New York Journal of Mathematics

New York J. Math. **14** (2008) 145–192.

Uniform estimates for some paraproducts

Xiaochun Li

Abstract. We establish $L^p \times L^q$ to L^r estimates for some general paraproducts that arise in the study of the bilinear Hilbert transform along curves.

Contents

1. Introduction 145
2. Main results 147
3. A telescoping argument 149
4. Time frequency analysis 153
   4.1. Definitions 154
   4.2. Reduction 158
   4.3. Principal lemmas 159
   4.4. The size estimate for a tree 165
   4.5. Proof of Lemma 4.11 174
5. Proof of Theorem 2.2 175
   5.1. Proof of Lemma 5.1 176
   5.2. A truncated trilinear form 181
   5.3. Preliminary lemmata 183
   5.4. Proof of Lemma 5.10 190
References 191

1. Introduction

It is an important theme of current research in analysis to decompose more complicated operators, such as the Cauchy integral on Lipschitz curves [1], as a sum of simpler operators. This theme has taken special prominence in

Received January 17, 2008.
Mathematics Subject Classification. Primary 42B20, 42B25. Secondary 46B70, 47B38.
Key words and phrases. paraproduct, uniform estimate.
Research was partially supported by NSF grant DMS-0456976.

ISSN 1076-9803/08

multilinear Harmonic Analysis, beginning with the work of Lacey and Thiele [12], which expressed the bilinear Hilbert transforms as a sum of modulated paraproducts. This theme has found much broader application as well.

The bilinear Hilbert transforms have a bilinear symbol given by restriction to a half-plane, with slope that depends upon the transform in question. In considering more complicated symbols, one is led to paraproducts which have a complicated underlying description. One then seeks certain estimates of these paraproducts that are *uniform* in the parametrizations. This line of investigation was started in [23], the results of which give a new, multilinear proof of the boundedness of the Calderon commutator, fulfilling a program of study of Calderon [1]. It was further extended in work of the author and Grafakos [8, 9, 14], in the study of the disc as a bilinear multiplier. Muscalu, Tao and Thiele [16, 15, 17] gave alternate proofs (and more general proofs) of these results in the multilinear operator setting.

In this paper, we continue this line of study, considering certain uniform estimates that are motivated by an analysis of a bilinear Hilbert transform along polynomial curves. Namely, consider the operators

$$(f,g) \longrightarrow \text{p.v.} \int_{-\infty}^{\infty} f(x-y)g(x-p(y))\,\frac{dy}{y}, \tag{1.1}$$

for some polynomial $p(y)$. The study of these operators leads to subtle questions in multilinear analysis, stationary phase methods, and paraproducts. An initial investigation into operators of this type is given in [6], where the polynomial is taken to be a square, and the singular kernel is mollified to $e^{i|y|^{-\beta}}/|y|$ for some $\beta > 0$. Without this modification, a significant difficulty might be encountered. There is a natural analogue of the bilinear Hilbert transform along parabolas in the ergodic theory setting, that is, the nonconventional ergodic average $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)g(T^{n^2}x)$. In [7], Furstenberg proved that the characteristic factor of the trilinear ergodic averages $\frac{1}{N}\sum_{n=0}^{N-1} f(T^{an})g(T^{bn})h(T^{cn})$ for all $a, b, c \in \mathbb{Z}$ is characteristic for the previous nonconventional ergodic average. We are indebted to M. Lacey for bringing these Furstenberg theorems to our attention. Thus a possible method for the bilinear Hilbert transform along a parabola is to understand the trilinear Hilbert transform first. Unfortunately, it turns out the trilinear Hilbert transform is very difficult to handle. It is very interesting to find a proof for the bilinear Hilbert transform along curves without using any information of the trilinear Hilbert transform. It might be possible to obtain such a way by combining time-frequency analysis and the known results for the trilinear oscillatory integrals. This investigation will appear in another paper.

The paraproducts that arise have a richer parametrization than what has been considered before. The question of uniform estimates is the main focus of this article. In the next section, a class of paraproducts are introduced. They are parametrized by:

- the *width* of the frequency window associated to the paraproducts, denoted by L_1 and L_2 below,
- the *overlap* of the frequency window associated to the paraproducts, denoted by M_1 and M_2 below,
- a *modulation* of the frequency window, denoted by the (lower case) parameters $n_1, n_2, 2^m$ below.

Prior results have concentrated on the uniformity of estimates with respect to M_1, M_2 from $L^p \times L^q$ to L^r for $r \geq 1$ and $L_1 = L_2$ [16]. The principal point of this article is to get the estimates for $1/2 < r < 1$ and arbitrary L_1, L_2. Another new point of this article is the (weak) uniformity that we establish in L_1, L_2 and the modulation parameters 2^m (see Theorem 2.2 below). This novelty is forced upon us by the stationary phase methods that one must use in the analysis of (1.1). One of anticipated applications of our theorems is the bilinear multiplier problems associated to the symbol defined by a characteristic function of a suitable domain with a smooth boundary.

Acknowledgements. The author would like to thank his wife, Helen, and his son, Justin, for being together through the hard times in the past two years. And he is also very thankful to Michael Lacey for his constant support and encouragement.

2. Main results

Let $j \in \mathbb{Z}$, L_1, L_2 be positive integers and M_1, M_2 be integers.

$$\omega_{1,j} = [2^{L_1 j + M_1}/2, 2 \cdot 2^{L_1 j + M_1}]$$

and

$$\omega_{2,j} = [-2^{L_2 j + M_2}, 2^{L_2 j + M_2}].$$

Let Φ_1 be a Schwartz function whose Fourier transform is a standard bump function supported on $[1/2, 2]$, and Φ_2 be a Schwartz function such that $\widehat{\Phi}_2$ is a standard bump function supported on $[-1, 1]$ and $\widehat{\Phi}_2(0) = 1$. For $\ell \in \{1, 2\}$ and $n_1, n_2 \in \mathbb{Z}$, define Φ_{ℓ,j,n_ℓ} by

$$\widehat{\Phi}_{\ell,j,n_\ell}(\xi) = \big(e^{2\pi i n_\ell (\cdot)} \widehat{\Phi}_\ell(\cdot)\big)\left(\frac{\xi}{2^{L_\ell j + M_\ell}}\right).$$

It is clear that $\widehat{\Phi}_{\ell,j,n_\ell}$ is supported on $\omega_{\ell,j}$. For locally integrable functions f_ℓ's, we define $f_{\ell,j}$'s by

$$f_{\ell,j,n_\ell}(x) = f_\ell * \Phi_{\ell,j,n_\ell}(x).$$

We define a paraproduct to be

$$\Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1, f_2)(x) = \sum_{j \in \mathbb{Z}} \prod_{\ell=1}^{2} f_{\ell,j,n_\ell}(x). \tag{2.1}$$

Another paraproduct we should introduce is the following. For $\ell \in \{1,2\}$, let $\omega'_{\ell,j}$ denote the set $\{\xi : 2^{L_\ell j+M_\ell}/2 \le |\xi| \le 2\cdot 2^{L_\ell j+M_\ell}\}$. Let m be a nonnegative integer and define $\Phi_{\ell,j,m}$ by

$$\widehat{\Phi}_{\ell,j,m}(\xi) = \big(e^{2\pi i 2^m(\cdot)}\widehat{\Phi}_1(\cdot)\big)\bigg(\frac{\xi}{2^{L_\ell j+M_\ell}}\bigg).$$

Let $f_{\ell,j,m}$ be the function defined by

$$f_{\ell,j,m}(x) = f_\ell * \Phi_{\ell,j,m}(x).$$

We define a paraproduct to be

$$\Pi_{L_1,L_2,M_1,M_2,m}(f_1,f_2)(x) = \sum_{j\in\mathbb{Z}}\prod_{\ell=1}^{2} f_{\ell,j,m}(x). \tag{2.2}$$

One reason we study these paraproducts is that one will encounter such paraproducts in the study of the bilinear Hilbert transforms along polynomial curves. We have the following uniform estimates for these paraproducts.

Theorem 2.1. *For any $p_1>1$, $p_2>1$ with $1/p_1+1/p_2=1/r$, there exists a constant C independent of M_1,M_2,n_1,n_2 such that*

$$\big\|\Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2)\big\|_r \le C\big(1+|n_1|\big)^{10}\big(1+|n_2|\big)^{10}\|f_1\|_{p_1}\|f_2\|_{p_2}, \tag{2.3}$$

for all $f_1\in L^{p_1}$ and $f_2\in L^{p_2}$.

Theorem 2.2. *Let $\Pi_{L_1,L_2,M_1,M_2,m}(f_1,f_2)$ be as in (2.2). Suppose that for all j,*

$$2^{L_2 j+M_2} \ge 2^{L_1 j+M_1+m}. \tag{2.4}$$

For any $\varepsilon>0$, $p_1>1$, $p_2>1$ with $1/p_1+1/p_2=1/r$, there exists a constant C independent of m,M_1,M_2,L_1,L_2 such that

$$\big\|\Pi_{L_1,L_2,M_1,M_2,m}(f_1,f_2)\big\|_r \le C2^{\varepsilon m}\|f_1\|_{p_1}\|f_2\|_{p_2}, \tag{2.5}$$

for all $f_1\in L^{p_1}$ and $f_2\in L^{p_2}$.

The case when $L_1=L_2$ and $r>1$ was proved in [16]. The constant C in Theorem 2.1 may depend on L_1,L_2. It is easy to see by the following argument that C is $O(\max\{2^{L_1},2^{L_2}\})$. It is possible to get a much better upper bound such as $O\big(\log(1+\max\{L_2/L_1,L_1/L_2\})\big)$ by tracking the constants carefully in the proof we will provide. But we do not pursue the sharp constant in this article. The independence of M_1,M_2 is the most important issue. In Sections 3, 4, we give a proof for Theorem 2.1. The proof of Theorem 2.2 will be given in Section 5. By using Theorem 2.1, we get the L^r bound for $\Pi_{L_1,L_2,M_1,M_2,m}$ with a operator norm $O(2^{10m})$. Unfortunately sometimes this is not enough for our application. The desired norm is $O(2^{\varepsilon m})$ for a very small positive number ε. It might be possible to remove the condition (2.4) or get the uniform estimate for $\Pi_{L_1,L_2,M_1,M_2,m}$ in which

the operator norm is independent of m. The uniform estimate from $L^2 \times L^2$ to L^1 is trivial and (2.4) is redundant for this case. In Section 5, we see that the uniform estimates for $\Pi_{L_1,L_2,M_1,M_2,m}$ can be achieved for $p_1, p_2 > 2$ and $1 < r < 2$ (see Proposition 5.7) and (2.4) is superfluous for Theorem 2.2 when $p_1, p_2 > 2$ and $1 < r < 2$ (see Corollary 5.3).

3. A telescoping argument

We now start to prove Theorem 2.1. To prove Theorem 2.1, we first introduce a definition of admissible trilinear form. And we should show that by a telescoping argument used in [8, 23], we can reduce the problem to estimates for an admissible trilinear form. And thus L^r estimates for $r > 1$ can be obtained by Littlewood–Paley theorem. The $r < 1$ case is more complicated. We have to use the time frequency analysis to deal with this case in Section 4.

Definition 3.1. An admissible trilinear form is a trilinear form

$$\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2,f_3) = \int \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} \widetilde{f}_{\ell,j,n_\ell}(x)dx, \tag{3.1}$$

where $n_3 = 0$, $\widetilde{f}_{\ell,j,n_\ell} = f_\ell * \widetilde{\Phi}_{\ell,j,n_\ell}$ and $\widetilde{\Phi}_{\ell,j,n_\ell}$ is a function whose Fourier transform is supported on $\widetilde{\omega}_{\ell,j}$ such that:

(1) Each $\widetilde{\omega}_{\ell,j}$ is an interval in $\mathbb{R}$ such that the distance from the origin to the interval is not more than $3|\widetilde{\omega}_{\ell,j}|$. And $\{\widetilde{\omega}_{\ell,j}\}_j$ forms a sequence of lacunary intervals, that is, $|\widetilde{\omega}_{\ell,j}|/|\widetilde{\omega}_{\ell,j+1}| \le 1/2$ for all $j \in \mathbb{Z}$. Moreover, $|\widetilde{\omega}_{3,j}| \ge C\max\{|\widetilde{\omega}_{1,j}|, |\widetilde{\omega}_{2,j}|\}$ for some constant C independent of M_1, M_2, n_1, n_2.

(2) There are at least two indices $\ell \in \{1,2,3\}$ such that $\widetilde{\Phi}_{\ell,j,n_\ell}$ satisfies

$$\widehat{\widetilde{\Phi}_{\ell,j,n_\ell}}(0) = 0 \tag{3.2}$$

$$\left| D^\alpha \left(\widehat{\widetilde{\Phi}_{\ell,j,n_\ell}}\big(|\widetilde{\omega}_{\ell,j}|\xi\big) \right) \right| \le \frac{C_N(1+|n_\ell|)^\alpha}{(1+|\xi|)^N}, \tag{3.3}$$

for all $\xi \in \mathbb{R}$ and all nonnegative integers α, N. If an index in $\{1,2,3\}$ satisfies (3.2) and (3.3), we call the index a good index in the trilinear form $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}$. For the index which is not a good index, we call it a bad index in the trilinear form $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}$.

(3) If $\ell \in \{2,3\}$ is a bad index, then $\widetilde{\Phi}_{\ell,j,n_\ell}$ satisfies (3.3). Moreover, among the other two good indices $\ell' \ne \ell$, at least one of them satisfies $|\widetilde{\omega}_{\ell',j}| \le C\min\{|\widetilde{\omega}_{1,j}|, |\widetilde{\omega}_{2,j}|, |\widetilde{\omega}_{3,j}|\}$ for some constant C independent of f_1, f_2, f_3, M_1, M_2, n_1, n_2.

(4) If 1 is a bad index, then $\widetilde{\Phi}_{1,j,n_1}$ satisfies

$$\widetilde{\Phi}_{1,j,n_1}(x) = \sum_{k=0}^{m'(j)} \Phi_{1,j+k,n_1}(x), \tag{3.4}$$

where $m'(j)$ is some nonnegative integer.

Lemma 3.2. *Let f_3 be a locally integrable function. Then*

$$\int \Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2)(x)f_3(x)dx$$

is a sum of finitely many admissible trilinear forms such that the number of admissible trilinear forms in the sum is no more than a constant C independent of M_1, M_2, n_1, n_2.

Proof. For $\ell \in \{1,2\}$, write $\omega_{\ell,j}$ as $[a_{\ell,j}, b_{\ell,j}]$. If $b_{2,j} < b_{1,j}/16$, then $|\omega_{2,j}| < |\omega_{1,j}|/6$ and the distance from $\omega_{1,j}+\omega_{2,j}$ to the origin is not less than $|\omega_{1,j}|/4$. In this case, simply let $\widetilde{\omega}_{3,j}$ be a small neighborhood of $-(\omega_{1,j}+\omega_{2,j})$ and the Fourier transform of $\widetilde{\Phi}_{3,j}$ is a suitable bump function adapted to $\widetilde{\omega}_{3,j}$, then we have the desired lemma. Thus we now only consider the case $b_{2,j} \geq b_{1,j}/16$. Let $\omega_{3,j}$ be $[-18b_{2,j}, 18b_{2,j}]$. And $\Phi_{3,j}$ be a Schwartz function such that its Fourier transform is a bump function adapted to $\omega_{3,j}$ and $\widehat{\Phi}_{3,j}(\xi) = 1$ for all $\xi \in [-17b_{2,j}, 17b_{2,j}]$. Then

$$\int \Pi(f_1,f_2)(x)f_3(x)dx = \int \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} f_{\ell,j,n_\ell}(x)dx,$$

where $f_{3,j,n_3}(x) = f_3 * \Phi_{3,j}(x)$ and $n_3 = 0$. Let $\widetilde{\Phi}_2$ be a Schwartz function such that $\widehat{\widetilde{\Phi}_2}$ is a bump function on $[-1,1]$ and $\widehat{\widetilde{\Phi}_2}(\xi) = 1$ for all $\xi \in [-3/4, 3/4]$. And define $\Phi_{2,j}$ by $\widehat{\Phi}_{2,j}(\xi) = \widehat{\widetilde{\Phi}_2}(\xi/b_{2,j})$. Let $f_{2,j} = f * \Phi_{2,j}$. We also denote f_{3,j,n_3} by $f_{3,j}$. We can replace f_{2,j,n_2} by $f_{2,j}$ because

$$\int \sum_{j\in\mathbb{Z}} f_{\ell,j,n_1}(x)\big(f_{2,j,n_2} - f_{2,j}\big)(x)f_{3,j}(x)dx$$

is an admissible trilinear form. Hence the only thing we need to show is that

$$\Lambda'(f_1,f_2,f_3) = \int \sum_{j\in\mathbb{Z}} f_{\ell,j,n_1}(x)f_{2,j}(x)f_{3,j}(x)dx$$

is admissible. For any real number x, let $[x]$ denote the largest integer not exceeding x. Let $m(j)$ be the integer defined by

$$m(j) = \left[\frac{(L_2 j + M_2) - (L_1 j + M_1) + 6}{L_2}\right].$$

By $b_{2,j} \geq b_{1,j}/16$, we see that $m(j) \geq 0$. By a telescoping argument, $\Lambda'(f_1, f_2, f_3)$ equals to

$$\int \sum_{j\in\mathbb{Z}} f_{1,j,n_1}(x) \sum_{k=0}^{m(j)} \Big(f_{2,j-k}(x) f_{3,j-k}(x) - f_{2,j-k-1}(x) f_{3,j-k-1}(x) \Big) dx,$$

since $\int f_{1,j,n_1}(x) f_{2,j-m(j)-1}(x) f_{3,j-m(j)-1}(x) dx = 0$ due to the following simple fact on the support of Fourier transform of each function in the integrand, i.e.,

$$\Big(\operatorname{supp} \widehat{f}_{1,j,n_1} + \operatorname{supp} \widehat{f}_{2,j-m(j)-1} \Big) \cap \Big(- \big(\operatorname{supp} \widehat{f}_{3,j-m(j)-1} \big) \Big) = \emptyset.$$

By a change of variables $j \to j+k$, we have that $\Lambda'(f_1, f_2, f_3)$ is equal to

$$\int \sum_{j\in\mathbb{Z}} \sum_{k=0}^{m'(j)} f_{1,j+k,n_1}(x) \Big(f_{2,j}(x) f_{3,j}(x) - f_{2,j-1}(x) f_{3,j-1}(x) \Big) dx,$$

where $m'(j)$ is the integer defined by

$$m'(j) = \left[\frac{(L_2 j + M_2) - (L_1 j + M_1) + 6}{L_1} \right].$$

We write this integral as a sum of three parts $\Lambda_1, \Lambda_2, \Lambda_3$, where

$$\Lambda_1 = \int \sum_{j\in\mathbb{Z}} \left(\sum_{k=0}^{m'(j)} f_{1,j+k,n_1}(x) \right) f_{2,j}(x) \big(f_{3,j}(x) - f_{3,j-1}(x) \big) dx,$$

$$\Lambda_2 = \int \sum_{j\in\mathbb{Z}} \left(\sum_{k=0}^{m'(j)} f_{1,j+k,n_1}(x) \right) \cdot \big(f_{2,j}(x) - f_{2,j-1}(x) \big) \big(f_{3,j-1}(x) - f_{3,j-8}(x) \big) dx,$$

$$\Lambda_3 = \int \sum_{j\in\mathbb{Z}} \left(\sum_{k=0}^{m'(j)} f_{1,j+k,n_1}(x) \right) \big(f_{2,j}(x) - f_{2,j-1}(x) \big) f_{3,j-8}(x) dx.$$

It is clear that Λ_2 is an admissible trilinear form. Write Λ_1 as $\Lambda_{11} + \Lambda_{12}$, where

$$\Lambda_{11} = \int \sum_{j\in\mathbb{Z}} \left(\sum_{k=0}^{m'(j)} f_{1,j+k,n_1}(x) \right) \big(f_{2,j}(x) - f_{2,j-1}(x) \big) \big(f_{3,j}(x) - f_{3,j-1}(x) \big) dx,$$

$$\Lambda_{12} = \int \sum_{j\in\mathbb{Z}} \left(\sum_{k=0}^{m'(j)} f_{1,j+k,n_1}(x) \right) f_{2,j-1}(x) \big(f_{3,j}(x) - f_{3,j-1}(x) \big) dx.$$

Clearly, Λ_{11} is an admissible trilinear form. Notice that

$$\operatorname{supp}\Bigg(\sum_{k=0}^{m'(j)-10-[L_2/L_1]} \widehat{f}_{1,j+k,n_1}\Bigg) \subseteq [0, 2^{-2}2^{L_2 j+M_2}] = [0, 2^{-2}2^{-L_2}b_{2,j}],$$

and

$$\operatorname{supp}\big(\widehat{f}_{3,j} - \widehat{f}_{3,j-1}\big) \subseteq [-18b_{2,j}, 18b_{2,j}]\backslash[-16\cdot 2^{-L_2}b_{2,j}, 16\cdot 2^{-L_2}b_{2,j}].$$

Thus Λ_{12} is equal to

$$\int \sum_{j\in\mathbb{Z}} \Bigg(\sum_{k=m'(j)-10-[L_2/L_1]}^{m'(j)} f_{1,j+k,n_1}(x)\Bigg) \cdot f_{2,j-1,n_2}(x)\big(f_{3,j,n_3}(x) - f_{3,j-1,n_3}(x)\big)dx,$$

which is obviously a finite sum of admissible trilinear forms. As for Λ_3, observe that

$$\operatorname{supp}\Bigg(\sum_{k=0}^{m'(j)-100-[L_2/L_1]} \widehat{f}_{1,j+k,n_1}\Bigg) \subseteq [0, 2^{-80}2^{L_2 j+M_2}] = [0, 2^{-80}2^{-L_2}b_{2,j}],$$

and

$$\operatorname{supp}\big(\widehat{f}_{2,j} - \widehat{f}_{2,j-1}\big) \subseteq [-b_{2,j}, b_{2,j}]\backslash[-2^{-L_2-1}b_{2,j}, 2^{-L_2-1}b_{2,j}].$$

Thus Λ_3 is equal to

$$\int \sum_{j\in\mathbb{Z}} \Bigg(\sum_{k=m'(j)-100-[L_2/L_1]}^{m'(j)} f_{1,j+k,n_1}(x)\Bigg)\big(f_{2,j} - f_{2,j-1}(x)\big) f_{3,j-8}(x)dx,$$

which is a finite sum of admissible trilinear forms. □

Lemma 3.3. *Let $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}$ be an admissible trilinear form. Then for any real numbers $p_1, p_2, p_3 > 1$ with $1/p_1 + 1/p_2 + 1/p_3 = 1$, there exists C independent of M_1, M_2, n_1, n_2 such that*

(3.5)

$$\big|\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1, f_2, f_3)\big| \le C(1+|n_1|)^{10}(1+|n_2|)^{10}\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3},$$

for all $f_1 \in L^{p_1}$, $f_2 \in L^{p_2}$ and $f_3 \in L^{p_3}$.

