Acta Mathematica Sinica, English Series
Jan., 2018, Vol. 34, No. 1, pp. 29–41
Published online: February 6, 2017
https://doi.org/10.1007/s10114-017-6415-9
http://www.ActaMath.com

Acta Mathematica Sinica, English Series

On a Hybrid of Bilinear Hilbert Transform and Paraproduct

Dong DONG **Xiao Chun LI**
Mathematics Department, University of Illinois at Urbana-Champaign, Urbana, IL 61801*, USA*
E-mail: *ddong3@illinois.edu* *xcli@math.uiuc.edu*

Abstract We introduce a class of tri-linear operators that combine features of the bilinear Hilbert transform and paraproduct. For two instances of these operators, we prove boundedness property in a large range $D = \{(p_1, p_2, p_3) \in \mathbb{R}^3 : 1 < p_1, p_2 < \infty, \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2}, 1 < p_3 < \infty\}$.

Keywords Bilinear Hilbert transform, paraproduct, tri-linear operators

MR(2010) Subject Classification 42A50, 47G10, 42B99

1 Introduction

We introduce the following class of tri-linear operators which can be viewed as a hybrid of the bilinear Hilbert transform (BHT for short) and paraproduct.

$$T(f_1, f_2, f_3)(x) = \sum_{k\in\mathbb{Z}} H_k(f_1, f_2)(x) f_{3k}(x), \quad x \in \mathbb{R}, \tag{1.1}$$

where

$$\begin{aligned} H_k(f_1, f_2)(x) &= \iint_{\mathbb{R}^2} \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2) \mathrm{e}^{2\pi \mathrm{i}(\xi_1+\xi_2)x} \widehat{\Phi_1}\left(\frac{\xi_1-\xi_2}{2^k}\right) d\xi_1 d\xi_2 \\ &= \int_{\mathbb{R}} f_1(x-t) f_2(x+t)\Phi_1(2^k t)2^k\, dt, \end{aligned} \tag{1.2}$$

and

$$f_{3k}(x) = \int_{\mathbb{R}} \widehat{f_3}(\xi) \mathrm{e}^{2\pi \mathrm{i}\xi x} \widehat{\Phi_2}\left(\frac{\xi}{2^{\alpha k}}\right) d\xi. \tag{1.3}$$

Here $\alpha \in \mathbb{R}$ and various conditions (about smoothness, support, etc) can be imposed on the cut-off functions Φ_1 and Φ_2.

In the study of BHT $B(f_1, f_2)(x) = \int f_1(x-t) f_2(x+t)\frac{dt}{t}$, the usual set up is to decompose the kernel $\frac{1}{t} = \sum_{k\in\mathbb{Z}} \Phi(2^k t)2^k$ using a nice function Φ (e.g. [4, 5]). The bilinear operator $H_k(f_1, f_2)$ then appears naturally as the piece of the decomposition at scale k. Meanwhile, a function with a cut-off in frequency as f_{3k} is a fundamental unit in many paraproducts (e.g. [1, 7, 11]). These two facts explain why T is a hybrid of BHT and paraproduct.

We became interested in such type of operators when it came out in our study of tri-linear Hilbert transform along curves such as operator $T_C(f_1, f_2, f_3)(x) = \int f_1(x-t) f_2(x+t) f_3(x-t^2)\frac{dt}{t}$, whose boundedness is still unknown. We may compare T_C with its bilinear analogue $B_C(f_1, f_2)(x) = \int f_1(x-t) f_2(x-t^2)\frac{dt}{t}$, whose boundedness is proved in [8] and [9]. B_C

Received August 26, 2016, accepted December 6, 2016

contains a paraproduct of the form $P(f_1, f_2)(x) = \sum_{k\in\mathbb{Z}} f_{1k}(x) f_{2k}(x)$ where f_{1k} and f_{2k} are defined the same way as f_{3k} in (1.3). Proving boundedness of P is the first step to study B_C (see [7]). T plays the same role in T_C as P in B_C.

Let us first investigate a simple case of T. Let $D = \{(p_1, p_2, p_3) \in \mathbb{R}^3 : 1 < p_1, p_2 < \infty, \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2}, 1 < p_3 < \infty\}$. We have

Proposition 1.1 *Let Φ_1 be a smooth function satisfying $\widehat{\Phi_1}(0) = 0$ and Φ_2 be an integrable C^1 function satisfying $\widehat{\Phi_2}(0) = 0$ and $|\Phi_2(x)| + |\Phi_2'(x)| \lesssim (1+|x|)^{-2}$. Assume $\alpha \neq 0$. Then the operator T defined by (1.1)–(1.3) is bounded from $L^{p_1} \times L^{p_2} \times L^{p_3}$ to L^p, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}$, whenever $(p_1, p_2, p_3) \in D$.*

The proof of Proposition 1.1 is straightforward. Under the hypotheses of the proposition, the two cut-off functions $\widehat{\Phi_1}$ and $\widehat{\Phi_2}$ are supported away from 0. Hence Proposition 1.1 follows from the Cauchy–Schwartz inequality, Hölder inequality, Littlewood–Paley theorem and its bilinear analogue (Theorem 6 in [2]).

T becomes more interesting and more difficult to handle when $\widehat{\Phi_2}$ is supported near 0. This paper is mainly devoted to this case.

Theorem 1.2 *Let Φ_1 and Φ_2 be a smooth functions satisfying $\operatorname{supp} \widehat{\Phi_1} \subset [L, L+1]$ for some large L and $\operatorname{supp} \widehat{\Phi_2} \subseteq [-1, 1]$. Assume $\alpha = 1$. Then the operator T defined by (1.1)–(1.3) is bounded from $L^{p_1} \times L^{p_2} \times L^{p_3}$ to L^p, $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}$, whenever $(p_1, p_2, p_3) \in D$.*

We believe that Theorem 1.2 holds for a larger class of operators and we have the following conjecture.

Conjecture 1.3 *The assumption $\alpha = 1$ in Theorem 1.2 can be dropped.*

Unfortunately our proof of Theorem 1.2 relies on the homogeneity of the scales and thus can not cover the $\alpha \neq 1$ case. A suitable use of the telescoping trick as in [3, 6] and [12] is expected to solve the conjecture.

We provide the proof of Theorem 1.2 in the rest of the paper. In Section 2, we simplify the study of T to a model form. Then we prove the boundedness of the model form in Section 3 using three lemmas whose proofs can be found in Sections 4–6. Throughout the paper we use C to denote a positive constant whose value may change from line to line. We may add one or more subscripts to C to emphasize dependence of C. $A \lesssim B$ is short for $A \le CB$ and $A \lesssim_N B$ means $A \le C_N B$. If $A \lesssim B$ and $B \lesssim A$, then we write $A \simeq B$. χ_E will be used to denote the characteristic function of the set E.

