CrossMark

# Quasi Pieces of the Bilinear Hilbert Transform Incorporated into a Paraproduct

**Dong Dong[1]** ⓘ

**Abstract** We prove the boundedness of a class of tri-linear operators consisting of a quasi piece of bilinear Hilbert transform whose scale equals to or dominates the scale of its linear counter part. Such type of operators is motivated by the tri-linear Hilbert transform and its curved versions.

**Keywords** Bilinear Hilbert transform · Paraproduct · Tri-linear operators

**Mathematics Subject Classification** 42A50 · 47G10 · 42B99

## 1 Introduction

### 1.1 Background

In a pair of breakthrough papers [4,5], Lacey and Thiele proved the boundedness property of the bilinear Hilbert transform (BHT)

$$B(f_1, f_2)(x) = p.v. \int f_1(x - t) f_2(x + t) \frac{1}{t} \, dt.$$

Many interesting results about multilinear operators have been established in the spirit of Lacey–Thiele's method. However, $L^p$-boundedness of tri-linear Hilbert transform (THT)

✉ Dong Dong
  ddong3@illinois.edu

[1] Mathematics Department, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

$$T(f_1, f_2, f_3)(x) = p.v. \int f_1(x-t) f_2(x-2t) f_3(x-3t) \frac{1}{t} dt.$$

is still unknown. One difficulty arises from certain non-linear issue hidden in the trilinear structure. This is one of the main reasons motivating Li to study BHT along curves [8], say

$$H_\Gamma(f_1, f_2)(x) = p.v. \int f_1(x-t) f_2(x-t^d) \frac{1}{t} dt, \quad \text{where} \quad d \geq 2 \text{ is an integer.}$$

In [8], $H_\Gamma$ is split into two operators according to the efficiency of some oscillatory integral estimate (stationary phase vs. non-stationary phase). One of the two operators is a paraproduct of the form $\Pi_\Gamma(f_1, f_2) = \sum_k f_{1k} f_{2k}$ [7] that is more complex than the classical Coifman–Meyer paraproduct [1]. Although it turns out $\Pi_\Gamma$ is slightly simpler than BHT, the proof of its boundedness already requires sophisticated multi-scale time-frequency analysis that is essential in the study of BHT. Hence it is reasonable to expect that tri-linear analogues of the paraproduct $\Pi_\Gamma$ would be easier to handle than THT, but at the same time the study of such tri-linear operators could provide some new insights to THT.

The definition of tri-linear correspondence of $\Pi_\Gamma(f_1, f_2)$ was given in [2], where the author and Li introduced the following class of operators $T^{\alpha,\beta}$ that can be viewed a hybrid of BHT and paraproduct:

$$T^{\alpha,\beta}(f_1, f_2, f_3)(x) = \sum_{k \in \mathbb{Z}} H^{\alpha,k}(f_1, f_2)(x) f_3^{\beta,k}(x), \qquad (1.1)$$

where

$$\begin{cases} H^{\alpha,k}(f_1, f_2)(x) = \iint_{\mathbb{R}^2} \widehat{f_1}(\xi_1) \widehat{f_2}(\xi_2) e^{2\pi i (\xi_1 + \xi_2) x} \widehat{\Phi_1}\left(\frac{\xi_1 - \xi_2}{2^{\alpha k}}\right) d\xi_1 d\xi_2, \\ f^{\beta,k}(x) = \int_{\mathbb{R}} \widehat{f}(\xi) e^{2\pi i \xi x} \widehat{\Phi_2}\left(\frac{\xi}{2^{\beta k}}\right) d\xi. \end{cases} \qquad (1.2)$$

Here $\alpha$, $\beta$ are non-zero positive real numbers, and various conditions (about smoothness, support, etc) can be imposed on the cut-off functions $\widehat{\Phi_1}$ and $\widehat{\Phi_2}$.

$T^{\alpha,\beta}$ is closely related with THT along curves. For example, one promising way to prove the boundedness of $T_C(f_1, f_2, f_3)(x) = p.v. \int f_1(x-t) f_2(x+t) f_3(x-t^d) \frac{dt}{t}$ is to study $T^{1,d}$ first (See [8] for a similar approach in the bilinear setting). The following theorem is proved in [2].

**Theorem 1.1** ([2], Theorem 1.2) *Let $\Phi_1$ and $\Phi_2$ be smooth functions satisfying supp $\widehat{\Phi_1} \subseteq [9, 10]$ and supp $\widehat{\Phi_2} \subseteq [-1, 1]$. Assume $\alpha = \beta \neq 0$. Then the operator $T^{\alpha,\beta}$ defined by (1.1) (1.2) 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 $(p_1, p_2, p_3) \in D = \left\{ (p_1, p_2, p_3) \in (1, \infty)^3 : \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2} \right\}$.*

**Fig. 1** Tile structure of $T^{\alpha,\beta}$, $\alpha < \beta, k \geq 2$
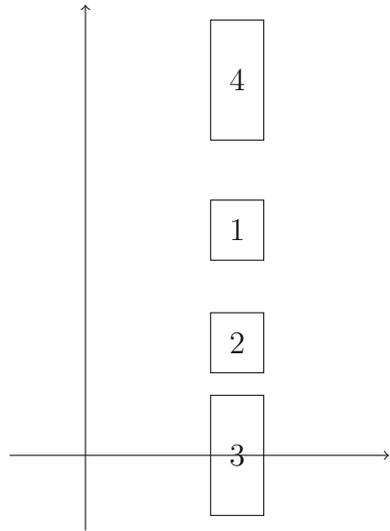


*Remarks* (1) Strictly speaking, this theorem is proved in [2] only in the case $\alpha = \beta = 1$, but this restriction is inessential. The proof given in [2] works for any homogeneous-scale case.

(2) The intervals [9, 10] and [−1, 1] in the assumptions of Theorem 1.1 are not essential. The point is that $\widehat{\Phi_1}$ should be supported away from 0 and $\widehat{\Phi_2}$ should be supported near 0.

(3) We conjectured that the condition $\alpha = \beta$ can be dropped in the above theorem, but the proof given in [2] relies on the homogeneity of the scales. Let us briefly analyze the difficulties in the case $\alpha \neq \beta$ here. Assume $0 < \alpha < \beta$ and let $k \geq 2$ be an integer. After wave packet decomposition, the tile associated with $f_3^{\beta,k}$ dominates the other two tiles (associated with $f_1$ and $f_2$) in frequency space as supp $\widehat{f_3^{\beta,k}}$ has a much larger scale $2^{\beta k}$. This will also introduce a long tile for the fourth function $f_4$ in the 4-linear form $\langle T^\alpha(f_1, f_2, f_3), f_4 \rangle$: see Fig. 1. As there are two long tiles and one of them contains the origin, the situation is difficult to handle even we use telescoping techniques that are powerful in some uniform estimates ([3,6,10]).

## 1.2 Main Result and Application

The purpose of this paper is to investigate other instances of $T^{\alpha,\beta}$, including some non-homogeneous scale cases. We would like to switch the roles of $\widehat{\Phi_1}$ and $\widehat{\Phi_2}$, i.e., assume that $\widehat{\Phi_1}$ is supported near the origin and $\widehat{\Phi_2}$ is supported away from 0 (instead of the other way around in Theorem 1.1). In this case, $H^{\alpha,k}$ is no longer a piece of BHT at certain scale: we may call it a *quasi piece* of BHT. Surprisingly we can obtain the same range of boundedness as before, even in some cases with non-homogeneous scales (See Theorem 1.3 below). More precisely, we have

**Theorem 1.2** *Let* $\Phi_1$ *and* $\Phi_2$ *be smooth bump functions satisfying supp* $\widehat{\Phi_1} \subseteq [-1, 1]$ *and supp* $\widehat{\Phi_2} \subseteq [9, 10]$. *Let* $\alpha = \beta \neq 0$. *Then the operator* $T^{\alpha,\beta}$ *defined by* (1.1) (1.2) *is bounded from* $L^{p_1} \times L^{p_2} \times L^{p_3}$ *to* $L^p$ *for any* $(p_1, p_2, p_3) \in D = \left\{ (p_1, p_2, p_3) \in (1, \infty)^3 : \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2} \right\}$, $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = \frac{1}{p}$.

