Boundedness of the twisted paraproduct

Vjekoslav Kovač

Abstract. We prove L^p estimates for a two-dimensional bilinear operator of paraproduct type. This result answers a question posed by Demeter and Thiele.

1. Introduction and overview of results

Let us denote dyadic martingale averages and differences by

$$\mathbb{E}_k f := \sum_{|I|=2^{-k}} \left(\tfrac{1}{|I|}\textstyle\int_I f\right) \mathbf{1}_I, \quad \Delta_k f := \mathbb{E}_{k+1}f - \mathbb{E}_k f,$$

for every $k \in \mathbb{Z}$, where the sum is taken over dyadic intervals $I \subseteq \mathbb{R}$ of length 2^{-k}. When we apply an operator in only one variable of a two-dimensional function, we indicate it with a superscript. For instance,

$$(\mathbb{E}_k^{(1)}F)(x,y) := \left(\mathbb{E}_k F(\cdot,y)\right)(x).$$

The *dyadic twisted paraproduct* is defined as

$$(1.1) \qquad T_{\mathrm{d}}(F,G) := \sum_{k\in\mathbb{Z}} (\mathbb{E}_k^{(1)}F)(\Delta_k^{(2)}G).$$

In the continuous case, let P_φ denote the Fourier multiplier with symbol $\hat{\varphi}$, i.e.,

$$\mathrm{P}_\varphi f := f * \varphi.$$

Take two functions $\varphi, \psi \in \mathrm{C}^1(\mathbb{R})$ satisfying[1]

$$(1.2) \qquad |\varphi(x)|, \ |\tfrac{d}{dx}\varphi(x)|, \ |\psi(x)|, \ |\tfrac{d}{dx}\psi(x)| \ \lesssim \ (1+|x|)^{-3},$$

and

$$\mathrm{supp}(\hat{\psi}) \subseteq \left\{\xi \in \mathbb{R} \ : \ \tfrac{1}{2} \leq |\xi| \leq 2\right\}.$$

---

[1]For two nonnegative quantities A and B, we write $A \lesssim B$ if there exists an absolute constant $C \geq 0$ such that $A \leq CB$, and we write $A \lesssim_P B$ if $A \leq C_P B$ holds for some constant $C_P \geq 0$ depending on a parameter P. Finally, we write $A \sim_P B$ if both $A \lesssim_P B$ and $B \lesssim_P A$.

For every $k \in \mathbb{Z}$, define $\varphi_k(x) := 2^k \varphi(2^k x)$ and $\psi_k(x) := 2^k \psi(2^k x)$. The associated *continuous twisted paraproduct* is defined as

$$(1.3) \qquad T_c(F, G) := \sum_{k \in \mathbb{Z}} (P^{(1)}_{\varphi_k} F)(P^{(2)}_{\psi_k} G).$$

We are interested in strong-type estimates

$$(1.4) \qquad \|T(F, G)\|_{\mathrm{L}^{pq/(p+q)}(\mathbb{R}^2)} \lesssim_{p,q} \|F\|_{\mathrm{L}^p(\mathbb{R}^2)} \|G\|_{\mathrm{L}^q(\mathbb{R}^2)},$$

and weak-type estimates

$$(1.5) \quad \alpha \left|\left\{(x, y) \in \mathbb{R}^2 : |T(F, G)(x, y)| > \alpha\right\}\right|^{(p+q)/pq} \lesssim_{p,q} \|F\|_{\mathrm{L}^p(\mathbb{R}^2)} \|G\|_{\mathrm{L}^q(\mathbb{R}^2)}$$

for (1.1) and (1.3). The exponent $\frac{pq}{p+q}$ is mandated by scaling invariance. When $p = \infty$ or $q = \infty$, we interpret it as q or p respectively.

The main result of the paper establishes (1.4) and (1.5) for certain ranges of (p, q).

Theorem 1. (a) *The operators T_d and T_c satisfy the strong bound* (1.4) *if*

$$1 < p, q < \infty, \quad \tfrac{1}{p} + \tfrac{1}{q} > \tfrac{1}{2}.$$

(b) *Additionally, the operators T_d and T_c satisfy the weak bound* (1.5) *when*

$$p = 1, \ 1 \le q < \infty \quad or \quad q = 1, \ 1 \le p < \infty.$$

(c) *The weak estimate* (1.5) *fails for $p = \infty$ or $q = \infty$.*

The name *twisted paraproduct* was suggested by Camil Muscalu because there is a "twist" in the variables in which the convolutions (or the martingale projections) are performed, as opposed to the case of the ordinary paraproduct. No bounds on (1.1) or (1.3) were known prior to this work. A conditional result was shown by Bernicot in [1], assuming boundedness in some range, and expanding the range towards lower exponents using a fiberwise Calderón–Zygmund decomposition. We repeat his argument in the dyadic setting in Section 5, for the purpose of extending the boundedness region established in Sections 3 and 4.

Figure 1 depicts the range of exponents in Theorem 1. The shaded region satisfies the strong estimate, while for two solid sides of the unit square we only establish the weak estimates. The two dashed sides of the square represent exponents for which we show that even the weak estimate fails. The white triangle in the lower left corner is the region we do not tackle in this paper.

The proof of Theorem 1 is organized as follows. Sections 3 and 4 prove estimates for T_d in the interior of the triangle ABC. In Section 5 the remaining bounds for T_d are obtained. Section 6 establishes bounds for T_c by relating T_c to T_d. Finally, in Section 7 we discuss the counterexamples. In the closing section we sketch a simpler proof for points D and E only.
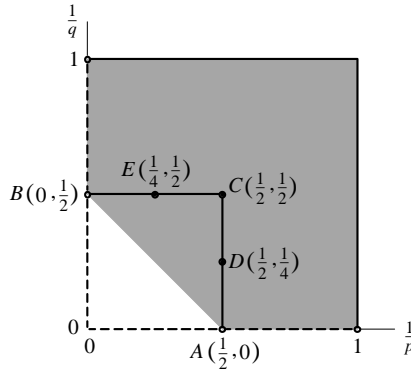
FIGURE 1. The range of exponents we discuss in this paper.

Several remarks. Before going into the proofs, we make several simple observations about T_{d}. Note that Theorem 1 also gives estimates for a family of shifted operators

$$(F, G) \mapsto \sum_{k \in \mathbb{Z}} (\mathbb{E}^{(1)}_{k+k_0} F)(\Delta^{(2)}_k G)$$

uniformly in $k_0 \in \mathbb{Z}$, because the last sum can be rewritten as

$$\mathrm{D}_{(2^{-k_0},1)} \, T_{\mathrm{d}}\big(\mathrm{D}_{(2^{k_0},1)}F, \, \mathrm{D}_{(2^{k_0},1)}G\big).$$

Here $\mathrm{D}_{(a,1)}$ denotes the non-isotropic dilation $(\mathrm{D}_{(a,1)}F)(x,y) := F(a^{-1}x, y)$.

If F and G are (say) compactly supported, then one can write

$$(1.6) \qquad T_{\mathrm{d}}(F, G) = FG - \sum_{k \in \mathbb{Z}} (\Delta^{(1)}_k F) \, (\mathbb{E}^{(2)}_{k+1} G).$$

Combining this with the previous remark and the fact that the pointwise product FG satisfies Hölder's inequality, we see that the set of estimates for $T_{\mathrm{d}}(F, G)$ is indeed symmetric under interchanging p and q, F and G. We use this fact to shorten some of the exposition below.

Furthermore, Theorem 1 implies bounds on more general dyadic operators of the following type:

$$(1.7) \qquad \Big\| \sum_{k \in \mathbb{Z}} c_k (\mathbb{E}^{(1)}_k F)(\Delta^{(2)}_k G) \Big\|_{\mathrm{L}^{pq/(p+q)}} \lesssim_{p,q} \|F\|_{\mathrm{L}^p} \|G\|_{\mathrm{L}^q},$$

for any numbers c_k such that $|c_k| \le 1$. Here we restrict ourselves to the interior range $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} > \frac{1}{2}$. One simply uses the known bound for $T_{\mathrm{d}}(F, \widetilde{G})$ with $\widetilde{G} := \sum_{k \in \mathbb{Z}} c_k \Delta^{(2)}_k G$, and the dyadic Littlewood–Paley inequality in the second variable. Note that the flexibility of having coefficients c_k is implicit in the definition of T_{c}, and indeed we will repeat a continuous variant of this argument in Section 6.

Some motivation. The one-dimensional bilinear Hilbert transform is an object that motivated most of the modern multilinear time-frequency analysis. Lacey and Thiele established its boundedness (in a certain range) in a pair of breakthrough papers, [8] and [9]. Recently, Demeter and Thiele investigated its two-dimensional analogue in [3]. For any two linear maps $A, B \colon \mathbb{R}^2 \to \mathbb{R}^2$ they considered

$$T_{A,B}(F,G)(x,y) := \text{p.v.} \int_{\mathbb{R}^2} F\big((x,y)-A(s,t)\big) G\big((x,y)-B(s,t)\big) K(s,t) ds dt,$$

where $K \colon \mathbb{R}^2 \setminus \{0,0\} \to \mathbb{C}$ is a Calderón–Zygmund kernel, i.e., $\hat{K}$ is a symbol satisfying

$$(1.8) \qquad |\partial^\alpha \hat{K}(\xi,\eta)| \lesssim_\alpha (\xi^2 + \eta^2)^{-|\alpha|/2}$$

for all derivatives ∂^α up to some large unspecified order. In [3], the bound

$$\|T_{A,B}(F,G)\|_{\mathrm{L}^{pq/(p+q)}(\mathbb{R}^2)} \lesssim_{A,B,p,q} \|F\|_{\mathrm{L}^p(\mathbb{R}^2)} \|G\|_{\mathrm{L}^q(\mathbb{R}^2)}$$

is proved in the range $2 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} > \frac{1}{2}$, and for most cases depending on A and B.