2 Reduction to the Model Form

The goal of this section is to reduce Theorem 1.2 to the study of a model form using *wave packet decomposition*. Let $\mathcal{S}(\mathbb{R})$ denote the class of Schwartz functions on $\mathbb{R}$. Given $f_j \in \mathcal{S}(\mathbb{R})$, $j \in \{1, 2, 3, 4\}$, consider the 4-linear form Λ associated with T

$$\begin{aligned}\Lambda(f_1, f_2, f_3, f_4) &:= \int T(f_1, f_2, f_3)(x) \overline{f_4}(x)\, dx \\ &= \sum_{k\in\mathbb{Z}} \iiint \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)\widehat{\Phi_1}\left(\frac{\xi_1 - \xi_2}{2^k}\right)\widehat{\Phi_2}\left(\frac{\xi_3}{2^k}\right)\overline{\widehat{f_4}}(\xi_1 + \xi_2 + \xi_3) d\xi_1 d\xi_2 d\xi_3. \qquad (2.1)\end{aligned}$$

Theorem 1.2 is then equivalent to

Theorem 2.1 *Under the hypotheses of Theorem* 1.2*, the* 4*-linear form* Λ *satisfies*

$$|\Lambda(f_1, f_2, f_3, f_4)| \lesssim \|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}\|f_4\|_{p_4}$$

for any $(p_1, p_2, p_3) \in D$, $\frac{1}{p_4} = 1 - \frac{1}{p_1} - \frac{1}{p_2} - \frac{1}{p_3}$, *and* $f_i \in \mathcal{S}(\mathbb{R})$, $i \in \{1, 2, 3, 4\}$.

We plan to localize each function in both time and frequency spaces. The following process is called *wave packet decomposition.* Pick a $\psi \in \mathcal{S}(\mathbb{R})$ such that $\operatorname{supp} \widehat{\psi} \subseteq [0, 1]$ and $\sum_{l\in\mathbb{Z}} \widehat{\psi}(\xi - \frac{l}{2}) = 1$ for any $\xi \in \mathbb{R}$. Define $\widehat{\psi_{k,l}}(\xi) := \widehat{\psi}(\frac{\xi - 2^{k-1}l}{2^k})$ for $(k, l) \in \mathbb{Z}^2$. Also pick a non-negative $\varphi \in \mathcal{S}(\mathbb{R})$ with $\operatorname{supp} \widehat{\varphi} \subset [-1, 1]$ and $\widehat{\varphi}(0) = 1$. Let $\varphi_k(x) := 2^k\varphi(2^kx)$, $k \in \mathbb{Z}$. For every $(k, n) \in \mathbb{Z}^2$, denote $I_{k,n} := [2^{-k}n, 2^{-k}(n + 1))$. Then for each scale $k \in \mathbb{Z}$ and function $f \in \mathcal{S}(\mathbb{R})$, we have

$$f = \sum_{(n,l)\in\mathbb{Z}^2} f_{k,n,l}, \tag{2.2}$$

where

$$f_{k,n,l}(x) := \chi^*_{I_{k,n}}(x) f * \psi_{k,l}(x), \text{ and} \tag{2.3}$$

$$\chi^*_I(x) := \chi_I * \varphi_k(x) \text{ for any interval } I. \tag{2.4}$$

Note that $\operatorname{supp} \widehat{f_{k,n,l}} \subset [2^k(\frac{l}{2} - 1), 2^k(\frac{l}{2} + 2)]$ and $f_{k,n,l}$ is essentially supported on $I_{k,n}$ in the sense that

$$|f_{k,n,l}(x)| \lesssim_{N,M} \left(1 + \frac{\operatorname{dist}(x, I_{k,n})}{|I_{k,n}|}\right)^{-N} \frac{1}{|I_{k,n}|} \int |f(y)|\left(1 + \frac{|x - y|}{|I_{k,n}|}\right)^{-M} dy. \tag{2.5}$$

So $f_{k,n,l}$ is indeed well-localized and the wave packet decomposition is complete.

Now we can apply the decomposition (2.2) to all the four functions in (2.1) and obtain

$$\begin{aligned}\Lambda(f_1, f_2, f_3, f_4) = \sum_{\substack{k\in\mathbb{Z}\\(n_1,n_2,n_3,n_4)\in\mathbb{Z}^4\\(l_1,l_2,l_3,l_4)\in\mathbb{Z}^4}} &\iiint \widehat{f_{1,k,n_1,l_1}},(\xi_1)\widehat{f_{2,k,n_2,l_2}}(\xi_2)\widehat{f_{3,k,n_3,l_3}}(\xi_3)\\ &\cdot \widehat{\Phi_1}\left(\frac{\xi_1 - \xi_2}{2^k}\right)\widehat{\Phi_2}\left(\frac{\xi_3}{2^k}\right)\overline{\widehat{f_{4,k,n_4,l_4}}}(\xi_1 + \xi_2 + \xi_3)\, d\xi_1 d\xi_2 d\xi_3.\end{aligned}$$

Here f_{j,k,n_j,l_j}, $j \in \{1, 2, 3, 4\}$, is defined as in (2.3) for $f = f_j$. By the support of functions, each term in the sum is non-zero only when

$$\begin{cases} \xi_i \in \left[2^k\left(\frac{l_i}{2} - 1\right), 2^k\left(\frac{l_i}{2} + 2\right)\right] \quad \text{for } i = 1, 2, 3;\\ \xi_1 - \xi_2 \in [L2^k, (L + 1)2^k), \quad |\xi_3| \lesssim 2^k;\\ \xi_1 + \xi_2 + \xi_3 \in \left[2^k\left(\frac{l_4}{2} - 1\right), 2^k\left(\frac{l_4}{2} + 2\right)\right]. \end{cases}$$

These imply that $|l_2 - (l_1 - 2L)| \lesssim 1$, $|l_3| \lesssim 1$, and $|l_4 - (2l_1 - 2L)| \lesssim 1$. In other words, among the four parameters l_1, l_2, l_3, l_4, only one is free (say l_1). Without loss of generality we can fix a dependence relation between l_2, l_3, l_4 and l_1, and think of $\Lambda(f_1, f_2, f_3, f_4)$ as

$$\sum_{\substack{k,l_1\\n_1,n_2,n_3,n_4}} \iiint \widehat{f_{1,k,n_1,l_1}},(\xi_1)\widehat{f_{2,k,n_2,l_2}}(\xi_2)\widehat{f_{3,k,n_3,l_3}}(\xi_3)$$

$$\cdot\, \widehat{\Phi_1}\left(\frac{\xi_1-\xi_2}{2^k}\right)\widehat{\Phi_2}\left(\frac{\xi_3}{2^k}\right)\overline{\widehat{f_{4,k,n_4,l_4}}}(\xi_1+\xi_2+\xi_3)\,d\xi_1 d\xi_2 d\xi_3.$$

We then drop the cut-off functions in the above expression by the Fourier expansion trick and ignore fast decay terms so that $\Lambda(f_1,f_2,f_3,f_4)$ becomes essentially as

$$\sum_{\substack{k,l_1\\ n_1,n_2,n_3,n_4}} \iiint f_{1,k,n_1,l_1}(x) f_{2,k,n_2,l_2}(x) f_{3,k,n_3,l_3}(x)\overline{f_{4,k,n_4,l_4}}(x)\,dx.$$