Proof. If there is no bad index in the trilinear form, take ℓ_0 to be any integer in $\{1, 2, 3\}$. Otherwise, let ℓ_0 be a bad index. Applying the Cauchy–Schwarz inequality, $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}$ is dominated by

$$\int \sup_{j\in\mathbb{Z}} |\widetilde{f}_{\ell_0,j,n_{\ell_0}}| \prod_{\ell\neq\ell_0} \Big(\sum_j |\widetilde{f}_{\ell,j,n_\ell}|^2\Big)^{1/2} dx.$$

Using Hölder's inequality, we dominate the trilinear form by

$$\Big\| \sup_{j\in\mathbb{Z}} |\widetilde{f}_{\ell_0,j,n_{\ell_0}}| \Big\|_{p_1} \prod_{\ell\neq\ell_0} \Big\| \Big(\sum_j |\widetilde{f}_{\ell,j,n_\ell}|^2\Big)^{1/2} \Big\|_{p_\ell}.$$

The Littlewood–Paley theorem yields that for $\ell \neq \ell_0$

$$\Big\| \Big(\sum_j |\widetilde{f}_{\ell,j,n_\ell}|^2 \Big)^{1/2} \Big\|_{p_\ell} \le C(1+|n_\ell|)^{10} \|f_\ell\|_{p_\ell}.$$

If $\ell_0 \in \{2,3\}$, then by (3.3), we have

$$\sup_{j\in\mathbb{Z}} |\widetilde{f}_{\ell_0,j,n_{\ell_0}}| \le (1+|n_{\ell_0}|)^{10} M(f_{\ell_0}),$$

which clearly yields the lemma. We now only need to consider the case $\ell_0 = 1$. It suffices to prove that

$$\Big\| \sup_j \Big| \sum_{k=0}^{m'(j)} f_1 * \Phi_{1,j+k,n_1} \Big| \Big\|_{p_1} \le C(1+|n_1|)^{10} \|f_1\|_{p_1}. \tag{3.6}$$

Notice that $\omega_{1,j}$'s are essentailly disjoint intervals and Fourier transform of $\sum_{k=0}^{m'(j)} \Phi_{1,j+k,n_1}$ is supported on a bounded interval depending on j. The left-hand side of (3.6) is less than

$$C \Big\| M\Big(\sum_j f_1 * \Phi_{1,j,n_1} \Big) \Big\|_{p_1}.$$

It is easy to verify that $\sum_j f_1 * \Phi_{1,j,n_1}$ is a bounded operator on L^2 associated to a standard Calderón–Zygmund kernel by paying at most a cost of $(1+|n_1|)^{10}$ in the corresponding estimates. Thus by a standard Calderón–Zygmund argument, we have for any real number $p > 1$, there is a constant C independent of M_1, M_2, n_1, n_2 such that

$$\Big\| \sum_j f * \Phi_{1,j,n_1} \Big\|_p \le C(1+|n_1|^{10}) \|f\|_p$$

holds for all $f \in L^p$, which yields (3.6). Therefore we complete the proof of the lemma. □

Combining Lemma 3.2 and Lemma 3.3, we obtain (2.3) for $p_1, p_2, r > 1$. To finish the proof of Theorem 2.1, we need to provide a proof of L^r estimate with $1/2 < r \le 1$ for (2.3), which will be given in Section 4.

4. Time frequency analysis

In this section we prove (2.3) with $1/2 < r \le 1$ for the paraproducts by time frequency analysis, which was used for establishing L^p (uniform) estimates for the bilinear Hilbert transforms in [9, 12, 13, 14, 15, 16, 17, 23].

Let F be a measurable set in $\mathbb{R}$. $X(F)$ denotes the set of all measurable functions supported on F such that the L^∞ norms of the functions are no more than 1. A function in $X(F)$ can be considered essentially as the characteristic function $\mathbf{1}_F$.

To obtain Theorem 2.1, by Lemma 3.3, an interpolation argument in [15], and the scaling invariance, it is sufficient to prove that for any $p_1, p_2 > 1$ such that $1/p_1 + 1/p_2 \geq 1$ and any measurable set $F_3 \subseteq \mathbb{R}$ with $|F_3| = 1$, there exists a subset $F_3' \subset F_3$ such that $|F_3'| \geq 1/2$ and

$$\left| \int \Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2)(x) f_3(x) dx \right| \leq C(1+|n_1|)^{10}(1+|n_2|)^{10}|F_1|^{1/p_1}|F_2|^{1/p_2} \tag{4.1}$$

holds for all $f_1 \in X(F_1), f_2 \in X(F_2), f_3 \in X(F_3')$, where C is a constant independent of $f_1, f_2, f_3, M_1, M_2, n_1, n_2$.

If $2^{L_2 j+M_2} < 2^{L_1 j+M_1}/8$, let $\omega_{3,j} = [-19 \cdot 2^{L_1 j+M_1}/8, -2^{L_1 j+M_1}/8]$ and $\Phi_{3,j}$ be a Schwartz function whose Fourier transform is a bump function adapted to $\omega_{3,j}$ such that $\widehat{\Phi}_{3,j}(\xi) = 1$ for all $\xi \in [-9\cdot 2^{L_1 j+M_1}/4, -2^{L_1 j+M_1}/4]$. If $2^{L_2 j+M_2} \geq 2^{L_1 j+M_1}/8$, let $\omega_{3,j} = [-18 \cdot 2^{L_2 j+M_2}, 18 \cdot 2^{L_2 j+M_2}]$ and $\Phi_{3,j}$ be a Schwartz function whose Fourier transform is a bump function adapted to $\omega_{3,j}$ such that $\widehat{\Phi}_{3,j}(\xi) = 1$ for all $\xi \in [-17 \cdot 2^{L_2 j+M_2}, 17 \cdot 2^{L_2 j+M_2}]$. Let $n_3 = 0$, $\Phi_{3,j,n_3} = \Phi_{3,j}$, $f_{3,j,n_3}(x) = f_3 * \Phi_{3,j,n_3}(x)$. Define a trilinear form $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}$ by

$$\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2,f_3) = \int \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} f_{\ell,j,n_\ell}(x) dx. \tag{4.2}$$

Clearly $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2} = \int \Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2)(x) f_3(x) dx$. Thus to prove (4.1), it suffices to prove the following lemma.

Lemma 4.1. *Let $p_1, p_2 > 1$ such that $1/p_1 + 1/p_2 \geq 1$ and $\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}$ be the trilinear form defined by (4.2). Let F_1, F_2, F_3 be measurable sets in $\mathbb{R}$ with $|F_3| = 1$. Then there exists a subset $F_3' \subseteq F_3$ such that $|F_3'| > 1/2$ and there exists a constant C independent of F_1, F_2, F_3, f_1, f_2, f_3, M_1, M_2, n_1, n_2 such that*

$$\left|\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2,f_3)\right| \leq C(1+|n_1|)^{10}(1+|n_2|)^{10}|F_1|^{1/p_1}|F_2|^{1/p_2} \tag{4.3}$$

holds for all $f_1 \in X(F_1), f_2 \in X(F_2), f_3 \in X(F_3')$.

Lemma 4.1 and Lemma 3.3 implies the estimates (2.3) by an interpolation argument in [15]. Therefore we obtain Theorem 2.1 once we finish a proof of Lemma 4.1. The following subsections are devoted to proof of Lemma 4.1.

4.1. Definitions. To prove Lemma 4.1, we introduce some definitions first. Let ψ be a nonnegative Schwartz function such that $\widehat{\psi}$ is supported in $[-1/100, 1/100]$ and satisfies $\widehat{\psi}(0) = 1$. Let $\psi_k(x) = 2^k \psi(2^k x)$ for any $k \in \mathbb{Z}$. For $j \in \mathbb{Z}$ and $\ell \in \{1,2,3\}$, define $k_{j\ell}$ to be an integer such that $|\omega_{\ell,j}| \sim 2^{k_{j\ell}}$.

Denote $\min_{\ell\in\{1,2,3\}} k_{j\ell}$ by k_j. And define

$$I_{k_j,n} = [2^{-k_j}n, 2^{-k_j}(n+1)].$$

Define

$$\mathbf{1}^*_{j,n}(x) = \mathbf{1}_{I_{k_j,n}} * \psi_{k_j}(x).$$

It is easy to see that

$$\Lambda_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1,f_2,f_3) = \int \sum_{j\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} \mathbf{1}^*_{j,n}(x)\prod_{\ell=1}^{3} f_{\ell,j,n_\ell}(x)dx.$$

For an integer γ with $0 \le \gamma < 2^{100}$, let $\mathbb{Z}(\gamma)$ be the set of all integers congruent to γ modulo 2^{100}. For $\mathbf{S} \subset \mathbb{Z}(\gamma)\times\mathbb{Z}$ we define

$$\Lambda_{\mathbf{S}}(f_1,f_2,f_3) = \int_{\mathbb{R}} \sum_{(j,n)\in\mathbf{S}} \mathbf{1}^*_{j,n}(x)\prod_{\ell=1}^{3} f_{\ell,j,n_\ell}(x)dx. \tag{4.4}$$

$\Lambda_{\mathbf{S}}$ depends on $L_1, L_2, M_1, M_1, n_1, n_2$. We suppress this dependence for notational convenience. Note that there are finite congruence classes modulo 2^{100}. We will therefore concentrate on proving Lemma 4.1 for the trilinear form $\Lambda_{\mathbf{S}}$.

In time-frequency space, each function $f_{\ell,j,n}$ for $\ell \in \{1,2,3\}$ corresponds to a box $I_{k_j,n}\times\omega_{\ell,j}$. The most difficult situation is when only one of boxes is the Heisenberg box, i.e., $|I_{k,j,n}||\omega_{\ell,j}| \sim 1$. In this situation, we can use the John–Nirenberg type argument to get the equivalence of L^p estimates of Littlewood–Paley type square functions for only one of functions. For other two functions, there is no such an equivalence and an extra cost for it has to been paid if one estimates the BMO norm. It turns out that the L^p equivalence for at least one of three functions is the most crucial key to solve the problem. Our proof will heavily rely on this equivalence for one of functions.

Let p be a positive number close to 1. To obtain the Lemma 4.1, it suffices to prove (4.3) for $p_1 \ge p$, $p_2 \ge p$ and $1/p_1 + 1/p_2 \ge 1$. For simplicity, we only deal with the case $n_1 = n_2 = n_3 = 0$. The general case can be handled in the same way by paying at most a cost of $(1+|n_1|)^{10}(1+|n_2|)^{10}$ in the constants.

We now start to prove that for $n_1 = n_2 = 0$, any $1 < p < 2$ and any measurable set F_3 with $|F_3| = 1$ in $\mathbb{R}$, there exists a subset F_3' of F_3 with $|F_3'| \ge 1/2$ such that

$$\big|\Lambda_{\mathbf{S}}(f_1,f_2,f_3)\big| \le C|F_1|^{1/p_1}|F_2|^{1/p_2} \tag{4.5}$$

holds for all $p_1 \ge p, p_2 \ge p$ with $1/p_1 + 1/p_2 \ge 1$, $f_1 \in X(F_1), f_2 \in X(F_2), f_3 \in X(F_3')$, where the constant C is independent of $\mathbf{S}$, F_1, F_2, F_3, f_1, f_2, f_3, M_1, M_2. Let us introduce some definitions first.

Definition 4.2. Let $p > 1$. Define the exceptional set Ω by

$$\Omega = \bigcup_{\ell=1}^{3} \big\{x \in \mathbb{R} : M_p\big(M\mathbf{1}_{F_\ell}\big)(x) > C_0|F_\ell|^{1/p}\big\} \tag{4.6}$$

where Mf is the Hardy–Littlewood maximal function of f and $M_p f$ equals to $\big(M(|f|^p)\big)^{1/p}$.

By this definition, for the measurable set F_3 with $|F_3| = 1$, we take $F_3' = F_3\backslash\Omega$. If C_0 is chosen sufficiently large we see that $|F_3'| \geq |F_3|/2$.

Definition 4.3. Given $\mathbf{S} \subset \mathbb{Z}(\gamma) \times \mathbb{Z}$ and $s = (j, n) \in \mathbf{S}$. Let $k_s = \min_{\ell\in\{1,2,3\}}\{k_{j\ell}\}$. The dyadic interval $[2^{-k_s}n, 2^{-k_s}(n+1)]$ is called the time interval of s. We denote it by I_s.

Definition 4.4. Let $\mathbf{S}$ be a subset of $\mathbb{Z}(\gamma) \times \mathbb{Z}$. We say that $\mathbf{S}$ is a convex set in $\mathbb{Z}(\gamma)\times\mathbb{Z}$ if for any $s \in \mathbb{Z}(\gamma)\times\mathbb{Z}$ with $I_{s_1} \subseteq I_s \subseteq I_{s_2}$ for some $s_1, s_2 \in \mathbf{S}$, we have $s \in \mathbf{S}$.

Definition 4.5. Let $\mathbf{T} \subset \mathbf{S}$. If there is $t \in \mathbf{T}$ such that $I_s \subset I_t$ holds for all $s \in T$, then $\mathbf{T}$ is called a tree with top t. $\mathbf{T}$ is called a maximal tree with top t in $\mathbf{S}$ if there does not exist a larger tree in $\mathbf{S}$ with the same top strictly containing $\mathbf{T}$.

Definition 4.6. Let $\mathbf{T}$ be a tree in $\mathbf{S}$. Define $\mathrm{scl}(\mathbf{T})$ the set of scale indices of $\mathbf{T}$ by

$$\mathrm{scl}(\mathbf{T}) = \{j \in \mathbb{Z} : \exists n \in \mathbb{Z}, \text{s. t. } (j, n) \in \mathbf{T}\}.$$

For $j \in \mathrm{scl}(\mathbf{T})$, the j-th shadow of $\mathbf{T}$ is defined by

$$\mathbf{Sh}_j(\mathbf{T}) = \bigcup \big\{I_s : s = (j, n) \in \mathbf{T}\big\}.$$

Define an approximation of $\mathbf{1}_{\mathbf{Sh}_j(\mathbf{T})}$ by

$$\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x) = \mathbf{1}_{\mathbf{Sh}_j(\mathbf{T})} * \psi_{k_j}(x).$$

Definition 4.7. Let $(j, n) = s \in \mathbf{S}$ and $\ell \in \{1, 2, 3\}$. And let

$$\mathbf{1}^{**}_{j,n}(x) = \int_{I_{k_j,n}} \frac{2^{k_j}}{(1 + 2^{2k_j}|x - y|^2)^{200}} dy.$$

Define a seminorm $\|f_\ell\|_{j,n}$ by

$$\big\|f_\ell\big\|_{j,n} = \big\|f_\ell\big\|_s = \frac{1}{|I_s|^{1/p}}\big\|\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}\big\|_p + \frac{1}{|I_s|^{1/p}}\big\|2^{-k_{j\ell}}\mathbf{1}^{**}_{j,n} Df_{\ell,j,n_\ell}\big\|_p \tag{4.7}$$

where Df_{ℓ,j,n_ℓ} is the derivative of f_{ℓ,j,n_ℓ}.

Define $\zeta(j, M, K)$ by

$$\zeta(j, M, K) = \left[\frac{L_1 j + M_1 - M_2 - 6}{L_2}\right] + \left[\frac{L_1}{L_2}\right] M + K, \tag{4.8}$$

where $L = 2^{100}$, K is an integer between $-10L$ and $10L$ and M is an integer between 0 and $6L$. For $\ell \in \{2,3\}$, we define a ζ seminorm $\big\|f_\ell\big\|_{j,n,\zeta}$ by

(4.9)

$$\big\|f_\ell\big\|_{j,n,\zeta} = \|f_\ell\|_{j,n} + \sup_{M,K} \frac{1}{|I_s|^{1/p}} \Big(\big\|\mathbf{1}^{**}_{j,n} f_{\ell,\zeta(j,M,K),0}\big\|_p + \big\||I_s|\mathbf{1}^{**}_{j,n} D f_{\ell,\zeta(j,M,K),0}\big\|_p \Big).$$

For $\ell = 1$, let the ζ seminorm $\big\|f_1\big\|_{j,n,\zeta} = \big\|f_1\big\|_{j,n}$.

Definition 4.8. Let $\mathbf{T} \subset \mathbf{S}$ be a tree and $t = (j_{\mathbf{T}}, n_{\mathbf{T}}) \in \mathbf{T}$ be the top of $\mathbf{T}$. Denote by $I_{\mathbf{T}}$ the time interval of the top of tree $\mathbf{T}$.

(a) In the case $|\omega_{2,j}| \le |\omega_{1,j}|/6$ for all $j \in \mathrm{scl}(\mathbf{T})$, define $\Delta^*_\ell(\mathbf{T})$ for $\ell \in \{1,3\}$ by

$$\Delta^*_\ell(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}} \big|\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}(x)\big|^2 \right)^{1/2}. \tag{4.10}$$

For $\ell = 2$, define

$$\Delta^*_2(\mathbf{T})(x) = \big|\mathbf{1}^{**}_{j_{\mathbf{T}},n_{\mathbf{T}}} f_{2,j_{\mathbf{T}},n_\ell}(x)\big|. \tag{4.11}$$

And in this case, for $\ell \in \{1,2,3\}$, define the ℓ-size of $\mathbf{T}$ by

$$\mathrm{size}_\ell(\mathbf{T}) = \frac{1}{|I_{\mathbf{T}}|^{1/p}} \big\|\Delta^*_\ell(\mathbf{T})\big\|_p + \big\|f_\ell\big\|_{j_{\mathbf{T}},n_{\mathbf{T}}}. \tag{4.12}$$

(b) In the case $|\omega_{2,j}| > |\omega_{1,j}|/6$ for all $j \in \mathrm{scl}(\mathbf{T})$, for $\ell = 2,3$, let $f_{\ell,j,\mathbf{T}} = f_{\ell,j,0}$ if $j \in \mathrm{scl}(\mathbf{T})$ and $f_{\ell,j,\mathbf{T}} \equiv 0$ if $j \notin \mathrm{scl}(\mathbf{T})$. Define:

$$\begin{aligned} \Delta^*_\ell(\mathbf{T}) = &\left(\sum_{(j,n)\in\mathbf{T}} \big|\mathbf{1}^{**}_{j,n}\big(f_{\ell,j,\mathbf{T}} - f_{\ell,j-L,\mathbf{T}}\big)(x)\big|^2 \right)^{1/2} \\ &+ \left(\sum_{(j,n)\in\mathbf{T}} \big|\mathbf{1}^{**}_{j,n}\big(f_{\ell,j,n_\ell} - f_{\ell,j,0}\big)(x)\big|^2 \right)^{1/2}. \end{aligned} \tag{4.13}$$

And define $\Delta^*_1(\mathbf{T})$ by

$$\Delta^*_1(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}} \big|\mathbf{1}^{**}_{j,n} f_{1,j,n_1}(x)\big|^2 \right)^{1/2}. \tag{4.14}$$

In this case, for $\ell \in \{1,2,3\}$, define the ℓ-size of $\mathbf{T}$ by

$$\mathrm{size}_\ell(\mathbf{T}) = \frac{1}{|I_{\mathbf{T}}|^{1/p}} \big\|\Delta^*_\ell(\mathbf{T})\big\|_p + \big\|f_\ell\big\|_{j_{\mathbf{T}},n_{\mathbf{T}},\zeta}. \tag{4.15}$$

Let $\mathbf{P}$ be a subset of $\mathbf{S}$. Define the ℓ-*size** of $\mathbf{P}$ by

$$\text{size}^*_\ell(\mathbf{P}) = \sup_{\mathbf{T}:\mathbf{T}\subset\mathbf{P}} \text{size}_\ell(\mathbf{T}), \tag{4.16}$$

where $\mathbf{T}$ ranges over all trees in $\mathbf{P}$.

