

## A $T(1)$ THEOREM FOR ENTANGLED MULTILINEAR DYADIC CALDERÓN–ZYGMUND OPERATORS

VJEKOSLAV KOVAČ AND CHRISTOPH THIELE

ABSTRACT. We prove a boundedness criterion for a class of dyadic multilinear forms acting on two-dimensional functions. Their structure is more general than the one of classical multilinear Calderón–Zygmund operators as several functions can now depend on the same one-dimensional variable. The study of this class is motivated by examples related to the two-dimensional bilinear Hilbert transform and to bilinear ergodic averages. This paper is a sequel to a prior paper by the first author.

### 1. Introduction

The recent papers [5], [3], [8], [7] initiated the study of multilinear singular integral operators and forms with certain general modulation symmetries which force the Schwartz kernels of the operators and forms to be supported on lower dimensional subspaces. A multilinear variant of the Littlewood–Paley theory combined with the method of Bellman functions was introduced in [7] and [8] and applied to proving  $L^p$  estimates for such multilinear forms acting on functions of two variables. The mentioned modulation symmetries appear when we can write the multilinear form schematically as

$$(1.1) \quad \Lambda(F_1, F_2, \dots) = \int_{\mathbb{R}^n} F_1(x_1, x_2) F_2(x_1, x_3) \cdots \\ \times K(x_1, \dots, x_n) dx_1 dx_2 dx_3 \cdots dx_n,$$

---

Received May 21, 2013; received in final form October 30, 2013.

The first author was partially supported by the MZOS Grant 037-0372790-2799 of the Republic of Croatia.

The second author was partially supported by NSF Grant DMS 1001535.

2010 *Mathematics Subject Classification.* 42B20.

with the exposed feature that  $F_1$  and  $F_2$  share the variable  $x_1$ . If, for some  $g \in L^\infty(\mathbb{R})$ , we have

$$\tilde{F}_1(x, y) := g(x)F_1(x, y), \quad \tilde{F}_2(x, y) := g(x)F_2(x, y),$$

then we have the modulation symmetry

$$\Lambda(\tilde{F}_1, F_2, \dots) = \Lambda(F_1, \tilde{F}_2, \dots).$$

Note that if two functions share both of their variables, which is tantamount to a symmetry under arbitrary modulations by two dimensional functions, then the two functions sharing both variables may be replaced by their pointwise product in the multilinear form. The multilinear form then trivially reduces to one of a lower degree. It is precisely the partial one-dimensional modulation symmetries that cause multilinear forms (1.1) to require techniques outside the scope of the classical Calderón–Zygmund theory.

Our main result, Theorem 1, establishes a T(1)-type criterion for boundedness of a class of multilinear forms of the above type. In this theorem, we specialize to a bipartite structure elaborated in the next section. This bipartite structure manifests itself in that the variables fall into two classes, namely the first entry variables  $x_i$  and the second entry variables  $y_j$ . Sharing of variables can only occur for variables of the same kind, so we rewrite (1.1) as

$$(1.2) \quad \Lambda(F_1, F_2, \dots) = \int_{\mathbb{R}^{m+n}} F_1(x_1, y_1) F_2(x_1, y_2) \cdots \\ \times K(x_1, \dots, x_m, y_1, \dots, y_n) dx_1 \cdots dx_m dy_1 \cdots dy_n.$$

If  $m$  equals  $n$ , then the kernel  $K$  should be thought of as the usual Coifman–Meyer or multilinear Calderón–Zygmund kernel in  $(\mathbb{R}^2)^n$ , which is singular along the diagonal  $(x_1, y_1) = \cdots = (x_n, y_n)$ . The requirement  $m = n$  is not necessary in our setting and the kernel is then singular along the set

$$(1.3) \quad x_1 = \cdots = x_m, \quad y_1 = \cdots = y_n.$$

We only discuss the dyadic model of the so-called perfect Calderón–Zygmund kernels, which are described in the next section. At present, we only have a partial understanding of the required modifications for extending our results to the case of continuous Calderón–Zygmund kernels.

