STUDIA MATHEMATICA 195 (1) (2009)

Littlewood–Paley g-functions with rough kernels on homogeneous groups

by

YONG DING and XINFENG WU (Beijing)

Abstract. Let $\mathbb{G}$ be a homogeneous group on $\mathbb{R}^n$ whose multiplication and inverse operations are polynomial maps. In 1999, T. Tao proved that the singular integral operator with $L\log^+L$ function kernel on $\mathbb{G}$ is both of type (p,p) and of weak type $(1,1)$. In this paper, the same results are proved for the Littlewood–Paley g-functions on $\mathbb{G}$.

1. Introduction. Let Ω be a function defined on the Euclidean unit sphere S^{n-1} in $\mathbb{R}^n$ $(n \geq 2)$ and satisfying the cancellation condition

$$\int_{S^{n-1}} \Omega(\theta)\, d\theta = 0. \tag{1.1}$$

Denote by T_Ω the singular integral operator defined by

$$T_\Omega f(x) = \text{p.v.} \int_{\mathbb{R}^n} \frac{\Omega(y/|y|)}{|y|^n} f(x-y)\, dy$$

for f in the Schwartz class $\mathcal{S}(\mathbb{R}^n)$. These operators, known as Calderón–Zygmund singular integral operators, were first studied by Calderón and Zygmund in their famous article [CZ1]. They proved L^p boundedness of T_Ω when the kernels are regular. Later in 1960, Hörmander [Hor] showed that the same results hold when the kernel only satisfies a weaker condition which is referred to as *Hörmander's condition* today.

An important and interesting question is whether the regularity conditions on the convolution kernels are necessary for the L^p $(1 < p < \infty)$ boundedness of the Calderón–Zygmund singular integral operators. In 1956, Calderón and Zygmund [CZ2] gave a negative answer. Using the method of rotations, Calderón and Zygmund proved that T_Ω is still bounded on L^p for $1 < p < \infty$ when Ω is odd in $L^1(S^{n-1})$, or even in $L\log^+L(S^{n-1})$ satisfy-

2000 *Mathematics Subject Classification*: Primary 42B25; Secondary 43A80, 43A99.
Key words and phrases: Littlewood–Paley g-function, rough kernel, homogeneous groups, TT^* method.

DOI: 10.4064/sm195-1-4

ing (1.1). In 1988, Hofmann [Hof] proved that the rough singular integral operator T_Ω is of weak type $(1,1)$ in $\mathbb{R}^2$ for $\Omega \in L^q(S^1)$, $q > 1$. In an unpublished work, M. Christ obtained a weak type $(1,1)$ inequality for T_Ω for $\Omega \in L\log^+ L(S^{n-1})$ in dimension $n \le 7$. In 1996, Seeger [Se] showed that T_Ω is of weak type $(1,1)$ if $\Omega \in L\log^+ L(S^{n-1})$ with the mean zero condition (1.1) for all dimensions $n \ge 2$. In [GS], Grafakos and Stefanov gave a nice survey, which contains a thorough discussion of the history of the operator T_Ω.

It is well known that a *homogeneous group* $\mathbb{G}$ is a Lie group equipped with multiplication, inverse, dilation, and norm structures

$$(1.2)\qquad (x,y) \mapsto xy, \quad x \mapsto x^{-1}, \quad (t,x) \mapsto t \circ x, \quad x \mapsto \rho(x)$$

for $x, y \in \mathbb{G}$ and $t > 0$. Here the multiplication and inverse operations are polynomial maps, while the dilation structure preserves the group operations and is given in coordinates by

$$(1.3)\qquad t \circ (x_1, \ldots, x_n) = (t^{\alpha_1} x_1, \ldots, t^{\alpha_n} x_n)$$

for some constants $0 < \alpha_1 \le \cdots \le \alpha_n$ and satisfies $\rho(t \circ x) = t\rho(x)$. It can be shown that Lebesgue measure dx is a Haar measure on $\mathbb{G}$ and that $\rho(x) \sim \rho(x^{-1})$. We call n the *Euclidean dimension* of $\mathbb{G}$, and $N = \alpha_1 + \cdots + \alpha_n$ the *homogeneous dimension* of $\mathbb{G}$. Denote by $\Sigma = \{x \in \mathbb{G} : \rho(x) = 1\}$ the "unit sphere" of $\mathbb{G}$.

When Ω satisfies a much stronger smoothness condition, T_Ω is a bounded operator on $L^p(\mathbb{G})$ for $1 < p < \infty$ (see [St2]). By analyzing the results on rough singular integral operators T_Ω mentioned above, it is natural to conjecture that analogous results also hold on the homogeneous group $\mathbb{G}$. However, there exist many difficulties in this generalization. In fact, the method based on Fourier transform estimates is not available on the general homogeneous group $\mathbb{G}$, and this method plays a key role in studying the L^2 and weak $(1,1)$ boundedness for the rough operator T_Ω on $\mathbb{R}^n$. In 1999, using a variant of Littlewood–Paley theory and an iterated TT^* method, Tao [T] generalized the results in [CZ2] and [Se] to the homogeneous group $\mathbb{G}$. To be precise, Tao proved the weak type $(1,1)$ and $(2,2)$ boundedness of T with kernels belonging to the class $L\log^+ L$ on $\mathbb{G}$, and hence the (p,p) boundedness of T for $1 < p < \infty$ follows easily by interpolation and duality. Tao's work in [T] is very significant because the ideas presented there pave the way to the theory of rough operators on general homogeneous groups.

On the other hand, it is well known that the Littlewood–Paley operators play an important role in harmonic analysis on $\mathbb{R}^n$, PDE, characterizing function spaces, etc. The Littlewood–Paley operators in high dimensions were first introduced by Stein in [St1]. If $\Omega \in L^1(S^{n-1})$ satisfies (1.1), set $\varphi(x) = \Omega(x/|x|)|x|^{-n+1}\chi_{\{|x|\le 1\}}(x)$ and $\varphi_t(x) = t^{-n}\varphi(x/t)$ for $t > 0$. Then the *Littlewood–Paley g-function* g_Ω with homogeneous kernel is defined by

$$g_\Omega(f)(x) = \Big(\int_0^\infty |\varphi_t * f(x)|^2 \, \frac{dt}{t} \Big)^{1/2}.$$

In [St1], Stein proved that if $\Omega \in \mathrm{Lip}_\alpha(S^{n-1}), 0 < \alpha \le 1$, then g_Ω is of weak type $(1,1)$ and of type (p,p) for $1 < p \le 2$. In 1962, Benedek, Calderón and Panzone [BCP] showed if $\Omega \in C^1(S^{n-1})$, then g_Ω is of type (p,p) for $1 < p < \infty$. In 2000, Ding, Fan and Pan [DFP2] proved that if $\Omega \in H^1(S^{n-1})$, the Hardy space on S^{n-1} (see [Con] or [RW] for the definition), then g_Ω is still of type (p,p) for $1 < p < \infty$. In 2001, Fan and Sato [FaSa] established the weak $(1,1)$ boundedness of g_Ω if $\Omega \in L\log^+ L(S^{n-1})$. The results in [DFP2] and [FaSa] show that the regularity condition imposed on Ω is not necessary for the $L^p\,(1 < p < \infty)$ and weak $(1,1)$ boundedness of g_Ω. See also [DFP1] and [AACP] for more results about the rough Littlewood–Paley g-function on $\mathbb{R}^n$.

Inspired by Tao's pioneering work, in this paper, we will discuss some mapping properties of the Littlewood–Paley g-functions with rough kernels on the homogeneous group $\mathbb{G}$.

Suppose that a function Ω defined on $\mathbb{G}$ satisfies the following conditions:

$$\Omega(t \circ x) = \Omega(x) \quad \text{for } t > 0 \text{ and } x \in \mathbb{G}, \tag{1.4}$$

$$\int_\Sigma \Omega(x)\, d\sigma(x) = 0, \tag{1.5}$$

and $\Omega \in L^1(\Sigma)$, that is,

$$\int_\Sigma |\Omega(x)|\, d\sigma(x) < \infty, \tag{1.6}$$

where σ is a Radon measure on Σ (see [FoSt, p. 14]). The Littlewood–Paley g-function on $\mathbb{G}$ is defined by

$$g_\Omega f(x) = \Big(\int_0^\infty |f * \kappa_t|^2 \, \frac{dt}{t} \Big)^{1/2},$$

where $\kappa_t(x) = t^{-1}\Omega(x)\rho(x)^{1-N}\chi_{\{\rho(x)\le t\}}(x)$ $(t > 0,\ x \in \mathbb{G})$ and Ω satisfies the conditions (1.4)–(1.6).

In this paper, we will prove the weak type $(1,1)$ and (p,p) boundedness of the Littlewood–Paley g-function on the homogeneous group $\mathbb{G}$ if the size condition (1.6) is replaced by the following $L\log^+ L(\Sigma)$ condition:

$$\int_\Sigma |\Omega(x) \log(2 + \Omega(x))|\, d\sigma(x) < \infty. \tag{1.7}$$

Our main results are as follows.

Theorem 1.1. *If $\Omega \in L\log^+ L(\Sigma)$ satisfies* (1.4) *and* (1.5), *then g_Ω is bounded on $L^2(\mathbb{G})$.*

Theorem 1.2. *If $\Omega \in L\log^+ L(\Sigma)$ and g_Ω is bounded on $L^2(\mathbb{G})$, then g_Ω is of weak type $(1,1)$ on $\mathbb{G}$.*

By an interpolation theorem, we get

Corollary 1.3. *Suppose $\Omega \in L\log^+ L(\Sigma)$ satisfies* (1.4) *and* (1.5). *Then g_Ω is bounded on $L^p(\mathbb{G})$ for $1 < p \le 2$ and is of weak type $(1,1)$ on $\mathbb{G}$.*

The basic idea of proving our theorems is to view the Littlewood–Paley g-function as a vector-valued singular integral (see Section 2). It should be pointed out that some ideas used in this work are borrowed from Tao's paper [T]. However, the extra integral in t causes some essential difficulties so that we do not use the method in [T] directly in many estimates. Here, we point out *three major differences*:

(1) The singular integral operator T and its adjoint operator T^* are essentially the same. Hence, the properties of T can be translated to T^*. However, the Littlewood–Paley g-function g_Ω and its adjoint g_Ω^* are essentially different, since g_Ω maps a scalar-valued function to a Hilbert-valued function and its adjoint operator g_Ω^* is a mapping from a Hilbert-valued function space to a scalar-valued function space.

(2) In the proof of Theorem 1.1, we view the Littlewood–Paley g-function as a vector-valued singular integral. However, in estimating the nondegenerate portion of the integral, there are some additional difficulties caused by the extra t integral coming from vector-valued duality (see Remark 4.4 at the end of Section 4).

To avoid this problem, we iterate T $m+1$ times instead of m times, and then "throw out" the extra t integral to reduce the estimate for a vector-valued integral to a scalar-valued one.

In the weak $(1,1)$ case, we shall iterate T $m+2$ times for symmetry considerations. Also more careful estimates are required for the reduction (see Section 7).

(3) In the proof of weak type $(1,1)$ boundedness, we are led to estimate the derivatives of homogeneous norms. We show that any homogeneous norm on the homogeneous group satisfies certain regularity conditions (see Theorem 7.1). Then we get the desired estimates by using left-invariant differentiation structures developed by Tao (see Lemma 7.2).

Remark 1.4. In 1960, Hörmander [Hor] proved the L^p boundedness of a parameterized Littlewood–Paley g-function. Using our argument, one can prove results similar to ours for the parameterized Littlewood–Paley

g-function $g_{\Omega,\nu}$ defined by

$$g_{\Omega,\nu} f(x) = \Big(\int_0^\infty |f * \kappa_t^\nu|^2 \,\frac{dt}{t}\Big)^{1/2},$$

where $\kappa_t^\nu(x) = t^{-\nu}\Omega(x)\rho(x)^{\nu-N}\chi_{\{\rho(x)\le t\}}(x)$ for $t,\nu>0$ and $x\in\mathbb{G}$, and Ω satisfies the conditions (1.4)–(1.6).

Throughout this paper, we will work exclusively with real-valued functions. The letters C (resp. c,ε) will always be used to denote large (resp. small) positive constants only depending on the homogeneous group $\mathbb{G}$ and any other specified quantities. The values of these constants are not necessarily the same at each occurrence. We use $A\lesssim B$ to denote $A\le CB$, and we write $A\sim B$ if $A\lesssim B\lesssim A$.

2. L^2 estimate I: Kernel truncation and frequency localization. Let $K(x)=\Omega(x)\rho(x)^{-N}$ and $h_t(x)=t^{-1}\rho(x)\chi_{\{\rho(x)\le t\}}(x)$. Set $K_0 := K\chi_{A_0}$ with $A_0 = \{x\in\mathbb{G} : 1\le\rho(x)\le 2\}$. Then we may normalize $\|K_0\|_{L\log^+ L(\mathbb{G})} = 1$ since $\Omega\in L\log^+ L(\Sigma)$. Let $\mathcal{H}$ be the Hilbert space $L^2(\mathbb{R}_+, dt/t)$ with the norm and inner product denoted by $|\cdot|_{\mathcal{H}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{H}}$, respectively. Thus, the Littlewood–Paley g-function g_Ω can be written as

$$g_\Omega(f)(x) = \Big(\int_0^\infty |f*(h_tK)(x)|^2\,\frac{dt}{t}\Big)^{1/2} = |f*(h_{(\cdot)}K)(x)|_{\mathcal{H}}. \tag{2.1}$$

For each $u>0$, define the scaling map $\Delta[u]$ by $\Delta[u]f(y) = u^{-N}f(u^{-1}\circ y)$. Then we have the identity

$$K(x) = \frac{1}{\log 2}\int_0^\infty \Delta[u]K_0(x)\,du. \tag{2.2}$$

Take a nonnegative function $\varphi\in C_0^\infty(\mathbb{R})$ such that $\mathrm{supp}(\varphi)\subset(1/4,4)$, $\varphi(u)\equiv 1$ for $1/2\le u\le 2$ and $\sum_{j\in\mathbb{Z}} 2^{-j}\varphi(2^{-j}u) = 1/\log 2$. For $j\in\mathbb{Z}$, define an operator S_j by

$$S_jF(x) = 2^{-j}\int_0^\infty \varphi(2^{-j}u)\Delta[u]F(x)\,du.$$

Then by (2.2), we have a dyadic decomposition $K=\sum_{j\in\mathbb{Z}} S_jK_0$. It is easy to see that

$$\|S_jf\|_{L^1(\mathbb{G})} \lesssim \|f\|_{L^1(\mathbb{G})} \quad \text{uniformly in } j \tag{2.3}$$

and

$$|h_{(\cdot)}(x)|_{\mathcal{H}} \lesssim 1 \quad \text{uniformly in } x\in\mathbb{G}. \tag{2.4}$$

We now define the $\mathcal{H}$-valued L^q spaces. For $1 \le q < \infty$, let

$$L^q(\mathcal{H}) = \left\{ f(x,t) : \|f\|_{L^q(\mathcal{H})} := \left\{ \int_{\mathbb{G}} \left(\int_0^\infty |f(x,t)|^2 \,\frac{dt}{t} \right)^{q/2} dx \right\}^{1/q} < \infty \right\}.$$

In view of Minkowski's inequality and the classical Young inequality, the following version of Young's inequality on $L^q(\mathcal{H})$ is obvious.