The proof of Theorem 1.2 uses Lacey–Thiele's ideas about BHT. However, it should be noted that because of the quasi pieces of BHT, the 4-tile structure of the operator $T^{\alpha,\alpha}$ quite different from the tri-tile structure of BHT (see Fig. 3 for a comparison): the loss of one tile (1-tile and 2-tile are identical) forces us to mainly work with only two tiles as opposed to three tiles in BHT. The presence of a Littlewood–Paley piece (3-tile), however, will be of great help (see the proof of Proposition 3.4).

Using Theorem 1.2 together with Theorem 1.1, we can derive the boundedness property of positive truncations of $T^{\alpha,\beta}$ in some non-homogeneous scale cases.

**Theorem 1.3** *Let* $\Phi_1$ *and* $\Phi_2$ *be smooth bump functions satisfying supp* $\widehat{\Phi_1} \subseteq [-1, 1]$ *and supp* $\widehat{\Phi_2} \subseteq [9, 10]$. *Assume* $\alpha > \beta > 0$. *Define a positive truncation of* $T^{\alpha,\beta}$ *by*

$$T_N^{\alpha,\beta}(f_1, f_2, f_3)(x) = \sum_{k \geq N} H^{\alpha,k}(f_1, f_2)(x) f_3^{\beta,k}(x), \ N \in \mathbb{N}, \qquad (1.3)$$

*where* $H^{\alpha,k}$ *and* $f_3^{\beta,k}$ *are given in* (1.2). *Then for any* $N \geq 10\alpha/\beta$, *the operator* $T_N^{\alpha,\beta}$ *is bounded from* $L^{p_1} \times L^{p_2} \times L^{p_3}$ *into* $L^p$ *for any* $(p_1, p_2, p_3) \in D = \left\{ (p_1, p_2, p_3) \in (1, \infty)^3 : \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2} \right\}$, $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = \frac{1}{p}$.

*Remarks* (1) The choice of intervals $[-1, 1]$ and $[9, 10]$ in the above two theorems are not important. The key is that $\widehat{\Phi_1}$ should be supported near 0 and $\widehat{\Phi_2}$ should be supported away from 0.

(2) One of anticipated applications of Theorem 1.3 is to use boundedness of $T_N^{d,1}$ to prove that of one prototype of THT along polynomial curves

$$T^C(f_1, f_2, f_3)(x) = p.v. \int_{-1}^{1} f_1(x - t) f_2(x - t^d) f_3(x + t^d) \frac{dt}{t}.$$

Just like the relationship between $H_\Gamma(f_1, f_2)(x) = p.v. \int f_1(x-t) f_2(x-t^d) \frac{1}{t} dt$ and the paraproduct $\Pi_\Gamma(f_1, f_2) = \sum_k f_{1k} f_{2k}$ studied in [7], $T^C$ can be written as the sum of finitely many operators of the form $T_N^{\alpha,\beta}$ (plus some other terms). The condition $N \geq 10\alpha/\beta$ in Theorem 1.3 is assumed only for technical reasons and it does not affect the application as each scale of $T^C$ (after the standard dyadic decomposition $\frac{1}{t} = \sum_k \rho_k(t)$) is trivially bounded. The reason that we only consider the positive truncation instead of $T^{\alpha,\beta}$ itself is that $|t| \leq 1$ in the definition of $T^C$.

(3) Under the assumptions on $\widehat{\Phi_1}$ and $\widehat{\Phi_2}$ in Theorem 1.3, Fig. 2 illustrates the worst case of the tri-tile structure of $T_N^{\alpha,\beta}$ with $\alpha > \beta$ at any positive scale $k$. The two identical long tiles seems to be very problematic. The key to resolve this issue is

**Fig. 2** Tri-tile structure of $T_N^{\alpha,\beta}$, $\alpha > \beta > 0$, $k \geq 1$
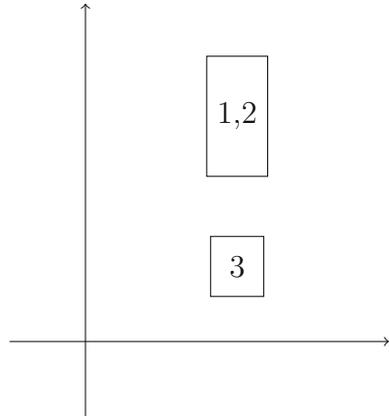


to reduce the study of $T_N^{\alpha,\beta}$ with $\alpha > \beta$ to that of $T^{\beta,\beta}$ (homogeneous case) by a telescoping argument. The details are provided in Sect. 6.

### 1.3 Notations

Throughout the paper we will 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 $A \leq CB$ and $A \lesssim_N B$ means $A \leq C_N B$. If $A \lesssim B$ and $B \lesssim A$, then we write $A \simeq B$. $\chi_E$ and $|E|$ will be used to denote the characteristic function and the Lebesgue measure of the set $E$, respectively.

## 2 Reduction to Model Form

The goal of this section is to reduce Theorem 1.2 to the study of a model form using standard wave packet decomposition process. For notational convenience, we assume $\alpha = \beta = 1$ in the proof. The general case can be handled the same way.

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 $\Lambda$ associated with $T^{1,1}$

$$
\begin{aligned}
\Lambda(f_1, f_2, f_3, f_4) &:= \int T^{1,1}(f_1, f_2, f_3)(x)\overline{f_4}(x)\,dx \\
&= \sum_{k \in \mathbb{Z}} \iiint \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\widehat{f_3}(\xi_3)\widehat{\Phi_1}\left(\frac{\xi_1 - \xi_2}{2^k}\right)\widehat{\Phi_2}\left(\frac{\xi_3}{2^k}\right)\overline{\widehat{f_4}}(\xi_1 + \xi_2 + \xi_3)\,d\xi_1 d\xi_2 d\xi_3,
\end{aligned}
\tag{2.1}
$$

where supp $\widehat{\Phi_1} \subseteq [-1, 1]$ and supp $\widehat{\Phi_2} \subseteq [9, 10]$.

To simplify the 4-linear form above, we use the wave packet decomposition. Choose a $\psi \in \mathcal{S}(\mathbb{R})$ such that supp $\widehat{\psi} \subseteq [0, 1]$ and

$$
\sum_{l \in \mathbb{Z}} \widehat{\psi}\left(\xi - \frac{l}{2}\right) = 1 \text{ for any } \xi \in \mathbb{R}.
$$

Define

$$\widehat{\psi_{k,l}}(\xi) := \widehat{\psi}\left(\frac{\xi - 2^{k-1}l}{2^k}\right) \text{ for } (k, l) \in \mathbb{Z}^2.$$

Pick a non-negative $\varphi \in \mathcal{S}(\mathbb{R})$ with $\mathrm{supp}\widehat{\varphi} \subseteq [-1, 1]$ and $\widehat{\varphi}(0) = 1$. Let

$$\varphi_k(x) := 2^k \varphi(2^k x), k \in \mathbb{Z}.$$

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 any 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

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

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

In sum, $f_{k,n,l}$ is well-localized, as $\mathrm{supp}\,\widehat{f_{k,n,l}} \subseteq [2^k(\frac{l}{2} - 1), 2^k(\frac{l}{2} + 2)]$ and $f_{k,n,l}$ is essentially supported on $I_{k,n}$ in the sense that

