COMMUNICATIONS ON
PURE AND APPLIED ANALYSIS
Volume **16**, Number **3**, May **2017**

doi:10.3934/cpaa.2017042

pp. **883–898**

L^p BOUNDEDNESS FOR MAXIMAL FUNCTIONS ASSOCIATED WITH MULTI-LINEAR PSEUDO-DIFFERENTIAL OPERATORS

Jiao Chen

School of Mathematical Sciences
Beijing Normal University, Beijing, 100875, China

Wei Dai

School of Mathematics and Systems Science
Beihang University (BUAA), Beijing, 100191, China

Guozhen Lu*

Department of Mathematics, University of Connecticut
Storrs, CT 06269, USA

(Communicated by Camil Muscalu)

Abstract. In this paper, we establish the L^p estimates for the maximal functions associated with the multilinear pseudo-differential operators. Our main result is Theorem 1.2. There are several major different ingredients and extra difficulties in our proof from those in Grafakos, Honzík and Seeger [15] and Honzík [22] for maximal functions generated by multipliers. First, in order to eliminate the variable x in the symbols, we adapt a non-smooth modification of the smooth localization method developed by Muscalu in [26, 30]. Then, by applying the inhomogeneous Littlewood-Paley dyadic decomposition and a discretization procedure, we can reduce the proof of Theorem 1.2 into proving the localized estimates for localized maximal functions generated by discrete paraproducts. The non-smooth cut-off functions in the localization procedure will be essential in establishing localized estimates. Finally, by proving a key localized square function estimate (Lemma 4.3) and applying the good-λ inequality, we can derive the desired localized estimates.

1. Introduction.

A n-linear Fourier multiplier T_m given by symbol m is defined as follows:

$$T_m(f_1,\cdots,f_n)(x)=\int_{\mathbb{R}^{nd}} m(\xi)e^{ix\cdot(\xi_1+\cdots+\xi_n)}\hat{f}_1(\xi_1)\cdots\hat{f}_n(\xi_n)d\xi, \tag{1}$$

where $\xi=(\xi_1,\cdots,\xi_n)\in\mathbb{R}^{nd}$ and $f_1,\cdots,f_n$ are Schwartz functions on $\mathbb{R}^d$.

From classical Coifman-Meyer theorem (see [6, 9, 18, 23]), we know if m satisfy Hörmander-Mikhlin conditions:

$$|\partial_\xi^\alpha m(\xi)|\leq C_\alpha|\xi|^{-|\alpha|} \tag{2}$$

2000 *Mathematics Subject Classification.* Primary: 35S05; Secondary: 42B15, 42B20.

Key words and phrases. Multi-linear pseudo-differential operators, L^p estimates, maximal functions, paraproducts.

The first two authors were partly supported by the NNSF of China (No. 11371056), the second author was also supported by the NNSF of China (No. 11501021), the third author was partly supported by a US NSF grant and a Simons Fellowship from the Simons Foundation.

* Corresponding author: Guozhen Lu.

for sufficiently many multi-indices α, then the operator T_m extends to a bounded n-linear operator from $L^{p_1}(\mathbb{R}^d) \times \cdots \times L^{p_n}(\mathbb{R}^d)$ into $L^p(\mathbb{R}^d)$, provided that $1 < p_1, \cdots, p_n \leq \infty$ and $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}$. Let T_a be the corresponding bilinear pseudo-differential operators defined by replacing m with a in (1), where $a \in BS^0_{1,0}(\mathbb{R}^{3d})$, that is, a satisfies the following conditions:

$$|\partial_x^\gamma \partial_\xi^\alpha \partial_\eta^\beta a(x, \xi, \eta)| \leq C_{d,\alpha,\beta,\gamma}(1 + |\xi| + |\eta|)^{-|\alpha|-|\beta|} \tag{3}$$

for sufficiently many multi-indices α, β, γ. Then by bilinear $T1$ theorem (see [6, 18]), T_a is bounded from $L^p(\mathbb{R}^d) \times L^q(\mathbb{R}^d)$ into $L^r(\mathbb{R}^d)$, provided that $1 < p, q \leq \infty$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ (see [4], and see [2, 30] for $d = 1$ case). In the multi-parameter settings, C. Muscalu, J. Pipher, T. Tao and C. Thiele [28, 29] proved the L^p boundedness for general multi-linear and multi-parameter Coifman-Meyer multipliers by using time-frequency analysis (see also [7]). The second and third author of the current paper proved in [12] that the same L^p estimates as in [28, 29] also holds for multi-linear and multi-parameter pseudo-differential operators. For more literature involving estimates for multi-linear, multi-parameter multiplier operators and pseudo-differential operators, see e.g. [1, 6, 8, 16, 18, 21, 19, 20, 23, 24, 25, 27, 30, 31] and references therein.

M. Christ, L. Grafakos, P. Honzík and A. Seeger [5] constructed an example which shows that a family of N Mikhlin-Hörmander multipliers on $\mathbb{R}^d$ that satisfy uniform estimates forms a maximal operator $\mathfrak{M}(f) := \sup_{1\leq i\leq N} |\mathcal{F}^{-1}[m_i \hat{f}]|$ whose L^p norm is at least $\mathcal{O}(\sqrt{\log(N+1)})$. Given N Hörmander-Mikhlin multipliers $m_1, \cdots, m_N$ with uniform differential estimates, L. Grafakos, P. Honzík and A. Seeger [15] also proved an optimal $\mathcal{O}(\sqrt{\log(N+1)})$ upper bound in L^p for the maximal function $\mathfrak{M}$. For more literature on the boundedness of maximal operators, please see e.g. [11, 13, 14, 33] and references therein.

In the bilinear setting, P. Honzík [22] considered the maximal bilinear operator

$$\mathfrak{M}(f, g)(x) = \sup_{1\leq j\leq N} |T_{m_j}(f, g)(x)|, \tag{4}$$

where T_{m_j} are the bilinear Coifman-Meyer operators with symbol m_j, and m_j satisfy

$$|\partial_\xi^\alpha m_j(\xi)| \leq C_\alpha |\xi|^{-|\alpha|} \tag{5}$$

for sufficiently many multi-indices α and uniformly in $j = 1, 2, \cdots, N$. He proved

Theorem 1.1 ([22]). *Let $1 < p, q < \infty$ and $\frac{1}{2} < r < \infty$ satisfy $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, then the bilinear maximal operator $\mathfrak{M}$ defined in (4) satisfies the estimate:*

$$\|\mathfrak{M}(f, g)\|_r \leq C\sqrt{\log(N+2)}\|f\|_p\|g\|_q \tag{6}$$

for all functions $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$. Conversely, for any $N \geq 1$ there is a family of symbols m_j satisfying (5) uniformly and two Schwartz functions f and g such that

$$\|\mathfrak{M}(f, g)\|_r \geq C\sqrt{\log(N+2)}\|f\|_p\|g\|_q, \tag{7}$$

where the constant C is independent of N.

The above Theorem 1.1 can also be extended to general n-linear case ($n \geq 3$).

The purpose of this paper is to prove the pseudo-differential variant of the L^p estimates for the maximal operator $\mathfrak{M}$. For simplicity, we will only consider the case $d = 1$ and $n = 2$ in this paper. However, it will be clear from the proof that we can extend the argument to the general n-linear settings straightforwardly.

Suppose that $a_j(x,\xi,\eta) \in C^\infty(\mathbb{R}^3)$ is a symbol satisfying

$$|\partial_x^\gamma \partial_\xi^\alpha \partial_\eta^\beta a_j(x,\xi,\eta)| \leq C_{\alpha,\beta,\gamma}(1+|\xi|+|\eta|)^{-|\alpha|-|\beta|} \tag{8}$$

uniformly in j, for $j=1,2,\cdots,N$. Let $\mathcal{M}$ be the bilinear maximal operator defined by

$$\mathcal{M}(f,g)(x) = \sup_{1\leq j\leq N} |T_{a_j}(f,g)(x)|. \tag{9}$$

Our main result in this article is the following theorem.

Theorem 1.2. *Let $1 < p, q < \infty$ and $1/p + 1/q = 1/r$. If a family of bilinear symbols $\{a_j\}_{j=1}^N$ satisfies (8) uniformly in j, then the associated maximal operator $\mathcal{M}$ satisfies the estimate:*

$$\|\mathcal{M}(f,g)\|_r \leq C\sqrt{\log(N+2)}\|f\|_p\|g\|_q \tag{10}$$

for all functions $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$. Moreover, the constant $\mathcal{O}(\sqrt{\log(N+2)})$ in the bound is optimal in the sense that we can find symbols $\{a_j\}$ satisfying (8) uniformly such that

$$\|\mathcal{M}(f,g)\|_r \geq C\sqrt{\log(N+2)}\|f\|_p\|g\|_q \tag{11}$$

for some function $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$.