The case  $m = 1$  or  $n = 1$  trivializes (see Section 4.2), hence the smallest non-trivial and motivational example of a form of the type (1.2) writes as follows:

$$(1.4) \quad \Lambda_{\square}(F_1, F_2, F_3) := \int_{\mathbb{R}^4} F_1(x_1, y_1) F_2(x_1, y_2) F_3(x_2, y_1) \\ \times K(x_1, x_2, y_1, y_2) dx_1 dx_2 dy_1 dy_2.$$

The translation invariant case

$$K(x_1, x_2, y_1, y_2) = \kappa(x_1 - x_2, y_1 - y_2)$$

has been studied previously. It is a special case of the two-dimensional bilinear Hilbert transform introduced in [5]. Estimates in various ranges of  $L^p$  spaces were proven by Bernicot [3] and by the first author [8]. The present paper particularly generalizes boundedness results for (1.4) beyond translation invariance, albeit only in the dyadic model of perfect Calderón–Zygmund kernels.

We view this paper as partial progress towards understanding some very challenging questions about operators with rather general structure (1.1). Here we address the bipartite case (1.2) only and it is much simpler than the general non-bipartite case. A profound example of the latter is the *triangular Hilbert transform*

$$(1.5) \quad \Lambda_{\Delta}(F_1, F_2, F_3) := \text{p.v.} \int_{\mathbb{R}^3} F_1(x, y)F_2(y, z)F_3(z, x) \frac{1}{x + y + z} dx dy dz.$$

Compared with (1.4) we have the additional shared variable  $x_2 = y_2$  in the notation of (1.4), destroying the bipartite structure. It appears possible that  $\Lambda_{\Delta}$  satisfies some  $L^p$  bounds, for example a bound  $L^3 \times L^3 \times L^3 \rightarrow \mathbb{C}$ . Proving such bounds remains out of reach at present. The form  $\Lambda_{\Delta}$  is closely related to and is stronger than both the one-dimensional bilinear Hilbert transform and the Carleson operator, which can be realized by choosing the functions  $F_i$  in (1.5) appropriately. The form  $\Lambda_{\Delta}$  is also stronger than instances of (1.4), as the method of rotation deduces bounds on these instances from the conjectural bounds on (1.5).

Part of the motivation for (1.5) comes from the desire to understand pointwise almost everywhere convergence and other questions for ergodic averages of the form

$$\frac{1}{n} \sum_{k=0}^{n-1} f(S^k \omega)g(T^k \omega), \quad \omega \in \Omega,$$

for two commuting measure preserving transformations  $S, T : \Omega \rightarrow \Omega$  on a probability space  $\Omega$ . A real-analytic approach suggested by Demeter and the second author in [5] is to first understand estimates for the bilinear operator

$$T_{\Delta}(F, G)(x, y) := \text{p.v.} \int_{\mathbb{R}} F(x + t, y)G(x, y + t) \frac{dt}{t}.$$

If we dualize with another function  $H$ , we obtain

$$\int_{\mathbb{R}^2} \left( \text{p.v.} \int_{\mathbb{R}} F(x + t, y)G(x, y + t) \frac{dt}{t} \right) H(x, y) dx dy.$$

After the change of variables  $z = -x - y - t$  and substitutions

$$F_1(x, y) = H(x, y), \quad F_2(y, z) = F(-y - z, y), \quad F_3(z, x) = G(x, -x - z),$$

we are left with the negative of the trilinear form (1.5).

The connection between  $\Lambda_\Delta$  and the paradigmatic objects in time frequency analysis, namely the bilinear Hilbert transform and Carleson's operator, suggest to look at time frequency analysis as an approach to  $\Lambda_\Delta$ . Time frequency analysis generally reduces estimation of more difficult operators to estimation of simpler operators called tree models. In the case of  $\Lambda_\Delta$  we expect these simpler operators to be forms of the type (1.4), albeit with kernels somewhat more general than those singular along the diagonal (1.3).