Some instances of A, B can be handled by an adaptation of the approach from [8], [9], while some cases lead the authors of [3] to invent a "one-and-a-half-dimensional" time-frequency analysis. At the other extreme, some instances of A, B degenerate to the one-dimensional bilinear Hilbert transform or the pointwise product. Up to the symmetry obtained by considering the adjoints, the only case of A, B that is left unresolved in [3] is

$$(1.9) \qquad T(F,G)(x,y) := \text{p.v.} \int_{\mathbb{R}^2} F(x-s,y)\, G(x,y-t)\, K(s,t)\, ds\, dt \, .$$

This case was denoted "Case 6", and as remarked there, it is largely degenerate but still nontrivial, so the usual wave-packet decompositions proved to be ineffective. It can also be viewed as the simplest example of higher-dimensional phenomena, i.e., complications not visible from the perspective of multilinear analysis arising in [8], [9], and even in quite general frameworks such as the ones in [12] or [2].

Theorem 1 establishes bounds on the *twisted bilinear multiplier* (1.9) for the special case of the symbol

$$\hat{K}(\xi,\eta) = \sum_{k \in \mathbb{Z}} \hat{\varphi}(2^{-k}\xi)\, \hat{\psi}(2^{-k}\eta) \, ,$$

i.e., the kernel

$$K(s,t) = \sum_{k \in \mathbb{Z}} 2^k \varphi(2^k s)\, 2^k \psi(2^k t) \, ,$$

with φ and ψ as in the introduction. A standard technique of "cone decomposition" (see [16]) then addresses general kernels K.

Our approach is to first work with the dyadic variant (1.1), and then use the square function introduced by Calderón and generalized by Jones, Seeger, and

Wright in [7] to transfer to the continuous case. Also, we dualize and prefer to consider the corresponding trilinear form

$$\Lambda_{\mathrm{d}}(F, G, H) := \int_{\mathbb{R}^2} T_{\mathrm{d}}(F, G)(x, y) \, H(x, y) \, dx \, dy \, .$$

Another reason why we call this object the twisted paraproduct is because the functions F, G, H are entwined in such a way that the trilinear form Λ_{d} does not split naturally into wavelet coefficients of each function separately, as it does for the ordinary paraproduct. As a substitute we introduce forms encoding "entwined wavelet coefficients", reminiscent of the Gowers box norm, which plays an important role in the proof. These forms keep functions intertwined, and we never attempt to break them but rather exploit their symmetries in an "induction on scales" type of argument.

A difference from the classical theory is that we gradually separate the functions F, G, H by repeated applications of the Cauchy–Schwarz inequality and a sort of telescoping identity that switches between the two variables. This is opposed to the usual approach to the ordinary paraproduct (even in the multiparameter case [10], [11]), where the Cauchy–Schwarz inequality is applied at once, and it immediately splits the form into governing operators like maximal and square functions (or their hybrids). This dominating procedure requires four steps for Λ_{d}, and generally finitely many steps for "more entwined" forms in higher dimensions, which are very briefly discussed in the closing section.

There seems to be many other higher dimensional phenomena worth studying. Another interesting two-dimensional object, more singular than the twisted paraproduct is

$$\mathrm{p.v.} \int_{\mathbb{R}} F(x - t, y) \, G(x, y - t) \, \frac{dt}{t} \, .$$

Its boundedness is still an open problem. One also has to notice that the yet more singular bi-parameter bilinear Hilbert transform

$$\mathrm{p.v.} \int_{\mathbb{R}^2} F(x - s, y - t) \, G(x + s, y + t) \, \frac{ds}{s} \frac{dt}{t}$$

does not satisfy any L^p estimates, as shown in [10].

2. A few words on the notation

A *dyadic interval* is an interval of the form $[2^k l, 2^k (l+1))$, for some integers k and l. For each dyadic interval I, we denote its left and right halves respectively by I_{left} and I_{right}. *Dyadic squares* and *dyadic rectangles* in $\mathbb{R}^2$ are defined in the obvious way. For any dyadic interval I, denote the Haar scaling function by $\varphi_I^{\mathrm{d}} := |I|^{-1/2} \mathbf{1}_I$ and the Haar wavelet by $\psi_I^{\mathrm{d}} := |I|^{-1/2}(\mathbf{1}_{I_{\mathrm{left}}} - \mathbf{1}_{I_{\mathrm{right}}})$.

Martingale averages and differences can be alternatively written in the Haar basis:
$$\mathbb{E}_k f = \sum_{|I|=2^{-k}} \left(\int_{\mathbb{R}} f \varphi_I^{\mathrm{d}} \right) \varphi_I^{\mathrm{d}}, \qquad \Delta_k f = \sum_{|I|=2^{-k}} \left(\int_{\mathbb{R}} f \psi_I^{\mathrm{d}} \right) \psi_I^{\mathrm{d}}.$$

In $\mathbb{R}^2$, every dyadic square Q partitions into four congruent dyadic squares that are called *children* of Q, and conversely, Q is said to be their *parent*.

In all of the following except in Section 6, the considered functions are assumed to be *nonnegative* dyadic step functions, i.e., positive finite linear combinations of characteristic functions of dyadic rectangles. This reduction is achieved by splitting into positive and negative, real and imaginary parts, and invoking density arguments.

Let $\mathcal{C}$ denote the collection of all dyadic squares in $\mathbb{R}^2$. Note that T_{d} and Λ_{d} can be rewritten as sums over $\mathcal{C}$:

$$T_{\mathrm{d}}(F,G)(x,y) = \sum_{I \times J \in \mathcal{C}} \int_{\mathbb{R}^2} F(u,y)G(x,v)\, \varphi_I^{\mathrm{d}}(u)\varphi_I^{\mathrm{d}}(x)\psi_J^{\mathrm{d}}(v)\psi_J^{\mathrm{d}}(y)\, du\, dv\,,$$

$$\Lambda_{\mathrm{d}}(F,G,H) = \sum_{I \times J \in \mathcal{C}} \int_{\mathbb{R}^4} F(u,y)G(x,v)H(x,y)$$
$$\varphi_I^{\mathrm{d}}(u)\varphi_I^{\mathrm{d}}(x)\psi_J^{\mathrm{d}}(v)\psi_J^{\mathrm{d}}(y)\, du\, dx\, dv\, dy\,.$$

In the subsequent discussion we will use a notion from additive combinatorics, namely the *Gowers box norm*. It is a two-dimensional variant of a series of norms introduced by Gowers in [4] and [5] to give quantitative bounds in Szemerédi's theorem, and was used by Shkredov in [13] to give bounds on sizes of sets that do not contain two-dimensional corners. Its occurrence in [14] is the one we find the most influential.

For any dyadic square $Q = I \times J$ we first define the *Gowers box inner product* of four functions F_1, F_2, F_3, F_4 as

$$[F_1, F_2, F_3, F_4]_{\square(Q)} := \frac{1}{|Q|^2} \int_I \int_I \int_J \int_J F_1(u,v)F_2(x,v)F_3(u,y)F_4(x,y)\, du\, dx\, dv\, dy.$$

Then for any function F we define the *Gowers box norm* by[2]

$$\|F\|_{\square(Q)} := [F, F, F, F]_{\square(Q)}^{1/4}.$$

It is easy to prove the *box Cauchy–Schwarz inequality*:

$$(2.1) \qquad [F_1, F_2, F_3, F_4]_{\square(Q)} \leq \|F_1\|_{\square(Q)} \|F_2\|_{\square(Q)} \|F_3\|_{\square(Q)} \|F_4\|_{\square(Q)}.$$

To see (2.1), one has to write $|Q|^2 [F_1, F_2, F_3, F_4]_{\square(Q)}$ as

$$\int_I \int_I \left(\int_J F_1(u,v)F_2(x,v)dv \right) \left(\int_J F_3(u,y)F_4(x,y)dy \right) du\, dx\,,$$

---

[2]If $F(x,y)$ restricted to Q is discretized and viewed as a matrix, then $\|F\|_{\square(Q)}$ can be recognized as its (properly normalized) Schatten 4-norm, i.e., the ℓ^4 norm of the sequence of its singular values. This comment gives yet one more immediate proof of inequality (2.2) below.

and apply the ordinary Cauchy–Schwarz inequality in $u, x \in I$. Then one rewrites the result as

$$\left(\int_J \int_J \left(\int_I F_1(u,v)F_1(u,y)du \right) \left(\int_I F_2(x,v)F_2(x,y)dx \right) dv\, dy \right)^{1/2}$$
$$\cdot \left(\int_J \int_J \left(\int_I F_3(u,v)F_3(u,y)du \right) \left(\int_I F_4(x,v)F_4(x,y)dx \right) dv\, dy \right)^{1/2},$$

and applies the Cauchy–Schwarz inequality again, this time in $v, y \in J$. From here it is also easily seen that $\|\cdot\|_{\square(Q)}$ is really a norm on functions supported on Q. On the other hand, a straightforward application of the (ordinary) Cauchy–Schwarz inequality yields

$$(2.2) \qquad \|F\|_{\square(Q)} \leq \left(\frac{1}{|Q|} \int_Q |F|^2 \right)^{1/2}.$$

An alternative way to verify (2.2) is to notice that it is a special case of the strong $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ estimate for the quadrilinear form

$$(F_1, F_2, F_3, F_4) \mapsto |Q|^2 [F_1, F_2, F_3, F_4]_{\square(Q)}.$$

Since $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is in the convex hull of $(1, 0, 0, 1)$ and $(0, 1, 1, 0)$, we can use complex interpolation, and it is enough to verify strong type estimates for the latter points, which is trivial.

3. Telescoping identities over trees

A *tree* is a collection $\mathcal{T}$ of dyadic squares in $\mathbb{R}^2$ such that there exists $Q_\mathcal{T} \in \mathcal{T}$, called the *root* of $\mathcal{T}$, satisfying $Q \subseteq Q_\mathcal{T}$ for every $Q \in \mathcal{T}$. A tree $\mathcal{T}$ is said to be *convex* if whenever $Q_1 \subseteq Q_2 \subseteq Q_3$, and $Q_1, Q_3 \in \mathcal{T}$, then also $Q_2 \in \mathcal{T}$. We will only be working with finite convex trees. A *leaf* of $\mathcal{T}$ is a square that is not contained in $\mathcal{T}$, but its parent is. The family of leaves of $\mathcal{T}$ will be denoted by $\mathcal{L}(\mathcal{T})$. Notice that for every finite convex tree $\mathcal{T}$, the squares in $\mathcal{L}(\mathcal{T})$ partition the root $Q_\mathcal{T}$.

