac{p}{r} > 2$ which would imply that $r < \frac{p}{2} \leqslant \frac{1}{2}$ and $\frac{r}{q} < \frac{1}{2q} < \frac{1}{2}$. Hence $1 = \frac{1}{p/r} + \frac{1}{q/r} < \frac{1}{2} + \frac{1}{2} = 1$ which gives a contradiction. Therefore $\frac{p}{r} \leqslant 2$ and we have that $\Big(\sum_Q \lambda_Q^r\Big)^{\frac{1}{r}} \leqslant c\Big(\sum_Q \lambda_Q^p\Big)^{\frac{1}{p}} = c\|f\|_{H^p}$.

To see this last inequality, we assume that $\sum_Q \lambda_Q^r \geqslant 1$ and that each $\lambda_Q \leqslant 1$. Then

$$\Big(\sum_Q \lambda_Q^r\Big)^{\frac{p}{r}} \leqslant \Big(\sum_Q \lambda_Q^r\Big)^2 \leqslant 2\sum_Q \lambda_Q^{2r} \leqslant 2\sum_Q \lambda_Q^p.$$

We have now shown that

$$\Big(\int \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) \sum_i \big(T_i^1 f\big)(x)\big(T_i^2 g\big)(x)\,\mathrm{d}x \Big|^r \mathrm{d}x_0 \Big)^{\frac{1}{r}} \leqslant c\|f\|_{H^p}\|g\|_{L^q},$$

and the proof of Theorem 1.6 is complete.

3. THE PROOF OF THEOREM 1.7

We are now going to prove that if B satisfies the hypotheses of Theorem 1.7, the L^r norm of the function

$$x_0 \to \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) B(f,g)(x)\,\mathrm{d}x \Big|$$

is controlled by $c\|f\|_{H^p}\|g\|_{H^q}$. Write $f = \sum \lambda_Q a_Q$, $g = \sum \mu_R b_R$, where $\lambda_Q, \mu_R > 0$, a_Q are p-atoms, and b_R are q-atoms. Denote

$$S(a_Q, b_R)(x_0) = \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) B(a_Q, b_R)(x)\,\mathrm{d}x \Big|.$$

Then

$$\begin{aligned} \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) B(f,g)(x)\,\mathrm{d}x \Big| &\leqslant \sum_{Q,R} \lambda_Q \mu_R S(a_Q, b_R)(x_0) \\ &\leqslant \Sigma_1(x_0) + \Sigma_2(x_0) + \Sigma_3(x_0) + \Sigma_4(x_0), \end{aligned}$$

where the $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$ are defined below.

$$\Sigma_1(x_0) = \sum_{\substack{Q,R \\ x_0 \in 5Q \\ x_0 \in 5R}} \lambda_Q \mu_R S(a_Q, b_R)(x_0), \qquad \Sigma_2(x_0) = \sum_{\substack{Q,R \\ x_0 \in 5Q \\ x_0 \notin 5R}} \lambda_Q \mu_R S(a_Q, b_R)(x_0)$$

$$\Sigma_3(x_0) = \sum_{\substack{Q,R \\ x_0 \notin 5Q \\ x_0 \in 5R}} \lambda_Q \mu_R S(a_Q, b_R)(x_0), \qquad \Sigma_4(x_0) = \sum_{\substack{Q,R \\ x_0 \notin 5Q \\ x_0 \notin 5R}} \lambda_Q \mu_R S(a_Q, b_R)(x_0).$$

It suffices to show that for each j, $\Sigma_j \in L^r(G)$, and $\|\Sigma_j\|_{L^r} \leqslant c\|f\|_{H^p}\|g\|_{H^q}$.

Case 1. $x_0 \in 5Q, x_0 \in 5R$. In this case we have

$$\begin{aligned} S(a_Q, b_R)(x_0) &\leqslant \sum_i \sup_{t>0} \Big| \int \varphi_{t,x_0}(x) \big(T_i^1 a_Q\big)(x) \big(T_i^2 b_R\big)(x)\,\mathrm{d}x \Big| \\ &\leqslant \sum_i \sup_{t>0} \Big(\int |\varphi_{t,x_0}(x)| \big|T_i^1 a_Q(x)\big|^2\,\mathrm{d}x \Big)^{\frac12} \Big(\int |\varphi_{t,x_0}(x)| \big|T_i^2 b_R(x)\big|^2\,\mathrm{d}x \Big)^{\frac12} \\ &\leqslant c \sum_i M\big(\big|T_i^1 a_Q\big|^2\big)^{\frac12}(x_0) M\big(\big|T_i^2 b_R\big|^2\big)^{\frac12}(x_0). \end{aligned}$$

Therefore

$$\begin{aligned} \int \Big(\sum\nolimits_1 (x_0)\Big)^r \mathrm{d}x_0 &\leqslant \int \Big(\sum_{\substack{Q,R \\ x_0 \in 5Q \cap 5R}} \lambda_Q \mu_R S(a_Q, b_R)(x_0) \Big)^r \mathrm{d}x_0 \\ &\leqslant c \int \Big(\sum_i \sum_{\substack{Q,R \\ x_0 \in 5Q \cap 5R}} \lambda_Q \mu_R M\big(\big|T_i^1 a_Q\big|^2\big)^{\frac12} M\big(\big|T_i^2 b_R\big|^2\big)^{\frac12} \Big)^r \mathrm{d}x_0 \\ &\leqslant c \sum_i \int \Big(\sum_{\substack{Q \\ x_0 \in 5Q}} \lambda_Q M\big(\big|T_i^1 a_Q\big|^2\big)^{\frac12} \Big)^r \Big(\sum_{\substack{R \\ x_0 \in 5R}} \mu_R M\big(\big|T_i^2 b_R\big|^2\big)^{\frac12} \Big)^r \mathrm{d}x_0 \end{aligned}$$

$$\leqslant c\sum_i\Big(\int\Big(\sum_{\substack{Q\\ x_0\in 5Q}}\lambda_Q M\big(|T_i^1a_Q|^2\big)^{\frac12}\Big)^p\,\mathrm{d}x_0\Big)^{\frac rp}$$

$$\cdot\Big(\int\Big(\sum_{\substack{R\\ x_0\in 5R}}\mu_R M\big(\big|T_i^2b_R\big|^2\big)^{\frac12}\Big)^q\,\mathrm{d}x_0\Big)^{\frac rq}$$

$$\leqslant c\sum_i\Big(\sum_Q\lambda_Q^p\int_{5Q}M\big(\big|T_i^1a_Q\big|^2\big)^{\frac p2}(x_0)\,\mathrm{d}x_0\Big)^{\frac rp}$$

$$\cdot\Big(\sum_R\mu_R^q\int_{5R}M\big(\big|T_i^2b_R\big|^2\big)^{\frac q2}(x_0)\,\mathrm{d}x_0\Big)^{\frac rq}.$$

It is a simple fact ([12], Chapter 1, 8.15) that

$$\begin{aligned}\int_{5Q}M\big(\big|T_i^1a_Q\big|^2\big)^{\frac p2}(x_0)\,\mathrm{d}x_0&\leqslant c|Q|^{1-p/2}\Big(\int\big|T_i^1a_Q\big|^2\,\mathrm{d}x\Big)^{\frac p2}\\&\leqslant c|Q|^{1-p/2}\Big(\int|a_Q|^2\,\mathrm{d}x\Big)^{\frac p2}\leqslant c.\end{aligned}\tag{4.1}$$