Since f_{j,k,n_j,l_j} is almost supported in $I_{k,n_j}=[2^{-k}n_j, 2^{-k}(n_j+1))$, there is no harm to assume $n_1=n_2=n_3=n_4$ due to the fast decay in other cases. Therefore the original 4-linear form has been simplified to the following model form (we still use Λ to denote the model 4-linear form by abuse of notation):

$$\Lambda(f_1,f_2,f_3,f_4)=\sum_{(k,n,l)\in\mathbb{Z}^3}\int\prod_{j=1}^{4} f_{j,k,n,l_j}(x)\,dx. \tag{2.6}$$

Here $l_1=l$, $l_2=l-2L$, $l_3=0$ and $l_4=2l-2L$. This model form is similar to that of BHT. [10] contains an elegant proof of *Lacey–Thiele* Theorem on BHT and thus we plan to adopt the ideas there to handle Λ. Of course we have to take extra care to the term f_{3k}.

We will prove directly that T is of restricted weak type (see [10, p. 312] for the definition) when (p_1,p_2,p_3) is in a smaller range $D_0 := \{(p_1,p_2,p_3) : 1<p_1,p_2<2, \frac{1}{p_1}+\frac{1}{p_2}<\frac{3}{2}, p_3\in(1,\infty)\}$. More precisely,

Theorem 2.2 *Let $(p_1,p_2,p_3)\in D_0$. For any measurable sets F_1,F_2,F_3,F of finite measure, there exists measurable set $F'\subseteq F$ with $|F'|\ge\frac{1}{2}|F|$ such that Λ defined in* (2.6) *satisfy*

$$|\Lambda(f_1,f_2,f_3,f_4)|\lesssim |F_1|^{\frac{1}{p_1}}|F_2|^{\frac{1}{p_2}}|F_3|^{\frac{1}{p_3}}|F'|^{\frac{1}{p'}} \tag{2.7}$$

for every $|f_1|\le\chi_{F_1}$, $|f_2|\le\chi_{F_2}$, $|f_3|\le\chi_{F_3}$ and $|f_4|\le\chi_{F'}$. Here $\frac{1}{p'}:=1-(\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3})$.

Theorem 2.1 follows from Theorem 2.2 by interpolation and the fact that three adjoints of T can be handled by the same method (see the Appendix of [10] for a similar interpolation treatment in the context of BHT). To prove Theorem 2.2, we pick up an arbitrary finite subset $S\subset\mathbb{Z}^3$ and aim to obtain (2.7) for

$$\Lambda_S(f_1,f_2,f_3,f_4):=\sum_{(k,n,l)\in S}\int\prod_{j=1}^{4} f_{j,k,n,l_j}(x)\,dx, \tag{2.8}$$

provided the bound does not depend on the set S. We can also assume $|F|=1$ by dilation invariance. Next we make the geometric structure of Λ_S clearer. To each tuple $s=(k,n,l)\in\mathbb{Z}^3$, we plan to assign a time-interval $I_s:=I_{k,n}$ and four frequency-intervals ω_{s_j}, $j\in\{1,2,3,4\}$, representing the localization of functions in the time-frequency space. More precisely, I_s and ω_{s_j}'s satisfy:

$$f_{j,k,n,l_j}(x) \text{ is dominated by} \tag{2.9}$$
$$C_{N,M}\left(1+\frac{\operatorname{dist}(x,I_s)}{|I_s|}\right)^{-N}\frac{1}{|I_s|}\int|f_j(y)|\left(1+\frac{|x-y|}{|I_s|}\right)^{-M}dy;$$

$$\text{The Fourier transform of } f_{j,k,n,l_j} \text{ is supported on } \omega_{s_j}. \tag{2.10}$$

We will call $s = (k, n, l)$ a 4-tile (or simply a tile) as it corresponds to 4 single-tiles $s_j := I_s \times \omega_{s_j}$, $j \in \{1, 2, 3, 4\}$. Write $f_{s_j} := f_{k,n,l_j}$ and $f_{j,s_j} := f_{j,k,n,l_j}$ for simplicity.

We can take finitely many sparse subsets of S and transform ω_{s_j}'s by fixed affine mappings if needed (since only relative locations of Fourier supports matter) so that I_s and ω_{s_j}'s enjoy nice geometric properties as follows:

$$|\omega_{s_1}| = |\omega_{s_2}| = |\omega_{s_3}| = |\omega_{s_4}| = C|I_s|^{-1}; \tag{2.11}$$

$$\text{dist}(\omega_{s_1}, \omega_{s_2}) = \text{dist}(\omega_{s_2}, \omega_{s_4}) = |\omega_{s_1}|; \tag{2.12}$$

$$c(\omega_{s_1}) > c(\omega_{s_2}) > c(\omega_{s_4}), \quad \text{where } c(I) \text{ is the center of the interval } I; \tag{2.13}$$

$$\{I_s\}_{s\in S} \text{ is a grid (defined below);} \tag{2.14}$$

$$\{\omega_{s_1} \cup \omega_{s_2} \cup \omega_{s_4}\}_{s\in S} \text{ is a gird;} \tag{2.15}$$

$$\omega_{s_i} \subsetneq J \text{ for some } i \in \{1, 2, 4\}, \quad J := \omega_{s'_1} \cup \omega_{s'_2} \cup \omega_{s'_4}, \quad s' \in S \Rightarrow \tag{2.16}$$
$$\omega_{s_j} \subseteq J \text{ for all } j \in \{1, 2, 4\}.$$

Here a grid is defined as a set of intervals having the property that if two different elements intersect then one must contain the other and the larger interval is at least twice as long as the smaller one. See [4] for a detailed construction of the time and frequency intervals. From now on we fix a finite set of tiles $S \subset \mathbb{Z}^3$ and assume the tiles satisfy (2.9)–(2.16). Theorem 2.2 has been reduced to the following theorem.

Theorem 2.3 *Let $p > 1$ be arbitrary. Given any $(p_1, p_2, p_3) \in D_0$ with $p_3 \geq p$ and any sets of finite measure F_1, F_2, F_3, F with $|F| = 1$, there exists $F' \subseteq F$ with $|F'| \geq \frac{1}{2}$ such that*

$$|\Lambda_S(f_1, f_2, f_3, f_4)| \lesssim |F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}} |F_3|^{\frac{1}{p_3}}$$

for every $|f_1| \leq \chi_{F_1}$, $|f_2| \leq \chi_{F_2}$, $|f_3| \leq \chi_{F_3}$ and $|f_4| \leq \chi_{F'}$.