For any finite convex tree $\mathcal{T}$ we define the local variant of Λ_{d} that sums over only the squares in $\mathcal{T}$, i.e.,

$$\Lambda_\mathcal{T}(F, G, H) := \sum_{I \times J \in \mathcal{T}} \int_{\mathbb{R}^4} F(u,y)G(x,v)H(x,y)$$
$$\cdot \varphi_I^{\mathrm{d}}(u)\, \varphi_I^{\mathrm{d}}(x)\, \psi_J^{\mathrm{d}}(v)\, \psi_J^{\mathrm{d}}(y)\, du\, dx\, dv\, dy.$$

It turns out to be handy to also introduce a slightly more general quadrilinear form:

$$\Theta_\mathcal{T}^{(2)}(F_1, F_2, F_3, F_4) := \sum_{I \times J \in \mathcal{T}} \int_{\mathbb{R}^4} F_1(u,v)F_2(x,v)F_3(u,y)F_4(x,y)$$
$$\cdot \varphi_I^{\mathrm{d}}(u)\, \varphi_I^{\mathrm{d}}(x)\, \psi_J^{\mathrm{d}}(v)\, \psi_J^{\mathrm{d}}(y)\, du\, dv\, dx\, dy,$$

and its modified counterpart

$$\Theta_{\mathcal{T}}^{(1)}(F_1, F_2, F_3, F_4) := \sum_{I \times J \in \mathcal{T}} \sum_{j \in \{\text{left,right}\}} \int_{\mathbb{R}^4} F_1(u,v) F_2(x,v) F_3(u,y) F_4(x,y)$$
$$\cdot \psi_I^{\mathrm{d}}(u) \psi_I^{\mathrm{d}}(x) \varphi_{J_j}^{\mathrm{d}}(v) \varphi_{J_j}^{\mathrm{d}}(y) \ du \, dv \, dx \, dy \,.$$

Note that in $\Theta_{\mathcal{T}}^{(1)}$ we actually sum over a certain collection of dyadic rectangles whose horizontal side is twice longer than the vertical one. This is just a technicality to make the arguments simpler at the cost of losing (geometric) symmetry. Also observe that $\Lambda_{\mathcal{T}}(F, G, H)$ can be recognized as $\Theta_{\mathcal{T}}^{(2)}(\mathbf{1}, G, F, H)$, where $\mathbf{1}$ is the constant function on $\mathbb{R}^2$.

Let us also denote for any collection $\mathcal{F}$ of dyadic squares:

$$(3.1) \qquad \Xi_{\mathcal{F}}(F_1, F_2, F_3, F_4) := \sum_{Q \in \mathcal{F}} |Q| \left[F_1, F_2, F_3, F_4 \right]_{\square(Q)},$$

or equivalently,

$$\Xi_{\mathcal{F}}(F_1, F_2, F_3, F_4) = \sum_{I \times J \in \mathcal{F}} \int_{\mathbb{R}^4} F_1(u,v) F_2(x,v) F_3(u,y) F_4(x,y)$$
$$(3.2) \qquad\qquad \cdot \varphi_I^{\mathrm{d}}(u) \varphi_I^{\mathrm{d}}(x) \varphi_J^{\mathrm{d}}(v) \varphi_J^{\mathrm{d}}(y) \ du \, dv \, dx \, dy \,.$$

The following lemma is the core of our method:

Lemma 2 (Telescoping identity). *For any finite convex tree $\mathcal{T}$ with root $Q_{\mathcal{T}}$ we have*

$$\Theta_{\mathcal{T}}^{(1)}(F_1, F_2, F_3, F_4) + \Theta_{\mathcal{T}}^{(2)}(F_1, F_2, F_3, F_4)$$
$$= \Xi_{\mathcal{L}(\mathcal{T})}(F_1, F_2, F_3, F_4) - \Xi_{\{Q_{\mathcal{T}}\}}(F_1, F_2, F_3, F_4) \,.$$

Proof. We first note that it is enough to verify the identity when $\mathcal{T}$ consists of only one square, as in general the right hand side can be expanded into a telescoping sum

$$\sum_{Q \in \mathcal{T}} \left(\sum_{\widetilde{Q} \text{ is a child of } Q} \Xi_{\{\widetilde{Q}\}} - \Xi_{\{Q\}} \right).$$

Here is where we use that $\mathcal{T}$ is convex, which means that each square $Q \in \mathcal{T} \setminus \{Q_{\mathcal{T}}\}$ has all four children and the parent in $\mathcal{T} \cup \mathcal{L}(\mathcal{T})$.

Second, observe that when $\mathcal{T}$ has only one square $I \times J$, then using (3.2) the identity reduces to showing

$$\sum_{j \in \{\text{left,right}\}} \psi_I^{\mathrm{d}}(u) \psi_I^{\mathrm{d}}(x) \varphi_{J_j}^{\mathrm{d}}(v) \varphi_{J_j}^{\mathrm{d}}(y) \ + \ \varphi_I^{\mathrm{d}}(u) \varphi_I^{\mathrm{d}}(x) \psi_J^{\mathrm{d}}(v) \psi_J^{\mathrm{d}}(y)$$
$$(3.3) \qquad = \sum_{i,j \in \{\text{left,right}\}} \varphi_{I_i}^{\mathrm{d}}(u) \varphi_{I_i}^{\mathrm{d}}(x) \varphi_{J_j}^{\mathrm{d}}(v) \varphi_{J_j}^{\mathrm{d}}(y) \ - \ \varphi_I^{\mathrm{d}}(u) \varphi_I^{\mathrm{d}}(x) \varphi_J^{\mathrm{d}}(v) \varphi_J^{\mathrm{d}}(y) \,,$$

multiplying by $F_1(u,v)F_2(x,v)F_3(u,y)F_4(x,y)$ and finally integrating. By adding and subtracting one extra term, equality (3.3) can be further rewritten as

$$\left(\sum_{j\in\{\text{left},\text{right}\}} \varphi^{\text{d}}_{J_j}(v)\varphi^{\text{d}}_{J_j}(y)\right)\left(\varphi^{\text{d}}_I(u)\varphi^{\text{d}}_I(x) + \psi^{\text{d}}_I(u)\psi^{\text{d}}_I(x) - \sum_{i\in\{\text{left},\text{right}\}} \varphi^{\text{d}}_{I_i}(u)\varphi^{\text{d}}_{I_i}(x)\right)$$
$$+ \varphi^{\text{d}}_I(u)\varphi^{\text{d}}_I(x)\left(\varphi^{\text{d}}_J(v)\varphi^{\text{d}}_J(y) + \psi^{\text{d}}_J(v)\psi^{\text{d}}_J(y) - \sum_{j\in\{\text{left},\text{right}\}} \varphi^{\text{d}}_{J_j}(v)\varphi^{\text{d}}_{J_j}(y)\right) = 0\,.$$

It remains to notice

$$\begin{aligned}
\varphi^{\text{d}}_I&(u)\varphi^{\text{d}}_I(x) + \psi^{\text{d}}_I(u)\psi^{\text{d}}_I(x)\\
&= |I|^{-1}\big(\mathbf{1}_{I_{\text{left}}}(u) + \mathbf{1}_{I_{\text{right}}}(u)\big)\big(\mathbf{1}_{I_{\text{left}}}(x) + \mathbf{1}_{I_{\text{right}}}(x)\big)\\
&\quad + |I|^{-1}\big(\mathbf{1}_{I_{\text{left}}}(u) - \mathbf{1}_{I_{\text{right}}}(u)\big)\big(\mathbf{1}_{I_{\text{left}}}(x) - \mathbf{1}_{I_{\text{right}}}(x)\big)\\
&= 2|I|^{-1}\mathbf{1}_{I_{\text{left}}}(u)\mathbf{1}_{I_{\text{left}}}(x) + 2|I|^{-1}\mathbf{1}_{I_{\text{right}}}(u)\mathbf{1}_{I_{\text{right}}}(x)\\
&= \varphi^{\text{d}}_{I_{\text{left}}}(u)\varphi^{\text{d}}_{I_{\text{left}}}(x) + \varphi^{\text{d}}_{I_{\text{right}}}(u)\varphi^{\text{d}}_{I_{\text{right}}}(x)\,,
\end{aligned}$$

and analogously

$$\varphi^{\text{d}}_J(v)\varphi^{\text{d}}_J(y) + \psi^{\text{d}}_J(v)\psi^{\text{d}}_J(y) = \varphi^{\text{d}}_{J_{\text{left}}}(v)\varphi^{\text{d}}_{J_{\text{left}}}(y) + \varphi^{\text{d}}_{J_{\text{right}}}(v)\varphi^{\text{d}}_{J_{\text{right}}}(y)\,.$$

$\square$

Let us remark that, since we assume $F_1, F_2, F_3, F_4 \geq 0$, we have

$$\Xi_{\{Q_{\mathcal{T}}\}}(F_1, F_2, F_3, F_4) \geq 0\,,$$

so tat the right hand side of the telescoping identity is less than or equal to $\Xi_{\mathcal{L}(\mathcal{T})}(F_1, F_2, F_3, F_4)$. We will use this observation without further mention.

The next lemma will be used to gradually control the forms $\Theta^{(1)}_{\mathcal{T}}$, $\Theta^{(2)}_{\mathcal{T}}$.

Lemma 3 (Reduction inequalities).

$$\begin{aligned}
\big|\Theta^{(1)}_{\mathcal{T}}(F_1, F_2, F_3, F_4)\big| &\leq \Theta^{(1)}_{\mathcal{T}}(F_1, F_1, F_3, F_3)^{1/2}\, \Theta^{(1)}_{\mathcal{T}}(F_2, F_2, F_4, F_4)^{1/2}\,,\\
\big|\Theta^{(2)}_{\mathcal{T}}(F_1, F_2, F_3, F_4)\big| &\leq \Theta^{(2)}_{\mathcal{T}}(F_1, F_2, F_1, F_2)^{1/2}\, \Theta^{(2)}_{\mathcal{T}}(F_3, F_4, F_3, F_4)^{1/2}\,.
\end{aligned}$$

Proof. Rewrite $\Theta^{(2)}_{\mathcal{T}}(F_1, F_2, F_3, F_4)$ as

$$\sum_{I\times J\in\mathcal{T}} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} F_1(u,v)F_2(x,v)\psi^{\text{d}}_J(v)dv\right)$$
$$\cdot \left(\int_{\mathbb{R}} F_3(u,y)F_4(x,y)\psi^{\text{d}}_J(y)dy\right) \varphi^{\text{d}}_I(u)\varphi^{\text{d}}_I(x)\, du\, dx\,,$$

and apply the Cauchy–Schwarz inequality, first over $(u,x)\in I\times I$, and then over $I\times J\in\mathcal{T}$. The inequality for $\Theta^{(1)}_{\mathcal{T}}$ is proved similarly. $\square$

Now we are ready to prove a local estimate, which will be "integrated" to a global one in the next section.