Lemma 2.1. *Suppose* $1 \le p,\ q, r \le \infty$, $1/p + 1/q = 1 + 1/r$, $f \in L^q(\mathcal{H})$ *and* $g \in L^p(\mathbb{G})$. *Then*

$$\|f * g\|_{L^r(\mathcal{H})} \le C\|g\|_{L^p(\mathbb{G})}\|f\|_{L^q(\mathcal{H})}.$$

With the above notations, by (2.3) and (2.4) we have

$$\|h_t S_j f\|_{L^1(\mathcal{H})} \lesssim \|f\|_{L^1(\mathbb{G})} \quad \text{uniformly in } j. \tag{2.5}$$

To prove Theorem 1.1, by (2.1), it suffices to show that

$$\Big\| f * \sum_j (h_t S_j K_0) \Big\|_{L^2(\mathcal{H})} \lesssim \|f\|_{L^2(\mathbb{G})}. \tag{2.6}$$

For $s \ge 0$, define

$$A_0^s = \{x \in A_0 : 2^{2^s} \le 2 + |K_0(x)| < 2^{2^{s+1}}\}$$

and $k_0^s = K_0 \chi_{A_0^s}$. Let

$$K_0^s = k_0^s - \frac{\chi_{A_0}}{|A_0|} \int_{A_0} k_0^s(x)\,dx.$$

Then $K_0 = \sum_{s\ge 0} K_0^s$ and each K_0^s has mean zero. Note that

$$\sum_{s\ge 0} 2^s \|K_0^s\|_{L^1(\mathbb{G})} \lesssim \sum_{s\ge 0} 2^s \|k_0^s\|_{L^1(\mathbb{G})} \lesssim \|K_0\|_{L\log^+ L(\mathbb{G})} = 1.$$

Thus, (2.6) will follow if we can show that for some $\varepsilon > 0$ and all $s \ge 0$,

$$\Big\| f * \sum_j (h_t S_j K_0^s) \Big\|_{L^2(\mathcal{H})} \lesssim \|f\|_{L^2(\mathbb{G})} (2^s \|K_0^s\|_{L^1(\mathbb{G})} + 2^s 2^{-\varepsilon 2^s}). \tag{2.7}$$

Now we fix s. For each integer k, let T_k denote the operator

$$T_k f = f * \sum_{j=k2^s}^{(k+1)2^s - 1} (h_t S_j K_0^s).$$

To get (2.7), it is sufficient to show the following operator norm estimate:

$$\Big\| \sum_k T_k \Big\|_{L^2(\mathbb{G}) \to L^2(\mathcal{H})} \lesssim 2^s \|K_0^s\|_{L^1(\mathbb{G})} + 2^s 2^{-\varepsilon 2^s}. \tag{2.8}$$

By Lemma 2.1 and (2.5), we have

$$\|T_k f\|_{L^2(\mathcal{H})} \lesssim \|f\|_{L^2(\mathbb{G})} 2^s \|K_0^s\|_{L^1(\mathbb{G})} \quad \text{uniformly in } k.$$

Hence, T_k is a bounded operator from $L^2(\mathbb{G})$ to $L^2(\mathcal{H})$. Then the adjoint operator T_k^* of T_k, defined by

$$T_k^* g_t(x) = \sum_{j=k2^s}^{(k+1)2^s-1} \int_0^\infty g_t * (h_t S_j \tilde{K}_0^s)(x)\, \frac{dt}{t},$$

is bounded from $L^2(\mathcal{H})$ to $L^2(\mathbb{G})$ with the same operator norm as T_k, where $\tilde{K}_0^s(x) = K_0^s(x^{-1})$ for each $x \in \mathbb{G}$. Thus for all k, k', we have

$$\max\{\|T_k T_{k'}^*\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})}, \|T_k^* T_{k'}\|_{L^2(\mathbb{G})\to L^2(\mathbb{G})}\} \lesssim (2^s \|K_0^s\|_{L^1(\mathbb{G})})^2.$$

Therefore, by the Cotlar–Stein lemma (see [St2]), to obtain the estimate (2.8), it suffices to show that there exists a large constant C such that when $|k-k'| \ge C$,

$$\max\left\{\|T_{k'} T_k^*\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})}, \|T_{k'}^* T_k\|_{L^2(\mathbb{G})\to L^2(\mathbb{G})}\right\} \lesssim 2^{2s} 2^{-\varepsilon 2^s |k-k'|}. \tag{2.9}$$

By the definitions of $T_{k'}$ and T_k^*, we need to prove that

$$\Big\| \sum_{j=k2^s}^{(k+1)2^s-1} \sum_{j'=k'2^s}^{(k'+1)2^s-1} \int_0^\infty g_{t_1} * (h_{t_1} S_j \tilde{K}_0^s) * (h_{t_2} S_{j'} K_0^s)\, \frac{dt_1}{t_1} \Big\|_{L^2(\mathcal{H})} \lesssim 2^{2s} 2^{-\varepsilon 2^s |k-k'|} \|g_{t_1}\|_{L^2(\mathcal{H})},$$

$$\Big\| \sum_{j=k2^s}^{(k+1)2^s-1} \sum_{j'=k'2^s}^{(k'+1)2^s-1} \int_0^\infty f * (h_t S_j K_0^s) * (h_t S_{j'} \tilde{K}_0^s)\, \frac{dt}{t} \Big\|_{L^2(\mathbb{G})} \lesssim 2^{2s} 2^{-\varepsilon 2^s |k-k'|} \|f\|_{L^2(\mathbb{G})}.$$

By the triangle inequality, it suffices to show that for all integers j, j' with $|j-j'| > C2^s$,

$$\Big\| \int_0^\infty g_{t_1} * (h_{t_1} S_j \tilde{K}_0^s) * (h_{t_2} S_{j'} K_0^s)\, \frac{dt_1}{t_1} \Big\|_{L^2(\mathcal{H})} \lesssim 2^{-\varepsilon|j-j'|} \|g_{t_1}\|_{L^2(\mathcal{H})}, \tag{2.10}$$

$$\Big\| \int_0^\infty f * (h_t S_j K_0^s) * (h_t S_{j'} \tilde{K}_0^s)\, \frac{dt}{t} \Big\|_{L^2(\mathbb{G})} \lesssim 2^{-\varepsilon|j-j'|} \|f\|_{L^2(\mathbb{G})}. \tag{2.11}$$

Now we use the Littlewood–Paley theory to show (2.10) and (2.11). Take a C^∞ function ϕ supported on the unit ball with $\|\phi\|_{C^1} \lesssim 1$ and $\int \phi = 1$. We may also assume that $\phi = \widetilde{\phi}$, where $\widetilde{\phi}(x) = \phi(x^{-1})$ for $x \in \mathbb{G}$. For each integer k, write

$$\psi_k = \Delta[2^{k-1}]\phi - \Delta[2^k]\phi.$$

Note that ψ_k is supported on the annulus of radius $C2^k$, that is, on the set $\{x \in \mathbb{G} : C2^{k-1} \le \rho(x) \le C2^{k+1}\}$ for some absolute constant C. Moreover,

ψ_k has mean zero and $\widetilde{\psi}_k = \psi_k$. Since

$$\sum_k \psi_k * f(x) = f(x) = \sum_k f * \psi_k(x)$$

for $x \neq 0$, we may write

$$f*(h_t S_j K_0^s)*(h_t S_{j'} \tilde{K}_0^s) = \sum_k \sum_{k'} f*(h_t S_j K_0^s)*\psi_k*\psi_{k'}*(h_t S_{j'} \tilde{K}_0^s). \tag{2.12}$$

We need the following lemma to prove (2.10) and (2.11).

Lemma 2.2. *For any integers j, k and any $L^\infty(\mathbb{G})$ function K_0 on the unit annulus with mean zero, we have*

$$\|f * (h_t S_j K_0) * \psi_k\|_{L^2(\mathcal{H})} \lesssim 2^{-\varepsilon|j-k|} \|f\|_{L^2(\mathbb{G})} \|K_0\|_{L^\infty(\mathbb{G})} \tag{2.13}$$

and

$$\|f * \psi_k * (h_t S_j K_0)\|_{L^2(\mathcal{H})} \lesssim 2^{-\varepsilon|j-k|} \|f\|_{L^2(\mathbb{G})} \|K_0\|_{L^\infty(\mathbb{G})}. \tag{2.14}$$

The proof of Lemma 2.2 will be postponed until the next section. Now let us complete the proof of (2.10) and (2.11) by applying Lemma 2.2. In fact, by (2.13), we have

$$\|f * (h_t S_j K_0^s) * \psi_k\|_{L^2(\mathcal{H})} \lesssim 2^{2^{s+1}} 2^{-\varepsilon|j-k|} \|f\|_{L^2(\mathbb{G})} \tag{2.15}$$

and (by duality)

$$\Big\| \int_0^\infty g_t * \psi_{k'} * (h_t S_{j'} \tilde{K}_0^s) \, \frac{dt}{t} \Big\|_{L^2(\mathbb{G})} \lesssim 2^{2^{s+1}} 2^{-\varepsilon|k'-j'|} \|g_t\|_{L^2(\mathcal{H})}. \tag{2.16}$$

Using the estimates (2.15) and (2.16), we get

$$\begin{aligned} \Big\| \int_0^\infty f * (h_t S_j K_0^s) * \psi_k * \psi_{k'} * (h_t S_{j'} \tilde{K}_0^s) \, \frac{dt}{t} \Big\|_{L^2(\mathbb{G})} & \\ \lesssim 2^{2^{s+2}} 2^{-\varepsilon|j-k|} 2^{-\varepsilon|k'-j'|} \|f\|_{L^2(\mathbb{G})}&. \end{aligned} \tag{2.17}$$

Similarly, it follows from (2.14) that

$$\begin{aligned} \Big\| \int_0^\infty f_{t_1} * (h_{t_1} S_j \tilde{K}_0^s) * \psi_k * \psi_{k'} * (h_{t_2} S_{j'} K_0^s) \, \frac{dt_1}{t_1} \Big\|_{L^2(\mathcal{H})} & \\ \lesssim 2^{2^{s+2}} 2^{-\varepsilon|j-k|} 2^{-\varepsilon|k'-j'|} \|f_{t_1}\|_{L^2(\mathcal{H})}&. \end{aligned} \tag{2.18}$$

On the other hand, from the smoothness and mean zero conditions on ψ_k, $\psi_{k'}$, we have

$$\|\psi_k * \psi_{k'}\|_{L^1(\mathbb{G})} \lesssim 2^{-\varepsilon|k-k'|}. \tag{2.19}$$

By Hölder's inequality and (2.4),

$$|\langle h_{(\cdot)}(x), h_{(\cdot)}(y) \rangle_{\mathcal{H}}| \le |h_{(\cdot)}(x)|_{\mathcal{H}} \cdot |h_{(\cdot)}(y)|_{\mathcal{H}} \lesssim 1$$

uniformly in $x, y \in \mathbb{G}$. This estimate together with Young's inequality, (2.19) and (2.3) yields

$$\begin{aligned}
(2.20)\qquad & \Big\| \int_0^\infty f * (h_t S_j K_0^s) * \psi_k * \psi_{k'} * (h_t S_{j'} \tilde{K}_0^s)\, \frac{dt}{t} \Big\|_{L^2(\mathbb{G})} \\
&\lesssim \Big\| \int_0^\infty (h_t S_j K_0^s) * \psi_k * \psi_{k'} * (h_t S_{j'} \tilde{K}_0^s)\, \frac{dt}{t} \Big\|_{L^1(\mathbb{G})} \|f\|_{L^2(\mathbb{G})} \\
&\lesssim \| |S_j K_0^s| * |\psi_k * \psi_{k'}| * |S_{j'} \tilde{K}_0^s| \|_{L^1(\mathbb{G})} \|f\|_{L^2(\mathbb{G})} \\
&\lesssim \|S_j K_0^s\|_{L^1(\mathbb{G})} \|\psi_k * \psi_{k'}\|_{L^1(\mathbb{G})} \|S_{j'} \tilde{K}_0^s\|_{L^1(\mathbb{G})} \|f\|_{L^2(\mathbb{G})} \\
&\lesssim 2^{-\varepsilon|k-k'|} \|f\|_{L^2(\mathbb{G})}.
\end{aligned}$$

By Minkowski's inequality and (2.4), we get

$$\begin{aligned}
(2.21)\qquad & \Big\| \int_0^\infty f_{t_1} * (h_{t_1} S_j \tilde{K}_0^s) * \psi_k * \psi_{k'} * (h_{t_2} S_{j'} K_0^s)\, \frac{dt_1}{t_1} \Big\|_{L^2(\mathcal{H})} \\
&\lesssim \Big\| \Big| \int_0^\infty f_{t_1} * (h_{t_1} S_j \tilde{K}_0^s)\, \frac{dt_1}{t_1} \Big| * |\psi_k * \psi_{k'}| * |S_{j'} K_0^s| \Big\|_{L^2(\mathbb{G})}.
\end{aligned}$$

By the Fubini theorem and Hölder's inequality, for every $x \in \mathbb{G}$,

$$\begin{aligned}
\Big| \int_0^\infty f_{t_1} * (h_{t_1} S_j \tilde{K}_0^s)(x)\, \frac{dt_1}{t_1} \Big| &\le \int_{\mathbb{G}} \Big| \int_0^\infty f_{t_1}(xy^{-1}) h_{t_1}(y)\, \frac{dt_1}{t_1} \Big| |S_j \tilde{K}_0^s(y)(y)|\, dy \\
&\le \int_{\mathbb{G}} |f_{t_1}(xy^{-1})|_{\mathcal{H}} |S_j \tilde{K}_0^s(y)|\, dy \\
&= (|f_{t_1}|_{\mathcal{H}} * |S_j \tilde{K}_0^s|)(x).
\end{aligned}$$

Inserting this estimate into (2.21) and using Young's inequality, (2.3) and (2.19), we get

$$\begin{aligned}
(2.22)\qquad & \Big\| \int_0^\infty f_{t_1} * (h_{t_1} S_j \tilde{K}_0^s) * \psi_k * \psi_{k'} * (h_{t_2} S_{j'} K_0^s)\, \frac{dt_1}{t_1} \Big\|_{L^2(\mathcal{H})} \\
&\lesssim \Big\| \Big(\int_0^\infty |f_{t_1}(\cdot)|^2\, \frac{dt_1}{t_1} \Big)^{1/2} * |S_j \tilde{K}_0^s| * |\psi_k * \psi_{k'}| * |S_{j'} K_0^s| \Big\|_{L^2(\mathbb{G})} \\
&\lesssim \|S_j \tilde{K}_0^s\|_{L^1(\mathbb{G})} \|\psi_k * \psi_{k'}\|_{L^1(\mathbb{G})} \|S_{j'} K_0^s\|_{L^1(\mathbb{G})} \|f_{t_1}\|_{L^2(\mathcal{H})} \\
&\lesssim 2^{-\varepsilon|k-k'|} \|f_{t_1}\|_{L^2(\mathcal{H})}.
\end{aligned}$$

Taking the geometric mean of (2.17) and (2.20) and then summing over k and k', we obtain

$$\Big\|\int_0^\infty f*(h_tS_jK_0)*(h_tS_{j'}\tilde{K}_0)\,\frac{dt}{t}\Big\|_{L^2(\mathbb{G})}\lesssim 2^{2^{s+1}}|j-j'|2^{-\varepsilon|j-j'|/2}\|f\|_{L^2(\mathbb{G})},$$

which gives (2.11) for some $\varepsilon>0$ if $|j-j'|>C2^s$ for a sufficiently large C. In the same way we can deduce (2.10) from (2.18) and (2.22). Thus, to finish the proof of Theorem 1.1, it remains to show Lemma 2.2.

3. L^2 estimate II: Iterated TT^* method. The proof of Lemma 2.2 will be given in this section. We only verify (2.13) here since (2.14) can be proved in a similar way. First we normalize K_0 so that $\|K_0\|_\infty=1$. By dilation invariance,

$$\|(h_tS_jK_0)*\psi_k\|_{L^1(\mathcal{H})}=\|(h_tS_0K_0)*\psi_{k-j}\|_{L^1(\mathcal{H})},$$

so we only need to show (2.13) for $j=0$. If $k\geq -C$, then by the mean zero condition on K_0 and the smoothness of ψ_k, we have

$$\|(h_tS_0K_0)*\psi_k\|_{L^1(\mathcal{H})}\lesssim 2^{-\varepsilon k}. \tag{3.1}$$

By Young's inequality (Lemma 2.1) and (3.1), we get (2.13) for $k\geq -C$, where C is a constant large enough to be determined later. Thus in the following we may assume that $k<-C$.