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

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

$$\Lambda(f_1, f_2, f_3, f_4) = \sum_{\substack{k \in \mathbb{Z} \\ (n_1,n_2,n_3,n_4) \in \mathbb{Z}^4 \\ (l_1,l_2,l_3,l_4) \in \mathbb{Z}^4}} \iiint \widehat{(f_1)_{k,n_1,l_1}}(\xi_1) \widehat{(f_2)_{k,n_2,l_2}}(\xi_2) \widehat{(f_3)_{k,n_3,l_3}}(\xi_3)$$

$$\widehat{\Phi_1}\left(\frac{\xi_1 - \xi_2}{2^k}\right) \widehat{\Phi_2}\left(\frac{\xi_3}{2^k}\right) \overline{\widehat{(f_4)_{k,n_4,l_4}}(\xi_1 + \xi_2 + \xi_3)}\, d\xi_1 d\xi_2 d\xi_3.$$

By the support of functions, each term in the sum is non-zero only when

$$\begin{cases} \xi_i \in [2^k(\frac{l_i}{2} - 1), 2^k(\frac{l_i}{2} + 2)] \text{ for } i = 1, 2, 3; \\ |\xi_1 - \xi_2| \lesssim 2^k, |\xi_3| \in [9 \cdot 2^k, 10 \cdot 2^k); \\ \xi_1 + \xi_2 + \xi_3 \in [2^k(\frac{l_4}{2} - 1), 2^k(\frac{l_4}{2} + 2)]. \end{cases}$$

These imply that

$$
\begin{cases}
|l_2 - l_1| \lesssim 1; \\
|l_3 - 9| \lesssim 1; \\
|l_4 - (2l_1 - 18)| \lesssim 1.
\end{cases}
$$

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 $l_1$. Then drop the cut-off functions by the Fourier expansion trick and ignore the fast decay terms so that $\Lambda(f_1, f_2, f_3, f_4)$ becomes essentially as

$$
\sum_{\substack{k, l_1 \\ n_1, n_2, n_3, n_4}} \int (f_1)_{k, n_1, l_1}(x)(f_2)_{k, n_2, l_2}(x)(f_3)_{k, n_3, l_3}(x)\overline{(f_4)_{k, n_4, l_4}}(x)\, dx.
$$

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 not too much loss 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 $\Lambda$ to denote the model 4-linear form by an 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 $l_1 = l$, $l_2 = l$, $l_3 = 18$ and $l_4 = 2l + 18$.

We will prove directly that $T$ is of restricted weak type (see [9] 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, we will prove

**Theorem 2.1** *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* $|F'| \geq \frac{1}{2}|F|$ *such that* $\Lambda$ *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* $|f_1| \leq \mathcal{X}_{F_1}$, $|f_2| \leq \mathcal{X}_{F_2}$, $|f_3| \leq \mathcal{X}_{F_3}$ *and* $|f_4| \leq \mathcal{X}_{F'}$. *Here* $\frac{1}{p'} := 1 - (\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3})$.

To prove Theorem 2.1 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 $\Lambda_S$ clearer. To each tuple $s = (k, n, l) \in \mathbb{Z}^3$ we assign a time-interval $I_s := I_{k, n}$ and four frequency-intervals

$\omega_{s_j}$, $j \in \{1, 2, 3, 4\}$, representing the localization of functions in the time-frequency space. More precisely, $I_s$ and $\omega_{s_j}$'s satisfy:

$(f_j)_{k,n,l_j}(x)$ is dominated by

$$C_{N,M} \left(1 + \frac{\mathrm{dist}(x, I_s)}{|I_s|}\right)^{-N} \frac{1}{|I_s|} \int |f_j(y)| \left(1 + \frac{|x - y|}{|I_s|}\right)^{-M} dy \qquad (2.9)$$

The Fourier transform of $(f_j)_{k,n,l_j}$ is supported on $\omega_{s_j}$. $\qquad (2.10)$

**Definition 2.2** We 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 $f_{s_j} := f_{k,n,l_j}$ for simplicity.

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

$$\omega_{s_1} = \omega_{s_2}; \qquad (2.11)$$

$$|\omega_{s_1}| = |\omega_{s_3}| = |\omega_{s_4}| = C|I_s|^{-1}; \qquad (2.12)$$

$$\mathrm{dist}(\omega_{s_1}, \omega_{s_4}) = |\omega_{s_1}|; \qquad (2.13)$$

$$c(\omega_{s_1}) > c(\omega_{s_4}), \text{ where } c(I) \text{ is the center of the interval } I; \qquad (2.14)$$

$$\{I_s\}_{s \in S} \text{ is a grid (defined below)}; \qquad (2.15)$$

$$\{\omega_{s_1} \cup \omega_{s_4}\}_{s \in S} \text{ is a gird}; \qquad (2.16)$$

$$\omega_{s_i} \not\subsetneq J \text{ for some } i \in \{1, 4\}, \ J := \omega_{s'_1} \cup \omega_{s'_2} \cup \omega_{s'_4}, s' \in S \Rightarrow$$

$$\omega_{s_j} \subseteq J \text{ for all } j \in \{1, 4\}. \qquad (2.17)$$

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)– (2.17). See Fig. 3 for a comparison between the tile structure of $T^{1,1}$ and that of BHT.

Theorem 2.1 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 $|f_4| \leq \chi_{F'}$.*
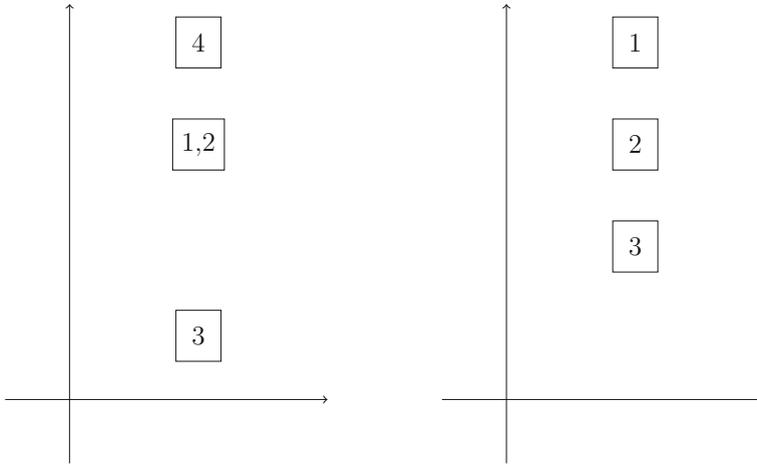
**Fig. 3** 4-tile of $T^{1,1}$ versus tri-tile of BHT

## 3 Proof of Theorem 1.2

In this section we prove Theorem 2.3 and hence Theorem 1.2, using some propositions whose proof will be given in subsequent sections. Fix $p > 1$, $(p_1, p_2, p_3) \in D_0 = \{(p_1, p_2, p_3) : 1 < p_1, p_2 < 2, \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2}, p_3 \in (1, \infty)\}$ 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

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

Then $|\Omega| \leq \frac{1}{4}$ when $C$ is large enough. Set $F' := F \setminus \Omega$ so that $|F'| \geq \frac{1}{2}$. For any dyadic number $\mu \geq 1$, define

$$S^{\mu} := \left\{ s \in S : 1 + \frac{\text{dist}(I_s, \Omega^c)}{|I_s|} \simeq \mu \right\}. \tag{3.1}$$

Then it suffices to obtain the estimate

$$|\Lambda_{S^{\mu}}(f_1, f_2, f_3, f_4)| \lesssim \mu^{-2}|F_1|^{\frac{1}{p_1}}|F_2|^{\frac{1}{p_2}}|F_3|^{\frac{1}{p_3}} \text{ for any dyadic } \mu \geq 1. \tag{3.2}$$

The main idea to obtain (3.2) is to group the tiles in $S^{\mu}$ appropriately, aiming to establish orthogonality among groups. The following definitions are needed.