Proposition 4 (Single tree estimate). *For any finite convex tree $\mathcal{T}$ we have*

$$(3.4) \qquad \big|\Theta^{(2)}_{\mathcal{T}}(F_1, F_2, F_3, F_4)\big| \ \leq \ 2|Q_{\mathcal{T}}| \prod_{j=1}^{4} \max_{Q \in \mathcal{L}(\mathcal{T})} \|F_j\|_{\square(Q)} \ .$$

In particular,

$$(3.5) \quad \big|\Lambda_{\mathcal{T}}(F, G, H)\big| \leq 2|Q_{\mathcal{T}}| \big(\max_{Q \in \mathcal{L}(\mathcal{T})} \|F\|_{\square(Q)} \big) \big(\max_{Q \in \mathcal{L}(\mathcal{T})} \|G\|_{\square(Q)} \big) \big(\max_{Q \in \mathcal{L}(\mathcal{T})} \|H\|_{\square(Q)} \big).$$

Proof. The proof of (3.4) consists of several alternating applications of Lemmas 2 and 3. Start with four nonnegative functions[3] G_1, G_2, G_3 and G_4, and normalize:

$$(3.6) \qquad \max_{Q \in \mathcal{L}(\mathcal{T})} \|G_j\|_{\square(Q)} = 1 \,,$$

for $j = 1, 2, 3, 4$, since the inequality is homogenous. By scale invariance, we may also assume $|Q_{\mathcal{T}}| = 1$. Observe that $\Theta^{(2)}_{\mathcal{T}}(G_j, G_j, G_j, G_j) \geq 0$, since it can be written as

$$\sum_{I \times J \in \mathcal{T}} \int_{\mathbb{R}^2} \Big(\int_{\mathbb{R}} G_j(u, y) G_j(x, y) \psi^{\mathrm{d}}_J(y) dy \Big)^2 \varphi^{\mathrm{d}}_I(u) \varphi^{\mathrm{d}}_I(x) \, du dx \,.$$

Thus, from the telescoping identity we get

$$\Theta^{(1)}_{\mathcal{T}}(G_j, G_j, G_j, G_j) \leq \Xi_{\mathcal{L}(\mathcal{T})}(G_j, G_j, G_j, G_j) \,,$$

and then from (3.6), (3.1), and the fact that $\mathcal{L}(\mathcal{T})$ partitions $Q_{\mathcal{T}}$:

$$\Theta^{(1)}_{\mathcal{T}}(G_j, G_j, G_j, G_j) \leq |Q_{\mathcal{T}}| = 1 \,.$$

The reduction inequalities now also give

$$(3.7) \qquad\qquad \big|\Theta^{(1)}_{\mathcal{T}}(G_1, G_2, G_1, G_2)\big| \leq 1 \,,$$
$$\big|\Theta^{(1)}_{\mathcal{T}}(G_3, G_4, G_3, G_4)\big| \leq 1 \,.$$

Next, from (3.6), (2.1), and (3.1) one gets

$$(3.8) \qquad\qquad \Xi_{\mathcal{L}(\mathcal{T})}(G_1, G_2, G_1, G_2) \leq 1 \,,$$

while Lemma 2 gives

$$\Theta^{(2)}_{\mathcal{T}}(G_1, G_2, G_1, G_2) \leq \Xi_{\mathcal{L}(\mathcal{T})}(G_1, G_2, G_1, G_2) - \Theta^{(1)}_{\mathcal{T}}(G_1, G_2, G_1, G_2) \,,$$

and combining with (3.7), (3.8) yields

$$\Theta^{(2)}_{\mathcal{T}}(G_1, G_2, G_1, G_2) \leq 2 \,.$$

Completely analogously,

$$\Theta^{(2)}_{\mathcal{T}}(G_3, G_4, G_3, G_4) \leq 2 \,.$$

---

[3]We have changed the notation in the proof from F_j to G_j to avoid confusion, since Lemmas 2 and 3 will be applied for various choices of F_j.
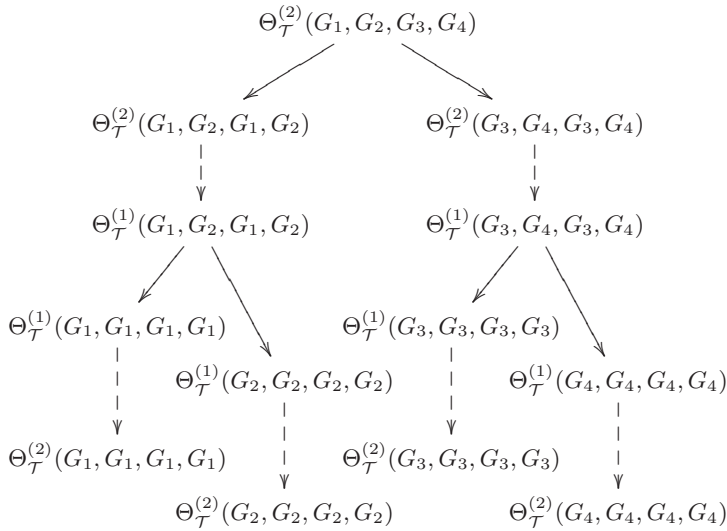
FIGURE 2. Schematic presentation of the proof of Proposition 4. A solid arrow denotes an application of the reduction inequality, while a broken arrow denotes an application of the telescoping identity.

By another application of Lemma 3 we end up with

$$\left|\Theta_{\mathcal{T}}^{(2)}(G_1, G_2, G_3, G_4)\right| \leq 2 \,,$$

and this establishes (3.4).

For the proof of (3.5) one has to substitute $F_1 = \mathbf{1}$, $F_2 = G$, $F_3 = F$, and $F_4 = H$ into (3.4). □

The above proof can be represented in the form of a tree diagram, as in Figure 2. We inductively bound terms starting from the bottom and proceeding to the top. The last row consists of nonnegative terms, allowing us to start the "induction". On every application of the telescoping identity we also get terms with $\Xi_{\mathcal{L}(\mathcal{T})}$, which we do not denote, and which are controlled by (3.6) and (2.1).

4. Proving the estimate in the local L^2 case

In this section we show the bound

$$(4.1) \qquad |\Lambda_{\mathrm{d}}(F, G, H)| \lesssim_{p,q,r} \|F\|_{\mathrm{L}^p}\|G\|_{\mathrm{L}^q}\|H\|_{\mathrm{L}^r}$$

for $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$, $2 < p, q, r < \infty$. By duality we get (1.4) for T_{d} in the range $2 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} > \frac{1}{2}$. The following material has become fairly standard, and indeed we follow closely ideas from [15], although in a much simpler setting.

Let us fix dyadic step functions $F, G, H \colon \mathbb{R}^2 \to [0, \infty)$, none of them identically 0. To make all arguments finite, in this section we restrict ourselves to considering only dyadic squares Q satisfying $|Q| \geq 2^{-2N}$ for some (large) fixed positive integer N. Since our bounds will be independent of N, letting $N \to \infty$ handles the whole collection $\mathcal{C}$.

We organize the family of dyadic squares in the following way. For any $k \in \mathbb{Z}$ we define the collection

$$\mathcal{P}_k^F := \Big\{ Q \ : \ 2^k \leq \sup_{Q' \supseteq Q} \|F\|_{\square(Q')} < 2^{k+1} \Big\},$$

and let $\mathcal{M}_k^F$ denote the family of maximal squares in $\mathcal{P}_k^F$ with respect to set inclusion. The collections $\mathcal{P}_k^G$, $\mathcal{M}_k^G$, $\mathcal{P}_k^H$, $\mathcal{M}_k^H$ are defined analogously. Furthermore, for any triple of integers k_1, k_2, k_3 we set

$$\mathcal{P}_{k_1, k_2, k_3} := \mathcal{P}_{k_1}^F \cap \mathcal{P}_{k_2}^G \cap \mathcal{P}_{k_3}^H,$$

and let $\mathcal{M}_{k_1, k_2, k_3}$ denote the family of maximal squares in $\mathcal{P}_{k_1, k_2, k_3}$.

For each $Q \in \mathcal{M}_{k_1, k_2, k_3}$ note that

$$\mathcal{T}_Q := \big\{ \widetilde{Q} \in \mathcal{P}_{k_1, k_2, k_3} \ : \ \widetilde{Q} \subseteq Q \big\}$$

is a finite convex tree with root Q, and that for different Q the corresponding trees $\mathcal{T}_Q$ occupy disjoint regions in $\mathbb{R}^2$. These trees decompose the collection $\mathcal{P}_{k_1, k_2, k_3}$, for each individual choice of k_1, k_2, k_3.