Fix $k=-s$ for some $s>C$. If we define an operator L_ψ by

$$L_\psi f(x,t)=f*(h_tS_0K_0)*\psi_{-s},$$

then it is easy to see that L_ψ is a bounded operator from $L^2(\mathbb{G})$ to $L^2(\mathcal{H})$. Denote by L_ψ^* the adjoint operator of L_ψ, that is, for $g_t\in L^2(\mathcal{H})$,

$$L_\psi^*(g_t)(x)=\int_0^\infty g_t*\psi_{-s}*(h_tS_0\tilde{K}_0)\,\frac{dt}{t}.$$

Thus, (2.13) will follow from the following estimates:

$$\begin{aligned}&\|L_\psi L_\psi^*(g_{t_0})\|_{L^2(\mathcal{H})}\\ &\qquad=\Big\|\Big(\int_0^\infty g_{t_0}*\psi_{-s}*(h_{t_0}S_0\tilde{K}_0)\,\frac{dt_0}{t_0}\Big)*(h_{t_1}S_0K_0)*\psi_{-s}\Big\|_{L^2(\mathcal{H})}\\ &\qquad\lesssim 2^{-\varepsilon s}\|g_{t_0}\|_{L^2(\mathcal{H})}\end{aligned}$$

and

$$\begin{aligned}\|L_\psi^*L_\psi(f)\|_{L^2(\mathbb{G})}&=\Big\|\int_0^\infty f*\psi_{-s}*(h_tS_0\tilde{K}_0)*(h_tS_0K_0)*\psi_{-s}\,\frac{dt}{t}\Big\|_{L^2(\mathbb{G})}\\ &\lesssim 2^{-\varepsilon s}\|f\|_{L^2(\mathbb{G})}.\end{aligned}$$

From the operator norm identity $\|L_\psi L_\psi^*\|=\|(L_\psi L_\psi^*)^{n+1}\|^{1/(n+1)}$, it suffices to show

$$\begin{aligned}(3.2)\qquad &\Big\|\int_0^\infty \cdots \int_0^\infty f_{t_0} * \psi_{-s} * (h_{t_0} S_0 \tilde{K}_0)\,\frac{dt_0}{t_0} * (h_{t_1} S_0 K_0) * \psi_{-s} \\ &\quad * \psi_{-s} * (h_{t_1} S_0 \tilde{K}_0)\,\frac{dt_1}{t_1} * (h_{t_2} S_0 K_0) * \psi_{-s} \\ &\quad \cdots \\ &\quad * \psi_{-s} * (h_{t_n} S_0 \tilde{K}_0)\,\frac{dt_n}{t_n} * (h_{t_{n+1}} S_0 K_0) * \psi_{-s}\Big\|_{L^2(\mathcal{H})} \\ &\quad \lesssim 2^{-\varepsilon s}\|f_{t_0}\|_{L^2(\mathcal{H})}\end{aligned}$$

and

$$\begin{aligned}(3.3)\qquad &\Big\|\int_0^\infty \cdots \int_0^\infty f * \psi_{-s} * (h_{t_1} S_0 \tilde{K}_0) * (h_{t_1} S_0 K_0) * \psi_{-s}\,\frac{dt_1}{t_1} \\ &\quad * \psi_{-s} * (h_{t_2} S_0 \tilde{K}_0) * (h_{t_2} S_0 K_0) * \psi_{-s}\,\frac{dt_2}{t_2} \\ &\quad \cdots \\ &\quad * \psi_{-s} * (h_{t_n} S_0 \tilde{K}_0) * (h_{t_n} S_0 K_0) * \psi_{-s}\,\frac{dt_n}{t_n}\Big\|_{L^2(\mathbb{G})} \\ &\quad \lesssim 2^{-\varepsilon s}\|f\|_{L^2(\mathbb{G})}.\end{aligned}$$

By Young's inequality, it is easy to see that (3.3) follows from

$$\begin{aligned}(3.4)\qquad &\Big\|\int_0^\infty \cdots \int_0^\infty \psi_{-s} * (h_{t_1} S_0 K_0) * (h_{t_1} S_0 \tilde{K}_0) * \psi_{-s}\,\frac{dt_1}{t_1} * \psi_{-s} * (h_{t_2} S_0 K_0) \\ &\quad * \cdots * \psi_{-s} * (h_{t_n} S_0 \tilde{K}_0) * (h_{t_n} S_0 K_0) * \psi_{-s}\,\frac{dt_n}{t_n}\Big\|_{L^1(\mathbb{G})} \lesssim 2^{-\varepsilon s}.\end{aligned}$$

Next, we want to show that (3.2) will follow from an $L^1(\mathbb{G})$ norm (not $L^1(\mathcal{H})$ norm) estimate similar to (3.4). This is a key step since if we worked with a vector-valued integral ($L^1(\mathcal{H})$ estimate), some essential difficulties would arise (see Remark 4.4 in the next section).

By Hölder's inequality,

$$\begin{aligned}\Big\|\int_0^\infty f_{t_0} * \psi_{-s} * (h_{t_0} S_0 \tilde{K}_0)\,\frac{dt_0}{t_0}\Big\|_{L^2(\mathbb{G})} &\le \|\,|f|_{\mathcal{H}} * |\psi_{-s}| * |h_{(\cdot)} S_0 \tilde{K}_0|_{\mathcal{H}}\|_{L^2(\mathbb{G})} \\ &\le \|f\|_{L^2(\mathcal{H})}\|\psi_{-s}\|_{L^1(\mathbb{G})}\|h_{(\cdot)} S_0 \tilde{K}_0\|_{L^1(\mathcal{H})} \\ &\lesssim \|f\|_{L^2(\mathcal{H})}.\end{aligned}$$

Thus, on applying Young's inequality again, (3.2) will follow from

$$(3.5)\qquad \Big\| \int_0^\infty \dots \int_0^\infty (h_{t_1} S_0 K_0) * \psi_{-s} * \psi_{-s} * (h_{t_1} S_0 \tilde{K}_0) \frac{dt_1}{t_1} * (h_{t_2} S_0 K_0) * \psi_{-s} * \dots * \psi_{-s} * (h_{t_n} S_0 \tilde{K}_0) \frac{dt_n}{t_n} * (h_{t_{n+1}} S_0 K_0) * \psi_{-s} \Big\|_{L^1(\mathcal{H})} \lesssim 2^{-\varepsilon s}.$$

Noting that $h_{t_{n+1}} S_0 K_0 * \psi_{-s} \in L^1(\mathcal{H})$, by Minkowski's inequality we see that (3.5) is implied by

$$(3.6)\qquad \Big\| \int_0^\infty \dots \int_0^\infty (h_{t_1} S_0 K_0) * \psi_{-s} * \psi_{-s} * (h_{t_1} S_0 \tilde{K}_0) \frac{dt_1}{t_1} * \dots * (h_{t_n} S_0 K_0) * \psi_{-s} * \psi_{-s} * (h_{t_n} S_0 \tilde{K}_0) \frac{dt_n}{t_n} \Big\|_{L^1(\mathbb{G})} \lesssim 2^{-\varepsilon s}.$$

Thus, to get (3.2) and (3.3), it remains to show (3.4) and (3.6).

Now we will use the following idea. Let ϕ be an L^1 function supported in B. Then

$$|\phi * f(x)| \le \|\phi\|_{L^1} \sup_{w \in B} |\delta_w * f(x)|,$$

where δ_w is the Dirac measure supported at w. Since the functions $\psi_{-s} * S_0 \tilde{K}_0$ and $\psi_{-s} * \psi_{-s} * S_0 \tilde{K}_0$ are bounded in $L^1(B(0,C))$, we have

$$|\psi_{-s} * (h_t S_0 \tilde{K}_0) * f(x)| \lesssim \sup_{w \in B(0,C),\, \xi \in A_0} |\delta_w * f(x) h_t(\xi)|.$$

Thus to show (3.6), it suffices to prove

$$(3.7)\qquad \Big\| \int_{[0,\infty)^n} ((h_{t_1} S_0 K_0) * \delta_{w_2} * \dots * \delta_{w_n} * (h_{t_n} S_0 K_0) * \psi_{-s} * \delta_{w_{n+1}}) \prod_{m=1}^n h_{t_m}(\xi_m) \frac{dt}{t} \Big\|_{L^1(\mathbb{G})} \lesssim 2^{-\varepsilon s}$$

uniformly for $(w_2, \dots, w_{n+1}) \in (B(0,C))^n$ and $(\xi_1, \dots, \xi_n) \in A_0^n$, where δ_{w_j} is the Dirac measure supported at w_j for $j = 2, \dots, n+1$. Moreover, $\frac{dt}{t} = \frac{dt_1}{t_1} \otimes \dots \otimes \frac{dt_n}{t_n}$. In the same way, to show (3.4), it suffices to verify

$$(3.8)\qquad \Big\| \int_{[0,\infty)^n} (\delta_{w_1} * (h_{t_1} S_0 K_0) * \delta_{w_2} * \dots * \delta_{w_n} * (S_0 K_0 h_{t_n}) * \psi_{-s}) \prod_{m=1}^n h_{t_m}(\xi_m) \frac{dt}{t} \Big\|_{L^1(\mathbb{G})} \lesssim 2^{-\varepsilon s}$$

uniformly for $(w_1, \dots, w_n) \in (B(0,C))^n$ and $(\xi_1, \dots, \xi_n) \in A_0^n$.

Below we only give the proof of (3.7), since (3.8) can be proved similarly. Fix $\vec{w} := (w_2, \ldots, w_{n+1}) \in (B(0,C))^n$ and $\vec{\xi} := (\xi_1, \ldots, \xi_n) \in A_0^n$. By duality, it suffices to show that the quantity

$$\Big|\Big\langle \int\limits_{[0,\infty)^n} ((h_{t_1} S_0 K_0) * \cdots * \delta_{w_n} * (h_{t_n} S_0 K_0) * \psi_{-s} * \delta_{w_{n+1}}) \prod_{m=1}^{n} h_{t_m}(\xi_m) \frac{dt}{t}, g\Big\rangle\Big| \tag{3.9}$$

is $\lesssim 2^{-\varepsilon s}$ for all test functions g with $\|g\|_{L^\infty(\mathbb{G})} \le 1$. For a fixed g, performing the t integration and taking $w_1 \in B(0,C)$, we can rewrite (3.9) as

$$\Big|\iiint \psi_{-s}(x) g(w_1^{-1} \Phi_{\vec{y}}(\vec{u}) x w_{n+1}) \times \prod_{i=1}^{n} \varphi(u_i) K_0(y_i) \min\Big\{\frac{u_i \rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i \rho(y_i)}\Big\} \, dx\, d\vec{u}\, d\vec{y}\Big|, \tag{3.10}$$

where $x \in \mathbb{G}$, $\vec{u} := (u_1, \ldots, u_n) \in [-C, C]^n$, $\vec{y} := (y_1, \ldots, y_n) \in A_0^n$ and

$$\Phi_{\vec{y}}(\vec{u}) := \prod_{i=1}^{n} w_i(u_i \circ y_i).$$

To see this, we only check this equality for $n = 2$. In this case, it is easy to see that (3.9) equals

$$\iint\limits_{[0,\infty)^n} \Big\{\int\limits_{\mathbb{G}} \psi_{-s}(x)((h_{t_2} S_0 \tilde{K}_0) * \delta_{w_2^{-1}} * (h_{t_1} S_0 \tilde{K}_0) * g * \delta_{w_3^{-1}})(x)\, dx\Big\} \times h_{t_1}(\xi_1) h_{t_2}(\xi_2) \frac{dt_1\, dt_2}{t_1 t_2}. \tag{3.11}$$

Expanding the first two convolutions, we get

$$\begin{aligned}
&(h_{t_2} S_0 \tilde{K}_0) * \delta_{w_2^{-1}} * (h_{t_1} S_0 \tilde{K}_0) * g * \delta_{w_3^{-1}}(x) \\
&\quad = \int\limits_{A_0} h_{t_2}(y_2) S_0 \tilde{K}_0(y_2) (\delta_{w_2^{-1}} * (h_{t_1} S_0 \tilde{K}_0) * g * \delta_{w_3^{-1}})(y_2^{-1} x)\, dy_2 \\
&\quad = \int\limits_{A_0} h_{t_2}(y_2) S_0 \tilde{K}_0(y_2) ((h_{t_1} S_0 \tilde{K}_0) * g * \delta_{w_3^{-1}})(w_2 y_2^{-1} x)\, dy_2.
\end{aligned}$$

Recalling

$$S_0 \tilde{K}_0(x) = \int_0^\infty \varphi(s) s^{-N} \tilde{K}_0(s^{-1} \circ x)\, ds$$

and $\tilde{K_0}(x) = K_0(x^{-1})$, we have

$$\begin{aligned}
&(h_{t_2} S_0 \tilde{K_0}) * \delta_{w_2^{-1}} * (h_{t_1} S_0 \tilde{K_0}) * g * \delta_{w_3^{-1}}(x) \\
&\qquad = \int\limits_{A_0} h_{t_2}(u_2 \circ y_2) \Big(\int\limits_0^\infty \varphi(u_2) \tilde{K_0}(y_2)\, du_2 \Big) \\
&\qquad\qquad \times ((h_{t_1} S_0 \tilde{K_0}) * g * \delta_{w_3^{-1}})(w_2 (u_2 \circ y_2^{-1}) x)\, dy_2 \\
&\qquad = \int\limits_{A_0} \int\limits_0^\infty h_{t_2}(u_2 \circ y_2^{-1}) \varphi(u_2) K_0(y_2) \\
&\qquad\qquad \times ((h_{t_1} S_0 \tilde{K_0}) * g * \delta_{w_3^{-1}})(w_2 (u_2 \circ y_2) x)\, du_2\, dy_2.
\end{aligned}$$

Similarly, we can expand the other two convolutions to obtain

$$\begin{aligned}
(3.12)\qquad &(h_{t_2} S_0 \tilde{K_0}) * \delta_{w_2^{-1}} * (h_{t_1} S_0 \tilde{K_0}) * g * \delta_{w_3^{-1}}(x) \\
&= \int\limits_{A_0 \times A_0} \int\limits_0^\infty \int\limits_0^\infty h_{t_2}(u_2 \circ y_2^{-1}) h_{t_1}(u_1 \circ y_1^{-1}) \varphi(u_2) \varphi(u_1) K_0(y_2) K_0(y_1) \\
&\qquad\qquad \times g((u_1 \circ y_1) w_2 (u_2 \circ y_2) x w_3)\, du_1\, du_2\, dy_1\, dy_2.
\end{aligned}$$

Inserting (3.12) into (3.11), we get (3.10) by noting that

$$\begin{aligned}
&\int\limits_0^\infty \int\limits_0^\infty h_{t_1}(u_1 \circ y_1) h_{t_2}(u_2 \circ y_2) h_{t_1}(\xi_1) h_{t_2}(\xi_2)\, \frac{dt_1}{t_1}\, \frac{dt_2}{t_2} \\
&\qquad = \Big(\int\limits_0^\infty h_{t_1}(u_1 \circ y_1) h_{t_1}(\xi_1)\, \frac{dt_1}{t_1} \Big) \Big(\int\limits_0^\infty h_{t_2}(u_2 \circ y_2) h_{t_2}(\xi_2)\, \frac{dt_2}{t_2} \Big) \\
&\qquad = \frac{u_1 \rho(y_1) \rho(\xi_1)}{2(\max\{u_1 \rho(y_1), \rho(\xi_1))\}^2} \cdot \frac{u_2 \rho(y_2) \rho(\xi_2)}{2(\max\{u_2 \rho(y_2), \rho(\xi_2)\}^2} \\
&\qquad = \frac{1}{4} \min\Big\{ \frac{u_1 \rho(y_1)}{\rho(\xi_1)}, \frac{\rho(\xi_1)}{u_1 \rho(y_1)} \Big\} \cdot \min\Big\{ \frac{u_2 \rho(y_2)}{\rho(\xi_2)}, \frac{\rho(\xi_2)}{u_2 \rho(y_2)} \Big\}.
\end{aligned}$$

The general case can be obtained by iterating the above process.