**Definition 3.1** Let $j \in \{1, 4\}$. 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 denote $I_T := I_t$. We call $T \subseteq S$ a **tree** (with top $t$) if for any $s \in T$ we have $I_s \subseteq I_t$ and $\omega_{s_j} \supseteq \omega_{t_j}$ for some $j \in \{1, 4\}$.

It is easy to see that any tree is a union of a 1-tree and a 4-tree.

**Definition 3.2** For any $P \subseteq S$ and $f \in \mathcal{S}(\mathbb{R})$, define

$$\mathbf{size}_j(P, f) := \sup_{\substack{T \subseteq P \\ T \text{ is a 4-tree}}} \left( \frac{1}{|I_T|} \sum_{s \in T} \|f_{s_j}\|_2^2 \right)^{\frac{1}{2}}, \quad j = 1 \text{ or } 2;$$

$$\mathbf{size}_4(P, f) := \sup_{\substack{T \subseteq P \\ T \text{ is a 1-tree}}} \left( \frac{1}{|I_T|} \sum_{s \in T} \|f_{s_4}\|_2^2 \right)^{\frac{1}{2}}.$$

Sizes can be controlled using the proposition below, whose proof will be given in Sect. 4.

**Proposition 3.3** *Fix a dyadic number $\mu \geq 1$. For any $P \subseteq S^\mu$, $j \in \{1, 2, 4\}$ and $f \in \mathcal{S}(\mathbb{R})$,*

$$size_j(P, f) \lesssim_M \sup_{s \in P} \left( \frac{1}{|I_s|} \|f\|_{L^1(\mu I_s)} + \mu^{-M} \inf_{y \in \mu I_s} Mf(y) \right).$$

If tiles form a tree, then we can control the corresponding 4-form by sizes, as suggested by the following proposition.

**Proposition 3.4** *Let $T \subseteq S^\mu$ be a tree. Then*

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

*Proof* First assume $T$ is a 1-tree. By Cauchy–Schwarz inequality, we have

$$|\Lambda_T(f_1, f_2, f_3, f_4)| \leq \int \sup_{s \in T} |(f_1)_{s_1}| \sup_{s \in T} |(f_2)_{s_2}| \left( \sum_{s \in T} |(f_3)_{s_3}|^2 \right)^{\frac{1}{2}} \left( \sum_{s \in T} |(f_4)_{s_4}|^2 \right)^{\frac{1}{2}}$$

$$\leq |I_T| \sup_{s \in T} \|(f_1)_{s_1}\|_\infty \sup_{s \in T} \|(f_2)_{s_2}\|_\infty \left( \frac{1}{|I_T|} \sum_{s \in T} \|(f_3)_{s_3}\|_2^2 \right)^{\frac{1}{2}} \left( \frac{1}{|I_T|} \sum_{s \in T} \|(f_4)_{s_4}\|_2^2 \right)^{\frac{1}{2}}.$$

Using the structure of the 1-tree and the definition of $S^\mu$,

$$\left( \frac{1}{|I_T|} \sum_{s \in T} \|(f_3)_{s_3}\|_2^2 \right)^{\frac{1}{2}} \lesssim \mu \min \left\{ 1, |F_3|^{\frac{1}{p}} \right\} \leq \mu |F_3|^{\frac{1}{p_3}} \tag{3.3}$$

Combine the above two estimates and can bound $|\Lambda_T(f_1, f_2, f_3, f_4)|$ by

$$\mu |I_T| \sup_{s \in T} \|(f_1)_{s_1}\|_\infty \sup_{s \in T} \|(f_2)_{s_2}\|_\infty \, size_4(T, f_4) |F_3|^{\frac{1}{p_3}}.$$

It remains to prove that for $i = 1$ or $i = 2$, $\|(f_i)_{s_i}\|_\infty \lesssim \mathrm{size}_i(T, f_i)$ for any $s \in T$. We will only consider $i = 1$ case as the other case can be handled similarly. We just need to prove the estimate

$$\|(f_1)_{s_1}\|_\infty \lesssim \|(f_1)_{s_1}\|_2 |I_s|^{-\frac{1}{2}} \tag{3.4}$$

since $\{s\}$ is a 4-tree. To prove (3.4), recall for $s = (k, n, l)$, $(f_1)_{s_1}(x) = \chi^*_{I_{k,n}}(x) f_1 * \psi_{k,l}(x)$, where $\psi_{k,l}(x) = 2^k \psi(2^k x) e^{-2\pi i \frac{l}{2} x}$. Let $b$ be a real number such that $|\frac{l}{2} - b| = 2^k$ and define $\widetilde{(f_1)_{s_1}}(x) := e^{2\pi i b x}(f_1)_{s_1}(x)$. Then $\widetilde{(f_1)_{s_1}}'(x) = \gamma (f_1)_{s_1}(x)$ for some $\gamma \lesssim 2^k$. 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.

Now assume $T$ is a 4-tree. By similar arguments, we have

$$|\Lambda_T(f_1, f_2, f_3, f_4)| \leq \int \left( \sum_{s \in T} |(f_1)_{s_1}|^2 \right)^{\frac{1}{2}} \sup_{s \in T} |(f_2)_{s_2}| \left( \sum_{s \in T} |(f_3)_{s_3}|^2 \right)^{\frac{1}{2}} \sup_{s \in T} |(f_4)_{s_4}|$$

$$\leq |I_T| \left( \frac{1}{|I_T|} \sum_{s \in T} \|(f_1)_{s_1}\|_2^2 \right)^{\frac{1}{2}} \sup_{s \in T} \|(f_2)_{s_2}\|_\infty \left( \frac{1}{|I_T|} \sum_{s \in T} \|(f_3)_{s_3}\|_2^2 \right)^{\frac{1}{2}} \sup_{s \in T} \|(f_4)_{s_4}\|_\infty$$

$$\lesssim \mu |I_T| \, \mathrm{size}_1(T, f_1) \sup_{s \in T} \|(f_2)_{s_2}\|_\infty \sup_{s \in T} \|(f_4)_{s_4}\|_\infty |F_3|^{\frac{1}{p_3}}$$

$$\lesssim \mu |I_T| \prod_{j \in \{1,2,4\}} \mathrm{size}_j(T, f_j) |F_3|^{\frac{1}{p_3}} .$$

This finishes the proof of Proposition 3.4.                                    □

The following proposition provides the algorithm to select trees and group tiles.

**Proposition 3.5** *Let $f \in L^2$. Suppose for some $j \in \{1, 2, 4\}$ and $P \subseteq S$, we have*

$$\mathrm{size}_j(P, f) \leq \sigma \|f\|_2 \text{ for some dyadic number } \sigma = 2^n, n \in \mathbb{Z}.$$

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

$$\mathrm{size}_j(P', f) \leq \frac{\sigma}{2} \|f\|_2 \tag{3.5}$$

*and $P''$ is a union of trees $T$ in some collection $\mathscr{F}$ with $\sum_{T \in \mathscr{F}} |I_T| \lesssim \frac{1}{\sigma^2}$.*

The proof of this organization proposition will be postponed to Sect. 5.

Now we are ready to prove our goal (3.2). By Proposition 3.3 and the definition of $S^\mu$, we have

$$\text{size}_j(S^\mu, f_j) \lesssim \begin{cases} \mu|F_j| & \text{when } j = 1, 2; \\ \mu^{-M} & \text{for any large } M > 0 \text{ when } j = 4. \end{cases} \quad (3.6)$$

Iterate the organization algorithm Proposition 3.5 for all $j = 1, 2, 4$ simultaneously, and we can decompose $S^\mu$ as

$$S^\mu = \bigcup_{\substack{\sigma \text{ is a} \\ \text{dyadic number}}} S_\sigma,$$

where

$$\text{size}_j(S_\sigma, f_j) \lesssim \begin{cases} \min\{\mu|F_j|, \sigma|F_j|^{\frac{1}{2}}\} & \text{when } j = 1, 2; \\ \min\{\mu^{-M}, \sigma\} & \text{for any large } M > 0 \text{ when } j = 4, \end{cases} \quad (3.7)$$

and $S_\sigma = \cup_{T \in \mathscr{F}_\sigma} T$ is a union of tees with $\sum_{T \in \mathscr{F}_\sigma} |I_T| \lesssim \frac{1}{\sigma^2}$.

