Illinois Journal of Mathematics
Volume 61, Numbers 1–2, Spring and Summer 2017, Pages 53–79
S 0019-2082

BI-PARAMETER LITTLEWOOD–PALEY OPERATORS WITH UPPER DOUBLING MEASURES

MINGMING CAO AND QINGYING XUE

ABSTRACT. Let $\mu = \mu_{n_1} \times \mu_{n_2}$, where μ_{n_1} and μ_{n_2} are upper doubling measures on $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Let the pseudo-accretive function $b = b_1 \otimes b_2$ satisfy a bi-parameter Carleson condition. In this paper, we established the $L^2(\mu)$ boundedness of non-homogeneous Littlewood–Paley g_λ^*-function with non-convolution type kernels on product spaces. This was mainly done by means of dyadic analysis and non-homogenous methods. The result is new even in the setting of Lebesgue measures.

1. Introduction

It is well known that the doubling condition of the measures are essential and necessary in the classical theory of Calderón–Zygmund operators. Certain operators, governed by non-doubling measures, have more recently been investigated. Among such achievements are the celebrated works of Tolsa [27] and Nazarov et al. [21]. In [27], the endpoint estimate and $L^2(\mu)$-boundedness of Cauchey's integral with non-doubling measures were given. In [21], by using a completely different method, a characterization of $L^2(\mu)$-boundedness of Cauchey's integral was presented. What is worth mentioning is that the proofs in [21] remained valid for a quite wide class of Calderón–Zygmund type operators. Additionally, the techniques that they used, including martingale difference decomposition, non-homogeneous analysis and dyadic-probabilistic methods, have been proved to be quite influential and powerful.

Still more recently, Tolsa [28] introduced and investigated systematically the spaces of BMO and Hardy space H^1 with non-doubling measures. More-

Received February 10, 2017; received in final form August 16, 2017.

Q. Xue is the corresponding author.

The authors were supported partly by NSFC (No. 11471041 and 11671039), NSFC-DFG (No. 11761131002).

2010 *Mathematics Subject Classification.* Primary 42B25. Secondary 42B20.

over, a John–Nirenberg inequality and Calderón–Zygmund decomposition suitable for non-doubling measures were presented. Later on, Tolsa [29] gave a $T(1)$ theorem for Calderón–Zygmund operators without doubling assumptions. Subsequently, the famous $T(b)$ theorem given by Christ [8] was extended to non-homogeneous spaces by Nazarov, Treil and Volberg [22]. This was done by showing that the boundedness of a Calderón–Zygmund operator on $L^2(\mu)$ is equivalent to the existence of an accrective system. Since then, there were several important applications for the probabilistic methods and the dyadic analysis in harmonic analysis. Among these applications is the celebrated works of Nazarov et al. [23], in the study of the non-homogeneous theory of Calderón–Zygmund operators. Achievements have also been made in the bi-parameter version of Tb theorem, such as the works of Ou [24], Han et al. [11], Hytönen and Martikainen [14] and the references therein.

This paper is devoted to investigate the Littlewood–Paley g_λ^*-function with non-doubling measures on product spaces. The probabilistic and dyadic analysis techniques will provide a foundation for our analysis. As far as we know, this is the first time to study g_λ^*-function in the simultaneous presence of two attributes: non-homogeneous and bi-parameter.

Before formulating our main result, we first recall some background. The classical higher dimensional Littlewood–Paley g_λ^*-function was first introduced by Stein [26] as follows:

$$g_\lambda^*(f)(x) = \left(\iint_{\mathbb{R}^{n+1}_+} \left(\frac{t}{t+|x-y|} \right)^{n\lambda} \left| \nabla P_t f(y,t) \right|^2 \frac{dy\, dt}{t^{n-1}} \right)^{1/2}, \quad \lambda > 1, n \geq 2,$$

where $\nabla = (\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial t})$, $P_t f(y,t) = p_t * f(y)$ and $p_t(y) = t^{-n} p(y/t)$ denotes the Poisson kernel. The weak $(1,1)$ estimate and strong (p,p) boundedness were obtained by Stein for $\lambda > 2$. In the same paper, Stein also pointed out that the weak $(1,1)$ estimate doesn't hold for $1 < \lambda \leq 2$. As a replacement of the weak $(1,1)$ estimate for $\lambda \leq 2$, the endpoint weak (p,p) estimate was established by Fefferman [10] for $1 < \lambda < 2$, $p > 1$ and $\lambda = 2/p$. After that, the above results were extended to the operators with convolution type kernels. Sharp weighted norm inequalities and even two-weight norm estimates were given by Lerner [15], [16], [17]. The core of Lerner's proofs lie in that the author established the point-wise estimates of Littlewood–Paley operators in terms of the sharp functions and also by certain dyadic sparse operators. It is worth pointing out that Lerner's local mean oscillation decomposition is also valid for the multilinear Littlewood–Paley g_λ^* function defined in [25]. This work was done by Bui and Hormozi in [2]. In the multilinear setting, Shi, Xue and Yabuta [25] showed that the operator g_λ^* is bounded from $L^1(\mathbb{R}^n) \times \cdots \times L^1(\mathbb{R}^n) \to L^{\frac{1}{m},\infty}(\mathbb{R}^n)$ for $\lambda > 2m$. The strong weighted bound and the weak weighted estimates were also be given. Recently, Xue and Yan [31] introduced and studied a more general type of multilinear Littlewood–Paley operators, where the non-convolution type kernel satisfies a class of integral smooth conditions which is much weaker than the standard Calderón–Zygmund kernel conditions. There is a very large literature devoted to the study of Littlewood–Paley operators, in both linear and multilinear cases; see [7], [9] and [30] for more details.

Our object of investigation is the following generalized bi-parameter Littlewood–Paley g_λ^*-function with non-doubling measures.

Definition 1.1. Let $\varepsilon=(\varepsilon_1,\varepsilon_2)$ and $\beta=(\beta_1,\beta_2)$ with $\varepsilon_i,\beta_i>0$ for $i=1,2$. For any $x=(x_1,x_2)\in\mathbb{R}^{n_1+n_2}$, we define the bi-parameter Littlewood–Paley g_λ^*-function by

$$g_{\varepsilon,\beta}^*(f)(x):=\bigg(\iint_{\mathbb{R}_+^{n_2+1}}\iint_{\mathbb{R}_+^{n_1+1}}\vartheta(x,y,t)\big|\theta_{t_1,t_2}f(y_1,y_2)\big|^2 \times\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{dt_1}{t_1}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\frac{dt_2}{t_2}\bigg)^{1/2},$$

where $\vartheta(x,y,t)=\vartheta_1(x_1,y_1,t_1)\vartheta_2(x_2,y_2,t_2)$, and

$$\vartheta_i(x_i,y_i,t_i):=\frac{t_i^{\varepsilon_i}\lambda_{n_i}(x_i,t_i)^{\beta_i}}{t_i^{\varepsilon_i}\lambda_{n_i}(x_i,t_i)^{\beta_i}+|x_i-y_i|^{\varepsilon_i}\lambda_{n_i}(x_i,|x_i-y_i|)^{\beta_i}},\quad i=1,2.$$

The linear term θ_{t_1,t_2} takes the form

$$\theta_{t_1,t_2}f(x)=\iint_{\mathbb{R}^{n_1+n_2}}K_{t_1,t_2}(x_1,x_2,y_1,y_2) \times f(y_1,y_2)\,d\mu_{n_1}(y_1)\,d\mu_{n_2}(y_2),\quad t_1,t_2>0.$$

Definition 1.1 means, of course, that $g_{\varepsilon,\beta}^*$ is a generalization of g_λ^*-function. In fact

(1) In the bi-parameter case, if $d\mu_{n_i}(x_i)=dx_i$, $\lambda_{n_i}(x_i,t_i)=t^{n_i}$, $\varepsilon_i=n_i(\lambda_i-2)$ and $\beta_i=2$, then $g_{\varepsilon,\beta}^*$ coincides with the operator defined and studied in [4].
(2) In the one-parameter case, if $\lambda_{n_1}(x_1,t_1)=t_1^{m_1}$, $\varepsilon_1=m_1(\lambda_1-2)$ and $\beta_1=2$, then the operator $g_{\varepsilon,\beta}^*$ with non-convolution type kernel is just the one introduced in [5] and [6].
(3) In the one-parameter case, if we take simply $d\mu_{n_1}(x_1)=dx_1$, $\lambda_{n_1}(x_1,t_1)=t^{n_1}$, $\varepsilon_1=n_1(\lambda_1-2)$, $\beta_1=2$ and $K_t(y,z)=p_t(y-z)$, where p is the classical Poisson kernel, then the above operator coincides with the classical g_λ^*-function introduced by Stein [26] in 1961 and later studied by Fefferman [10] in 1970, Muckenhoupt and Wheeden [20] in 1974, etc.

A technical condition which helps to underly the analysis, and which may be of some independent interest, is the following bi-parameter Carleson condition.

Definition 1.2. Let $\mu = \mu_{n_1} \times \mu_{n_2}$, $\mathscr{D} = \mathscr{D}_{n_1} \times \mathscr{D}_{n_2}$, where $\mathscr{D}_{n_1}$ is a dyadic grid on $\mathbb{R}^{n_1}$ and $\mathscr{D}_{n_2}$ is a dyadic grid on $\mathbb{R}^{n_2}$. Denote

$$\mathcal{C}_{IJ}^b = \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \big|\theta_{t_1,t_2} b(y)\big|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2}.$$

Then, b is said to satisfy the *bi-parameter Carleson condition* if: for every $\mathscr{D}$ there holds that

$$\sum_{\substack{I\times J \in \mathscr{D} \\ I \times J \subset \Omega}} \mathcal{C}_{IJ}^b \lesssim \mu(\Omega) \tag{1.1}$$

for all sets $\Omega \subset \mathbb{R}^{n_1+n_2}$ such that $\mu(\Omega) < \infty$ and such that for every $x \in \Omega$ there exists $I \times J \in \mathscr{D}$ so that $x \in I \times J \subset \Omega$.

The main result of this paper is the following theorem.

Theorem 1.3. *Let $\varepsilon_i > 0$, $\beta_i \geq 3$ and $0 < \alpha_i < \varepsilon_i/2$ for $i = 1, 2$. Let $\mu = \mu_{n_1} \times \mu_{n_2}$, where μ_{n_1} and μ_{n_2} are upper doubling measures on $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Let b be a pseudo-accretive function defined on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Suppose that $\{K_{t_1,t_2}\}$ satisfy the Assumptions 2.3–2.4 in Section 2 and b satisfies the bi-parameter Carleson condition. Then there holds that*

$$\big\| g_{\varepsilon,\beta}^*(f) \big\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}. \tag{1.2}$$

Remark 1.4. We should point out that the bi-parameter Carleson condition is necessary in the following sense : $K_{t_1,t_2} = K_{t_1} \otimes K_{t_2}$, the one-parameter kernels satisfy the size condition and corresponding square operators are bounded on $L^2(\mu)$. Thus, in this sense, Theorem 1.3 is a characterization of L^2 boundedness of $g_{\varepsilon,\beta}^*$. We only give the outline of the proof of the necessity in Section 7. We adopt the similar strategy as the proof of the necessity [18] with slight modifications. One needs to employ Journés covering lemma with general product measures, which was already known. Some non-homogeneous calculations needed are essentially contained in Section 3.

Notation 1.5. We write $A \lesssim B$, if there is a constant $C > 0$ so that $A \leq CB$. We may also write $A \simeq B$ if $B \lesssim A \lesssim B$.