We apply Proposition 4 to each of the trees $\mathcal{T}_Q$. Consider any leaf $\widetilde{Q} \in \mathcal{L}(\mathcal{T}_Q)$, and denote its parent by Q'. From $Q' \in \mathcal{T}_Q \subseteq \mathcal{P}_{k_1, k_2, k_3}$ we get

$$\tfrac{1}{2} \|F\|_{\square(\widetilde{Q})} \leq \|F\|_{\square(Q')} < 2^{k_1+1} .$$

Thus, $\|F\|_{\square(\widetilde{Q})} \lesssim 2^{k_1}$, and similarly $\|G\|_{\square(\widetilde{Q})} \lesssim 2^{k_2}$, and $\|H\|_{\square(\widetilde{Q})} \lesssim 2^{k_3}$, so the "single tree estimate" (3.5) implies

$$\big| \Lambda_{\mathcal{T}_Q}(F, G, H) \big| \ \lesssim \ 2^{k_1+k_2+k_3} |Q| .$$

We split Λ_{d} into a sum of $\Lambda_{\mathcal{T}_Q}$ over all $k_1, k_2, k_3 \in \mathbb{Z}$ and all $Q \in \mathcal{M}_{k_1, k_2, k_3}$. In order to finish the proof of (4.1), it remains to show

$$(4.2) \qquad \sum_{k_1, k_2, k_3 \in \mathbb{Z}} 2^{k_1+k_2+k_3} \sum_{Q \in \mathcal{M}_{k_1, k_2, k_3}} |Q| \ \lesssim_{p,q,r} \ \|F\|_{\mathrm{L}^p} \|G\|_{\mathrm{L}^q} \|H\|_{\mathrm{L}^r} .$$

The trick from [15] is to observe that for any fixed triple $k_1, k_2, k_3 \in \mathbb{Z}$, squares in $\mathcal{M}_{k_1}^F$ cover squares in $\mathcal{M}_{k_1, k_2, k_3}$, and the latter are disjoint. The same is true for $\mathcal{M}_{k_2}^G$ and $\mathcal{M}_{k_3}^H$. Thus, it suffices to prove

$$\sum_{k_1, k_2, k_3 \in \mathbb{Z}} 2^{k_1+k_2+k_3} \min \Big(\sum_{Q \in \mathcal{M}_{k_1}^F} |Q|, \sum_{Q \in \mathcal{M}_{k_2}^G} |Q|, \sum_{Q \in \mathcal{M}_{k_3}^H} |Q| \Big)$$

$$(4.3) \qquad\qquad\qquad\qquad \lesssim_{p,q,r} \|F\|_{\mathrm{L}^p} \|G\|_{\mathrm{L}^q} \|H\|_{\mathrm{L}^r} .$$

Consider the following version of the dyadic maximal function:

$$\mathrm{M}_2 F := \sup_{Q \in \mathcal{C}} \Big(\frac{1}{|Q|} \int_Q |F|^2 \Big)^{1/2} \mathbf{1}_Q \,.$$

For each $Q \in \mathcal{M}_k^F$, from (2.2) and $\|F\|_{\Box(Q)} \geq 2^k$ we have $Q \subseteq \{\mathrm{M}_2 F \geq 2^k\}$, and by disjointness,

$$\sum_{Q \in \mathcal{M}_k^F} |Q| \leq |\{\mathrm{M}_2 F \geq 2^k\}| \,.$$

Also note that

$$\sum_{k \in \mathbb{Z}} 2^{pk} |\{\mathrm{M}_2 F \geq 2^k\}| \ \sim_p \ \|\mathrm{M}_2 F\|_{\mathrm{L}^p}^p \ \lesssim_p \ \|F\|_{\mathrm{L}^p}^p \,,$$

because M_2 is bounded on $\mathrm{L}^p(\mathbb{R}^2)$ for $2 < p < \infty$. Therefore,

(4.4) $$\sum_{k \in \mathbb{Z}} 2^{pk} \sum_{Q \in \mathcal{M}_k^F} |Q| \ \lesssim_p \ \|F\|_{\mathrm{L}^p}^p \,,$$

and completely analogously we get

$$\sum_{k \in \mathbb{Z}} 2^{qk} \sum_{Q \in \mathcal{M}_k^G} |Q| \ \lesssim_q \ \|G\|_{\mathrm{L}^q}^q \,, \quad \sum_{k \in \mathbb{Z}} 2^{rk} \sum_{Q \in \mathcal{M}_k^H} |Q| \ \lesssim_r \ \|H\|_{\mathrm{L}^r}^r \,.$$

A purely algebraic "integration lemma" stated and proved in [15] deduces (4.3) from these three estimates. The idea is to split the sum in (4.3) into three parts, depending on which of the numbers

$$\frac{2^{pk_1}}{\|F\|_{\mathrm{L}^p}^p}, \ \frac{2^{qk_2}}{\|G\|_{\mathrm{L}^q}^q}, \ \frac{2^{rk_3}}{\|H\|_{\mathrm{L}^r}^r}$$

is the largest. For instance, the part of the sum over

$$S_1 := \Big\{ (k_1, k_2, k_3) \in \mathbb{Z}^3 : \frac{2^{pk_1}}{\|F\|_{\mathrm{L}^p}^p} \geq \frac{2^{qk_2}}{\|G\|_{\mathrm{L}^q}^q}, \ \frac{2^{pk_1}}{\|F\|_{\mathrm{L}^p}^p} \geq \frac{2^{rk_3}}{\|H\|_{\mathrm{L}^r}^r} \Big\}$$

is controlled as

$$\sum_{k_1 \in \mathbb{Z}} \frac{2^{pk_1}}{\|F\|_{\mathrm{L}^p}^p} \Big(\sum_{Q \in \mathcal{M}_{k_1}^F} |Q| \Big) \sum_{\substack{k_2, k_3 \, \in \, \mathbb{Z} \\ (k_1, k_2, k_3) \, \in \, S_1}} \Big(\frac{2^{qk_2}/\|G\|_{\mathrm{L}^q}^q}{2^{pk_1}/\|F\|_{\mathrm{L}^p}^p} \Big)^{\frac{1}{q}} \Big(\frac{2^{rk_3}/\|H\|_{\mathrm{L}^r}^r}{2^{pk_1}/\|F\|_{\mathrm{L}^p}^p} \Big)^{\frac{1}{r}} \lesssim_{p,q,r} 1 \,,$$

which follows from (4.4) and by summing two convergent geometric series with their largest terms at most 1, and ratios equal to $\frac{1}{2}$.

5. Extending the range of exponents

The extension of the main estimate to the range $p \leq 2$ or $q \leq 2$ follows from the conditional result of Bernicot, [1]. Here we repeat his argument in the dyadic case, where it is a bit simpler. His idea is to use the one-dimensional Calderón–Zygmund decomposition in each fiber $F(\cdot, y)$ or $G(x, \cdot)$.

We start with an estimate obtained in the previous section:

$$(5.1) \qquad \|T_{\mathrm{d}}(F, G)\|_{\mathrm{L}^{pq/(p+q), \infty}} \leq \|T_{\mathrm{d}}(F, G)\|_{\mathrm{L}^{pq/(p+q)}} \lesssim_{p,q} \|F\|_{\mathrm{L}^p} \|G\|_{\mathrm{L}^q} ,$$

for some $2 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} > \frac{1}{2}$. If we prove the weak estimate

$$(5.2) \qquad \|T_{\mathrm{d}}(F, G)\|_{\mathrm{L}^{p/(p+1), \infty}} \lesssim_{p,q} \|F\|_{\mathrm{L}^p} \|G\|_{\mathrm{L}^1} ,$$

then T_{d} will be bounded in the whole range of Theorem 1, by real interpolation of multilinear operators, as stated for instance in [6] or [16]. We first cover the case $p > 2$, $q \leq 2$, then use (1.6) for $p \leq 2$, $q > 2$, and finally repeat the argument to tackle the case $p, q \leq 2$.

By homogeneity we may assume $\|F\|_{\mathrm{L}^p} = \|G\|_{\mathrm{L}^1} = 1$. For each $x \in \mathbb{R}$ denote by $\mathcal{J}_x$ the collection of all maximal dyadic intervals J with the property

$$\frac{1}{|J|} \int_J |G(x, y)| \, dy > 1 .$$

Furthermore, set

$$E := \bigcup_{x \in \mathbb{R}} \bigcup_{J \in \mathcal{J}_x} (\{x\} \times J) .$$

By our qualitative assumptions on G, the set E is simply a finite union of dyadic rectangles. Using the disjointness of the $J \in \mathcal{J}_x$,

$$(5.3) \qquad |E| = \int_{\mathbb{R}} \sum_{J \in \mathcal{J}_x} |J| \, dx \leq \int_{\mathbb{R}} \Big(\sum_{J \in \mathcal{J}_x} \int_J |G(x, y)| \, dy \Big) dx \leq 1 .$$

Next, we define "the good part" of G by

$$\widetilde{G}(x, y) := \begin{cases} \frac{1}{|J|} \int_J G(x, v) dv, & \text{for } y \in J \in \mathcal{J}_x \\ G(x, y), & \text{for } (x, y) \notin E \end{cases}$$

By the construction of $\mathcal{J}_x$ we have $\|\widetilde{G}\|_{\mathrm{L}^\infty} \leq 2$, and from $\|\widetilde{G}\|_{\mathrm{L}^1} \leq 1$ we also get $\|\widetilde{G}\|_{\mathrm{L}^q} \leq 2$, so using the known estimate (5.1) we obtain

$$(5.4) \qquad \big| \{ (x, y) : |T_{\mathrm{d}}(F, \widetilde{G})(x, y)| > 1 \} \big| \lesssim_{p,q} 1 .$$

As the last ingredient, we show that

$$(5.5) \qquad \Big(\int_{\mathbb{R}} \big(G(x, v) - \widetilde{G}(x, v) \big) \psi_{J'}^{\mathrm{d}}(v) dv \Big) \psi_{J'}^{\mathrm{d}}(y) = 0$$

for every $J' \in \mathcal{D}$, whenever $(x, y) \notin E$. Since $G(x, \cdot) - \widetilde{G}(x, \cdot)$ is supported on $\bigcup_{J \in \mathcal{J}_x} J$, this in turn will follow from

$$(5.6) \qquad \Big(\int_{\mathbb{R}} \big(G(x, v) - \widetilde{G}(x, v) \big) \, \psi_{J'}^{\mathrm{d}}(v) \, \mathbf{1}_J(v) \, dv \Big) \, \psi_{J'}^{\mathrm{d}}(y) = 0$$

for every $J \in \mathcal{J}_x$. In order to verify (5.6) it is enough to consider $J \cap J' \neq \emptyset$ and $y \in J'$, and since $y \notin J$, we conclude that J is strictly contained in J'. In this case $\psi_{J'}^{\mathrm{d}}(v) \, \mathbf{1}_J(v) = \pm |J'|^{-1/2} \mathbf{1}_J(v)$, so we only have to observe $\int_J \big(G(x, v) - \widetilde{G}(x, v) \big) dv = 0$, by the definition of $\widetilde{G}$.

Equation (5.5) immediately gives $T_{\mathrm{d}}(F, G - \widetilde{G})(x, y) = 0$ for $(x, y) \notin E$, so

$$\big\{ (x, y) : |T_{\mathrm{d}}(F, G)(x, y)| > 1 \big\} \subseteq E \cup \big\{ (x, y) : |T_{\mathrm{d}}(F, \widetilde{G})(x, y)| > 1 \big\},$$

and then, from (5.3) and (5.4),

$$\big| \big\{ (x, y) : |T_{\mathrm{d}}(F, G)(x, y)| > 1 \big\} \big| \lesssim_{p, q} 1.$$

This establishes (5.2) by dyadic scaling.