Similarly for $M\big(\big|T_i^2b_R\big|^2\big)^{\frac q2}$. We now have that

$$\int\Big(\Sigma_1(x_0)\Big)^r\,\mathrm{d}x_0\leqslant c\Big(\sum_Q\lambda_Q^p\Big)^{\frac rp}\Big(\sum_R\mu_R^q\Big)^{\frac rq}\leqslant c\|f\|_{H^p}^r\|g\|_{H^q}^r.\tag{4.2}$$

Case 2. $x_0\in 5Q, x_0\notin 5R$. For any integer $j\geqslant 0$, we denote by $P_a^j(y)$ the Taylor polynomial of homogeneous degree j of $\varphi_{t,x_0}(\,\cdot\,)$ at the point a. Let c_R be the center of the ball R and $c_R'=c_R^{-1}x_0$. We have

$$\begin{aligned}S(a_Q,b_R)(x_0)&\leqslant\sup_{0<t\leqslant\frac12\rho(c_R')}\Big|\int\varphi_{t,x_0}B(a_Q,b_R)\,\mathrm{d}x\Big|+\sup_{t>\frac12\rho(c_R')}\Big|\int\varphi_{t,x_0}B(a_Q,b_R)\,\mathrm{d}x\Big|\\&=S_1(a_Q,b_R)(x_0)+S_2(a_Q,b_R)(x_0).\end{aligned}$$

Consider S_1 first. We have

$$\begin{aligned}S_1(a_Q,b_R)(x_0)\leqslant&\sup_{0<t\leqslant\frac12\rho(c_R')}\Big|\int_{2R}\varphi_{t,x_0}B(a_Q,b_R)\,\mathrm{d}x\Big|\\&+\sup_{0<t\leqslant\frac12\rho(c_R')}\Big|\int_{(2R)^c}\varphi_{t,x_0}B(a_Q,b_R)\,\mathrm{d}x\Big|=S_{11}(x_0)+S_{12}(x_0).\end{aligned}$$

Since $\rho(c_R')\geqslant 2t$, note that c_R is not in the support of φ_{t,x_0} and thus $P_{c_R}^l\equiv 0$. Here l is a large integer to be determined later. Thus

$$\begin{aligned}&S_{11}(x_0)\\&\leqslant\sum_i\sup_{0<t\leqslant\frac12\rho(c_R')}\Big|\int_{\{x\in 2R\,:\,\rho(x^{-1}x_0)\leqslant t\}}\big(\varphi_{t,x_0}(x)-P_{c_R}^l(c_R^{-1}x)\big)T_i^1a_Q(x)T_i^2b_R(x)\,\mathrm{d}x\Big|\\&\leqslant c\sum_i\sup_{0<t\leqslant\frac12\rho(c_R')}\int_{\{x\in 2R\,:\,\rho(x^{-1}x_0)\leqslant t\}}t^{-D-l-1}\rho(c_R^{-1}x)^{l+1}\big|T_i^1a_Q(x)\big|\big|T_i^2b_R(x)\big|\,\mathrm{d}x\end{aligned}$$

(using that $\frac{1}{t} \leqslant \frac{1}{\rho(x^{-1}x_0)} \leqslant c\frac{1}{\rho(c_R^{-1}x_0)}$ since $x \in 2R$, and $x_0 \notin 5R$)

$$S_{11}(x_0) \leqslant c\sum_i \frac{1}{\rho(c_R^{-1}x_0)^{\frac{D}{2}+l+1}} \sup_{t>0}\Big(\frac{1}{t^D}\int\limits_{\rho(x^{-1}x_0)\leqslant t} \big|T_i^1 a_Q(x)\big|^2\,\mathrm{d}x\Big)^{\frac12} \cdot \Big(\int\limits_{2R} \rho(c_R^{-1}x)^{2(l+1)}\big|T_i^2 b_R(x)\big|^2\,\mathrm{d}x\Big)^{\frac12}$$

$$\leqslant c\sum_i M\big(\big|T_i^1 a_Q\big|^2\big)^{\frac12}(x_0)\frac{|R|^{-\frac1q+\frac12+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}, \tag{4.3}$$

since it can be easily seen that

$$\Big(\int\limits_{(2R)^c} \rho(c_R^{-1}x)^{2(l+1)}\big|T_i^2 b_R(x)\big|^2\,\mathrm{d}x\Big)^{\frac12} \leqslant c|R|^{\frac{l+1}{D}}\|b_R\|_2 \leqslant\ c|R|^{-\frac1q+\frac12+\frac{l+1}{D}}.$$

Next, observe that

$$\int\limits_{x_0\notin 5R}\Bigg(\frac{|R|^{-\frac1q+\frac12+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}\Bigg)^q\,\mathrm{d}x_0 \leqslant c, \tag{4.4}$$

provided that we have $\big(\frac{D}{2}+l+1\big)q > D$. Fix l to be the least nonnegative integer such that $l > D\big(\frac1q - 1\big) + \frac{D}{2} - 1$. We now deduce the inequality

$$\int S_{11}^r(x_0)\,\mathrm{d}x_0 \leqslant c\|f\|_{H^p}^r\|g\|_{H^q}^r$$

as a consequence of (4.1), (4.4), and Hölder's inequality.

Next consider $S_{12}(x_0)$. We have

$$S_{12}(x_0) \leqslant \sum_i \sup_{0<t\leqslant\frac12\rho(c_R')}\Big(\int\limits_{(2R)^c}|\varphi_{t,x_0}|\big|T_i^1a_Q\big|^2\,\mathrm{d}x\Big)^{\frac12}\Big(\int\limits_{(2R)^c}|\varphi_{t,x_0}|\,\big|T_i^2b_R\big|^2\,\mathrm{d}x\Big)^{\frac12}$$

$$\leqslant c\sum_i M\big(\big|T_i^1a_Q\big|^2\big)^{\frac12}(x_0)\sup_{0<t\leqslant\frac12\rho(c_R')}\Bigg(\int\limits_{(2R)^c}|\varphi_{t,x_0}|\Bigg(\frac{|R|^{-\frac1q+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x\big)^{D+l+1}}\Bigg)^2\mathrm{d}x\Bigg)^{\frac12},$$

where we use (3.6) for b_R. Now, since $\rho\big(c_R^{-1}x_0\big) \geqslant 2t$ and $\rho(x^{-1}x_0) \leqslant t$, we have $\rho\big(c_R^{-1}x\big) \geqslant \rho\big(c_R^{-1}x_0\big) - \rho(x^{-1}x_0) \geqslant \frac12\rho\big(c_R^{-1}x_0\big)$. Therefore

$$S_{12}(x_0) \leqslant c\sum_i M\big(\big|T_i^1a_Q\big|^2\big)^{\frac12}(x_0)\frac{|R|^{-\frac1q+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}}.$$

It is easy to check that

$$\int\limits_{(5R)^c}\Bigg(\frac{|R|^{-\frac1q+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}}\Bigg)^q\,\mathrm{d}x_0 \leqslant c. \tag{4.5}$$

So we have the inequality

$$\int S_{12}^r(x_0)\,\mathrm{d}x_0 \leqslant c\|f\|_{H^p}^r\|g\|_{H^q}^r$$

as a consequence of (4.1), (4.5), and Hölder's inequality.