This paper, as a sequel to [7], will use the same graph-theoretic setup as there and occasionally refer to results of [7]. Theorems 1 and 2 below generalize [7, Theorem 1.1], apart from changes in the exponent range. We restrict attention to the two-dimensional case, that is to functions  $F_i$  of two real variables, since we do not have phenomena to report in higher dimensions other than straightforward generalizations. We refer to [8, Section 8] for a brief outline of some aspects of the higher dimensional theory.

## 2. Formulation of the results

We write  $A \lesssim B$  for two non-negative quantities  $A$  and  $B$  if  $A \leq CB$  holds with some constant  $0 \leq C < \infty$ . We will always specify if  $C$  is an absolute constant or if it depends on some parameters. In order to aid reading the text, functions of one real variable are denoted by lowercase letters, while functions of several real variables are denoted by uppercase letters. The characteristic function of a set  $S \subseteq \mathbb{R}^n$  will be denoted by  $\mathbf{1}_S$ .

A *dyadic interval* will always mean an interval of the form  $[2^{-k}\ell, 2^{-k}(\ell+1))$  for  $k, \ell \in \mathbb{Z}$ . An  *$n$ -dimensional dyadic cube* is any set  $I_1 \times I_2 \times \cdots \times I_n$ , where  $I_1, I_2, \dots, I_n$  are dyadic intervals of the same length. We will write  $\mathcal{C}_n$  for the collection of all  $n$ -dimensional dyadic cubes. Particularly important are *dyadic squares* in  $\mathbb{R}^2$ , obtained by specializing  $n = 2$ , and their collection will be denoted simply by  $\mathcal{C}$ .

Fix two positive integers  $m$  and  $n$ . The *diagonal* in  $\mathbb{R}^{m+n}$  is

$$D := \left\{ \underbrace{(x_1, \dots, x_m)}_m, \underbrace{(y_1, \dots, y_n)}_n : x, y \in \mathbb{R} \right\}.$$

The definition of *perfect dyadic Calderón–Zygmund kernel* will be adapted from [1, Section 6]. It is a locally integrable function  $K : \mathbb{R}^{m+n} \rightarrow \mathbb{C}$  satisfying the bound

$$(2.1) \quad \left| K(x_1, \dots, x_m, y_1, \dots, y_n) \right| \lesssim \left( \sum_{1 \leq i_1 < i_2 \leq m} |x_{i_1} - x_{i_2}| + \sum_{1 \leq j_1 < j_2 \leq n} |y_{j_1} - y_{j_2}| \right)^{2-m-n}$$

for each  $(x_1, \dots, x_m, y_1, \dots, y_n) \in \mathbb{R}^{m+n} \setminus D$  and such that  $K$  is constant on each  $(m+n)$ -dimensional dyadic cube  $I_1 \times \cdots \times I_m \times J_1 \times \cdots \times J_n$  that does not intersect the diagonal. Finally, we impose the qualitative technical condition

that  $K$  is bounded and compactly supported, but without any quantitative information on either its  $L^\infty$ -norm, or the size of its support.

Even though (2.1) is a variant of the usual size estimate from multilinear Calderón–Zygmund theory, see the paper by Grafakos and Torres [6], we emphasize that the operators we study in this paper can have more complicated structure due to the modulation symmetries.

The rest of the setup is similar to the one in [7]. Choose

$$E \subseteq \{1, \dots, m\} \times \{1, \dots, n\}.$$

We can view  $E$  as the set of edges of some simple bipartite undirected graph  $G$  on the sets of vertices  $\{x_1, \dots, x_m\}$  and  $\{y_1, \dots, y_n\}$ . More precisely,  $x_i$  and  $y_j$  are connected by an edge in this graph if and only if  $(i, j) \in E$ . Finally, we define a multilinear form  $\Lambda_E$  by

$$(2.2) \quad \Lambda_E((F_{i,j})_{(i,j) \in E}) := \int_{\mathbb{R}^{m+n}} K(x_1, \dots, x_m, y_1, \dots, y_n) \times \prod_{(i,j) \in E} F_{i,j}(x_i, y_j) dx_1 \cdots dx_m dy_1 \cdots dy_n.$$

Thus, the graph structure determines which functions occur in the above expression. To avoid degeneracy, we assume that there are no isolated vertices in  $G$ , that is, each vertex is incident to some edge. This forces each integration variable to appear in at least one of the functions. In all of the following, we assume that the functions  $F_{i,j}$  are bounded and compactly supported, so that (2.2) is a priori well-defined.