6. Transition to the continuous case

Now we turn to the task of proving strong estimates for T_{c} in the range from part (a) of Theorem 1:

$$\| T_{\mathrm{c}}(F, G) \|_{\mathrm{L}^{pq/(p+q)}} \lesssim_{p, q} \| F \|_{\mathrm{L}^p} \| G \|_{\mathrm{L}^q}$$

for $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} > \frac{1}{2}$. In order to get the boundary weak estimates, one can proceed as in [1].

Let φ and ψ be as in the Introduction. If $\int_{\mathbb{R}} \varphi = 0$, then $T_{\mathrm{c}}(F, G)$ is dominated by

$$\Big(\sum_{k \in \mathbb{Z}} |\mathrm{P}_{\varphi_k}^{(1)} F|^2 \Big)^{1/2} \Big(\sum_{k \in \mathbb{Z}} |\mathrm{P}_{\psi_k}^{(2)} G|^2 \Big)^{1/2},$$

and it is enough to use bounds for the two square functions. Otherwise, we have $0 < |\int_{\mathbb{R}} \varphi| \lesssim 1$, so let us normalize $\int_{\mathbb{R}} \varphi = 1$.

A tool that comes in very handy here is the square function appearing in the work of Calderón and generalized by Jones, Seeger, and Wright [7]. It effectively compares convolutions to martingale averages, allowing us to make the transition easily.

Proposition 5 ([7]). *Let φ be a function satisfying (1.2) and $\int_{\mathbb{R}} \varphi = 1$. The square function*

$$\mathcal{S}_{\mathrm{JSW}, \varphi} f := \Big(\sum_{k \in \mathbb{Z}} \big| \mathrm{P}_{\varphi_k} f - \mathbb{E}_k f \big|^2 \Big)^{1/2}$$

is bounded from $\mathrm{L}^p(\mathbb{R})$ to $\mathrm{L}^p(\mathbb{R})$ for $1 < p < \infty$, with a constant depending only on p.

Let ϕ be a nonnegative C^∞ function such that $\hat{\phi}(\xi) = 1$ for $|\xi| \leq 2^{-0.6}$, and $\hat{\phi}(\xi) = 0$ for $|\xi| \geq 2^{-0.4}$. We regard it as fixed, so we do not keep track of the dependence of the constants on ϕ. For any $a \in \mathbb{R}$ define ϕ_a, ϑ_a, ρ_a by

$$\hat{\phi}_a(\xi) := \hat{\phi}(2^{-a}\xi),$$
$$\hat{\vartheta}_a(\xi) := \hat{\phi}(2^{-a-1}\xi) - \hat{\phi}(2^{-a}\xi) = \hat{\phi}_{a+1}(\xi) - \hat{\phi}_a(\xi),$$
$$\hat{\rho}_a(\xi) := \hat{\phi}(2^{-a-0.6}\xi) - \hat{\phi}(2^{-a-0.5}\xi),$$

so that in particular

(6.1)
$$\hat{\vartheta}_a = 1 \quad \text{on } \operatorname{supp}(\hat{\rho}_a),$$

(6.2)
$$\sum_{i=-20}^{20} \hat{\rho}_{k+0.1i} = 1 \quad \text{on } \operatorname{supp}(\hat{\psi}_k),$$

(6.3)
$$\sum_{i=-20}^{20} \hat{\rho}_{k+0.1i} = 0 \quad \text{on } \operatorname{supp}(\hat{\psi}_{k'}) \text{ if } |k' - k| \geq 10.$$

We first use Proposition 5 to obtain bounds for a special case of our continuous twisted paraproduct:

(6.4)
$$T_{\varphi,\vartheta,b}(F, G) := \sum_{k\in\mathbb{Z}} (\mathrm{P}^{(1)}_{\varphi_k} F)\,(\mathrm{P}^{(2)}_{\vartheta_{k+b}} G),$$

where $b \in \mathbb{R}$ is a fixed parameter. The constants can depend on b, as later b will take only finitely many concrete values. Since we have already established estimates for (1.1), it is enough to bound their difference:

(6.5)
$$\big\| T_{\varphi,\vartheta,b}(F, G) - T_{\mathrm{d}}(F, G) \big\|_{\mathrm{L}^{pq/(p+q)}} \lesssim_{p,q,b} \|F\|_{\mathrm{L}^p}\, \|G\|_{\mathrm{L}^q}.$$

We introduce a mixed-type operator

$$T_{\mathrm{aux},b}(F, G) := \sum_{k\in\mathbb{Z}} (\mathbb{E}^{(1)}_k F)\,(\mathrm{P}^{(2)}_{\vartheta_{k+b}} G).$$

Using the Cauchy–Schwarz inequality in $k \in \mathbb{Z}$, one gets

$$\big| T_{\varphi,\vartheta,b}(F, G) - T_{\mathrm{aux},b}(F, G) \big| \leq \Big(\sum_{k\in\mathbb{Z}} \big|\mathrm{P}^{(1)}_{\varphi_k} F - \mathbb{E}^{(1)}_k F\big|^2 \Big)^{1/2} \Big(\sum_{k\in\mathbb{Z}} \big|\mathrm{P}^{(2)}_{\vartheta_{k+b}} G\big|^2 \Big)^{1/2}.$$

The first term on the right hand side is $\mathcal{S}^{(1)}_{\mathrm{JSW},\varphi} F$, while the second term is the ordinary square function in the second variable, as $\int_{\mathbb{R}} \vartheta_b = 0$. Next, one can rewrite $T_{\mathrm{aux},b}$ and T_{d} as

$$T_{\mathrm{aux},b}(F, G) = FG - \sum_{k\in\mathbb{Z}} (\Delta^{(1)}_k F)(\mathrm{P}^{(2)}_{\phi_{k+1+b}} G),$$
$$T_{\mathrm{d}}(F, G) = FG - \sum_{k\in\mathbb{Z}} (\Delta^{(1)}_k F)(\mathbb{E}^{(2)}_{k+1} G).$$

Subtracting and using the Cauchy–Schwarz inequality in $k \in \mathbb{Z}$, this time we obtain

$$\left| T_{\text{aux},b}(F,G) - T_{\text{d}}(F,G) \right| \leq \left(\sum_{k \in \mathbb{Z}} \left| \Delta_k^{(1)} F \right|^2 \right)^{1/2} \left(\sum_{k \in \mathbb{Z}} \left| \text{P}_{\phi_{k+b}}^{(2)} G - \mathbb{E}_k^{(2)} G \right|^2 \right)^{1/2}.$$

The first term on the right hand side is just the dyadic square function in the first variable, while the second term is $\mathcal{S}_{\text{JSW},\phi_b}^{(2)} G$. The estimate (6.5) now follows from Proposition 5 and bounds on the two common square functions.

Actually, we need a paraproduct "sparser" than the one in (6.4):

$$(6.6) \qquad T_{\varphi,\rho,b,l}^{10\mathbb{Z}}(F,G) := \sum_{j \in \mathbb{Z}} (\text{P}_{\varphi_{10j+l}}^{(1)} F)\,(\text{P}_{\rho_{10j+l+b}}^{(2)} G),$$

for any $l = 0, 1, \ldots, 9$. To see that (6.6) is bounded too, we define

$$\widetilde{G}_{b,l} := \sum_{j \in \mathbb{Z}} \text{P}_{\rho_{10j+l+b}}^{(2)} G.$$

Notice that because of (6.1) we have

$$\text{P}_{\vartheta_{k+b}}^{(2)} \widetilde{G}_{b,l} = \begin{cases} \text{P}_{\rho_{10j+l+b}}^{(2)} G & \text{for } k = 10j + l \in 10\mathbb{Z} + l\,, \\ 0 & \text{for } k \in \mathbb{Z},\ k \notin 10\mathbb{Z} + l\,, \end{cases}$$

and the Littlewood–Paley inequality gives

$$\|\widetilde{G}_{b,l}\|_{\text{L}^q} \lesssim_{q,b,l} \|G\|_{\text{L}^q}.$$

It remains to write

$$T_{\varphi,\rho,b,l}^{10\mathbb{Z}}(F,G) = T_{\varphi,\vartheta,b}(F,\widetilde{G}_{b,l})\,,$$

and use the boundedness of (6.4).

Finally, we tackle the original operator (1.3). The following computation is possible because of (6.2) and (6.3):

$$\sum_{k \in \mathbb{Z}} \hat{\varphi}_k(\xi)\,\hat{\psi}_k(\eta) = \sum_{l=0}^{9} \sum_{j \in \mathbb{Z}} \hat{\varphi}_{10j+l}(\xi)\,\hat{\psi}_{10j+l}(\eta)$$

$$= \sum_{l=0}^{9} \sum_{i=-20}^{20} \sum_{j \in \mathbb{Z}} \hat{\varphi}_{10j+l}(\xi)\,\hat{\rho}_{10j+l+0.1i}(\eta)\,\hat{\psi}_{10j+l}(\eta)$$

$$= \sum_{l=0}^{9} \sum_{i=-20}^{20} \sum_{j \in \mathbb{Z}} \hat{\varphi}_{10j+l}(\xi)\,\hat{\rho}_{10j+l+0.1i}(\eta)\,\hat{\Psi}_l(\eta)$$

Above we have set $\Psi_l := \sum_{m \in \mathbb{Z}} \psi_{10m+l}$. This "symbol identity" leads us to

$$(6.7) \qquad T_{\text{c}}(F,G) = \sum_{l=0}^{9} \sum_{i=-20}^{20} T_{\varphi,\rho,\,0.1i,\,l}^{10\mathbb{Z}}(F, \text{P}_{\Psi_l}^{(2)} G)\,.$$

Since $\hat\psi$ has compact support and $|\hat\psi(\eta)|, |\frac{d}{d\eta}\hat\psi(\eta)| \lesssim 1$ by (1.2), scaling gives $|\hat\Psi_l(\eta)| \lesssim 1$, $|\frac{d}{d\eta}\hat\Psi_l(\eta)| \lesssim |\eta|^{-1}$, and thus the Hörmander–Mikhlin multiplier theorem (in one variable) implies

$$\big\|\mathrm{P}^{(2)}_{\Psi_l}G\big\|_{\mathrm{L}^q} \lesssim_{q,l} \|G\|_{\mathrm{L}^q}.$$

It remains to use (6.7) and the boundedness of (6.6).