3 Boundedness of the Model Form

In this section we prove Theorem 2.3 using three lemmas to be proved in the remaining sections. Fix $p > 1$, $(p_1, p_2, p_3) \in D_0$ with $p_3 > p$, and measurable sets F_1, F_2, F_3, F with $|F| = 1$. Let $\mathcal{M}$ denote the maximal operator. Define the exceptional set

$$\Omega := \left(\bigcup_{j=1}^{2} \{x : \mathcal{M}(\chi_{F_j})(x) > C|F_j|\} \right) \cup \{x : \mathcal{M}(\chi_{F_3})(x) > C|F_3|^{\frac{1}{p}}\}.$$

Then $|\Omega| \leq \frac{1}{4}$ when C is large enough. Set $F' := F \setminus \Omega$ so that $|F'| \geq \frac{1}{2}$. Decompose $S = \bigcup_{d\in\mathbb{N}} S^d$, where

$$S^d := \{s \in S : 1 + \text{dist}(I_s, \Omega^c)|I_s|^{-1} \simeq 2^d\}. \tag{3.1}$$

Clearly, it suffices to obtain the estimate

$$|\Lambda_{S^d}(f_1, f_2, f_3, f_4)| \lesssim 2^{-2d} |F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}} |F_3|^{\frac{1}{p_3}} \quad \text{for } d \in \mathbb{N}. \tag{3.2}$$

The main idea is to group the tiles in S^d carefully so that there is orthogonality among the groups. The following definitions will be used in the grouping algorithm.

Definition 3.1 *Let $j \in \{1, 2, 4\}$. Given two 4-tiles s and s', we write $s_j < s'_j$ if $I_s \subseteq I_{s'}$ and $\omega_{s_j} \supseteq \omega_{s'_j}$. We call $T \subseteq S$ a j-tree if there exists a $t \in T$ such that $s_j < t_j$ for all $s \in T$. t is called the top of T and $I_T := I_t$.*

Definition 3.2 *Let $i, j \in \{1, 2, 4\}$ and $i \neq j$. A finite sequence of i-trees $T_1, T_2, \dots, T_M$ is called a chain of strongly j-disjoint trees if*

$$l_1 \neq l_2, \quad s \in T_{l_1}, \quad s' \in T_{l_2} \Rightarrow I_s \times \omega_{s_j} \neq I_{s'} \times \omega_{s'_j}; \tag{3.3}$$

$$l_1 \neq l_2, \quad s \in T_{l_1}, \quad s' \in T_{l_2}, \quad \omega_{s_j} \subsetneqq \omega_{s'_j} \Rightarrow I_{s'} \cap I_{T_{l_1}} = \emptyset; \tag{3.4}$$

$$l_1 < l_2, \quad s \in T_{l_1}, \quad s' \in T_{l_2}, \quad \omega_{s_j} = \omega_{s'_j} \Rightarrow I_{s'} \cap I_{T_{l_1}} = \emptyset. \tag{3.5}$$

Definition 3.3 *For any $P \subseteq S$, $j \in \{1, 2, 4\}$ and $f \in \mathcal{S}(\mathbb{R})$, define*

$$\mathrm{size}_j(P, f) := \sup_{\substack{T \subseteq P \\ T \text{ is an } i\text{-tree for some } i \neq j}} \left(\frac{1}{|I_T|} \sum_{s \in T} \|f_{s_j}\|_2^2 \right)^{\frac{1}{2}}$$

and

$$\mathrm{energy}_j(P, f) := \sup_{n \in \mathbb{Z}} \sup_{\mathrm{T}} 2^n \left(\sum_{T \in \mathrm{T}} |I_T| \right)^{\frac{1}{2}},$$

where T *ranges over all chains of strongly j-disjoint trees in P having the property that*

$$\left(\sum_{s \in T} \|f_{s_j}\|_2^2 \right)^{\frac{1}{2}} \geq 2^n |I_T|^{\frac{1}{2}} \quad \text{for all } T \in \mathrm{T} \text{ and such that}$$

$$\left(\sum_{s \in T'} \|f_{s_j}\|_2^2 \right)^{\frac{1}{2}} \leq 2^{n+1} |I_{T'}|^{\frac{1}{2}} \quad \text{for all subtrees } T' \subseteq T \in \mathrm{T}.$$

We need the following three lemmas, whose proof will be given in Sections 4, 5, 6, respectively.

Lemma 3.4 *For any $0 \leq \theta_1, \theta_2, \theta_4 < 1$ with $\theta_1 + \theta_2 + \theta_4 = 1$ and $d \in \mathbb{N}$,*

$$|\Lambda_{S^d}(f_1, f_2, f_3, f_4)| \lesssim_{\theta_1, \theta_2} 2^d \prod_{j \in \{1,2,4\}} \mathrm{size}_j(S^d, f_j)^{\theta_j} \mathrm{energy}_j(S^d, f_j)^{1-\theta_j} |F_3|^{\frac{1}{p_3}}.$$

Lemma 3.5 *For any $P \subseteq S$, $j \in \{1, 2, 4\}$, $f \in \mathcal{S}(\mathbb{R})$, and $d \in \mathbb{N}$,*

$$\mathrm{size}_j(P, f) \lesssim_M \sup_{s \in P} \left(\frac{1}{|I_s|} \|f\|_{L^1(5 \cdot 2^d I_s)} + 2^{-Md} \inf_{y \in 2^d I_s} Mf(y) \right).$$

Lemma 3.6 *For any $P \subseteq S$, $j \in \{1, 2, 4\}$ and $f \in \mathcal{S}(\mathbb{R})$,*

$$\mathrm{energy}_j(P, f) \lesssim \|f\|_2.$$

By the above lemmas and the definitions of S^d and the exceptional set Ω, we have

$$|\Lambda_{S^d}(f_1, f_2, f_3, f_4)| \lesssim_M 2^{-M} |F_1|^{\frac{1+\theta_1}{2}} |F_2|^{\frac{1+\theta_2}{2}} |F_3|^{\frac{1}{p_3}},$$

from which (3.2) follows.

4 4-linear Form Estimate

We prove Lemma 3.4 in this section. For notational convenience, in the rest of this paper we write

$$S_j := \mathrm{size}_j(S^d, f_j) \quad \text{and} \quad E_j := \mathrm{energy}_j(S^d, f_j), \quad j \in \{1, 2, 4\}. \tag{4.1}$$

The following lemma is the main ingredient of the "grouping" algorithm we will use.

Lemma 4.1 *Let $j \in \{1,2,4\}$ and $P \subseteq S^d$ satisfy*

$$\text{size}_j(P, f_j) \le 2^{-n} E_j \quad \text{for some } n \in \mathbb{Z}.$$

Then we can decompose $P = P' \cup P''$ such that

$$\text{size}_j(P', f_j) \le \; 2^{-n-1} E_j \tag{4.2}$$

and P'' is a union of a collection $\mathscr{F}$ of trees T with $\sum_{T\in\mathscr{F}} |I_T| \lesssim 2^{2n}$.