The next step is to split the integral in (3.10) into two parts. We need the left-invariant differentiation structures exploited by Tao in [T]. Let $f(t)$ be smooth functions from $\mathbb{R}$ to $\mathbb{G}$. The left-invariant derivative $\partial_t^L f(t)$ is defined by Newton's approximation:

$$f(t + \varepsilon) = f(t)(\varepsilon \partial_t^L f(t)) + \varepsilon^2 O(1), \qquad \text{for } \varepsilon \text{ small.}$$

If $F(x)$ is a smooth function from $\mathbb{R}^n$ to $\mathbb{G}$, the left-invariant derivative $D_x^L F(x)$ is defined to be the matrix with columns given by

$$D_x^L f(x) = (\partial_{x_1}^L F(x), \dots, \partial_{x_n}^L F(x)).$$

Using the above notions, we split the integral in (3.10) into

$$\Big|\int_{A_0^n}\int_{[C^{-1},C]^n}\int_{\mathbb{G}} \psi_{-s}(x)g(w_1^{-1}\Phi_{\vec y}(\vec u)xw_{n+1})\eta(2^{n\varepsilon s}\det D^L_{\vec u}(\Phi_{\vec y})(\vec u)) \times\prod_{i=1}^n\varphi(u_i)K_0(y_i)\min\Big\{\frac{u_i\rho(y_i)}{\rho(\xi_i)},\frac{\rho(\xi_i)}{u_i\rho(y_i)}\Big\}\,dx\,d\vec u\,d\vec y\Big| \tag{3.13}$$

and

$$\Big|\int_{A_0^n}\int_{[C^{-1},C]^n}\int_{\mathbb{G}} \psi_{-s}(x)g(w_1^{-1}\Phi_{\vec y}(\vec u)xw_{n+1})[1-\eta(2^{n\varepsilon s}\det D^L_{\vec u}(\Phi_{\vec y})(\vec u))] \times\prod_{i=1}^n\varphi(u_i)K_0(y_i)\min\Big\{\frac{u_i\rho(y_i)}{\rho(\xi_i)},\frac{\rho(\xi_i)}{u_i\rho(y_i)}\Big\}\,dx\,d\vec u\,d\vec y\Big|, \tag{3.14}$$

where η is a smooth nonnegative bump function which equals 1 near $0\in\mathbb{G}$.

To estimate (3.13), we simply replace everything by absolute values, and observe that it can be controlled by the "degenerate portion" of the integral (see Section 6 in [T]). More precisely, using the bounds on g, φ and K_0, we see that the left hand side of (3.13) does not exceed

$$\int_{\mathbb{G}}\int_{[C^{-1},C]^n}\int_{A_0^n} |\psi_{-s}(x)|\eta(2^{n\varepsilon s}\det D^L_{\vec u}(\Phi_{\vec y})(\vec u))\,d\vec y\,d\vec u\,dx,$$

which is $\le C2^{-\varepsilon s}$ (see [T]). So it remains to deal with (3.14).

4. L^2 estimate III: Nondegenerate portion of the integral. We first give the following result due to Tao.

LEMMA 4.1 ([T, Lemma 7.1]). *Let f be a function on $B(0,C)$ with mean zero and $\|f\|_1\lesssim 1$. Then there exist functions $f_1,\dots,f_n$ supported on a slightly larger ball $B(0,C)$ with $\|f_i\|_1\lesssim 1$ and $f(x)=\sum_{i=1}^n\partial_{x_i}f_i(x)$.*

Let us continue the proof of the L^2 estimate. Note that $K_0\in L^\infty(A_0)$ implies $K_0\in L^1(\mathbb{G})$. Thus to prove (3.14) $\lesssim 2^{-\varepsilon s}$, it suffices to show

$$\Big|\int_{[C^{-1},C]^n}\int_{\mathbb{G}} \psi_{-s}(x)\,g(w_1^{-1}\Phi_{\vec y}(\vec u)xw_{n+1})[1-\eta(2^{n\varepsilon s}\det D^L_{\vec u}(\Phi_{\vec y})(\vec u))] \times\prod_{i=1}^n\varphi(u_i)\min\Big\{\frac{u_i\rho(y_i)}{\rho(\xi_i)},\frac{\rho(\xi_i)}{u_i\rho(y_i)}\Big\}\,dx\,d\vec u\Big|\lesssim 2^{-\varepsilon s}$$

uniformly in $\vec y\in A_0^n$.

Since ψ_0 is supported on $B(0,16)$ with mean zero and $\|\psi\|_{L^1(\mathbb{G})}\lesssim 1$, by Lemma 4.1 we have $\psi_{-s}(x)=\sum_{j=1}^n\partial_{x_j}f_j(x)$, where the functions f_j are supported on $B(0,2^{4-s})$ and satisfy

$$\|f_j\|_1 \lesssim 2^{-\alpha_j s}. \tag{4.1}$$

It thus suffices to bound the quantity

$$\left|\iint \partial_{x_j} f_j(x) g(w_1^{-1}\Phi_{\vec{y}}(\vec{u}) x w_{n+1}) a(\vec{u}) \prod_{i=1}^{n} \min\left\{\frac{u_i\rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i\rho(y_i)}\right\} d\vec{u}\, dx\right| \tag{4.2}$$

by $C2^{-\varepsilon s}$ for all $j=1,\dots,n$, where

$$a(\vec{u}) = [1-\eta(2^{n\varepsilon s}\det D_{\vec{u}}^L(\Phi_{\vec{y}})(\vec{u}))]\prod_{i=1}^{n}\varphi(u_i).$$

If we integrate by parts in the x_j variable, (4.2) can be rewritten as

$$\left|\iint f_j(x)\partial_{x_j} g(\tilde{\Phi}) a(\vec{u}) \prod_{i=1}^{n} \min\left\{\frac{u_i\rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i\rho(y_i)}\right\} d\vec{u}\, dx\right|.$$

where we use $\tilde{\Phi}$ to denote $w_1^{-1}\Phi_{\vec{y}}(\vec{u})xw_{n+1}$ for simplicity. By (4.1), if ε is small enough, we only need to show

$$\left|\int \partial_{x_j} g(\tilde{\Phi}) a(\vec{u}) \prod_{i=1}^{n} \min\left\{\frac{u_i\rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i\rho(y_i)}\right\} d\vec{u}\right| \lesssim 2^{C\varepsilon s}$$

uniformly in $x\in B(0,16)$ for some constant C.

The following result was proved in [T].

Lemma 4.2 ([T, Lemma 7.2]). *Let $f:\mathbb{R}\times\mathbb{R}^n\to\mathbb{G}$ and $F:\mathbb{G}\to\mathbb{R}$ be smooth functions. Then*

$$\partial_s F(f(s,t)) = \nabla_t F(f(s,t))\cdot(D_t^L f(s,t))^{-1}\partial_s^L f(s,t)$$

whenever $\det D_t^L f(s,t)$ *is nonzero.*

By Lemma 4.2, it suffices to show

$$\left|\int \nabla_{\vec{u}} g(\tilde{\Phi})\cdot(D_{\vec{u}}^L\tilde{\Phi})^{-1}\partial_{x_j}^L(\tilde{\Phi}) a(\vec{u}) \prod_{i=1}^{n} \min\left\{\frac{u_i\rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i\rho(y_i)}\right\} d\vec{u}\right| \lesssim 2^{C\varepsilon s}, \tag{4.3}$$

where $\nabla_{\vec{u}} = (\partial_{u_1},\dots,\partial_{u_n})$. To show (4.3), it is equivalent to show

$$\left|\int \partial_{u_k} g(\tilde{\Phi})((D_{\vec{u}}^L\tilde{\Phi})^{-1}\partial_{x_j}^L(\tilde{\Phi}))_k a(\vec{u}) \prod_{i=1}^{n} \min\left\{\frac{u_i\rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i\rho(y_i)}\right\} d\vec{u}\right| \lesssim 2^{C\varepsilon s}$$

for each $k=1,\dots,n$, where $(\cdot)_k$ denotes the kth component of a vector.

It is not hard to see that the above inequality follows from

$$\left|\partial_{u_k}\left[((D_{\vec{u}}^L\tilde{\Phi})^{-1}\partial_{x_j}^L(\tilde{\Phi}))_k a(\vec{u}) \prod_{i} \min\left\{\frac{u_i\rho(y_i)}{\rho(\xi_i)}, \frac{\rho(\xi_i)}{u_i\rho(y_i)}\right\}\right]\right| \lesssim 2^{C\varepsilon s},$$

where ∂_{u_k} is understood to be the weak derivative when acting on the minimum function. Since

$$\Big|(1+\nabla_{\vec{u}})\prod_{i\neq k}\min\Big\{\frac{u_i\rho(y_i)}{\rho(\xi_i)},\frac{\rho(\xi_i)}{u_i\rho(y_i)}\Big\}\Big|\lesssim 1,$$

it suffices to show

$$|\partial_{u_k}[((D^L_{\vec{u}}\tilde{\Phi})^{-1}\partial^L_{x_j}(\tilde{\Phi}))_k a(\vec{u})]|\lesssim 2^{C\varepsilon s}.$$

Noticing that all functions appearing in the definition of a are smooth and compactly supported, we can easily see that

$$|a(\vec{u})|\lesssim 1 \quad\text{and}\quad |\nabla_{\vec{u}}a(\vec{u})|\lesssim 2^{n\varepsilon s}.$$

Thus it suffices to show that, on the support of $a(\vec{u})$,

$$|(1+\partial_{u_k})\cdot[(D^L_{\vec{u}}(\tilde{\Phi})^{-1}\partial^L_{x_j}(\tilde{\Phi})]_k|\lesssim 2^{C\varepsilon s}.$$

By Cramer's rule, it is equivalent to show

$$\Big|(1+\partial_{u_k})\frac{\det(\partial^L_{u_1}(\tilde{\Phi}),\ldots,\partial^L_{x_j}(\tilde{\Phi}),\ldots,\partial^L_{u_n}(\tilde{\Phi}))}{\det D^L_{\vec{u}}(\tilde{\Phi})}\Big|\lesssim 2^{C\varepsilon s}, \tag{4.4}$$

where $(\partial^L_{u_1}(\tilde{\Phi}),\ldots,\partial^L_{x_j}(\tilde{\Phi}),\ldots,\partial^L_{u_n}(\tilde{\Phi}))$ denotes the matrix whose j'th column vector is $\partial^L_{u_{j'}}\tilde{\Phi}$ for $j'=1,\ldots,n$, $j'\neq k$ and whose kth column vector is $\partial^L_{x_j}(\tilde{\Phi})$.

Now we summarize some useful conclusions in the following lemma; see [T] for the proofs.

Lemma 4.3 (see [T]). *Let $f(t)$, $g(t)$ be a smooth functions from $\mathbb{R}$ to $\mathbb{G}$ and $s(t)$ be a smooth function from $\mathbb{R}$ to $\mathbb{R}$. Then*

(a) $|\partial_t f(t)|\sim|\partial^L_t f(t)|$ *whenever* $|f(t)|\lesssim 1$;

(b) $\partial^L_t(f(t)g(t))=\partial^L_t g(t)+C[g(t)]\partial^L_t f(t)$;

(c) $C[t\circ x](t\circ v)=t\circ(C[x]v),\quad C[x]^{-1}=C[x^{-1}]$;

(d) $|C[x]v|\sim|v|$ *whenever* $|x|\lesssim 1$;

(e) $X(t\circ x)=t\circ X(x)$ *and* $\rho(X(x))\lesssim\rho(x)$;

(f) $\rho(\partial_t(t\circ x))\sim\rho(t\circ x)$ *for* $t\lesssim 1$;

(g) $\partial^L_t(s(t)\circ f(t))=s(t)\circ\partial^L_t f(t)+(s'(t)/s(t))(s(t)\circ X[f(t)])$,

where $X(x)$ is the vector field defined by $X(x)=\partial^L_t(t\circ x)|_{t=1}$.

By Lemma 4.3(b),

$$D^L_{\vec{u}}(\tilde{\Phi})=D^L_{\vec{u}}(\Phi_{\vec{y}}(\vec{u})xw_{n+1})=C[xw_{n+1}]D^L_{\vec{u}}(\Phi_{\vec{y}}(\vec{u})). \tag{4.5}$$

Since $\rho(xw_{n+1})\lesssim\rho(x)+\rho(w_{n+1})\lesssim 1$ (cf. [FoSt, p. 9]),

$$|\det D^L_{\vec{u}}(\tilde{\Phi})|=|\det D^L_{\vec{u}}(\Phi_{\vec{y}}(\vec{u})xw_{n+1})|\sim|\det D^L_{\vec{u}}(\Phi_{\vec{y}}(\vec{u}))|\gtrsim 2^{-n\varepsilon s} \tag{4.6}$$

on the support of $a(\vec{u})$. On the other hand, since $\tilde{\Phi}$ is smooth and compactly supported in all variables, we can readily see that

$$|(1+\partial_{u_k})\det(\partial_{u_1}^L(\tilde{\Phi}),\dots,\partial_{x_j}^L(\tilde{\Phi}),\dots,\partial_{u_n}^L(\tilde{\Phi}))| \lesssim 1.$$

This inequality together with (4.6) yields (4.4). Thus we have completed the proof of Theorem 1.1.

We make the following remark to demonstrate why we need to pass from the vector-valued integral to a scalar-valued one.

Remark 4.4. If we worked with the vector-valued integral, instead of (4.3), we need to show

$$\Big|\iint_0^\infty \nabla_{\vec{u}} g_{t_n}(w_1^{-1}\Phi_{\vec{y}}(\vec{u})x)\cdot(D_{\vec{u}}^L(w_1^{-1}\Phi_{\vec{y}}(\vec{u})x))^{-1}\partial_{x_j}^L(w_1^{-1}\Phi_{\vec{y}}(\vec{u})x)h_{t_n}(u_n|y_n|)$$
$$\times a(\vec{u})\prod_{i=1}^{n-1}\min\bigg\{\frac{u_i\rho(y_i)}{\rho(\xi_i)},\frac{\rho(\xi_i)}{u_i\rho(y_i)}\bigg\}\,\frac{dt_n}{t_n}\,d\vec{u}\Big| \lesssim 2^{C\varepsilon s}.$$

Take a look at the terms depending on t_n and observe that if we perform integration by parts in u_n, the boundary term which depends on t_n is

$$\int_0^\infty g_{t_n}(w_1^{-1}\Phi_{\vec{y}}(\vec{u})x)|_{u_n=t_n/\rho(y_n)}\,\frac{dt_n}{t_n},$$

which is not bounded in $L^\infty(\mathbb{G})$ although $g_{t_n}\in L^\infty(\mathcal{H})$.