We then set some dyadic notation. For cubes I and J we denote

- $\ell(I)$ is the side-length of I;
- $d(I,J)$ denotes the distance between the cubes I and J;
- $D(I,J) := \ell(I) + \ell(J) + d(I,J)$ is the long distance;
- $W_I := I \times (\ell(I)/2, \ell(I)]$ is the Whitney region associated with I;
- $\widehat{I} = I \times (0, \ell(I)]$ is the Carleson box over I;
- $I^{(k)}$ denotes the unique dyadic cube for which $\ell(I^{(k)}) = 2^k \ell(I)$ and $I \subset I^{(k)}$.

The organizations of this paper are as follows: In Section 2, we first introduce some notations and definitions which will be used later. Then, after defining the b-adapted Haar functions, we will make the initial reductions for the proof of Theorem 1.3. Some useful estimates and key lemmas will be presented in Section 3. Sections 4–6 will be devoted to dealing with the L^2 estimate that is needed to complete the corresponding simplified proof of our main theorem. In Section 7, we discuss the necessity of bi-parameter Carleson condition.

2. Preliminaries

In this section, our goal is to introduce some fundamental tools which will be used later in the proof of Theorem 1.3. With these main tools in hand, we will try to give the reduction of the initial estimate. We begin by considering the following class of measures.

DEFINITION 2.1. Let $\lambda : \mathbb{R}^n \times (0,\infty) \to (0,\infty)$ be a function so that $r \mapsto \lambda(x,r)$ is non-decreasing and $\lambda(x,2r) \le C_\lambda \lambda(x,r)$ for all $x \in \mathbb{R}^n$ and $r > 0$. We say that a Borel measure μ in $\mathbb{R}^n$ is upper doubling with the dominating function λ, if $\mu(B(x,r)) \le \lambda(x,r)$ for all $x \in \mathbb{R}^n$ and $r > 0$. We set $d_\lambda = \log_2 C_\lambda$.

The property $\lambda(x,|x-y|) \simeq \lambda(y,|x-y|)$ can be assumed without loss of generality. Moreover, we may always assume that dominating functions λ satisfy the additional symmetry property $\lambda(x,r) \le C\lambda(y,r)$ if $|x-y| \le r$.

The concept of upper doubling measures was first introduced by Hytönen [12]. In addition, the theory of Calderón–Zygmund singular integrals in this setting was investigated by Bui and Duong [1]. In terms of square functions with upper doubling measures, the authors in [19] gave a characterization of L^2-boundedness.

From now on, let $\mu = \mu_{n_1} \times \mu_{n_2}$, where μ_{n_1} and μ_{n_2} are upper doubling measures on $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$ respectively. The corresponding dominating functions are denoted by λ_{n_1} and λ_{n_2}. We use, for minor convenience, ℓ^∞ metrics on $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$.

DEFINITION 2.2. A function $b \in L^\infty(\mu)$ is called pseudo-accretive if there is a positive constant C such that for any rectangle $R \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ with sides parallel to axes,

$$\frac{1}{\mu(R)} \left| \int_R b(x)\, d\mu(x) \right| > C.$$

In this paper, we will only discuss the case $b = b_1 \otimes b_2$, where b_1 and b_2 are in $L^\infty(\mu_{n_1})$ and $L^\infty(\mu_{n_2})$, respectively. Then, the pseudo-accretivity and boundedness of b imply that there exists a constant C such that for any cubes

$I \subset \mathbb{R}^{n_1}$, $J \subset \mathbb{R}^{n_2}$, the following inequalities are valued.

$$\frac{1}{\mu_{n_1}(I)}\left|\int_I b_1\, d\mu_{n_1}\right| > C \quad \text{and} \quad \frac{1}{\mu_{n_2}(J)}\left|\int_J b_2\, d\mu_{n_2}\right| > C.$$

That is, b_1 and b_2 are both pseudo-accretive in the classical sense.

Next, we introduce some appropriate assumptions on the kernels that we need throughout the argument. We always assume that the fixed numbers α_1 and α_2 are positive.

Assumption 2.3. The kernel $K_{t_1,t_2} : \mathbb{R}^{n_1+n_2} \times \mathbb{R}^{n_1+n_2} \to \mathbb{C}$ is assumed to satisfy the following estimates:

(1) Size condition:

$$\begin{aligned}\left|K_{t_1,t_2}(x,y)\right| \lesssim{}& \frac{t_1^{\alpha_1}}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1) + |x_1-y_1|^{\alpha_1}\lambda_{n_1}(x_1,|x_1-y_1|)}\\ &\times \frac{t_2^{\alpha_2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2) + |x_2-y_2|^{\alpha_2}\lambda_{n_2}(x_2,|x_2-y_2|)}.\end{aligned}$$

(2) Hölder condition:

$$\begin{aligned}&\left|K_{t_1,t_2}(x,y) - K_{t_1,t_2}\big(x,(y_1,y_2')\big) - K_{t_1,t_2}\big(x,(y_1',y_2)\big) + K_{t_1,t_2}\big(x,y'\big)\right|\\ &\lesssim \frac{|y_1-y_1'|^{\alpha_1}}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1) + |x_1-y_1|^{\alpha_1}\lambda_{n_1}(x_1,|x_1-y_1|)}\\ &\quad\times \frac{|y_2-y_2'|^{\alpha_2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2) + |x_2-y_2|^{\alpha_2}\lambda_{n_2}(x_2,|x_2-y_2|)},\end{aligned}$$

whenever $|y_1-y_1'| < t_1/2$ and $|y_2-y_2'| < t_2/2$.

(3) Mixed Hölder and size conditions:

$$\begin{aligned}&\left|K_{t_1,t_2}(x,y) - K_{t_1,t_2}\big(x,(y_1,y_2')\big)\right|\\ &\lesssim \frac{t_1^{\alpha_1}}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1) + |x_1-y_1|^{\alpha_1}\lambda_{n_1}(x_1,|x_1-y_1|)}\\ &\quad\times \frac{|y_2-y_2'|^{\alpha_2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2) + |x_2-y_2|^{\alpha_2}\lambda_{n_2}(x_2,|x_2-y_2|)},\end{aligned}$$

whenever $|y_2-y_2'| < t_2/2$. And

$$\begin{aligned}&\left|K_{t_1,t_2}(x,y) - K_{t_1,t_2}\big(x,(y_1',y_2)\big)\right|\\ &\lesssim \frac{|y_1-y_1'|^{\alpha_1}}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1) + |x_1-y_1|^{\alpha_1}\lambda_{n_1}(x_1,|x_1-y_1|)}\\ &\quad\times \frac{t_2^{\alpha_2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2) + |x_2-y_2|^{\alpha_2}\lambda_{n_2}(x_2,|x_2-y_2|)},\end{aligned}$$

whenever $|y_1-y_1'| < t_1/2$.

Assumption 2.4. For every cube $I \subset \mathbb{R}^{n_1}$ and $J \subset \mathbb{R}^{n_2}$, there holds that

(1) Mixed Carleson and size conditions:

$$\bigg(\iint_{\widehat{I}}\int_{\mathbb{R}^{n_1}} \vartheta_1(x_1,y_1,t_1)\bigg|\int_I b_1(z_1)K_{t_1,t_2}(y,z_1,z_2)\,d\mu_{n_1}(z_1)\bigg|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\bigg)^{1/2} \lesssim \frac{t_2^{\alpha_2}\mu_{n_1}(I)^{1/2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2)+|y_2-z_2|^{\alpha_2}\lambda_{n_2}(x_2,|y_2-z_2|)}$$

and

$$\bigg(\iint_{\widehat{J}}\int_{\mathbb{R}^{n_2}} \vartheta_2(x_2,y_2,t_2)\bigg|\int_J b_2(z_2)K_{t_1,t_2}(y,z_1,z_2)\,d\mu_{n_2}(z_2)\bigg|^2 \times \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2}\bigg)^{1/2} \lesssim \frac{t_1^{\alpha_1}\mu_{n_2}(J)^{1/2}}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1)+|y_1-z_1|^{\alpha_1}\lambda_{n_1}(x_1,|y_1-z_1|)}.$$

(2) Mixed Carleson and Hölder conditions:

$$\bigg(\iint_{\widehat{I}}\int_{\mathbb{R}^{n_1}} \vartheta_1(x_1,y_1,t_1)\bigg|\int_I b_1(z_1)\big[K_{t_1,t_2}(y,z_1,z_2)-K_{t_1,t_2}\big(y,z_1,z_2'\big)\big]\,d\mu_{n_1}(z_1)\bigg|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\bigg)^{1/2} \lesssim \frac{|z_2-z_2'|^{\alpha_2}\mu_{n_1}(I)^{1/2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2)+|y_2-z_2|^{\alpha_2}\lambda_{n_2}(x_2,|y_2-z_2|)},$$

whenever $|z_2-z_2'|<t_2/2$. And

$$\bigg(\iint_{\widehat{J}}\int_{\mathbb{R}^{n_2}} \vartheta_2(x_2,y_2,t_2)\bigg|\int_J b_2(z_2)\big[K_{t_1,t_2}(y,z_1,z_2)-K_{t_1,t_2}\big(x,z_1',z_2\big)\big]\,d\mu_{n_2}(z_2)\bigg|^2 \times \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2}\bigg)^{1/2} \lesssim \frac{|z_1-z_1'|^{\alpha_1}\mu_{n_2}(J)^{1/2}}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1)+|y_1-z_1|^{\alpha_1}\lambda_{n_1}(x_1,|y_1-z_1|)},$$

whenever $|z_1-z_1'|<t_1/2$.

2.5. Random dyadic grids. We are now in the position to introduce the fundamental technique, random dyadic grids. Let $\omega_n=\{\omega_n^j\}_{j\in\mathbb{Z}}$, where $\omega_n^j\in\{0,1\}^n$. Let $\mathscr{D}_n^0$ be the standard dyadic grids on $\mathbb{R}^n$. That is,

$$\mathscr{D}_n^0:=\bigcup_{k\in\mathbb{Z}}\mathscr{D}_{n,k}^0,\quad \mathscr{D}_{n,k}^0:=\big\{2^k\big([0,1)^n+m\big);m\in\mathbb{Z}^n\big\}.$$

In $\mathbb{R}^n$, we define the new dyadic grid

$$\mathscr{D}_n := \mathscr{D}_{\omega_n} = \big\{I + \omega_n; I \in \mathscr{D}_n^0\big\} := \bigg\{I + \sum_{j:2^{-j}<\ell(I)} 2^{-j}\omega_n^j; I \in \mathscr{D}_n^0\bigg\}.$$

There is a natural product probability structure on $(\{0,1\}^n)^{\mathbb{Z}}$. So we have independent random dyadic grids $\mathscr{D}_{n_1}$ and $\mathscr{D}_{n_2}$ in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively. Even for $n_1 = n_2$, we still need two independent grids.

2.6. Good cubes. A cube $I \in \mathscr{D}_n$ is said to be *bad* if there exists a $J \in \mathscr{D}_n$ with $\ell(J) \geq 2^r \ell(I)$ such that $\operatorname{dist}(I, \partial J) \leq \ell(I)^{\gamma_n} \ell(J)^{1-\gamma_n}$. Otherwise, I is called *good*. Here $r \in \mathbb{Z}_+$ and $\gamma_n \in (0, \frac{1}{2})$ are given parameters.