Proof An i-tree T with top t is called "upward" if $c(\omega_{s_j}) \ge c(\omega_{t_j})$ for all $s \in T$. We can define "downward" i-tree with obvious modifications. Note that every i-tree is a union of an upward tree and a downward tree and the two trees intersect only at the top. Initially set $\mathscr{F} = \emptyset$. Consider all upward i-trees T for some $i \ne j$ satisfying $\sum_{s\in T} \|f_{s_j}\|^2 > |I_T|(2^{-n-1}|E_j|)^2$. If there are no such trees, then we stop. Otherwise, we choose a tree T satisfying two conditions: firstly it is maximal with respect to set inclusion; secondly, if t denotes the top of T, then $c(\omega_{t_j})$ is as large as possible. After selecting such a T, we define a j-tree by $T' := \{s \in P\backslash T : s_j < t_j\}$. Add T and T' to the collection $\mathscr{F}$ and remove tiles of T and T' from P. Repeat this process and when we stop we have $T_1, T_1', T_2, T_2', \ldots, T_M, T_M' \in \mathscr{F}$.

We claim that $T_1, T_2, \ldots, T_M$ form a chain of strongly j-disjoint tress. To see this, let us verify (3.3)–(3.5). (3.3) is automatically satisfied by construction. Let $l_1 \ne l_2$, $s \in T_{l_1}$, $s' \in T_{l_2}$ with $\omega_{s_j} \subsetneqq \omega_{s'_j}$. Then T_{l_1} is selected first by (2.16) and "upward" tree structure. Suppose $I_{T_{l_1}} \cap I_{s'} \ne \emptyset$. Then $I_{T_{l_1}} \supset I_{s'}$ and thus $s' \in T'_{l_1}$, which contradicts $s' \in T_{l_2}$. This proves (3.4). (3.5) can be verified in the similar way.

By the claim and the definition of energy$_j$, we have $\sum_{i=1}^M |I_{T_i}| \lesssim 2^{2n}$. Since T_i and T_i' have the same top, $\sum_{T\in\mathscr{F}} |I_T| \lesssim 2^{2n}$ as well. The termination of the previous algorithm implies that for any upward i-tree T in the remaining set $P \setminus (T_1 \cup T_1' \cup \cdots \cup T_M \cup T_M')$, we must have $\sum_{s\in T} \|f_{s_j}\|_2^2 < 2^{-2n-2}|I_T||E_j|^2$. We can run the algorithm again for downward trees with apparent modifications and the remaining set P' satisfy (4.2) by definition of *size*$_j$. The estimate $\sum_{T\in\mathscr{F}} |I_T| \lesssim 2^{2n}$ still holds after adding downward trees to $\mathscr{F}$. This finishes the proof of the lemma. □

Lemma 4.1 provides a way to reorganize the trees according to their sizes. To sum up all the trees with different sizes, we need to estimate the contribution from a single tree, which leads to the following lemma.

Lemma 4.2 *Let $T \subseteq S^d$ be a tree. Then*

$$|\Lambda_T(f_1, f_2, f_3, f_4)| \lesssim 2^d |I_T| \prod_{j\in\{1,2,4\}} \text{size}_j(T, f_j) |F_3|^{\frac{1}{p_3}}.$$

Proof Without loss of generality, assume T is a 1-tree. By the definition of Λ_S in (2.8) with S being the single tree T, Cauchy–Schwartz inequality, (2.9), and the definition of tree, we have

$$\begin{aligned}|\Lambda_T(f_1, f_2, f_3, f_4)| &\le \int \sup_{s\in T} |f_{1,s_1}| \Big(\sum_{s\in T} |f_{2,s_2}|^2\Big)^{\frac12} \Big(\sum_{s\in T} |f_{4,s_4}|^2\Big)^{\frac12} \sup_{s\in T} |f_{3,s_3}| \\ &\le |I_T| \sup_{s\in T} \|f_{1,s_1}\|_\infty \Big(\frac{1}{|I_T|} \sum_{s\in T} \|f_{2,s_2}\|_2^2\Big)^{\frac12}\end{aligned}$$

$$\cdot\left(\frac{1}{|I_T|}\sum_{s\in T}\|f_{4,s_4}\|_2^2\right)^{\frac{1}{2}}\sup_{s\in T}\frac{1}{|I_s|}\int|f_3(y)|\left(1+\frac{\operatorname{dist}(y,I_s)}{|I_s|}\right)^{-N}$$
$$\lesssim 2^d|I_T|\sup_{s\in T}\|f_{1,s_1}\|_\infty\ \mathrm{size}_2(T,f_2)\ \mathrm{size}_4(T,f_4)|F_3|^{\frac{1}{p_3}}.$$

It remains to prove $\|f_{1,s_1}\|_\infty\lesssim\mathrm{size}_1(T,f_1)$ for any $s\in T$. This follows from the estimate

$$\|f_{1,s_1}\|_\infty\lesssim\|f_{1,s_1}\|_2|I_s|^{-\frac{1}{2}},\tag{4.3}$$

since $\{s\}$ is a 2-tree. To prove (4.3), recall for $s=(k,n,l)$, $f_{1,s_1}(x)=\chi^*_{I_{k,n}}(x)f_1*\psi_{k,l}(x)$, where $\psi_{k,l}(x)=2^k\psi(2^kx)\mathrm{e}^{-2\pi\mathrm{i}\frac{l}{2}x}$. Let b be a real number such that $|\frac{l}{2}-b|=2^k$ and define $\widetilde{f_{1,s_1}}(x):=\mathrm{e}^{2\pi\mathrm{i}bx}f_{1,s_1}(x)$. Then $\widetilde{f_{1,s_1}}'(x)=\beta f_{1,s_1}(x)$ for some $\beta\lesssim 2^k$. Hence

$$\|f_{1,s_1}\|_\infty=\|\widetilde{f_{1,s_1}}\|_\infty\lesssim\sqrt{\|\widetilde{f_{1,s_1}}\|_2\|\widetilde{f_{1,s_1}}'\|_2}\lesssim 2^{\frac{k}{2}}\|f_{1,s_1}\|_2\lesssim\|f_{1,s_1}\|_2|I_s|^{-\frac{1}{2}},$$

as desired. □

We are now ready to prove Lemma 3.4, which is equivalent to

$$\left|\Lambda_{S^d}\left(\frac{f_1}{E_1},\frac{f_2}{E_2},f_3,\frac{f_4}{E_4}\right)\right|\lesssim 2^d\prod_{j\in\{1,2,4\}}\left(\frac{S_j}{E_j}\right)^{\theta_j}|F_3|^{\frac{1}{p_3}},\tag{4.4}$$

where S_j and E_j are defined as in (4.1). Apply Lemma 4.1 inductively to decompose $S^d=\bigcup_{n\in\mathbb{Z}}\bigcup_{T\in\mathscr{F}_n}T$, where $\mathscr{F}_n$ is a collection of trees satisfying $\sum_{T\in\mathscr{F}_n}|I_T|\lesssim 2^{2n}$ and *j-size* $(\bigcup_{T\in\mathscr{F}_n}T,f_j)\le\min(\frac{E_j}{2^n},S_j)$ for any $j\in\{1,2,4\}$. So

$$\begin{aligned}\left|\Lambda_{S^d}\left(\frac{f_1}{E_1},\frac{f_2}{E_2},f_3,\frac{f_4}{E_4}\right)\right|&\le\sum_n\sum_{T\in\mathscr{F}_n}\left|\Lambda_T\left(\frac{f_1}{E_1},\frac{f_2}{E_2},f_3,\frac{f_4}{E_4}\right)\right|\\&\lesssim 2^d\sum_n 2^{2n}\prod_{j\in\{1,2,4\}}\min\left(2^{-n},\frac{S_j}{E_j}\right)|F_3|^{\frac{1}{p_3}}\quad\text{(by Lemma 4.2)}\\&\lesssim 2^d\prod_{j\in\{1,2,4\}}\left(\frac{S_j}{E_j}\right)^{\theta_j}|F_3|^{\frac{1}{p_3}}.\end{aligned}$$