7. Endpoint counterexamples

We give the arguments in the dyadic setting, the continuous case being similar. First we show that T_d does not map boundedly

$$\mathrm{L}^\infty(\mathbb{R}^2) \times \mathrm{L}^q(\mathbb{R}^2) \to \mathrm{L}^{q,\infty}(\mathbb{R}^2)$$

for $1 \le q < \infty$. Take G to be

$$G(x,y) := \mathbf{1}_{[0,2^{-n})}(x) \sum_{k=1}^{n} R_k(y),$$

for some positive integer n, where R_k denotes the k-th Rademacher function[4] on $[0,1)$, i.e.,

$$R_k := \sum_{J \subseteq [0,1),\, |J|=2^{-k+1}} \big(\mathbf{1}_{J_{\mathrm{left}}} - \mathbf{1}_{J_{\mathrm{right}}}\big).$$

Recall Khintchine's inequality, which can be formulated as:

$$\Big\|\sum_{k=1}^{n} c_k R_k\Big\|_{\mathrm{L}^q} \sim_q \Big(\sum_{k=1}^{n} |c_k|^2\Big)^{1/2}, \quad \text{for } 0 < q < \infty,$$

giving us $\|G\|_{\mathrm{L}^q} \sim_q 2^{-n/q}n^{1/2}$. Observe that

$$(\Delta^{(2)}_k G)(x,y) = \mathbf{1}_{[0,2^{-n})}(x)R_{k+1}(y), \quad \text{for } k = 0,1,\dots,n-1.$$

We choose F supported in the unit square $[0,1)^2$ and defined by

$$F(x,y) := \begin{cases} 2R_j(y) - R_{j+1}(y) & \text{for } x \in [2^{-j}, 2^{-j+1}),\ j = 1,\dots,n-1, \\ R_n(y) & \text{for } x \in [0, 2^{-n+1}). \end{cases}$$

Note that $\|F\|_{\mathrm{L}^\infty} \le 3$ and $(\mathbb{E}^{(1)}_k F)(x,y) = R_{k+1}(y)$ for $x \in [0,2^{-n})$, $k = 0,1,\dots,n-1$. Since the output function is now simply $T_\mathrm{d}(F,G) = n\,\mathbf{1}_{[0,2^{-n})\times[0,1)}$, we have

$$\frac{\|T_\mathrm{d}(F,G)\|_{\mathrm{L}^{q,\infty}}}{\|F\|_{\mathrm{L}^\infty}\|G\|_{\mathrm{L}^q}} \gtrsim_q \frac{2^{-n/q}n}{2^{-n/q}n^{1/2}} = n^{1/2},$$

which shows unboundedness.

---

[4] Linear combinations of Rademacher functions $\sum_k c_k R_k(t)$ are dyadic analogues of lacunary trigonometric series $\sum_k c_k e^{i2^k t}$.

The remaining estimate $\|T_{\mathrm{d}}(F,G)\|_{\mathrm{L}^\infty} \lesssim \|F\|_{\mathrm{L}^\infty}\|G\|_{\mathrm{L}^\infty}$ is even easier to disprove. For a positive integer n, take

$$F(x,y) := \begin{cases} 1 & \text{for } x \in \bigcup_{j=0}^{n-1}[2^{-2j-1}, 2^{-2j}),\ y \in [0,1)\,, \\ 0 & \text{otherwise,} \end{cases}$$

and $G(x,y) := F(y,x)$. It is easy to see that $|T_{\mathrm{d}}(F,G)(x,y)| \sim n$ on the square $(x,y) \in [0, 2^{-2n})^2$.

8. Closing remarks

The primary purpose of the decomposition into trees in Section 4 is for proving the estimate for a larger range of exponents. If one is content with just having estimates in some nontrivial range, then a simpler proof can be given. Using Lemma 3,

$$(8.1) \qquad |\Lambda_{\mathrm{d}}(F,G,H)| \leq \Theta_{\mathcal{C}}^{(2)}(\mathbf{1},G,\mathbf{1},G)^{1/2}\, \Theta_{\mathcal{C}}^{(2)}(F,H,F,H)^{1/2}\,,$$

$$(8.2) \qquad \left|\Theta_{\mathcal{C}}^{(1)}(F,H,F,H)\right| \leq \Theta_{\mathcal{C}}^{(1)}(F,F,F,F)^{1/2}\, \Theta_{\mathcal{C}}^{(1)}(H,H,H,H)^{1/2}\,.$$

If in Lemma 2 one lets a single tree $\mathcal{T}$ exhaust the family of all dyadic squares, then the telescoping identity becomes simply

$$\Theta_{\mathcal{C}}^{(1)}(F_1,F_2,F_3,F_4) + \Theta_{\mathcal{C}}^{(2)}(F_1,F_2,F_3,F_4) = \int_{\mathbb{R}^2} F_1 F_2 F_3 F_4\,.$$

Particular instances of this equality are:

$$(8.3) \qquad \Theta_{\mathcal{C}}^{(2)}(F,H,F,H) = \|FH\|_{\mathrm{L}^2}^2 - \Theta_{\mathcal{C}}^{(1)}(F,H,F,H)\,,$$

$$(8.4) \qquad \Theta_{\mathcal{C}}^{(2)}(\mathbf{1},G,\mathbf{1},G) = \|G\|_{\mathrm{L}^2}^2 - \Theta_{\mathcal{C}}^{(1)}(\mathbf{1},G,\mathbf{1},G) = \|G\|_{\mathrm{L}^2}^2\,,$$

$$(8.5) \qquad \Theta_{\mathcal{C}}^{(1)}(F,F,F,F) = \|F\|_{\mathrm{L}^4}^4 - \Theta_{\mathcal{C}}^{(2)}(F,F,F,F) \leq \|F\|_{\mathrm{L}^4}^4\,,$$

$$(8.6) \qquad \Theta_{\mathcal{C}}^{(1)}(H,H,H,H) = \|H\|_{\mathrm{L}^4}^4 - \Theta_{\mathcal{C}}^{(2)}(H,H,H,H) \leq \|H\|_{\mathrm{L}^4}^4\,.$$

Combining (8.1)–(8.6) one ends up with

$$|\Lambda_{\mathrm{d}}(F,G,H)| \leq \|G\|_{\mathrm{L}^2}\Big(\|FH\|_{\mathrm{L}^2}^2 + \|F\|_{\mathrm{L}^4}^2\|H\|_{\mathrm{L}^4}^2\Big)^{1/2}\,,$$

which establishes the estimate for $(p,q,r) = (4,2,4)$. By symmetry one also gets the point $(p,q,r) = (2,4,4)$, and then uses interpolation and the method from Section 5. However, that way we would leave out the larger part of the Banach triangle, including the "central" point $(p,q,r) = (3,3,3)$.

Starting from the single tree estimate (3.4) and adjusting the arguments from Section 4 in an obvious way, we also obtain estimates for an even more "entwined" form:

$$\Theta_{\mathcal{C}}^{(2)}(F_1,F_2,F_3,F_4) = \sum_{I \times J \in \mathcal{C}} \int_{\mathbb{R}^4} F_1(u,v)F_2(x,v)F_3(u,y)F_4(x,y)$$
$$\cdot \varphi_I^{\mathrm{d}}(u)\,\varphi_I^{\mathrm{d}}(x)\,\psi_J^{\mathrm{d}}(v)\,\psi_J^{\mathrm{d}}(y)\ du\,dv\,dx\,dy\,.$$

The bound we get is

$$\bigl|\Theta_{\mathcal{C}}^{(2)}(F_1, F_2, F_3, F_4)\bigr| \ \lesssim_{p_1,p_2,p_3,p_4} \ \prod_{j=1}^{4} \|F_j\|_{\mathrm{L}^{p_j}},$$

whenever $\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}+\frac{1}{p_4} = 1$, $2 < p_1, p_2, p_3, p_4 < \infty$. This time we do not know of any arguments from the Calderón–Zygmund theory that could help expand the range of exponents.

Let us conclude with several words on a straightforward generalization of the method presented in Sections 3 and 4 to higher dimensions. For notational simplicity we only state the result in $\mathbb{R}^3$.

Theorem 6. *For any $S \subseteq \{0, 1, \ldots, 7\}$ we define a multilinear form Λ_S, acting on $|S|$ functions $F_j \colon \mathbb{R}^3 \to \mathbb{C}$ by*

$$\Lambda_S\bigl((F_j)_{j \in S}\bigr) := \sum_Q \int_{\mathbb{R}^6} \prod_{j \in S} F_j\bigl(x_1^{j_1}, x_2^{j_2}, x_3^{j_3}\bigr) \ \varphi_{I_1}^{\mathrm{d}}(x_1^0) \, \varphi_{I_1}^{\mathrm{d}}(x_1^1)$$
$$\cdot \ \varphi_{I_2}^{\mathrm{d}}(x_2^0) \, \varphi_{I_2}^{\mathrm{d}}(x_2^1) \, \psi_{I_3}^{\mathrm{d}}(x_3^0) \, \psi_{I_3}^{\mathrm{d}}(x_3^1) \ dx_1^0 \, dx_1^1 \, dx_2^0 \, dx_2^1 \, dx_3^0 \, dx_3^1 \, ,$$

where $Q = I_1 \times I_2 \times I_3$ is a dyadic cube, and $j = j_1 + 2j_2 + 4j_3$, $j_1, j_2, j_3 \in \{0, 1\}$. Then Λ_S satisfies the bound

$$\bigl|\Lambda_S\bigl((F_j)_{j \in S}\bigr)\bigr| \ \lesssim_{(p_j)_{j \in S}} \ \prod_{j \in S} \|F_j\|_{\mathrm{L}^{p_j}(\mathbb{R}^3)} \, ,$$

whenever the exponents $(p_j)_{j \in S}$ are such that $\sum_{j \in S} \frac{1}{p_j} = 1$, and $4 < p_j < \infty$ for every $j \in S$.

The result is nontrivial only when $|S| \geq 5$. We sketch a proof of Theorem 6, which uses the same ingredients as before.