Before starting the proof of our main result (Theorem 1.2), we would like to give a brief overview of the ingredients in our proof strategy and indicate its additional difficulties due to our variable coefficient settings compared with the L^p boundedness proved by L. Grafakos, P. Honzík and A. Seeger [15] for maximal functions of Mikhlin-Hörmander multipliers and P. Honzík [22] for maximal functions of multi-linear Coifman-Meyer multipliers. First, since the derivatives with respect to variable x do not affect the uniform estimates (8) for symbols $\{a_j\}_{j=1}^N$, by using the idea from C. Muscalu in [26, 30] (see also [12] in multi-parameter settings), we can essentially reduce the proof of Theorem 1.2 for $r>1$ into proving a localized estimate (see (30)) for the localized maximal function $\mathcal{M}^{\vec{0}}$ generated by bilinear Coifman-Meyer multipliers. What deserves to be mentioned is that, we use non-smooth cut-off functions in the localization procedure, which is clearly different from the localization used by C. Muscalu [26, 30] and will be essential in our subsequent proof (for instance, the localized square function estimate, see Lemma 4.3). Then, by applying the inhomogeneous Littlewood-Paley dyadic decomposition, we can bound the localized maximal function $\mathcal{M}^{\vec{0}}$ pointwisely by a summation of four localized bilinear maximal functions $\mathcal{M}_i^{\vec{0}}$ ($i=1,\cdots,4$, see (31)), in which the localized maximal functions $\mathcal{M}_1^{\vec{0}}$ and $\mathcal{M}_3^{\vec{0}}$ can be reduced further into localized maximal functions $\widetilde{\mathcal{M}}_1^{\vec{0}}$ and $\widetilde{\mathcal{M}}_3^{\vec{0}}$ generated by discrete bilinear paraproducts (see (37) and (38)) by a standard discretization procedure (see [28, 30]). Therefore, the proof of Theorem 1.2 can be finally reduced into proving the localized estimates (39) and (40) consisting of a auxiliary bilinear operator G_s (see (21)) for the "high-high" frequency part $\widetilde{\mathcal{M}}_3^{\vec{0}}$, "low-low" frequency part $\mathcal{M}_4^{\vec{0}}$ and the "high-low" frequency part $\widetilde{\mathcal{M}}_1^{\vec{0}}$. Second, we apply the localized bilinear paraproduct estimates (Proposition 1, for the proof, see [26, 30] and see also [12] for bi-parameter case) and its variants (see Remark 1) to prove estimate (39) for the localized bilinear maximal functions $\widetilde{\mathcal{M}}_3^{\vec{0}}$ and $\mathcal{M}_4^{\vec{0}}$, moreover, we also use the nonnegativity assumption on the non-lacunary family of L^2-normalized bump functions $(\varphi_I^3)_{I\in\mathcal{I}}$ in the estimate of $\widetilde{\mathcal{M}}_3^{\vec{0}}$. The third key ingredient in our proof is that, through a careful observation and analysis, we can establish a localized square function estimate (Lemma 4.3), which indicates that the martingale square functions (see (57)) of the localized "high-low" frequency paraproducts Λ_j (see (53)) are monotone with respect to the corresponding starting levels and can be controlled pointwisely (and also uniformly in j) by the localized auxiliary operator G_s. Then, by using the key Lemma 4.3, the good-λ inequality (69) (see [10]) and a refined estimate on the measure of the set $\{x \in I_0 : \sup_{1\leq j\leq N} |\mathbb{E}_{-N}(\Lambda_j(f,g)\chi_{I_0})| > 2^{1-N}\lambda\}$ (see (71) and (72)), we can finally derive the localized estimate (40) for $\widetilde{\mathcal{M}}_1^{\vec{0}}$. Once the upper bounds in Theorem 1.2 have been established for $r > 1$, we can cover the $\frac{1}{2} < r \leq 1$ cases by using the endpoint weak type estimates (see Proposition 2) and multi-linear interpolations.

The rest of this paper is organized as follows. In Section 2 we give some useful notations and preliminary knowledge, in particular, we reduce the proof of the main Theorem 1.2 into proving a localized estimate for the localized maximal function $\mathcal{M}^{\vec{0}}$ generated by bilinear Coifman-Meyer multipliers. In Section 3 we carry out the proof of the localized estimates (39) for the "high-high" frequency part $\widetilde{\mathcal{M}}_3^{\vec{0}}$ and "low-low" frequency part $\mathcal{M}_4^{\vec{0}}$ in the decomposition (31). Section 4 is devoted to proving the localized estimates (40) consisting of a auxiliary operator G_s for the "high-low" frequency part $\widetilde{\mathcal{M}}_1^{\vec{0}}$ in the decomposition (31). In Section 5 we will first derive the upper bound by using the endpoint weak type estimates and multi-linear interpolations, then we give a counter-example that indicates the upper bound $\mathcal{O}(\sqrt{\log(N+2)})$ is also optimal, which completes the proof of our main theorem, Theorem 1.2.

2. **Notations and preliminary results.**

Let $\mathcal{S}(\mathbb{R})$ denotes the space of Schwartz functions, and $\mathcal{S}'(\mathbb{R})$ denotes the space of tempered distributions. The Fourier transform $\hat{f}$ and the inverse Fourier transform $\check{f}$ of $f \in \mathcal{S}(\mathbb{R})$ are defined by

$$\mathcal{F}f(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} e^{-2\pi i x\cdot\xi} f(x)dx \quad \textit{and} \quad \mathcal{F}^{-1}f(x) = \check{f}(x) = \int_{\mathbb{R}} e^{2\pi i x\cdot\xi} f(\xi)d\xi.$$

Let $\varphi \in \mathcal{S}(\mathbb{R})$ such that $\mathit{supp}\hat{\varphi} \subseteq [-\frac{4}{3}, \frac{4}{3}]$ and $\hat{\varphi} = 1$ on $[-\frac{3}{4}, \frac{3}{4}]$, and define $\hat{\psi}(\xi) = \hat{\varphi}(\frac{\xi}{2}) - \hat{\varphi}(\xi)$. Then $\mathit{supp}\hat{\psi} \subseteq [-\frac{8}{3}, -\frac{3}{4}] \cup [\frac{3}{4}, \frac{8}{3}]$. For every integer $k \geq 0$, we define $\hat{\varphi}_k, \hat{\psi}_k$ by

$$\hat{\varphi}_k(\xi) := \hat{\varphi}\Big(\frac{\xi}{2^k}\Big), \quad \hat{\psi}_k(\xi) := \hat{\psi}\Big(\frac{\xi}{2^k}\Big). \tag{12}$$

We use the convention $\hat{\psi}_{-1}(\xi) := \hat{\varphi}(\xi)$, then it is easy to see

$$\sum_{k\geq -1} \hat{\psi}_k(\xi) = 1. \tag{13}$$

Then we have the following inhomogeneous Littlewood-Paley dyadic decomposition for arbitrary function $f, g \in \mathcal{S}'(\mathbb{R})$:

$$f = \sum_{k_1\geq -1} f * \psi_{k_1}, \quad g = \sum_{k_2\geq -1} g * \psi_{k_2}. \tag{14}$$

Furthermore, we have Bony's paraproducts decomposition of the product $f \cdot g$:

$$\begin{aligned} f \cdot g &= \sum_{k_1,k_2\geq -1} (f*\psi_{k_1})(g*\psi_{k_2}) \quad (15)\\ &= \Big\{ \sum_{-1\leq k_1\leq k_2-2} + \sum_{-1\leq k_2\leq k_1-2} + \sum_{k_1,k_2\geq -1, |k_1-k_2|\leq 1} \Big\} (f*\psi_{k_1})(g*\psi_{k_2}) \\ &= \sum_{k\geq 1} (f*\tilde{\varphi}_k)(g*\psi_k) + \sum_{k\geq 1} (f*\psi_k)(g*\tilde{\varphi}_k) + \sum_{k\geq 0} (f*\psi_k)(g*\tilde{\psi}_k) \\ &\quad + \{(f*\varphi)(g*\psi) + (f*\phi)(g*\varphi) + (f*\varphi)(g*\varphi)\} \\ &=: \Pi_{lh}(f,g) + \Pi_{hl}(f,g) + \Pi_{hh}(f,g) + \Pi_{ll}(f,g), \end{aligned}$$

where $\hat{\tilde{\varphi}}_k(\xi) := \hat{\varphi}_{k-1}(\xi) = \hat{\tilde{\varphi}}(\frac{\xi}{2^k})$ for any $k \geq 1$, $\hat{\tilde{\varphi}}(\xi) := \hat{\varphi}(2\xi)$, and

$$\tilde{\psi}_k := \sum_{|k-k'|\leq 1, k'\geq 0} \psi_{k'} \quad (16)$$

for any $k \geq 0$.

Definition 2.1. For $J \subseteq \mathbb{R}$ an arbitrary interval, a smooth function Φ_J is called a bump function adapted to J, if and only if the following inequalities hold[1]:

$$|\Phi_J^{(l)}(x)| \lesssim_{l,M} \frac{1}{|J|^l} \cdot \frac{1}{\Big(1 + \frac{dist(x,J)}{|J|}\Big)^M} \quad (17)$$

for every integer $M \in \mathbb{N}$ and for sufficiently many derivatives $l \in \mathbb{N}$. If Φ_J is a bump adapted to J, we say that $|J|^{-\frac{1}{2}}\Phi_J$ is an L^2-normalized bump function adapted to J.