Denote $\pi_{\text{good}}^{n_i} = \mathbb{P}_{\omega_{n_i}}(I + \omega_{n_i} \text{ is good}) = \mathbb{E}_{\omega_{n_i}}(\mathbf{1}_{\text{good}}(I + \omega_{n_i}))$. Then $\pi_{\text{good}}^{n_i}$ is independent of $I \in \mathscr{D}_{n_i}^0$, and the parameter r is a fixed constant so that $\pi_{\text{good}}^{n_1}, \pi_{\text{good}}^{n_2} > 0$. Throughout this article, we take $\gamma_{n_i} = \frac{\alpha_i}{2(d_{\lambda_{n_i}} + \alpha_i)}$, where $\alpha_i > 0$ appears in the kernel estimates. It is important to observe that the position and goodness of a cube $I \in \mathscr{D}_{n_1}^0$ are independent.

2.7. b-adapted Haar functions. The abbreviation $b_1(E) := \int_E b_1 \, d\mu_{n_1}$ will be used. For each $I \in \mathscr{D}_{n_1}$, we denote its dyadic children by $I_1, \ldots, I_{2^{n_1}}$. We index $\{I_j\}$ in such a way that

$$\big|b_1\big(I_j^*\big)\big| \geq \big[1 - (k-1)2^{-n_1}\big]\mu_{n_1}(I), \quad I_j^* = \bigcup_{k=j}^{2^{n_1}} I_k, j = 1, \ldots, 2^{n_1}.$$

The existence of such way was shown in Lemma 4.2 [13]. The b_1-adapted Haar function is defined by

$$\varphi_{I,j}^{b_1} := \left(\frac{b_1(I_j) b_1(I_{j+1}^*)}{b_1(I_j^*)}\right)^{1/2} \left(\frac{\mathbf{1}_{I_j}}{b_1(I_j)} - \frac{\mathbf{1}_{I_{j+1}^*}}{b_1(I_{j+1}^*)}\right), \quad j = 1, \ldots, 2^{n_1} - 1.$$

Similarly, we can define the function $\psi_{J,k}^{b_2}$ with respect to b_2 and $J \in \mathscr{D}_{n_2}$.

The adapted Haar functions enjoy the following properties:

(1) $\int_{\mathbb{R}^{n_1}} b_1 \varphi_{I,j}^{b_1} \, d\mu_{n_1} = 0$.
(2) $|\varphi_{I,j}^{b_1}| \simeq \mu_{n_1}(I_j)^{1/2}\big(\frac{\mathbf{1}_{I_j}}{b_1(I_j)} + \frac{\mathbf{1}_{I_{j+1}^*}}{b_1(I_{j+1}^*)}\big)$.
(3) $\|\varphi_{I,j}^{b_1}\|_{L^p(\mu_{n_1})} \simeq \mu_{n_1}(I_j)^{1/p-1/2}$, $p \in [1, \infty]$.
(4) The similar above properties hold for $\psi_{J,k}^{b_2}$ as well.
(5) For any $f \in L^2(\mu)$, there holds that

$$f = \sum_{j=1}^{2^{n_1}-1} \sum_{k=1}^{2^{n_2}-1} \sum_{I \in \mathscr{D}_{n_1}} \sum_{J \in \mathscr{D}_{n_2}} \big\langle f, \varphi_{I,j}^{b_1} \otimes \psi_{J,k}^{b_2} \big\rangle b \cdot \varphi_{I,j}^{b_1} \otimes \psi_{J,k}^{b_2}.$$

The properties (1)–(4) can be found in Proposition 4.3 [13]. Property (5) can be verified by iteration of the one-parameter argument.

2.8. INITIAL REDUCTIONS. Let $f \in L^2(\mu)$, $I_1, I_2 \in \mathscr{D}_{n_1}$ and $J_1, J_2 \in \mathscr{D}_{n_2}$. Note that the position and goodness of $I + \omega_{n_1}$ are independent. Therefore, one can write

$$\left\|g^*_{\varepsilon,\beta}(f)\right\|^2_{L^2(\mu)} = c_{n_1,n_2}\mathbb{E}_{\omega_{n_1}}\mathbb{E}_{\omega_{n_2}}\mathscr{G}_{\omega_{n_1},\omega_{n_2}},$$

where $c_{n_1,n_2} = (\pi^{n_1}_{\text{good}} \cdot \pi^{n_2}_{\text{good}})^{-1}$ and

$$\mathscr{G}_{\omega_{n_1},\omega_{n_2}} := \sum_{I_2,J_2\colon \text{ good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t)\left|\theta_{t_1,t_2}f(y)\right|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2}.$$

Indeed, to get this equality, we only need to apply the similar argument to the one-parameter case twice. For more details in the one-parameter setting, see [3]. Then, applying b-adapted Haar decomposition of f (suppressing the finite j,k summation), we may further write

$$\mathscr{G}_{\omega_{n_1},\omega_{n_2}} := \sum_{I_2,J_2\colon \text{ good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \times \left|\sum_{I_1,J_1} f_{I_1J_1}\theta_{t_1,t_2}\big(b\cdot\varphi^{b_1}_{I_1}\otimes\psi^{b_2}_{J_1}\big)(y)\right|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2},$$

where $f_{I_1J_1} = \langle f, \varphi^{b_1}_{I_1}\otimes\psi^{b_2}_{J_1}\rangle$. When ω_{n_1} and ω_{n_2} are fixed, we denote $\mathscr{G}_{\omega_{n_1},\omega_{n_2}}$ by $\mathscr{G}$. Consequently, it is enough to show $\mathscr{G} \lesssim \|f\|^2_{L^2(\mu)}$, where the implied constant is independent of ω_{n_1} and ω_{n_2}.

We can perform the decomposition

$$\mathscr{G} \lesssim \mathscr{G}_{<,<} + \mathscr{G}_{<,\geq} + \mathscr{G}_{\geq,<} + \mathscr{G}_{\geq,\geq},$$

where

$$\mathscr{G}_{<,<} := \sum_{I_2,J_2\colon \text{ good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta_1(x,y,t) \times \left|\sum_{\substack{I_1,J_1\\ \ell(I_1)<\ell(I_2)\\ \ell(J_1)<\ell(J_2)}} f_{I_1J_1}\theta_{t_1,t_2}\big(b\cdot\varphi^{b_1}_{I_1}\otimes\psi^{b_2}_{J_1}\big)(y)\right|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2}$$

and the others are completely similar.

Sequentially, it suffices to focus on controlling the four pieces: $\mathscr{G}_{<,<}$, $\mathscr{G}_{<,\geq}$, $\mathscr{G}_{\geq,<}$, $\mathscr{G}_{\geq,\geq}$ in the following sections.

3. Some standard estimates

This section will be devoted to establishing some general and useful calculations, which will be employed at certain steps of the proof of Theorem 1.3. Some estimates with new measures μ are essentially different from those of Lesbegue measures. For example, the measure of two balls $B(x,t)$ and $B(y,t)$ may have no relationship and thus usually are not equal.

LEMMA 3.1. *Let μ_n be an upper doubling measure with the dominating function λ_n. Then for any $\varepsilon_1 > 0$ and $\beta_1 \geq 1$, there holds that*

$$\int_{|y-y_0|\geq r} \frac{|y-y_0|^{-\varepsilon_1}}{\lambda_n(y_0,|y-y_0|)^{\beta_1}}\, d\mu_n(y) \lesssim r^{-\varepsilon_1}\lambda_n(y_0,r)^{1-\beta_1}, \tag{3.1}$$

$$\int_{\mathbb{R}^{n_1}} \vartheta_1(x_1,y_1,t_1)\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \lesssim 1. \tag{3.2}$$

Proof. Since (3.2) follows from (3.1) and (3.1) can be obtained by a standard argument, we omit the proof. □

LEMMA 3.2. *Let $0 < \alpha_1 < \varepsilon_1/2$ and $\beta_1 \geq 3$. Assume that $I_1, I_2 \in \mathscr{D}_{n_1}$ and $E \subset \mathbb{R}^{n_1}$. There holds that*

$$\mathscr{F}_{n_1,\alpha_1}(E,x_1,t_1) \lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1),$$

where

$$\mathscr{F}_{n_1,\alpha_1}(E,x_1,t_1) := \left(\int_{\mathbb{R}^{n_1}} \vartheta_1(x_1,y_1,t_1)\mathcal{F}_{n_1,\alpha_1}(E,x_1,y_1,t_1)^2 \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\right)^{1/2},$$

$$\mathcal{F}_{n_1,\alpha_1}(E,x_1,y_1,t_1) := \int_E \frac{d\mu_{n_1}(z_1)}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1) + |y_1-z_1|^{\alpha_1}\lambda_{n_1}(x_1,|y_1-z_1|)}.$$

Proof. For given $y_1 \in \mathbb{R}^n$, denote

$$E_1 := \{z_1 \in E; |z_1-x_1| \geq 2|x_1-y_1|\},$$
$$E_2 := \{z_1 \in E; |z_1-x_1| < 2|x_1-y_1|\}.$$

Then it holds that

$$\mathscr{F}_{n_1,\alpha_1}(E,x_1,t_1) \leq \mathscr{F}_{n_1,\alpha_1}(E_1,x_1,t_1) + \mathscr{F}_{n_1,\alpha_1}(E_2,x_1,t_1).$$

If $|z_1-x_1| \geq 2|x_1-y_1|$, then

$$|y_1-z_1| \geq |z_1-x_1| - |x_1-y_1| \geq \frac{1}{2}|z_1-x_1|,$$

which implies that

$$\begin{aligned}\mathscr{F}_{n_1,\alpha_1}(E_1,x_1,t_1) &\lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1)\left(\int_{\mathbb{R}^{n_1}} \vartheta_1(x_1,y_1,t_1)\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\right)^{1/2} \\ &\lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1).\end{aligned}$$

If $|z_1 - x_1| < 2|x_1 - y_1|$ and $|y_1 - x_1| \le t_1$, then $|z_1 - x_1| \lesssim t_1$ and

$$t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1) + |x_1 - z_1|^{\alpha_1}\lambda_{n_1}\big(x_1,|x_1 - z_1|\big) \simeq t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1).$$

Accordingly, it yields that

$$\mathcal{F}_{n_1,\alpha_1}(E_2,x_1,y_1,t_1) \lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1).$$

This inequality together with (3.2), yields that

$$\mathscr{F}_{n_1,\alpha_1}(E_2,x_1,t_1) \lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1).$$

If $|z_1 - x_1| < 2|x_1 - y_1|$ and $|y_1 - x_1| > t_1$, then by (3.1), one may deduce that

$$\begin{aligned}
\mathscr{F}_{n_1,\alpha_1}(E_2,x_1,t_1) &\lesssim \frac{\mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1)}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1)} \\
&\quad\times\bigg(\int_{|y_1-x_1|>t_1} \vartheta_1(x_1,y_1,t_1)|y_1 - x_1|^{2\alpha_1}\lambda_{n_1}\big(x_1,|y_1-x_1|\big)^2 \\
&\quad\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\bigg)^{1/2} \\
&\lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1)t_1^{(\varepsilon_1-2\alpha_1)/2}\lambda_{n_1}(x_1,t_1)^{(\beta_1-3)/2} \\
&\quad\times\bigg(\int_{|y_1-x_1|>t_1} \frac{|y_1-x_1|^{-(\varepsilon_1-2\alpha_1)}}{\lambda_{n_1}(x_1,|y_1-x_1|)^{\beta_1-2}}\,d\mu_{n_1(y_1)}\bigg)^{1/2} \\
&\lesssim \mathcal{F}_{n_1,\alpha_1}(E,x_1,x_1,t_1),
\end{aligned}$$

where in the last step we have used the condition that $\varepsilon_1 > 2\alpha_1$ and $\beta_1 \ge 3$. This completes the proof. □

LEMMA 3.3. *Let $k \ge 1$, $I \in \mathscr{D}_{n_1}$ be a good cube and $(x_1,t_1) \in W_I$. Set*

$$\mathscr{S}_k(x_1) := \ell(I)^{\alpha_1}\mathscr{F}_{n_1,\alpha_1}\Big(\big(I^{(k-1)}\big)^c, x_1, t_1\Big).$$

Then we have the geometric decay $\mathscr{S}_k(x_1) \lesssim 2^{-\alpha_1 k/2}$.