5. Weak $(1,1)$ estimate I: Reduction to a strong type estimate. We turn to the proof of Theorem 1.2. Let us begin with some definitions and notations. A left-invariant quasi-distance d on $\mathbb{G}$ is defined by $d(x,y)=\rho(x^{-1}y)$. A ball $J:=B(x_J,2^j)$ with center x_J and radius 2^j is a set of the form $J=\{x: d(x,x_J)<2^j\}$ for some $x_J\in\mathbb{G}$ and $j\in\mathbb{Z}$. For some $C>1$ (only depending on the constant B_0 in the quasi-triangle inequality $\rho(xy^{-1})\le B_0[\rho(x)+\rho(y)]$), denote by $J_\triangle$ the annulus $CJ\setminus C^{-1}J$, where $rJ:=\{x: d(x,x_J)<r2^j\}$ for $r>0$. Moreover, let $K(x)=\Omega(x)\rho(x)^{-N}$, $h_t(x)=t^{-1}\rho(x)\chi_{\{\rho(x)\le t\}}(x)$ and $K_0=K\chi_{A_0}$ with $A_0=\{x\in\mathbb{G}: 1\le \rho(x)\le 2\}$. We wish to show that

$$\bigg|\bigg\{x\in\mathbb{G}:\bigg(\int_0^\infty |f*(Kh_t)(x)|^2\,\frac{dt}{t}\bigg)^{1/2}>\alpha\bigg\}\bigg| \lesssim \alpha^{-1}\|f\|_{L^1}\|K_0\|_{L\log L}.$$

We may assume that $f\in C_0^\infty$ and $\alpha=1$, $\|K_0\|_{L\log L}=1$ by linearity. We perform the standard Calderón–Zygmund decomposition of f at height 1 to obtain $f=g+\sum_J b_J$, where $\|g\|_{L^1}\lesssim\|f\|_{L^1}$ and $\|g\|_{L^\infty}\lesssim\alpha=1$. The J range over a collection of disjoint balls with $\sum_J|J|\lesssim\|f\|_{L^1}$ and for each J,

$$\operatorname{supp} b_J\subset CJ \tag{5.1}$$

and

$$\|b_J\|_{L^1(CJ)} \lesssim |J|, \quad \int b_J = 0. \tag{5.2}$$

Since f is smooth, we may arrange matters so that the b_J are smooth. We decompose $K = \sum_{j\in\mathbb{Z}} S_j K_0$, where

$$S_j K_0(x) = 2^{-j} \int_0^\infty \varphi(2^{-j}s)s^{-N} K_0(s^{-1}\circ x)\, ds, \qquad j \in \mathbb{Z}.$$

(See Section 2 for the definition of φ). In the following, we use 2^j to denote the radius of J, where $j = j(J)$ is an integer. Now we write

$$\begin{aligned}\Big(\int_0^\infty |f * (Kh_t)(x)|^2 \frac{dt}{t}\Big)^{1/2} \\ &\le |g * (Kh_{(\cdot)})(x)|_{\mathcal{H}} + \Big|\sum_{s\le C}\sum_J b_J * (h_{(\cdot)} S_{j+s}K_0)(x)\Big|_{\mathcal{H}} \\ &\quad + \Big|\sum_{s>C}\sum_J b_J * (h_{(\cdot)} S_{j+s}K_0)(x)\Big|_{\mathcal{H}} \\ &=: I_1 + I_2 + I_3.\end{aligned}$$

For the first term, the L^2 boundedness of g_Ω (Theorem 1.1) and Chebyshev's inequality imply that

$$|\{x : I_1 > 1\}| \le \|g\|^2_{L^2(\mathbb{G})} \lesssim \|g\|_{L^1(\mathbb{G})} \lesssim \|f\|_{L^1(\mathbb{G})}.$$

The second term is supported in $\bigcup_J CJ$, thus

$$|\{x : I_2 > 1\}| \le \Big|\bigcup_J CJ\Big| \le \sum_J |CJ| \lesssim \|f\|_{L^1(\mathbb{G})}.$$

To handle the remaining term, it suffices to show that

$$\bigg|\bigg\{\sum_{s>C}\Big(\int_0^\infty \Big|\sum_J b_J * (h_t S_{j+s}K_0)(x)\Big|^2 \frac{dt}{t}\Big)^{1/2} > 1\bigg\}\bigg| \lesssim \sum_J |J|.$$

Note that $\operatorname{supp}(S_{j+s}K_0) \subset \{x \in \mathbb{G} : 2^{j+s-2} \le \rho(x) \le 2^{j+s+3}\}$. Hence, $b_J * (h_t S_{j+s}K_0)$ is supported on the annulus $(2^s J)_\Delta$ provided that C in the above inequality is large enough (only depending on B_0 and C in (5.1)). Take a suitable smooth cutoff function ψ such that

$$\operatorname{supp}\psi \subset \{x : C^{-1}/2 < \rho(x) < 4C\}$$

and $\psi \equiv 1$ on $\{x : C^{-1} < \rho(x) < 2C\}$. Let $\psi_J^s(x) = \psi(2^{-j-s}\circ(x_J^{-1}x))$. Then to finish the proof of Theorem 1.2, it suffices to show that

$$\bigg|\bigg\{\sum_{s>C}\Big(\int_0^\infty \Big|\sum_J \psi_J^s(x)(b_J * (h_t S_{j+s}K_0)(x))\Big|^2 \frac{dt}{t}\Big)^{1/2} > 1\bigg\}\bigg| \lesssim \sum_J |J|. \tag{5.3}$$

Below we show that (5.3) is a consequence of the following proposition:

Proposition 5.1. *Suppose $s > C$ and $1 < p < 2$. Let $\mathcal{J}$ be a non-empty finite collection of disjoint balls such that*

$$\sum_J |J| \lesssim 1 \tag{5.4}$$

and

$$\Big\|\sum_J \chi_{C2^s J}\Big\|_{L^\infty(\mathbb{G})} \lesssim 2^{Ns}. \tag{5.5}$$

Let b_J be a collection of smooth functions satisfying (5.2) *and let $\psi_J = \psi_J^s$. Then*

$$\Big\|\sum_J \psi_J(b_J * (h_t S_{j+s} F_J))\Big\|_{L^p(\mathcal{H})} \lesssim 2^{-\varepsilon s}\Big(\sum_J |J|\,\|F_J\|^2_{L^2(\mathbb{G})}\Big)^{1/2} \tag{5.6}$$

for all functions F_J in $L^2(\mathbb{G})$.

The proof of Proposition 5.1 will be postponed until Section 6. Now let us complete the proof of (5.3) by applying Proposition 5.1. We need the following

Lemma 5.2 ([T, Lemma 9.2]). *Let $B \subset B(0, C)$ be any Euclidean ball of radius at least $2^{-\varepsilon s}$, and define the functions $\psi_{J,B}$ by*

$$\psi_{J,B}(x) = \psi_B(2^{-j-s} \circ (x_J^{-1} x)),$$

where ψ_B is any bump function adapted to B (this means that $0 \le \psi_B \le 1$, $\psi_B \in C^\infty$ with $\operatorname{supp}\psi_B \subset 2B$ and $\psi_B \equiv 1$ on B). Then

$$\Big|\Big\{\sum_J \psi_{J,B}(x) > s^3 2^{Ns}|B|\Big\}\Big| \lesssim 2^{-\varepsilon s^2}.$$

Applying Lemma 5.2 with a ball of size roughly 1 and a nonnegative cutoff, we obtain

$$\Big|\Big\{\sum_J \chi_{C2^s J} > s^3 2^{Ns}|B|\Big\}\Big| \lesssim 2^{-\varepsilon s^2}.$$

Then we use a sieving argument of Córdoba (see [Cor, p. 11]). For any ball $J \in \mathcal{J}$, define the *height* $h(J)$ to be the number

$$h(J) = \#\{J' \in \mathcal{J} : 2J \subset 2J'\},$$

where $\#E$ denotes the cardinality of the set E.

Define the exceptional set E_s by

$$E_s = \bigcup_{h(J) \ge s^3 2^{Ns}} J.$$

Then $|E_s| \lesssim 2^{-\varepsilon s}$. For each $a = 0, 1, \ldots, s^3 - 1$, the collection of balls with height between $a2^{Ns}$ and $(a+1)2^{Ns}$ (denoted by J_a) satisfies (5.6), and by

s^3 applications of (5.5) and the triangle inequality, we obtain

$$\Big\|\sum_J \psi_J(b_J * (h_t S_{j+s} F_J))\Big\|_{L^p(E_s^c,\mathcal{H})} \lesssim 2^{-\varepsilon s}\Big(\sum_J |J|\,\|F_J\|^2_{L^2(\mathbb{G})}\Big)^{1/2} \tag{5.7}$$

for all functions F_J in $L^2(\mathcal{H})$ (see [T]).

We adapt Tao's argument in [T]. By dilation invariance, it suffices to verify (5.3) in the case when $\sum_J |J| \sim 1$. In particular, we may assume that (5.4) holds.

For each $s > C$ we decompose K_0 as $K_0 = K^{\le s} + K^{>s}$, where $K^{\le s}(x) = K_0(x)\chi_{\{|K_0| \lesssim 2^{\varepsilon s/2}\}}(x)$. It suffices to show

$$\bigg|\bigg\{\sum_{s>C}\Big(\int_0^\infty \Big|\sum_J \psi_J(b_J * (h_t S_{j+s} K^{>s}))\Big|^2 \frac{dt}{t}\Big)^{1/2} \gtrsim 1\bigg\}\bigg| \lesssim 1, \tag{5.8}$$

$$\bigg|\bigg\{\sum_{s>C}\Big(\int_0^\infty \Big|\sum_J \psi_J(b_J * (h_t S_{j+s} K^{\le s}))\Big|^2 \frac{dt}{t}\Big)^{1/2} \gtrsim 1\bigg\}\bigg| \lesssim 1. \tag{5.9}$$

By Lemma 2.1, (2.5) and (5.2), we have

$$\begin{aligned}\|\psi_J(b_J * (h_t S_{j+s} K^{>s}))\|_{L^1(\mathcal{H})} &\lesssim \|b_J\|_{L^1(\mathbb{G})}\|h_t S_{j+s} K^{>s}\|_{L^1(\mathcal{H})}\\ &\lesssim |J|\,\|K^{>s}\|_{L^1(\mathbb{G})},\end{aligned}$$

Note that

$$\sum_{s>C}\|K^{>s}\|_{L^1(\mathbb{G})} \lesssim \|K_0\|_{L\log L(\mathbb{G})} \lesssim 1.$$

Then by (5.4) and the triangle inequality,

$$\sum_{s>C}\Big\|\sum_J \psi_J(b_J * (h_t S_{j+s} K^{>s}))\Big\|_{L^1(\mathcal{H})} \lesssim \sum_{j>C} |J|\,\|K^{>s}\|_{L^1(\mathbb{G})} \lesssim 1,$$

which implies (5.8) by Chebyshev's inequality.

As for (5.9), applying the estimate (5.7) with $F_J = K^{\le s}$ for all J, we have, for each s,

$$\begin{aligned}\Big\|\Big|\sum_J \psi_J(b_J * (h_t S_{j+s} K^{\le s}))\Big|_{\mathcal{H}}\Big\|_{L^p(E_s^c)} &\lesssim 2^{-\varepsilon s}\Big(\sum_J |J|\,\|K^{\le s}\|^2_{L^2(\mathbb{G})}\Big)^{1/2}\\ &\lesssim 2^{-\varepsilon s/2}.\end{aligned}$$

Summing this over $s > C$ yields

$$\bigg\|\sum_{s>C}\Big(\int_0^\infty \Big|\sum_J \psi_J(b_J * (h_t S_{j+s} K^{\le s}))\Big|^2 \frac{dt}{t}\Big)^{1/2}\bigg\|_{L^p((\bigcup_{s>C} E_s)^c)} \lesssim 1.$$

Thus (5.9) follows from Chebyshev's inequality and the fact $|\bigcup_{s>C} E_s| \lesssim 1$. This completes the derivation of (5.3) from Proposition 5.1.

6. Weak $(1,1)$ estimate II: Proof of Proposition 5.1 (I). The following sections are devoted to proving Proposition 5.1. To get (5.6), by duality, it suffices to show

$$\Big(\sum_J |J|^{-1}\Big\|\int_0^\infty S^*_{j+s}[(\tilde{b}_J)*(\psi_J g_t)h_t]\,\frac{dt}{t}\Big\|^2_{L^2(\mathbb{G})}\Big)^{1/2}\lesssim 2^{-\varepsilon s}\|g_t\|_{L^{p'}(\mathcal{H})}$$

for all test functions $g_t\in L^{p'}(\mathcal{H})$, where S^*_{j+s} denotes the adjoint of S_{j+s} and $\tilde{b}_J(x)=b_J(-x)$. By the TT^* method, we need to show

$$\Big\|\sum_J |J|^{-1}\psi_J\Big(b_J * S_{j+s}\Big[\int_0^\infty S^*_{j+s}([\tilde{b}_J*(\psi_J g_{t_1})]h_{t_1})\,\frac{dt_1}{t_1}\,h_{t_2}\Big]\Big)\Big\|_{L^p(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|g_t\|_{L^{p'}(\mathcal{H})}. \tag{6.1}$$

Define the self-adjoint operators T_J and T by

$$T_J f_t(x)=\frac{b_J}{|J|}*\Big(S_0\Big(\int_0^\infty S_0^*[(\tilde{b}_J*f_{t_1})h_{t_1}]\,\frac{dt_1}{t_1}\Big)h_{t_2}\Big)(x)$$

and $T=2^{-Ns}\sum_J\psi_J T_J\psi_J$, respectively, where ψ_J denotes the multiplier defined by

$$\psi_J f(x):=\psi_J(x)f(x).$$

Then the desired estimate (6.1) takes the form

$$\|Tf_t\|_{L^p(\mathcal{H})}\lesssim 2^{-\varepsilon s}\|f_t\|_{L^{p'}(\mathcal{H})}, \tag{6.2}$$

since $S_{j+s}S^*_{j+s}=2^{-N(s+j)}S_0S_0^*$. Let $d_J:B(0,C)\to CJ$ denote the map

$$d_J(v)=x_J(2^j\circ v). \tag{6.3}$$

Define the smooth functions c_J supported on the ball $B(0,C)$ by $c_J(v)=|J|^{-1}b_J(d_J(v))$. Then by (5.2),

$$\|c_J\|_{L^1(B(0,C))}\lesssim 1 \quad\text{and}\quad \int_{B(0,C)} c_J=0.$$

Note that

$$S_0S_0^*F(x)=\int\bar{\varphi}(u)F(u\circ x)\,du,$$

where $\bar{\varphi}(u)=\int\varphi(v)\varphi(uv)v^{-(N-1)}dv$ is a bump function adapted to $\{u\sim 1\}$. Then T_J takes the form

$$T_JF_{t_1}(t_2,x)=\iiiint c_J(v)\bar{\varphi}(u)c_J(w)F_{t_1}(d_J(w)u\circ(d_J(v)^{-1}x)) \times h_{t_1}(u\circ(d_J(v)^{-1}x))h_{t_2}(d_J(v)^{-1}x)\,dv\,du\,dw\,\frac{dt_1}{t_1}.$$

We need to define a slightly larger and noncancellative version of T_J. For each J, let ψ_J^+ be a slight enlargement of ψ_J which is positive on the support of ψ_J. Also, applying Lemma 4.1, we may find functions $c_J^1, \ldots, c_J^n$ supported on $B(0, C)$ such that

$$c_J = \sum_{i=1}^{n} \partial_{x_i} c_J^i, \qquad \|c_J^i\|_{L^1(\mathbb{G})} \lesssim 1. \tag{6.4}$$

Define $c_J^+ = |c_J| + \sum_{i=1}^n |c_J^i|$. Then c_J^+ is a nonnegative function on $B(0, C)$ with

$$\|c_J^+\|_{L^1(\mathbb{G})} \lesssim 1. \tag{6.5}$$

Finally, we choose φ^+ to be any enlargement of $\bar{\varphi}$ which is strictly positive on the support of $\bar{\varphi}$, and satisfies $\varphi^+(u) = \varphi^+(u^{-1})u^{2-N}$. We then define the self-adjoint operator T_J^+ by

$$\begin{aligned} T_J^+ F_{t_1}(t_2, x) = \iiiint & c_J^+(v)\varphi^+(u)c_J^+(w)F_{t_1}(d_J(w)u \circ (d_J(v)^{-1}x)) \\ & \times h_{t_1}(u \circ (d_J(v)^{-1}x))\, h_{t_2}(d_J(v)^{-1}x)\, dv\, du\, dw\, \frac{dt_1}{t_1}, \end{aligned}$$

and

$$T^+ = 2^{-Ns} \sum_J \psi_J^+ T_J^+ \psi_J^+,$$

where ψ_J^+ denotes the multiplier defined by $\psi_J^+ f(x) := \psi_J^+(x) f(x)$. Clearly, for all J and non-negative F, we have the pointwise bounds

$$T_J F_{t_1}(t_2, x) \le T_J^+ F_{t_1}(t_2, x) \quad \text{and} \quad T F_{t_1}(t_2, x) \le T^+ F_{t_1}(t_2, x). \tag{6.6}$$

Before estimating (6.2), we first show

$$\|T^+ F_t\|_{L^p(\mathcal{H})} \lesssim \|F_t\|_{L^q(\mathcal{H})} \qquad \text{for all } 1 \le p \le q \le \infty. \tag{6.7}$$

By interpolation and duality, it suffices to verify (6.7) for $q = \infty$. By Hölder's inequality and (5.5), we have

$$\begin{aligned} \Big\| \sum_J \psi_J^+ T_J^+ \psi^+ F_{t_1} \Big\|_{L^p(\mathcal{H})} &= \bigg\| \bigg(\int_0^\infty \Big| \sum_J (\psi_J^+ T_J^+ \psi^+ F_{t_1})(t_2, \cdot) \Big|^2 \frac{dt_2}{t_2} \bigg)^{1/2} \bigg\|_{L^p(\mathbb{G})} \\ \le \bigg\| \sum_J |\psi_J^+| \bigg(\int_0^\infty |(T_J^+ \psi^+ F_{t_1})(t_2, \cdot)|^2 \frac{dt_2}{t_2} \bigg)^{1/2} \bigg\|_{L^p(\mathbb{G})} \\ \lesssim \Big\| \Big(\sum_J |\chi_{C2^s J}| \Big)^{1/p'} \Big\|_{L^\infty(\mathbb{G})} \bigg\| \bigg\{ \sum_J \bigg(\int_0^\infty |(T_J^+ \psi^+ F_{t_1})(t_2, \cdot)|^2 \frac{dt_2}{t_2} \bigg)^{p/2} \bigg\}^{1/p} \bigg\|_{L^p(\mathbb{G})} \\ \lesssim 2^{Ns/p'} \Big(\sum_J \|T_J^+ \psi_J^+ F_{t_1}\|_{L^p(\mathcal{H})}^p \Big)^{1/p}. \end{aligned}$$

By Hölder's and Minkowski's inequalities, using $\sum |J| \lesssim 1$ and (5.5) we get

$$\begin{aligned}\sum_J \|T_J^+\psi_J^+F_t\|^p_{L^p(\mathcal{H})} &\lesssim \|F_{t_1}\|^p_{L^\infty(\mathcal{H})}\sum_J \Big\| \iiint c_J^+(v)\varphi^+(u)c_J^+(w) \\ &\qquad \times \psi_J^+(d_J(w)u\circ(d_J(v)^{-1}x))\,dv\,du\,dw\Big\|^p_{L^p(\mathbb{G},dx)} \\ &\lesssim \|F_{t_1}\|^p_{L^\infty(\mathcal{H})}\sum_J \|\psi_J^+\|^p_{L^p(\mathbb{G})} \lesssim 2^{Ns}\|F_{t_1}\|^p_{L^\infty(\mathcal{H})}.\end{aligned}$$

Now combining these two estimates yields (6.7).