In the definition of $\mathbf{1}^{**}_{j,n}$, we can replace the exponent 200 by a larger number 2^{100} to define a new function. We denote this function by $\widetilde{\mathbf{1}}^*_{j,n}$. If $\mathbf{1}^{**}_{j,n}$ is replaced by $\widetilde{\mathbf{1}}^*_{j,n}$ in the definition of $\Delta^*_\ell(\mathbf{T})$, we denote the corresponding function by $\Delta_\ell(\mathbf{T})$.

Definition 4.9. Let $\mathbf{S}$ be a subset of $\mathbb{Z}(\gamma)\times\mathbb{Z}$. Suppose that $\mathbf{S}$ is a union of trees $\mathbf{T}\in\mathcal{F}$. Define count($\mathbf{S}$) by

$$\text{count}(\mathbf{S}) = \sum_{\mathbf{T}\in\mathcal{F}} |I_{\mathbf{T}}|. \tag{4.17}$$

4.2. Reduction. Let $\mathbf{S}$ be a subset of $\mathbb{Z}(\gamma)\times\mathbb{Z}$. For Ω defined in (4.6), we define

$$\mathbf{S}(\Omega) = \{s\in\mathbf{S} : I_s \nsubseteq \Omega\}. \tag{4.18}$$

The following lemma indicates that we only need to seek the upper bound for the trilinear form $\Lambda_{\mathbf{S}(\Omega)}$.

Lemma 4.10. *Let $n_1 = n_2 = 0$ and $f_3 \in X(F_3')$. For all functions $f_1 \in X(F_1)$, $f_2\in X(F_2)$, the following inequality holds.*

$$\begin{aligned}\big|\Lambda_{\mathbf{S}}(f_1,f_2,f_3) - \Lambda_{\mathbf{S}(\Omega)}(f_1,f_2,f_3)\big| \\ \le C\min\big\{1,|F_1|^{1/p}\big\}\min\big\{1,|F_2|^{1/p}\big\},\end{aligned} \tag{4.19}$$

where C is a constant independent of $\mathbf{S}$, F_1, F_2, F_3, f_1, f_2, f_3, M_1, M_2.

Proof. Notice that if $s=(j,n)\in\mathbf{S}(\Omega)^c$, then $I_s\subseteq\Omega$. Let $\mathbf{S}_L(\Omega)$ be defined by

$$\mathbf{S}_L(\Omega) = \{s\in\mathbf{S}(\Omega)^c : 2^L I_s\subseteq\Omega, \text{ but } 2^{L+1}I_s\nsubseteq\Omega\}.$$

We see that $\mathbf{S}(\Omega)^c = \cup_{L=0}^\infty \mathbf{S}_L(\Omega)$. Let $\mathcal{J}_L$ be the set of all time intervals I_s's for $s\in\mathbf{S}_L(\Omega)$. It is easy to see that $\mathcal{J}_L$ is a collection of disjoint intervals and $\sum_{J\in\mathcal{J}_L}|J|\le|\Omega|<1$. Hence, it suffices to show that for any $J\in\mathcal{J}_L$ and any $(j,n)=s\in\mathbf{S}_L(\Omega)$ such that $I_s=J$, we have

$$\begin{aligned}\Big|\int \mathbf{1}^*_{j,n}(x)\prod_\ell f_{\ell,j,n_\ell}(x)dx\Big| \\ \le C2^{-L}\min\big\{1,|F_1|^{1/p}\big\}\min\big\{1,|F_2|^{1/p}\big\}|J|,\end{aligned} \tag{4.20}$$

where C is a constant independent of f_1, f_2, f_3, M_1, M_2, since (4.19) follows by summing all L's and J's together.

We now prove (4.20). Since $F_3' = F_3\backslash\Omega$ and $f_3 \in X(F_3')$, we get for any $(j,n) \in \mathbf{S}$ and any positive integer N,

$$\left|\mathbf{1}^*_{j,n}(x) f_{3,j,n_\ell}(x)\right| \le \frac{C_N}{\left(1+2^{k_j}\operatorname{dist}(x,I_s)\right)^{3N}\left(1+2^{k_{j3}}\operatorname{dist}(x,\Omega^c)\right)^{3N}}. \tag{4.21}$$

Clearly we have for $\ell \in \{1,2\}$ and $(j,n) \in \mathbf{S}$,

$$\left|f_{\ell,j,n_\ell}(x)\right| \le \int \frac{C_N |f_\ell(y)| 2^{k_{j\ell}}\, dy}{\left(1+2^{k_{j\ell}}|x-y|\right)^N}. \tag{4.22}$$

By the definition of Ω, we have for $\ell \in \{1,2\}$ and $(j,n) \in \mathbf{S}$,

$$\left|f_{\ell,j,n_\ell}(x)\right| \le C_N \min\left\{1, |F_\ell|^{1/p}\right\}\left(1+2^{k_{j\ell}}\operatorname{dist}(x,\Omega^c)\right)^2. \tag{4.23}$$

Thus (4.21), (4.23) and the fact $2^{k_{j3}} \sim 2^{\max\{k_{j\ell}\}}$ yield that the left-hand side of (4.20) is no more than

$$C_N 2^{-LN} \prod_{\ell=1}^{2} \min\left\{1, |F_\ell|^{1/p}\right\} |J|$$

for any positive integer $N \ge 2$, which is the desired estimate. □

Hence, to prove (4.5), we only need to prove the following lemma for $\Lambda_{\mathbf{S}(\Omega)}$. The details of the proof of Lemma 4.11 will be given in the next few subsections.

Lemma 4.11. *Let $n_1 = n_2 = 0$, $1 < p < 2$, $F_3 \subset \mathbb{R}$, and $\mathbf{S}(\Omega)$ be the set defined in* (4.18) *and $F_3' = F\backslash\Omega$. For all $p_1, p_2 \ge p$ with $1/p_1 + 1/p_2 \ge 1$, and all functions $f_1 \in X(F_1)$, $f_2 \in X(F_2)$, $f_3 \in X(F_3')$, the following inequality holds.*

$$\left|\Lambda_{\mathbf{S}(\Omega)}(f_1,f_2,f_3)\right| \le C|F_1|^{1/p_1}|F_2|^{1/p_2}, \tag{4.24}$$

where C is a constant independent of $\mathbf{S}$, F_1, F_2, F_3, f_1, f_2, f_3, M_1, M_2.

4.3. Principal lemmas. We now state some lemmata which will be used in proof of Lemma 4.11.

Lemma 4.12. *Let $1 < q < \infty$, $\ell \in \{1,2,3\}$ and $\mathbf{T}$ be a tree in $\mathbf{S}$. Then*

$$\left\|\Delta^*_\ell(\mathbf{T})\right\|_q \le C \inf_{x\in I_{\mathbf{T}}} M_q(Mf_\ell)(x)|I_{\mathbf{T}}|^{1/q}, \tag{4.25}$$

$$\operatorname{size}_\ell(\mathbf{T}) \le C \inf_{x\in I_{\mathbf{T}}} M_p(Mf_\ell)(x), \tag{4.26}$$

where C is a constant independent of $f_\ell, \mathbf{T}$, $\mathbf{S}$, M_1, M_2.

Proof. (4.25) is a consequence of the following L^q estimates of $\Delta_\ell(\mathbf{T})$.

$$\left\|\Delta^*_\ell(\mathbf{T})\right\|_q \le C\|f_\ell\|_q. \tag{4.27}$$

In fact, one can decompose f_ℓ into $f_\ell \mathbf{1}_{2I_{\mathbf{T}}}$ and $f_\ell \mathbf{1}_{(2I_{\mathbf{T}})^c}$. For the first function, apply (4.27) to get the desired estimates. For the second function,

the desired estimates follow by the fast decay due to $\Delta_\ell^*(\mathbf{T})$ is essentially supported on $I_{\mathbf{T}}$.

Note that we consider only the case $n_\ell = 0$. For $n_\ell \neq 0$, the following argument still works if one changes the constant C to $C(1+|n_\ell|)^5$. We only give the details for the case $|\omega_{2,j}| \leq |\omega_{1,j}|/2$ and $\ell \in \{1,3\}$ since other cases can be done in the same way. In this case, we have

$$\Delta_\ell^*(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}} \left|\mathbf{1}_{j,n}^{**} f_{\ell,j,0}(x)\right|^2 \right)^{1/2}.$$

Notice that $\Delta_\ell^*(\mathbf{T})(x)$ is dominated by

$$\left(\sum_{j\in\mathbb{Z}} \left|f_{\ell,j,0}(x)\right|^2 \right)^{1/2},$$

where $f_{\ell,j,0}$ is defined by $\widehat{f_{\ell,j,0}} = \widehat{f_\ell}\widehat{\Phi}_{\ell,j,0}$. Note that $\widehat{\Phi}_{\ell,j,0}$ is supported on $\omega_{\ell,j}$ and $\omega_{\ell,j}$'s are disjoint. Thus the Littlewood–Paley theorem then yields the L^q estimates (4.27). To get (4.25), it suffices to show that

$$\left\|\Delta_{\ell,\mathrm{out}}^*(\mathbf{T})\right\|_q \leq C \inf_{x\in I_{\mathbf{T}}} M_q(Mf_\ell)(x)|I_{\mathbf{T}}|^{1/q},$$

where $\Delta_{\ell,\mathrm{out}}^*(\mathbf{T})$ is defined by

$$\Delta_{\ell,\mathrm{out}}^*(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}} \left|\mathbf{1}_{j,n}^{**}(x)\big((f\mathbf{1}_{(2I_{\mathbf{T}})^c}) * \Phi_{\ell,j,0}\big)(x)\right|^2 \right)^{1/2}.$$

By the definition of $\mathbf{1}_{j,n}^{**}$ and $\Phi_{\ell,j,0}$, we have that for any positive integer N,

$$\begin{aligned}&\left|\mathbf{1}_{j,n}^{**}(x)\big((f\mathbf{1}_{(2I_{\mathbf{T}})^c}) * \Phi_{\ell,j,0}\big)(x)\right| \\ &\qquad\leq \frac{C_N}{\big(1+2^{k_j}\mathrm{dist}(x,I_s)\big)^{100}} \int_{(2I_{\mathbf{T}})^c} \frac{|f_\ell(y)|2^{k_{j\ell}}}{\big(1+2^{k_{j\ell}}|x-y|\big)^N} dy.\end{aligned}$$

which is clearly dominated by

$$\frac{CMf_\ell(x)}{\big(1+2^{k_j}\mathrm{dist}(x,I_s)\big)^{50}\big(1+2^{k_j}\mathrm{dist}(I_s,(2I_{\mathbf{T}})^c)\big)^{50}}.$$

Thus for $s \in \mathbf{T}$,

$$\begin{aligned}&\left\|\mathbf{1}_{j,n}^{**}\big((f\mathbf{1}_{(2I_{\mathbf{T}})^c}) * \Phi_{\ell,j,0}\big)\right\|_q^q \\ &\qquad\leq \frac{C|I_s|}{\big(1+2^{k_j}\mathrm{dist}(I_s,(2I_{\mathbf{T}})^c)\big)^{25q}} \Big(\inf_{x\in I_{\mathbf{T}}} M_q(Mf_\ell)(x)\Big)^q.\end{aligned}$$

By the triangle inequality, we obtain that

$$\left\|\Delta_{\ell,\mathrm{out}}^*(\mathbf{T})\right\|_q \leq \sum_{s\in\mathbf{T}} \frac{C|I_s|^{1/q}}{\big(1+|I_s|^{-1}\mathrm{dist}(I_s,(2I_{\mathbf{T}})^c)\big)^{25}} \inf_{x\in I_{\mathbf{T}}} M_q(Mf_\ell)(x),$$

which yields the desired estimate (4.25). Notice that

$$\big\|\mathbf{1}^{**}_{j_{\mathbf{T}},n_{\mathbf{T}}}f_{\ell,j_{\mathbf{T}},n_\ell}\big\|_p+\big\|2^{-k_{j_{\mathbf{T}}\ell}}\mathbf{1}^{**}_{j_{\mathbf{T}},n_{\mathbf{T}}}Df_{\ell,j_{\mathbf{T}},n_\ell}\big\|_p\le\Big\|\frac{CMf_\ell(\cdot)}{\big(1+|I_{\mathbf{T}}|^{-1}\mathrm{dist}(\cdot,I_{\mathbf{T}})\big)^N}\Big\|_p,$$

which is clearly dominated by $\inf_{x\in I_{\mathbf{T}}}M_p(Mf_\ell)(x)|I_{\mathbf{T}}|^{1/p}$. Therefore we obtain (4.26). □

Lemma 4.13. *Suppose that $s=(j,n)\in\mathbf{S}$.*
If $2^{k_{j\ell}}\sim 2^{k_j}$, then

$$\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_\infty\le C\big\|f_\ell\big\|_{j,n} \tag{4.28}$$

holds for $\ell\in\{1,2,3\}$, where C is a constant independent of s,f_ℓ,n_ℓ.
If $2^{k_{j1}}\sim 2^{k_j}$, then

$$\big\|\mathbf{1}^{**}_{j,n}f_{\ell,\zeta(j,M,K),n_\ell}\big\|_\infty\le C\big\|f_\ell\big\|_{j,n,\zeta} \tag{4.29}$$

holds for $\ell\in\{2,3\}$, where $\zeta(j,M,K)$ is defined in Definition 4.7 *and C is a constant independent of $s,f_\ell,n_\ell,\zeta,M,K$.*

Proof. We only prove (4.28) since (4.29) essentially is a consequence of (4.28). Let $\mu=\big\|f_\ell\big\|_{j,n}$. By the definition of the seminorm, we have

$$\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_p+\big\||I_s|\mathbf{1}^{**}_{j,n}Df_{\ell,j,n_\ell}\big\|_p\le\mu|I_s|^{1/p}. \tag{4.30}$$

First we prove the BMO estimate for the function, that is

$$\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_{\mathrm{BMO}}\le C\mu. \tag{4.31}$$

If $|I_s|\le|J|$, by (4.30) we have

$$\inf_c\int_J\big|\mathbf{1}^{**}_{j,n}(x)f_{\ell,j,n_\ell}(x)-c\big|dx\le\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_p|J|^{1-\frac1p}\le\mu|I_s|^{\frac1p}|J|^{1-\frac1p}\le\mu|J|.$$

If $|I_s|\ge|J|$, by (4.30) we obtain that

$$\begin{aligned}
&\inf_c\int_J\big|\mathbf{1}^{**}_{j,n}(x)f_{\ell,j,n_\ell}(x)-c\big|dx\\
&\le|J|\int_J\big|\big(\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big)'(x)\big|dx\\
&\le|J|\int_J\big|\big(\mathbf{1}^{**}_{j,n}\big)'(x)f_{\ell,j,n_\ell}(x)\big|dx+|J|\int_J\big|\mathbf{1}^{**}_{j,n}(x)Df_{\ell,j,n_\ell}(x)\big|dx\\
&\le C|J||I_s|^{-1}\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_p|J|^{1-\frac1p}+|J|\big\|\mathbf{1}^{**}_{j,n}Df_{\ell,j,n_\ell}\big\|_p|J|^{1-\frac1p}\\
&\le C\mu|J|^{2-\frac1p}|I_s|^{\frac1p-1}\ \le\ C\mu|J|.
\end{aligned}$$

Thus we get the BMO estimate (4.31). Interpolating (4.31) and (4.30), we have for any $p\le q<\infty$,

$$\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_q\le C\mu|I_s|^{1/q}.$$

Notice that an integration by parts and Hölder inequality yield that

$$\big\|\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}\big\|_\infty \le \big\|\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}\big\|_{p'}^{1/2} \big\|\big(\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}\big)'\big\|_p^{1/2},$$

where $1/p + 1/p' = 1$. Hence the desired estimate (4.28) follows by (4.30) and $L^{p'}$ estimates for the functions. □

Lemma 4.14. *Suppose that $2^{k_{j\ell}} \sim 2^{k_j}$ holds for all $(j,n) \in \mathbf{S}$. Then for any tree $\mathbf{T}$ in $\mathbf{S}$, we have*

$$\big\|\Delta_\ell(\mathbf{T})\big\|_{\mathrm{BMO}} \le C \mathrm{size}^*_\ell(\mathbf{T}), \tag{4.32}$$

where C is a constant independent of $\mathbf{T}, \mathbf{S}, L_1, L_2, M_1, M_2, f_\ell, n_\ell$.

Proof. We only give the a proof for $\ell = 1$. Other cases can be handled in the same way. Let $\mu = \mathrm{size}^*_\ell(\mathbf{S})$. Let J be a dyadic interval and $\mathbf{T}_J = \{s \in \mathbf{T} : I_s \subseteq J\}$. We then dominate $\inf_c \int_J \big|\Delta_\ell(\mathbf{T})(x) - c\big| dx$ by a sum of the following three parts.

$$\int_J \left(\sum_{s\in\mathbf{T}_J} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 \right)^{1/2} dx,$$

$$\int_J \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 \right)^{1/2} dx,$$

and

$$\inf_c \int_J \left| \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|>|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 \right)^{1/2} - c \right| dx.$$

The first part is clearly dominated by $\mu|J|$ because of the Hölder inequality and the fact that μ is the ℓ-size* of $\mathbf{S}$.

Since $p \le 2$ we estimate the second part by

$$\left\| \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n} f_{\ell,j,n_\ell}\big|^2 \right)^{1/2} \right\|_{L^p(J)} |J|^{1-\frac{1}{p}}$$

$$\le \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le|J|}} \big\|\widetilde{\mathbf{1}}^*_{j,n} f_{\ell,j,n_\ell}\big\|^p_{L^p(J)} \right)^{1/p} |J|^{1-\frac{1}{p}}$$

$$\leq \left(\sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|\leq |J|}} \frac{C\big\|\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}\big\|_p^p}{\big(1+|I_s|^{-1}\mathrm{dist}(J,I_s)\big)^{100}} \right)^{1/p} |J|^{1-\frac{1}{p}}$$

$$\leq \mu \left(\sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|\leq |J|}} \frac{C|I_s|}{\big(1+|I_s|^{-1}\mathrm{dist}(J,I_s)\big)^{100}} \right)^{1/p} |J|^{1-\frac{1}{p}} \ \leq \ C\mu|J|.$$

The third part is estimated by

$$\left(\inf_c \int_J \left| \left(\sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|>|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 \right)^{1/2} - c \right|^2 dx \right)^{1/2} |J|^{1/2}$$

$$\leq \left(\inf_c \int_J \left| \sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|>|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 - c \right| dx \right)^{1/2} |J|^{1/2}$$

$$\leq C \left(\int_J \sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|>|J|}} \left| \Big(\big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 \Big)' \right| dx \right)^{1/2} |J|,$$

which is dominated by a sum of the following two terms,

$$R_1 = C \left(\int_J \sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|>|J|}} |I_s|^{-1} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big|^2 dx \right)^{1/2} |J|,$$

and

$$R_2 = C \left(\int_J \sum_{\substack{s\in \mathbf{T}\setminus \mathbf{T}_J \\ |I_s|>|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}(x) f_{\ell,j,n_\ell}(x)\big| \big|\widetilde{\mathbf{1}}^*_{j,n}(x) D f_{\ell,j,n_\ell}(x)\big| dx \right)^{1/2} |J|,$$

By Lemma 4.13, we see that for any $q \geq p$,

$$\big\|\mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}\big\|_q \leq C\mu |I_{\mathbf{T}}|^{1/q}.$$

Thus, by Hölder's inequality, the first term R_1 is estimated by

$$C\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\frac{|I_s|^{-1}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,I_s)\big)^{100}}\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_4^2|J|^{1/2}\right)^{1/2}\quad|J|$$

$$\le C\mu\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\frac{|I_s|^{-1/2}|J|^{1/2}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,I_s)\big)^{100}}\right)^{1/2}\quad|J|\ \le\ C\mu|J|,$$

and the second term R_2 is estimated by

$$C\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\big\|\widetilde{\mathbf{1}}^{*}_{j,n}f_{\ell,j,n_\ell}\big\|_{L^{p'}(J)}\big\|\mathbf{1}^{**}_{j,n}Df_{\ell,j,n_\ell}\big\|_p\right)^{1/2}\quad|J|$$

$$\le C\left(\mu\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\frac{|I_s|^{\frac{1}{p}-1}\big\|\mathbf{1}^{**}_{j,n}f_{\ell,j,n_\ell}\big\|_{p'+1}|J|^{\frac{1}{p'(p'+1)}}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,I_s)\big)^{100}}\right)^{1/2}\quad|J|$$

$$\le C\mu\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\frac{|I_s|^{-\frac{1}{p'(p'+1)}}|J|^{\frac{1}{p'(p'+1)}}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,I_s)\big)^{100}}\right)^{1/2}\quad|J|\ \le\ C\mu|J|.$$

This completes the proof of (4.32). □

The principal lemma is the following organization lemma.