Proof. If $k \le r$, by Lemma 3.2 and the inequality (3.1), one may obtain that

$$\begin{aligned}
\mathscr{S}_k(x_1) &\le \ell(I)^{\alpha_1}\int_{2I}\frac{d\mu_{n_1}(z_1)}{t_1^{\alpha_1}\lambda_{n_1}(x_1,t_1)} + \ell(I)^{\alpha_1}\int_{(2I)^c}\frac{|z_1-x_1|^{-\alpha_1}}{\lambda_{n_1}(x_1,|z_1-x_1|)}\,d\mu_{n_1}(z_1) \\
&\lesssim 1 \simeq 2^{-\alpha_1 k/2}.
\end{aligned}$$

If $k > r$, Lemma 3.2 implies that

$$\mathscr{S}_k(x_1) \lesssim \ell(I)^{\alpha_1}\int_{(I^{(k-1)})^c}\frac{|z_1-x_1|^{-\alpha_1}}{\lambda_{n_1}(x_1,|z_1-x_1|)}\,d\mu_{n_1}(z_1) := A(x_1).$$

By the goodness of I, it yields that

$$d\Big(I,\big(I^{(k-1)}\big)^c\Big) > \ell(I)^{\gamma_{n_1}}\ell\Big(I^{(k-1)}\Big)^{1-\gamma_{n_1}} = 2^{(k-1)(1-\gamma_{n_1})}\ell(I) \gtrsim 2^{k/2}\ell(I).$$

Thus together with (3.1), we get

$$\begin{aligned}\mathscr{S}_k(x_1) \lesssim A(x_1) &\le \ell(I)^{\alpha_1} \int_{B(x_1,d(I,(I^{(k-1)})^c))^c} \frac{|z_1-x_1|^{-\alpha_1}}{\lambda_{n_1}(x_1,|z_1-x_1|)}\, d\mu_{n_1}(z_1)\\ &\lesssim \ell(I)^{\alpha_1} d\big(I,\big(I^{(k-1)}\big)^c\big)^{-\alpha_1} \lesssim 2^{-\alpha_1 k/2}. \qquad \square\end{aligned}$$

LEMMA 3.4. *Let $J_1 \in \mathscr{D}_{n_2}$ be a fixed dyadic cube. Denote*

$$\begin{aligned}a_I := \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} &\vartheta(x,y,t)\big|\theta_{t_1,t_2}\big(b_1 \otimes \big(b_2 \psi_{J_1}^{b_2}\big)\big)(y)\big|^2\\ &\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1}.\end{aligned}$$

Then $\{a_I\}_{I\in\mathscr{D}_{n_1}}$ is a Carleson sequence. Rather, there holds for any good cube $I \in \mathscr{D}_{n_1}$

$$\sum_{I':I'\subset I} a_{I'} \lesssim \mu_{n_1}(I)\big(\mu_{n_2}(J_1)^{-1/2}\ell(J_1)^{\alpha_2}\mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\big)^2. \tag{3.3}$$

Proof. The first step is to split

$$\sum_{I':I'\subset I} a_{I'} \lesssim \mathcal{S}_1(x_2,t_2) + \mathcal{S}_2(x_2,t_2),$$

where

$$\begin{aligned}\mathcal{S}_1(x_2,t_2) = \iint_{\widehat{3I}} \iint_{\mathbb{R}^{n_1+n_2}} &\vartheta(x,y,t)\big|\theta_{t_1,t_2}\big((b_1\mathbf{1}_{3I}) \otimes \big(b_2 \psi_{J_1}^{b_2}\big)\big)(y)\big|^2\\ &\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1}\end{aligned}$$

and

$$\begin{aligned}\mathcal{S}_2(x_2,t_2) = \iint_{\widehat{I}} \iint_{\mathbb{R}^{n_1+n_2}} &\vartheta(x,y,t)\big|\theta_{t_1,t_2}\big((b_1\mathbf{1}_{(3I)^c}) \otimes \big(b_2 \psi_{J_1}^{b_2}\big)\big)(y)\big|^2\\ &\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1}\\ := \iint_{\widehat{I}} &H(x,t)\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1}.\end{aligned}$$

Using the Minkowski's inequality and Carleson and Hölder conditions, it follows that

$$\begin{aligned}\mathcal{S}_1(x_2,t_2) \lesssim \mu_{n_2}(J_1)^{-1} \int_{\mathbb{R}^{n_2}} \iint_{\widehat{3I}} \int_{\mathbb{R}^{n_1}} &\bigg| \int_{3I\times J_1} b_1(z_1)\\ &\times \big[K_{t_1,t_2}\big(y,(z_1,z_2)\big) - K_{t_1,t_2}\big(y,(z_1,z_2+c_{J_1})\big)\big]\, d\mu(z)\bigg|^2\\ &\times \vartheta(x,y,t) \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\end{aligned}$$

$$
\begin{aligned}
&\leq \mu_{n_2}(J_1)^{-1} \int_{\mathbb{R}^{n_2}} \bigg\{ \int_{J_1} \bigg(\iint_{\widehat{3I}} \int_{\mathbb{R}^{n_1}} \bigg| \int_{3I} b_1(z_1) \\
&\quad \times \big[K_{t_1,t_2}\big(y,(z_1,z_2)\big) - K_{t_1,t_2}\big(y,(z_1,z_2+c_{J_1})\big)\big]\, d\mu_{n_1}(z_1) \bigg|^2 \\
&\quad \times \vartheta_1(x_1,y_1,t_1) \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \bigg)^{1/2} d\mu_{n_2}(z_2) \bigg\}^2 \\
&\quad \times \vartheta_2(x_2,y_2,t_2) \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \\
&\lesssim \mu_{n_1}(I) \big(\mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\big)^2.
\end{aligned}
$$

The mixed Hölder and size estimate give that

$$
\begin{aligned}
&\big|\theta_{t_1,t_2}\big((b_1 \mathbf{1}_{(3I)^c}) \otimes \big(b_2 \psi_{J_1}^{b_2}\big)\big)(y)\big| \\
&\lesssim \int_{(3I)^c} \frac{t_1^{\alpha_1}\, d\mu_{n_1}(z_1)}{t_1^{\alpha_1} \lambda_{n_1}(x_1,t_1) + |y_1 - z_1|^{\alpha_1} \lambda_{n_1}(x_1,|y_1-z_1|)} \\
&\quad \times \mu_{n_2}(J_1)^{-1/2} \int_{J_1} \frac{\ell(J_1)^{\alpha_2}\, d\mu_{n_2}(z_2)}{t_2^{\alpha_2} \lambda_{n_2}(x_2,t_2) + |y_2 - z_2|^{\alpha_2} \lambda_{n_2}(x_2,|y_2-z_2|)}.
\end{aligned}
$$

Thus, by Lemma 3.3, one can deduce that

$$
\begin{aligned}
H(x,t) &\lesssim \mathscr{S}_k(x_1)^2 \big(\mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\big)^2 \\
&\lesssim t_1^{2\alpha_1} \ell(I)^{-2\alpha_1} \big(\mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\big)^2.
\end{aligned}
$$

Therefore, one obtains that

$$
\begin{aligned}
\mathcal{S}_2(x_2,t_2) &= \iint_{\widehat{I}} H(x,t)\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \\
&\lesssim \mu_{n_1}(I) \ell(I)^{-2\alpha_1} \int_0^{\ell(I)} t_1^{2\alpha_1 - 1} dt_1 \\
&\quad \cdot \big(\mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\big)^2 \\
&\lesssim \mu_{n_1}(I) \big(\mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\big)^2.
\end{aligned}
$$

This finishes the proof of Lemma 3.4. □

LEMMA 3.5. *Let $k \geq 1$ and $I \in \mathscr{D}_{n_1}$ be a good cube and and $(x_1,t_1) \in W_I$. We have the following Carleson estimate*:

$$
\sum_{J' : J' \subset J} a_{J'} \lesssim 2^{-\alpha_1 k} \mu_{n_1}\big(I^{(k)}\big)^{-1} \mu_{n_2}(J), \tag{3.4}
$$

where ξ_I^k *is given in* (5.4) *below and*

$$a_J := \iint_{W_J} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \big|\theta_{t_1,t_2}\big(\big(b_1 \xi_I^k\big) \otimes b_2\big)(y)\big|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2}.$$

The proof of Lemma 3.5 is similar to Lemma 3.4. The size condition and mixed Carleson and size estimate need to be used. In addition, Lemma 3.3 is needed.

We will also need the following lemma, which can be found in [22].

LEMMA 3.6. *We set*

$$A_{I_1 I_2} = \frac{\ell(I_1)^{\alpha/2} \ell(I_2)^{\alpha/2}}{D(I_1,I_2)^\alpha \sup_{z_1 \in I_1 \cup I_2} \lambda_n(z_1, D(I_1,I_2))} \mu_n(I_1)^{1/2} \mu_n(I_2)^{1/2},$$

where $\alpha > 0$ *and* $D(I_1,I_2) = \ell(I_1) + \ell(I_2) + d(I_1,I_2)$, $I_1, I_2 \in \mathcal{D}_n$. *Then for any* $x_{I_1}, y_{I_2} \geq 0$, *we have the following estimate*

$$\bigg(\sum_{I_1,I_2} A_{I_1 I_2} x_{I_1} y_{I_2}\bigg)^2 \lesssim \sum_{I_1} x_{I_1}^2 \times \sum_{I_2} y_{I_2}^2.$$

In particular, there holds that

$$\sum_{I_2} \bigg(\sum_{I_1} A_{I_1 I_2} x_{I_1}\bigg)^2 \lesssim \sum_{I_1} x_{I_1}^2.$$

Finally, we present a dyadic Carleson embedding theorem, which was proved in [21].