We now turn our attention to $S_2(a_Q,b_R)$. We have

$$\begin{aligned}
S_2(a_Q,b_R)(x_0) &= \sup_{t>\frac12\rho(c_R')}\Big|\int\Big(\varphi_{t,x_0}(x)-P_{c_R}^l(c_R^{-1}x)\Big)\sum_i(T_i^1a_Q)(x)(T_i^2b_R)(x)\,\mathrm{d}x\Big| \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(by assumption (1.5))} \\
&\leqslant c\sup_{t>\frac12\rho(c_R')}\int\frac{\rho\big(c_R^{-1}x\big)^{l+1}}{t^{D+l+1}}\sum_i\big|T_i^1a_Q(x)\big|\big|T_i^2b_R(x)\big|\,\mathrm{d}x \\
&\leqslant c\sup_{t>\frac12\rho(c_R')}\Big(\int_{\rho(x^{-1}x_0)\geqslant 4t}\mathrm{d}x\Big)+c\sup_{t>\frac12\rho(c_R')}\Big(\int_{\rho(x^{-1}x_0)<4t}\mathrm{d}x\Big) \\
&= S_{21}(x_0)+S_{22}(x_0).
\end{aligned}$$

Consider first S_{21}. The inequality $\rho(x^{-1}x_0)\geqslant 4t$ implies $\rho\big(c_R^{-1}x\big)\geqslant\rho(x^{-1}x_0)-\rho\big(c_R^{-1}x_0\big)\geqslant 2t>\rho\big(c_R^{-1}x_0\big)$, thus $x\notin 5R$ and $\rho(x^{-1}x_0)\leqslant 2\rho\big(c_R^{-1}x\big)$. Using (3.6) for b_R, we obtain

$$\begin{aligned}
S_{21}(x_0) &\leqslant c\sum_i\frac{1}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}} \\
&\qquad\cdot\sup_{t>\frac12\rho(c_R')}\int_{\rho(x^{-1}x_0)\geqslant 4t}\rho\big(c_R^{-1}x\big)^{l+1}\big|T_i^1a_Q(x)\big|\frac{|R|^{-\frac1q+1+\frac{l+2}{D}}}{\rho\big(c_R^{-1}x\big)^{D+l+2}}\,\mathrm{d}x \\
&\leqslant c\sum_i\frac{1}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}}\sup_{t>0}\int_{\rho(x^{-1}x_0)\geqslant 4t}\frac{|T_i^1a_Q(x)|}{\rho(x^{-1}x_0)^{\frac D2}}\frac{|R|^{-\frac1q+1+\frac{l+2}{D}}}{\rho\big(c_R^{-1}x\big)^{\frac D2+1}}\,\mathrm{d}x \\
&\leqslant c\sum_i\frac{|R|^{-\frac1q+1+\frac{l+2}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}} \\
&\qquad\cdot\sup_{t>0}\Big(\int_{\rho(x^{-1}x_0)\geqslant 4t}\frac{|T_i^1a_Q(x)|^2}{\rho(x^{-1}x_0)^D}\,\mathrm{d}x\Big)^{\frac12}\Big(\int_{(5R)^c}\frac{1}{\rho\big(c_R^{-1}x\big)^{D+2}}\,\mathrm{d}x\Big)^{\frac12} \\
&\leqslant c\sum_i M\big(\big|T_i^1a_Q\big|^2\big)^{\frac12}(x_0)\frac{|R|^{-\frac1q+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}}.
\end{aligned}$$

By (4.1), (4.5), and Hölder's inequality, it follows that

$$\int S_{21}^r(x_0)\,\mathrm{d}x_0\leqslant c\|f\|_{H^p}^r\|g\|_{H^q}^r.$$

Now consider S_{22}. We have $S_{22}(x_0) \leqslant S_{221}(x_0) + S_{222}(x_0)$, where

$$S_{221}(x_0) = c \sum_i \sup_{t > \frac{1}{2}\rho(c'_R)} \Big(\int_{\{x \in 2R \,:\, \rho(x^{-1}x_0) \leqslant 4t\}} \frac{\rho\big(c_R^{-1}x\big)^{l+1}}{t^{D+l+1}} \big|T_i^1 a_Q(x)\big|\big|T_i^2 b_R(x)\big| \,\mathrm{d}x \Big)$$

$$S_{222}(x_0) = c \sum_i \sup_{t > \frac{1}{2}\rho(c'_R)} \Big(\int_{\{x \in (2R)^c \,:\, \rho(x^{-1}x_0) \leqslant 4t\}} \frac{\rho\big(c_R^{-1}x\big)^{l+1}}{t^{D+l+1}} \big|T_i^1 a_Q(x)\big|\big|T_i^2 b_R(x)\big| \,\mathrm{d}x \Big).$$

Arguing similarly as for the term S_{11}, we obtain estimate (4.3) for S_{221}. Next

$$\begin{aligned} S_{222}(x_0) &\leqslant c \sum_i \frac{|R|^{-\frac{1}{q}+1+\frac{l+2}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}} \\ &\qquad \cdot \sup_{t>0} \Big(\frac{1}{t^D} \int_{\rho(x^{-1}x_0) \leqslant 4t} \big|T_i^1 a_Q\big|^2 \,\mathrm{d}x \Big)^{\frac{1}{2}} \Big(\int_{(2R)^c} \frac{1}{\rho\big(c_R^{-1}x\big)^{2(D+1)}} \,\mathrm{d}x \Big)^{\frac{1}{2}} \\ &\leqslant c \sum_i M\big(\big|T_i^1 a_Q\big|^2\big)^{\frac{1}{2}}(x_0) \frac{|R|^{-\frac{1}{q}+\frac{1}{2}+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}. \end{aligned}$$

As before using (4.1), (4.4), and Hölder's inequality, we conclude that

$$\int (\Sigma_2(x_0))^r \,\mathrm{d}x_0 \leqslant c \|f\|_{H^p}^r \|g\|_{H^q}^r. \tag{4.6}$$

This finishes the proof of Case 2.

Case 3. $x_0 \notin 5Q, x_0 \in 5R$. This case is the same as case 2. Let us denote by m the least nonnegative integer satisfying $m > D(\frac{1}{p} - 1) + \frac{D}{2} - 1$. ($m$ in this case plays the role of l in Case 2.)

Case 4. $x_0 \notin 5Q, x_0 \notin 5R$. Without loss of generality, we assume that $\operatorname{supp}\varphi \subseteq [-\frac{9}{10}, \frac{9}{10}]$. Divide this case into the following four nonmutually exclusive subcases:

1° $0 < t \leqslant \rho(c'_Q)$, $0 < t \leqslant \rho(c'_R)$;
2° $0 < t \leqslant \frac{1}{2}\rho(c'_Q)$, $t > \rho(c'_R)$;
3° $t > \rho(c'_Q)$, $0 < t \leqslant \frac{1}{2}\rho(c'_R)$;
4° $t > \frac{1}{2}\rho(c'_Q)$, $t > \frac{1}{2}\rho(c'_R)$.

Subcase 3° is similar to 2°. So we consider Subcases 1°, 2°, and 4° only. Let $k = m + l + 1$, where l and m are the integers that appeared in Cases 2 and 3.