Using this decomposition and the estimate on a single tree (Proposition 3.4), we see that

$$|\Lambda_{S^\mu}(f_1, f_2, f_3, f_4)| \lesssim \sum_{\sigma \text{ is dyadic}} \sum_{T \in \mathscr{F}_\sigma} |\Lambda_T(f_1, f_2, f_3, f_4)|$$

$$\lesssim \mu \sum_\sigma \sum_{T \in \mathscr{F}_\sigma} |I_T| \prod_{j \in \{1,2,4\}} \text{size}_j(T, f_j)|F_3|^{\frac{1}{p_3}}$$

$$\lesssim \mu^3 |F_3|^{\frac{1}{p_3}} \sum_\sigma \frac{1}{\sigma^2} \min\left\{|F_1|, \sigma|F_1|^{\frac{1}{2}}\right\} \min\left\{|F_2|, \sigma|F_2|^{\frac{1}{2}}\right\} \min\left\{\mu^{-M}, \sigma\right\}.$$

Apply the elementary inequality $\min\{X, Y\} \leq X^\theta Y^{1-\theta}$, and we can bound $|\Lambda_{S^\mu}(f_1, f_2, f_3, f_4)|$ by

$$\mu^3 |F_3|^{\frac{1}{p_3}} \sum_\sigma \frac{1}{\sigma^2} \sigma^{2\left(1-\frac{1}{p_1}\right)+2\left(1-\frac{1}{p_2}\right)}|F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}} \min\left\{\mu^{-M}, \sigma\right\}$$

$$\lesssim \mu^{-2} |F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}} |F_3|^{\frac{1}{p_3}},$$

where we used the fact $\frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2}$ in the last inequality. This proves (3.2).

## 4 Size Estimates

In this section, we prove Proposition 3.3. The proofs of some variants of this proposition already appear in [2] and [9]. For the convenience of the reader, we include the details here. First we need the following lemma which is another form of the John-Nirenberg inequality.

**Lemma 4.1** *For any $P \subseteq S$ and $f \in \mathcal{S}(\mathbb{R})$,*

$$size_j(P, f) \lesssim \sup_{\substack{T \subseteq P \\ T \text{ is a 4-tree}}} \frac{1}{|I_T|} \left\| \left( \sum_{s \in T} \frac{\|f_{s_j}\|_2^2}{|I_s|} \chi_{I_s} \right)^{\frac{1}{2}} \right\|_{1,\infty} , \quad j \in \{1, 2\},$$

$$size_4(P, f) \lesssim \sup_{\substack{T \subseteq P \\ T \text{ is a 1-tree}}} \frac{1}{|I_T|} \left\| \left( \sum_{s \in T} \frac{\|f_{s_j}\|_2^2}{|I_s|} \chi_{I_s} \right)^{\frac{1}{2}} \right\|_{1,\infty} .$$

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

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

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

$$\left( \frac{1}{|I_T|} \sum_{s \in T} a_s^2 \right)^{\frac{1}{2}} \lesssim \frac{1}{|I_T|} \left\| \left( \sum_{s \in T} \frac{a_s^2}{|I_s|} \chi_{I_s} \right)^{\frac{1}{2}} \right\|_{1,\infty} . \tag{4.1}$$

Denote the left-hand side (LHS) and the right-hand side (RHS) of (4.1) by $A$ and $B$, respectively. Let $C$ be a large constant and define the set

$$E := \left\{ x : \left( \sum_{s \in T} \frac{a_s^2}{|I_s|} \chi_{I_s}(x) \right)^{\frac{1}{2}} > CB \right\} \subseteq I_T. \tag{4.2}$$

By the definition of weak 1 norm,

$$|E| \leq \frac{B|I_T|}{CB} = \frac{|I_T|}{C} \tag{4.3}$$

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

$$\mathcal{J} := \left\{ I = I_{s_0} \text{ for some } s_0 \in T : \left( \sum_{s \in T, I_s \supseteq I} a_s^2 |I_s|^{-1} \right)^{\frac{1}{2}} > CB \right\}. \tag{4.4}$$

By the definition of $A$,

$$A^2 |I_T| = \sum_{s \in T} a_s^2 = \int_E \sum_{s \in T} \frac{a_s^2}{|I_s|} \chi_{I_s} + \int_{I_T \setminus E} \sum_{s \in T} \frac{a_s^2}{|I_s|} \chi_{I_s} =: H + K. \tag{4.5}$$

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

$$H = \sum_{I^m \in \mathcal{J}^M} \int_{I^m} \sum_{s \in T, I_s \supsetneq I^m} \frac{a_s^2}{|I_s|} \chi_{I_s} + \sum_{I^m \in \mathcal{J}^M} \int_{I^m} \sum_{s \in T, I_s \subseteq I^m} \frac{a_s^2}{|I_s|} \chi_{I_s} =: H_1 + H_2.$$

(4.6)

Since each $I^m$ is maximal in $\mathcal{J}$ defined by (4.4),

$$H_1 \leq \sum_{I^m \in \mathcal{J}^M} (CB)^2 |I^m| = (CB)^2 |E| \leq (CB)^2 |I_T|.$$

(4.7)

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 $size_j(P, f)$ and (4.3) give

$$H_2 = \sum_{I^m \in \mathcal{J}^M} |I^m| \left( \frac{1}{|I^m|} \sum_{s \in T, I_s \subseteq I^m} a_s^2 \right) \leq \sum_{I^m \in \mathcal{J}^M} |I^m| A^2 = A^2 |E| \leq A^2 \frac{|I_T|}{C}$$

(4.8)

Since the integrand in $K$ is dominated by $CB$ by (4.2), we have

$$K \leq (CB)^2 |I_T|.$$

(4.9)

Putting (4.5)–(4.9) together, we obtain

$$A^2 |I_T| = H_1 + H_2 + K \leq (CB)^2 |I_T| + A^2 \frac{|I_T|}{C} + (CB)^2 |I_T|,$$

(4.10)

from which we obtain $A \lesssim B$. This proves (4.1) and thus Lemma 4.1. $\qquad\square$

We now turn to the proof of Proposition 3.3. Without loss of generality, assume $j = 1$. By Lemma 4.1, it suffices to show for any 4-tree $T$,

$$\left\| \left( \sum_{s \in T} \frac{\|f_{s_1}\|_2^2}{|I_s|} \chi_{I_s} \right)^{\frac{1}{2}} \right\|_{1,\infty} \lesssim_M \|f\|_{L^1(\mu I_T)} + \mu^{-M} \inf_{y \in \mu I_T} \mathcal{M} f(y) |I_T|.$$

(4.11)

Write $f = f \chi_{\mu I_T} + f \chi_{(\mu I_T)^c}$. LHS of (4.11) is bounded by

$$\left\| \left( \sum_{s \in T} \frac{\|(f\chi_{\mu I_T})_{s_1}\|_2^2}{|I_s|} \chi_{I_s} \right)^{\frac{1}{2}} \right\|_{1,\infty} + \left\| \left( \sum_{s \in T} \frac{\|(f\chi_{(\mu I_T)^c})_{s_1}\|_2^2}{|I_s|} \chi_{I_s} \right)^{\frac{1}{2}} \right\|_1 =: I + II.$$

By the conditions (2.11)–(2.17) of the tiles, in a 4-tree, $s_1$ tiles are Littlewood–Paley pieces as illustrated in Fig. 4. Thus term $I$ is bounded by $C\|f\|_{L^1(\mu I_T)}$ since the discrete square-function operator is of weak type $(1, 1)$ by the $L^2$ estimate and Calderón-Zygmund decomposition.