To finish the proof of Theorem 1.2, we must obtain an attenuation factor $2^{-2\varepsilon s}$ for (6.2). To this end, we want to iterate T $m+1$ times with $m = 2^{2n-3}$, as in the L^2 case. But symmetry considerations require iterating one more time, i.e., iterating $m+2$ times is sufficient for our purpose (see Remark 6.3 below). Now we turn to the details.

Below we show that, to get (6.2), it suffices to prove

$$\|T^{m+2}F_t\|_{L^p(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|F_t\|_{L^{p'}(\mathcal{H})} \tag{6.8}$$

for $1 \le p < 2$ and some $\varepsilon > 0$. Indeed, by (6.8) and $m-2$-fold application of (6.7) for $q = p$ and replacing T^+ by T (see (6.6)), we get

$$\|T^{2m}F_t\|_{L^p(\mathcal{H})} \lesssim \|T^{m+2}F_t\|_{L^p(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|F_t\|_{L^{p'}(\mathcal{H})}. \tag{6.9}$$

By the TT^* method and the self-adjointness of T, it is easy to see that (6.9) implies

$$\|T^{m}F_t\|_{L^p(\mathcal{H})} \lesssim 2^{-\varepsilon' s}\|F_t\|_{L^{2}(\mathcal{H})} \tag{6.10}$$

for some $\varepsilon' > 0$. On the other hand, by repeated application of (6.7) and (6.6),

$$\|T^{m}F_t\|_{L^p(\mathcal{H})} \lesssim \|F_t\|_{L^{q}(\mathcal{H})} \quad \text{for all } q \ge p \ge 1. \tag{6.11}$$

By interpolation between (6.10) and (6.11), we thus obtain

$$\|T^{m}F_t\|_{L^p(\mathcal{H})} \lesssim 2^{-\varepsilon'' s}\|F_t\|_{L^{p'}(\mathcal{H})} \tag{6.12}$$

for some small $\varepsilon'' > 0$. Iterating this argument $2^{2n-4}-1$ times we thus obtain

$$\|T^{2}F_t\|_{L^p(\mathcal{H})} \lesssim 2^{-\theta s}\|F_t\|_{L^{p'}(\mathcal{H})}$$

for some $\theta > 0$, and (6.2) follows from this easily (see the process of deriving (6.12) from (6.9)).

Thus, we only need to show (6.8). Applying (6.7) $m+2$ times for $p = q = 2$, we see that T^{m+2} is bounded on $L^2(\mathcal{H})$. By interpolation and duality, it suffices to prove (6.8) for $p = 1$. By expanding T^m, we thus only need to show that

$$\Big\|2^{-Nms}\sum_{J_1,\dots,J_m\in\mathcal{J}} T\Big(\bigotimes_{i=1}^m \psi_{J_i}T_{J_i}\psi_{J_i}\Big)TF_{t_m}\Big\|_{L^1(\mathcal{H})}\lesssim 2^{-\varepsilon s}\|F_{t_m}\|_{L^\infty(\mathcal{H})}; \tag{6.13}$$

here, $\bigotimes_{i=1}^m Q_i$ denotes the composition of the operators $Q_1,\dots,Q_m$, defined by $\bigotimes_{i=1}^m Q_i(f) := Q_m(\cdots(Q_2(Q_1 f))\cdots)$. The balls $\{J_1,\dots,J_m\}$ may have different sizes. We shall extract a subsequence of n ball, whose sizes increase monotonically. First we need the following definition:

DEFINITION 6.1. Let $\mathbf{J}=(J_1,\dots,J_m)$ be an m-tuple of balls $(m\ge n)$ and $\mathbf{k}=\{k_1,\dots,k_n\}$ be a strictly increasing n-tuple of integers in $\{1,\dots,m\}$. We say that $\mathbf{J}$ is *ascending* with respect to $\mathbf{k}$ if

$$j_{k_q}\le j_l \quad \text{for all } k_q\le l\le k_n;$$

we then write $\mathbf{J}\nearrow\mathbf{k}$. Similarly, we say that $\mathbf{J}$ is *descending* with respect to $\mathbf{k}$ if

$$j_{k_q}\le j_l \quad \text{for all } k_1\le l\le k_q,$$

and write $\mathbf{J}\searrow\mathbf{k}$, where 2^{j_i} is the radius of J_i $(i=1,\dots,m)$.

The following lemma is due to Tao.

LEMMA 6.2 ([T, Lemma 10.2]). *If $m=2^{2n-3}$ and $\mathbf{J}\in\mathcal{J}^m$, then there exists a sequence $\mathbf{k}$ such that either $\mathbf{J}\nearrow\mathbf{k}$ or $\mathbf{J}\searrow\mathbf{k}$.*

For all $\mathbf{k}=\{k_1,\dots,k_n\}\subset\{1,\dots,m\}$, we say $\mathbf{k}<\mathbf{k}'$ if $k_1<k_1'$, or $k_j=k_j'$ for $j=1,\dots,i-1$, but $k_i<k_i'$ $(i=2,\dots,n)$. This defines an order on $\mathbf{k}$. Thus, when $m=2^{2n-3}$, for all $\mathbf{J}\in\mathcal{J}^m$ we can choose a largest sequence $\mathbf{k}:=\mathbf{k}_{\max}(\mathbf{J})$ with respect to this order so that either $\mathbf{J}\nearrow\mathbf{k}$ or $\mathbf{J}\searrow\mathbf{k}$. Clearly, Lemma 6.2 implies that $\mathbf{k}_{\max}(\mathbf{J})$ is well-defined. Since the number of choices of $\mathbf{k}$ is finite, it suffices to show that

$$2^{-Nms}\Big\|\sum_{\substack{\mathbf{J}=(J_1,\dots,J_m)\in\mathcal{J}^m\\ \mathbf{k}_{\max}(\mathbf{J})=\mathbf{k}}} T\Big(\bigotimes_{i=1}^m \psi_{J_i}T_{J_i}\psi_{J_i}\Big)TF_{t_m}\Big\|_{L^1(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|F_{t_m}\|_{L^\infty(\mathcal{H})} \tag{6.14}$$

for each $\mathbf{k}$.

Fix $\mathbf{k}$. Using the following discussion we can reduce to the case when $k_1=1$ and $k_n=m$. Since

$$\mathbf{k}_{\max}(J_1,\dots,J_m)=\mathbf{k}$$

is independent of the choices of J_i for $k_n<i\le m$, we may write

$$\mathbf{k}_{\max}(J_1,\dots,J_{k_n})=\mathbf{k} \quad \text{instead of} \quad \mathbf{k}_{\max}(J_1,\dots,J_m)=\mathbf{k}.$$

Then the left hand side of (6.14) can be rewritten as

$$2^{-N(m-k_n)s}\Big\|\sum_{\substack{J_1,\dots,J_{k_n}\in\mathcal{J}\\ \mathbf{k}_{\max}(J_1,\dots,J_{k_n})=\mathbf{k}}} T\Big(\bigotimes_{i=1}^{k_n}\psi_{J_i}T_{J_i}\psi_{J_i}\Big)TT^{m-k_n}F_{t_m}\Big\|_{L^1(\mathcal{H})}.$$

By (6.7), T is bounded on $L^\infty(\mathcal{H})$, and it is sufficient to show that

$$2^{-N(m-k_n)s}\Big\|\sum_{\substack{J_1,\dots,J_{k_n}\in\mathcal{J}\\ \mathbf{k}_{\max}(J_1,\dots,J_{k_n})=\mathbf{k}}} T\Big(\bigotimes_{i=1}^{k_n}\psi_{J_i}T_{J_i}\psi_{J_i}\Big)TF_{t_m}\Big\|_{L^1(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|F_{t_m}\|_{L^\infty(\mathcal{H})}.$$

Using a similar argument to the above, the desired estimate can be further reduced to

$$2^{-N(k_n-k_1+1)s}\Big\|T^{k_1-1}\sum_{\substack{J_{k_1},\dots,J_{k_n}\in\mathcal{J}\\ \mathbf{k}_{\max}(J_{k_1},\dots,J_{k_n})=\mathbf{k}}} T\Big(\bigotimes_{i=k_1}^{k_n}\psi_{J_i}T_{J_i}\psi_{J_i}\Big)TF_{t_m}\Big\|_{L^1(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|F_{t_m}\|_{L^\infty(\mathcal{H})}.$$

The left hand side is majorized by

$$2^{-N(k_n-k_1+1)s}\Big\|(T^+)^{k_1-1}\sum_{\substack{J_{k_1},\dots,J_{k_n}\in\mathcal{J}\\ \mathbf{J}\nearrow\mathbf{k}\ \mathrm{or}\ \mathbf{J}\searrow\mathbf{k}}}\Big|T\Big(\bigotimes_{i=k_1}^{k_n}\psi_{J_i}T_{J_i}\psi_{J_i}\Big)TF_{t_m}\Big|\Big\|_{L^1(\mathcal{H})}.$$

By (6.7), T^+ is bounded on $L^1(\mathcal{H})$, so we may discard the $(T^+)^{k_1-1}$ operator. Relabelling $\mathbf{J}$ and $\mathbf{k}$, and reducing m to k_n-k_1+1, we only need to show that

$$\Big\|2^{-Nms}\sum_{\substack{J_1,\dots,J_m\in\mathcal{J}\\ \mathbf{J}\nearrow\mathbf{k}\ \mathrm{or}\ \mathbf{J}\searrow\mathbf{k}}}\Big|T\Big(\bigotimes_{i=1}^{m}\psi_{J_i}T_{J_i}\psi_{J_i}\Big)TF_{t_m}\Big|\Big\|_{L^1(\mathcal{H})} \lesssim 2^{-\varepsilon s}\|F_{t_m}\|_{L^\infty(\mathcal{H})} \tag{6.15}$$

for all $n\le m\le 2^{2n-3}$ and all $\mathbf{k}$ with $k_1=1$ and $k_n=m$.

Remark 6.3. Recall that in the estimates of the L^2 case, we iterate L $n+1$ times (see Section 3) because of an extra integral in the variable t. Here we need to iterate T $m+2$ times in (6.8) in order to guarantee that the left hand side of (6.15) is symmetric in the sense that the case $\mathbf{J}\nearrow\mathbf{k}$ is dual to $\mathbf{J}\searrow\mathbf{k}$. Thus we only need to prove (6.15) for $\mathbf{J}\nearrow\mathbf{k}$.

Below we prove (6.15) for $\mathbf{J}\nearrow\mathbf{k}$. Since T is bounded on $L^\infty(\mathcal{H})$, it suffices to show

(6.16)

$$\Big\| 2^{-Nms} \sum_{J_1,\dots,J_m\in\mathcal{J}\,:\,\mathbf{J}\nearrow\mathbf{k}} \Big| T\Big(\bigotimes_{i=1}^{m} \psi_{J_i} T_{J_i} \psi_{J_i}\Big) F_{t_m} \Big| \Big\|_{L^1(\mathcal{H})} \lesssim 2^{-\varepsilon s} \|F_{t_m}\|_{L^\infty(\mathcal{H})}$$

for all $n \le m \le 2^{2n-3}$ and all $\mathbf{k}$ such that $k_1 = 1,\ k_n = m$.

Next, we shall remove the t integral in the $\mathcal{H}$ norm on the left hand side of (6.16). Then the problem reduces to estimating a scalar-valued integral. To be more precise, we will use the following conclusion:

Lemma 6.4.

$$\|T^+ F_{t_0}\|_{L^1(\mathcal{H})} \lesssim \sup_{J\in\mathcal{J},\, w\in B(0,C)} \Big\| \int_0^\infty h_{t_0}(|d_J(w)^{-1}\cdot|) F_{t_0} \psi_J^+ \frac{dt_0}{t_0} \Big\|_{L^1(\mathbb{G})}. \tag{6.17}$$

Proof. We first show that (6.17) is a consequence of (5.5) and the inequality

$$\|(T_J^+ \psi_J^+ F_{t_0})(\cdot,\cdot)\|_{L^1(\mathcal{H})} \lesssim \sup_{w\in B(0,C)} \Big\| \int_0^\infty h_{t_0}(d_J(w)^{-1}\cdot) \psi_J^+ F_{t_0} \frac{dt_0}{t_0} \Big\|_{L^1(\mathbb{G})}. \tag{6.18}$$

In fact, the above estimate together with (5.5) yields

$$\begin{aligned}
\|T^+ F_{t_0}\|_{L^1(\mathcal{H})} &\lesssim 2^{-Ns} \Big(\sum_J \|\psi_J^+\|_{L^\infty(\mathbb{G})} \cdot \|T_J^+(\psi_J^+ F_{t_0})\|_{L^1(\mathcal{H})} \Big) \\
&\lesssim 2^{-Ns} \sum_J \|\chi_{C2^sJ}\|_{L^\infty(\mathbb{G})} \sup_{w\in B(0,C)} \Big\| \int_0^\infty h_{t_0}(|d_J(w)^{-1}\cdot|) F_{t_0} \psi_J^+ \frac{dt_0}{t_0} \Big\|_{L^1(\mathbb{G})} \\
&\lesssim \sup_{J\in\mathcal{J},\, w\in B(0,C)} \Big\| \int_0^\infty h_{t_0}(|d_J(w)^{-1}\cdot|) F_{t_0} \psi_J^+ \frac{dt_0}{t_0} \Big\|_{L^1(\mathbb{G})}.
\end{aligned}$$

Let us turn to the proof of (6.18). For any $G_s \in L^\infty(\mathcal{H})$ with $\|G_s\|_{L^\infty(\mathcal{H})} \le 1$, we write $|\langle (T_J^+ \psi_J^+ F_{t_0})(\cdot, s), G(\cdot, s)\rangle_{L^1(\mathcal{H})}|$ as

$$\begin{aligned}
\Big| \iiiint & c_J^+(v) \varphi^+(u) c_J^+(w) \langle F_{t_0}(d_J(w)u \circ (d_J(v)^{-1}x)), h_{t_0}(u \circ (d_J(v)^{-1}x))\rangle_{\mathcal{H}} \\
&\times \psi_J^+(d_J(w)u \circ (d_J(v)^{-1}x)) \langle h_s(d_J(v)^{-1}x), G_s(x)\rangle_{\mathcal{H}}\, dx\, du\, dv\, dw \Big|.
\end{aligned}$$

Now we use the change of variables $x = d_J(v)\frac{1}{u} \circ (d_J(w)^{-1}y)$ (so $dx = u^{-N} dy$) and $u \sim 1$ to deduce

$$\begin{aligned}
&|\langle (T_J^+ \psi_J^+ F_{t_0})(\cdot, s), G(\cdot, s)\rangle_{L^1(\mathcal{H})}| \\
&\quad\lesssim \Big| \iiiint c_J^+(v) \varphi^+(u) c_J^+(w) \psi_J^+(y) \langle F_{t_0}(y), h_{t_0}(d_J(w)^{-1}y)\rangle_{\mathcal{H}} \\
&\qquad \times \Big\langle h_s\Big(\frac{1}{u} \circ (d_J(w)^{-1}y)\Big), G_s\Big(d_J(v)\,\frac{1}{u} \circ (d_J(w)^{-1}y)\Big) \Big\rangle_{\mathcal{H}} dy\, du\, dv\, dw \Big|
\end{aligned}$$

$$\lesssim \iiiint c_J^+(v)\varphi^+(u)c_J^+(w)\psi_J^+(y)|\langle F_{t_0}(y), h_{t_0}(d_J(w)^{-1}y)\rangle_{\mathcal{H}}| \times \|h_s\|_{L^\infty(\mathcal{H})}\|G_s\|_{L^\infty(\mathcal{H})}\,dy\,du\,dv\,dw$$

$$\lesssim \sup_{w\in B(0,C)} \|\langle h_{t_0}(d_J(w)^{-1}\cdot), \psi_J^+F_{t_0}\rangle_{\mathcal{H}}\|_{L^1(\mathbb{G})}\|c_J^+\|_{L^1(\mathbb{G})}^2\|\psi^+\|_{L^\infty(\mathbb{G})}\|\varphi^+\|_{L^1(\mathbb{G})}$$

$$\lesssim \sup_{w\in B(0,C)} \bigg\|\int_0^\infty h_{t_0}(d_J(w)^{-1}\cdot)\psi_J^+F_{t_0}\,\frac{dt_0}{t_0}\bigg\|_{L^1(\mathbb{G})},$$

where (6.5) is used in the last inequality. By duality, we finally obtain (6.18). ∎

Remark 6.5. From the pointwise inequality (6.6), we see that Lemma 6.4 still holds on replacing T^+ by T.