Subcase 1°. In this subcase, $P^j_{c_Q}(\,\cdot\,) \equiv P^j_{c_R}(\,\cdot\,) \equiv 0$ for all j, since both c_Q and c_R are not in the support of φ_{t,x_0}. We have

$$\begin{aligned} \sup_{1°} \Big| \int \varphi_{t,x_0}(x) \sum_i \big(T_i^1 a_Q\big)(x)\big(T_i^2 b_R\big)(x) \,\mathrm{d}x \Big| &\leqslant \sup_{1°} \Big| \int_{2Q \cap 2R} \Big| + \sup_{1°} \Big| \int_{2Q \cap (2R)^c} \Big| \\ &\quad + \sup_{1°} \Big| \int_{(2Q)^c \cap 2R} \Big| + \sup_{1°} \Big| \int_{(2Q)^c \cap (2R)^c} \Big| = I_1(x_0) + I_2(x_0) + I_3(x_0) + I_4(x_0). \end{aligned}$$

Note that $I_1(x_0) \neq 0$ when $\rho(x^{-1}x_0) \leqslant t$. So

$$\frac{1}{t} \leqslant \frac{1}{\rho(x^{-1}x_0)} \leqslant c\frac{1}{\rho\big(c_Q^{-1}x_0\big)} \quad \text{and similarly} \quad \frac{1}{t} \leqslant c\frac{1}{\rho\big(c_R^{-1}x_0\big)}$$

since $x \in 2Q \cap 2R$ but $x_0 \notin 5Q \cup 5R$. When $|R| \leqslant |Q|$, since $P^k_{c_R} \equiv 0$, we have

$$\begin{aligned}
I_1(x_0) &= \sup_{1^\circ}\Big|\int\limits_{2Q\cap 2R}\Big(\varphi_{t,x_0}(x) - P^k_{c_R}(c_R^{-1}x)\Big)\sum_i (T_i^1 a_Q)(T_i^2 b_R)\,\mathrm{d}x\Big| \\
&\leqslant c\sum_i \sup_{1^\circ}\int\limits_{2Q\cap 2R}\frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}}\big|T_i^1 a_Q\big|\big|T_i^2 b_R\big|\,\mathrm{d}x \\
&\leqslant c\sum_i \sup_{1^\circ}\int\limits_{2Q\cap 2R}\frac{|Q|^{\frac{m+1}{D}}}{t^{\frac{D}{2}+m+1}}\big|T_i^1 a_Q\big|\frac{|R|^{\frac{l+1}{D}}}{t^{\frac{D}{2}+l+1}}\big|T_i^2 b_R\big|\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{1}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}\Big(\int\limits_{2Q}|Q|^{2(m+1)}\big|T_i^1 a_Q\big|^2\,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\qquad\cdot\frac{1}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}\Big(\int\limits_{2R}|R|^{2(l+1)}\big|T_i^2 b_R\big|^2\,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}\frac{|R|^{-\frac{1}{q}+\frac{1}{2}+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}.
\end{aligned}$$

If $|R| > |Q|$, we use $P^k_{c_Q}(c_Q^{-1}x)$ instead of $P^k_{c_R}(c_R^{-1}x)$ above to get the same estimate. The inequality

$$\int I_1^r(x_0)\,\mathrm{d}x_0 \leqslant c\|f\|^r_{H^p}\|g\|^r_{H^q}$$

follows as before. Now consider I_2. Since $P^m_{c_Q} \equiv 0$,

$$\begin{aligned}
I_2(x_0) &\leqslant c\sum_i \sup_{1^\circ}\Big|\int\limits_{2Q\cap(2R)^c}\Big(\varphi_{t,x_0}(x) - P^m_{c_Q}(c_Q^{-1}x)\Big)\chi_{\rho(x^{-1}x_0)\leqslant t}\big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big| \\
&\leqslant c\sum_i \sup_{1^\circ}\int\limits_{2Q\cap(2R)^c}\frac{\rho\big(c_Q^{-1}x\big)^{m+1}}{t^{\frac{D}{2}+m+1}}\big|T_i^1 a_Q(x)\big|\frac{1}{t^{\frac{D}{2}}}\chi_{\rho(x^{-1}x_0)\leqslant t}\big|T_i^2 b_R(x)\big|\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{1}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}\Big(\int\limits_{2Q}\rho\big(c_Q^{-1}x\big)^{2(m+1)}\big|T_i^1 a_Q(x)\big|^2\,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\qquad\cdot\sup_{1^\circ}\Bigg(\int\limits_{(2R)^c}\frac{1}{t^D}\chi_{\rho(x^{-1}x_0)\leqslant t}\Bigg(\frac{|R|^{-\frac{1}{q}+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x\big)^{D+l+1}}\Bigg)^2\,\mathrm{d}x\Bigg)^{\frac{1}{2}}.
\end{aligned}$$

Now $\rho\big(c_R^{-1}x\big) \geqslant \rho\big(c_R^{-1}x_0\big) - \rho(x^{-1}x_0) \geqslant \frac{1}{10}\rho\big(c_R^{-1}x_0\big)$ whenever $x \in \operatorname{supp}\varphi_{t,x_0}$.

Therefore

$$I_2(x_0) \leqslant c \frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}}$$

as desired. Term I_3 is similar to I_2. We now treat term I_4.

$$\begin{aligned} I_4(x_0) &\leqslant \sum_i \sup_{1^\circ} \int\limits_{(2Q)^c\cap(2R)^c} |\varphi_{t,x_0}(x)|\big|T_i^1 a_Q(x)\big|\big|T_i^2 b_R(x)\big| \,\mathrm{d}x \\ &\leqslant c \sum_i \sup_{1^\circ} \Bigg(\int\limits_{(2Q)^c} |\varphi_{t,x_0}(x)| \bigg(\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x\big)^{D+m+1}} \bigg)^2 \mathrm{d}x \Bigg)^{\frac{1}{2}} \\ &\quad \cdot \sup_{1^\circ} \Bigg(\int\limits_{(2R)^c} |\varphi_{t,x_0}(x)| \bigg(\frac{|R|^{-\frac{1}{q}+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x\big)^{D+l+1}} \bigg)^2 \mathrm{d}x \Bigg)^{\frac{1}{2}}. \end{aligned}$$

By the previous argument we have that $\rho\big(c_Q^{-1}x\big) \geqslant \frac{1}{10}\rho\big(c_Q^{-1}x_0\big), \rho\big(c_R^{-1}x\big) \geqslant \frac{1}{10}\rho\big(c_R^{-1}x_0\big)$. Thus

$$I_4(x_0) \leqslant c \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+l+1}}$$

as before.

Subcase 4°. In this subcase, because of symmetry we may assume that $|R| \leqslant |Q|$.

Using $\frac{1}{t} \leqslant \frac{2}{\rho\left(c_Q^{-1}x_0\right)}$ and $\frac{1}{t} \leqslant \frac{2}{\rho\left(c_R^{-1}x_0\right)}$ and (1.5) we obtain

$$\begin{aligned} &\sup_{4^\circ}\Big| \int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big| \\ &= \sup_{4^\circ}\Big| \int \Big(\varphi_{t,x_0}(x) - P^k_{c_R}(c_R^{-1}x)\Big) \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big| \\ &\leqslant c \sum_i \sup_{4^\circ} \int \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}} \big|T_i^1 a_Q\big|\big|T_i^2 b_R\big|\,\mathrm{d}x \\ &\leqslant c \sum_i \Bigg(\sup_{4^\circ} \int\limits_{2Q\cap 2R} + \sup_{4^\circ} \int\limits_{2Q\cap(2R)^c} + \sup_{4^\circ} \int\limits_{(2Q)^c\cap 2R} + \sup_{4^\circ} \int\limits_{(2Q)^c\cap(2R)^c} \Bigg) \\ &= J_1(x_0) + J_2(x_0) + J_3(x_0) + J_4(x_0). \end{aligned}$$