**Fig. 4** $s_1$ tiles in a 4-tree

Using the fact $l^2$ norm is no more than $l^1$ norm, we estimate $II$ by

$$\sum_{s \in T} \|(f \chi_{(\mu I_T)^c})_{s_1}\|_2 |I_s|^{\frac{1}{2}}.$$

It remains to show

$$\sum_{s \in T} \|(f \chi_{(\mu I_T)^c})_{s_1}\|_2 |I_s|^{\frac{1}{2}} \lesssim_M \mu^{-M} \inf_{y \in \mu I_T} \mathcal{M} f(y) |I_T|. \tag{4.12}$$

Using (2.9) we see that control the function $|(f \chi_{(\mu I_T)^c})_{s_1}(x)|$ is bounded above by

$$\left(1 + \frac{\text{dist}(I_s, (\mu I_T)^c)}{|I_s|}\right)^{-N} \left(1 + \frac{\text{dist}(x, I_s)}{|I_s|}\right)^{-N} \inf_{y \in \mu I_T} \mathcal{M} f(y).$$

Hence $\sum_{s \in T} \|(f \chi_{(\mu I_T)^c})_{s_1}\|_2 |I_s|^{\frac{1}{2}}$ is dominated by

$$\inf_{y \in \mu I_T} \mathcal{M} f(y) \sum_{s \in T} |I_s| \left(1 + \frac{\text{dist}(I_s, (\mu I_T)^c)}{|I_s|}\right)^{-N} \lesssim_M \mu^{-M} \inf_{y \in \mu I_T} \mathcal{M} f(y) |I_T|,$$

as desired.

## 5 Organizing Tiles

We provide the proof of Proposition 3.5 in this section. Without loss of generality, let $j = 1$. By the assumptions of Proposition 3.5,

$$\sup_{\substack{T \subseteq P \\ T \text{ is a 4-tree}}} \left(\frac{1}{|I_T|} \sum_{s \in T} \|f_{s_1}\|_2^2\right)^{\frac{1}{2}} \leq \sigma \|f\|_2. \tag{5.1}$$

Now we begin the tree selection algorithm. Initially set $S_0 = P$ and $\mathscr{F} = \emptyset$. Let

$$\mathscr{F}_0 = \left\{ T \subseteq S_0 : T \text{ is a 4-tree such that } \left( \frac{1}{|I_T|} \sum_{s \in T} \|f_{s_1}\|_2^2 \right)^{\frac{1}{2}} \geq \frac{\sigma}{2} \|f\|_2 \right\}. \quad (5.2)$$

If $\mathscr{F}_0 \neq \emptyset$, then take $T_1$ to be the 4-tree in $\mathscr{F}_0$ with top $t$ such that $c(\omega_{t_4}) \geq c(\omega_{t_4'})$ for any other $T \in \mathscr{F}_0$ with top $t'$. Let

$$\begin{cases} T_1^{(4)} := \text{maximal 4-tree in } S_0 \text{ with top } t, \\ T_1^{(1)} := \text{maximal 1-tree in } S_0 \text{ with top } t, \\ T_1^* := T_1^{(1)} \cup T_1^{(4)} \text{ (This is a tree with top } t). \end{cases}$$

Update $S_0$ and $\mathscr{F}$ by setting $S_0 := S_0 \setminus T_1^*$ and $\mathscr{F} := \mathscr{F} \cup \{T_1^*\}$.
Repeat this algorithm until there is no 4-tree in the updated $S_0$ satisfying

$$\left( \frac{1}{|I_T|} \sum_{s \in T} \|f_{s_1}\|_2^2 \right)^{\frac{1}{2}} \geq \frac{\sigma}{2} \|f\|_2.$$

When the algorithm terminates, we obtain

$$S_0 = P \setminus \{T_1^*, T_2^*, \dots, T_l^*\},$$
$$\mathscr{F} = \{T_1^*, T_2^*, \dots, T_l^*\}.$$

Simply let $P' = S_0$ and $P'' = \cup_{T \in \mathscr{F}} T$. Then Clearly $\text{size}_1(P', f) \leq \frac{\sigma}{2} \|f\|_2$.
Now we turn to the proof of $\sum_{T \in \mathscr{F}} |I_T| \lesssim \frac{1}{\sigma^2}$. We can assume that each $T \in \mathscr{F}$ is a 4-tree. By the definition of $\mathscr{F}_0$ (5.2), for any $T \in \mathscr{F}$,

$$\left( \frac{1}{|I_T|} \sum_{s \in T} \|f_{s_1}\|_2^2 \right)^{\frac{1}{2}} \geq \frac{\sigma}{2} \|f\|_2. \quad (5.3)$$

Therefore,

$$\sum_{T \in \mathscr{F}} |I_T| \lesssim \frac{1}{\sigma^2 \|f\|_2^2} \sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_1}\|_2^2.$$

It will suffice to prove

$$\sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_1}\|_2^2 \lesssim \|f\|_2^2. \quad (5.4)$$

For each 4-tile $s$, define an operator $A_s$ by $A_s f(x) = f_{s_1}(x)$. By Cauchy–Schwarz inequality,

$$\sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_j}\|_2^2 = \left\langle \sum_{T \in \mathscr{F}} \sum_{s \in T} A_s^* A_s f, f \right\rangle \leq \left\| \sum_{T \in \mathscr{F}} \sum_{s \in T} A_s^* A_s f \right\|_2 \|f\|_2.$$

Hence (5.4) follows from the following estimate:

$$\left\| \sum_{T \in \mathscr{F}} \sum_{s \in T} A_s^* A_s f \right\|_2 \lesssim \left( \sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_j}\|_2^2 \right)^{\frac{1}{2}}. \tag{5.5}$$

To prove (5.5), write

$$(\text{LHS of (5.5)})^2 = \sum_{T, T' \in \mathscr{F}} \sum_{\substack{s \in T \\ s' \in T'}} \left\langle A_s^* A_s f, A_{s'}^* A_{s'} f \right\rangle = I + II,$$

where

$$\begin{cases} I := \sum_{T \neq T' \in \mathscr{F}} \sum_{\substack{s \in T \\ s' \in T'}} \left\langle A_s^* A_s f, A_{s'}^* A_{s'} f \right\rangle, \\ II := \sum_{T \in \mathscr{F}} \sum_{s, s' \in T} \left\langle A_s^* A_s f, A_{s'}^* A_{s'} f \right\rangle. \end{cases}$$

Therefore, (5.5) follows from the estimate

$$\max\{I, II\} \lesssim \sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_1}\|_2^2. \tag{5.6}$$

We will only provide the estimate for $I$, as $II$ is easier to control so we omit the proof. Apply Cauchy–Schwarz inequality,

$$I \leq \sum_{T \neq T' \in \mathscr{F}} \sum_{\substack{s \in T \\ s' \in T'}} \|A_s f\|_2 \|A_s A_{s'}^*\| \|A_{s'} f\|_2.$$

Hence (5.6) is a consequence of the inequality below.

$$\sum_{T \neq T' \in \mathscr{F}} \sum_{\substack{s \in T \\ s' \in T'}} \|A_s f\|_2 \|A_s A_{s'}^*\| \|A_{s'} f\|_2 \lesssim \sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_1}\|_2^2. \tag{5.7}$$

The following estimate for $\|A_s A_{s'}^*\|$ is the key to sum up all the terms in the LHS of (5.7).