Dyadic cubes $Q = I_1 \times I_2 \times I_3 \subseteq \mathbb{R}^3$ are again organized into families of trees. For each tree $\mathcal{T}$ we define the three local forms $\Theta_{\mathcal{T}}^{(1)}$, $\Theta_{\mathcal{T}}^{(2)}$, $\Theta_{\mathcal{T}}^{(3)}$. For instance,

$$\Theta_{\mathcal{T}}^{(1)}(F_0, \ldots, F_7) := \sum_{Q \in \mathcal{T}} \int_{\mathbb{R}^6} \prod_{j=0}^{7} F_j\bigl(x_1^{j_1}, x_2^{j_2}, x_3^{j_3}\bigr) \sum_{\alpha, \beta \in \{\text{left}, \text{right}\}} \psi_{I_1}^{\mathrm{d}}(x_1^0) \psi_{I_1}^{\mathrm{d}}(x_1^1)$$
$$\cdot \ \varphi_{I_{2,\alpha}}^{\mathrm{d}}(x_2^0) \, \varphi_{I_{2,\alpha}}^{\mathrm{d}}(x_2^1) \, \varphi_{I_{3,\beta}}^{\mathrm{d}}(x_3^0) \, \varphi_{I_{3,\beta}}^{\mathrm{d}}(x_3^1) \ dx_1^0 \, dx_1^1 \, dx_2^0 \, dx_2^1 \, dx_3^0 \, dx_3^1 \, .$$

The form $\Xi_{\mathcal{F}}$ is defined analogously, with $[\cdot]_{\square(Q)}$ replaced by the three-dimensional Gowers box inner product:

$$[F_0, \ldots, F_7]_{\square^3(Q)} := \mathbb{E}\Bigl(\prod_{j=0}^{7} F_j\bigl(x_1^{j_1}, x_2^{j_2}, x_3^{j_3}\bigr) \Big| \, x_1^0, x_1^1 \in I_1, \ x_2^0, x_2^1 \in I_2, \ x_3^0, x_3^1 \in I_3 \Bigr),$$

in the probabilistic notation. However, the inequality (2.2) has to be replaced with

$$\|F\|_{\square^3(Q)} \ \leq \ \Bigl(\frac{1}{|Q|} \int_Q |F|^4 \Bigr)^{1/4},$$

which is the reason why the range of exponents is severely restricted.

$$\Theta_{\mathcal{T}}^{(3)}(F_0, F_1, F_2, F_3, F_4, F_5, F_6, F_7)$$

$$\Theta_{\mathcal{T}}^{(3)}(F_0, F_1, F_2, F_3, F_0, F_1, F_2, F_3) \qquad \Theta_{\mathcal{T}}^{(3)}(F_4, F_5, F_6, F_7, F_4, F_5, F_6, F_7)$$

$$\Theta_{\mathcal{T}}^{(1)}(F_0, F_1, F_2, F_3, F_0, F_1, F_2, F_3) \qquad \Theta_{\mathcal{T}}^{(2)}(F_0, F_1, F_2, F_3, F_0, F_1, F_2, F_3)$$

$$\Theta_{\mathcal{T}}^{(1)}(F_0, F_0, F_2, F_2, F_0, F_0, F_2, F_2) \qquad \Theta_{\mathcal{T}}^{(1)}(F_1, F_1, F_3, F_3, F_1, F_1, F_3, F_3)$$

$$\Theta_{\mathcal{T}}^{(2)}(F_0, F_0, F_2, F_2, F_0, F_0, F_2, F_2) \qquad \Theta_{\mathcal{T}}^{(3)}(F_0, F_0, F_2, F_2, F_0, F_0, F_2, F_2) \geq 0$$

$$\Theta_{\mathcal{T}}^{(2)}(F_0, F_0, F_0, F_0, F_0, F_0, F_0, F_0) \qquad \Theta_{\mathcal{T}}^{(2)}(F_2, F_2, F_2, F_2, F_2, F_2, F_2, F_2)$$

$$\Theta_{\mathcal{T}}^{(1)}(F_0, F_0, F_0, F_0, F_0, F_0, F_0, F_0) \geq 0 \qquad \Theta_{\mathcal{T}}^{(3)}(F_0, F_0, F_0, F_0, F_0, F_0, F_0, F_0) \geq 0$$
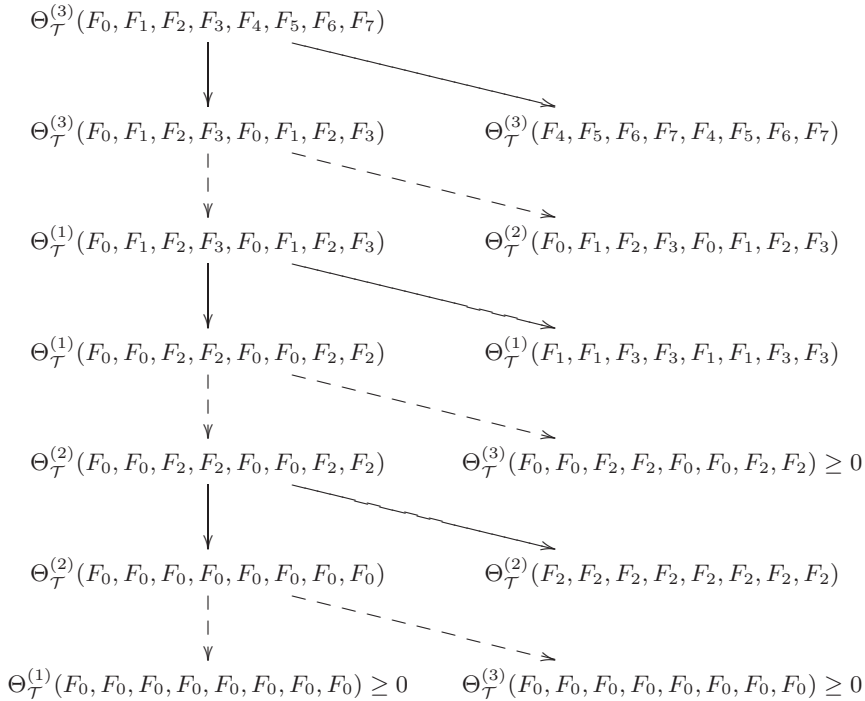
FIGURE 3. The proof of the single tree estimate in $\mathbb{R}^3$. A solid arrow denotes an application of the Cauchy–Schwarz inequality, while a broken arrow denotes an application of identity (8.7).

The telescoping identity now has three terms on the left hand side:

$$(8.7) \qquad\qquad \Theta_{\mathcal{T}}^{(1)} + \Theta_{\mathcal{T}}^{(2)} + \Theta_{\mathcal{T}}^{(3)} = \Xi_{\mathcal{L}(\mathcal{T})} - \Xi_{\{Q_{\mathcal{T}}\}} \,.$$

The proof of the single tree estimate is inductive, with alternating applications of the identity (8.7) and the Cauchy–Schwarz inequality. The telescoping identity reduces the problem of controlling a particular "theta-term"

$$\Theta^{(i)}(F_{k_0}, \ldots, F_{k_7})$$

to bounding two other theta-terms. If either of the latter is nonnegative, it can be ignored. On the other hand, any term that is not nonnegative can be estimated, using an analogue of Lemma 3, by two nonnegative terms with a smaller number of different functions involved. The induction starts with theta-terms containing only one function, $\Theta^{(i)}(F_k, \ldots, F_k)$, but these are obviously nonnegative.

Figure 3 presents these steps in the form of a tree-diagram. We draw essentially different branches only, i.e., omit the ones that can be treated by analogy.

References

[1] Bernicot, F.: Fiber-wise Calderón–Zygmund decomposition and application to a bi-dimensional paraproduct. To appear in *Illinois J. Math.*

[2] Demeter, C., Pramanik, M., and Thiele, C. M.: Multilinear singular operators with fractional rank. *Pacific J. Math.* **246** (2010), no. 2, 293–324.

[3] Demeter, C. and Thiele, C. M.: On the two-dimensional bilinear Hilbert transform. *Amer. J. Math.* **132** (2010), no. 1, 201–256.

[4] Gowers, T.: A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* **8** (1998), no. 3, 529–551.

[5] Gowers, T.: A new proof of Szemerédi's theorem. *Geom. Funct. Anal.* **11** (2001), no. 3, 465–588.

[6] Grafakos, L. and Tao, T.: Multilinear interpolation between adjoint operators. *J. Funct. Anal.* **199** (2003), no. 2, 379–385.

[7] Jones, R. L., Seeger, A., and Wright, J.: Strong variational and jump inequalities in harmonic analysis. *Trans. Amer. Math. Soc.* **360** (2008), no. 12, 6711–6742.

[8] Lacey, M. T. and Thiele, C. M.: L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. *Ann. of Math. (2)* **146** (1997), no. 3, 693–724.

[9] Lacey, M. T. and Thiele, C. M.: On Calderón's conjecture. *Ann. of Math. (2)* **149** (1999), no. 2, 475–496.

[10] Muscalu, C., Pipher, J., Tao, T., and Thiele, C. M.: Bi-parameter paraproducts. *Acta Math.* **193** (2004), no. 2, 269–296.

[11] Muscalu, C., Pipher, J., Tao, T., and Thiele, C. M.: Multi-parameter paraproducts. *Rev. Mat. Iberoam.* **22** (2006), no. 3, 963–976.

[12] Muscalu, C., Tao, T., and Thiele, C. M.: Multi-linear operators given by singular multipliers. *J. Amer. Math. Soc.* **15** (2002), no. 2, 469–496.

[13] Shkredov, I. D.: On a problem of Gowers. *Izv. Ross. Akad. Nauk Ser. Mat.* **70** (2006), no. 2, 179–221. Translation in *Izv. Math.* **70** (2006), no. 2, 385–425.

[14] Tao, T.: The ergodic and combinatorial approaches to Szemerédi's theorem. In *Additive combinatorics,* 145–193. CRM Proc. Lecture Notes 43, Amer. Math. Soc., Providence, RI, 2007.

[15] Thiele, C. M.: Time-frequency analysis in the discrete phase plane. In *Topics in analysis and its applications,* 99–152. Ph.D. thesis, Yale University, 1995. World Sci. Publ., River Edge, NJ, 2000.

[16] Thiele, C. M.: *Wave packet analysis.* CBMS Regional Conference Series in Mathematics 105, American Mathematical Society, Providence, RI, 2006.

Vjekoslav Kovač: Department of Mathematics, University of Zagreb, Bijenička cesta 30, 10000 Zagreb, Croatia.
E-mail: vjekovac@math.hr