Applying Lemma 6.4 to the function

$$F(t_0,\cdot) = \Big[\Big(\bigotimes_{i=1}^m \psi_{J_i}T_{J_i}\psi_{J_i}\Big)F_{t_m}\Big](t_0,\cdot)$$

and $J = J_0$, we see that to show (6.16), it suffices to verify

$$2^{-Nms}\sum_{\mathbf{J}\in\mathcal{J}^m:\mathbf{J}\nearrow\mathbf{k}} \Big\|\Big\langle\Big[\Big(\bigotimes_{i=1}^m \psi_{J_i}T_{J_i}\psi_{J_i}\Big)F_{t_m}\Big](t_0,\cdot)\psi_{J_0}^+(\cdot), h_{t_0}(d_{J_0}(w_0)^{-1}\cdot)\Big\rangle_{\mathcal{H}}\Big\|_{L^1(\mathbb{G})} \lesssim 2^{-\varepsilon s}\|F_{t_m}\|_{L^\infty(\mathcal{H})} \tag{6.19}$$

uniformly in J_0 and $w_0 \in B(0,C)$ for all $n \le m \le 2^{2n-3}$ and all $\mathbf{k}$ with $k_1 = 1$ and $k_n = m$.

Fix m and $\mathbf{k}$. By duality, it suffices to prove that the quantity

$$\sum_{\mathbf{J}\nearrow\mathbf{k}}\Big|\Big\langle\int_0^\infty\Big[\Big(\bigotimes_{i=1}^m \psi_{J_i}T_{J_i}\psi_{J_i}\Big)F_{t_m}\Big](t_0,\cdot)h_{t_0}(d_{J_0}(w_0)^{-1}\cdot)\psi_{J_0}^+\,\frac{dt_0}{t_0}, G\Big\rangle\Big| \tag{6.20}$$

is $\lesssim 2^{-\varepsilon s}2^{Nms}$ for all functions F_t in the unit ball of $L^\infty(\mathcal{H})$ and G in the unit ball of $L^\infty(\mathbb{G})$.

For each $\mathbf{J}\in\mathcal{J}^m$ and $\mathbf{J}\nearrow\mathbf{k}$, we expand the inner product in (6.20) as

$$\int\cdots\int F_{t_m}(x_m)G(x_0)\psi_{J_0}^+(x_0)\prod_{i=1}^m\big(\psi_{J_i}(x_{i-1})c_{J_i}(v_i)\bar\varphi(u_i)c_{J_i}(w_i)\psi_{J_i}(x_i)\big) \times\prod_{j=0}^{m-1}R_jh_{t_m}(u_m\circ(d_{J_m}(v_m)^{-1}x_{m-1}))\,\frac{dt_m}{t_m}\,d\vec w\,d\vec u\,d\vec v\,dx_0,$$

where $\vec v = (v_1,\dots,v_m)$, $\vec w = (w_1,\dots,w_m)$ range over $B(0,C)^m$, and $\vec u = (u_1,\dots,u_m)$ ranges over $[C^{-1},C]^m$, $d\vec w = \prod_{i=1}^m dw_i$, $d\vec v = \prod_{i=1}^m dv_i$, x_0

ranges over $\mathbb{G}$, and $x_1,\dots,x_m$ are defined recursively by

$$x_i = d_{J_i}(w_i)(u_i\circ(d_{J_i}(v_i)^{-1}x_{i-1})) \quad \text{for } i=1,\dots,m. \tag{6.21}$$

Moreover, R_j $(j=0,\dots,m-1)$ is defined by

$$R_0 := \min\left\{\frac{\rho(d_{J_1}(v_1)^{-1}x_0)}{\rho(d_{J_0}(w_0)^{-1}x_0)}, \frac{\rho(d_{J_0}(w_0)^{-1}x_0)}{\rho(d_{J_1}(v_1)^{-1}x_0)}\right\},$$

and

$$R_j := \min\left\{\frac{\rho(u_j\circ(d_{J_j}(v_j)^{-1}x_{j-1}))}{\rho(d_{J_{j+1}}(v_{j+1})^{-1}x_j)}, \frac{\rho(d_{J_{j+1}}(v_{j+1})^{-1}x_j)}{\rho(u_j\circ(d_{J_j}(v_j)^{-1}x_{j-1}))}\right\}$$

for $j=1,\dots,m-1$.

We define new variables $\vec{\tau}=(\tau_1,\dots,\tau_n)$ by $\tau_q=u_{k_q}$ and $y=v_1$, since only these variables are actively used below. As in the proof of Theorem 1.1, we shall decompose the $\mathbf{J}\nearrow\mathbf{k}$ portion of (6.20) into

$$\begin{aligned}\sum_{\mathbf{J}\in\mathcal{J}^m:\mathbf{J}\nearrow\mathbf{k}}\bigg|\int\cdots\int F_{t_m}(x_m)G(x_0)\psi^+_{J_0}(x_0)\prod_{i=1}^m\big(\psi_{J_i}(x_{i-1})c_{J_i}(v_i)\bar{\varphi}(u_i)\\ \times c_{J_i}(w_i)\psi_{J_i}(x_i)\big)\prod_{j=0}^{m-1}R_j\eta(2^{\delta s}2^{-M_n}\det D^L_{\vec{\tau}}(x_m))\\ \times h_{t_m}(u_m\circ(d_{J_m}(v_m)^{-1}x_{m-1}))\,\frac{dt_m}{t_m}\,d\vec{w}\,d\vec{u}\,d\vec{v}\,dx_0\bigg|\end{aligned} \tag{6.22}$$

and

$$\begin{aligned}\sum_{\mathbf{J}\in\mathcal{J}^m:\mathbf{J}\nearrow\mathbf{k}}\bigg|\int\cdots\int F_{t_m}(x_m)G(x_0)\psi^+_{J_0}(x_0)\prod_{i=1}^m\big(\psi_{J_i}(x_{i-1})c_{J_i}(v_i)\bar{\varphi}(u_i)\\ \times c_{J_i}(w_i)\psi_{J_i}(x_i)\big)\prod_{j=0}^{m-1}R_j[1-\eta(2^{\delta s}2^{-M_n}\det D^L_{\vec{\tau}}(x_m))]\\ \times h_{t_m}(u_m\circ(d_{J_m}(v_m)^{-1}x_{m-1}))\,\frac{dt_m}{t_m}\,d\vec{w}\,d\vec{u}\,d\vec{v}\,dx_0\bigg|\end{aligned} \tag{6.23}$$

where $M_n=\sum_{i=1}^n\alpha_i(j_{k_i}+s)$ (recall that $2^{j_{k_i}}$ is the radius of J_{k_i}) and $\delta>0$ is a small constant to be chosen later, and η is a bump function which equals 1 near the identity.

The degenerate portion (6.22), as in the L^2 case, can be majorized by the corresponding portion of the singular integral. Thus it suffices to verify that

$$\text{the nondegenerate portion (6.23)} \lesssim 2^{-\varepsilon s}2^{Nms}. \tag{6.24}$$

7. Weak $(1,1)$ estimate III: Nondegenerate portion of the integral. To finish the proof of Proposition 5.1, we only need to show (6.24). Define

$$\tilde{T}_J^+ F(x) = \iiint c_J^+(v)\varphi^+(u)c_J^+(w)F(d_J(w)u \circ (d_J(v)^{-1}x))\, dw\, du\, dv$$

and

$$\tilde{T}^+ = 2^{-Ns} \sum_J \psi_J^+ \tilde{T}_J^+ \psi_J^+,$$

where c_J^+ and φ^+ are as in the definition of T_J^+ in Section 6. In [T], the following estimate is proved:

$$\|\tilde{T}^+ F\|_{L^p(\mathbb{G})} \lesssim \|F\|_{L^q(\mathbb{G})} \quad \text{for all } 1 \le p \le q \le \infty. \tag{7.1}$$

Therefore, to show (6.24), it is sufficient to show that

$$\begin{aligned}\sum_{\mathbf{J}\in\mathcal{J}^m:\mathbf{J}\nearrow\mathbf{k}} \bigg| \int \cdots \int F_{t_m}(x_m)G(x_0)\psi_{J_0}^+(x_0) \prod_{i=1}^m \big(\psi_{J_i}(x_{i-1})c_{J_i}(v_i)\bar{\varphi}(u_i)c_{J_i}(w_i) \\ \times \psi_{J_i}(x_i)\big) \prod_{j=0}^{m-1} R_j(1-\eta(2^{\delta s}2^{-M_n}\det D_{\vec{\tau}}^L(x_m))) \\ \times h_{t_m}(u_m \circ (d_{J_m}(v_m)^{-1}x_{m-1}))\, \frac{dt_m}{t_m}\, d\vec{w}\, d\vec{u}\, d\vec{v}\, dx_0 \bigg| \\ \lesssim 2^{-\varepsilon s}2^{Nms}\langle (\tilde{T}^+)^m 1, 1\rangle.\end{aligned}$$

By expanding out $\tilde{T}^+$, we rewrite the above estimate as

$$\begin{aligned}\bigg| \int \cdots \int F_{t_m}(x_m)G(x_0)\psi_{J_0}^+(x_0) \prod_{i=1}^m \big(\psi_{J_i}(x_{i-1})c_{J_i}(v_i)\bar{\varphi}(u_i)c_{J_i}(w_i) \\ \times \psi_{J_i}(x_i)\big) \prod_{j=0}^{m-1} R_j h_{t_m}(u_m \circ (d_{J_m}(v_m)^{-1}x_{m-1})) \\ \times (1-\eta(2^{\delta s}2^{-M_n}\det D_{\vec{\tau}}^L(x_m)))\, \frac{dt_m}{t_m}\, d\vec{w}\, d\vec{u}\, d\vec{v}\, dx_0 \bigg| \\ \lesssim 2^{-\varepsilon s} \iiiint \prod_{i=1}^m \psi_{J_i}^+(x_{i-1})c_{J_i}^+(v_i)\varphi^+(u_i)c_{J_i}^+(w_i)\psi_{J_i}^+(x_i)\, dx_0\, d\vec{w}\, d\vec{u}\, d\vec{v}.\end{aligned} \tag{7.2}$$

Fix all the frozen variables. Removing all the factors in the above expression which do not depend on the variables $y := v_1$ and $\vec{\tau}$, we reduce (7.2) to

$$\iint G(x_0)\Big(\int_0^\infty F_{t_m}(x_m)h_{t_m}(u_m\circ(d_{J_m}(v_m)^{-1}x_{m-1}))\,\frac{dt_m}{t_m}\Big)c_{J_1}(y) \times a(y,\vec{\tau})\Big(\prod_{i=0}^{m-1}R_i\Big)(1-\eta(2^{\delta s}2^{-M_n}\det D^L_{\vec{\tau}}(x_m)))\,dy\,d\vec{\tau} \lesssim 2^{-\varepsilon s}\iint c^+_{J_1}(v_1)a^+(y,\vec{\tau})\,dy\,d\vec{\tau}, \tag{7.3}$$

where

$$a(y,\vec{\tau})=\prod_{l=1}^m\psi_{J_l}(x_{l-1})\psi_{J_l}(x_l)(1-\eta(2^{\delta s}2^{-M_n}\det D^L_{\vec{\tau}}(x_m)))\prod_{q=1}^n\bar{\varphi}(\vec{\tau}_q)\psi^+_{J_0}(x_0)$$

$$a^+(y,\vec{\tau})=\prod_{l=1}^m\psi^+_{J_l}(x_{l-1})\psi^+_{J_l}(x_l)\prod_{q=1}^n\varphi^+(\vec{\tau}_q).$$

Since $d_{J_m}(w_m)$ is independent of y and $\vec{\tau}$, we may set

$$f(x_m)=\int_0^\infty F_{t_m}(x_m)h_{t_m}(u_m\circ(d_{J_m}(v_m)^{-1}x_{m-1}))\,\frac{dt_m}{t_m} =\int_0^\infty F_{t_m}(x_m)h_{t_m}(d_{J_m}(w_m)^{-1}x_m)\,\frac{dt_m}{t_m}.$$

By Cauchy–Schwarz's inequality, we can easily see $\|f\|_{L^\infty(\mathbb{G})}\le 1$.

To show (7.3), by (6.4) and integration by parts, it suffices to verify that

$$I_1+I_2+I_3:=\Big|\iint c^i_{J_1}(y)\partial_{y_i}f(x_m)a(y,\vec{\tau})\prod_{\varsigma=0}^{m-1}R_\varsigma\,dy\,d\vec{\tau}\Big| +\Big|\iint c^i_{J_1}(y)f(x_m)\partial_{y_i}a(y,\vec{\tau})\prod_{\varsigma=0}^{m-1}R_\varsigma\,dy\,d\vec{\tau}\Big| +\Big|\iint c^i_{J_1}(y)f(x_m)a(y,\vec{\tau})\Big(\partial_{y_i}\prod_{\varsigma=0}^{m-1}R_\varsigma\Big)\,dy\,d\vec{\tau}\Big| \lesssim 2^{-\varepsilon s}\iint c^+_{J_1}(v_1)a^+(y,\vec{\tau})\,dy\,d\vec{\tau} \tag{7.4}$$

for each $i=1,\dots,n$. The desired estimate for I_2 follows from Lemma 13.1 in [T] and the obvious bound $\prod_{\varsigma=0}^{m-1}R_\varsigma\le 1$. To estimate I_1 and I_3, the estimates for the derivatives of homogeneous norms are required.

We give an explicit estimate for the Euclidean derivative of the homogeneous norm which is new on general homogeneous groups.

Theorem 7.1. *The homogeneous norm ρ satisfies the estimate*

$$|\partial_{x_i}\rho(x)|\lesssim\rho(x)^{-\alpha_i+1}, \tag{7.5}$$

where α_i is given by (1.3).