**Claim 5.1** $\|A_s A_{s'}^*\| \neq 0$ *only when* $\omega_{s_1} \cap \omega_{s_1'} \neq \emptyset$. *Moreover,*

$$\|A_s A_{s'}^*\| \lesssim_N \frac{|I_{s'}|^{\frac{1}{2}}}{|I_s|^{\frac{1}{2}}} \left(1 + \frac{dist(I_s, I_{s'})}{|I_s|}\right)^{-N} \quad if \quad \omega_{s_1} \subseteq \omega_{s_1'}. \tag{5.8}$$

*Proof* Write $A_s A_{s'}^* f(x) = \int K(x, y) f(y) \, dy$, 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 $\widetilde{\psi_{s_j'}} * \psi_{s_j}(t) = \int \overline{\widehat{\psi_{s'}}(\xi)} \widehat{\psi_s}(\xi) e^{2\pi i \xi t} \, d\xi$ is non-zero only when $\omega_{s_j} \cap \omega_{s_j'} \neq \emptyset$ by (2.3) and (2.10). Assume $\omega_{s_1} \subseteq \omega_{s_1'}$. By the definitions of $\chi_I^*$ (2.4) and $\psi_{k,l}$ and using the triangle inequality $(1 + |a|)^{-1} + (1 + |b|)^{-1} \leq (1 + |a + b|)^{-1}$,

$$|K(x, y)| \lesssim_N \left(1 + \frac{dist(x, I_s)}{|I_s|}\right)^{-2N} \left(1 + \frac{dist(y, I_{s'})}{|I_{s'}|}\right)^{-N}$$
$$\frac{1}{|I_s||I_{s'}|} \int \left(1 + \frac{|x - y - z|}{|I_{s'}|}\right)^{-2N} \left(1 + \frac{|z|}{|I_s|}\right)^{-N} dz$$
$$\lesssim_N \left(1 + \frac{dist(I_s, I_{s'})}{|I_s|}\right)^{-N} \frac{1}{|I_s|} \left(1 + \frac{dist(x, I_s)}{|I_s|}\right)^{-N}.$$

Hence

$$\int |K(x, y)| dx \lesssim_N \left(1 + \frac{dist(I_s, I_{s'})}{|I_s|}\right)^{-N}. \tag{5.9}$$

Similarly,

$$\int |K(x, y)| dy \lesssim_N \left(1 + \frac{dist(I_s, I_{s'})}{|I_s|}\right)^{-N} \frac{|I_{s'}|}{|I_s|}. \tag{5.10}$$

(5.9) and (5.10) imply (5.8) by Schur's lemma. $\qquad \square$

By the claim and symmetry, in the proof of (5.7) we will assume without loss of generality $\omega_{s_1} \subseteq \omega_{s_1'}$. We will also assume that $\omega_{s_1} \subsetneqq \omega_{s_1'}$, as the case $\omega_{s_1} = \omega_{s_1'}$ can be handled the same way. Under these assumptions, (5.7) has been reduced to

$$\sum_{\substack{T \neq T' \in \mathscr{F} \\ \omega_{s_1} \subsetneqq \omega_{s_1'}}} \sum_{s \in T, s' \in T'} \|A_s f\|_2 \|A_s A_{s'}^*\| \|A_{s'} f\|_2 \lesssim \sum_{T \in \mathscr{F}} \sum_{s \in T} \|f_{s_1}\|_2^2. \tag{5.11}$$

Since $\{s\}$ is a 4-tree and $\mathrm{size}_1(f, P) \leq \sigma \|f\|_2$,

$$\|A_s f\|_2 \leq |I_s|^{\frac{1}{2}} \sigma \|f\|_2. \tag{5.12}$$

Also notice that by (5.3)

$$\sigma \|f\|_2 \lesssim \left(|I_T|^{-1} \sum_{s_0 \in T} \|f_{(s_0)_1}\|_2^2\right)^{\frac{1}{2}}. \tag{5.13}$$

Combine (5.12) and (5.13), and we see that

$$\|A_s f\|_2 \lesssim |I_s|^{\frac{1}{2}} |I_T|^{-\frac{1}{2}} \left( \sum_{s_0 \in T} \|f_{(s_0)_1}\|_2^2 \right)^{\frac{1}{2}}. \qquad (5.14)$$

Similarly,

$$\|A_{s'} f\|_2 \lesssim |I_{s'}|^{\frac{1}{2}} |I_T|^{-\frac{1}{2}} \left( \sum_{s_0 \in T} \|f_{(s_0)_1}\|_2^2 \right)^{\frac{1}{2}}. \qquad (5.15)$$

Using (5.14) and (5.15), LHS of (5.11) is bounded by

$$\sum_{T \in \mathbf{T}} \left( \sum_{s_0 \in T} \|f_{(s_0)_1}\|_2^2 \right) \left( \sum_{\substack{s \in T, T' \neq T \\ s' \in T', \omega_{s_1} \subsetneq \omega_{s'_1}}} |I_s|^{\frac{1}{2}} |I_{s'}|^{\frac{1}{2}} |I_T|^{-1} \|A_s A_{s'}^*\| \right).$$

Therefore, (5.11) will be established once we show that for any $T \in \mathscr{F}$,

$$\sum_{\substack{s \in T, T' \neq T \\ s' \in T', \omega_{s_1} \subsetneq \omega_{s'_1}}} |I_s|^{\frac{1}{2}} |I_{s'}|^{\frac{1}{2}} |I_T|^{-1} \|A_s A_{s'}^*\| \lesssim 1.$$

By (5.8), this can be reduced to the estimate that for any $T \in \mathscr{F}$,

$$\sum_{\substack{s \in T, T' \neq T \\ s' \in T', \omega_{s_1} \subsetneq \omega_{s'_1}}} \left( 1 + \frac{\text{dist}(I_s, I_{s'})}{|I_s|} \right)^{-N} |I_{s'}| \lesssim |I_T|. \qquad (5.16)$$

To prove (5.16), we need a crucial observation.

**Claim 5.2** *If* $T_1 \neq T_2 \in \mathscr{F}$, $s \in T_1$, *and* $s' \in T_2$, *then*

$$\omega_{s_1} \subseteq \omega_{s'_1} \Rightarrow I_{s'} \cap I_{T_1} = \emptyset.$$

*Proof* Let $t$ and $t'$ denote the top of $T_1$ and $T_2$, respectively. Assume otherwise $I_{s'} \cap I_{T_1} \neq \emptyset$. Then $I_{s'} \subseteq I_t$. By (2.17) and the definition of tree, $\omega_{s'_1} \supseteq \omega_{s_4} \supseteq \omega_{t_4}$. Then $T_1$ is selected before $T_2$ as $c(\omega_{t_4}) > c(\omega_{t'_4})$. However, $s'_1 < t_1$ indicates that $s'$ should be selected together with $T_1$ according to the algorithm (See Fig. 5). This contradicts with the assumption that $s' \in T_2$.                                    $\square$

**Fig. 5** A crucial geometric observation
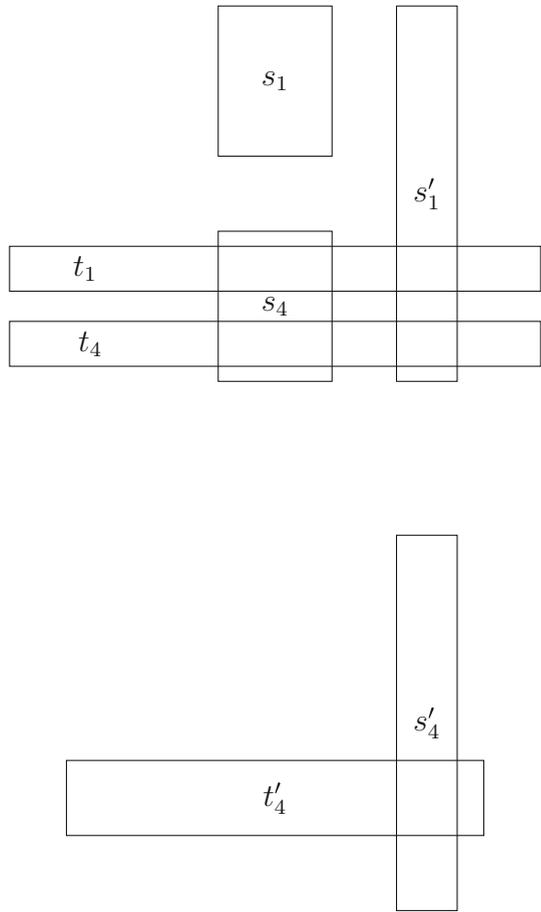


Now we are ready to prove (5.16). It is easy to see that

$$\text{LHS of (5.16)} \lesssim \sum_{s \in T} \sum_{\substack{T' \neq T \\ s' \in T', \omega_{s_1} \subsetneq \omega_{s_1'}}} \int_{I_{s'}} \left(1 + \frac{\text{dist}(I_s, x)}{|I_s|}\right)^{-N} dx.$$