Definition 2.2. A family of L^2-normalized adapted bump functions $(\varphi_I)_I$ is said to be nonlacunary if and only if for every I one has

$$supp\widehat{\varphi_I} \subseteq [-4|I|^{-1}, 4|I|^{-1}].$$

A family of L^2-normalized adapted bump functions $(\varphi_I)_I$ is said to be lacunary if and only if for every I one has

$$supp\widehat{\varphi_I} \subseteq \big[-4|I|^{-1}, -\frac{1}{4}|I|^{-1}\big] \cup \big[\frac{1}{4}|I|^{-1}, 4|I|^{-1}\big].$$

Definition 2.3. Let $\mathcal{I}$ be a finite set of dyadic intervals. A bilinear expression of the type

$$\Pi_{\mathcal{I}}(f,g) = \sum_{I\in\mathcal{I}} c_I \frac{1}{|I|^{\frac{1}{2}}} \langle f, \varphi_I^1\rangle\langle g, \varphi_I^2\rangle \varphi_I^3 \quad (18)$$

is called a bilinear discretized paraproduct if and only if $(c_I)_I$ is a bounded sequence of complex numbers and at least two of the families of L^2-normalized bump functions $(\varphi_I^i)_I$ for $i = 1, 2, 3$ are lacunary in the sense of Definition 2.2.

---

[1]Throughout this paper, $A \lesssim B$ means that there exists a universal constant $C > 0$ such that $A \leqslant CB$. If necessary, we use explicitly $A \lesssim_{\star,\cdots,\star} B$ to indicate that there exists a positive constant $C_{\star,\cdots,\star}$ depending only on the quantities appearing in the subscript continuously such that $A \leq C_{\star,\cdots,\star}B$.

Now we will use the idea from C. Muscalu in [26, 30] that the proof of Theorem 1.2 can be essentially reduced to establishing a localized variant of the L^p boundedness for maximal function of bilinear Coifman-Meyer multipliers proved by P. Honzík [22]. We will proceed as follows. First, pick a sequence of characteristic functions $\{\chi_{I_n}\}_{n\in\mathbb{Z}}$ with $I_n=[n,n+1)$. Then one single bilinear pseudo-differential operator T_{a_j} can be split as follows:

$$T_{a_j}=\sum_{n\in\mathbb{Z}}T^n_{a_j}, \tag{19}$$

where

$$T^n_{a_j}(f,g)(x):=T_{a_j}(f,g)(x)\chi_{I_n}(x). \tag{20}$$

We define the auxiliary operator (see [22])

$$G_s(f,g)(x):=\left(\sum_k\Big(M_sM\big((M(|\widetilde{\psi_k}*f|))(Mg)\big)\Big)^2\right)^{1/2}, \tag{21}$$

where M denotes the Hardy-Littlewood maximal function and $M_s(f):=\big(M(|f|^s)\big)^{\frac{1}{s}}$.

By the Littlewood-Paley theory and Fefferman-Stein inequality, we can derive the following estimates for the bilinear operator G_s.

Lemma 2.4 ([22]). *Assume that $1<s<\min\{r,2\}$. Then we have*

$$\|G_s(f,g)\|_{L^r}\lesssim\|f\|_{L^p}\cdot\|g\|_{L^q}, \tag{22}$$

where $\frac{1}{r}=\frac{1}{p}+\frac{1}{q}<1$ and $p>1$, $q>1$.

Now we define

$$\mathcal{M}^n(f,g)(x):=\sup_{1\le j\le N}|T^n_{a_j}(f,g)(x)|$$

and claim that for every $n\in\mathbb{Z}$, one has the following localized estimates:

$$\|\mathcal{M}^n(f,g)\|_{L^r(\mathbb{R})}\lesssim\|f\tilde{\chi}_{I_n}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_n}\|_{L^q(\mathbb{R})}+\sqrt{\log N}\|G_s(f,g)\chi_{I_n}\|_{L^r(\mathbb{R})}, \tag{23}$$

provided that $1<p,q<\infty$ and $\frac{1}{r}=\frac{1}{p}+\frac{1}{q}$, where $\tilde{\chi}_{I_n}(x):=\Big(1+\frac{dist(x,I_n)}{|I_n|}\Big)^{-100}$ and constant C in the bounds are independent of N and n. Suppose that we have proved the claim (23), then by Hölder inequality and Lemma 2.4, we have for $r>1$,

$$\begin{aligned}\|\mathcal{M}(f,g)\|_{L^r(\mathbb{R})}\lesssim&\Big(\sum_{n\in\mathbb{Z}}\|\mathcal{M}^n(f,g)\|^r_{L^r(\mathbb{R})}\Big)^{\frac{1}{r}} \qquad (24)\\ \lesssim&\Big(\sum_{n\in\mathbb{Z}}\|f\tilde{\chi}_{I_n}\|^r_{L^p}\|g\tilde{\chi}_{I_n}\|^r_{L^q}\Big)^{\frac{1}{r}}+\sqrt{\log N}\Big(\sum_{n\in\mathbb{Z}}\|G_s(f,g)\chi_{I_n}\|^r_{L^r}\Big)^{\frac{1}{r}}\\ \lesssim&\Big(\sum_{n\in\mathbb{Z}}\|f\tilde{\chi}_{I_n}\|^p_{L^p}\Big)^{\frac{1}{p}}\Big(\sum_{n\in\mathbb{Z}}\|g\tilde{\chi}_{I_n}\|^q_{L^q}\Big)^{\frac{1}{q}}+\sqrt{\log N}\|G_s(f,g)\|_{L^r(\mathbb{R})}\\ \lesssim&\sqrt{\log(N+2)}\|f\|_{L^p(\mathbb{R})}\|g\|_{L^q(\mathbb{R})},\end{aligned}$$

where we have used the convergence of series $\sum_{k\ge1}k^{-s}$ for $s>1$ to obtain the last inequality. The estimate (24) yields the upper bound in our main Theorem 1.2 for $r>1$. Therefore, from now on, we only need to prove the claim (23).

To this end, fix some $n_0\in\mathbb{Z}$, we have

$$T^{n_0}_{a_j}(f,g)(x)=\int_{\mathbb{R}^2}a_j(x,\xi,\eta)\tilde{\varphi}_{n_0}(x)\chi_{I_{n_0}}(x)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi ix(\xi+\eta)}d\xi d\eta \tag{25}$$

for every $j = 1, \cdots, N$, where $\tilde{\varphi}_{n_0}$ is a smooth function supported on the interval $[n_0-1, n_0+2]$ and equals 1 on I_{n_0}. Then we can rewrite the symbols $a_j(x,\xi,\eta)\tilde{\varphi}_{n_0}(x)$ by using Fourier expansions with respect to the x variable:

$$a_j(x,\xi,\eta)\tilde{\varphi}_{n_0}(x) = \sum_{l\in\mathbb{Z}} m_{j,l}(\xi,\eta)e^{2\pi i x l}. \tag{26}$$

By integration by parts, the condition (8) guarantees that

$$|\partial^\alpha m_{j,l}(\xi,\eta)| \lesssim \frac{1}{(1+|l|)^M} \cdot \frac{1}{(1+|\xi|+|\eta|)^{|\alpha|}} \tag{27}$$

for a sufficiently large number M and sufficiently many multi-indices α. Observe that the rapid decay in l in the estimates (27) for Fourier coefficients is acceptable for summation, thus we only need to treat the maximal operator corresponding to $l = 0$, which is given by

$$\begin{aligned} \mathcal{M}_0^{n_0}(f,g)(x) &= \sup_{1\le j\le N} |T^{n_0}_{m_{j,0}}(f,g)(x)| \\ &= \sup_{1\le j\le N} \Big| \int_{\mathbb{R}^2} m_{j,0}(\xi,\eta)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi i x(\xi+\eta)}d\xi d\eta \cdot \chi_{I_{n_0}}(x)\Big|, \end{aligned} \tag{28}$$

where the multipliers $m_{j,0}$ satisfy

$$|\partial^\alpha m_{j,0}(\xi,\eta)| \lesssim \frac{1}{(1+|\xi|+|\eta|)^{|\alpha|}} \tag{29}$$

uniformly in $j = 1, \cdots, N$.

The operator $\mathcal{M}_0^{n_0}$ is simply a localization of the maximal Coifman-Meyer bilinear operator investigated by P. Honzík in [22]. By translation invariance, we can also assume that $n_0 = 0$, that is, in order to prove our claim (23), we only need to prove the following localized estimates for the localized maximal operator $\mathcal{M}^{\vec{0}} := \mathcal{M}_0^0$:

$$\|\mathcal{M}^{\vec{0}}(f,g)\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_0}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_0}\|_{L^q(\mathbb{R})} + \sqrt{\log N}\|G_s(f,g)\chi_{I_0}\|_{L^r(\mathbb{R})}, \tag{30}$$

provided that $1 < p, q < \infty$ and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$.