The last inequality follows from the observation $2^{-n}\le\max_{j\in\{1,2,4\}}(\frac{S_j}{E_j})$ (see Section 6.7 in [10] for a similar trick).

5 Estimates on Size

We prove Lemma 3.5 in this section. The following lemma is closely related with the John–Nirenberg inequality.

Lemma 5.1 *For any $P\subseteq S$, $j\in\{1,2,4\}$ and $f\in\mathcal{S}(\mathbb{R})$,*

$$\mathrm{size}_j(P,f)\lesssim\sup_{\substack{T\subseteq P\\ T\text{ is an }i\text{-tree for some }i\neq j}}\frac{1}{|I_T|}\left\|\left(\sum_{s\in T}\frac{\|f_{s_j}\|_2^2}{|I_s|}\chi_{I_s}\right)^{\frac{1}{2}}\right\|_{1,\infty}.$$

Proof Fix $j\in\{1,2,4\}$, $P\subset S$ and $f\in\mathcal{S}(\mathbb{R})$. Let $T\subseteq P$ be an i-tree for some $i\neq j$ such that

$$\mathrm{size}_j(P,f)=\left(\frac{1}{|I_T|}\sum_{s\in T}\|f_{s_j}\|_2^2\right)^{\frac{1}{2}}.$$

For simplicity write $a_s := \|f_{s_j}\|_2$ for $s \in T$ and we aim to show

$$\left(\frac{1}{|I_T|}\sum_{s\in T} {a_s}^2\right)^{\frac{1}{2}} \lesssim \frac{1}{|I_T|}\left\|\left(\sum_{s\in T}\frac{{a_s}^2}{|I_s|}\chi_{I_s}\right)^{\frac{1}{2}}\right\|_{1,\infty}. \tag{5.1}$$

Denote the left-hand side (LHS) and the right-hand side (RHS) of (5.1) by A and B, respectively. Let C be a large constant and define the set

$$E := \left\{x : \left(\sum_{s\in T}\frac{{a_s}^2}{|I_s|}\chi_{I_s}(x)\right)^{\frac{1}{2}} > CB\right\} \subseteq I_T. \tag{5.2}$$

By the definition of weak 1 norm,

$$|E| \leq \frac{B|I_T|}{CB} = \frac{|I_T|}{C} \tag{5.3}$$

Write E as a joint union of intervals $E = \bigcup_{I^m\in\mathcal{J}^M} I^m$, where $\mathcal{J}^M$ is the set of maximal elements in

$$\mathcal{J} := \left\{I = I_{s_0} \text{ for some } s_0 \in T : \left(\sum_{\substack{s\in T\\ I_s\supseteq I}}\frac{{a_s}^2}{|I_s|}\right)^{\frac{1}{2}} > CB\right\}. \tag{5.4}$$

By the definition of A,

$$A^2|I_T| = \sum_{s\in T} {a_s}^2 = \int_E \sum_{s\in T}\frac{{a_s}^2}{|I_s|}\chi_{I_s} + \int_{I_T\setminus E}\sum_{s\in T}\frac{{a_s}^2}{|I_s|}\chi_{I_s} =: \mathrm{H} + \mathrm{K}. \tag{5.5}$$

Use the decomposition $E = \bigcup_{I^m\in\mathcal{J}^M} I^m$ to split H further as

$$\mathrm{H} = \sum_{I^m\in\mathcal{J}^M}\int_{I^m}\sum_{\substack{s\in T\\ I_s\supsetneq I^m}}\frac{{a_s}^2}{|I_s|}\chi_{I_s} + \sum_{I^m\in\mathcal{J}^M}\int_{I^m}\sum_{\substack{s\in T\\ I_s\subseteq I^m}}\frac{{a_s}^2}{|I_s|}\chi_{I_s} =: \mathrm{H}_1 + \mathrm{H}_2. \tag{5.6}$$

Since each I^m is maximal in $\mathcal{J}$ defined by (5.4),

$$\mathrm{H}_1 \leq \sum_{I^m\in\mathcal{J}^M}(CB)^2|I^m| = (CB)^2|E| \leq (CB)^2|I_T|. \tag{5.7}$$

For each $I^m \in \mathcal{J}^M$, $\{s \in T : I_s \subseteq I^m\}$ is still an i-tree by the grid structure. So the definition of $\text{size}_j(P, f)$ and (5.3) give

$$\mathrm{H}_2 = \sum_{I^m\in\mathcal{J}^M}|I^m|\left(\frac{1}{|I^m|}\sum_{\substack{s\in T\\ I_s\subseteq I^m}}{a_s}^2\right) \leq \sum_{I^m\in\mathcal{J}^M}|I^m|A^2 = A^2|E| \leq A^2\frac{|I_T|}{C}. \tag{5.8}$$

Since the integrand in K is dominated by CB by (5.2), we have

$$\mathrm{K} \leq (CB)^2|I_T|. \tag{5.9}$$

Putting (5.5)–(5.9) together, we obtain

$$A^2|I_T| = \mathrm{H}_1 + \mathrm{H}_2 + \mathrm{K} \leq (CB)^2|I_T| + A^2\frac{|I_T|}{C} + (CB)^2|I_T|, \tag{5.10}$$

from which (5.1) follows. □

By the previous lemma, to prove Lemma 3.5 it suffices to show for any i-tree T with $i \neq j$,

$$\left\|\left(\sum_{s\in T}\frac{\|f_{s_j}\|_2^2}{|I_s|}\chi_{I_s}\right)^{\frac{1}{2}}\right\|_{1,\infty} \lesssim_M \|f\|_{L^1(5\cdot 2^d I_s)} + 2^{-Md}\inf_{y\in 2^d I_s}\mathcal{M}f(y)|I_T|. \tag{5.11}$$