To estimate J_1 note that $\rho\big(c_R^{-1}x\big) \leqslant 2|R|^{\frac{1}{D}} \leqslant 2|Q|^{\frac{1}{D}}$. Thus

$$\begin{aligned}
J_1(x_0) &\leqslant c\sum_i \frac{|Q|^{\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \frac{|R|^{\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}} \Big(\int \big|T_i^1 a_Q\big|^2 \,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\qquad\cdot \Big(\int \big|T_i^2 b_R\big|^2 \,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \frac{|R|^{-\frac{1}{q}+\frac{1}{2}+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}
\end{aligned}$$

as desired. Using (3.6) for b_R, we estimate J_2 as follows:

$$\begin{aligned}
J_2(x_0) &\leqslant c\sum_i \sup_{4^\circ} \int\limits_{2Q\cap(2R)^c} \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}} \big|T_i^1 a_Q\big| \frac{|R|^{-\frac{1}{q}+1+\frac{l+m+3}{D}}}{\rho\big(c_R^{-1}x\big)^{D+m+l+3}} \,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|Q|^{\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \Big(\int \big|T_i^1 a_Q\big|^2 \,\mathrm{d}x\Big)^{\frac{1}{2}} \frac{|R|^{-\frac{1}{q}+1+\frac{l+2}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}} \\
&\qquad\cdot \Big(\int\limits_{(2R)^c} \frac{1}{\rho\big(c_R^{-1}x\big)^{2(D+1)}} \,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \frac{|R|^{-\frac{1}{q}+\frac{1}{2}+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}
\end{aligned}$$

as desired. To estimate J_3 note that $\rho\big(c_R^{-1}x\big) \leqslant 2|R|^{\frac{1}{D}} \leqslant 2|Q|^{\frac{1}{D}} \leqslant \rho\big(c_Q^{-1}x\big)$. Using (3.6) for a_Q we obtain

$$\begin{aligned}
J_3(x_0) &\leqslant c\sum_i \sup_{4^\circ} \int\limits_{(2Q)^c\cap 2R} \frac{\rho\big(c_R^{-1}x\big)^{l+1}}{t^{\frac{D}{2}+l+1}} \big|T_i^2 b_R(x)\big| \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{t^{\frac{D}{2}+m+1}} \frac{\rho\big(c_R^{-1}x\big)^{m+1}}{\rho\big(c_Q^{-1}x\big)^{D+m+1}} \,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|R|^{\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}} \Big(\int \big|T_i^2 b_R\big|^2 \,\mathrm{d}x\Big)^{\frac{1}{2}} \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \\
&\qquad\cdot \Big(\int\limits_{(2Q)^c} \frac{1}{\rho\big(c_Q^{-1}x\big)^{2D}} \,\mathrm{d}x\Big)^{\frac{1}{2}} \\
&\leqslant c\frac{|R|^{-\frac{1}{q}+\frac{1}{2}+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}} \frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}
\end{aligned}$$

as desired. Finally using (3.6) for both a_Q and b_R we obtain

$$\begin{aligned}
J_4(x_0) &\leqslant c\sum_i \sup_{4^\circ} \int\limits_{(2Q)^c\cap(2R)^c} \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x\big)^{D+m+1}} \frac{1}{t^{\frac{D}{2}+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x\big)^{D+k+1}} \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{\frac{D}{2}+l+1}}\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} |Q|^{\frac{m+1}{D}} \Bigg(\int\limits_{(2Q)^c} \frac{1}{\rho\big(c_Q^{-1}x\big)^{2(D+m+1)}}\,\mathrm{d}x\Bigg)^{\frac{1}{2}} \\
&\qquad \cdot \frac{|R|^{-\frac{1}{q}+1+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}} \Bigg(\int\limits_{(2R)^c} \frac{1}{\rho\big(c_R^{-1}x\big)^{2D}}\,\mathrm{d}x\Bigg)^{\frac{1}{2}} \\
&\leqslant c \frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}} \frac{|R|^{-\frac{1}{q}+\frac{1}{2}+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}
\end{aligned}$$

as desired.

This concludes Subcase 4°. We are now left with Subcases 2° and 3°. Because of symmetry we only consider the former.

Subcase 2°. We break up Subcase 2° of Case 4 into two subsubcases:

Subsubcase $|R| \leqslant |Q|$. In this subsubcase write

$$\begin{aligned}
&\sup_{2^\circ}\Big|\int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big| \\
&\quad = \sup_{2^\circ}\Big|\int \big(\varphi_{t,x_0}(x) - P^k_{c_R}(c_R^{-1}x)\big) \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big| \qquad \text{by (1.5)} \\
&\quad \leqslant c\sum_i \sup_{2^\circ} \int \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}} \big|T_i^1 a_Q\big|\,\big|T_i^2 b_R\big|\,\mathrm{d}x \\
&\quad \leqslant c\sum_i \Big(\sup_{2^\circ}\int\limits_{2Q\cap 2R} + \sup_{2^\circ}\int\limits_{2Q\cap(2R)^c} + \sup_{2^\circ}\int\limits_{(2Q)^c\cap 2R} + \sup_{2^\circ}\int\limits_{(2Q)^c\cap(2R)^c}\Big) \\
&\quad = K_1(x_0) + K_2(x_0) + K_3(x_0) + K_4(x_0).
\end{aligned}$$

We begin with term K_1. If $2Q$ and $2R$ intersect, it follows that $\rho(c_R^{-1}c_Q) < 2|R|^{\frac{1}{D}} + 2|Q|^{\frac{1}{D}} < 4|Q|^{\frac{1}{D}}$. Observe that $t > \rho\big(c_R^{-1}x_0\big) \geqslant \rho\big(c_Q^{-1}x_0\big) - \rho(c_R^{-1}c_Q) \geqslant 5|Q|^{\frac{1}{D}} - 4|Q|^{\frac{1}{D}} = |Q|^{\frac{1}{D}}$. Hence $\rho\big(c_Q^{-1}x_0\big) \leqslant \rho\big(c_R^{-1}x_0\big) + 4|Q|^{\frac{1}{D}} < 5t$. Using that both $\rho(c_Q^{-1}x_0)$ and $\rho(c_R^{-1}x_0)$ are less than a multiple of t we obtain

$$\begin{aligned}
K_1(x_0) &\leqslant c\sum_i \sup_{2^\circ} \int\limits_{2Q\cap 2R} \frac{\rho\big(c_R^{-1}x\big)^{m+1}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}\big|T_i^1 a_Q\big| \frac{\rho\big(c_R^{-1}x\big)^{l+1}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}\big|T_i^2 b_R\big|\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|R|^{\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}\Big(\int\big|T_i^1 a_Q\big|^2\,\mathrm{d}x\Big)^{\frac12}\frac{|R|^{\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}}\Big(\int\big|T_i^2 b_R\big|^2\,\mathrm{d}x\Big)^{\frac12} \\
&\leqslant c\frac{|Q|^{-\frac1p+\frac12+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{\frac{D}{2}+m+1}}\frac{|R|^{-\frac1q+\frac12+\frac{l+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{\frac{D}{2}+l+1}},
\end{aligned}$$

since $\rho\big(c_R^{-1}x\big)\leqslant 2|R|^{\frac1D}$ and $|R|^{\frac1D}\leqslant|Q|^{\frac1D}$.