Next, we will decompose the localized bilinear maximal function $\mathcal{M}^{\vec{0}}(f,g)$ into a finite summation of maximal functions of localized bilinear discrete paraproducts (see Definition 2.3) by applying the inhomogeneous Littlewood-Paley decomposition (see (14), (15)) and a standard discretization procedure (see [28, 30, 12]). We will proceed this procedure as follows. First, by using the inhomogeneous Littlewood-Paley decomposition, we can split one single symbol $m_{j,0}(\xi)$ into four terms:

$$\begin{aligned} m_{j,0}(\xi,\eta) &= m_{j,0}(\xi,\eta)\sum_{l\ge 1}\hat{\tilde{\varphi}}_l(\xi)\hat{\psi}_l(\eta) + m_{j,0}(\xi,\eta)\sum_{l\ge 1}\hat{\psi}_l(\xi)\hat{\tilde{\varphi}}_l(\eta) \\ &+ m_{j,0}(\xi,\eta)\sum_{l\ge 0}\hat{\psi}_l(\xi)\hat{\tilde{\psi}}_l(\eta) + m_{j,0}(\xi,\eta)\{\hat{\varphi}(\xi)\hat{\psi}(\eta) + \hat{\psi}(\xi)\hat{\varphi}(\eta) + \hat{\varphi}(\xi)\hat{\varphi}(\eta)\}. \end{aligned}$$

Therefore, by splitting the symbol $m_{j,0}$ as above, we can decompose the localized bilinear operator $T_{m_j,\vec{0}} := T^0_{m_{j,0}}$ into a sum of four localized bilinear operators as follows:

$$\begin{aligned}
T_{m_j,\vec{0}}(f,g) &= \sum_{l\geq 1}\int_{\mathbb{R}^2}\Big(m_{j,0}(\xi,\eta)\hat{\psi}_l(\xi)\hat{\tilde{\varphi}}_l(\eta)\Big)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi i x(\xi+\eta)}d\xi d\eta\cdot\chi_{I_0}(x)\\
&+\sum_{l\geq 1}\int_{\mathbb{R}^2}\Big(m_{j,0}(\xi,\eta)\hat{\tilde{\varphi}}_l(\xi)\hat{\psi}_l(\eta)\Big)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi i x(\xi+\eta)}d\xi d\eta\cdot\chi_{I_0}(x)\\
&+\sum_{l\geq 0}\int_{\mathbb{R}^2}\Big(m_{j,0}(\xi,\eta)\hat{\psi}_l(\xi)\hat{\tilde{\psi}}_l(\eta)\Big)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi i x(\xi+\eta)}d\xi d\eta\cdot\chi_{I_0}(x)\\
&+\int_{\mathbb{R}^2}m_{j,0}(\xi,\eta)\Big[\hat{\varphi}(\xi)\hat{\psi}(\eta)+\hat{\psi}(\xi)\hat{\varphi}(\eta)+\hat{\varphi}(\xi)\hat{\varphi}(\eta)\Big]\hat{f}(\xi)\hat{g}(\eta)e^{2\pi i x(\xi+\eta)}d\xi d\eta\chi_{I_0}(x)\\
&=: T^1_{m_j,\vec{0}}(f,g)+T^2_{m_j,\vec{0}}(f,g)+T^3_{m_j,\vec{0}}(f,g)+T^4_{m_j,\vec{0}}(f,g).
\end{aligned}$$

As a consequence,

$$\begin{aligned}
\mathcal{M}^{\vec{0}}(f,g)(x) &\leq \sup_{1\leq j\leq N}|T^1_{m_j,\vec{0}}(f,g)(x)|+\sup_{1\leq j\leq N}|T^2_{m_j,\vec{0}}(f,g)(x)|\\
&\quad+\sup_{1\leq j\leq N}|T^3_{m_j,\vec{0}}(f,g)(x)|+\sup_{1\leq j\leq N}|T^4_{m_j,\vec{0}}(f,g)(x)| \qquad (31)\\
&=: \mathcal{M}^{\vec{0}}_1(f,g)(x)+\mathcal{M}^{\vec{0}}_2(f,g)(x)+\mathcal{M}^{\vec{0}}_3(f,g)(x)+\mathcal{M}^{\vec{0}}_4(f,g)(x).
\end{aligned}$$

Since the role of variables ξ and η are symmetric in the definition of $\mathcal{M}^{\vec{0}}_1(f,g)$ and $\mathcal{M}^{\vec{0}}_2(f,g)$, by exchanging ξ and η, we can treat $\mathcal{M}^{\vec{0}}_1(f,g)$ and $\mathcal{M}^{\vec{0}}_2(f,g)$ similarly. Therefore, we only need to deal with the localized maximal functions $\mathcal{M}^{\vec{0}}_1(f,g)$, $\mathcal{M}^{\vec{0}}_3(f,g)$ and $\mathcal{M}^{\vec{0}}_4(f,g)$ and prove localized estimates for them respectively.

Since $supp\hat{\psi}_l(\xi)\hat{\tilde{\varphi}}_l(\eta)$ lies inside a cube of side length about 2^l whose distance to the origin is also of size 2^l, the smooth restriction of the symbol $m_{j,0}(\xi,\eta)$ to that cube (maybe supported on a slightly larger cube, and equals to $m_{j,0}$ on $supp\hat{\psi}_l(\xi)\hat{\tilde{\varphi}}_l(\eta)$), which is denoted by $m_{j,0,l}(\xi,\eta)$, can be decomposed as a double Fourier series:

$$m_{j,0,l}(\xi,\eta)=\sum_{n_1,n_2\in\mathbb{Z}}C^{j,l}_{n_1,n_2}e^{2\pi i n_1\xi/2^l}e^{2\pi i n_2\eta/2^l}, \qquad (32)$$

where the Fourier coefficient $C^{j,l}_{n_1,n_2}$ are given by

$$C^{j,l}_{n_1,n_2}=2^{-2l}\int_{\mathbb{R}^2}m_{j,0,l}(\xi,\eta)e^{-2\pi i n_1\xi/2^l}e^{-2\pi i n_2\eta/2^l}d\xi d\eta. \qquad (33)$$

By taking advantage of (29) and integrating by parts, one can see that

$$|C^{j,l}_{n_1,n_2}|\lesssim\frac{1}{(1+|n_1|+|n_2|)^M}, \qquad (34)$$

where the constant is independent of j and M is sufficiently large.

If we apply the double Fourier expansions to the smoothly restricted symbols $m_{j,0,l}(\xi,\eta)$ for every $l\geq 1$, and insert the corresponding double Fourier series (32) into the definition of $T^1_{m_j,\vec{0}}$, we can obtain

$$T^1_{m_j,\vec{0}}(f,g)(x)=\sum_{n_1,n_2\in\mathbb{Z}}\sum_{l\geq 1}C^{j,l}_{n_1,n_2}\big((f*\psi_{l,n_1})(g*\tilde{\varphi}_{l,n_2})\big)*\psi'_l\cdot\chi_{I_0}(x), \qquad (35)$$

where $\hat{\psi}_{l,n_1}(\xi) := \hat{\psi}_l(\xi)e^{2\pi i n_1 \xi/2^l}$, $\hat{\tilde{\varphi}}_{l,n_2}(\eta) := \hat{\tilde{\varphi}}_l(\eta)e^{2\pi i n_2 \eta/2^l}$ and $\widehat{\psi'_l}(\gamma) := \widehat{\psi'}(\frac{\gamma}{2^l})$, ψ' is a Schwartz function such that *supp* $\widehat{\psi'} \subseteq [-4, -\frac{1}{16}] \cup [\frac{1}{16}, 4]$ and $\widehat{\psi'} = 1$ on $[-\frac{10}{3}, -\frac{1}{12}] \cup [\frac{1}{12}, \frac{10}{3}]$. We can deduce in a similar way that

$$T^3_{m_j,\vec{0}}(f,g)(x) = \sum_{n_1,n_2\in\mathbb{Z}} \sum_{l\geq 0} C^{j,l}_{n_1,n_2}\big((f*\psi_{l,n_1})(g*\tilde{\psi}_{l,n_2})\big) * \varphi'_l \cdot \chi_{I_0}(x), \tag{36}$$

where $\hat{\tilde{\psi}}_{l,n_2}(\eta) := \hat{\tilde{\psi}}_l(\eta)e^{2\pi i n_2 \eta/2^l}$ and $\widehat{\varphi'_l}(\gamma) := \widehat{\varphi'}(\frac{\gamma}{2^l})$, $\widehat{\varphi'}(\gamma) := \hat{\varphi}(\frac{\gamma}{2^4})$.