Proof. The homogeneity of ρ implies that

$$\rho(r^{\alpha_1}x_1,\dots,r^{\alpha_n}x_n)=r\rho(x_1,\dots,x_n).$$

Taking the derivative in x_i, we have

$$r^{\alpha_i}\partial_{x_i}\rho(r\circ x)=r\partial_{x_i}\rho(x),$$

which shows that $\partial x_i\rho(x)$ is homogeneous of order $-\alpha_i+1$. Since the homogeneous norm is C^∞ away from the origin (see [FoSt]), we know that $\partial_{x_i}\rho(x)\lesssim 1$ for $x\in\Sigma$. Then (7.5) follows by a dilation argument. ∎

Using the above theorem, we next show the following

Lemma 7.2. *For $y=v_1$ and $x_{\varsigma-1}$ $(\varsigma=1,\dots,m)$ defined by* (6.21), *the following estimates hold*:

(a) $$|\partial_{y_i}[\rho(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})]|\lesssim 2^{-\varepsilon s}\rho(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})$$
for $\varsigma=1,\dots,m$ and $i=1,\dots,n$

(b) $$|\partial_{\vec{\tau}_q}[\rho(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})]|\lesssim \rho(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})$$
for $\varsigma=1,\dots,m$ and $q=1,\dots,n$.

Proof. Since $x_{\varsigma-1}\in(2^sJ_\varsigma)_\Delta$, we see that

$$\rho(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})\approx 2^{s+j_\varsigma}.$$

For (a), we only need to show that

$$|\partial_{y_i}[(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})_k]|\lesssim 2^{-\varepsilon s}2^{(s+j_\varsigma)\alpha_k},\tag{7.6}$$

where $(x)_k$ denotes the kth component of $x\in\mathbb{R}^n$. Indeed, let

$$z=(d_{J_\varsigma}(v_\varsigma)^{-1}\cdot x_{\varsigma-1}).$$

Since ρ is regular, by the chain rule, this estimate together with (7.5) yields

$$|\partial_{y_i}[\rho(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})]|=\left|\sum_{m=1}^{n}\frac{\partial\rho(z)}{\partial z_m}\,\frac{\partial z_m}{y_i}\right|\lesssim 2^{-\varepsilon s}2^{s+j_\varsigma},$$

which gives (a). Now we prove the inequality (7.6). Since

$$|2^{-s-j_\varsigma}\circ(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})|\sim\rho(2^{-s-j_\varsigma}\circ(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))\lesssim 1,$$

by Lemma 4.3(a), we have

$$|\partial_{y_i}(2^{-s-j_\varsigma}\circ(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))|\sim|\partial^L_{y_i}(2^{-s-j_\varsigma}\circ(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))|.$$

Applying Lemma 4.3(b), (g), we get

$$\begin{aligned}|\partial^L_{y_i}(2^{-s-j_\varsigma}\circ(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))|&=|\partial^L_{y_i}((2^{-s-j_\varsigma}u_{\varsigma-1})\circ(d_{J_{\varsigma-1}}(v_{\varsigma-1})^{-1}x_{\varsigma-2}))|\\&=\cdots\\&=|\partial^L_{y_i}((2^{-s-j_\varsigma}u_{\varsigma-1}\dots u_1)\circ(d_{J_1}(y)^{-1}x_0))|.\end{aligned}$$

Using the fact $j_1 \le j_\varsigma$ arising from $\mathbf{J} \nearrow \mathbf{k}$ and $u_l \sim 1$ for $l = 1, \dots, m$, and then applying Lemma 4.3(b), (c), (d), we have

$$\begin{aligned}|\partial_{y_i}(2^{-s-j_\varsigma} \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))| &\lesssim |\partial^L_{y_i}(2^{-s-j_\varsigma} \circ (d_{J_1}(y)^{-1}x_0))| \\ &\lesssim |\partial^L_{y_i}(2^{-s-j_1} \circ (d_{J_1}(y)^{-1}x_0))| \\ &\lesssim |C[2^{-j_1-s} \circ (x_{J_1}^{-1}x_0)](2^{-s} \circ \partial^L_{y_i}(y^{-1}))| \\ &\lesssim |(2^{-s} \circ \partial^L_{y_i}(y^{-1}))| \lesssim 2^{-\alpha_1 s},\end{aligned}$$

where we use $d_{J_1}(y)^{-1} = (2^{j_1} \circ y^{-1})x_J^{-1}$ in the third inequality, and in the last inequality we use the fact that y^{-1} depends on y polynomially. Hence, by (1.3),

$$\begin{aligned}2^{-(s+j_\varsigma)\alpha_k}|\partial_{y_i}(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})_{(k)}| &= |\partial_{y_i}(2^{-(s+j_\varsigma)\alpha_k}(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))_{(k)}| \\ &\le |\partial_{y_i}(2^{-s-j_i} \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))| \le 2^{-\alpha_1 s},\end{aligned}$$

which implies (7.6) .

The proof of (b) is similar. We may assume that $k_q \le \varsigma$ (otherwise (b) is trivial). By the chain rule, it suffices to show

$$|\partial_{\vec{\tau}_q}[(d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})_k]| \lesssim 2^{(s+j_\varsigma)\alpha_k}. \tag{7.7}$$

Using Lemma 4.3(a), (b), (g), (e), we have

$$\begin{aligned}|\partial_{\vec{\tau}_q}(2^{-s-j_\varsigma} \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))| &\sim |\partial^L_{\vec{\tau}_q}(2^{-s-j_\varsigma} \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))| \\ &\sim |\partial^L_{\vec{\tau}_q}(2^{-s-j_\varsigma} \circ (\vec{\tau}_q \circ d_{J_q}(v_q)^{-1}x_{q-1}))| \\ &\lesssim |\vec{\tau}_q^{\,-1}(\vec{\tau}_q \circ X[2^{-s-j_q} \circ (d_{J_q}(v_q)^{-1}x_{q-1})])| \\ &\sim |X[2^{-s-j_q} \circ (d_{J_q}(v_q)^{-1}x_{q-1})]| \\ &\sim \rho(2^{-s-j_q} \circ (d_{J_q}(v_q)^{-1}x_{q-1})) \lesssim 1,\end{aligned}$$

which, by (1.3), implies (7.7). ∎

Let us continue to estimate I_1 and I_3 in (7.4). We consider I_3 first. To begin, we prove that

$$|\partial_{y_i}(R_\varsigma)| \lesssim 2^{-\varepsilon s} \quad \text{for } i = 1, \dots, n \text{ and } \varsigma = 0, 1, \dots, m-1. \tag{7.8}$$

When $\varsigma = 0$, if $\rho(d_{J_1}(v_1)^{-1}x_0) \le \rho(d_{J_0}(w_0)^{-1}x_0)$, then by Lemma 7.2(a),

$$|\partial_{y_i}(R_0)| \lesssim \frac{|\partial_{y_i}[\rho(d_{J_1}(v_1)^{-1}x_0)]|}{\rho(d_{J_0}(w_0)^{-1}x_0)} \lesssim 2^{-\varepsilon s}.$$

If $\rho(d_{J_0}(w_0)^{-1}x_0) \le \rho(d_{J_1}(v_1)^{-1}x_0)$, the proof is similar.

Next we estimate $\partial_{y_i}(R_\varsigma)$ for $\varsigma = 1, \dots, m-1$. If $\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})) \le \rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)$, then by Lemma 7.2(a), we have

$$\begin{aligned}|\partial_{y_i}(R_\varsigma)| &\le \left|\frac{\partial_{y_i}[\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))]\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)}{\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)^2}\right| \\ &\quad + \left|\frac{\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))\partial_{y_i}[\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)]}{\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)^2}\right| \\ &\lesssim \frac{|\partial_{y_i}[\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))]| + |\partial_{y_i}[\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)]|}{\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)} \lesssim 2^{-\varepsilon s}.\end{aligned}$$

The case when $\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma) \le \rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))$ can be treated similarly. Thus (7.8) is proved. From (7.8) and $R_\varsigma \le 1\,(\varsigma = 0, \dots, m-1)$, we have

$$\left|\partial_{y_i}\prod_{\varsigma=0}^{m-1} R_\varsigma\right| \lesssim 2^{-\varepsilon s}. \tag{7.9}$$

From (7.9) and $\|f\|_{L^\infty(\mathbb{G})} \le 1$, the estimate for I_3 follows easily.

Finally, we estimate I_1. It suffices to show that

$$\left|\int \partial_{y_i} f(x_m)a(y,\vec\tau)\prod_{\varsigma=0}^{m-1} R_\varsigma\, d\vec\tau\right| \lesssim 2^{-\varepsilon s}\int a^+(y,\vec\tau)\, d\vec\tau \tag{7.10}$$

uniformly in y. Fix y. By Lemma 4.2, we can rewrite the left hand side as

$$\left|\int \nabla_{\vec\tau} f(x_m)\cdot (D^L_{\vec\tau}x_m)^{-1}\partial^L_{y_i}x_m a(y,\vec\tau)\prod_{\varsigma=0}^{m-1} R_\varsigma\, d\vec\tau\right|.$$

Integrating by parts, we see that this is equal to

$$\left|\int f(x_m)\nabla_{\vec\tau}\cdot ((D^L_{\vec\tau}x_m)^{-1}\partial^L_{y_i}x_m a(y,\vec\tau)\prod_{\varsigma=0}^{m-1} R_\varsigma) d\vec\tau\right|.$$

Thus to show (7.10), it suffices to prove the pointwise bound

$$\left\|(1+\nabla_{\vec\tau})\Big(\prod_{\varsigma=0}^{m-1} R_\varsigma\Big)\right\| \, \|(1+\nabla_{\vec\tau})((D^L_{\vec\tau}x_m)^{-1}\partial^L_{y_i}x_m a(y,\vec\tau))\| \lesssim 2^{-\varepsilon s}a^+(y,\vec\tau),$$

where $\|(1+\nabla)f\| := |\nabla f| + |f|$. The estimate

$$\|(1+\nabla_{\vec\tau})((D^L_{\vec\tau}x_m)^{-1}\partial^L_{y_i}x_m a(y,\vec\tau))\| \lesssim 2^{-\varepsilon s}a^+(y,\vec\tau)$$

is shown in [T, p. 1583]. Thus it suffices to show

$$\left\|(1+\nabla_{\vec\tau})\Big(\prod_{\varsigma=0}^{m-1} R_\varsigma\Big)\right\| \lesssim 1.$$

From the obvious fact $|R_\varsigma| \lesssim 1$, we reduce the proof to

$$|\partial_{\tau_q}(R_\varsigma)| \lesssim 1 \tag{7.11}$$

for each $\varsigma = 0, \ldots, m-1$ and $q = 1, \ldots, n$.

We first estimate $\partial_{\tau_q}(R_0)$. If $\rho(d_{J_1}(v_1)^{-1}x_0) \le \rho(d_{J_0}(w_0)^{-1}x_0)$, then by Lemma 7.2(b), we get

$$|\partial_{\tau_q}(R_0)| = \frac{|\partial_{\tau_q}[\rho(d_{J_1}(v_1)^{-1}x_0)]|}{\rho(d_{J_0}(w_0)^{-1}x_0)} \lesssim 1.$$

If $\rho(d_{J_0}(w_0)^{-1}x_0) \le \rho(d_{J_1}(v_1)^{-1}x_0)$, the estimate is similar.

Next we estimate $\partial_{\tau_q}(R_\varsigma)$ for $\varsigma = 1, \ldots, m-1$. If $\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1})) \le \rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)$, then by Lemma 7.2(b), we have

$$\begin{aligned}|\partial_{\tau_q}(R_\varsigma)| &\le \left|\frac{\partial_{\tau_q}[\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))]\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)}{\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)^2}\right| \\ &\quad + \left|\frac{\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))\partial_{\tau_q}[\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)]}{\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)^2}\right| \\ &\lesssim \frac{|\partial_{\tau_q}[\rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))]| + |\partial_{\tau_q}[\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)]|}{\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma)} \\ &\lesssim 1.\end{aligned} \tag{7.12}$$

The case $\rho(d_{J_{\varsigma+1}}(v_{\varsigma+1})^{-1}x_\varsigma) \le \rho(u_\varsigma \circ (d_{J_\varsigma}(v_\varsigma)^{-1}x_{\varsigma-1}))$ can be verified in a similar way. Thus (7.11) is proved. This ends the proof of Proposition 5.1 and, therefore, Theorem 1.2.

Acknowledgements. The research was supported by NSF of China (Grant no. 10931001) and SRFDP of China (Grant no. 20050027025). The authors would like to express their deep gratitude to the referee for his/her valuable comments and suggestions.

References

[AACP] A. Al-Salman, H. Al-Qassem, L. Cheng and Y. Pan, *L^p bounds for the function of Marcinkiewicz*, Math. Res. Lett. 9 (2002), 697–700.

[BCP] A. Benedek, A. Calderón and R. Panzone, *Convolution operators on Banach space valued functions*, Proc. Nat. Acad. Sci. U.S.A. 48 (1962), 356–365.

[CZ1] A. P. Calderón and A. Zygmund, *On the existence of certain singular integrals*, Acta Math. 88 (1952), 85–139.

[CZ2] —, —, *On singular integrals*, Amer. J. Math. 78 (1956), 289–309.

[Ch] M. Christ, *Weak type $(1,1)$ bounds for rough operators*, Ann. of Math. 128 (1988), 19–42.

[CR] M. Christ and J. L. Rubio de Francia, *Weak-type $(1,1)$ bounds for rough operators II*, Invent. Math. 93 (1988), 225–237.

[CLW] W. Cohn, G. Lu and P. Wang, *Sub-elliptic global high order Poincaré inequalities in stratified Lie groups and applications*, J. Funct. Anal. 249 (2007), 393–424.

[Con] W. Connett, *Singular integrals near L^1*, in: Proc. Symposia Pure Math. 35, Part I, Amer. Math. Soc., 1979, 163–165.

[Cor] A. Córdoba, *The Kakeya maximal function and the spherical summation multipliers*, Amer. J. Math. 99 (1977), 1–22.

[DFP1] Y. Ding, D. Fan and Y. Pan, *Weighted boundedness for a class of rough Marcinkiewicz integrals*, Indiana Univ. Math. J. 48 (1999), 1037–1055.

[DFP2] —, —, —, *L^p boundedness of Marcinkiewicz integral with Hardy space function kernels*, Acta Math. Sinica (English Ser.) 16 (2000), 593–600.

[FR] E. Fabes and N. Rivière, *Singular integrals with mixed homogeneity*, Studia Math. 27 (1966), 19–38.

[FaSa] D. Fan and S. Sato, *Weak type* $(1,1)$ *estimates for Marcinkiewicz integrals with rough kernels*, Tohoku Math. J. 53 (2001), 265–284.

[FoSt] G. Folland and E. M. Stein, *Hardy Spaces on Homogeneous Groups*, Math. Notes 28, Princeton Univ. Press, Princeton, NJ, 1982.

[GS] L. Grafakos and A. Stefanov, *Convolution Calderón–Zygmund singular integral operators with rough kernels*, in: Analysis of Divergence (Orono, ME, 1997), Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 1999, 119–143.

[Hof] S. Hofmann, *Weak type* $(1,1)$ *boundedness of singular integrals with nonsmooth kernels*, Proc. Amer. Math. Soc. 103 (1988), 260–264.

[Hor] L. Hörmander, *Estimates for translation invariant operators in L^p spaces*, Acta Math. 104 (1960), 93–139.

[RW] F. Ricci and G. Weiss, *A characterization of* $H^1(\Sigma_{n-1})$, in: Proc. Symposia Pure Math. 35, Part I, Amer. Math. Soc., 1979, 289–294.

[Se] A. Seeger, *Singular integral operators with rough convolution kernels*, J. Amer. Math. Soc. 9 (1996), 95–105.

[St1] E. M. Stein, *On the functions of Littlewood–Paley*, *Lusin*, *and Marcinkiewicz*, Trans. Amer. Math. Soc. 88 (1958), 430–466.

[St2] —, *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, 1993.

[T] T. Tao, *The weak-type* $(1,1)$ *of* $L\log L$ *homogeneous convolution operators*, Indiana Univ. Math. J. 48 (1999), 1547–1584.

Yong Ding
School of Mathematical Sciences
Beijing Normal University
Laboratory of Mathematics
and Complex Systems (BNU)
Beijing, 100875, China
E-mail: dingy@bnu.edu.cn

Xinfeng Wu
Department of Mathematics
China University of Mining
and Technology (Beijing)
Beijing, 100083, China
E-mail: wuxf@cumtb.edu.cn

Received December 8, 2008
Revised version August 8, 2009 (6485)