We now control term K_2. Note that when $x\in 2Q\cap(2R)^c$, then $\rho\big(c_R^{-1}x\big)\geqslant \rho\big(c_Q^{-1}x_0\big)-\rho\big(c_Q^{-1}x\big)-\rho\big(c_R^{-1}x_0\big)>\frac35\rho\big(c_Q^{-1}x_0\big)-\rho\big(c_R^{-1}x_0\big)\geqslant\frac1{10}\rho\big(c_Q^{-1}x_0\big)$, since $2\rho\big(c_Q^{-1}x_0\big)>10|Q|^{\frac1D}>5\rho\big(c_Q^{-1}x\big)$ and $(\frac35-\frac1{10})\rho\big(c_Q^{-1}x_0\big)>\rho\big(c_R^{-1}x_0\big)$. We now use that $t>\rho(c_R^{-1}x_0)$ and $\rho(c_R^{-1}x)\geqslant\frac1{10}\rho(c_Q^{-1}x_0)$ to obtain

$$\begin{aligned}
K_2(x_0) &\leqslant c\sum_i \sup_{2^\circ} \int\limits_{2Q\cap(2R)^c} \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}}\big|T_i^1 a_Q\big|\frac{|R|^{-\frac1q+1+\frac{k+m+2}{D}}}{\rho\big(c_R^{-1}x\big)^{D+k+m+2}}\,\mathrm{d}x \\
&\leqslant c\frac{|R|^{-\frac1q+1+\frac{k+1}{D}}}{\rho(c_R^{-1}x_0)^{D+k+1}}\frac{|R|^{\frac{m+1}{D}}}{\rho(c_Q^{-1}x_0)^{D+m+1}}\int\big|T_i^1 a_Q\big|\,\mathrm{d}x \\
&\leqslant c\frac{|R|^{-\frac1q+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}}\frac{|Q|^{-\frac1p+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \qquad \text{since}|R|\leqslant|Q|.
\end{aligned}$$

This concludes the estimate for K_2.

Consider K_3 now. Since $5|R|^{\frac1D}\leqslant\rho\big(c_R^{-1}x_0\big)\leqslant t$, it follows that $|R|^{\frac1D}\leqslant\frac15 t$. Hence $\rho(x^{-1}x_0)\leqslant\rho\big(c_R^{-1}x_0\big)+\rho\big(c_R^{-1}x\big)\leqslant\frac75 t\leqslant\frac7{10}\rho\big(c_Q^{-1}x_0\big)$ and $\rho\big(c_Q^{-1}x\big)\geqslant\rho\big(c_Q^{-1}x_0\big)-\rho(x^{-1}x_0)\geqslant\frac3{10}\rho\big(c_Q^{-1}x_0\big)$. Now we have

$$\begin{aligned}
K_3(x_0) &\leqslant c\sum_i \sup_{2^\circ} \int\limits_{(2Q)^c\cap 2R} \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}}\big|T_i^2 b_R\big|\frac{|Q|^{-\frac1p+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x\big)^{D+m+1}}\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|Q|^{-\frac1p+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}}\frac{|R|^{\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}}\int\limits_{2R}\big|T_i^2 b_R(x)\big|\,\mathrm{d}x \\
&\leqslant c\frac{|Q|^{-\frac1p+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}}\frac{|R|^{\frac{k+1}{D}-\frac1q+1}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}},
\end{aligned}$$

which is the desired estimate for K_3.

Finally, we discuss K_4. Let $A(x_0,Q,R)$ be the set of all $x\in(2Q)^c\cap(2R)^c$ with $\rho\big(c_Q^{-1}x\big)\geqslant\frac1{10}\rho\big(c_Q^{-1}x_0\big)$ and $B(x_0,Q,R)$ be the set of all $x\in(2Q)^c\cap(2R)^c$

with $\rho\big(c_Q^{-1}x\big) < \frac{1}{10}\rho\big(c_Q^{-1}x_0\big)$. Then,

$$\begin{aligned}K_{41}(x_0) &\equiv c\sum_i \sup_{2^\circ} \int\limits_{A(x_0,Q,R)} \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x\big)^{D+m+1}} \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+l+2}{D}}}{\rho\big(c_R^{-1}x\big)^{D+k+l+2}}\,\mathrm{d}x\\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}} |R|^{\frac{l+1}{D}} \int\limits_{(2R)^c} \frac{1}{\rho\big(c_R^{-1}x\big)^{D+l+1}}\,\mathrm{d}x\\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}}.\end{aligned}$$

Now suppose that x is in the set $B(x_0,Q,R)$. Then, $\rho(x^{-1}x_0) \geqslant \rho\big(c_Q^{-1}x_0\big) - \rho\big(c_Q^{-1}x\big) > \frac{9}{10}\rho\big(c_Q^{-1}x_0\big) > \frac{9}{5}t$. Hence $\rho\big(c_R^{-1}x\big) \geqslant \rho(x^{-1}x_0) - \rho\big(c_R^{-1}x_0\big) > \frac{4}{9}\rho(x^{-1}x_0) > \frac{2}{5}\rho\big(c_Q^{-1}x_0\big)$. We have

$$\begin{aligned}K_{42}(x_0) &\equiv c\sum_i \sup_{2^\circ} \int\limits_{B(x_0,Q,R)} \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x\big)^{D+m+1}} \frac{\rho\big(c_R^{-1}x\big)^{k+1}}{t^{D+k+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+m+2}{D}}}{\rho\big(c_R^{-1}x\big)^{D+k+m+2}}\,\mathrm{d}x\\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}} |R|^{\frac{m+1}{D}} \int\limits_{(2Q)^c} \frac{1}{\rho\big(c_Q^{-1}x\big)^{D+m+1}}\,\mathrm{d}x\\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}},\end{aligned}$$

since $|R|^{\frac{m+1}{D}} \cdot |Q|^{-\frac{m+1}{D}} \leqslant 1$. Since $K_4 \leqslant K_{41} + K_{42}$, we conclude the estimates for K_4.

We now come to the second Subsubcase of Subcase 2° of Case 4:

Subsubcase $|R| > |Q|$. We use the hypothesis to subtract a suitable term and we then split things in four parts as before.

$$\begin{aligned}&\sup_{2^\circ}\Big| \int \varphi_{t,x_0} \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big|\\
&\quad= \sup_{2^\circ}\Big| \int \big(\varphi_{t,x_0} - P^k_{c_Q}(c_Q^{-1}x)\big) \sum_i \big(T_i^1 a_Q\big)\big(T_i^2 b_R\big)\,\mathrm{d}x\Big|\\
&\quad\leqslant c\Big(\sup_{2^\circ}\int\limits_{2Q\cap 2R} + \sup_{2^\circ}\int\limits_{2Q\cap(2R)^c} + \sup_{2^\circ}\int\limits_{(2Q)^c\cap 2R} + \sup_{2^\circ}\int\limits_{(2Q)^c\cap(2R)^c}\Big)\\
&\quad= L_1(x_0) + L_2(x_0) + L_3(x_0) + L_4(x_0).\end{aligned}$$

We observe that term $L_1 \equiv 0$. In fact, if there were some x in the intersection of the doubles of Q and R that appear in term L_1, then $\rho\big(c_Q^{-1}x_0\big) \leqslant \rho\big(c_R^{-1}x_0\big) + \rho\big(c_Q^{-1}c_R\big) \leqslant t + \rho\big(c_Q^{-1}x\big) + \rho\big(c_R^{-1}x\big) \leqslant t + 2|R|^{\frac{1}{D}} + 2|Q|^{\frac{1}{D}} \leqslant t + 4|R|^{\frac{1}{D}} < t + \frac{4}{5}t < 2t$, which is impossible. Therefore $L_1 \equiv 0$.

We now proceed with term L_2. As we showed for term K_2, we have that $\rho\big(c_R^{-1}x\big) \geqslant \frac{1}{10}\rho\big(c_Q^{-1}x_0\big)$. Using this fact and that $t > \rho\big(c_R^{-1}x_0\big)$, we obtain

$$\begin{aligned}
L_2(x_0) &\leqslant c\sum_i \sup_{2^\circ} \int\limits_{2Q\cap(2R)^c} \frac{\rho\big(c_Q^{-1}x\big)^{k+1}}{t^{D+k+1}} \big|T_i^1 a_Q\big| \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x\big)^{D+k+1}}\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|Q|^{\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+k+1}} \int \big|T_i^1 a_Q\big|\,\mathrm{d}x \\
&\leqslant c\frac{|Q|^{-\frac{1}{p}+1+\frac{k+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+k+1}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}}.
\end{aligned}$$

This concludes the estimate for L_2.