Since by (34), the Fourier coefficients satisfy a rapid decay $|C^{j,l}_{n_1,n_2}| \lesssim \frac{1}{(1+|n_1|+|n_2|)^M}$ uniformly in j for sufficiently large integer M, which is acceptable for summation. Therefore, we can fix $n_1 = n_2 = 0$ from now on. Then, by inserting (35) and (36) into (31), one can easily verify that localized maximal functions $\mathcal{M}^{\vec{0}}_1(f,g)$, $\mathcal{M}^{\vec{0}}_3(f,g)$ can be reduced to the following localized bilinear discrete maximal function $\widetilde{\mathcal{M}}^{\vec{0}}_1(f,g)$ and $\widetilde{\mathcal{M}}^{\vec{0}}_3(f,g)$:

$$\widetilde{\mathcal{M}}^{\vec{0}}_1(f,g)(x) := \sup_{1\leq j\leq N}\Big|\sum_{l\geq 1} C^{j,l}_{0,0}\big((f*\psi_l)(g*\tilde{\varphi}_l)\big) * \psi'_l \cdot \chi_{I_0}(x)\Big|, \tag{37}$$

$$\widetilde{\mathcal{M}}^{\vec{0}}_3(f,g)(x) := \sup_{1\leq j\leq N}\Big|\sum_{l\geq 0} C^{j,l}_{0,0}\big((f*\psi_l)(g*\tilde{\psi}_l)\big) * \varphi'_l \cdot \chi_{I_0}(x)\Big|. \tag{38}$$

Finally, we can reduce the proof of the upper bound for $r>1$ in our main Theorem 1.2 (or equivalently, the proof of (30)) into proving the following localized estimates for the localized bilinear maximal operators $\widetilde{\mathcal{M}}^{\vec{0}}_1$, $\widetilde{\mathcal{M}}^{\vec{0}}_3$ and $\mathcal{M}^{\vec{0}}_4$ respectively:

$$\|\widetilde{\mathcal{M}}^{\vec{0}}_3(f,g)\|_{L^r(\mathbb{R})} + \|\mathcal{M}^{\vec{0}}_4(f,g)\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_0}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_0}\|_{L^q(\mathbb{R})}, \tag{39}$$

$$\|\widetilde{\mathcal{M}}^{\vec{0}}_1(f,g)\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_0}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_0}\|_{L^q(\mathbb{R})} + \sqrt{\log N}\|G_s(f,g)\chi_{I_0}\|_{L^r(\mathbb{R})}, \tag{40}$$

provided that $1<p,q<\infty$ and $\frac{1}{r} = \frac{1}{p}+\frac{1}{q}$.

3. **Estimates for $\widetilde{\mathcal{M}}^{\vec{0}}_3$ and $\mathcal{M}^{\vec{0}}_4$.** In this section, we will carry out the proof of estimate (39) for the localized bilinear maximal operators $\widetilde{\mathcal{M}}^{\vec{0}}_3$ and $\mathcal{M}^{\vec{0}}_4$.

For the estimate of $\widetilde{\mathcal{M}}^{\vec{0}}_3(f,g)$, we will use the following Proposition from [26, 30], one can also see [12].

Proposition 1. *For a bilinear discrete paraproduct*

$$\Pi_{\mathcal{I}}(f,g) := \sum_{I\in\mathcal{I},|I|\leq 1} \frac{1}{|I|^{\frac{1}{2}}}\langle f,\varphi^1_I\rangle\langle g,\varphi^2_I\rangle\varphi^3_I,$$

pick a sequence of smooth functions $(\varphi_n)_{n\in\mathbb{Z}}$ such that supp$\varphi_n \subseteq [n-1,n+1]$ and

$$\sum_{n\in\mathbb{Z}} \varphi_n = 1, \tag{41}$$

then we have the following localized estimates:

$$\|\Pi_{\mathcal{I}}(f,g)\varphi_n\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_n}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_n}\|_{L^q(\mathbb{R})}, \tag{42}$$

where the constant in bound is independent of the cardinality of $\mathcal{I}$ and n.

Remark 1. From the proof of Proposition 1 (see [26, 30, 12]), one can easily verify that the following localized estimates also hold for every $n\in\mathbb{Z}$:

$$\Big\|\Big\{\sum_{I\in\mathcal{I},|I|\leq 1} \frac{1}{|I|^{\frac{1}{2}}}|\langle f,\varphi^1_I\rangle||\langle g,\varphi^2_I\rangle|\varphi^3_I\Big\}\varphi_n\Big\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_n}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_n}\|_{L^q(\mathbb{R})}. \tag{43}$$

Remark 2. In fact, if we replace smooth functions φ_n by characteristic functions χ_{I_n} in (42) and (43) respectively, then the conclusions in Proposition 1 and Remark 1 still hold true.

By a standard discretization procedure (see [28, 30, 12]), we can discretize the classical paraproduct $\sum_{l\geq 0} C_{0,0}^{j,l}\big((f*\psi_l)(g*\tilde{\psi}_l)\big)*\varphi_l'$ further into a discrete paraproduct:

$$\Pi_{\mathcal{I}}^{j}(f,g)=\sum_{I\in\mathcal{I},|I|\leq 1} C_I^j \frac{1}{|I|^{\frac{1}{2}}}\langle f,\varphi_I^1\rangle\langle g,\varphi_I^2\rangle\varphi_I^3, \tag{44}$$

where $\{C_I^j\}_{I\in\mathcal{I}}$ is uniformly bounded in j, $(\varphi_I^i)_I$ is lacunary for $i=1,2$ and non-lacunary for $i=3$. Therefore, in order to obtain localized estimate for $\widetilde{\mathcal{M}}_3^{\vec{0}}(f,g)$, it's enough for us to investigate the following maximal function of localized bilinear discrete paraproducts:

$$\overline{\mathcal{M}}_3^{\vec{0}}(f,g):=\sup_{1\leq j\leq N}\Big|\sum_{I\in\mathcal{I},|I|\leq 1} C_I^j \frac{1}{|I|^{\frac{1}{2}}}\langle f,\varphi_I^1\rangle\langle g,\varphi_I^2\rangle\varphi_I^3\cdot\chi_{I_0}\Big|. \tag{45}$$

Since $(\varphi_I^3)_I$ are non-lacunary, we may assume that bump functions φ_I^3 are non-negative for every $I\in\mathcal{I}$, and hence we can deduce from Proposition 1, (43) in Remark 1 and Remark 2 that

$$\begin{aligned}\|\overline{\mathcal{M}}_3^{\vec{0}}(f,g)\|_{L^r(\mathbb{R})}\leq \sup_{1\leq j\leq N}\sup_{I\in\mathcal{I}} C_I^j\cdot\Big\|\sum_{I\in\mathcal{I},|I|\leq 1}\frac{1}{|I|^{\frac{1}{2}}}|\langle f,\varphi_I^1\rangle||\langle g,\varphi_I^2\rangle|\varphi_I^3\cdot\chi_{I_0}\Big\|_{L^r(\mathbb{R})}\\ \lesssim\|f\tilde{\chi}_{I_0}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_0}\|_{L^q(\mathbb{R})},\end{aligned} \tag{46}$$

where the constant in bound is independent of the cardinality of $\mathcal{I}$ and N. We can infer from (46) immediately the following localized estimate for $\widetilde{\mathcal{M}}_3^{\vec{0}}$:

$$\|\widetilde{\mathcal{M}}_3^{\vec{0}}(f,g)\|_{L^r(\mathbb{R})}\lesssim\|f\tilde{\chi}_{I_0}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_0}\|_{L^q(\mathbb{R})}. \tag{47}$$

Now we will prove localized estimate for the localized bilinear maximal function $\mathcal{M}_4^{\vec{0}}(f,g)$. For simplicity, we may consider w.l.g. only one term in the definition of the operator $T^4_{m_j,\vec{0}}$ $(j=1,\cdots,N)$, for instance, the term with symbol $m_{j,0}(\xi,\eta)\hat{\varphi}(\xi)\hat{\psi}(\eta)$, which will still be denoted by $T^4_{m_j,\vec{0}}$, that is,

$$T^4_{m_j,\vec{0}}(f,g)(x)=\int_{\mathbb{R}^2} m_{j,0}(\xi,\eta)\hat{\varphi}(\xi)\hat{\psi}(\eta)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi ix(\xi+\eta)}d\xi d\eta\cdot\chi_{I_0}(x). \tag{48}$$

We can deal with the other two terms in definition of $T^4_{m_j,\vec{0}}$ $(j=1,\cdots,N)$ in a completely similar way.

We first decompose the symbol $m_{j,0}(\xi,\eta)\hat{\varphi}(\xi)\hat{\psi}(\eta)$ into double Fourier series:

$$m_{j,0}(\xi,\eta)\hat{\varphi}(\xi)\hat{\psi}(\eta)=\sum_{n_1,n_2\in\mathbb{Z}} C^j_{n_1,n_2}e^{2\pi in_1\xi}e^{2\pi in_2\eta}\hat{\varphi}(\xi)\hat{\psi}(\eta). \tag{49}$$

By inserting the above (49) into (48), we get

$$T^4_{m_j,\vec{0}}(f,g)(x)=\sum_{n_1,n_2\in\mathbb{Z}} C^j_{n_1,n_2}(f*\varphi_{n_1})(g*\psi_{n_2})\chi_{I_0}(x), \tag{50}$$

where $\hat{\varphi}_{n_1}(\xi):=\hat{\varphi}(\xi)e^{2\pi in_1\xi}$ and $\hat{\psi}_{n_2}(\eta):=\hat{\psi}(\eta)e^{2\pi in_2\eta}$.