The  $|E|$ -linear form  $\Lambda_E$  uniquely determines  $|E|$  mutually adjoint  $(|E| - 1)$ -linear operators  $T_{u,v}$ ,  $(u, v) \in E$ . These are explicitly defined by

$$(2.3) \quad T_{u,v}((F_{i,j})_{(i,j) \in E \setminus \{(u,v)\}})(x_u, y_v) := \int_{\mathbb{R}^{m+n-2}} K(x_1, \dots, x_m, y_1, \dots, y_n) \times \prod_{(i,j) \in E \setminus \{(u,v)\}} F_{i,j}(x_i, y_j) \prod_{i \neq u} dx_i \prod_{j \neq v} dy_j,$$

so that

$$(2.4) \quad \Lambda_E((F_{i,j})_{(i,j) \in E}) = \int_{\mathbb{R}^2} T_{u,v}((F_{i,j})_{(i,j) \neq (u,v)}) F_{u,v}.$$

Let us also recall the definition of the *dyadic BMO-seminorm*,

$$(2.5) \quad \|F\|_{\text{BMO}(\mathbb{R}^2)} := \sup_{Q \in \mathcal{C}} \left( \frac{1}{|Q|} \int_Q \left| F - \frac{1}{|Q|} \int_Q F \right|^2 \right)^{1/2}.$$

We are now ready to state the main result of this paper.

THEOREM 1 (Entangled T(1) theorem).

(a) Suppose that  $m, n \geq 2$ . There exist positive integers  $(d_{i,j})_{(i,j) \in E}$ , depending on  $m, n, E$ , satisfying

$$(2.6) \quad \sum_{(i,j) \in E} \frac{1}{d_{i,j}} > 1,$$

such that the following holds. If one has

$$(2.7) \quad |\Lambda_E(\mathbf{1}_Q, \dots, \mathbf{1}_Q)| \lesssim |Q| \quad \text{for each } Q \in \mathcal{C}$$

and

$$(2.8) \quad \|T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2})\|_{\text{BMO}(\mathbb{R}^2)} \lesssim 1 \quad \text{for each } (u, v) \in E,$$

then the form  $\Lambda_E$  satisfies the estimate

$$(2.9) \quad |\Lambda_E((F_{i,j})_{(i,j) \in E})| \lesssim \prod_{(i,j) \in E} \|F_{i,j}\|_{L^{p_{i,j}}(\mathbb{R}^2)}$$

whenever exponents  $p_{i,j}$  are taken from the region determined by

$$\sum_{(i,j) \in E} \frac{1}{p_{i,j}} = 1 \quad \text{and} \quad d_{i,j} < p_{i,j} \leq \infty \quad \text{for each } (i, j) \in E.$$

In particular, it is guaranteed by (2.6) that estimate (2.9) holds in a non-empty open range of exponents. The implicit constant in (2.9) depends on implicit constants from (2.1), (2.7), (2.8), on the graph structure (which was given by  $m, n, E$ ), and on the exponents  $p_{i,j}$ .

(b) Conversely, if inequality (2.9) holds for some choice of exponents  $1 < p_{i,j} \leq \infty$  such that  $\sum_{(i,j) \in E} p_{i,j}^{-1} = 1$ , then conditions (2.7) and (2.8) must also be satisfied, with constants depending on the constants from (2.1), (2.9), on the graph, and on the exponents.