Write $f = f\chi_{5\cdot 2^d I_T} + f\chi_{(5\cdot 2^d I_T)^c}$. LHS of (5.11) is bounded by

$$\Big\|\Big(\sum_{s\in T}\frac{\|(f\chi_{5\cdot 2^d I_T})_{s_j}\|_2^2}{|I_s|}\chi_{I_s}\Big)^{\frac12}\Big\|_{1,\infty}+\Big\|\Big(\sum_{s\in T}\frac{\|(f\chi_{(5\cdot 2^d I_T)^c})_{s_j}\|_2^2}{|I_s|}\chi_{I_s}\Big)^{\frac12}\Big\|_{1} =: \mathrm{I}+\mathrm{II}.$$

Term I is bounded by $C\|f\|_{L^1(5\cdot 2^d I_s)}$ since the discrete square-function operator is of weak type $(1,1)$ by the L^2 estimate and Calderón–Zygmund decomposition. Using the fact l^2 norm is no more than l^1 norm, we estimate II by

$$\sum_{s\in T}\|(f\chi_{(5\cdot 2^d I_T)^c})_{s_j}\|_2|I_s|^{\frac12}.$$

It remains to show

$$\sum_{s\in T}\|(f\chi_{(5\cdot 2^d I_T)^c})_{s_j}\|_2|I_s|^{\frac12}\lesssim_M 2^{-Md}\inf_{y\in 2^d I_s}\mathcal{M}f(y)|I_T|. \tag{5.12}$$

Using (2.9) we control the function $(f\chi_{(5\cdot 2^d I_T)^c})_{s_j}$ by

$$\begin{aligned}&|(f\chi_{(5\cdot 2^d I_T)^c})_{s_j}(x)|\\&\quad\lesssim_N \Big(1+\frac{\mathrm{dist}(I_s,(5\cdot 2^d I_T)^c)}{|I_s|}\Big)^{-N}\Big(1+\frac{\mathrm{dist}(x,I_s)}{|I_s|}\Big)^{-N}\inf_{y\in 2^d I_s}\mathcal{M}f(y).\end{aligned}$$

Hence

$$\begin{aligned}\text{LHS of (5.12)}&\lesssim_N \inf_{y\in 2^d I_s}\mathcal{M}f(y)\sum_{s\in T}|I_s|\Big(1+\frac{\mathrm{dist}(I_s,(5\cdot 2^d I_T)^c)}{|I_s|}\Big)^{-N}\\&\lesssim_M 2^{-Md}\inf_{y\in 2^d I_s}\mathcal{M}f(y)|I_T|,\end{aligned}$$

as desired.

6 Estimates on Energy

We prove Lemma 3.6 in this section. Fix $j\in\{1,2,4\}$. Let n and T be as in the definition of energy_j in Definition 3.3 where the supremum is attained, i.e.,

$$\mathrm{energy}_j(P,f) = 2^n\Big(\sum_{T\in\mathrm{T}}|I_T|\Big)^{\frac12}.$$

Here T is a *chain of strongly j-disjoint* trees in P having the property that

$$\Big(\sum_{s\in T}\|f_{s_j}\|_2^2\Big)^{\frac12}\ge 2^n|I_T|^{\frac12}\quad \text{for all } T\in\mathrm{T}, \tag{6.1}$$

and such that

$$\Big(\sum_{s\in T'}\|f_{s_j}\|_2^2\Big)^{\frac12}\le 2^{n+1}|I_{T'}|^{\frac12}\quad \text{for all subtrees } T'\subseteq T\in\mathrm{T}. \tag{6.2}$$

We need to show

$$2^n\Big(\sum_{T\in\mathrm{T}}|I_T|\Big)^{\frac12}\lesssim\|f\|_2. \tag{6.3}$$

By (6.1),

$$(\text{LHS of } (6.3))^2 = 2^{2n}\Big(\sum_{T\in \mathrm{T}} |I_T|\Big) \le 2^{2n}2^{-2n}\sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2 = \sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2,$$

and thus it remains to prove

$$\sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2 \lesssim \|f\|_2^2. \tag{6.4}$$

For each 4-tile s, define an operator A_s by $A_s f(x) = f_{s_j}(x)$. By the Cauchy–Schwarz inequality,

$$\sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2 = \Big\langle \sum_{T\in\mathrm{T}}\sum_{s\in T} A_s^*A_s f, f\Big\rangle \le \Big\|\sum_{T\in\mathrm{T}}\sum_{s\in T} A_s^*A_s f\Big\|_2 \|f\|_2.$$

Hence (6.4) follows from the following estimate:

$$\Big\|\sum_{T\in\mathrm{T}}\sum_{s\in T} A_s^*A_s f\Big\|_2 \lesssim \Big(\sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2\Big)^{\frac12}. \tag{6.5}$$

To prove (6.5), write

$$(\text{LHS of } (6.5))^2 = \sum_{T,T'\in\mathrm{T}}\sum_{\substack{s\in T\\ s'\in T'}} \langle A_s^*A_s f, A_{s'}^*A_{s'} f\rangle =: \mathrm{I} + \mathrm{II},$$

where I contains all off-diagonal terms in which $T\neq T'$, and II contains all diagonal terms where $T = T'$. Finally We only need to prove the following estimates:

Lemma 6.1 $\mathrm{I} \lesssim \sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2$ *and* $\mathrm{II} \lesssim \sum_{T\in\mathrm{T}}\sum_{s\in T}\|f_{s_j}\|_2^2$.

Proof We only prove the estimate for I. II is easier to control and we omit the proof. Apply Cauchy–Schwartz inequality,

$$\mathrm{I} \le \sum_{T\neq T'}\sum_{\substack{s\in T\\ s'\in T'}} \|A_s f\|_2 \|A_s A_{s'}^*\| \|A_{s'} f\|_2. \tag{6.6}$$

The following estimate for $\|A_sA_{s'}^*\|$ is the key to sum up all the terms in (6.6).

Lemma 6.2 $\|A_sA_{s'}^*\| \neq 0$ *only when* $\omega_{s_j}\cap\omega_{s'_j}\neq\emptyset$. *Moreover,*

$$\|A_sA_{s'}^*\| \lesssim_N \frac{|I_{s'}|^{\frac12}}{|I_s|^{\frac12}}\Big(1+\frac{\mathrm{dist}(I_s,I_{s'})}{|I_s|}\Big)^{-N} \quad \text{if } \omega_{s_j}\subseteq\omega_{s'_j}. \tag{6.7}$$