LEMMA 3.7. *Let* ν *be a measure on* $\mathbb{R}^n$. *If the numbers* $a_Q \geq 0$, $Q \in \mathscr{D} \subset \mathbb{R}^n$ *satisfy the following Carleson measure condition*

$$\sum_{Q' \subset Q} a_{Q'} \leq \nu(Q), \quad \text{for each } Q \in \mathscr{D},$$

then for any $f \in L^2(\nu)$

$$\sum_{Q \in \mathscr{D}} a_Q \big|\langle f \rangle_Q^\nu\big|^2 \leq 4 \|f\|_{L^2(\nu)}^2.$$

4. The case: $\ell(I_1) < \ell(I_2)$ and $\ell(J_1) < \ell(J_2)$

The kernel $K_{t_1,t_2}(y,z)$ can be changed into

$$K_{t_1,t_2}(y,z) - K_{t_1,t_2}\big(y,(z_1,c_{J_1})\big) - K_{t_1,t_2}\big(y,(c_{I_1},z_2)\big) + K_{t_1,t_2}\big(y,(c_{I_1},c_{J_1})\big),$$

which is provided by the cancellation properties of the adapted Haar functions

$$\int_{\mathbb{R}^{n_1}} b_1 \varphi_{I_1}^{b_1} \, d\mu_{n_1} = \int_{\mathbb{R}^{n_2}} b_2 \psi_{J_1}^{b_2} \, d\mu_{n_2} = 0.$$

From the full Hölder condition of the kernel K_{t_1,t_2}, it follows that

$$\begin{aligned}\big|\theta_{t_1,t_2}\big(b\cdot\varphi_{I_1}^{b_1}\otimes\psi_{J_1}^{b_2}\big)(y)\big| &\lesssim \mu_{n_1}(I_1)^{-1/2}\ell(I_1)^{\alpha_1}\mathcal{F}_{n_1,\alpha_1}(I_1,x_1,y_1,t_1)\\ &\quad\times\mu_{n_2}(J_1)^{-1/2}\ell(J_1)^{\alpha_2}\mathcal{F}_{n_2,\alpha_2}(J_1,x_2,y_2,t_2).\end{aligned}$$

If $\ell(I_1)<\ell(I_2)$ and $\ell(J_1)<\ell(J_2)$, we get

$$\begin{aligned}\mathcal{G}(x,t) &:= \bigg(\iint_{\mathbb{R}^{n_1+n_2}}\vartheta(x,y,t)\big|\theta_{t_1,t_2}\big(b\cdot\varphi_{I_1}^{b_1}\otimes\psi_{J_1}^{b_2}\big)(y)\big|^2\\ &\qquad\times\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\bigg)^{1/2}\\ &\lesssim \mu_{n_1}(I_1)^{-1/2}\ell(I_1)^{\alpha_1}\mathscr{F}_{n_1,\alpha_1}(I_1,x_1,t_1)\\ &\qquad\cdot\mu_{n_2}(J_1)^{-1/2}\ell(J_1)^{\alpha_2}\mathscr{F}_{n_2,\alpha_2}(I_2,x_2,t_2)\\ &\lesssim A_{I_1I_2}\mu_{n_1}(I_2)^{-1/2}\cdot A_{J_1J_2}\mu_{n_2}(J_2)^{-1/2},\end{aligned}$$

where Lemma 3.2 has been used. Therefore, by Lemma 3.6 we deduce that

$$\begin{aligned}\mathscr{G}_{<,<} &\lesssim \sum_{I_2,J_2:\text{ good}}\iint_{W_{J_2}}\iint_{W_{I_2}}\bigg\{\sum_{\substack{\ell(I_1)<\ell(I_2)\\ \ell(J_1)<\ell(J_2)}}|f_{I_1J_1}|\mathcal{G}(x,t)\bigg\}^2 d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2}\\ &\lesssim \sum_{J_2}\sum_{I_2}\bigg[\sum_{I_1}A_{I_1I_2}\sum_{J_1}A_{J_1J_2}|f_{I_1J_1}|\bigg]^2\\ &\lesssim \sum_{J_2}\sum_{I_1}\bigg[\sum_{J_1}A_{J_1J_2}|f_{I_1J_1}|\bigg]^2\\ &\lesssim \sum_{I_1}\sum_{J_1}|f_{I_1J_1}|^2\lesssim\|f\|_{L^2(\mu)}^2.\end{aligned}$$

5. The case: $\ell(I_1)\ge\ell(I_2)$ and $\ell(J_1)<\ell(J_2)$

Noting that the mixed Carleson and Hölder estimates, and the mixed Hölder and size conditions are symmetric, we omit the control of of $\mathscr{G}<,\ge$, since it is handled almost symmetrically as the term $\mathscr{G}_{\ge,<}$.

After the splitting

$$\begin{aligned}\sum_{\ell(I_1)\ge\ell(I_2)} &= \sum_{\substack{\ell(I_1)\ge\ell(I_2)\\ d(I_1,I_2)>\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}}}\\ &\quad+\sum_{\substack{\ell(I_1)>2^r\ell(I_2)\\ d(I_1,I_2)\le\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}}}\\ &\quad+\sum_{\substack{\ell(I_2)\le\ell(I_1)\le2^r\ell(I_2)\\ d(I_1,I_2)\le\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}}},\end{aligned}$$

what remains to be done is to bound the following three terms:

$$\mathscr{G}_{\text{out},<} := \sum_{I_2,J_2:\ \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \times \Bigg| \sum_{\substack{I_1:\ell(I_1)\geq \ell(I_2) \\ d(I_1,I_2)>\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}}} \sum_{J_1:\ell(J_1)<\ell(J_2)} f_{I_1J_1}\theta_{t_1,t_2}\big(b\cdot \varphi_{I_1}^{b_1}\otimes \psi_{J_1}^{b_2}\big)(y)\Bigg|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\, d\mu_{n_2}(x_2)\frac{dt_2}{t_2},$$

$$\mathscr{G}_{\text{in},<} := \sum_{I_2,J_2:\ \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \times \Bigg| \sum_{\substack{I_1:\ell(I_1)> 2^r\ell(I_2) \\ d(I_1,I_2)\leq\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}}} \sum_{J_1:\ell(J_1)<\ell(J_2)} f_{I_1J_1}\theta_{t_1,t_2}\big(b\cdot \varphi_{I_1}^{b_1}\otimes \psi_{J_1}^{b_2}\big)(y)\Bigg|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\, d\mu_{n_2}(x_2)\frac{dt_2}{t_2},$$

and

$$\mathscr{G}_{\text{near},<} := \sum_{I_2,J_2:\ \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \times \Bigg| \sum_{\substack{I_1:\ell(I_2)\leq\ell(I_1)\leq 2^r\ell(I_2) \\ d(I_1,I_2)\leq\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}}} \sum_{J_1:\ell(J_1)<\ell(J_2)} f_{I_1J_1}\theta_{t_1,t_2}\big(b\cdot \varphi_{I_1}^{b_1}\otimes \psi_{J_1}^{b_2}\big)(y)\Bigg|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\, d\mu_{n_2}(x_2)\frac{dt_2}{t_2}.$$

The above three terms will be analyzed, respectively.

5.1. Part $\mathscr{G}_{\text{out},<}$. We begin by showing the following inequality:

$$\frac{\ell(I_2)^\alpha}{d(I_1,I_2)^\alpha \lambda_{n_1}(x_1,d(I_1,I_2))} \lesssim \frac{\ell(I_1)^{\alpha/2}\ell(I_2)^{\alpha/2}}{D(I_1,I_2)^\alpha \lambda_{n_1}(x_1,D(I_1,I_2))}. \tag{5.1}$$

If $\ell(I_1)\leq d(I_1,I_2)$, then $D(I_1,I_2)\simeq d(I_1,I_2)$. So, the inequality (5.1) holds. If $\ell(I_1)> d(I_1,I_2)$, then $D(I_1,I_2)\simeq \ell(I_1)$. The doubling condition of $\lambda_{n_1}(x_1,t)$ gives that

$$\begin{aligned}\lambda_{n_1}\big(x_1,\ell(I_1)\big) &= \lambda_{n_1}\big(x_1,\big(\ell(I_1)/\ell(I_2)\big)^{\gamma_{n_1}}\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}\big)\\ &\lesssim C_{\lambda_{n_1}}^{\log_2(\ell(I_1)/\ell(I_2))^{\gamma_{n_1}}}\lambda_{n_1}\big(x_1,\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}\big)\\ &= \big(\ell(I_1)/\ell(I_2)\big)^{\gamma_{n_1}d_{n_1}}\lambda_{n_1}\big(x_1,\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}\big).\end{aligned}$$

Note that $\gamma_{n_1}(d_{\lambda_{n_1}}+\alpha_1)=\alpha_1/2$ and $d(I_1,I_2)>\ell(I_2)^{\gamma_{n_1}}\ell(I_1)^{1-\gamma_{n_1}}$. Hence, one may conclude that

$$\begin{aligned}\frac{\ell(I_2)^{\alpha_1}}{d(I_1,I_2)^{\alpha_1}\lambda_{n_1}(x_1,d(I_1,I_2))}&\lesssim\frac{\ell(I_2)^{\alpha_1/2}}{\ell(I_1)^{\alpha_1/2}\lambda_{n_1}(x_1,\ell(I_1))}\\&\lesssim\frac{\ell(I_1)^{\alpha_1/2}\ell(I_2)^{\alpha_1/2}}{D(I_1,I_2)^{\alpha_1}\lambda_{n_1}(x_1,D(I_1,I_2))}.\end{aligned}$$

This demonstrates the inequality (5.1).

Now we turn to $\mathscr{G}_{\mathrm{out},<}$. By the cancellation property, we replace the kernel $K_{t_1,t_2}(y,z)$ by

$$K_{t_1,t_2}(y,z)-K_{t_1,t_2}\big(y,(z_1,c_{J_1})\big).$$

Then, the mixed Hölder and size condition gives that

$$\begin{aligned}\big|\theta_{t_1,t_2}\big(b\cdot\varphi_{I_1}^{b_1}\otimes\psi_{J_1}^{b_2}\big)(y)\big|&\lesssim\mu_{n_1}(I_1)^{-1/2}t_1^{\alpha_1}\mathcal{F}_{n_1,\alpha_1}(I_1,x_1,y_1,t_1)\\&\quad\times\mu_{n_2}(J_1)^{-1/2}\ell(J_1)^{\alpha_2}\mathcal{F}_{n_2,\alpha_2}(J_1,x_2,y_2,t_2).\end{aligned}$$

The above inequality, together with Lemma 3.2 and (5.1) yields that

$$\begin{aligned}\mathcal{G}(x,t)&\lesssim\mu_{n_1}(I_1)^{-1/2}t_1^{\alpha_1}\mathscr{F}_{n_1,\alpha_1}(I_1,x_1,t_1)\\&\quad\cdot\mu_{n_2}(J_1)^{-1/2}\ell(J_1)^{\alpha_2}\mathscr{F}_{n_2,\alpha_2}(J_1,x_2,t_2)\\&\lesssim\frac{\ell(I_2)^{\alpha_1}\mu_{n_1}(I_1)^{1/2}}{\ell(I_2)^{\alpha_1}\lambda_{n_1}(x_1,\ell(I_2))+d(I_1,I_2)^{\alpha_1}\lambda_{n_1}(x_1,d(I_1,I_2))}\\&\quad\cdot A_{J_1J_2}\mu_{n_2}(J_2)^{-1/2}\\&\lesssim\frac{\ell(I_2)^{\alpha_1}}{d(I_1,I_2)^{\alpha_1}\lambda_{n_1}(x_1,d(I_1,I_2))}\mu_{n_1}(I_1)^{1/2}\cdot A_{J_1J_2}\mu_{n_2}(J_2)^{-1/2}\\&\lesssim A_{I_1I_2}\mu_{n_1}(I_2)^{-1/2}\cdot A_{J_1J_2}\mu_{n_2}(J_2)^{-1/2}.\end{aligned}\tag{5.2}$$

This allows us to estimate $\mathscr{G}_{\mathrm{out},<}$ with similar steps to what we have used with $\mathscr{G}_{<,<}$. Accordingly, there holds that

$$\mathscr{G}_{\mathrm{out},<}\lesssim\|f\|_{L^2(\mu)}^2.$$

5.2. Part $\mathscr{G}_{\mathrm{near},<}$. In this case, there holds that $\ell(I_1)\simeq\ell(I_2)\simeq D(I_1,I_2)$. For convenience, we write $I_1\simeq I_2$ in this case. It immediately yields that

$$\frac{\mu_{n_1}(I_1)^{1/2}}{\lambda_{n_2}(x_1,\ell(I_2))}\simeq\frac{\mu_{n_1}(I_1)^{1/2}}{\lambda_{n_1}(c_{I_1},\ell(I_1))^{1/2}}\lambda_{n_1}\big(c_{I_2},\ell(I_2)\big)^{-1/2}\le\mu_{n_1}(I_2)^{-1/2}.\tag{5.3}$$