For a first insight into concrete choices of  $d_{i,j}$  the reader is referred to (3.1), which applies whenever the  $d_{i,j}$  such defined satisfy (2.6). This happens in most instances, the few exceptional cases are discussed in Section 4. Generally, the exponent region increases as  $|E|$  gets larger with fixed  $m, n$ . The actual range of estimates is not known even for very simple instances of entangled forms, such as  $\Lambda_{\square}$ , and it is clear that known techniques are insufficient to address the question of optimal range. Disheartened by this fact, we did not attempt to make the region of exponents as large as possible within our set of available techniques.

Condition (2.7) is a version of a *weak boundedness property*. Usually weak boundedness properties involve testing with bump functions (see [4]), but here it is natural to test characteristic functions of dyadic squares. The T(1)-type conditions in (2.8) are similar to requirements of the classical T(1) theorem from [4], but we need to perform  $|E|$  verifications, instead of only two.

One might compare this situation with the T(1) theorem in [2], dealing with modulation-invariant trilinear forms and having three conditions of that kind.

The emphasis of Theorem 1 is on the qualitative  $L^p$  bound for  $\Lambda_E$  and on constant dependencies, since mere continuity of  $\Lambda_E$  is automatic by our assumptions on  $K$ . In applications one would apply the theorem to truncated and localized kernels  $K$ .

We can restate Theorem 1 in terms of *local T(1) conditions*, compare with [1, Corollary 6.3]. Their continuous analogues are the well-known *restricted boundedness conditions* from [9]. In the following formulation, we will also emphasize the generalized modulation invariance we mentioned in the Introduction. Suppose that we are given two one-dimensional functions  $a_Q^{i,j}, b_Q^{i,j} \in L^\infty(\mathbb{R})$  for each dyadic square  $Q = I \times J \in \mathcal{C}$  and for each  $(i, j) \in E$ . We require that

$$(2.10) \quad \prod_{j:(i,j) \in E} a_Q^{i,j} = \mathbf{1}_I \quad \text{for every } i, \quad \prod_{i:(i,j) \in E} b_Q^{i,j} = \mathbf{1}_J \quad \text{for every } j,$$

each  $a_Q^{i,j}$  vanishes outside  $I$ , and each  $b_Q^{i,j}$  vanishes outside  $J$ . To simplify the notation, we also write

$$B_Q^{i,j} := a_Q^{i,j} \otimes b_Q^{i,j},$$

where  $f \otimes g$  denotes the elementary tensor defined by  $(f \otimes g)(x, y) := f(x)g(y)$ .

**THEOREM 2 (Entangled T(1) theorem, reformulation).** *For  $m, n \geq 2$  there exist positive integers  $(d_{i,j})_{(i,j) \in E}$  depending on  $m, n, E$ , satisfying (2.6), and such that the following holds. If one can verify*

$$(2.11) \quad \|T_{u,v}((B_Q^{i,j})_{(i,j) \in E \setminus \{(u,v)\}})B_Q^{u,v}\|_{L^1(\mathbb{R}^2)} \lesssim |Q|$$

for each  $Q = I \times J$  and for each  $(u, v) \in E$ , then estimate (2.9) holds in the same range of exponents as in Theorem 1. The implicit constant depends on implicit constants from (2.1), (2.11), on  $m, n, E$ , and on the exponents  $p_{i,j}$ .

Note that (2.10) can be equivalently written as

$$\prod_{(i,j) \in E} a_Q^{i,j}(x_i)b_Q^{i,j}(y_j) = \prod_{i=1}^m \mathbf{1}_I(x_i) \prod_{j=1}^n \mathbf{1}_J(y_j).$$

By the one-dimensional modulation invariance of (2.3), we instantly observe

$$\begin{aligned} T_{u,v}(a_Q^{i,j} \otimes b_Q^{i,j})_{(i,j) \in E \setminus \{(u,v)\}}(a_Q^{u,v} \otimes b_Q^{u,v}) \\ = T_{u,v}(\mathbf{1}_I \otimes \mathbf{1}_J)_{(i,j) \in E \setminus \{(u,v)\}}(\mathbf{1}_I \otimes \mathbf{1}_J), \end{aligned}$$

so that condition (2.11) equivalently reads

$$(2.12) \quad \|T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)\|_{L^1(Q)} \lesssim |Q|.$$

In the following sections, we can assume that this reduction is already performed. An interesting consequence of Theorem 2, as it is formulated, is that



FIGURE 1. Bipartite graphs associated with  $\Lambda_{\sqcup}$  and  $\Lambda_{\square}$ , respectively.

in verification of the bound (2.9), it is enough to assume that all functions  $F_{i,j}$  but one are indeed elementary tensors, in which case the entangled structure disappears. On the other hand, Theorem 2 is not really a T(b)-type theorem, as the functions  $B_Q^{i,j}$  have to be chosen under quite special constraints.