Lemma 4.15. *Let $\ell\in\{1,2,3\}$ and $\mathbf{S}$ be a subset of $\mathbb{Z}(\gamma)\times\mathbb{Z}$. $\mathbf{S}$ can be partitioned to two parts $\mathbf{S}_1$ and $\mathbf{S}_2$ such that $\mathbf{S}_1$ is a union of maximal trees with*

$$\mathrm{count}(\mathbf{S}_1)\le C\big(\mathrm{size}^*_\ell(\mathbf{S})\big)^{-p}|F_\ell|, \tag{4.33}$$

and

$$\mathrm{size}^*_\ell(\mathbf{S}_2)\le\frac{1}{2}\mathrm{size}^*_\ell(\mathbf{S}), \tag{4.34}$$

where C is a constant independent of $\mathbf{S}, M_1, M_2, f_\ell, F_\ell$.

Proof. Let $\mathcal{F}_0$ be the set of all trees $\mathbf{T}\subset\mathbf{S}$ such that $\mathrm{size}_\ell(\mathbf{T})>\mathrm{size}^*_\ell(\mathbf{S})/2$. Recall that $I_{\mathbf{T}}$ is the time interval for the top of $\mathbf{T}$. Let $\mathcal{I}$ denote the collection of all possible $I_{\mathbf{T}}$'s for trees $\mathbf{T}\in\mathcal{F}_0$. Initially, set $\mathbf{S}_1:=\emptyset$, $\mathcal{I}_{\mathrm{stock}}:=\mathcal{I}$, and $\mathbf{S}_{\mathrm{stock}}:=\mathbf{S}$. Take a longest interval J in $\mathcal{I}_{\mathrm{stock}}$. By the defintion of $\mathcal{I}$, there must be a tree $\mathbf{T}\in\mathcal{F}_0$ whose top is J. Let $\widetilde{\mathbf{T}}$ be the maximal tree

in $\mathbf{S}_{\text{stock}}$ with the top J. Obviously $\text{size}_\ell(\widetilde{\mathbf{T}}) \geq \text{size}^*_\ell(\mathbf{S})/2$. We remove this maximal tree from $\mathbf{S}_{\text{stock}}$. Update $\mathbf{S}_{\text{stock}} := \mathbf{S}_{\text{stock}} \backslash \widetilde{\mathbf{T}}$, $\mathbf{S}_1 := \mathbf{S}_1 \cup \widetilde{\mathbf{T}}$, and

$$\mathcal{I}_{\text{stock}} := \mathcal{I}_{\text{stock}} \backslash \{I \in \mathcal{I}_{\text{stock}} : I \subseteq J\}.$$

Repeat this procedure until $\mathcal{I}_{\text{stock}} = \emptyset$. Clearly when this process terminates, $\mathbf{S}_1$ is a union of a trees $\widetilde{\mathbf{T}}$'s and $I_{\widetilde{\mathbf{T}}}$'s are disjoint due to the maximality of trees. By (4.26) and the size condition on $\widetilde{\mathbf{T}}$, we have

$$\inf_{x \in I_{\widetilde{\mathbf{T}}}} M_p(Mf_\ell)(x) \geq \text{size}^*_\ell(\mathbf{S})/2,$$

which implies that

$$\bigcup_{\widetilde{\mathbf{T}}} I_{\widetilde{\mathbf{T}}} \subseteq \big\{x \in \mathbb{R} : M_p(Mf_\ell)(x) \geq \text{size}^*_\ell(\mathbf{S})/2\big\}.$$

Thus the disjointness property of $I_{\widetilde{\mathbf{T}}}$'s and (weak) L^q estimates for $1 \leq q \leq \infty$ of Hardy–Littlewood maximal functions yield (4.33). Let $\mathbf{S}_2 = \mathbf{S} \backslash \mathbf{S}_1$. Clearly $\mathbf{S}_2$ satisfies (4.34). Therefore we complete the proof of Lemma 4.15. □

4.4. The size estimate for a tree. Let $\mathbf{S}$ be a convex subset of $\mathbb{Z}(\gamma) \times \mathbb{Z}$. By the definition of $\mathbf{S}(\Omega)$ in (4.18), it is clear that $\mathbf{S}(\Omega)$ is convex. Partition $\mathbf{S}(\Omega)$ into two subsets $\mathbf{S}^{(1)}(\Omega)$ and $\mathbf{S}^{(2)}(\Omega)$, where

$$\mathbf{S}^{(1)}(\Omega) = \big\{(j,n) \in \mathbf{S}(\Omega) : |\omega_{2,j}| \leq |\omega_{1,j}|/6\big\} \tag{4.35}$$

$$\mathbf{S}^{(2)}(\Omega) = \big\{(j,n) \in \mathbf{S}(\Omega) : |\omega_{2,j}| > |\omega_{1,j}|/6\big\}. \tag{4.36}$$

For any $(j,n) \in \mathbf{S}^{(1)}(\Omega)$, $k_{j2} = k_j$ by the definition of k_j. And for any $(j,n) \in \mathbf{S}^{(2)}(\Omega)$, $2^{k_{j1}} \sim 2^{k_j}$.

Lemma 4.16. *For $\kappa \in \{1,2\}$, $\mathbf{S}^{(\kappa)}(\Omega)$ is convex.*

Proof. We only prove the lemma for $\kappa = 2$. One can prove the lemma for $\kappa = 1$ similarly. Let $s_1 = (j_1, n_1), s_2 = (j_2, n_2)$ in $\mathbf{S}^{(2)}(\Omega)$. And $s = (j,n) \in \mathbb{Z}(\gamma) \times \mathbb{Z}$ such that $I_{s_2} \subseteq I_s \subseteq I_{s_1}$. By the convexity of $\mathbf{S}(\Omega)$ we get $s \in \mathbf{S}(\Omega)$. In order to get $s \in \mathbf{S}^{(2)}(\Omega)$, we need to show that $|\omega_{2,j}| > |\omega_{1,j}|/6$. The simple case is the case $2^{k_j} = |\omega_{1,j}|$. In this case,

$$|\omega_{1,j_2}|/10 \leq |\omega_{1,j}| \leq 10|\omega_{1,j_1}|,$$

which implies $j_2 \leq j \leq j_1$. Since $|\omega_{2,j_1}| > |\omega_{1,j_1}|/6$ and $|\omega_{2,j_2}| > |\omega_{1,j_2}|/6$, the linearity of the function $f(j) = (L_1 j + M_1) - (L_2 j + M_2)$ yields that $|\omega_{2,j}| > |\omega_{1,j}|/6$.

We now turn to another case $2^{k_j} = |\omega_{2,j}|$. Since I_s is nested between I_{s_1} and I_{s_2}, we get $|\omega_{1,j_2}|/10 \leq |\omega_{2,j}| \leq 10|\omega_{1,j_1}|$. The first half part of this inequality and the definition of k_j imply $j_2 \leq j$. And the second half part of the inequality and the fact $(j_1, n_1) \in \mathbf{S}^{(2)}(\Omega)$ yield $j \leq j_1$. Thus we get

$|\omega_{2,j}| > |\omega_{1,j}|/6$ by the linearity of the function $f(j)$. Hence s must be in $\mathbf{S}^{(2)}(\Omega)$ in either case. This proves the lemma. □

Lemma 4.17. *Let $\kappa \in \{1,2\}$, $\mathbf{T}$ be a convex tree in $\mathbf{S}^{(\kappa)}(\Omega)$ with the top $t = (j_{\mathbf{T}}, n_{\mathbf{T}})$ and $\partial\mathbf{Sh}_j(\mathbf{T})$ be the boundary of the j-th shadow of $\mathbf{T}$. Let $\mathrm{Card}(\partial\mathbf{Sh}_j(\mathbf{T}))$ denote the cardinality of the boundary of the j-th shadow. Then*

$$\sum_{j\geq j_{\mathbf{T}}} 2^{-k_j}\mathrm{Card}(\partial\mathbf{Sh}_j(\mathbf{T})) \leq C|I_{\mathbf{T}}|, \tag{4.37}$$

where C is a constant independent of $\mathbf{T}$.

Proof. This lemma is similar to one technical lemma (Lemma 4.8) in [17]. We give a similar proof. Note that the j-th shadow consists of finite disjoint intervals and its boundary thus contains all endpoints of the intervals. It is sufficient to consider only all left endpoints since the right endpoints can be handled in the same way. Let $\partial_{\mathrm{left}}(\mathbf{Sh}_j(\mathbf{T}))$ denote the collection of all left endpoints of the intervals in the j-th shadow. Let $z \in \partial_{\mathrm{left}}(\mathbf{Sh}_j(\mathbf{T}))$ and $I_j(z) = (z - 2^{-k_j}, z - 2^{-k_j}/2)$. To prove (4.37), it suffices to show that the intervals $I_j(z)$'s are disjoint for all possible j, z. Assume that there are $j, j' \in \mathrm{scl}(\mathbf{T})$, $z \in \partial_{\mathrm{left}}(\mathbf{Sh}_j(\mathbf{T}))$ and $z' \in \partial_{\mathrm{left}}(\mathbf{Sh}_{j'}(\mathbf{T}))$ such that $(j,z) \neq (j',z')$ and $I_j(z) \cap I_{j'}(z') \neq \emptyset$. By the nesting property of dyadic intervals and the fact that $z - 2^{-k_j}$ is an endpoint of some dyadic intervals, we see that $j \neq j'$. Without loss of generality, suppose that $j < j'$. The fact that $I_j(z)$ and $I_{j'}(z')$ have nonempty intersection then implies $z' \in (z - 2^{-k_j}, z)$. Since z is a left endpoint of some intervals in the j-th shadow, z' can not be in $\mathbf{Sh}_j(\mathbf{T})$. However, the convexity of $\mathbf{T}$ yields that $\mathbf{Sh}_{j'}(\mathbf{T}) \subseteq \mathbf{Sh}_j(\mathbf{T})$. This is a contradiction. Therefore we obtain the lemma. □

Lemma 4.18. *Let $\kappa \in \{1,2\}$, let $\mathbf{T}$ be a convex tree in $\mathbf{S}^{(\kappa)}(\Omega)$ and let $\widetilde{\Lambda}_{\mathbf{T}}(f_1,f_2,f_3)$ be defined by*

$$\widetilde{\Lambda}_{\mathbf{T}}(f_1,f_2,f_3) = \sum_j \int \prod_{\ell=1}^{3} \sum_{n\in\mathbf{T}_j} F_{\ell,j,n}(x)dx, \tag{4.38}$$

where $\mathbf{T}_j = \{n \in \mathbb{Z} : (j,n) \in \mathbf{T}\}$ and $F_{\ell,j,n}$ is defined by

$$F_{\ell,j,n}(x) = \mathbf{1}^*_{j,n}(x) f_{\ell,j,n_\ell}(x). \tag{4.39}$$

Then we have

$$\big|\Lambda_{\mathbf{T}}(f_1,f_2,f_3) - \widetilde{\Lambda}_{\mathbf{T}}(f_1,f_2,f_3)\big| \leq C\mathrm{size}_1^*(\mathbf{T})\mathrm{size}_2^*(\mathbf{T})|I_{\mathbf{T}}|, \tag{4.40}$$

where C is a constant independent of $\mathbf{T}, \mathbf{S}, f_1, f_2, f_3$.

Proof. Observe that the difference $|\Lambda_{\mathbf{T}} - \widetilde{\Lambda}_{\mathbf{T}}|$ by

$$\sum_{j\in\mathrm{scl}(\mathbf{T})} \int \Big|\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x) - \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}\big)^3(x)\Big| \prod_{\ell=1}^{3} \big|f_{\ell,j,n_\ell}(x)\big|dx,$$

which is dominated by

$$\sum_{j\in \mathrm{scl}(\mathbf{T})}\sum_{I:|I|=2^{-k_j}} \int_I \left|\left(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x)-\left(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}\right)^3(x)\right)\left(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_j(\mathbf{T})}(x)\right)^{-\frac{1}{10}}\right|\Pi_{j,\mathbf{T}}(f_1,f_2,f_3)(x)dx,$$

where

$$\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_j(\mathbf{T})}(x)=\int_{\mathbf{Sh}_j(\mathbf{T})}\frac{2^{k_j}}{\left(1+2^{2k_j}|x-y|^2\right)^{2^{1000}}}dx \tag{4.41}$$

and

$$\Pi_{j,\mathbf{T}}(f_1,f_2,f_3)(x)=\prod_{\ell=1}^3\left|\left(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_j(\mathbf{T})}\right)^{1/30}f_{\ell,j,n_\ell}(x)\right|.$$

Hölder's inequality, Lemma 4.13 and (4.25) then yield that

$$\left\|\Pi_{j,\mathbf{T}}(f_1,f_2,f_3)\right\|_{L^1(I)}\le C\mathrm{size}^*_1(\mathbf{T})\mathrm{size}^*_2(\mathbf{T})2^{-k_j}. \tag{4.42}$$

Thus we estimate the difference $|\Lambda_{\mathbf{T}}-\widetilde{\Lambda}_{\mathbf{T}}|$ by

$$C\mathrm{size}^*_1(\mathbf{T})\mathrm{size}^*_2(\mathbf{T})\cdot\sum_{j\in\mathrm{scl}(\mathbf{T})}\sum_{I:|I|=2^{-k_j}}|I|\left\|\left(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}-\left(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}\right)^3\right)\left(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_j(\mathbf{T})}\right)^{-1/10}\right\|_{L^\infty(I)},$$

By the definition of $\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}$, it is easy to see that it is a smooth approximation of $\mathbf{1}_{\mathbf{Sh}_j(\mathbf{T})}$ and for any positive interger N the following inequality holds.

$$|I|\left\|\left(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}-\left(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}\right)^3\right)\left(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_j(\mathbf{T})}\right)^{-1/10}\right\|_{L^\infty(I)}\le\frac{C_N|I|}{\left(1+|I|^{-1}\mathrm{dist}(I,\partial\mathbf{Sh}_j(\mathbf{T}))\right)^N}.$$

Summing up all I's with $|I|=2^{-k_j}$, we estimate the difference by

$$C\mathrm{size}^*_1(\mathbf{T})\mathrm{size}^*_2(\mathbf{T})\sum_{j\in\mathrm{scl}(\mathbf{T})}2^{-k_j}\mathrm{Card}(\partial\mathbf{Sh}_j(\mathbf{T})).$$

Hence the lemma follows by Lemma 4.17. □

Lemma 4.19. *Let* $\mathbf{T}$ *be a convex tree in* $\mathbf{S}^{(2)}(\Omega)$. *For* $\ell\in\{2,3\}$, *let* $F_{\ell,j}$ *be defined by*

$$F_{\ell,j}(x)=\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x)f_{\ell,j,0}(x), \tag{4.43}$$

if $\mathbf{T}_j \neq \emptyset$, *and* $F_{\ell,j} \equiv 0$ *if* $\mathbf{T}_j = \emptyset$. *Then we have*

$$\sup_M \Big\| \Big(\sum_j \big| F_{\ell,j-M} - F_{\ell,j-M-L} \big|^2 \Big)^{1/2} \Big\|_p \le C \mathrm{size}^*_\ell(\mathbf{T}) |I_{\mathbf{T}}|^{1/p}, \tag{4.44}$$

where $L = 2^{100}$, M *ranges over all integers between* 0 *and* $6L$ *and* C *is a constant independent of* $f_\ell, \mathbf{T}$.

Proof. For simplicity, we only prove the lemma for $M = 0$. It is easy to see that $|F_{\ell,j} - F_{\ell,j-L}(x)|$ is dominated by

$$\big| \mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x) \big(f_{\ell,j,0}(x) - f_{\ell,j-L,0}(x) \big) \big| + \big| \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x) - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}(x) \big) f_{\ell,j-L}(x) \big|.$$

Clearly, by the definition of $\Delta^*_\ell(\mathbf{T})$ and $\mathrm{size}^*_\ell(\mathbf{T})$, we get

$$\Big\| \Big(\sum_j \big| \mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} \big(f_{\ell,j,0} - f_{\ell,j-L,0} \big) \big|^2 \Big)^{1/2} \Big\|_p \le C \big\| \Delta^*_\ell(\mathbf{T}) \big\|_p \le C \mathrm{size}^*(\mathbf{T}) |I_{\mathbf{T}}|^{1/p}.$$

Thus to obtain (4.44), it suffices to show that

$$\Big\| \Big(\sum_j \big| \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})} \big) f_{\ell,j-L,0} \big|^2 \Big)^{1/2} \Big\|_p \le C \mathrm{size}^*(\mathbf{T}) |I_{\mathbf{T}}|^{1/p}. \tag{4.45}$$

Heuristically one can consider $\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}$ as $\mathbf{1}_{\mathbf{Sh}_j(\mathbf{T})}$. Then by the nesting property of the j-th shadows due to the convexity of the tree, we see that $\mathbf{Sh}_{j-L}(\mathbf{T}) \backslash \mathbf{Sh}_j(\mathbf{T})$'s are disjoint and this is the reason why we have such an estimate.

Now we go to the technical details. Since $p \le 2$, we estimate the left-hand side of (4.45) by

$$\Big(\sum_{j \in \mathrm{scl}(\mathbf{T})} \big\| \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})} \big) f_{\ell,j-L,0} \big\|_p^p \Big)^{1/p}.$$

This is dominated by

$$\Big(\sum_{j \in \mathrm{scl}(\mathbf{T})} \sum_{I : |I| = 2^{-k_j}} \int_I \Big| \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})} \big)(x) (\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}(x))^{-\frac{1}{10}} \Pi^*_j(f_\ell)(x) \Big|^p dx \Big)^{1/p},$$

where $\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_j(\mathbf{T})}$ is the function defined in (4.41) and

$$\Pi^*_j(f_\ell) = (\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})})^{1/10} f_{\ell,j-L,0}.$$

Hölder's inequality, Lemma 4.13 and (4.25) then yield that

$$\big\| \Pi^*_j(f_\ell) \big\|_{L^p(I)} \le C \mathrm{size}^*_1(\mathbf{T}) \mathrm{size}^*_2(\mathbf{T}) |I|^{1/p}. \tag{4.46}$$

Thus we dominate the left-hand side of (4.45) by

$$C\mathrm{size}_1^*(\mathbf{T})\mathrm{size}_2^*(\mathbf{T}) \cdot \left(\sum_{j\in \mathrm{scl}(\mathbf{T})} \sum_{I:|I|=2^{-k_j}} \left\| \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}\big)\big(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}\big)^{-\frac{1}{10}} \right\|_{L^\infty(I)} |I| \right)^{1/p}$$

Since $\mathbf{Sh}_j(\mathbf{T}) \subset \mathbf{Sh}_{j-L}(\mathbf{T})$, it is easy to see that

$$\big|\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})}(x) - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}(x)\big| \le C\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}(x).$$

On the other hand, observe that $|\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}|$ is dominated by

$$\mathbf{dSh}^*_j(x) = \mathbf{1}_{\mathbf{Sh}_{j-L}(\mathbf{T})\backslash\mathbf{Sh}_j(\mathbf{T})} * \psi_{k_{j-L}}(x) + \frac{C_N}{\big(1+2^{k_j}\mathrm{dist}(x,\partial(\mathbf{Sh}_j(\mathbf{T})))\big)^N},$$

for any positive integer N. Hence the $L^\infty(I)$ norm of

$$\big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}\big)\big(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}\big)^{-\frac{1}{10}}$$

is estimated by

$$\frac{C_N}{\big(1+|I|^{-1}\mathrm{dist}(I,\mathbf{Sh}_{j-L}(\mathbf{T})\backslash\mathbf{Sh}_j(\mathbf{T}))\big)^N} + \frac{C_N}{\big(1+|I|^{-1}\mathrm{dist}(I,\partial(\mathbf{Sh}_j(\mathbf{T})))\big)^N}.$$

For those I's contained in $\mathbf{Sh}_j(\mathbf{T})$, we have

$$\frac{1}{\big(1+|I|^{-1}\mathrm{dist}(I,\mathbf{Sh}_{j-L}(\mathbf{T})\backslash\mathbf{Sh}_j(\mathbf{T}))\big)^N} \le \frac{1}{\big(1+|I|^{-1}\mathrm{dist}(I,\partial(\mathbf{Sh}_j(\mathbf{T})))\big)^N}.$$

For those I's contained in $(\mathbf{Sh}_{j-L}(\mathbf{T}))^c$, we get

$$\frac{1}{\big(1+|I|^{-1}\mathrm{dist}(I,\mathbf{Sh}_{j-L}(\mathbf{T})\backslash\mathbf{Sh}_j(\mathbf{T}))\big)^N} \le \frac{1}{\big(1+|I|^{-1}\mathrm{dist}(I,\partial(\mathbf{Sh}_{j-L}(\mathbf{T})))\big)^N}.$$

Thus we have

$$\sum_{I:|I|=2^{-k_j}} \frac{1}{\big(1+|I|^{-1}\mathrm{dist}(I,\mathbf{Sh}_{j-L}(\mathbf{T})\backslash\mathbf{Sh}_j(\mathbf{T}))\big)^N} \le |I|^{-1}\big|\mathbf{Sh}_{j-L}(\mathbf{T})\backslash\mathbf{Sh}_j(\mathbf{T})\big| + \mathrm{Card}\big(\partial\mathbf{Sh}_j(\mathbf{T})\big) + \mathrm{Card}\big(\partial\mathbf{Sh}_{j-L}(\mathbf{T})\big).$$

By the nesting property of j-th shadows, the fact $2^{k_j} \sim 2^{k_{j-L}}$, and Lemma 4.17, we obtain that

$$\sum_{j\in \mathrm{scl}(\mathbf{T})} \sum_{I:|I|=2^{-k_j}} \left\| \big(\mathbf{1}^*_{\mathbf{Sh}_j(\mathbf{T})} - \mathbf{1}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}\big)\big(\widetilde{\mathbf{1}}^*_{\mathbf{Sh}_{j-L}(\mathbf{T})}\big)^{-\frac{1}{10}} \right\|_{L^\infty(I)} |I| \le C|I_{\mathbf{T}}|,$$

which yields the desired estimate (4.45). □

Lemma 4.20. *Let $\kappa \in \{1,2\}$ and $\mathbf{T}$ be a convex tree in $\mathbf{S}^{(\kappa)}(\Omega)$. Then we have*

$$\big|\Lambda_{\mathbf{T}}(f_1,f_2,f_3)\big| \le C\text{size}_1^*(\mathbf{T})\text{size}_2^*(\mathbf{T})|I_{\mathbf{T}}|, \tag{4.47}$$

where C is a constant independent of $\mathbf{T},\mathbf{S},f_1,f_2,f_3$.