We will only consider the case $n_1 = n_2 = 0$, since the Fourier coefficients $C^j_{n_1,n_2}$ have rapid decay in n_1, n_2 uniformly in j. The corresponding maximal operator satisfies the following localized estimates:

$$\begin{aligned}\Big\| \sup_{1\le j\le N} |C^j_{0,0}(f*\varphi)(g*\psi)\chi_{I_0}|\Big\|_{L^r} &\le \sup_{1\le j\le N} |C^j_{0,0}|\cdot \big\|(f*\varphi)(g*\psi)\varphi_0\tilde{\varphi}_0\big\|_{L^r}\\ &\lesssim \|(f*\varphi)\varphi_0\|_{L^p}\|(g*\psi)\tilde{\varphi}_0\|_{L^q} \lesssim \|f\tilde{\chi}_{I_0}\|_{L^p}\|g\tilde{\chi}_{I_0}\|_{L^q}, \qquad (51)\end{aligned}$$

where the last inequality is obtained by Minkowski's inequality, $\varphi_0, \tilde{\varphi}_0$ are bump functions adapted to $[-1,1]$, which are supported in a interval slightly larger than $[-1,1]$ and equal to 1 on $[-1,1]$. Estimate (51) yields that

$$\|\mathcal{M}_4^{\vec{0}}(f,g)\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_0}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_0}\|_{L^q(\mathbb{R})}. \qquad (52)$$

Combining (52) with (47), we get (39) immediately.

4. **Estimate for $\widetilde{\mathcal{M}}_1^{\vec{0}}$.** In this section, we will prove the localized estimate (40) for the localized bilinear maximal operator $\widetilde{\mathcal{M}}_1^{\vec{0}}$. To this end, we define bilinear operators

$$\Lambda_j(f,g)(x) := \sum_{l\ge 1} C^{j,l}_{0,0}\big((f*\psi_l)(g*\tilde{\varphi}_l)\big)*\psi_l'(x). \qquad (53)$$

We have the following localized estimate for $\Lambda_j(f,g)$ from [26, 30].

Lemma 4.1. *The bilinear operators $\Lambda_j(f,g)$ satisfy*

$$\|\Lambda_j(f,g)\chi_{I_n}\|_{L^r(\mathbb{R})} \lesssim \|f\tilde{\chi}_{I_n}\|_{L^p(\mathbb{R})}\|g\tilde{\chi}_{I_n}\|_{L^q(\mathbb{R})}, \qquad (54)$$

where the constant in bound is independent of n and $j = 1,\cdots,N$, $1 < p,q < \infty$ and $\frac{1}{r} = \frac{1}{p}+\frac{1}{q}$.

For the proof of Lemma 4.1 in multi-parameter settings, please refer to [12].

Before carrying out the proof for estimate (40), let us first introduce some useful definitions and lemmas.

We define the conditional expectation operator

$$\mathbb{E}_k f(x) := 2^k \sum_{I\in\mathcal{D}_k} \chi_I(x)\int_I f(y)dy, \qquad (55)$$

the martingale difference operator

$$\mathbb{D}_k f(x) := \mathbb{E}_{k+1}f(x) - \mathbb{E}_k f(x), \qquad (56)$$

and also define the family of martingale square functions with starting levels $k_0 \in \mathbb{Z}$:

$$S_{k_0}(f)(x) := \Big(\sum_{k\ge k_0} |\mathbb{D}_k f(x)|^2\Big)^{1/2}, \qquad (57)$$

where $\mathcal{D}_k$ denotes the family of dyadic intervals with length 2^{-k}. The martingale square functions $S_{k_0}(f)$ are L^p bounded (see [3]). The maximal martingale operator $\sup_k |\mathbb{E}_k f|$ is pointwise bounded by Mf and thus it is bounded in L^p norm.

Following [15], we will introduce two convolution operators which have some connection with the dyadic martingale difference (see [15, 22]). Suppose $\widehat{\psi*\tilde{\varphi}}$ is supported in $\{\xi\in\mathbb{R} : {C_0}^{-1} \le |\xi| \le C_0\}$. Let b be a radial Schwartz function supported in $[-\frac{1}{4},\frac{1}{4}]$ with $\hat{b}(\xi)\neq 0$ for ${C_0}^{-1}\le|\xi|\le C_0$ and $\int_{\mathbb{R}} b(x)dx = 0$ and set $\hat{b}_i(\xi) = \hat{b}(2^{-i}\xi)$. Pick a function $\hat{\gamma}\in C_c^\infty(\mathbb{R})$ such that $(\hat{b}(\xi))^2\hat{\gamma}(\xi) = 1$ for

$\xi \in \{\xi \in \mathbb{R} : {C_0}^{-1} \leq |\xi| \leq C_0\}$ and set $\hat{\gamma}_i(\xi) = \hat{\gamma}(2^{-i}\xi)$. We shall define the notations:

$$B_i f := f * b_i \quad \text{and} \quad \Gamma_i f := f * \gamma_i \quad \text{for } i \in \mathbb{Z}. \tag{58}$$

The operators B_i and Γ_i satisfy the following pointwise estimates.

Lemma 4.2 ([15, 22]). *For $i, k \in \mathbb{Z}$, $j \in \mathbb{N} \cup \{0\}$ and $s > 1$, we have*

$$|(B_i \Gamma_i f)(x)| \lesssim Mf(x), \tag{59}$$

$$|(\mathbb{D}_k B_{k+i} f)(x)| \lesssim 2^{-|i|/s'} M_s(f)(x) \tag{60}$$

and

$$|\mathbb{E}_{k+1} B_{k+j} f(x)| + |\mathbb{E}_k B_{k+j} f(x)| \lesssim 2^{-j/s'} M_s(f)(x). \tag{61}$$

By using the estimates for B_i and Γ_i in Lemma 4.2, P. Honzík proved a pointwise estimate for the square function in Lemma 6.2 [22], that is, the square function S_{k_0} of the "high-low" frequency part Λ_j can be controlled by the auxiliary operator G_s globally in $x \in \mathbb{R}$. Through a careful observation and analysis, we can establish the following key estimate, i.e., a localized square function estimate.

Lemma 4.3 (Localized square function estimate). *For any $1 < s < \infty$ and for $1 \leq j \leq N$, we have a localized pointwise estimate:*

$$S_0\big(\Lambda_j(f,g)\chi_{I_0}\big)(x) \leq S_{-N}\big(\Lambda_j(f,g)\chi_{J_N}\big)(x) \lesssim_s G_s(f,g)(x)\chi_{J_N}(x), \tag{62}$$

where the constant in bound is independent of $j = 1, \cdots, N$ and the interval $J_N := [0, 2^N]$.

Proof. Since $k \geq -N$ in the definition of the martingale square function (57), we have $I \cap J_N = \emptyset$ unless $I \subseteq J_N$ for any $I \in \mathcal{D}_k$. By the definitions of the martingale difference operator $\mathbb{D}_k$ and the conditional expectation operator $\mathbb{E}_k$ in (55) and (56), we have for $k \geq -N$,

$$\mathbb{D}_k(\Lambda_j(f,g)\chi_{J_N}) = 2^k \left[\left(2\sum_{I\in\mathcal{D}_{k+1}} - \sum_{I\in\mathcal{D}_k}\right)\chi_I(x)\int_I \Lambda_j(f,g)\chi_{J_N}\,dy\right] \tag{63}$$
$$= 2^{k+1}\sum_{I\in\mathcal{D}_{k+1},\, I\subseteq J_N}\chi_I(x)\int_I \Lambda_j(f,g)dy - 2^k \sum_{I\in\mathcal{D}_k,\, I\subseteq J_N}\chi_I(x)\int_I \Lambda_j(f,g)dy$$
$$= \mathbb{E}_{k+1}(\Lambda_j(f,g))\chi_{J_N} - \mathbb{E}_k(\Lambda_j(f,g))\chi_{J_N} = \mathbb{D}_k(\Lambda_j(f,g))\chi_{J_N}.$$

Thus we can infer from the definition of the martingale square function (57) and (63) that

$$S_{-N}(\Lambda_j(f,g)\chi_{J_N})(x) = S_{-N}(\Lambda_j(f,g))(x)\cdot\chi_{J_N}(x), \tag{64}$$

combining this with the pointwise estimate $S_{-N}(\Lambda_j(f,g))(x) \lesssim_s G_s(f,g)(x)$ proved in Lemma 6.2 [22], we get the second inequality in the localized pointwise estimate (62):

$$S_{-N}(\Lambda_j(f,g)\chi_{J_N})(x) \lesssim_s G_s(f,g)(x)\chi_{J_N}(x). \tag{65}$$