Let us turn to particular instances of our general result, assuming that conditions (2.7) and (2.8) are satisfied. If we take  $m = n = 2$  and  $E = \{(1,1), (1,2), (2,1)\}$ , we obtain (1.4), written as

$$\begin{aligned} & \Lambda_{\sqcup}(F_{1,1}, F_{1,2}, F_{2,1}) \\ &= \int_{\mathbb{R}^4} F_{1,1}(x_1, y_1) F_{1,2}(x_1, y_2) F_{2,1}(x_2, y_1) K(x_1, x_2, y_1, y_2) dx_1 dx_2 dy_1 dy_2. \end{aligned}$$

The corresponding graph is depicted in the left half of Figure 1. By taking a look at the numbers  $d_{i,j}$  defined by (3.1), we see that estimate (2.9) holds in the range  $p_{1,1}^{-1} + p_{1,2}^{-1} + p_{2,1}^{-1} = 1$ ,  $p_{1,1}, p_{1,2}, p_{2,1} > 2$ . On the other hand, there is no good dyadic model for (1.5) that would be covered by Theorem 1. The closest in spirit is the “four-cyclic” quadrilinear form

$$\begin{aligned} \Lambda_{\square}(F_{1,1}, F_{1,2}, F_{2,1}, F_{2,2}) &= \int_{\mathbb{R}^4} F_{1,1}(x_1, y_1) F_{1,2}(x_1, y_2) F_{2,1}(x_2, y_1) F_{2,2}(x_2, y_2) \\ &\quad \times K(x_1, x_2, y_1, y_2) dx_1 dx_2 dy_1 dy_2. \end{aligned}$$

It is determined by  $m = n = 2$  and  $E = \{1,2\} \times \{1,2\}$ ; see the right half of Figure 1. This time we establish the bound (2.9) for  $p_{1,1}^{-1} + p_{1,2}^{-1} + p_{2,1}^{-1} + p_{2,2}^{-1} = 1$ ,  $p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2} > 2$ .

We also comment that if  $m = n$  and each component of  $G$  contains only one edge, then (after relabeling) we can write

$$\begin{aligned} & \Lambda_E(F_{1,1}, \dots, F_{n,n}) \\ &= \int_{(\mathbb{R}^2)^n} K(x_1, \dots, x_n, y_1, \dots, y_n) \prod_{j=1}^n F_{j,j}(x_j, y_j) dx_1 dy_1 \cdots dx_n dy_n. \end{aligned}$$

In this case, Theorems 1 and 2 recover dyadic variants of classical results from multilinear Calderón–Zygmund theory, as it was developed by Grafakos and Torres [6]. As is seen from (3.1), we reprove estimate (2.9) in this case whenever  $\sum_{j=1}^n p_{j,j}^{-1} = 1$  and  $p_{j,j} > 1$  for each  $j$ .



As for the organization of the paper, Sections 3 and 4 establish part (a) of Theorem 1 in all possible cases, while Section 5 proves Theorem 2 and then reduces part (b) of Theorem 1 to it. In Section 4, we also explain why we require  $m, n \geq 2$  by giving an appropriate counterexample in the case  $\min\{m, n\} = 1$ .

### 3. Sufficiency of the testing conditions

The most substantial argument is the proof of part (a) of Theorem 1 and we present it in this section. By multilinearity of  $\Lambda_E$ , we can assume that functions  $F_{i,j}$  are non-negative.