We now consider term L_3. As we showed for term K_3, we have that $\rho\big(c_Q^{-1}x\big) \geqslant \frac{3}{10}\rho\big(c_Q^{-1}x_0\big)$. Thus

$$\begin{aligned}
L_3(x_0) &\leqslant c\sum_i \sup_{2^\circ} \int\limits_{(2Q)^c\cap 2R} \frac{\rho\big(c_Q^{-1}x\big)^{k+1}}{t^{D+k+1}} \big|T_i^2 b_R(x)\big| \frac{|Q|^{-\frac{1}{p}+1+\frac{m+k+2}{D}}}{\rho\big(c_Q^{-1}x\big)^{D+m+k+2}}\,\mathrm{d}x \\
&\leqslant c\sum_i \frac{|Q|^{\frac{k+1}{D}}}{\rho\big(c_R^{-1}x_0\big)^{D+k+1}} \frac{|Q|^{-\frac{1}{p}+\frac{1}{2}+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}} \int\limits_{(2Q)^c\cap 2R} \big|T_i^2 b_R(x)\big|\,\mathrm{d}x \\
&\leqslant c\frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho(c_R^{-1}x_0)^{D+k+1}} \frac{|Q|^{-\frac{1}{p}+1}|Q|^{\frac{m+1}{D}}}{\rho(c_Q^{-1}x_0)^{D+m+1}} \qquad \text{since } |Q| < |R|.
\end{aligned}$$

Finally, we discuss L_4. We have

$$\begin{aligned}
L_4(x_0) &\leqslant c\sum_i \int\limits_{(2Q)^c\cap(2R)^c} \frac{\rho(c_Q^{-1}x)^{k+1}}{t^{D+k+1}} \frac{|Q|^{-\frac{1}{p}+1+\frac{m+k+2}{D}}}{\rho(c_Q^{-1}x)^{D+m+k+2}} \frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho(c_R^{-1}x)^{D+k+1}}\,\mathrm{d}x \\
&\leqslant c\frac{|R|^{-\frac{1}{q}+1+\frac{k+1}{D}}}{\rho(c_R^{-1}x_0)^{D+k+1}} \int\limits_{(2Q)^c\cap(2R)^c} \frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}+\frac{k+1}{D}}}{\rho(c_Q^{-1}x)^{D+m+1}\rho(c_R^{-1}x)^{D+k+1}}\,\mathrm{d}x.
\end{aligned}$$

As we did with term K_4, we consider the sets $A(x_0, Q, R)$ and $B(x_0, Q, R)$. For $x \in A(x_0, Q, R)$, use the estimate $\rho(c_Q^{-1}x) \geqslant \frac{1}{10}\rho\big(c_Q^{-1}x_0\big)$ to bound the integral above by

$$\frac{|Q|^{-\frac{1}{p}+1+\frac{m+1}{D}}}{\rho\big(c_Q^{-1}x_0\big)^{D+m+1}}.$$

For $x \in B(x_0, Q, R)$, we showed before that $\rho\big(c_R^{-1}x\big) > \frac{2}{5}\rho\big(c_Q^{-1}x_0\big)$. Then use this estimate to bound the integral above by

$$\frac{|Q|^{-\frac{1}{p}+1+\frac{k+1}{D}}}{\rho(c_Q^{-1}x_0)^{D+k+1}}.$$

In both cases, we have proved the desired pointwise estimate for term K_4, thus concluding the proof of the second Subsubcase of Subcase 2° of Case 4. Subcase 3° of Case 4 is similar. Theorem 1.7 is now completely proved. ∎

Note that when D is even, m can be $\big[D\big(\frac{1}{p}-1\big)\big]+\frac{D}{2}$ and l can be $\big[D\big(\frac{1}{q}-1\big)\big]+\frac{D}{2}$. Thus k can be as small as $\big[D\big(\frac{1}{p}-1\big)\big]+\big[D\big(\frac{1}{q}-1\big)\big]+D+1$. When D is odd, m can be $\big[D\big(\frac{1}{p}-1\big)\big]+\frac{D+1}{2}$ and l can be $\big[D\big(\frac{1}{q}-1\big)\big]+\frac{D+1}{2}$. In this case k can be as small as $\big[D\big(\frac{1}{p}-1\big)\big]+\big[D\big(\frac{1}{q}-1\big)\big]+D+2$. Moreover, it is easy to see that $r>\frac{D}{D+k+1}$.

5. APPLICATIONS AND EXAMPLES

We can use Theorem 1.5 to extend, in the context of stratified homogeneous groups, the result of [3] which says that the commutator of a Calderón-Zygmund operator and multiplication by a BMO function maps $L^p(\mathbb{R}^n)$ into itself. More precisely, we have the following:

Corollary 5.1. *Let $b\in \mathrm{BMO}(G)$ and T be a Calderón-Zygmund operator as in Section* 1. *Then the operator*

$$[b,T](f)=bT(f)-T(bf)$$

maps $L^p(G)$ boundedly into itself for $1<p<\infty$, and

$$\|[b,T](f)\|_{L^p}\leqslant c\|f\|_{L^p}\|b\|_{\mathrm{BMO}}.$$

Proof. Let p' be the conjugate idex of p. From Theorem 1.5, we know that for all $g\in L^{p'}(G)$, $g(Tf)-f(T^*g)\in H^1(G)$ since assumption (1.3) is obviously satisfied with $k=0$. Moreover, $\|g(Tf)-f(T^*g)\|_{H^1}\leqslant c\|f\|_{L^p}\|g\|_{L^{p'}}$. Using the duality between H^1 and BMO on homogeneous groups ([6]), we obtain

$$\begin{aligned}\Big|\int [b,T](f)(x)g(x)\,\mathrm{d}x\Big| &= \Big|\int b(x)[g(x)(Tf)(x)-f(x)(T^*g)(x)]\,\mathrm{d}x\Big|\\ &\leqslant \|b\|_{\mathrm{BMO}}\|g(Tf)-f(T^*g)\|_{H^1}\leqslant c\|b\|_{\mathrm{BMO}}\|f\|_{L^p}\|g\|_{L^{p'}}.\end{aligned}$$ ∎

Next we discuss another application of our results. Consider the Heisenberg group $\mathcal{H}^n$ which is a stratified homogeneous group of homogeneous dimension $2n+2$.

The Cauchy-Szegö projection on $\mathcal{H}^n$ is defined as the following principal value convolution

$$C(f)(x)=\int\limits_{H^n} K(y^{-1}x)f(y)\,\mathrm{d}y,\quad f\in L^2(\mathcal{H}^n)\text{ with compact support,}$$

where K is a homogeneous distribution of degree $-(2n+2)$ which equals the smooth function $c(t+\mathrm{i}|\xi|^2)^{-n-1}$ away from the origin, $x=[\xi,t]\in\mathcal{H}^n$, $\xi\in R^{2n}$, $t\in R$, and $c=2^{n-1}\mathrm{i}^{n+1}n!/\pi^{n+1}$, $\mathrm{i}^2=-1$. Let C^* denote the adjoint of C. It is easy to see that $C=C^*$.

The following is a consequence of Theorem 1.5 and Theorem 1.6. As before, we set $H^p=L^p$ for $p>1$.