In particular, for any $x \in I_0 \subseteq J_N$, we can deduce from (55) that for $k \geq 0$,

$$\mathbb{E}_k(\Lambda_j(f,g)\chi_{J_N})(x) = 2^k \sum_{I\in\mathcal{D}_k,\, I\subseteq J_N,\, I\cap I_0\neq\emptyset}\chi_I(x)\int_I \Lambda_j(f,g)\chi_{J_N}\,dy \tag{66}$$
$$= 2^k \sum_{I\in\mathcal{D}_k,\, I\subseteq I_0}\chi_I(x)\int_I \Lambda_j(f,g)\chi_{I_0}\,dy = \mathbb{E}_k(\Lambda_j(f,g)\chi_{I_0})(x),$$

and hence, (66) combining with definitions (56) and (57) yield that

$$S_0(\Lambda_j(f,g)\chi_{I_0})(x) = S_0(\Lambda_j(f,g)\chi_{J_N})(x) \tag{67}$$

for every $x \in I_0$. One can observe from the definition (57) that $S_0(\Lambda_j(f,g)\chi_{J_N}) \leq S_{-N}(\Lambda_j(f,g)\chi_{J_N})$, so we can deduce from (65) and (67) that

$$S_0\big(\Lambda_j(f,g)\chi_{I_0}\big)(x) \leq S_{-N}\big(\Lambda_j(f,g)\chi_{J_N}\big)(x) \lesssim_s G_s(f,g)(x)\chi_{J_N}(x) \tag{68}$$

for every $x \in I_0$. Since one can also verify that $S_0(\Lambda_j(f,g)\chi_{I_0})(x) \equiv 0$ for any $x \in I_0^c$, combining this fact with (65) and (68) conclude the proof of Lemma 4.3. □

Now we are ready to prove the localized estimate (40) for $\widetilde{\mathcal{M}_1^0}(f,g)$ in the case $r > 1$. For that purpose, we will make use of the following good-λ inequality from [10], which states that

$$|\{\sup_k |\mathbb{E}_k f - \mathbb{E}_{k_0} f| > 2\lambda\} \cap \{S_{k_0}(f) < \epsilon\lambda\}| \leq Ce^{-C_d/\epsilon^2} |\{\sup_k |\mathbb{E}_k f| > \lambda\}| \tag{69}$$

for any $\lambda > 0$ and $0 < \varepsilon < 1$. The inequality is stated for a martingale inside unit cube in $\mathbb{R}^d$, but it is clear that it can be extended to one-sided martingale starting at any level k_0.

First, notice that for $x \in I_0$,

$$\mathbb{E}_0(\Lambda_j(f,g)\chi_{I_0})(x) = 2^N \mathbb{E}_{-N}(\Lambda_j(f,g)\chi_{I_0})(x), \tag{70}$$

we can deduce from Lemma 4.3 the following estimate:

$$\begin{aligned}
&\big|\{\widetilde{\mathcal{M}_1^0}(f,g) > 4\lambda\}\big| \qquad (71)\\
\leq &\sum_{j=1}^{N} \Big|\{x \in I_0 : |\Lambda_j(f,g)\chi_{I_0} - 2^N \mathbb{E}_{-N}(\Lambda_j(f,g)\chi_{I_0})| > 2\lambda,\, G_s(f,g)(x) < \epsilon\lambda\}\Big| \\
&+ \big|\{G_s(f,g)\chi_{I_0} \geq \epsilon\lambda\}\big| + \Big|\Big\{x \in I_0 : \sup_{1\leq j\leq N} |\mathbb{E}_{-N}(\Lambda_j(f,g)\chi_{I_0})| > 2^{1-N}\lambda\Big\}\Big| \\
\leq &\sum_{1\leq j\leq N} \Big|\Big\{\sup_{k\geq 0}\Big|\sum_{l=0}^{k-1} \mathbb{D}_l(\Lambda_j(f,g)\chi_{I_0})\Big| > 2\lambda\Big\} \cap \{S_{-N}(\Lambda_j(f,g)\chi_{J_N}) < C_s\epsilon\lambda\}\Big| \\
&+ \big|\{G_s(f,g)\chi_{I_0} \geq \epsilon\lambda\}\big| + \Big|\Big\{\sup_{1\leq j\leq N} |\mathbb{E}_{-N}(\Lambda_j(f,g)\chi_{I_0})\chi_{I_0}| > 2^{1-N}\lambda\Big\}\Big| \\
\leq &\sum_{1\leq j\leq N} \Big|\Big\{\sup_{k\geq 0}\Big|\sum_{l=0}^{k-1} \mathbb{D}_l(\Lambda_j(f,g)\chi_{I_0})\Big| > 2\lambda\Big\} \cap \{S_0(\Lambda_j(f,g)\chi_{I_0}) < C_s\epsilon\lambda\}\Big| \\
&+ \big|\{G_s(f,g)\chi_{I_0} \geq \epsilon\lambda\}\big| + \Big|\Big\{\sup_{1\leq j\leq N}\Big|\int_{I_0} \Lambda_j(f,g)(y)dy \cdot \chi_{I_0}\Big| > 2\lambda\Big\}\Big|.
\end{aligned}$$

If we take $\epsilon = \frac{\sqrt{C_d}}{C_s\sqrt{r\log N}}$, multiply $r4^r\lambda^{r-1}$ and then integrate on λ in both sides of (71), then we can deduce from the good-λ inequality (69) and Lemma 4.1 the following L^r norm estimate for $r > 1$:

$$\begin{aligned}
\|\widetilde{\mathcal{M}}_1^0(f,g)\|_{L^r} &\lesssim \frac{1}{N}\sum_{1\le j\le N}\|M(\Lambda_j(f,g)\chi_{I_0})\|_{L^r} + \sqrt{\log N}\|G_s(f,g)\chi_{I_0}\|_{L^r} \\
&+ \Big\|\sup_{1\le j\le N}\Big|\int_{I_0}\Lambda_j(f,g)(y)dy\cdot\chi_{I_0}\Big|\Big\|_{L^r} \\
&\lesssim \sqrt{\log N}\|G_s(f,g)\chi_{I_0}\|_{L^r} + \sup_{1\le j\le N}\|\Lambda_j(f,g)\chi_{I_0}\|_{L^r} + \sup_{1\le j\le N}\Big|\int_{I_0}\Lambda_j(f,g)dy\Big| \\
&\lesssim \sqrt{\log N}\|G_s(f,g)\chi_{I_0}\|_{L^r} + \|f\tilde{\chi}_{I_0}\|_{L^p}\|g\tilde{\chi}_{I_0}\|_{L^q} + \sup_{1\le j\le N}\|\Lambda_j(f,g)\chi_{I_0}\|_{L^r} \\
&\lesssim \sqrt{\log N}\|G_s(f,g)\chi_{I_0}\|_{L^r} + \|f\tilde{\chi}_{I_0}\|_{L^p}\|g\tilde{\chi}_{I_0}\|_{L^q}, \qquad (72)
\end{aligned}$$

from which the localized estimate (40) follows.

5. End of the proof for Theorem 1.2.

5.1. Upper bound. Combining the estimates (39) and (40), we have proved the upper bound part of Theorem 1.2 in the cases $r > 1$. In order to cover the cases $\frac{1}{2} < r \le 1$ in Theorem 1.2, we need to prove an endpoint weak type estimate for $L^1 \times L^1 \to L^{1/2,\infty}$. Once such endpoint weak type estimate was obtained, by standard multi-linear interpolation argument (see e.g. [17, 30, 32]), we can finish the proof of the upper bound in Theorem 1.2 for general $r > \frac{1}{2}$.

The following proposition allows us to deduce the endpoint weak type estimate with the same upper bound $\mathcal{O}(\sqrt{\log(N+2)})$ from the L^r $(r > 1)$ estimates derived in Sections 2, 3 and 4.

Proposition 2. *Assume that we have a countable family of bilinear smooth symbols $\{a_j\}$ such that the condition (8) is satisfied uniformly for all a_j and the associated bilinear maximal operator $\mathcal{M}$ is bounded from $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ into $L^r(\mathbb{R})$ by a constant A for some $1 < p, q, r < \infty$ satisfying $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Then, the maximal operator $\mathcal{M}$ is bounded from $L^1(\mathbb{R}) \times L^1(\mathbb{R})$ into $L^{1/2,\infty}(\mathbb{R})$ with the norm at most a multiple of $A + B$, where B is a constant depends only on the constants from the condition (8).*

Proposition 2 has been proved in the case of maximal function given by a family of bilinear Coifman-Meyer multipliers $\{m_j\}$ satisfying estimate (5) uniformly in [22] (see Theorem 8.1 therein). The main tools of the proof in [22] are Caldrón-Zygmund decomposition and the following estimates for distribution kernels $K_j = \mathcal{F}^{-1}(m_j)$, that is,

$$|\partial^\beta K_j(x_1, x_2)| \lesssim (|x_1| + |x_2|)^{-2d+|\beta|} \qquad (73)$$

for any j and multi-index $|\beta| \le 1$. Now, in the case $d = 1$, if we define distribution kernels

$$\tilde{K}_j = \mathcal{F}^{-1}(a_j), \qquad (74)$$

where $\{a_j\}$ is the family of symbols of bilinear pseudo-differential operators satisfying (8) uniformly given by Proposition 2, then we can also obtain the same estimate as (73), that is,

$$|\partial^\beta \tilde{K}_j(x_1, x_2)| \lesssim (|x_1| + |x_2|)^{-2+|\beta|} \qquad (75)$$

for any j and multi-index $|\beta| \le 1$. For the proof of estimate (75) in details, please refer to [1]. Therefore, the proof of Proposition 2 will be completely similar to the proof of Theorem 8.1 in [22], so we omit the details here.