It follows from (5.2) that

$$\mathcal{G}(x,t)\lesssim\frac{\mu_{n_1}(I_1)^{1/2}}{\lambda_{n_1}(x_1,\ell(I_2))}\cdot A_{J_1J_2}\mu_{n_2}(J_2)^{-1/2}\lesssim\mu_{n_1}(I_2)^{-1/2}\cdot A_{J_1J_2}\mu_{n_2}(J_2)^{-1/2}.$$

It is worth pointing out that for a given I_2, there are finite cubes I_1 such that $I_1 \simeq I_2$. That also holds for a given I_1. Consequently, we deduce that

$$\begin{aligned}\mathscr{G}_{\mathrm{near},<} &\lesssim \sum_{I_2}\sum_{J_2}\sum_{I_1:I_1\simeq I_2}\bigg(\sum_{J_1} A_{J_1J_2}|f_{I_1J_1}|\bigg)^2 \\ &\lesssim \sum_{I_1}\sum_{J_2}\bigg(\sum_{J_1} A_{J_1J_2}|f_{I_1J_1}|\bigg)^2 \sum_{I_2:I_2\simeq I_1} 1 \\ &\lesssim \sum_{I_1}\sum_{J_1}|f_{I_1J_1}|^2 \simeq \|f\|_{L^2(\mu)}^2.\end{aligned}$$

5.3. Part $\mathscr{G}_{in,<}$. In this case, the goodness of I_2 indicates $I_2 \subsetneq I_1$. We use $I^{(k)} \in \mathscr{D}_{n_1}$ to denote the unique cube for which $\ell(I^{(k)}) = 2^k\ell(I)$ and $I \subset I^{(k)}$. This enables us to write

$$\begin{aligned}\mathscr{G}_{\mathrm{in},<} = &\sum_{I,J_2:\ \mathrm{good}}\iint_{W_{J_2}}\iint_{W_I}\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \\ &\times \bigg|\sum_{k=1}^{\infty}\sum_{J_1:\ell(J_1)<\ell(J_2)} f_{I^{(k)}J_1}\theta_{t_1,t_2}\big(b\cdot\varphi_{I_1}^{b_1}\otimes\psi_{J_1}^{b_2}\big)(y)\bigg|^2 \\ &\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\, d\mu_{n_2}(x_2)\frac{dt_2}{t_2}.\end{aligned}$$

Set

$$\xi_I^k := -\big\langle\varphi_{I^{(k)}}^{b_1}\big\rangle_{I^{(k-1)}}\mathbf{1}_{(I^{(k-1)})^c} + \sum_{\substack{I'\in ch(I^{(k)})\\ I'\neq I^{(k-1)}}}\varphi_{I^{(k)}}^{b_1}\mathbf{1}_{I'}. \tag{5.4}$$

It is easy to check that $\operatorname{supp}\xi_I^k \subset (I^{(k-1)})^c$, $\|\xi_I^k\|_{L^\infty(\mu_{n_1})} \lesssim \mu_{n_1}(I^{(k)})^{-1/2}$, and

$$\varphi_{I^{(k)}}^{b_1} = \xi_I^k + \big\langle\varphi_{I^{(k)}}^{b_1}\big\rangle_{I^{(k-1)}}. \tag{5.5}$$

Denote $f_{J_1} = \langle f, \psi_{J_1}^{b_2}\rangle$ so that $f_{J_1}(y_1) = \int_{\mathbb{R}^{n_2}} f(y_1,y_2)\psi_{J_1}^{b_2}(y_2)\, d\mu_{n_2}(y_2)$, $y_1 \in \mathbb{R}^{n_1}$.

We are reduced to dominating

$$\begin{aligned}\mathscr{G}_{\mathrm{mod},<} = &\sum_{I,J_2:\ \mathrm{good}}\iint_{W_{J_2}}\iint_{W_I}\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \\ &\times \bigg|\sum_{k=1}^{\infty}\sum_{J_1:\ell(J_1)<\ell(J_2)} f_{I^{(k)}J_1}\theta_{t_1,t_2}\big(b\cdot\xi_I^k\otimes\psi_{J_1}^{b_2}\big)(y)\bigg|^2 \\ &\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\, d\mu_{n_2}(x_2)\frac{dt_2}{t_2}\end{aligned}$$

and

$$\begin{aligned}
\mathscr{G}_{\mathrm{Car},<} &= \sum_{I,J_2:\ \text{good}} \iint_{W_{J_2}} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \\
&\quad \times \bigg| \sum_{J_1:\ell(J_1)<\ell(J_2)} f_{I^{(k)}J_1} \theta_{t_1,t_2}\big(b_1 \otimes \big(b_2 \psi_{J_1}^{b_2}\big)\big)(y) \sum_{k=1}^{\infty} f_{I^{(k)}J_1} \big\langle \varphi_{I^{(k)}}^{b_1} \big\rangle_{I^{(k-1)}} \bigg|^2 \\
&\quad \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2} \\
&= \sum_{I,J_2:\ \text{good}} \iint_{W_{J_2}} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \\
&\quad \times \bigg| \sum_{J_1:\ell(J_1)<\ell(J_2)} \langle f_{J_1} \rangle_I \theta_{t_1,t_2}\big(b_1 \otimes \big(b_2 \psi_{J_1}^{b_2}\big)\big)(y) \bigg|^2 \\
&\quad \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2}.
\end{aligned}$$

We will consider two cases.

- *Case 1.* We first control the following integral

$$\begin{aligned}
\mathcal{H}(x,t) := \bigg(\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \big| \theta_{t_1,t_2}\big(b \cdot \xi_I^k \otimes \psi_{J_1}^{b_2}\big)(y) \big|^2 \\
\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \bigg)^{1/2}.
\end{aligned}$$

Using the cancellation property again, we can change the kernel to $K_{t_1,t_2}(y,z) - K_{t_1,t_2}(y,(z_1,c_{J_1}))$. It follows from the mixed Hölder and size condition that

$$\begin{aligned}
\big| \theta_{t_1,t_2}\big(b \cdot \xi_I^k \otimes \psi_{J_1}^{b_2}\big)(y) \big| \lesssim \mu_{n_1}\big(I^{(k)}\big)^{-1/2} t_1^{\alpha_1} \mathcal{F}_{n_1,\alpha_1}\big(\big(I^{(k-1)}\big)^c, x_1, y_1, t_1\big) \\
\times \mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathcal{F}_{n_2,\alpha_2}(J_1, x_2, y_2, t_2).
\end{aligned}$$

Combining Lemma 3.2 with Lemma 3.3, we see that

$$\begin{aligned}
\mathcal{H}(x,t) &\lesssim \mu_{n_1}\big(I^{(k)}\big)^{-1/2} \mathscr{S}_k(x_1) \cdot \mu_{n_2}(J_1)^{-1/2} \ell(J_1)^{\alpha_2} \mathscr{F}_{n_2,\alpha_2}(J_1, x_2, t_2) \\
&\lesssim 2^{-\alpha_1 k/2} \mu_{n_1}\big(I^{(k)}\big)^{-1/2} \cdot A_{J_1 J_2} \mu_{n_2}(J_2)^{-1/2}.
\end{aligned}$$

Thereby, Minkowski's integral inequality implies that

$$\begin{aligned}
\mathscr{G}_{\mathrm{mod},<} \lesssim \sum_{I,J_2:\ \text{good}} \iint_{W_{J_2}} \iint_{W_I} \bigg\{ \sum_{k=1}^{\infty} \sum_{J_1:\ell(J_1)<\ell(J_2)} |f_{I^{(k)}J_1}| \mathcal{H}(x,t) \bigg\}^2 \\
\times d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2}
\end{aligned}$$

$$
\begin{aligned}
&\lesssim \sum_{I}\sum_{J_2}\Bigg\{\sum_{k=1}^{\infty} 2^{-\alpha_1 k/2}\bigg(\frac{\mu_{n_1}(I)}{\mu_{n_1}(I^{(k)})}\sum_{J_1:\ell(J_1)<\ell(J_2)} A_{J_1J_2}|f_{I^{(k)}J_1}|\bigg)^{1/2}\Bigg\}^2 \\
&\le \Bigg\{\sum_{k=1}^{\infty} 2^{-\alpha_1 k/4}\cdot 2^{-\alpha_1 k/4} \\
&\quad \times \bigg(\sum_{I}\frac{\mu_{n_1}(I)}{\mu_{n_1}(I^{(k)})}\sum_{J_2}\bigg(\sum_{J_1:\ell(J_1)<\ell(J_2)} A_{J_1J_2}|f_{I^{(k)}J_1}|\bigg)^2\bigg)^{1/2}\Bigg\}^2 \\
&\lesssim \sum_{k=1}^{\infty} 2^{-\alpha_1 k/2}\sum_{I}\frac{\mu_{n_1}(I)}{\mu_{n_1}(I^{(k)})}\sum_{J_2}\bigg(\sum_{J_1:\ell(J_1)<\ell(J_2)} A_{J_1J_2}|f_{I^{(k)}J_1}|\bigg)^2 \\
&\lesssim \sum_{k=1}^{\infty} 2^{-\alpha_1 k/2}\sum_{Q,J_1}|f_{QJ_1}|^2\mu_{n_1}(Q)^{-1}\sum_{I:I^{(k)}=Q}\mu_{n_1}(I) \lesssim \|f\|_{L^2(\mu)}^2.
\end{aligned}
$$

- *Case 2.* Applying Lemma 3.4, we obtain that for $\ell(J_1)<\ell(J_2)$,

$$
a_I \lesssim \mu_{n_1}(I)\big(A_{J_1,J_2}\mu_{n_2}(J_2)^{-1/2}\big)^2,
$$

where $\{a_I\}$ was defined in Lemma 3.4. Consequently, from the Carleson Embedding Theorem 3.7, it follows that

$$
\begin{aligned}
\mathscr{G}_{\mathrm{Car},<} &\le \sum_{J_2}\iint_{W_{J_2}}\sum_{I}\Bigg\{\sum_{J_1:\ell(J_1)<\ell(J_2)}\bigg(\iint_{W_I}\iint_{\mathbb{R}^{n_1+n_2}}\vartheta(x,y,t) \\
&\quad \times \big|\langle f_{J_1}\rangle_I\theta_{t_1,t_2}\big(b_1\otimes\big(b_2\psi_{J_1}^{b_2}\big)\big)(y)\big|^2 \\
&\quad \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\bigg)^{1/2}\Bigg\}^2 d\mu_{n_2}(x_2)\frac{dt_2}{t_2} \\
&\le \sum_{J_2}\iint_{W_{J_2}}\Bigg\{\sum_{J_1:\ell(J_1)<\ell(J_2)}\bigg(\sum_{I}\big|\langle f_{J_1}\rangle_I\big|^2\iint_{W_I}\iint_{\mathbb{R}^{n_1+n_2}}\vartheta(x,y,t) \\
&\quad \times \big|\theta_{t_1,t_2}\big(b_1\otimes\big(b_2\psi_{J_1}^{b_2}\big)\big)(y)\big|^2 \\
&\quad \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\bigg)^{1/2}\Bigg\}^2 d\mu_{n_2}(x_2)\frac{dt_2}{t_2} \\
&\lesssim \sum_{J_2}\bigg(\sum_{J_1:\ell(J_1)<\ell(J_2)} A_{J_1J_2}\|f_{J_1}\|_{L^2(\mu_{n_1})}\bigg)^2 \lesssim \sum_{J_1}\|f_{J_1}\|_{L^2(\mu_{n_1})}^2 \\
&\lesssim \|f\|_{L^2(\mu)}^2.
\end{aligned}
$$

This finishes the estimates of $\mathscr{G}_{\ge,<}$.