Proof. By Lemma 4.18, it is sufficient to show that

$$\big|\widetilde{\Lambda}_{\mathbf{T}}(f_1,f_2,f_3)\big| \le C\text{size}_1^*(\mathbf{T})\text{size}_2^*(\mathbf{T})|I_{\mathbf{T}}|, \tag{4.48}$$

where C is a constant independent of $\mathbf{T},\mathbf{S},f_1,f_2,f_3$.

We first prove the simple case $\kappa = 1$. In this case, $k_{j2} = k_j$ for all $(j,n) \in \mathbf{T}$. We thus dominate $|\widetilde{\Lambda}_{\mathbf{T}}|$ by

$$\int_{\mathbb{R}} \sup_j \left| \sum_{n\in\mathbf{T}_j} F_{2,j,n}(x)\right| \prod_{\ell\neq 2}\left(\sum_{(j,n)\in\mathbf{T}} \big|F_{\ell,j,n}(x)\big|^2\right)^{1/2} dx.$$

By the definition of Δ_ℓ and Hölder's inequality, we estimate $|\Lambda_{\mathbf{T}}|$ by

$$\left\| \sup_{(j,n)\in\mathbf{T}} \big|F^*_{2,j,n_2}\big|\right\|_\infty \big\|\Delta_1(\mathbf{T})\big\|_p \big\|\Delta_3(\mathbf{T})\big\|_{p'},$$

where $1/p + 1/p' = 1$ and $F^*_{\ell,j,n} = \mathbf{1}^{**}_{j,n} f_{\ell,j,n_\ell}$. Lemma 4.13 yields that

$$\big\|F^*_{2,j,n}\big\|_\infty \le \text{size}_2^*(\mathbf{T}).$$

Clearly the definition of size yields

$$\|\Delta_1(\mathbf{T})\|_p \le \text{size}_1^*(\mathbf{T})|I_{\mathbf{T}}|^{1/p}.$$

And (4.25) yields

$$\big\|\Delta_3(\mathbf{T})\big\|_{p'} \le C|I_{\mathbf{T}}|^{1/p'}.$$

Putting all of them together, we obtain (4.47) for the case $\kappa = 1$.

We now prove the case $\kappa = 2$. In this case, $2^{k_j} \sim 2^{k_{j1}}$ for all $(j,n) \in \mathbf{T}$. For simplicity, we only consider the case $n_\ell = 0$. The general case can be done in the same way by paying a cost of $(1+|n_\ell|)^{10}$ in the constant. Then we write the trilinear form $\widetilde{\Lambda}_{\mathbf{T}}$ as

$$\widetilde{\Lambda}_{\mathbf{T}}(f_1,f_2,f_3) = \sum_{j\in\mathbb{Z}} \int \prod_{\ell=1}^3 F_{\ell,j}(x)dx,$$

where $F_{\ell,j}$ is defined in (4.43). Here we take a convenient notation that $F_{\ell,j}$ is identically zero if $j \notin \text{scl}(\mathbf{T})$. Let $L = 2^{100}$. By the telescoping argument used in Lemma 3.2, we can write $\widetilde{\Lambda}_{\mathbf{T}}$ as a finite sum of two types of trilinear forms. One type of them is defined by

$$\Lambda_{\mathbf{T},1}(f_1,f_2,f_3) = \int \sum_{j\in\mathbb{Z}} F_{1,j+m'(j)-M}(x)\Pi_{j,L}(F_{2,j},F_{3,j})(x)dx, \tag{4.49}$$

where $m'(j) = [(L_2 j + M_2 - L_1 j - M_1 + 6)/L_1]$, M is an integer between 0 and $6L$, and $\Pi_{j,L}(F_{2,j}, F_{3,j})$ equals either $(F_{2,j} - F_{2,j-L})F_{3,j-8L}$ or $F_{2,j-L}(F_{3,j} - F_{3,j-L})$. Another type is defined by

(4.50)
$$\int \sum_{j\in\mathbb{Z}} \Bigg(\sum_{k=0}^{m'(j)} F_{1,j+k}(x) \Bigg) \big(F_{2,j}(x) - F_{2,j-L}(x)\big)\big(F_{3,j-M}(x) - F_{3,j-M-L}(x)\big)\,dx,$$

which is denoted by $\Lambda_{\mathbf{T},2}(f_1, f_2, f_3)$.

We now prove the estimate for the first type trilinear form $\Lambda_{\mathbf{T},1}$. Let us first consider the case

$$\Lambda_{\mathbf{T},1}(f_1, f_2, f_3) = \int \sum_{j\in\mathbb{Z}} F_{1,j+m'(j)-M}(x)(F_{2,j} - F_{2,j-L})(x)F_{3,j-8L}(x)dx.$$

In this case, by the Cauchy–Schwarz inequality, $|\Lambda_{\mathbf{T},1}|$ is estimated by

$$\int \Big(\sum_j \big|F_{1,j+m'(j)-M}(x)F_{3,j-8L}(x)\big|^2 \Big)^{1/2} \Big(\sum_j \big|F_{2,j}(x) - F_{2,j-L}(x)\big|^2 \Big)^{1/2} dx.$$

Using Hölder's inequality, we dominate it by

$$\Big\| \Big(\sum_j \big|F_{1,j+m'(j)-M}F_{3,j-8L}\big|^2 \Big)^{1/2} \Big\|_{p'} \Big\| \Big(\sum_j \big|F_{2,j} - F_{2,j-L}\big|^2 \Big)^{1/2} \Big\|_p.$$

The first factor in this expression is no more than

$$\Big\| \Big(\sum_j \Big| \sum_{n\in\mathbf{T}_{j+m'(j)-M}} \mathbf{1}^*_{j+m'(j)-M,n} f_{1,j+m'(j)-M,n_1} f_{3,j-8L,0} \Big|^2 \Big)^{1/2} \Big\|_{p'},$$

which is dominated by

$$\Big\| \Big(\sum_j \sum_{n\in\mathbf{T}_{j+m'(j)-M}} \Big| \big(\widetilde{\mathbf{1}}^*_{j+m'(j)-M,n}\big)^2 f_{1,j+m'(j)-M,n_1} f_{3,j-8L,0} \Big|^2 \Big)^{1/2} \Big\|_{p'}.$$

We estimate it by

$$\Big\| \Big(\sum_{(j,n)\in\mathbf{T}} \big|\widetilde{\mathbf{1}}^*_{j,n} f_{1,j,n_1}\big|^2 \Big)^{1/2} \Big\|_{p'} \sup_{(j,n)\in\mathbf{T}} \big\|\widetilde{\mathbf{1}}^*_{j,n} f_{3,\zeta(j,M,K),0}\big\|_\infty,$$

where K is some integer between $-10L$ and $10L$ and $\zeta(j, M, K)$is defined as in (4.8). Clearly, $\widetilde{\mathbf{1}}^*_{j,n} f_{3,\zeta(j,M,K),0}$ is bounded. Also by Lemma 4.14 and an interpolation, we have

$$\Big\| \Big(\sum_j \big|\widetilde{\mathbf{1}}^*_{j,n} f_{1,j,n_1}\big|^2 \Big)^{1/2} \Big\|_{p'} \le C \mathrm{size}^*_1(\mathbf{T}) |I_{\mathbf{T}}|^{1/p'}. \tag{4.51}$$

And Lemma 4.19 yields that

$$\Big\| \Big(\sum_j \big| F_{2,j} - F_{2,j-L} \big|^2 \Big)^{1/2} \Big\|_p \le \mathrm{size}_2^*(\mathbf{T}) |I_{\mathbf{T}}|^{1/p}. \tag{4.52}$$

(4.51) and (4.52) give us the desired estimate for $\Lambda_{\mathbf{T},1}$ in the first case.

We now consider the case

$$\Lambda_{\mathbf{T},1}(f_1, f_2, f_3) = \int \sum_{j \in \mathbb{Z}} F_{1,j+m'(j)-M}(x) F_{2,j-L}(x) \big(F_{3,j} - F_{3,j-L} \big)(x) dx.$$

In this case, using the Cauchy–Schwarz inequality, we have that $|\Lambda_{\mathbf{T},1}|$ is estimated by

$$\int \Big(\sum_j \big| F_{1,j+m'(j)-M}(x) F_{2,j-L}(x) \big|^2 \Big)^{1/2} \Big(\sum_j \big| F_{3,j}(x) - F_{3,j-L}(x) \big|^2 \Big)^{1/2} dx.$$

By Hölder's inequality, we dominate it by

$$\Big\| \Big(\sum_j \big| F_{1,j+m'(j)-M} F_{2,j-L} \big|^2 \Big)^{1/2} \Big\|_{p'} \Big\| \Big(\sum_j \big| F_{3,j} - F_{3,j-L} \big|^2 \Big)^{1/2} \Big\|_p.$$

The first factor in this expression is no more than

$$\Big\| \Big(\sum_j \sum_{n \in \mathbf{T}_{j+m'(j)-M}} \Big| \big(\widetilde{\mathbf{1}}^*_{j+m'(j)-M,n} \big)^2 f_{1,j+m'(j)-M,n_1} f_{2,j-L,0} \Big|^2 \Big)^{1/2} \Big\|_{p'}.$$

We estimate it by

$$\Big\| \Big(\sum_{(j,n) \in \mathbf{T}} |\widetilde{\mathbf{1}}^*_{j,n} f_{1,j,n_1}|^2 \Big)^{1/2} \Big\|_{p'} \sup_{(j,n) \in \mathbf{T}} \big\| \widetilde{\mathbf{1}}^*_{j,n} f_{2,\zeta(j,M,K),0} \big\|_\infty,$$

where K is some integer between $-10L$ and $10L$ and $\zeta(j, M, K)$is defined as in (4.8). By (4.29) and the definition of size, we see that

$$\sup_{(j,n) \in \mathbf{T}} \big\| \widetilde{\mathbf{1}}^*_{j,n} f_{2,\zeta(j,M,K),0} \big\|_\infty \le C \mathrm{size}_2^*(\mathbf{T}). \tag{4.53}$$

Lemma 4.19 and (4.26) yield that

$$\Big\| \Big(\sum_j \big| F_{3,j} - F_{3,j-L} \big|^2 \Big)^{1/2} \Big\|_p \le |I_{\mathbf{T}}|^{1/p}. \tag{4.54}$$

Putting (4.51), (4.53) and (4.54) together, we thus get the desired estimate for $\Lambda_{\mathbf{T},1}$ in the second case.

Finally let us estimate $\Lambda_{\mathbf{T},2}$. The integrand in (4.50) is dominated by

$$\sup_j \left| \sum_{k=0}^{m'(j)} F_{1,j+k}(x) \right| \left(\sum_{j\in\mathbb{Z}} \left| \big(F_{2,j} - F_{2,j-L}\big)(x) \right|^2 \right)^{\frac{1}{2}} \cdot \left(\sum_{j\in\mathbb{Z}} \left| \big(F_{3,j-M} - F_{3,j-M-L}\big)(x) \right|^2 \right)^{\frac{1}{2}}.$$

There exist $p_1, p_3 \in \mathbb{R}$ such that $1/p_1 + 1/p + 1/p_3 = 1$ and $p_1 > p', p_3 > 1$. By Hölder's inequality we dominate $\Lambda_{\mathbf{T},2}$ by

$$\left\| \sup_j \left| \sum_{k=0}^{m'(j)} F_{1,j+k}(x) \right| \right\|_{p_1} \left\| \left(\sum_{j\in\mathbb{Z}} \left| F_{2,j} - F_{2,j-L} \right|^2 \right)^{\frac{1}{2}} \right\|_p \cdot \left\| \left(\sum_{j\in\mathbb{Z}} \left| F_{3,j-M} - F_{3,j-M-L} \right|^2 \right)^{\frac{1}{2}} \right\|_{p_3}.$$

Just notice that one can simply define the size with respect to any number p_3 by using L^{p_3}, then (4.26) and Lemma 4.19 still hold. Thus we have

$$\left\| \left(\sum_{j\in\mathbb{Z}} \left| F_{3,j-M} - F_{3,j-M-L} \right|^2 \right)^{\frac{1}{2}} \right\|_{p_3} \le C|I_{\mathbf{T}}|^{1/p_3} \tag{4.55}$$

Notice that the supports of Fourier transform of $F_{1,j+k}$'s are essentially disjoint. We thus have

$$\left\| \sup_j \left| \sum_{k=0}^{m'(j)} F_{1,j+k}(x) \right| \right\|_{p_1} \le C \left\| \sum_j F_{1,j} \right\|_{p_1}.$$

Clearly,

$$\left\| \sum_j F_{1,j} \right\|_2 \le \big\| \Delta_1(\mathbf{T}) \big\|_2.$$

By Lemma 4.14 and an interpolation, we have that

$$\big\| \Delta_1(\mathbf{T}) \big\|_2 \le C \mathrm{size}_1^*(\mathbf{T}) |I_{\mathbf{T}}|^{1/2}.$$

Thus we get

$$\left\| \sum_j F_{1,j} \right\|_2 \le C \mathrm{size}_1^*(\mathbf{T}) |I_{\mathbf{T}}|^{1/2}.$$

A routine argument as we did in Lemma 4.14 yields

$$\left\| \sum_j F_{1,j} \right\|_{\mathrm{BMO}} \le C \mathrm{size}_1^*(\mathbf{T}). \tag{4.56}$$

Now by an interpolation, we obtain that

$$\Big\| \sum_j F_{1,j} \Big\|_{p_1} \le C \mathrm{size}_1^*(\mathbf{T}) |I_{\mathbf{T}}|^{1/p_1}. \tag{4.57}$$

Hence the desired estimate for $\Lambda_{\mathbf{T},2}$ now follows by (4.57), (4.52) and (4.55). Therefore we obtain Lemma 4.20. □

4.5. Proof of Lemma 4.11. We now prove Lemma 4.11. Without loss of generality, we can assume that $\mathbf{S}$ is a convex set. Lemma 4.16 then yields that $\mathbf{S}^{(1)}(\Omega)$ and $\mathbf{S}^{(2)}(\Omega)$ are convex. By the definition of convexity, we see that the convexity is preserved for a maximal tree in a convex set and the remaining set obtained by removing a maximal tree from a convex set. Thus, applying the organization Lemma 4.15 for $\mathbf{S}^{(\kappa)}(\Omega)$ inductively, we decompose

$$\mathbf{S}^{(\kappa)}(\Omega) = \bigcup_{\sigma} \mathbf{S}_\sigma^{(\kappa)}, \tag{4.58}$$

where $\kappa \in \{1, 2\}$, σ ranges over all possible dyadic numbers, $\mathbf{S}_\sigma^{(\kappa)} = \cup_{\mathbf{T} \in \mathcal{F}_\sigma^{(\kappa)}} \mathbf{T}$ such that $\mathcal{F}_\sigma^\kappa$ is a collection of convex trees with

$$\mathrm{count}(\mathbf{S}_\sigma^{(\kappa)}) \le C\sigma^{-p}, \tag{4.59}$$

and for both $\ell = 1$ and $\ell = 2$,

$$\mathrm{size}_\ell^*(\mathbf{S}_\sigma^{(\kappa)}) \le \sigma |F_\ell|^{1/p}. \tag{4.60}$$

By Lemma 4.12 and the definition of $\mathbf{S}(\Omega)$, we know that $\sigma \le 1$ in order to make $\mathbf{S}_\sigma^{(\kappa)}$ nonempty and we can also sharpen the upper bound in the size estimate for $\mathbf{S}_\sigma^{(\kappa)}$ by

$$\mathrm{size}_\ell^*(\mathbf{S}_\sigma^{(\kappa)}) \le \min\{1, \sigma |F_\ell|^{1/p}\}. \tag{4.61}$$

Hence we estimate $\Lambda_{\mathbf{S}(\Omega)}$ by

$$\big|\Lambda_{\mathbf{S}(\Omega)}(f_1, f_2, f_3)\big| \le \sum_{\kappa=1}^{2} \sum_{\sigma \le 1} \sum_{\mathbf{T} \in \mathcal{F}_\sigma^{(\kappa)}} \big|\Lambda_{\mathbf{T}}(f_1, f_2, f_3)\big|.$$

Lemma 4.20 yields that

$$\big|\Lambda_{\mathbf{S}(\Omega)}(f_1, f_2, f_3)\big| \le \sum_{\kappa=1}^{2} \sum_{\sigma \le 1} \sum_{\mathbf{T} \in \mathcal{F}_\sigma^{(\kappa)}} \mathrm{size}_1^*(\mathbf{S}_\sigma^{(\kappa)}) \mathrm{size}_2^*(\mathbf{S}_\sigma^{(\kappa)}) |I_{\mathbf{T}}|.$$

Applying (4.61) and (4.59), we thus obtain

$$\big|\Lambda_{\mathbf{S}(\Omega)}(f_1, f_2, f_3)\big| \le C \sum_{\sigma \le 1} \min\{1, \sigma |F_1|^{1/p}\} \min\{1, \sigma |F_2|^{1/p}\} \sigma^{-p}, \tag{4.62}$$

which clearly implies (4.24). Therefore we have proven Lemma 4.11.

5. Proof of Theorem 2.2

We now prove Theorem 2.2. The uniform estimate from $L^2 \times L^2$ to L^1 follows immediately by a change of variables and Littlewood–Paley theory and (2.4) is superfluous. Take this simple idea and we can get the uniform estimate for $p_1, p_2 > 2$ and $1 < r < 2$ in Proposition 5.7 for the case $2^{L_2 j+M_2} < 2^{L_1 j+M_1}/8$ or $2^{L_1 j+M_1} < 2^{L_2 j+M_2}/8$. For the general case, we pay a cost of m in the operator norm in this range of p_1, p_2, p to get Lemma 5.4.

For the $r < 1$ case, we use some idea from Section 4 and one can see that technically it is much simpler than what we did in Section 4. We have to assume (2.4) and pay a little more for the operator norm such as $2^{\varepsilon m}$ (see Lemma 5.10). The uniform estimate might be true but $2^{\varepsilon m}$ for a small $\varepsilon > 0$ is good enough for our application.

As we did in Section 4, we set up a trilinear form first. Let us ignore the condition (2.4) for a while. If $2^{L_2 j+M_2} < 2^{L_1 j+M_1}/8$, let

$$\omega'_{3,j} = \big\{\xi : 2^{L_1 j+M_1}/8 \le |\xi| \le 19 \cdot 2^{L_1 j+M_1}/8\big\}$$

and let $\Phi_{3,j}$ be a Schwartz function whose Fourier transform is a bump function adapted to $\omega'_{3,j}$ such that $\widehat{\Phi}_{3,j}(\xi) = 1$ for all $2^{L_1 j+M_1}/4 \le |\xi| \le 9 \cdot 2^{L_1 j+M_1}/4$.

If $2^{L_1 j+M_1} < 2^{L_2 j+M_2}/8$, let $\omega'_{3,j} = \{\xi : 2^{L_2 j+M_2}/8 \le |\xi| \le 19 \cdot 2^{L_2 j+M_2}/8\}$ and $\Phi_{3,j}$ be a Schwartz function whose Fourier transform is a bump function adapted to $\omega'_{3,j}$ such that $\widehat{\Phi}_{3,j}(\xi) = 1$ for all $2^{L_2 j+M_2}/4 \le |\xi| \le 9 \cdot 2^{L_2 j+M_2}/4$.