Proof Write $A_sA_{s'}^* f(x) = \int K(x,y)f(y)\,dy$, where $K(x,y) = \chi_{I_s}^*(x)\chi_{I_{s'}}^*(y)\widetilde{\psi_{s'_j}} * \psi_{s_j}(x-y)$, $\psi_{s_j} := \psi_{k,l_j}$ for $s=(k,n,l)$ and $\widetilde{g}(x) := \overline{g(-x)}$ for any function g. Note that $\widetilde{\psi_{s'_j}} * \psi_{s_j}(t) = \int \overline{\widehat{\psi_{s'}}}(\xi)\widehat{\psi_s}(\xi)\mathrm{e}^{2\pi\mathrm{i}\xi t}\,d\xi$ is non-zero only when $\omega_{s_j}\cap\omega_{s'_j}\neq\emptyset$ by (2.3) and (2.10). Assume $\omega_{s_j}\subseteq\omega_{s'_j}$. By definitions of χ_I^* (2.4) and $\psi_{k,l}$ and using the triangle inequality $(1+|a|)^{-1}+(1+|b|)^{-1} \le (1+|a+b|)^{-1}$,

$$\begin{aligned}|K(x,y)| &\lesssim_N \Big(1+\frac{\mathrm{dist}(x,I_s)}{|I_s|}\Big)^{-2N}\Big(1+\frac{\mathrm{dist}(y,I_{s'})}{|I_{s'}|}\Big)^{-N}\\ &\quad\cdot\frac{1}{|I_s||I_{s'}|}\int\Big(1+\frac{|x-y-z|}{|I_{s'}|}\Big)^{-2N}\Big(1+\frac{|z|}{|I_s|}\Big)^{-N}dz\\ &\lesssim_N \Big(1+\frac{\mathrm{dist}(I_s,I_{s'})}{|I_s|}\Big)^{-N}\frac{1}{|I_s|}\Big(1+\frac{\mathrm{dist}(x,I_s)}{|I_s|}\Big)^{-N}.\end{aligned}$$

Hence

$$\int |K(x,y)|dx \lesssim_N \left(1+\frac{\mathrm{dist}(I_s,I_{s'})}{|I_s|}\right)^{-N}. \tag{6.8}$$

Similarly,

$$\int |K(x,y)|dy \lesssim_N \left(1+\frac{\mathrm{dist}(I_s,I_{s'})}{|I_s|}\right)^{-N}\frac{|I_{s'}|}{|I_s|}. \tag{6.9}$$

(6.8) and (6.9) imply (6.7) by Schur's lemma. □

By Lemma 6.2 and symmetry, we can assume without loss of generality $\omega_{s_j} \subseteq \omega_{s'_j}$. In the following we assume further that $\omega_{s_j} \subsetneqq \omega_{s'_j}$. The case $\omega_{s_j} = \omega_{s'_j}$ can be handled the same way. For any $T, T' \in \mathrm{T}$, $T \neq T'$ and $s \in T$, $s' \in T'$, by (6.1) and (6.2), we have

$$\|A_s f\|_2 |I_s|^{-\frac{1}{2}} \le 2^{n+1} \lesssim 2^n \quad \text{and} \quad 2^n \le \left(|I_T|^{-1}\sum_{s_0\in T}\|f_{(s_0)_j}\|_2^2\right)^{\frac{1}{2}}.$$

These imply

$$\|A_s f\|_2 \lesssim |I_s|^{\frac{1}{2}}|I_T|^{-\frac{1}{2}}\left(\sum_{s_0\in T}\|f_{(s_0)_j}\|_2^2\right)^{\frac{1}{2}}. \tag{6.10}$$

Similarly,

$$\|A_{s'} f\|_2 \lesssim |I_{s'}|^{\frac{1}{2}}|I_T|^{-\frac{1}{2}}\left(\sum_{s_0\in T}\|f_{(s_0)_j}\|_2^2\right)^{\frac{1}{2}}. \tag{6.11}$$

Using (6.10) and (6.11), RHS of (6.6) can be estimated by

$$\begin{aligned}
&\sum_{T\neq T'}\sum_{\substack{s\in T, s'\in T'\\ \omega_{s_j}\subsetneqq\omega_{s'_j}}}\|A_s f\|_2\|A_sA_{s'}^*\|\|A_{s'}f\|_2\\
&\lesssim \sum_{T\neq T'}\sum_{\substack{s\in T, s'\in T'\\ \omega_{s_j}\subsetneqq\omega_{s'_j}}}|I_s|^{\frac{1}{2}}|I_{s'}|^{\frac{1}{2}}|I_T|^{-1}\left(\sum_{s_0\in T}\|f_{(s_0)_j}\|_2^2\right)\|A_sA_{s'}^*\|\\
&=\sum_{T\in\mathrm{T}}\left(\sum_{s_0\in T}\|f_{(s_0)_j}\|_2^2\right)\left(\sum_{\substack{s\in T, T'\neq T\\ s'\in T', \omega_{s_j}\subsetneqq\omega_{s'_j}}}|I_s|^{\frac{1}{2}}|I_{s'}|^{\frac{1}{2}}|I_T|^{-1}\|A_sA_{s'}^*\|\right).
\end{aligned}$$

Therefore, the estimate for I will be done once we show that for any $T \in \mathrm{T}$,

$$\sum_{\substack{s\in T, T'\neq T\\ s'\in T', \omega_{s_j}\subsetneqq\omega_{s'_j}}}|I_s|^{\frac{1}{2}}|I_{s'}|^{\frac{1}{2}}|I_T|^{-1}\|A_sA_{s'}^*\| \lesssim 1.$$

By (6.7), this amounts to

$$\sum_{\substack{s\in T, T'\neq T\\ s'\in T', \omega_{s_j}\subsetneqq\omega_{s'_j}}}\left(1+\frac{\mathrm{dist}(I_s,I_{s'})}{|I_s|}\right)^{-N} \lesssim |I_T|, \tag{6.12}$$

which is an immediate consequence of the separation condition (3.4). □

References

[1] Bernicot, F.: Uniform estimates for paraproducts and related multilinear multipliers. *Rev. Mat. Iberoam*, **25**, 1055–1088 (2009)

[2] Bernicot, F.: L^p estimates for non-smooth bilinear Littlewood–Paley square functions on $\mathbb{R}$. *Math. Ann.*, **351**, 1–49 (2011)

[3] Grafakos, L., Li, X.: Uniform bounds for the bilinear Hilbert transforms. I. *Ann. of Math.*, **159**, 889–933 (2004)

[4] Lacey, M., Thiele, C.: L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. *Ann. of Math.*, **146**, 693–724 (1997)

[5] Lacey, M., Thiele, C.: On Calderón's conjecture. *Ann. of Math.*, **149**, 475–496 (1999)

[6] Li, X.: Uniform bounds for the bilinear Hilbert transforms. II. *Rev. Mat. Iberoam.*, **22**, 1069–1126 (2006)

[7] Li, X.: Uniform estimates for some paraproducts. *New York J. Math.*, **14**, 145–192 (2008)

[8] Li, X.: Bilinear Hilbert transforms along curves I: The monomial case. *Anal. PDE*, **6**, 197–220 (2013)

[9] Li, X., Xiao, L. : Uniform estimates for bilinear Hilbert transform and bilinear maximal functions associated to polynomials. *Amer. J. Math.*, **138**, 907–962 (2016)

[10] Muscalu, C., Schlag, W.: Classical and Multilinear Harmonic Analysis. Vol. II, Cambridge University Press, Cambridge, 2013

[11] Muscalu, C., Tao, T., Thiele, C. : Uniform estimates on paraproducts. *J. Anal. Math.*, **87**, 369–384 (2002)

[12] Thiele, C.: A uniform estimate. *Ann. of Math.*, **156**, 519–563 (2002)