By Claim 5.2, $I_{s'}$'s are pairwise disjoint and the union of these intervals is contained in $(I_T)^c$. Therefore,

$$\sum_{s \in T} \sum_{\substack{T' \neq T \\ s' \in T', \omega_{s_1} \subsetneq \omega_{s_1'}}} \int_{I_{s'}} \left(1 + \frac{\mathrm{dist}(I_s, x)}{|I_s|}\right)^{-N} dx$$

$$\leq \sum_{s \in T} \int_{(I_T)^c} \left(1 + \frac{\mathrm{dist}(I_s, x)}{|I_s|}\right)^{-N} dx \lesssim \sum_{s \in T} \left(1 + \frac{\mathrm{dist}(I_s, (I_T)^c)}{|I_s|}\right)^{-N} |I_s|$$

Using the tree structure of $T$ and the grid structure of tiles, it is easy to see that

$$\sum_{s \in T} \left(1 + \frac{\mathrm{dist}(I_s, (I_T)^c)}{|I_s|}\right)^{-N} |I_s| \lesssim |I_T|.$$

This proves (5.16).

## 6 Telescoping

We prove Theorem 1.3 by a telescoping argument. In what follows, $[x]$ will be used to denote the integer part of $x \in \mathbb{R}$.

Since $k \geq N \geq 10\alpha/\beta$, $[\frac{\beta}{\alpha}k]$ is large and essentially we have

$$T_N^{\alpha,\beta}(f_1, f_2, f_3)(x) = \sum_{k \geq N} H^{\alpha,k}(f_1, f_2)(x) f_3^{\beta,k}(x) =: A + B,$$

where

$$\begin{cases} A := \sum_{k \geq N} \left( \sum_{j=0}^{[(1-\frac{\beta}{\alpha})k]-1} \left( H^{\alpha,k-j}(f_1, f_2)(x) - H^{\alpha,k-j-1}(f_1, f_2)(x) \right) \right) f_3^{\beta,k}(x), \\ B := \sum_{k \geq N} H^{\alpha,[\frac{\beta}{\alpha}k]}(f_1, f_2)(x) f_3^{\beta,k}(x) = \sum_{k \geq N} H^{\beta,k}(f_1, f_2)(x) f_3^{\beta,k}(x). \end{cases}$$

$B$ has a much better tile structure than $T_N^{\alpha,\beta}$: See Figs. 2 and 6 for a comparison. Since $B$ is a part of $T^{\beta,\beta}$ and the proof of Theorem 1.2 is valid for any collection of scales $k$, boundedness of $B$ is obtained.

It remains to analyze the operator $A$. By a change of variable $k \to k + j$, we can write $A = I + II$, where

$$\begin{cases} I := \sum_{k \geq N} \left( H^{\alpha,k}(f_1, f_2)(x) - H^{\alpha,k-1}(f_1, f_2)(x) \right) \left( \sum_{j=0}^{[(\frac{\alpha}{\beta}-1)k]} f_3^{\beta,k+j}(x) \right), \\ II := \sum_{(k,j) \in P} \left( H^{\alpha,k}(f_1, f_2)(x) - H^{\alpha,k-1}(f_1, f_2)(x) \right) f_3^{\beta,k+j}(x). \end{cases}$$

Here $P$ is a finite set of indices, and $II$ should be considered as an error term, whose boundedness follows from Hölder and Lacey–Thiele's Theorem ([4,5]). To prove the

**Fig. 6** Tri-tile structure of $T^{\beta,\beta}$
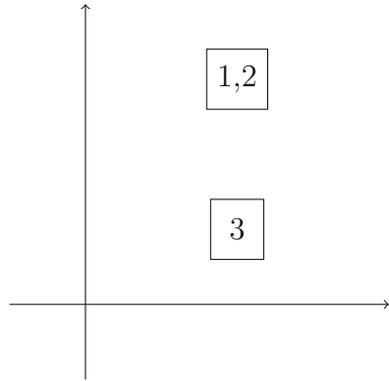


boundedness of the main term $I$, first note that

$$\sum_{j=0}^{[(\frac{\alpha}{\beta}-1)k]} f_3^{\beta,k+j}(x) = f_3^{\alpha k}(x) - f_3^{\beta k}(x),$$

where

$$f^l(x) := \int \hat{f}(\xi)\phi_0\left(\frac{\xi}{2^l}\right) d\xi, \; l \in \mathbb{R},$$

for some bump function $\phi_0$ supported in $[-1, 1]$. Hence we can write $I$ as the difference of two parts:

$$I = \sum_{k\geq1} \left(H^{\alpha,k}(f_1, f_2)(x) - H^{\alpha,k-1}(f_1, f_2)(x)\right) f_3^{\alpha k}(x)$$

$$- \sum_{k\geq1} \left(H^{\alpha,k}(f_1, f_2)(x) - H^{\alpha,k-1}(f_1, f_2)(x)\right) f_3^{\beta,k}(x).$$

Note that $H^{\alpha,k}(f_1, f_2)(x) - H^{\alpha,k-1}(f_1, f_2)(x)$ is a piece of BHT at scale $k$. Since $\alpha > \beta$ and $k > 0$, the supports of $\widehat{f_3^{\alpha k}}$ and $\widehat{f_3^{\beta k}}$ are at most as large as $2^{\alpha k}$. We can introduce a fourth function and do the wave packet decomposition to $f_1, f_2, f_4$. Then the tiles associated with these functions have structures similar to that of the tri-tiles as in the study of BHT. Therefore, the proof of Theorem 1.1 given in [2] still applies to $I$, and we omit the details. This finishes the proof of Theorem 1.3.

# References

1. Coifman, R., Meyer, Y.: Ondelettes et oprateurs. III. Hermann, Paris (1991)
2. Dong, D., Li, X.: On a hybrid of bilinear Hilbert transform and paraproduct, Acta. Math. Sin.-English Ser. https://doi.org/10.1007/s10114-017-6415-9
3. Grafakos, L., Li, X.: Uniform bounds for the bilinear Hilbert transforms. I. Ann. Math. (2) **159**(3), 889–933 (2004)
4. Lacey, M., Thiele, C.: $L^p$ estimates on the bilinear Hilbert transform for $2 < p < \infty$. Ann. Math. **146**, 693–724 (1997)
5. Lacey, M., Thiele, C.: On Calderón's conjecture. Ann. Math. (2) **149**(2), 475–496 (1999)
6. Li, X.: Uniform bounds for the bilinear Hilbert transforms. II. Rev. Mat. Iberoam. **22**(3), 1069–1126 (2006)
7. Li, X.: Uniform estimates for some paraproducts. N. Y. J. Math. **14**, 145–192 (2008)
8. Li, X.: Bilinear Hilbert transforms along curves I: The monomial case. Anal. PDE **6**(1), 197–220 (2013)
9. Muscalu, C., Schlag, W.: Classical and Multilinear Harmonic Analysis, Vol. II. Cambridge Studies in Advanced Mathematics, vol. 138. Cambridge University Press, Cambridge (2013)
10. Thiele, C.: A uniform estimate. Ann. Math. (2) **156**(2), 519—563 (2002)