6. The case: $\ell(I_1) \geq \ell(I_2)$ and $\ell(J_1) \geq \ell(J_2)$

As we see in the preceding section, for the relative position of I_1 and I_2, there are three different cases. Similarly, there are also three different cases for the second variable. This leads to

$$\begin{aligned}\mathscr{G}_{\geq,\geq} \lesssim \mathscr{G}_{\text{out,out}} + \mathscr{G}_{\text{out,in}} + \mathscr{G}_{\text{out,near}} + \mathscr{G}_{\text{in,out}} + \mathscr{G}_{\text{in,in}}\\ + \mathscr{G}_{\text{in,near}} + \mathscr{G}_{\text{near,out}} + \mathscr{G}_{\text{near,in}} + \mathscr{G}_{\text{near,near}}.\end{aligned}$$

6.1. $\mathscr{G}_{\text{out,out}}$. We first treat the term $\Sigma_{\text{out,out}}$, where the new bi-parameter phenomena will appear. Using the similar decomposition to (5.5), we can split the function $\psi_J^{b_2}$ with η_J^i. Hence, it is enough to dominate the following terms:

$$\begin{aligned}\mathscr{G}_{\text{mod,mod}} = \sum_{I,J\colon \text{ good}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \Big| \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} f_{I^{(k)}J^{(i)}} \theta_{t_1,t_2}\big(b\cdot \xi_I^k \otimes \eta_J^i\big)(y)\Big|^2\\ \times \vartheta(x,y,t)\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2},\end{aligned}$$

$$\begin{aligned}\mathscr{G}_{\text{mod,Car}} = \sum_{I,J\colon \text{ good}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \Big| \sum_{k=1}^{\infty} \langle f_{I^{(k)}}\rangle_J \theta_{t_1,t_2}\big(\big(b_1 \xi_I^k\big) \otimes b_2\big)(y)\Big|^2\\ \times \vartheta(x,y,t)\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2},\end{aligned}$$

$$\begin{aligned}\mathscr{G}_{\text{Car,mod}} = \sum_{I,J\colon \text{ good}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \Big| \sum_{i=1}^{\infty} \langle f_{J^{(i)}}\rangle_I \theta_{t_1,t_2}\big(b_1 \otimes \big(b_2 \eta_J^i\big)\big)(y)\Big|^2\\ \times \vartheta(x,y,t)\frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2},\end{aligned}$$

and

$$\begin{aligned}\mathscr{G}_{\text{Car,Car}} = \sum_{I,J\colon \text{ good}} \big|\langle f\rangle_{I\times J}\big|^2 \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t)\big|\theta_{t_1,t_2} b(y)\big|^2\\ \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2}.\end{aligned}$$

First, to consider $\mathscr{G}_{\text{mod,mod}}$, by size condition and (3.3), we have

$$\begin{aligned}\bigg(\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t)\big|\theta_{t_1,t_2}\big(b\cdot \xi_I^k \otimes \eta_J^i\big)(y)\big|^2 \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\bigg)^{1/2}\\ \lesssim 2^{-\alpha_1 k/2}\mu_{n_1}\big(I^{(k)}\big)^{-1/2}\cdot 2^{-\alpha_2 i}\mu_{n_2}\big(J^{(i)}\big)^{-1/2}.\end{aligned}$$

Applying the techniques in the estimates of $\mathscr{G}_{\mathrm{mod},<}$ to analyze $\Sigma_{\mathrm{mod},\mathrm{mod}}$, we conclude that

$$\begin{aligned}\mathscr{G}_{\mathrm{mod},\mathrm{mod}} &\lesssim \sum_{k,i} 2^{-\alpha k/2} 2^{-\beta i} \sum_{Q,R} |f_{QR}|^2 \frac{1}{\mu_{n_1}(Q)} \sum_{I:I^{(k)}=Q} \mu_{n_1}(I) \\ &\quad \cdot \frac{1}{\mu_{n_2}(R)} \sum_{J:J^{(i)}=R} \mu_{n_2}(J) \\ &\lesssim \|f\|_{L^2(\mu)}^2.\end{aligned}$$

Secondly, to control $\mathscr{G}_{\mathrm{Car},\mathrm{Car}}$, we employ the bi-parameter Carleson condition

$$\begin{aligned}\mathscr{G}_{\mathrm{Car},\mathrm{Car}} &= \sum_{I,J} \big|\langle f\rangle_{I\times J}\big|^2 \mathcal{C}_{IJ}^b = 2\int_0^\infty \sum_{\substack{I,J \\ |\langle f\rangle_{I\times J}|>t}} \mathcal{C}_{IJ}^b t\,dt \\ &\lesssim \int_0^\infty \sum_{\substack{I,J \\ I\times J\subset\{M_s^{\mathcal{D}} f>t\}}} \mathcal{C}_{IJ}^b t\,dt \lesssim \int_0^\infty \mu\big(\{M_s^{\mathcal{D}} f>t\}\big) t\,dt \\ &\lesssim \big\|M_s^{\mathcal{D}} f\big\|_{L^2(\mu)}^2 \lesssim \|f\|_{L^2(\mu)}^2,\end{aligned}$$

where we have used the $L^p(\mu)$ $(1<p<\infty)$ boundedness of the strong maximal function associated with rectangles.

Next, we treat $\mathscr{G}_{\mathrm{Car},\mathrm{mod}}$. There holds that

$$\mathscr{G}_{\mathrm{mod},\mathrm{Car}} \leq \sum_{I:\ \mathrm{good}} \iint_{W_I} \Bigg[\sum_{k=1}^\infty \Big(\sum_J \big|\langle f_{I^{(k)}}\rangle_J\big|^2 a_J\Big)^{1/2}\Bigg]^2 d\mu_{n_1}(x_1)\frac{dt_1}{t_1},$$

where

$$\begin{aligned}a_J = \iint_{W_J} \iint_{\mathbb{R}^{n_1+n_2}} &\vartheta(x,y,t)\big|\theta_{t_1,t_2}\big((b_1\xi_I^k)\otimes b_2\big)(y)\big|^2 \\ &\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\, d\mu_{n_2}(x_2)\frac{dt_2}{t_2}.\end{aligned}$$

Combining Lemma 3.5 with Lemma 3.7, it yields that

$$\begin{aligned}\mathscr{G}_{\mathrm{mod},\mathrm{Car}} &\lesssim \sum_I \Bigg(\sum_{k=1}^\infty 2^{-\alpha_1 k/2} \mu_{n_1}\big(I^{(k)}\big)^{-1/2} \|f_{I^{(k)}}\|_{L^2(\mu_{n_2})}^2\Bigg)^2 \mu_{n_1}(I) \\ &\lesssim \sum_{k=1}^\infty 2^{-\alpha_1 k/2} \sum_Q \|f_Q\|_{L^2(\mu_{n_2})}^2 \frac{1}{\mu_{n_1}(Q)} \sum_{I:I^{(k)}=Q} \mu_{n_1}(I) \lesssim \|f\|_{L^2(\mu)}^2.\end{aligned}$$

$\mathscr{G}_{\mathrm{Car},\mathrm{mod}}$ is symmetric to $\mathscr{G}_{\mathrm{mod},\mathrm{Car}}$.

Finally, as for the estimates of the remaining terms, the decompositions and calculations needed are contained in the above sections essentially. Simply, the combinations of the techniques that we have used will lead to the desired results. □

7. Necessity of bi-parameter Carleson condition

In this section, we will show that the bi-parameter Carleson condition is necessary for $g^*_{\varepsilon,\beta}$-function bound on $L^2(\mu)$. The argument below follows along the lines of the proofs in [14] and [18].

Suppose that $\theta_{t_1,t_2} = \theta^{n_1}_{t_1} \otimes \theta^{n_2}_{t_2}$ is bounded on $L^2(\mu)$, where $\theta^{n_i}_{t_i}$ has a kernel $K^{n_i}_{t_i}(x_i, y_i)$, $x_i, y_i \in \mathbb{R}^{n_i}$ and $t_i > 0$, $i = 1, 2$. We assume that these satisfy the size condition and the corresponding L^2 bounds in $\mathbb{R}^{n_i}$, $i = 1, 2$. We shall show that the bi-parameter Carleson condition (1.1) holds.

Let $M_{\mathscr{D}}$ be the strong maximal function related to the grid $\mathscr{D}$ and M denote the strong maximal function. Let $\Omega \subset \mathbb{R}^{n_1+n_2}$ be such a set that $\mu(\Omega) < \infty$ and that for every $x \in \Omega$ there exists $I \times J \in \mathscr{D}$ so that $x \in I \times J \subset \Omega$. Define $\widetilde{\Omega} = \{M_{\mathscr{D}} \mathbf{1}_{\Omega} > 1/2\}$ and $\widehat{\Omega} = \{M \mathbf{1}_{\widetilde{\Omega}} > c\}$ for a small enough dimensional constant $c = c(n_1, n_2)$. Then we have

$$\mu(\widetilde{\Omega}) \le 4\|M_{\mathcal{D}} \mathbf{1}_{\Omega}\|^2_{L^2(\mu)} \lesssim \|\mathbf{1}_{\Omega}\|^2_{L^2(\mu)} = \mu(\Omega).$$

Similarly, there holds that $\mu(\widehat{\Omega}) \lesssim \mu(\widetilde{\Omega}) \lesssim \mu(\Omega)$. Consequently, it suffices to show

$$\begin{aligned} \sum_{\substack{I\times J \in \mathscr{D} \\ I\times J \subset \Omega}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \big|\theta_{t_1,t_2}(b\mathbf{1}_{\widehat{\Omega}^c})(y_1,y_2)\big|^2 \\ \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2} \\ \lesssim \mu(\Omega). \end{aligned}$$

For every $J \in \mathscr{D}_{n_2}$ we let $\mathcal{F}_J$ consist of the maximal $F \in \mathscr{D}_{n_1}$ for which $F \times J \subset \widetilde{\Omega}$. Then we define $F_J := \bigcup_{F \in \mathcal{F}_J} 2F$. Moreover, for fixed $I \in \mathscr{D}_{n_1}$, let $\mathcal{G}_I$ be the family of the maximal $G \in \mathscr{D}_{n_2}$ for which $I \times G \subset \Omega$, and $I_G \in \mathscr{D}_{n_1}$ be the maximal cube for which $I_G \supset I$ and $I_G \times G \subset \widetilde{\Omega}$. Thus, it is enough to show the following.

$$\begin{aligned} \mathscr{G}_1 := \sum_{\substack{I\times J \in \mathscr{D} \\ I\times J \subset \Omega}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \big|\theta_{t_1,t_2}(b\mathbf{1}_{\widehat{\Omega}^c}\mathbf{1}_{F_J})(y_1,y_2)\big|^2 \\ \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)} \, d\mu_{n_1}(x_1) \frac{dt_1}{t_1} \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2} \\ := \sum_J \iint_{W_J} \mathscr{G}_J(x_2,t_2) \, d\mu_{n_2}(x_2) \frac{dt_2}{t_2} \lesssim \mu(\Omega), \end{aligned}$$

and

$$\begin{aligned}\mathscr{G}_2 &:= \sum_{\substack{I\times J\in\mathscr{D}\\ I\times J\subset\Omega}} \iint_{W_J}\iint_{W_I}\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t)\big|\theta_{t_1,t_2}(b\mathbf{1}_{\widehat{\Omega}^c}\mathbf{1}_{F_J^c})(y_1,y_2)\big|^2 \\ &\quad\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\,d\mu_{n_2}(x_2)\frac{dt_2}{t_2} \\ &:= \sum_I \iint_{W_I} \mathscr{G}_I(x_1,t_1)\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1} \lesssim \mu(\Omega).\end{aligned}$$