If $2^{L_1 j+M_1}/8 \le 2^{L_2 j+M_2} \le 8 \cdot 2^{L_1 j+M_1}$, let

$$\omega'_{3,j} = \{\xi : |\xi| \le 18 \cdot \max\{2^{L_1 j+M_1}, 2^{L_2 j+M_2}\}\}$$

and $\Phi_{3,j}$ be a Schwartz function whose Fourier transform is a bump function adapted to $\omega'_{3,j}$ such that $\widehat{\Phi}_{3,j}(\xi) = 1$ for all $|\xi| \le 17 \cdot \max\{2^{L_1 j+M_1}, 2^{L_2 j+M_2}\}$. Let $\Phi_{3,j,m} = \Phi_{3,j}$, $f_{3,j,m}(x) = f_{3,j,0}(x) = f_3 * \Phi_{3,j,0}(x)$. Define a trilinear form $\Lambda_{L_1,L_2,M_1,M_2,m}$ by

$$\Lambda_{L_1,L_2,M_1,M_2,m}(f_1, f_2, f_3) = \int \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} f_{\ell,j,m}(x)dx. \tag{5.1}$$

Clearly $\Lambda_{L_1,L_2,M_1,M_2,m} = \int \Pi_{L_1,L_2,M_1,M_2,m}(f_1, f_2)(x) f_3(x)dx$.

We will prove the following two lemmata.

Lemma 5.1. *Let $p_1, p_2 > 2$ and $1 < r < 2$ such that $1/p_1 + 1/p_1 = 1/r$. Let F_1, F_2, F_3 be measurable sets in $\mathbb{R}$. There exists a constant C independent of $F_1, F_2, F_3, f_1, f_2, f_3$, M_1, M_2, m such that*

$$\big|\Lambda_{L_1,L_2,M_1,M_2,m}(f_1, f_2, f_3)\big| \le Cm|F_1|^{1/p_1}|F_2|^{1/p_2}|F_3|^{1/r'} \tag{5.2}$$

holds for all $f_1 \in X(F_1)$, $f_2 \in X(F_2)$ and $f_3 \in X(F_3)$.

Lemma 5.2. *Let ε be any positive number, $1 < p < 2$ and F_1, F_2, F_3 be measurable sets in $\mathbb{R}$ such that $|F_3| = 1$. Suppose (2.4) holds for all j's.*

Then there is a subset $F_3' \subset F_3$ with $|F_3'| \ge |F_3|/2$ such that for all $p_1, p_2 \ge p$ with $1/p_1+1/p_2 \ge 1$, and all functions $f_1 \in X(F_1)$, $f_2 \in X(F_2)$, $f_3 \in X(F_3)$, the following inequality holds.

$$\big|\Lambda_{L_1,L_2,M_1,M_2,m}(f_1,f_2,f_3)\big| \le C2^{\varepsilon m}|F_1|^{1/p_1}|F_2|^{1/p_2}\,, \tag{5.3}$$

where C is a constant independent of $\mathbf{S}$, $F_1, F_2, F_3, f_1, f_2, f_3, M_1, M_2, m$.

Theorem 2.2 is a consequence of these two lemmas by using interpolation and duality. We also have a corollary from Lemma 5.1 by a simple interpolation.

Corollary 5.3. *Let $p_1, p_2 > 2$ and $1 < r < 2$ such that $1/p_1 + 1/p_1 = 1/r$. There exists a constant C independent of $F_1, F_2, F_3, f_1, f_2, f_3$, M_1, M_2, m such that*

$$\big\|\Pi_{L_1,L_2,M_1,M_2,m}(f_1,f_2)\big\|_r \le Cm\|f_1\|_{p_1}\|f_2\|_{p_2} \tag{5.4}$$

holds for all $f_1 \in L^{p_1}$ and $f_2 \in L^{p_2}$.

5.1. Proof of Lemma 5.1. For $\ell \in \{1,2,3\}$, let $\mathrm{Tr}_{\ell,j,m}$ be a translation function defined by

$$\mathrm{Tr}_{\ell,j,m}(x) = x + m_{j\ell}, \tag{5.5}$$

where $m_{j\ell} = 2^{m-jL_\ell - M_\ell}$ if $\ell \in \{1,2\}$ and $m_{j3} = 0$. Notice that $f_{\ell,j,m}(x) = f_{\ell,j,0}(\mathrm{Tr}_{\ell,j,m}(x))$. Write $\Lambda_{L_1,L_2,M_1,M_2,m}$ as

$$\Lambda_{L_1,L_2,M_1,M_2,m}(f_1,f_2,f_3) = \int_{\mathbb{R}} \prod_{\ell=1}^{3} \sum_{(j,n)\in\mathbb{Z}\times\mathbb{Z}} \mathbf{1}^*_{j,n}\big(\mathrm{Tr}_{\ell,j,m}(x)\big) f_{\ell,j,0}\big(\mathrm{Tr}_{\ell,j,m}(x)\big)dx.$$

For $\mathbf{S} \subset \mathbb{Z}(\gamma)\times\mathbb{Z}$ we define

$$\Lambda_{\mathbf{S},m}(f_1,f_2,f_3) = \int_{\mathbb{R}} \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} \sum_{n\in\mathbf{S}_j} F_{\ell,j,n,m}(x)dx, \tag{5.6}$$

where $\mathbf{S}_j = \{n : (j,n) \in \mathbf{S}\}$ and $F_{\ell,j,n,m}$ is defined by

$$F_{\ell,j,n,m}(x) = \big((\mathbf{1}^*_{j,n} f_{\ell,j,0}) \circ \mathrm{Tr}_{\ell,j,m}\big)(x). \tag{5.7}$$

Let $k_{j\ell}$ be an integer such that $|\omega'_{\ell,j}| \sim 2^{k_{j\ell}}$. For $s = (j,n) \in \mathbf{S}$, let $k_s = k_j = \min_\ell k_{j\ell}$. The time interval of s is defined by $I_s = [2^{-k_s}n, 2^{-k_s}(n+1)]$. We then can define a tree in $\mathbf{S}$ as in Section 4. To prove Lemma 5.1, it is sufficient to prove the following lemma.

Lemma 5.4. *Let $p_1, p_2 > 2$ and $1 < r < 2$ such that $1/p_1 + 1/p_1 = 1/r$. Let F_1, F_2, F_3 be measurable sets in $\mathbb{R}$. There exists a constant C independent of $F_1, F_2, F_3, f_1, f_2, f_3$, M_1, M_2, m such that*

$$\big|\Lambda_{\mathbf{S},m}(f_1,f_2,f_3)\big| \le Cm|F_1|^{1/p_1}|F_2|^{1/p_2}|F_3|^{1/r'} \tag{5.8}$$

holds for all $f_1 \in X(F_1)$, $f_2 \in X(F_2)$ and $f_3 \in X(F_3)$.

By scaling invariance, we can assume that $|F_3| = 1$. We partition $\mathbf{S}$ into two subsets $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$, where

$$\mathbf{S}^{(1)} = \{(j,n) \in \mathbf{S} : |\omega'_{2,j}| \le |\omega'_{1,j}|/10 \text{ or } |\omega'_{1,j}| \le |\omega'_{2,j}|/10\} \tag{5.9}$$

$$\mathbf{S}^{(2)} = \mathbf{S} \backslash \mathbf{S}^{(1)}. \tag{5.10}$$

We should change the definitions of sizes of trees in $\mathbf{S}$.

Definition 5.5. Let $(j,n) \in \mathbf{S}$ and $\ell \in \{1,2,3\}$. Define a seminorm $\big\|f_\ell\big\|_{j,n}$ by

$$\big\|f_\ell\big\|_{j,n} = \frac{1}{|I_s|^{1/2}}\big\|\mathbf{1}^{**}_{j,n} f_{\ell,j,0}\big\|_2 + \frac{1}{|I_s|^{1/2}}\big\|2^{-k_{j\ell}}\mathbf{1}^{**}_{j,n} Df_{\ell,j,0}\big\|_2, \tag{5.11}$$

where $Df_{\ell,j,0}$ is the derivative of $f_{\ell,j,0}$.

Definition 5.6. For $\ell \in \{1,2,3\}$ and a tree $\mathbf{T}$, let $(j_{\mathbf{T}}, n_{\mathbf{T}})$ be the top of the tree $\mathbf{T}$. And define

$$\Delta^*_\ell(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}} \big|\mathbf{1}^*_{j,n} f_{\ell,j,0}(x)\big|^2\right)^{1/2}. \tag{5.12}$$

If $\mathbf{T}$ is a tree in $\mathbf{S}^{(1)}$, we define

$$\mathrm{size}_\ell(\mathbf{T}) = \frac{1}{|I_{\mathbf{T}}|^{1/2}}\big\|\Delta^*_\ell(\mathbf{T})\big\|_2 + \big\|f_\ell\big\|_{j_{\mathbf{T}},n_{\mathbf{T}}}, \tag{5.13}$$

for all $\ell \in \{1,2,3\}$.

If $\mathbf{T}$ is a tree in $\mathbf{S}^{(2)}$, define $\mathrm{size}_\ell(\mathbf{T})$ by (5.13) only for $\ell \in \{1,2\}$. For $\ell = 3$, we define the size by

$$\mathrm{size}_3(\mathbf{T}) = \big\|f_3\big\|_{j_{\mathbf{T}},n_{\mathbf{T}}}, \tag{5.14}$$

Let $\mathbf{P}$ be a subset of $\mathbf{S}$. Define the ℓ-size* of $\mathbf{T}$ by

$$\mathrm{size}^*_\ell(\mathbf{P}) = \sup_{\mathbf{T}:\mathbf{T}\subset\mathbf{P}} \mathrm{size}_\ell(\mathbf{T}), \tag{5.15}$$

where $\mathbf{T}$ ranges over all trees in $\mathbf{P}$.

One should notice that for $\Lambda_{\mathbf{S}^{(1)},m}$ we have a uniform estimate for $p_1, p_2 > 2$ and $1 < r < 2$. We state it as follow

Proposition 5.7. *Let $p_1, p_2 > 2$ and $1 < r < 2$ with $1/p_1 + 1/p_2 = 1/r$. Let $f_1 \in L^{p_1}$, $f_2 \in L^{p_2}$ and $f_3 \in L^{r'}$. Then*

$$\big|\Lambda_{\mathbf{S}^{(1)},m}(f_1,f_2,f_3)\big| \le C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{r'}, \tag{5.16}$$

where C is independent of m, f_1, f_2, f_3.

Proof. We do not need time frequency analysis for this proposition. The key point is that when $s \in \mathbf{S}^{(1)}$ the support of Fourier transform of $f_{3,j,0}$ is away from the origin so that we can apply the Littlewood–Paley Theorem for the square function generated by $f_{3,j,0}$'s. Clearly $|\Lambda_{\mathbf{S}^{(1)},m}|$ is estimated by

$$\int_{\mathbb{R}} \sum_j \prod_{\ell=1}^{3} f_{\ell,j,0}(\mathrm{Tr}_{\ell,j,m}(x))dx.$$

By Hölder's inequality, we dominate $|\Lambda_{\mathbf{S}^{(1)},m}|$ by

$$\Big\| \Big(\sum_j |f_{1,j,0} \circ \mathrm{Tr}_{1,j,m}|^{p_1} \Big)^{1/p_1} \Big\|_{p_1} \Big\| \Big(\sum_j |f_{2,j,0} \circ \mathrm{Tr}_{2,j,m}|^{p_2} \Big)^{1/p_2} \Big\|_{p_2} \cdot \Big\| \Big(\sum_j |f_{3,j,0}|^{r'} \Big)^{1/r'} \Big\|_{r'}.$$

By a change of variables, it is clear that for $\ell = 1, 2$,

$$\Big\| \Big(\sum_j |f_{\ell,j,0} \circ \mathrm{Tr}_{\ell,j,m}|^{p_\ell} \Big)^{1/p_\ell} \Big\|_{p_\ell} = \Big\| \Big(\sum_j |f_{\ell,j,0}|^{p_\ell} \Big)^{1/p_\ell} \Big\|_{p_\ell}.$$

Notice the elementary inequality

$$\Big(\sum_j |a_j|^q \Big)^{1/q} \le \Big(\sum_j |a_j|^2 \Big)^{1/2}$$

holds for $q \ge 2$. We thus dominate $|\Lambda_{\mathbf{S}^{(1)},m}|$ by

$$\Big\| \Big(\sum_j |f_{1,j,0}|^2 \Big)^{1/2} \Big\|_{p_1} \Big\| \Big(\sum_j |f_{2,j,0}|^2 \Big)^{1/2} \Big\|_{p_2} \Big\| \Big(\sum_j |f_{3,j,0}|^2 \Big)^{1/2} \Big\|_{r'}.$$

Now the Littlewood–Paley theorem yields the desired estimate (5.16). This proves the proposition. □

We now use time frequency analysis to prove Lemma 5.4. Although we only need to estimate $\Lambda_{\mathbf{S}^{(2)},m}$ due to Proposition 5.7, we still write a proof for both of $\Lambda_{\mathbf{S}^{(1)},m}$ and $\Lambda_{\mathbf{S}^{(2)},m}$.

We first prove the size estimate for a single tree, that is,

$$\big|\Lambda_{\mathbf{T},m}(f_1, f_2, f_3)\big| \le C \prod_{\ell=1}^{3} \mathrm{size}_\ell^*(\mathbf{T}) |I_{\mathbf{T}}|. \tag{5.17}$$

We only prove the case when $\mathbf{T}$ is a tree in $\mathbf{S}^{(2)}$ for (5.17) since the other case is similar. In this case $2^{k_{j\ell}} \sim 2^{k_j}$ for all ℓ in $\{1, 2, 3\}$. We thus dominate

$|\Lambda_{\mathbf{T},m}|$ by

$$\int_{\mathbb{R}} \sup_{(j,n)\in\mathbf{T}} \left|(\mathbf{1}^{**}_{j,n} f_{3,j,0})\circ \mathrm{Tr}_{\ell,j,m}(x)\right| \prod_{\ell\neq 3} \left(\sum_{(j,n)\in\mathbf{T}} \left|(\mathbf{1}^{**}_{j,n} f_{\ell,j,0})\circ \mathrm{Tr}_{\ell,j,m}(x)\right|^2 \right)^{1/2} dx.$$

By the definition of Δ_ℓ and Hölder's inequality, we estimate $|\Lambda_{\mathbf{T},m}|$ by

$$\sup_{(j,n)\in\mathbf{T}} \left\|F^*_{3,j,0}\right\|_\infty \left\|\Delta^*_1(\mathbf{T})\right\|_2 \left\|\Delta^*_2(\mathbf{T})\right\|_2,$$

where $F^*_{3,j,0} = \mathbf{1}^{**}_{j,n} f_{3,j,0}$. Notice that Lemma 4.13 holds for the seminorm. Thus we have

$$\left\|F^*_{3,j,0}\right\|_\infty \le \mathrm{size}^*_3(\mathbf{T}).$$

Clearly the definition of size yields

$$\|\Delta_\ell(\mathbf{T})\|_2 \le \mathrm{size}^*_\ell(\mathbf{T}) |I_{\mathbf{T}}|^{1/2}$$

for $\ell \in \{1,2\}$. Putting all of them together, we obtain (5.17).

Lemma 5.8. *Let $\kappa \in \{1,2\}$, $\mathbf{T}$ be a tree in $\mathbf{S}^{(\kappa)}$ and $\mathbf{P}$ be a subset of $\mathbf{S}^{(\kappa)}$. Suppose that $\mathbf{P}\cap\mathbf{T} = \emptyset$ and $\mathbf{T}$ is a maximal tree in $\mathbf{P}\cup\mathbf{T}$. Then we have*

$$\left|\Lambda_{\mathbf{P}\cup\mathbf{T},m}(f_1,f_2,f_3) - \Lambda_{\mathbf{P},m}(f_1,f_2,f_3)\right| \le Cm \prod_{\ell=1}^3 \mathrm{size}^*_\ell(\mathbf{T}\cup\mathbf{P}) |I_{\mathbf{T}}|, \tag{5.18}$$

where C is independent of $f_1, f_2, f_3, \mathbf{P}$, $\mathbf{T}$.

Proof. Clearly the difference $|\Lambda_{\mathbf{P}\cup\mathbf{T},m} - \Lambda_{\mathbf{P},m}|$ is dominated by a sum of $C\Lambda_{\mathbf{T},m}$ and at most finite many following trilinear forms

$$\left| \int \sum_{j\in \mathrm{scl}(\mathbf{T})} \left(\sum_{n\in\mathbf{T}_j} F_{\ell_1,j,n,m}(x) \right) \left(\sum_{n\in\mathbf{P}_j} F_{\ell_2,j,n,m}(x) \right) \cdot \left(\sum_{n\in(\mathbf{P}\cup\mathbf{T})_j} F_{\ell_3,j,n,m}(x) \right) dx \right|,$$

where (ℓ_1,ℓ_2,ℓ_3) is a permutation of $(1,2,3)$. By (5.17), it sufficient to show that this trilinear form can be estimated by the right-hand side of (5.8). We only handle the most difficult case $\ell_1 = 1, \ell_2 = 2$. Other cases are similar. We estimate the trilinear form by

$$\sum_{j\in\mathrm{scl}(\mathbf{T})} \sum_{I:|I|=2^{-k_j}} \left\| \left(\sum_{n\in\mathbf{T}_j} F_{1,j,n,m} \right) \left(\sum_{n\in\mathbf{P}_j} F_{2,j,n,m} \right) \left(\sum_{n\in(\mathbf{P}\cup\mathbf{T})_j} F_{3,j,n,m} \right) \right\|_{L^1(I)}. \tag{5.19}$$

There is at least one index $\ell \in \{1,2\}$ satisfying $k_{j\ell} = k_j$. Without loss of generality, assume $k_{j1} = k_j$. We have that for any positive integer N,

$$\left\| \sum_{n\in\mathbf{T}_j} F_{1,j,n,m} \right\|_{L^\infty(I)} \le \frac{C_N}{\left(1 + 2^{k_j} \mathrm{dist}(I(m_{j1}), I_{\mathbf{T}})\right)^N} \left\|\mathbf{1}^{**}_{j,n'} f_{\ell,j,0}\right\|_\infty,$$

where $I(m_{j1}) = I + m_{j1}$ is an interval generated by shifting I to the right by m_{j1} and $n' \in (\mathbf{P} \cup \mathbf{T})_j$ which minimizes the distance between $I_{j,n}$ and $I(m_{j1})$. Since Lemma 4.13 holds for the seminorm, we get

$$\Big\| \sum_{n \in \mathbf{T}_j} F_{1,j,n,m} \Big\|_{L^\infty(I)} \leq \frac{C_N \mathrm{size}_1^*(\mathbf{P} \cup \mathbf{T})}{\big(1 + 2^{k_j} \mathrm{dist}(I(m_{j1}), I_{\mathbf{T}})\big)^N}.$$

And since $\mathbf{P} \cap \mathbf{T} = \emptyset$ and $\mathbf{T}$ is a maximal tree in $\mathbf{P} \cup \mathbf{T}$, we have

$$\Big\| \sum_{n \in \mathbf{P}_j} F_{2,j,n,m} \Big\|_{L^2(I)} \leq \frac{C_N}{\big(1 + 2^{k_j} \mathrm{dist}(I(m_{j2}), (I_{\mathbf{T}})^c)\big)^N} \big\| \mathbf{1}_{j,n'}^{**} f_{\ell,j,0} \big\|_2,$$

which is obviously bounded by

$$\frac{C_N \mathrm{size}_2^*(\mathbf{P} \cup \mathbf{T}) |I|^{1/2}}{\big(1 + 2^{k_j} \mathrm{dist}(I(m_{j2}), (I_{\mathbf{T}})^c)\big)^N}.$$

Similarly, we also have

$$\Big\| \sum_{n \in (\mathbf{P} \cup \mathbf{T})_j} F_{3,j,n,m} \Big\|_{L^2(I)} \leq C \mathrm{size}_3^*(\mathbf{P} \cup \mathbf{T}) |I|^{1/2}.$$

Thus we estimate (5.19) by

$$\sum_{j \in \mathrm{scl}(\mathbf{T})} \sum_{I : |I| = 2^{-k_j}} \frac{C_N \mathrm{size}_1^*(\mathbf{P} \cup \mathbf{T}) \mathrm{size}_2^*(\mathbf{P} \cup \mathbf{T}) \mathrm{size}_3^*(\mathbf{P} \cup \mathbf{T}) |I|}{\big(1 + 2^{k_j} \mathrm{dist}(I(m_{j1}), I_{\mathbf{T}})\big)^N \big(1 + 2^{k_j} \mathrm{dist}(I(m_{j2}), (I_{\mathbf{T}})^c)\big)^N}.$$

Let $j_{\mathbf{T}}$ be the index for the top of $\mathbf{T}$. If $j_{\mathbf{T}} + 10m \geq j \geq j_{\mathbf{T}}$, we only have at most $10m$ different values for j. Notice that if $I(m_{j1}) \subset (I_{\mathbf{T}})^c$, then we can replace $\mathrm{dist}(I(m_{j1}), I_{\mathbf{T}})$ by $\mathrm{dist}(I(m_{j1}), \partial I_{\mathbf{T}})$. Thus if we only sum j from $j_{\mathbf{T}}$ to $j_{\mathbf{T}} + 10m$ we get that (5.19) is dominated by

$$Cm \prod_{\ell=1}^{3} \mathrm{size}_\ell^*(\mathbf{P} \cup \mathbf{T}) |I_{\mathbf{T}}|.$$

The remaining thing we need to deal with is to sum all $j > j_{\mathbf{T}} + 10m$. The main difficulty is the case $I(m_{j1}) \nsubseteq (I_{\mathbf{T}})^c$ and $I(m_{j2}) \nsubseteq I_{\mathbf{T}}$, because in other cases we gain $\big(1 + 2^{k_j} \mathrm{dist}(I(m_{j\ell}), \partial I_{\mathbf{T}})\big)^{-100}$ in the estimate for at least one of $\ell \in \{1, 2\}$, which trivializes the estimate. We also know from the definition of $m_{j\ell}$ that $\mathrm{dist}(I(m_{j1}), I(m_{j2})) \leq 2^m |I|$. To make the difficult case happen, the interval I must satisfy $\mathrm{dist}(I(m_{j\ell}), \partial I_{\mathbf{T}}) \leq 10 \cdot 2^m |I|$ for both $\ell = 1, 2$. Sum $|I(m_{j\ell})|$ for all such I's to get a upper bound $C 2^m 2^{-k_j}$. Then summing these upper bounds for all $j > j_{\mathbf{T}} + 10m$ we get a bound $C 2^{-8m} |I_{\mathbf{T}}|$. Therefore we estimate (5.19) by $Cm \prod_{\ell=1}^3 \mathrm{size}_\ell^*(\mathbf{P} \cup \mathbf{T}) |I_{\mathbf{T}}|$. This proves the lemma. □

Lemma 4.15 still holds for the sizes of trees defined in Subsection 5.1. Let $\kappa \in \{1,2\}$. Applying this organization lemma inductively for $\mathbf{S}^{(\kappa)}$, we decompose

$$\mathbf{S}^{(\kappa)} = \bigcup_{\sigma} \mathbf{S}_{\sigma}^{(\kappa)}, \tag{5.20}$$

where σ ranges over all possible dyadic numbers, $\mathbf{S}_{\sigma}^{(\kappa)} = \cup_{\mathbf{T}\in\mathcal{F}_{\sigma}^{(\kappa)}}\mathbf{T}$ such that $\mathcal{F}_{\sigma}^{\kappa}$ is a collection of maximal trees with

$$\operatorname{count}(\mathbf{S}_{\sigma}^{(\kappa)}) \le C\sigma^{-2}, \tag{5.21}$$

and

$$\operatorname{size}_{\ell}^{*}(\mathbf{S}_{\sigma}^{(\kappa)}) \le \sigma|F_{\ell}|^{1/2} \tag{5.22}$$

holds for all $\ell \in \{1,2,3\}$.