COROLLARY 5.2. *Let $0 < p, q < \infty$ and assume that at least one of the p, q is bigger than one. Let $r > \frac{2n+2}{2n+3}$ satisfy $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Then for $f \in H^p(\mathcal{H}^n)$ and $g \in H^q(\mathcal{H}^n)$ we have that $B(f,g) = fC(g) - gC^*(f) \in H^r(\mathcal{H}^n)$, and*

$$\|fC(g) - gC^*(f)\|_{H^r(\mathcal{H}^n)} \leqslant C\|f\|_{H^p(\mathcal{H}^n)}\|g\|_{H^q(\mathcal{H}^n)},$$

for some constant C independent of f and g.

It can be seen that this bilinear operator B does not have higher order moments vanishing and thus we can not expect it to map into $H^r(\mathcal{H}^n)$, for $r \leqslant \frac{2n+2}{2n+3}$.

Examples of bilinear operators with vanishing moments of all orders are given by

$$B_1(f,g) = f(Hg) + (Hf)g \quad \text{and} \quad B_2(f,g) = (Hf)(Hg) - fg,$$

where H is the usual Hilbert transform on $\mathbb{R}^1$. For these operators we obtain $H^p \times H^q \to H^r$ boundedness for all $0 < p, q, r < \infty$, with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$.

Suitable combinations of Riesz transforms give examples of operators with only a finite number of vanishing moments. For instance, if R_1 and R_2 are the usual Riesz transforms in $\mathbb{R}^2$, the operator

$$D(f,g) = (R_1^2 f)(R_2^2 g) - 2(R_1R_2 f)(R_1R_2 g) + (R_2^2 f)(R_1^2 g)$$

has integral and first order moments vanishing (but no higher order moments vanishing). This operator is naturally obtained from the determinant of the $2\times 2\times 2$ matrix of all second order partial derivatives of the function $(f,g) : \mathbb{R}^2 \to \mathbb{R}^2$. Theorem 1.7 gives that the operator D is bounded from $H^p \times H^q$ into H^r for $r > \frac{1}{2}$. This result is analogous to the theorem in [2] on the H^1 boundedness of the Jacobian.

6. APPENDIX : EXISTENCE OF COMPACTLY SUPPORTED COMMUTATIVE APPROXIMATE IDENTITIES ON STRATIFIED GROUPS

The following construction was communicated to us by J. Dziubański.

Let G be a connected and simply connected stratified homogeneous Lie group. Let $\Delta = -\sum_j X_j^2$ be the sublaplacian on G. By [9], for any Schwartz function m on $[0,\infty)$, there exists a Schwartz function $\widetilde{m}$ on G such that

$$\int_0^\infty m(\lambda)\,\mathrm{d}E(\lambda)f = f * \widetilde{m},$$

where $\int_0^\infty \lambda\,\mathrm{d}E(\lambda)f = \Delta f$ is the spectral decomposition of the sublaplacian on G. Moreover, $\int_G \widetilde{m}(x)\,\mathrm{d}x = m(0)$, and if $m^t(\lambda) = m(t\lambda)$, $t > 0$, then $(\widetilde{m}^t)(x) = t^{-D/2}\widetilde{m}(\delta_{t^{-1/2}}x)$. Also, for m, η being Schwartz functions on $[0,\infty)$, we have $\widetilde{m} * \widetilde{\eta} = \widetilde{\eta} * \widetilde{m}$.

Let ψ be a real smooth even function on the line supported in the interval $[-1,1]$ with integral 1. Let $m = \widehat{\psi}$. Since m can be extended to an even holomorphic function on the complex plane, the function $\eta(\lambda) = m(\sqrt{\lambda})$ is also a Schwartz function on $[0,\infty)$ and $\eta(0) = 1$.

We claim that the Schwartz function $\widetilde{\eta}$ on G is compactly supported. To see the claim, consider the fundamental solution Γ_t of the wave equation $(\Delta+\partial_t^2)u = 0$ on $G\times\mathbb{R}$ with initial conditions $u(x,0) = f$ and $\partial_t u(x,0) = 0$. An easy calculation shows that $f * \Gamma_t = \int_0^\infty \cos(t\sqrt{\lambda})\,\mathrm{d}E(\lambda)f$. Furthermore,

$$
\begin{aligned}
f*\widetilde{\eta} &= \int_0^\infty \widehat{\psi}(\sqrt{\lambda})\,\mathrm{d}E(\lambda)f = \int_0^\infty\int_{-1}^1 \psi(t)\cos(t\sqrt{\lambda})\,\mathrm{d}t\,\mathrm{d}E(\lambda)f \\
&= \int_{-1}^1 \psi(t)\Big[\int_0^\infty \cos(t\sqrt{\lambda})\,\mathrm{d}E(\lambda)f\Big]\,\mathrm{d}t = \int_{-1}^1 \psi(t)(f*\Gamma_t)\,\mathrm{d}t = f*\Big[\int_{-1}^1 \psi(t)\Gamma_t(\,\cdot\,)\,\mathrm{d}t\Big].
\end{aligned}
$$

By [10], the support of Γ_t is contained in the "cone" $\big\{(x,t) : \|x\| \leqslant |t|\big\}$, where $\|\cdot\|$ is the distance associated with the vector fields X_j as in the work of [11]. The support properties of Γ_t and the identities above imply that $\widetilde{\eta}$ is compactly supported as a distribution and hence as a function. By the properties of $\widetilde{m}$, it follows that $\widetilde{\eta}_t(x) = t^{-D}\widetilde{\eta}(\delta_{\frac{1}{t}}x)$ is a compactly supported commutative approximate identity on G.

Acknowledgements. The first named author's research was partially supported by the National Science Foundation.

REFERENCES

1. R. Coifman, L. Grafakos, Hardy space estimates for multilinear operators. I, *Rev. Mat. Iberoamericana* **8**(1992), 45–62.
2. R.R. Coifman, P.-L. Lions, Y. Meyer, S. Semmes, Compacité par compensation et éspaces de Hardy, *J. Math. Pures Appl. (9)* **72**(1993), 247–286.
3. R.R. Coifman, R. Rochberg, G. Weiss, Factorization theorems for Hardy spaces in several variables, *Ann. of Math.* **103**(1976), 611–635.
4. R. R. Coifman, G. Weiss, Extension of Hardy spaces and their use in analysis, *Bull. Amer. Math. Soc.* **83**(1977), 569–645.
5. G. Folland, Subelliptic estimates and function spaces on nilpotent Lie groups, ıt Ark. Mat. **13**(1975), 161–207.
6. G. Folland, E. Stein, *Hardy Spaces on Homogeneous Groups,* Math. Notes, vol. 28, Princeton Univ. Press, Princeton 1982.
7. L. Grafakos, Hardy space estimates for multilinear operators. II, *Rev. Mat. Iberoamericana* **8**(1992), 69–92.
8. W. Hebisch, A. Sikora, A smooth subadditive homogeneous norm on a homogeneous group, *Studia Math.* **96**(1990), 231–236.
9. A. Hulanicki, A functional calculus for Rockland operators on nilpotent Lie groups, *Studia Math.* **78**(1984), 253–266.

10. R. MELROSE, *Propagation for the Wave Group of Positive Subelliptic Second Order Differential Operators in Hyperbolic Equations and Related Topics*, Academic Press, Boston 1986.
11. A. NAGEL, E.M. STEIN, S. WAINGER, Balls and metrics defined by vector fields. I: Basic properties, *Acta Math.* **155**(1985), 103–147.
12. E. STEIN, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, Princeton 1993.

LOUKAS GRAFAKOS
Department of Mathematics
University of Missouri
Columbia, MO 65211
USA

E-mail: loukas@math.missouri.edu

XINWEI LI
Department of Mathematics
Washington University
St. Louis, MO 63130
USA

E-mail: xli@math.wustl.edu

Received May 15, 1998; revised August 31, 1998.