5.2. **An example with sharp lower bound.** In [22], the author also provided an example of a countable family $\{m_j\}$ of bilinear Coifman-Meyer multipliers, such that the L^r norm of the corresponding maximal operator $\mathfrak{M}$ larger than $C\sqrt{\log N}$ with C independent on N. More precisely, fix a smooth function ψ, such that $supp\psi \subseteq [-\frac{1}{4}, \frac{1}{4}]$ and $\psi(\xi) \equiv 1$ for every $|\xi| \leq 1/8$, then one can verify the symbols

$$m_j(\xi_1, \xi_2) := \sum_{k=1}^{\infty} j(k)\psi(2^{-k}\xi_2)\psi(2^{-k}\xi_1 - 1) \tag{76}$$

satisfy the estimate (5) uniformly in j, where $j(k)$ denotes the k-th digit of binary representation of j. Then, one can take a smooth non-zero function ϕ with $supp\hat{\phi} \subseteq [-\frac{1}{8}, \frac{1}{8}]$ and $f := \mathcal{F}^{-1}(\sum_{k=1}^{l} \hat{\phi}(\xi - 2^k))$ as test functions, where l is the integer such that $2^l \leq N < 2^{l+1}$. For the maximal operator $\mathfrak{M}$ associated with the first N symbols of $\{m_j\}$ and text functions ϕ and f, the following sharp lower bounds hold:

$$\|\mathfrak{M}(f, \phi)\|_{L^r} \gtrsim \sqrt{\log N}\|f\|_{L^p}\|\phi\|_{L^q}. \tag{77}$$

In our variable coefficient settings, if we simply consider a countable family $\{a_j\}$ of smooth symbols that are independent of the variable x, more precisely, if we take $a_j(x, \xi_1, \xi_2) = m_j(\xi_1, \xi_2)$ $(j = 1, 2, \cdots)$ and the test functions f, ϕ constructed as above, then one can easily infer from (76) that such $\{a_j\}$ do not have singularity at the origin of $\mathbb{R}^2_\xi$ and hence satisfy the estimate (8) uniformly in j, and it follows from (77) that the maximal functions $\mathcal{M}(f, \phi)$ associated with the first N symbols of $\{a_j\}$ also satisfy the following sharp lower bounds:

$$\|\mathcal{M}(f, \phi)\|_{L^r(\mathbb{R})} := \Big\| \sup_{1 \leq j \leq N} |T_{a_j}(f, \phi)(x)| \Big\|_{L^r(\mathbb{R})} \gtrsim \sqrt{\log N}\|f\|_{L^p(\mathbb{R})}\|\phi\|_{L^q(\mathbb{R})}. \tag{78}$$

This concludes the proof of Theorem 1.2.

REFERENCES

[1] Á. Bényi, D. Maldonado, V. Naibo and H. Torres, On the Hörmander classes of bilinear pseudodifferential operators, *Integr. Equ. oper. Theory*, **67** (2010), 341–364.
[2] F. Bernicot, Local estimates and global continuities in Lebesgue spaces for bilinear operators, *Anal. PDE*, **1** (2008), 1–27.
[3] D. L. Burkholder, Distribution function inequalities for martingales, *Ann. Prob.*, **1** (1973), 19–42.
[4] Á. Bényi and R. H. Torres, Symbolic calculus and the transposes of bilinear pseudodifferential operators, *Comm. Partial Differential Equations*, **28** (2003), 1161–1181.
[5] M. Christ, L. Grafakos, P. Honzík and A. Seeger, Maximal functions associated with multipliers of Mikhlin-Hörmander type, *Math. Zeit.*, **249** (2005), 223–240.
[6] M. Christ and J. L. Journé, Polynomial growth estimates for multilinear singular integral operators, *Acta Math.*, **159** (1987), 51–80.
[7] J. Chen and G. Lu, Hörmander type theorems for multi-linear and multi-parameter Fourier multiplier operators with limited smoothness, *Nonlinear Anal.*, **101** (2014), 98–112.
[8] R. Coifman and Y. Meyer, On commutators of singular integrals and bilinear singular integrals, *Trans. Amer. Math. Soc.*, **212** (1975), 315–331.
[9] R. Coifman and Y. Meyer, Au delá des opérateurs pseudo-différentiels, *Astérisque*, **57** (1978).
[10] S. Y. A. Chang, M. Wilson and T. Wolff, Some weighted norm inequalities concerning the Schrödinger operator, *Comment. Math. Helv.*, **60** (1985), 217–246.
[11] J. Duoandikoetxea, *Fourier Analysis*, Grad. Stud. Math., vol. **29**, Amer. Math. Soc., Providence, RI, 2001.
[12] W. Dai and G. Lu, L^p estimates for multi-linear and multi-parameter pseudo-differential operators, *Bull. Soc. Math. France*, **143** (2015), 567–597.
[13] H. Dappa and W. Trebels, On maximal functions generated by Fourier multipliers, *Ark. Mat.*, **23** (1985), 241–259.

[14] C. Fefferman and E. M. Stein, Some maximal inequalities, *Amer. J. Math.*, **93** (1971), 107–115.

[15] L. Grafakos, P. Honzík and A. Seeger, On maximal functions for Mikhlin-Hörmander multipliers, *Adv. Math.*, **204** (2006), 363–378.

[16] L. Grafakos and N. J. Kalton, The Marcinkiewicz multiplier condition for bilinear operators, *Studia Math.*, **146** (2001), 115–156.

[17] L. Grafakos and T. Tao, Multilinear interpolation between adjoint operators, *J. Funct. Anal.*, **199** (2003), 379–385.

[18] L. Grafakos and R. Torres, Multilinear Calderón-Zygmund theory, *Adv. Math.*, **165** (2002), 124–164.

[19] Q. Hong and G. Lu, Symbolic calculus and boundedness of multi-parameter and multi-linear pseudo-differential operators, *Adv. Nonlinear Stud.*, **14** (2014), 1055–1082.

[20] Q. Hong and L. Zhang, L^p estimates for bi-parameter and bilinear Fourier integral operators, *Acta Mathematica Sinica - English Series*, (2016).

[21] L. Hörmander, Estimates for translation invariant operators in L^p spaces, *Acta Math.*, **104** (1960), 93–140.

[22] P. Honzík, Maximal functions of multilinear multipliers, *Math. Res. Lett.*, **16** (2009), 995–1006.

[23] C. E. Kenig and E. M. Stein, Multilinear estimates and fractional integration, *Math. Res. Lett.*, **6** (1999), 1–15.

[24] G. Lu and L. Zhang, L^p estimates for a trilinear pseudo-differential operator with flag symbols, *Indiana University Mathematics Journal*, to appear.

[25] G. Lu and P. Zhang, Multilinear Calderón-Zygmund operators with kernels of Dini's type and applications, *Nonlinear Anal.*, **107** (2014), 92–117.

[26] C. Muscalu, Unpublished notes I, IAS Princeton, 2003.

[27] C. Muscalu, Paraproducts with flag singularities I. A case study, *Rev. Mat. Iberoam.*, **23** (2007), 705–742.

[28] C. Muscalu, J. Pipher, T. Tao and C. Thiele, Bi-parameter paraproducts, *Acta Math.*, **193** (2004), 269–296.

[29] C. Muscalu, J. Pipher, T. Tao and C. Thiele, Multi-parameter paraproducts, *Rev. Mat. Iberoam.*, **23** (2007), 705–742.

[30] C. Muscalu and W. Schlag, *Classical and multilinear Harmonic Analysis, II*, Cambridge Studies in Advanced Mathematics, vol. **138**, Cambridge University Press, Cambridge, 2013.

[31] C. Muscalu, T. Tao and C. Thiele, Multilinear operators given by singular multipliers, *J. Amer. Math. Soc.*, **15** (2002), 469–496.

[32] C. Muscalu, T. Tao and C. Thiele, L^p estimates for the biest II. The Fourier case, *Math. Ann.*, **329** (2004), 427–461.

[33] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Priceton Univ. Press, Princeton, NJ, 1970.

Received July 2016; revised December 2016.

E-mail address: chenjiao@mail.bnu.edu.cn
E-mail address: weidai@buaa.edu.cn
E-mail address: guozhen.lu@uconn.edu