Notice that Lemma 4.12 holds for the new sizes of trees defined in Subsection 5.1. We thus can also sharpen the upper bound in the size estimate for $\mathbf{S}_{\sigma}^{(\kappa)}$ by

$$\operatorname{size}_{\ell}^{*}(\mathbf{S}_{\sigma}^{(\kappa)}) \le \min\{1, \sigma|F_{\ell}|^{1/2}\}. \tag{5.23}$$

Hence by Lemma 5.8 we estimate $\Lambda_{\mathbf{S},m}$ by

$$\sum_{\kappa=1}^{2}\sum_{\sigma}\sum_{\mathbf{T}\in\mathcal{F}_{\sigma}^{(\kappa)}} m\prod_{\ell=1}^{3}\operatorname{size}_{\ell}^{*}(\mathbf{S}_{\sigma}^{(\kappa)})|I_{\mathbf{T}}|.$$

Applying (5.23) and (5.21), we thus obtain

$$\begin{aligned}\big|\Lambda_{\mathbf{S},m}(f_1,f_2,f_3)\big| &\\ \le Cm\sum_{\sigma}\sigma^{-2}\min\{1,\sigma|F_1|^{1/2}\}\min\{1,\sigma|F_2|^{1/2}\}\min\{1,\sigma\},&\end{aligned} \tag{5.24}$$

which clearly implies (5.8). Therefore we have proven Lemma 5.4.

5.2. A truncated trilinear form. First by a change of variable, we write $\Lambda_{L_1,L_2,M_1,M_2,m}$ as

$$\Lambda_{L_1,L_2,M_1M_2,m}(f_1,f_2,f_3) = \int\sum_{j}\prod_{\ell=1}^{3} f_{\ell,j,0}\big(\widetilde{\operatorname{Tr}}_{\ell,j,m}(x)\big)dx, \tag{5.25}$$

where $\widetilde{\operatorname{Tr}}_{1,j,m}(x) = \operatorname{Tr}_{1,j,m} - m_{j2}$, $\widetilde{\operatorname{Tr}}_{2,j,m}(x) = x$, $\widetilde{\operatorname{Tr}}_{3,j,m}(x) = x - m_{j2}$.

To prove Lemma 5.2, we have to set up our time-frequency decomposition in a slightly different way for technical reasons. Recall that ψ is a nonnegative Schwartz function such that $\widehat{\psi}$ is supported in $[-1/100, 1/100]$ and $\widehat{\psi}(0) = 1$. And $\psi_k(x) = 2^k\psi(2^k x)$. Let Ω be the set defined as in (4.6).

As before, $k_{j\ell}$ is an integer such that $2^{k_{j\ell}} \sim |\omega'_{\ell,j}|$ for for $\ell \in \{1,2,3\}$ and $k_j = \min\{k_{j\ell}\}$. For a very small positive number ε, we define

$$\Omega_j = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) \geq 2^{\varepsilon^2 m} 2^{-k_j}\}. \tag{5.26}$$

$$\psi_{j1} = \psi_{j2} = \psi_{j3} = \mathbf{1}_{(\Omega_j)^c} * \psi_{k_j}(x). \tag{5.27}$$

Ω_j, $\psi_{j\ell}$ depend on m, ε but this dependence is suppressed for notational convenience. A truncated trilinear form is defined by

$$\Lambda_{\Omega,m}(f_1, f_2, f_3) = \int \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} \psi_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\operatorname{Tr}}_{\ell,j,m}(x)\big) dx. \tag{5.28}$$

Heuristically, $\psi_{j\ell}$ can be considered as $\mathbf{1}_{(\Omega_j)^c}$ since it is a smooth approximation of $\mathbf{1}_{(\Omega_j)^c}$. In time space, Ω_j is an exceptional set which can be removed. we can handle it well. The technical details about this can be found in Section 4. In order to get $2^{\varepsilon m}$ instead of 2^m in the estimates, we have to remove only a smaller set. Here is the lemma which allows us to do so.

Lemma 5.9. *Let F_1, F_2, F_3 be measurable sets. Let $F'_3 = F_3 \backslash \Omega$. Then*

$$\begin{aligned} &\big|\big(\Lambda_{L_1,L_2,M_1,M_2,m} - \Lambda_{\Omega,m}\big)(f_1, f_2, f_3)\big| \\ &\qquad \leq C 2^{-100m} \min\big\{1, |F_1|^{1/p}\big\} \min\big\{1, |F_2|^{1/p}\big\} \end{aligned} \tag{5.29}$$

holds for all functions $f_1 \in X(F_1), f_2 \in X(F_2), f_3 \in X(F'_3)$, where C is a constant independent of L_1, L_2, M_1, M_2, m, f_1, f_2, f_3, F_1, F_2, F_3.

Proof. The difference $|\Lambda_{L_1,L_2,M_1,M_2,m} - \Lambda_{\Omega,m}|$ is dominated by

$$\int \sum_j \Big|1 - \prod_{\ell=1}^{3} \psi_{j\ell}(x)\Big| \Big|\prod_{\ell=1}^{3} f_{\ell,j,0}\big(\widetilde{\operatorname{Tr}}_{\ell,j,m}(x)\big)\Big| dx.$$

Clearly,

$$\Big|1 - \prod_{\ell=1}^{3} \psi_{j\ell}(x)\Big| \leq 3 \sum_{\ell=1}^{3} \big|1 - \psi_{j\ell}(x)\big|$$

For $\ell = \{1,2\}$, by the definition of Ω, we have for any positive integer N,

$$\begin{aligned} \big|f_{\ell,j,0}\big(\widetilde{\operatorname{Tr}}_{\ell,j,m}(x)\big)\big| &\leq \int \frac{C_N |f_\ell(y)| 2^{k_{j\ell}}}{\big(1 + 2^{k_{j\ell}} |\widetilde{\operatorname{Tr}}_{\ell,j,m}(x) - y|\big)^N} dy \\ &\leq C 2^{2m} (1 + 2^{k_{j\ell}} \operatorname{dist}(\widetilde{\operatorname{Tr}}_{\ell,j,m}(x), \Omega^c))^2 \min\{1, |F_\ell|^{1/p}\}. \end{aligned}$$

Since $f_3 \in X(F'_3)$, we obtain that

$$\big|f_{3,j,0}(x)\big| \leq \frac{C_N}{\big(1 + 2^{k_{j3}} \operatorname{dist}(\widetilde{\operatorname{Tr}}_{3,j,m}(x), \Omega^c)\big)^N}. \tag{5.30}$$

Thus by the fact that $2^{k_{j3}} \sim \max\{2^{k_{j\ell}}\}$, $k_{j2} > k_{j1} + m$ and the definition of Ω_j, the difference in the left-hand side of (5.29) is estimated by

$$\begin{aligned}
&\sum_j \int\int_{\Omega_j} \frac{2^{k_j}}{\big(1+2^{k_j}|x-y|\big)^N} dy \frac{C_N 2^{4m} \min\{1,|F_1|^{1/p}\} \min\{1,|F_2|^{1/p}\}}{\big(1+2^{k_{j3}} \mathrm{dist}(\widetilde{\mathrm{Tr}}_{3,j,m}(x), \Omega^c)\big)^N} dx \\
&\le \sum_j \int_{\Omega_j} \frac{C_N 2^{4m} \min\{1,|F_1|^{1/p}\} \min\{1,|F_2|^{1/p}\}}{\big(1+2^{k_j} \mathrm{dist}(y,\Omega^c)\big)^N} dy \\
&\le C 2^{-100m} \min\{1,|F_1|^{1/p}\} \min\{1,|F_2|^{1/p}\}.
\end{aligned}$$

Therefore we have finished the proof. □

By this lemma, we only need to consider $\Lambda_{\Omega,m}$. For $\mathbf{S} \subset \mathbb{Z}(\gamma) \times \mathbb{Z}$ we define

$$\Lambda_{\mathbf{S},\Omega,m}(f_1,f_2,f_3) = \int_{\mathbb{R}} \sum_{j\in\mathbb{Z}} \prod_{\ell=1}^{3} \sum_{n\in\mathbf{S}_j} \widetilde{F}_{\ell,j,n,m}(x) dx, \tag{5.31}$$

where $\widetilde{F}_{\ell,j,n,m}$ is defined by

$$\widetilde{F}_{\ell,j,n,m}(x) = \psi_{j\ell}(x) \mathbf{1}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big). \tag{5.32}$$

As before we only need to consider the trilinear form (5.31). To prove Lemma 5.2, it is sufficient to show the following lemma due to Lemma 5.9.

Lemma 5.10. *Let ε be any positive number, $1 < p < 2$ and F_1, F_2, F_3 be measurable sets in $\mathbb{R}$ such that $|F_3| = 1$. There is a subset $F_3' \subset F_3$ with $|F_3'| \ge |F_3|/2$ such that for all $p_1, p_2 \ge p$ with $1/p_1 + 1/p_2 \ge 1$, and all functions $f_1 \in X(F_1)$, $f_2 \in X(F_2)$, $f_3 \in X(F_3)$, the following inequality holds.*

$$\big|\Lambda_{\mathbf{S},\Omega,m}(f_1,f_2,f_3)\big| \le C 2^{\varepsilon m} |F_1|^{1/p_1} |F_2|^{1/p_2}, \tag{5.33}$$

where C is a constant independent of $\mathbf{S}$, F_1, F_2, F_3, f_1, f_2, f_3, L_1, L_2, M_1, M_2, m.

5.3. Preliminary lemmata. To prove Lemma 5.10, we should change the definitions of size of a tree in $\mathbf{S}$ and set up some lemmata first.

Definition 5.11. Let $(j,n) \in \mathbf{S}$ and $\ell \in \{1,2,3\}$. Let $\psi^*_{j\ell}$ be the function

$$\psi^*_{j\ell}(x) = \int_{(\Omega_j)^c} \frac{2^{k_j}}{\big(1+2^{2k_j}|x-y|^2\big)^{200}} dy \tag{5.34}$$

Define a seminorm $\big\|f_\ell\big\|_{j,n,m}$ by

$$\begin{aligned}
\big\|f_\ell\big\|_{j,n,m} = {} & \frac{1}{|I_s|^{1/p}} \big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big) f_{\ell,j,0}\big\|_p \\
& + \frac{1}{|I_s|^{1/2}} \big\|2^{-k_{j\ell}} \mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big) D f_{\ell,j,0}\big\|_p,
\end{aligned} \tag{5.35}$$

where $\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}$ is the inverse of $\widetilde{\mathrm{Tr}}_{\ell,j,m}$ and $Df_{\ell,j,0}$ is the derivative of $f_{\ell,j,0}$.

Definition 5.12. For $\ell \in \{1,2\}$ and a tree $\mathbf{T}$, let $(j_{\mathbf{T}}, n_{\mathbf{T}})$ be the top of the tree $\mathbf{T}$. And let $\Delta^*_{\ell,m}(\mathbf{T})$ be defined by

$$\Delta^*_{\ell,m}(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}} \left|\mathbf{1}^*_{j,n}(x)\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)(x) f_{\ell,j,0}(x)\right|^2 \right)^{1/2}. \tag{5.36}$$

If $\mathbf{T}$ is a tree in $\mathbf{S}$, we define

$$\mathrm{size}_{\ell,m}(\mathbf{T}) = \frac{1}{|I_{\mathbf{T}}|^{1/2}} \big\|\Delta^*_{\ell,m}(\mathbf{T})\big\|_p + \big\|f_\ell\big\|_{j_{\mathbf{T}},n_{\mathbf{T}},m}, \tag{5.37}$$

for all $\ell \in \{1,2\}$.

Let $\mathbf{P}$ be a subset of $\mathbf{S}$. Define the (ℓ,m)-size* of $\mathbf{T}$ by

$$\mathrm{size}^*_{\ell,m}(\mathbf{P}) = \sup_{\mathbf{T}:\mathbf{T}\subset\mathbf{P}} \mathrm{size}_{\ell,m}(\mathbf{T}), \tag{5.38}$$

where $\mathbf{T}$ ranges over all trees in $\mathbf{P}$.

In the definition of $\psi^*_{j\ell}$, we can replace the exponent 200 by a larger number 2^{100} to define a new function. Denote this function by $\widetilde{\psi}^*_{j\ell}$. If $\mathbf{1}^*_{j,n}$ and $\psi^*_{j\ell}$ are replaced by $\widetilde{\mathbf{1}}^*_{j,n}$ and $\widetilde{\psi}^*_{j\ell}$ respectively in the definition $\Delta^*_{\ell,m}(\mathbf{T})$, we denote the corresponding function by $\Delta_{\ell,m}(\mathbf{T})$.

Lemma 5.13. *Let $1 < q < \infty$, $\ell \in \{1,2,3\}$ and $\mathbf{T}$ be a tree in $\mathbf{S}$. Then*

$$\big\|\Delta^*_{\ell,m}(\mathbf{T})\big\|_q \le C \inf_{x\in I_{\mathbf{T}}} M_q(Mf_\ell)(x)|I_{\mathbf{T}}|^{1/q}, \tag{5.39}$$

$$\mathrm{size}_{\ell,m}(\mathbf{T}) \le C \min\{2^{\beta_\ell m}|F_\ell|^{1/p}, \inf_{x\in I_{\mathbf{T}}} M_p(Mf_\ell)(x)\}, \tag{5.40}$$

where $\beta_\ell = 1$ if $\ell = 1$, $\beta_\ell = \varepsilon^2$ if $\ell = 2$, and C is a constant independent of $f_\ell, \mathbf{T}, \mathbf{S}, L_1, L_2, M_1, M_2$.

Proof. Repeating an argument similar to that used in the proof of (4.25) and (4.26), we easily obtain (5.39) and part of (5.40). The only thing we need to prove is

$$\mathrm{size}_{\ell,m}(\mathbf{T}) \le C 2^{\beta_\ell m}|F_\ell|^{1/p}. \tag{5.41}$$

Assume $2^{\beta_\ell m+10} I_{\mathbf{T}} \subset \Omega$: otherwise the upper bound $\inf_{x\in I_{\mathbf{T}}} M_p(Mf_\ell)(x)$ implies (5.41). Let $\mathbf{T}_L$ be a collection of all $s = (j,n) \in \mathbf{T}$ such that $2^L I_s \subset \Omega$ but $2^{L+1} I_s \nsubseteq \Omega$. Then

$$\mathbf{T} = \bigcup_{L=[\beta_\ell m+10]}^{\infty} \mathbf{T}_L$$

Let $\mathbb{J}_L$ be the set of all time intervals I_s's for $s \in \mathbf{T}_L$. Clearly, $\mathbb{J}_L$ is a set of disjoint intervals and $\sum_{J\in\mathbb{J}_L} |J| \le \min\{|I_{\mathbf{T}}|, 1\}$. Thus it is sufficient to show that for any $J \in \mathbb{J}_L$ and any $(j,n) \in \mathbf{T}$ such that $I_s = J$,

$$\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)\widetilde{f}_{\ell,j,0}\big\|_p^p \le C_N\big(\inf_{x\in J} M_p(Mf_\ell)(x)\big)^p L^{-N}|J| \tag{5.42}$$

holds for a large integer N, where $\widetilde{f}_{\ell,j,0}$ is $f_{\ell,j,0}$ or $2^{-k_{j\ell}}Df_{\ell,j,0}$, since the desired estimate follows by summing up all L's and J's. By the definition of $\psi^*_{j\ell}$, we have

$$\begin{aligned}&\big|\mathbf{1}^{**}_{j,n}(x)\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(x)\big| \\ &\qquad\le \frac{C}{\big(1+2^{k_j}\mathrm{dist}(x,J)\big)^{200}\big(1+2^{k_j}\mathrm{dist}\big(\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(x),(\Omega_j)^c\big)\big)^{200}},\end{aligned}$$

which is clearly dominated by

$$\frac{C}{\big(1+2^{k_j}\mathrm{dist}(x,J)\big)^{100}\big(1+2^{k_j}\mathrm{dist}\big(J_{j,m},(\Omega_j)^c\big)\big)^{100}},$$

where $J_{j,m}$ is the interval $\{\widetilde{\mathrm{Tr}}_{\ell,j,m}(x) : x \in J\}$. Since $L \ge \beta_\ell m + 9$, by the definition of $\widetilde{\mathrm{Tr}}_{\ell,j,m}$ we thus dominate $\big|\mathbf{1}^{**}_{j,n}(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m})\big|$ by

$$\frac{C}{\big(1+2^{k_j}\mathrm{dist}(x,J)\big)^{100}\big(1+2^{k_j}\mathrm{dist}\big(J,(\Omega)^c\big)\big)^{100}}.$$

Thus we have

$$\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)\widetilde{f}_{\ell,j,0}\big\|_p^p \le C\big(\inf_{x\in J} M_p(Mf_\ell)(x)\big)^p L^{-100p}|J|,$$

which yields (5.42). □

Lemma 5.14. *Suppose that $s = (j,n) \in \mathbf{S}$. If $2^{k_{j\ell}} \sim 2^{k_j}$, then*

$$\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,0}\big\|_\infty \le C\big\|f_\ell\big\|_{j,n,m} \tag{5.43}$$

holds for $\ell \in \{1,2,3\}$, where C is a constant independent of s, f_ℓ, m, L_1, L_2, M_1, M_2.