To attain the goal, we need to first bound $\mathscr{G}_J(x_2,t_2)$ and $\mathscr{G}_I(x_1,t_1)$. Actually, Minkowski's integral inequality and size estimate yield that

$$\begin{aligned}\mathscr{G}_J(x_2,t_2) &\lesssim \iint_{\mathbb{R}^{n_1+1}_+}\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \\ &\quad\times\bigg|\int_{\mathbb{R}^{n_2}} K^{n_2}_{t_2}(y_2,z_2)\theta^{n_1}_{t_1}\big((b\mathbf{1}_{\widehat{\Omega}^c})(\cdot,z_2)\mathbf{1}_{F_J}\big)(y_1)\,d\mu_{n_2}(z_2)\bigg|^2 \\ &\quad\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1} \\ &\lesssim \bigg[\int_{\mathbb{R}^{n_2}}\bigg(\iint_{\mathbb{R}^{n_1+1}_+}\iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x,y,t) \\ &\quad\times \big|K^{n_2}_{t_2}(y_2,z_2)\big|^2\big|\theta^{n_1}_{t_1}\big((b\mathbf{1}_{\widehat{\Omega}^c})(\cdot,z_2)\mathbf{1}_{F_J}\big)(y_1)\big|^2 \\ &\quad\times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1,t_1)}\frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\,d\mu_{n_1}(x_1)\frac{dt_1}{t_1}\bigg)^{1/2} d\mu_{n_2}(z_2)\bigg]^2 \\ &\lesssim \bigg[\int_{\mathbb{R}^{n_2}}\bigg(\int_{\mathbb{R}^{n_2}} \frac{t_2^{\alpha_2}\vartheta_2(x_2,y_2,t_2)}{(t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2)+|y_2-z_2|^{\alpha_2}\lambda_{n_2}(x_2,|y_2-z_2|))^2} \\ &\quad\times \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2,t_2)}\bigg)^{1/2} \\ &\quad\times \big\|(b\mathbf{1}_{\widehat{\Omega}^c})(\cdot,z_2)\mathbf{1}_{F_J}\big\|_{L^2(\mu_{n_1})}\,d\mu_{n_2}(z_2)\bigg]^2 \\ &\lesssim \bigg[\int_{\mathbb{R}^{n_2}} \frac{t_2^{\alpha_2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2)+|x_2-z_2|^{\alpha_2}\lambda_{n_2}(x_2,|x_2-z_2|)} \\ &\quad\times \big\|(b\mathbf{1}_{\widehat{\Omega}^c})(\cdot,z_2)\mathbf{1}_{F_J}\big\|_{L^2(\mu_{n_1})}\,d\mu_{n_2}(z_2)\bigg]^2 \\ &\lesssim \int_{\mathbb{R}^m} \frac{t_2^{\alpha_2}}{t_2^{\alpha_2}\lambda_{n_2}(x_2,t_2)+|x_2-z_2|^{\alpha_2}\lambda_{n_2}(x_2,|x_2-z_2|)} \\ &\quad\times \big\|(b\mathbf{1}_{\widehat{\Omega}^c})(\cdot,z_2)\mathbf{1}_{F_J}\big\|^2_{L^2(\mu_{n_1})}\,d\mu_{n_2}(z_2)\end{aligned}$$

$$\lesssim \int_{\mathbb{R}^{n_2}} \frac{\ell(J)^{\alpha_2}}{|z_2 - c_J|^{\alpha_2} \lambda_{n_2}(x_2, |z_2 - c_J|)} \times \left\| (b\mathbf{1}_{\widehat{\Omega}^c})(\cdot, z_2)\mathbf{1}_{F_J} \right\|^2_{L^2(\mu_{n_1})} d\mu_{n_2}(z_2)$$

$$\lesssim \int_{\mathbb{R}^{n_1}} \mathbf{1}_{F_J}(z_1) \int_{\mathbb{R}^{n_2}} \frac{\ell(J)^{\alpha_2}}{|z_2 - c_J|^{\alpha_2} \lambda_{n_2}(x_2, |z_2 - c_J|)} \times \mathbf{1}_{\widehat{\Omega}^c}(z_1, z_2)\, d\mu_{n_2}(z_2)\, d\mu_{n_1}(z_1).$$

Similarly, we may estimate

$$\mathscr{G}_I(x_1, t_1) = \sum_{G \in \mathcal{G}_I} \sum_{J : J \subset G} \iint_{W_J} \iint_{\mathbb{R}^{n_1+n_2}} \vartheta(x, y, t) \left| \theta_{t_1,t_2}(b\mathbf{1}_{\widehat{\Omega}^c}\mathbf{1}_{F_J^c})(y_1, y_2) \right|^2 \times \frac{d\mu_{n_1}(y_1)}{\lambda_{n_1}(x_1, t_1)} \frac{d\mu_{n_2}(y_2)}{\lambda_{n_2}(x_2, t_2)}\, d\mu_{n_1}(x_1) \frac{dt_1}{t_1}\, d\mu_{n_2}(x_2) \frac{dt_2}{t_2}$$

$$\lesssim \left[\int_{\mathbb{R}^{n_1}} \frac{\ell(I)^{\alpha_1}}{\ell(I)^{\alpha_1} \lambda_{n_1}(x_1, \ell(I)) + |x_1 - z_1|^{\alpha_1} \lambda_{n_1}(x_1, |x_1 - z_1|)} \times \left(\sum_{G \in \mathcal{G}_I} \mathbf{1}_{(2I_G)^c}(z_1) \mu_{n_2}(G) \right)^{1/2} d\mu_{n_1}(z_1) \right]^2$$

$$\lesssim \sum_{G \in \mathcal{G}_I} \mu_{n_2}(G) \int_{I_G^c} \frac{\ell(I)^{\alpha_1}\, d\mu_{n_1}(z_1)}{|x_1 - z_1|^{\alpha_1} \lambda_{n_1}(x_1, |x_1 - z_1|)}$$

$$\lesssim \sum_{G \in \mathcal{G}_I} \mu_{n_2}(G) \left(\frac{\ell(I)}{\ell(I_G)} \right)^{\alpha_1}.$$

The remaining calculation is a routine application of the idea of [18]. We here omit the details. Finally, we obtain

$$\mathscr{G}_1 \lesssim \mu(\Omega) \quad \text{and} \quad \mathscr{G}_2 \lesssim \mu(\Omega).$$

Thus, we have proved the the necessity. □

Acknowledgments. The authors want to express their sincere thanks to the referee for his or her valuable remarks and suggestions, which made this paper more readable.

References

[1] T. A. Bui and X. T. Duong, *Hardy spaces, regularized BMO spaces and the boundedness of Calderón–Zygmund operators on non-homogeneous spaces*, J. Geom. Anal. **23** (2013), no. 2, 895–932. MR 3023861

[2] T. A. Bui and M. Hormozi, *Weighted bounds for multilinear quare functions*, Potential Anal. **46** (2017), 135–148. MR 3595966

[3] M. Cao, K. Li and Q. Xue, *A characterization of two weight norm inequality for Littlewood–Paley g^*_λ-function*, to appear in J. Geom. Anal. (2018).

[4] M. Cao and Q. Xue, *On the boundedness of bi-parameter Littlewood–Paley g^*_λ-function*, J. Funct. Spaces **2016** (2016), Art. ID 3605639, 15 pp.

[5] M. Cao and Q. Xue, *A non-homogeneous local Tb theorem for Littlewood–Paley g^*_λ-function with L^p-testing condition*, to appear in Forum Math. (2018).
[6] M. Cao and Q. Xue, *L^p boundedness of non-homogeneous Littlewood–Paley $g^*_{\lambda,\mu}$-function with non-doubling measures*; available at arxiv:1605.04649.
[7] X. Chen, Q. Xue and K. Yabuta, *On multilinear Littlewood–Paley operators*, Nonlinear Anal. **115** (2015), 25–40.
[8] M. Christ, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), 601–628.
[9] Y. Ding, S. Lu and Q. Xue, *Parametrized Littlewood–Paley operators on Hardy and weak Hardy spaces*, Math. Nachr. **280** (2007), 351–363.
[10] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36.
[11] Y. Han, M. Lee and C. Lin, *Tb theorem on product spaces*, Math. Proc. Cambridge Philos. Soc. **161** (2016), no. 1, 117–141. MR 3505674
[12] T. Hytönen, *A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa*, Publ. Mat. **54** (2010), no. 2, 485–504.
[13] T. Hytönen, *The vector-valued non-homogeneous Tb theorem*, Int. Math. Res. Not. IMRN **2** (2014), 451–511.
[14] T. Hytönen and H. Martikainen, *Non-homogeneous $T1$ theorem for bi-parameter singular integrals*, Adv. Math. **261** (2014), 220–273.
[15] A. K. Lerner, *On some sharp weighted norm inequalities*, J. Funct. Anal. **232** (2006), 477–494.
[16] A. K. Lerner, *On some weighted norm inequalities for Littlewood–Paley operators*, Illinois J. Math. **52** (2008), no. 2, 653–666.
[17] A. K. Lerner, *On sharp aperture-weighted estimates for square functions*, J. Fourier Anal. Appl. **20** (2014), no. 4, 784–800.
[18] H. Martikainen, *Boundedness of a class of bi-parameter square function in the upper half-space*, J. Funct. Anal. **267** (2014), 3580–3597. MR 3266241
[19] H. Martikainen and M. Mourgoglou, *Square functions with general measures*, Proc. Amer. Math. Soc. **142** (2014), no. 11, 3923–3931.
[20] B. Muckenhoupt and R. L. Wheeden, *Norm inequalities for the Littlewood–Paley function g^*_λ*, Trans. Amer. Math. Soc. **191** (1974), 95–111.
[21] F. Nazarov, S. Treil and A. Volberg, *Cauchy integral and Calderón–Zygmund operators on nonhomogeneous spaces*, Int. Math. Res. Not. IMRN **15** (1997), 703–726.
[22] F. Nazarov, S. Treil and A. Volberg, *Accretive system Tb-theorems on nonhomogeneous spaces*, Duke Math. J. **113** (2002), no. 2, 259–312.
[23] F. Nazarov, S. Treil and A. Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math. **190** (2003), no. 2, 151–239.
[24] Y. Ou, *A $T(b)$ theorem on product spaces*, Trans. Amer. Math. Soc. **367** (2015), no. 9, 6159–6197.
[25] S. Shi, Q. Xue and K. Yabuta, *On the boundedness of multilinear Littlewood–Paley g^*_λ function*, J. Math. Pures Appl. **101** (2014), 394–413.
[26] E. M. Stein, *On some function of Littlewood–Paley and Zygmund*, Bull. Amer. Math. Soc. **67** (1961), 99–101.
[27] X. Tolsa, *L^2-boundedness of the Cauchy integral operator for continuous measures*, Duke Math. J. **98** (1999), 269–304.
[28] X. Tolsa, *BMO, H^1, and Calderón–Zygmund operators for non doubling measures*, Math. Ann. **319** (2001), 89–149.
[29] X. Tolsa, *Littlewood–Paley theory and the $T(1)$ theorem with non-doubling measures*, Adv. Math. **164** (2001), 57–116.
[30] Q. Xue and Y. Ding, *Weighted estimates for the multilinear commutators of the Littlewood–Paley operators*, Sci. China Ser. A **52** (2009), 1849–1868.

[31] Q. Xue and J. Yan, *On multilinear square function and its applications to multilinear Littlewood–Paley operators with non-convolution type kernels*, J. Math. Anal. Appl. **422** (2015), 1342–1362.

Mingming Cao, School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, People's Republic of China

E-mail address: m.cao@mail.bnu.edu.cn

Qingying Xue, School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, People's Republic of China

E-mail address: qyxue@bnu.edu.cn