Proof. Let $\mu = \big\|f_\ell\big\|_{j,n,m}$. By the definition of the seminorm, we have

$$\begin{aligned}&\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,0}\big\|_p + \big\||I_s|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)Df_{\ell,j,0}\big\|_p \\ &\qquad\qquad\le \mu|I_s|^{1/p}.\end{aligned} \tag{5.44}$$

First we prove the BMO estimate for the function, that is

$$\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell} \circ \widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,0}\big\|_{\mathrm{BMO}} \le C\mu. \tag{5.45}$$

If $|I_s| \le |J|$, by (5.44) we have

$$\inf_c \int_J \big|\mathbf{1}^{**}_{j,n}(x)\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)(x)f_{\ell,j,0}(x)-c\big|dx \\ \le \big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,0}\big\|_p|J|^{1-\frac{1}{p}} \le \mu|I_s|^{\frac{1}{p}}|J|^{1-\frac{1}{p}} \le \mu|J|.$$

If $|I_s| \ge |J|$, by (5.44) we obtain that

$$\begin{aligned}
&\inf_c \int_J \big|\mathbf{1}^{**}_{j,n}(x)\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)(x)f_{\ell,j,n_\ell}(x)-c\big|dx\\
&\le |J|\int_J\left|\left(\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,n_\ell}\right)'(x)\right|dx\\
&\le C|J||I_s|^{-1}\int_J\big|\mathbf{1}^{**}_{j,n}(x)\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)(x)f_{\ell,j,n_\ell}(x)\big|dx\\
&\quad+|J|\int_J\big|\mathbf{1}^{**}_{j,n}(x)\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)(x)Df_{\ell,j,n_\ell}(x)\big|dx\\
&\le C|J||I_s|^{-1}\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,n_\ell}\big\|_p|J|^{1-\frac{1}{p}}\\
&\quad+|J|\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)Df_{\ell,j,n_\ell}\big\|_p|J|^{1-\frac{1}{p}}\\
&\le C\mu|J|^{2-\frac{1}{p}}|I_s|^{\frac{1}{p}-1} \le C\mu|J|.
\end{aligned}$$

Thus we get the BMO estimate (5.45). Interpolating (5.45) and (5.44), we have for any $p \le q < \infty$,

$$\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,n_\ell}\big\|_q \le C\mu|I_s|^{1/q}.$$

Notice that an integration by parts and Hölder's inequality yield that

$$\big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,n_\ell}\big\|_\infty \\ \le \big\|\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,n_\ell}\big\|^{1/2}_{p'}\big\|\big(\mathbf{1}^{**}_{j,n}\big(\psi^*_{j\ell}\circ\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}\big)f_{\ell,j,n_\ell}\big)'\big\|^{1/2}_p,$$

where $1/p+1/p'=1$. Hence the desired estimate (5.43) follows by (5.44) and the $L^{p'}$ estimate for the functions. □

Lemma 5.15. *For any tree* $\mathbf{T}$ *in* $\mathbf{S}$*, let*

$$\widetilde{\Delta}_{\ell,m}(\mathbf{T})(x) = \left(\sum_{(j,n)\in\mathbf{T}}\big|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x)f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big|^2\right)^{1/2}. \tag{5.46}$$

Then for $\ell=1$ *we have*

$$\big\|\Delta_{\ell,m}(\mathbf{T})\big\|_{\mathrm{BMO}} \le C\mathrm{size}^*_\ell(\mathbf{T}), \tag{5.47}$$

$$\big\|\widetilde{\Delta}_{\ell,m}(\mathbf{T})\big\|_{\mathrm{BMO}} \le Cm\,\mathrm{size}^*_\ell(\mathbf{T}), \tag{5.48}$$

$$\left\|\widetilde{\Delta}_{\ell,m}(\mathbf{T})\right\|_q \le Cm^{1-2/q}\mathrm{size}^*_\ell(\mathbf{T})|I_{\mathbf{T}}|^{1/q}, \tag{5.49}$$

where $q \ge 2$ and C is a constant independent of $\mathbf{T}, \mathbf{S}, L_1, L_2, M_1, M_2, f_\ell, n_\ell$.

Proof. (5.47) can be obtained routinely as we did for Lemma 4.14. We omit the details. We need only prove (5.48). (5.49) is a simple consequence of (5.47), (5.48) and an interpolation argument.

Clearly by a change of variables $\left\|\Delta_{\ell,m}(\mathbf{T})\right\|_2 = \left\|\widetilde{\Delta}_{\ell,m}(\mathbf{T})\right\|_2$. Thus (5.47) and an interpolation yield

$$\left\|\widetilde{\Delta}_{\ell,m}(\mathbf{T})\right\|_2 \le C\mathrm{size}^*_\ell(\mathbf{T})|I_{\mathbf{T}}|^{1/2}. \tag{5.50}$$

Let $\mu = \mathrm{size}^*_\ell(\mathbf{T})$. Let J be a dyadic interval and $\mathbf{T}_J = \{s \in \mathbf{T} : I_s \subseteq 3J\}$. We then dominate $\inf_c \int_J \left|\Delta_\ell(\mathbf{T})(x) - c\right|dx$ by a sum of the following three parts.

$$\int_J \left(\sum_{s\in\mathbf{T}_J} \left|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\right|^2 \right)^{1/2} dx,$$

$$\int_J \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le|J|}} \left|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\right|^2 \right)^{1/2} dx,$$

and

$$\inf_c \int_J \left| \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|>|J|}} \left|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\right|^2 \right)^{1/2} - c \right| dx.$$

$\mathbf{T}_J$ can be decomposed to a union of trees $\mathbf{T}_{J,k}$'s such that the time intervals $I_{\mathbf{T}_{J,k}}$'s are disjoint and all of them are contained in $3J$. Using the Cauchy–Schwarz inequality, the first part is estimated by

$$\left(\sum_k \left\|\widetilde{\Delta}_{\ell,m}(\mathbf{T}_{J,k})\right\|_2^2\right)^{1/2} |J|^{1/2}.$$

Appying (5.50), we dominated the first part by $C\mu|J|$.

Since $p \le 2$ we estimate the second part by

$$\left\| \left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le|J|}} \left|(\widetilde{\mathbf{1}}^*_{j,n} \circ \widetilde{\mathrm{Tr}}_{\ell,j,m})\widetilde{\psi}^*_{j\ell}(f_{\ell,j,n_\ell} \circ \widetilde{\mathrm{Tr}}_{\ell,j,m})\right|^2 \right)^{1/2} \right\|_{L^p(J)} |J|^{1-\frac{1}{p}}$$

$$\le \left(\sum_{\substack{s\in \mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le |J|}} \big\| (\widetilde{\mathbf{1}}^*_{j,n}\circ \widetilde{\mathrm{Tr}}_{\ell,j,m})\widetilde{\psi}^*_{j\ell}(f_{\ell,j,n_\ell}\circ \widetilde{\mathrm{Tr}}_{\ell,j,m})\big\|^p_{L^p(J)} \right)^{1/p} |J|^{1-\frac{1}{p}}$$

$$\le \left(\sum_{\substack{s\in \mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le |J|}} \frac{C\big\| (\mathbf{1}^{**}_{j,n}\circ \widetilde{\mathrm{Tr}}_{\ell,j,m})\psi^*_{j\ell}(f_{\ell,j,n_\ell}\circ \widetilde{\mathrm{Tr}}_{\ell,j,m})\big\|^p_p}{\big(1+|I_s|^{-1}\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\big)^{100}} \right)^{1/p} |J|^{1-\frac{1}{p}}$$

$$\le \mu \left(\sum_{\substack{s\in \mathbf{T}\setminus\mathbf{T}_J \\ |I_s|\le |J|}} \frac{C|I_s|}{\big(1+|I_s|^{-1}\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\big)^{100}} \right)^{1/p} |J|^{1-\frac{1}{p}},$$

where $\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s)$ is the interval $\{\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(x) : x\in I_s\}$. Observe that if $|I_s|\le 2^{-m-10}|J|$ and $s\in \mathbf{T}\setminus\mathbf{T}_J$, then $\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\sim \mathrm{dist}(J,I_s)$. Thus summing for all s in this case, we get the desired estimate $C\mu|J|$. In the remaining case, there are only $10m$ different scales for $|I_s|$'s since s's satisfy $2^{-m-10}|J|<|I_s|\le |J|$. The worst situation is that when $\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s)\cap J\neq\emptyset$, because otherwise $\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))$ can be replaced by $\mathrm{dist}(\partial J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))$ and thus the desired estimate follows. But in this situation, $\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s)$ must be a subset of $3J$ since $|I_s|\le |J|$. For all $\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s)\subset 3J$ with a fixed scale, the sum of $|I_s|$'s is no more than $3|J|$. Summing for at most $10m$ different scales, we thus get the upper bound $Cm\mu|J|$. Hence the second part is dominated by $Cm\mu|J|$.

The third part is estimated by

$$\left(\inf_c \int_J \left| \left(\sum_{\substack{s\in \mathbf{T}\setminus\mathbf{T}_J \\ |I_s|>|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x)f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\big|^2 \right)^{1/2} - c \right|^2 dx \right)^{1/2} \cdot |J|^{1/2}$$

$$\le \left(\inf_c \int_J \left| \sum_{\substack{s\in \mathbf{T}\setminus\mathbf{T}_J \\ |I_s|>|J|}} \big|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x)f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\big|^2 - c \right| dx \right)^{1/2} \cdot |J|^{1/2}$$

$$\leq C\left(\int_J \sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}} \left|\left(\left|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\right|^2\right)'\right| dx\right)^{1/2} |J|,$$

which is dominated by a sum of following two terms,

$$R_1 = C\left(\int_J \sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}} |I_s|^{-1}\left|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\right|^2 dx\right)^{1/2} |J|,$$

and

$$R_2 = C\left(\int_J \sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}} \left|\widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) f_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\right| \left|G_{\ell,j,m}(x)\right| dx\right)^{1/2} |J|,$$

where $G_{\ell,j,m}$ is the function defined by

$$G_{\ell,j,m}(x) = \widetilde{\mathbf{1}}^*_{j,n}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)\widetilde{\psi}^*_{j\ell}(x) Df_{\ell,j,0}\big(\widetilde{\mathrm{Tr}}_{\ell,j,m}(x)\big)$$

By Lemma 5.14, we see that for any $q \geq p$,

$$\big\|(\mathbf{1}^{**}_{j,n}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m})\psi^*_{j\ell}(f_{\ell,j,n_\ell}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m})\big\|_q \leq C\mu|I_s|^{1/q}.$$

Thus, by Hölder's inequality, the first term R_1 is estimated by

$$C\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}} \frac{|I_s|^{-1}\big\|(\mathbf{1}^{**}_{j,n}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m})\psi^*_{j\ell}(f_{\ell,j,n_\ell}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m})\big\|_4^2 |J|^{1/2}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\big)^{100}}\right)^{1/2} |J|$$

$$\leq C\mu\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}} \frac{|I_s|^{-1/2}|J|^{1/2}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\big)^{100}}\right)^{1/2} |J| \;\leq\; C\mu|J|.$$

It is obvious by the fact $2^{k_{j\ell}} \sim 2^{k_j}$ when $\ell = 1$ and the definition of the seminorm that

$$\|G_{\ell,j,m}\|_p \leq \|f_\ell\|_{j,n,m}|I_s|^{1/p-1}. \tag{5.51}$$

Thus the second term R_2 is estimated by

$$C\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\big\|\big(\widetilde{\mathbf{1}}^*_{j,n}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m}\big)\widetilde{\psi}^*_{j\ell}\big(f_{\ell,j,0}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m}\big)\big\|_{L^{p'}(J)}\big\|G_{\ell,j,m}\big\|_p\right)^{1/2}\quad|J|$$

$$\le C\left(\mu\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\frac{|I_s|^{\frac{1}{p}-1}\big\|(\mathbf{1}^{**}_{j,n}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m})\psi^*_{j\ell}(f_{\ell,j,n_\ell}\circ\widetilde{\mathrm{Tr}}_{\ell,j,m})\big\|_{p'+1}|J|^{\frac{1}{p'(p'+1)}}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\big)^{100}}\right)^{1/2}$$

$$\cdot|J|$$

$$\le C\mu\left(\sum_{\substack{s\in\mathbf{T}\setminus\mathbf{T}_J\\ |I_s|>|J|}}\frac{|I_s|^{-\frac{1}{p'(p'+1)}}|J|^{\frac{1}{p'(p'+1)}}}{\big(1+|I_s|^{-1}\mathrm{dist}(J,\widetilde{\mathrm{Tr}}^{-1}_{\ell,j,m}(I_s))\big)^{100}}\right)^{1/2}\quad|J|\ \le\ C\mu|J|.$$

This completes the proof of (5.48). □

Lemma 5.16. *Let* $\mathbf{T}$ *be a tree in* $\mathbf{S}$ *and* $\mathbf{P}$ *be a subset of* $\mathbf{S}$. *Suppose that* $\mathbf{P}\cap\mathbf{T}=\emptyset$ *and* $\mathbf{T}$ *is a maximal tree in* $\mathbf{P}\cup\mathbf{T}$. *Then we have*

$$\big|\Lambda_{\mathbf{P}\cup\mathbf{T},\Omega,m}(f_1,f_2,f_3)-\Lambda_{\mathbf{P},\Omega,m}(f_1,f_2,f_3)\big|\le\big|\Lambda_{\mathbf{T},\Omega,m}(f_1,f_2,f_3)\big|+Cm\prod_{\ell=1}^{2}\mathrm{size}^*_\ell(\mathbf{T}\cup\mathbf{P})|I_{\mathbf{T}}|,\tag{5.52}$$

where C *is independent of* $f_1,f_2,f_3,L_1,L_2,M_1,M_2,\mathbf{P},\ \mathbf{T}$.

The proof is similar to the proof of Lemma 5.16. We omit the details and leave it as an exercise to the readers.

5.4. Proof of Lemma 5.10. It is easy to prove a size estimate for the trilinear form on a single tree, that is, for any tree $\mathbf{T}$,

$$\big|\Lambda_{\mathbf{T},\Omega,m}(f_1,f_2,f_3)\big|\le Cm^{2/p-1}\prod_{\ell=1}^{2}\mathrm{size}^*_\ell(\mathbf{T})|I_{\mathbf{T}}|,\tag{5.53}$$

where C is independent of $L_1,L_2,M_1,M_2,m,f_1,f_2,f_3,\mathbf{T}$.

In fact, by Hölder's inequality, we estimate $|\Lambda_{\mathbf{T},\Omega,m}|$ by

$$\big\|\widetilde{\Delta}^*_{1,m}(\mathbf{T})\big\|_{p'}\big\|\Delta^*_{2,m}(\mathbf{T})\big\|_p.$$

By (5.49) and the definition of size, we obtain (5.53) immediately.

Lemma 4.15 still holds for our new sizes of trees and $\mathbf{S}$. Applying this organization lemma inductively for $\mathbf{S}$, we decompose

$$\mathbf{S}=\bigcup_\sigma\mathbf{S}_\sigma,\tag{5.54}$$

where σ ranges over all possible dyadic numbers, $\mathbf{S}_\sigma = \cup_{\mathbf{T}\in\mathcal{F}_\sigma}\mathbf{T}$ such that $\mathcal{F}_\sigma$ is a collection of maximal trees with

$$\operatorname{count}(\mathbf{S}_\sigma) \le C\sigma^{-p}, \tag{5.55}$$

and

$$\operatorname{size}^*_\ell(\mathbf{S}_\sigma) \le \sigma|F_\ell|^{1/p} \tag{5.56}$$

holds for all $\ell \in \{1,2\}$.

By (5.40), the upper bound in the size estimates for $\mathbf{S}_\sigma$ can be sharpened by

$$\operatorname{size}^*_\ell(\mathbf{S}_\sigma) \le \min\{1, 2^{\beta_\ell m}|F_\ell|^{1/p}, \sigma|F_\ell|^{1/p}\}. \tag{5.57}$$

Hence by Lemma 5.16 and (5.53) we estimate $\Lambda_{\mathbf{S},\Omega,m}$ by

$$\sum_\sigma \sum_{\mathbf{T}\in\mathcal{F}_\sigma} m \prod_{\ell=1}^{2} \operatorname{size}^*_\ell(\mathbf{S}_\sigma)|I_{\mathbf{T}}|.$$

Applying (5.57) and (5.55), we thus dominate $\big|\Lambda_{\mathbf{S},\Omega,m}(f_1,f_2,f_3)\big|$ by

$$Cm\sum_\sigma \sigma^{-p}\min\{1, 2^m|F_1|^{1/p}, \sigma|F_1|^{1/p}\}\min\{1, 2^{\varepsilon^2 m}|F_2|^{1/p}, \sigma|F_2|^{1/p}\}, \tag{5.58}$$

which clearly implies (5.33). Therefore we have completed the proof of Lemma 5.10.

References

[1] CALDERÓN, ALBERTO P. Commutators of singular integral operators. *Proc. Natl. Acad. Sci. USA* **53** (1965) 1092–1099. MR0177312 (31 #1575), Zbl 0151.16901.

[2] CARBERY, ANTHONY; CHRIST, MICHAEL; WRIGHT, JAMES. Multidimensional van der Corput and sublevel set estimates. *J. Amer. Math. Soc.* **12** (1999), 981–1015. MR1683156 (2000h:42010), Zbl 0938.42008.

[3] CHRIST, MICHAEL. Hilbert transforms along curves. I. Nilpotent groups. *Ann. of Math* (2) **122** (1985) 575–596. MR0819558 (87f:42039a), Zbl 0593.43011.

[4] CHRIST, MICHAEL. Hilbert transforms along curves. II. A flat case. *Duke Math. J.* **52** (1985) 887–894. MR0816390 (87f:42039b), Zbl 0677.42011.

[5] CHRIST, MICHAEL; LI, XIAOCHUN; TAO, TERENCE; THIELE, CHRISTOPH. On multilinear oscillatory integrals, nonsingular and singular. *Duke Math. J.* **130** (2005) 321–351. MR2181092 (2006g:42012), Zbl pre05011512.

[6] FAN, DASHAN; LI, XIAOCHUN. A bilinear oscillatory integrals along parabolas. arXiv:0709.2907.

[7] FURSTENBERG, HILLEL. Nonconventional ergodic averages. *The legacy of John von Neumann* (Hempstead, NY, 1988), 43–56. Proceedings of Symposia in Pure Math., 50, *Amer. Math. Soc., Providence, RI*, 1990. MR1067751 (91i:28011), Zbl 0711.28006.

[8] GRAFAKOS, LOUKAS; LI, XIAOCHUN. The disc as a bilinear multiplier. *Amer. J. Math.* **128** (2006) 91–119. MR2197068 (2006i:42012), Zbl pre05015888.

[9] GRAFAKOS, LOUKAS; LI, XIAOCHUN. Uniform bounds for the bilinear Hilbert transforms. I. *Ann. of Math* (2) **159** (2004) 889–993. MR2113017 (2006e:42011), Zbl 1071.44004.

[10] DUOANDIKOETXEA, JAVIER; RUBIO DE FRANCIA, JOSÉ L. Maximal and singular integral operators via Fourier transform estimates. *Invent. Math.* **84** (1986) 541–561. MR0837527 (87f:42046), Zbl 0568.42012.

[11] LACEY, MICHAEL T. The bilinear maximal functions map into L^p for $2/3 < p \leq 1$. *Ann. of Math.* (2) **151** (2000) 35–57. MR1745019 (2001b:42015), Zbl 0967.47031.

[12] LACEY, MICHAEL T.; THIELE, CHRISTOPH M. L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. *Ann. Math.* **146** (1997), 693–724. MR1491450 (99b:42014), Zbl 0914.46034.

[13] LACEY, MICHAEL T.; THIELE, CHRISTOPH M. On Calderón's conjecture. *Ann. of Math.* **149** (1999) 475–496. MR1689336 (2000d:42003), Zbl 0934.42012.

[14] LI, XIAOCHUN. Uniform bounds for the bilinear Hilbert transforms. II. *Rev. Mat. Iberoamericana* **22** (2006), no.3, 1067–1124. MR2320411 (2008c:42014).

[15] MUSCALU, CAMIL; TAO, TERENCE; THIELE, CHRISTOPH M. Multilinear operators given by singular multipliers. *J. Amer. Math. Soc.* **15** (2002) 469–496. MR1887641 (2003b:42017), Zbl 0994.42015.

[16] MUSCALU, CAMIL; TAO, TERENCE; THIELE, CHRISTOPH M. Uniform estimates on paraproducts. *Journal d'Analyse de Jerusalem* **87** (2002) 369–384. MR1945289 (2004a:42023), Zbl 1043.42012.

[17] MUSCALU, CAMIL; TAO, TERENCE; THIELE, CHRISTOPH M. Uniform estimates for multilinear operators with modulation symmetry. *Journal d'Analyse de Jerusalem* **88** (2002) 255–309. MR1979774 (2004d:42032), Zbl 1041.42013.

[18] NAGEL, ALEXANDER; VANCE, JAMES; WAINGER, STEPHEN; WEINBERG, DAVID. Hilbert transforms for convex curves. *Duke. Math. J.* **50** (1983) 735–744. MR0833365 (88b:42025), Zbl 0524.44001.

[19] SOGGE, CHRISTOPHER D. Fourier integrals in classical analysis. Cambridge Tracts in Mathematics, 105. *Cambridge University Press*, 1993. MR1205579 (94c:35178), Zbl 0783.35001.

[20] STEIN, ELIAS M. Harmonic analysis, real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. *Princeton University Press, Princeton, NJ*, 1993. MR1232192 (95c:42002), Zbl 0821.42001.

[21] STEIN, ELIAS M. Oscillatory integrals related to Radon-like transforms. Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993). *J. Fourier Analysis Appl.* (1995), Special issue, 535–551. MR1364908 (96i:42013), Zbl 0971.42009.

[22] STEIN, ELIAS M.; WAINGER, STEPHEN. Problems in harmonic analysis related to curvature. *Bull. Amer. Math. Soc.* **84** (1978) 1239–1295. MR0508453 (80k:42023), Zbl 0393.42010.

[23] THIELE, CHRISTOPH M. A uniform estimate. *Ann. of Math.* (2) **157** (2002) 1–45. MR1933076 (2003i:47036), Zbl 1038.42019.

XIAOCHUN LI, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL, 61801, USA
xcli@math.uiuc.edu

This paper is available via http://nyjm.albany.edu/j/2008/14-6.html.