We begin by introducing quantities that will determine the range of exponents in which  $L^p$  estimates for  $\Lambda_E$  will be proven. They are obtained by splitting  $G$  into connected components  $G_1, G_2, \dots, G_k$ . Vertices of each component  $G_\ell$  of  $G$  are conveniently indexed by the sets

$$\mathcal{X}_\ell := \{i \in \{1, \dots, m\} : x_i \text{ is a vertex in } G_\ell\}$$

and

$$\mathcal{Y}_\ell := \{j \in \{1, \dots, n\} : y_j \text{ is a vertex in } G_\ell\},$$

while its edges are determined by  $E_\ell := E \cap (\mathcal{X}_\ell \times \mathcal{Y}_\ell)$ . Take a connected component  $G_\ell$  and for each  $(i, j) \in E_\ell$  set

$$(3.1) \quad d_{i,j} := \begin{cases} \max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\} & \text{if } E_\ell = \mathcal{X}_\ell \times \mathcal{Y}_\ell \text{ or } \max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\} \leq 2, \\ \max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\} + 1 & \text{otherwise.} \end{cases}$$

Observe that  $E_\ell = \mathcal{X}_\ell \times \mathcal{Y}_\ell$  simply means that the component  $G_\ell$  is a complete bipartite graph. We keep this choice throughout this section and it will be sufficient for “most” graphs  $G$ , that is, for all but a single series of exceptional cases. In Section 4, we will discuss non-triviality of that exponent range and treat those exceptional instances differently.

**3.1. Decomposition into paraproducts.** We write  $\mathbf{h}_I$  for the  $L^\infty$ -normalized Haar wavelet on a dyadic interval  $I$ , that is,  $\mathbf{h}_I = \mathbf{1}_{I_{\text{left}}} - \mathbf{1}_{I_{\text{right}}}$ , where  $I_{\text{left}}$  and  $I_{\text{right}}$  denote the two halves of  $I$ . Consider an orthonormal basis of  $L^2(\mathbb{R}^{m+n})$  consisting of the functions

$$\frac{\mathbf{a}_{I_1}^{(1)}}{|I_1|^{1/2}} \otimes \dots \otimes \frac{\mathbf{a}_{I_m}^{(m)}}{|I_m|^{1/2}} \otimes \frac{\mathbf{b}_{J_1}^{(1)}}{|J_1|^{1/2}} \otimes \dots \otimes \frac{\mathbf{b}_{J_n}^{(n)}}{|J_n|^{1/2}},$$

where  $I_1 \times \dots \times I_m \times J_1 \times \dots \times J_n$  ranges over all  $(m+n)$ -dimensional dyadic cubes and each of the letters  $\mathbf{a}^{(i)}$  and  $\mathbf{b}^{(j)}$  stands for either the characteristic function  $\mathbf{1}$  or the Haar function  $\mathbf{h}$ , excluding the possibility that all of them are simultaneously  $\mathbf{1}$ . Since  $K$  is a square integrable function by our qualitative

assumptions, it can be decomposed in the above basis as

$$K(x_1, \dots, x_m, y_1, \dots, y_n) = \sum_{\substack{I_1 \times \dots \times I_m \times J_1 \times \dots \times J_n \in \mathcal{C}_{m+n} \\ \mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)} \in \{\mathbf{1}, \mathbf{h}\} \\ \mathbf{a}^{(i)}, \mathbf{b}^{(j)} \text{ are not all equal } \mathbf{1}}} \nu_{I_1 \times \dots \times I_m \times J_1 \times \dots \times J_n}^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}} \times \mathbf{a}_{I_1}^{(1)}(x_1) \cdots \mathbf{a}_{I_m}^{(m)}(x_m) \mathbf{b}_{J_1}^{(1)}(y_1) \cdots \mathbf{b}_{J_n}^{(n)}(y_n),$$

where

$$(3.2) \quad \nu_{I_1 \times \dots \times I_m \times J_1 \times \dots \times J_n}^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}} := \left\langle K, \frac{\mathbf{a}_{I_1}^{(1)}}{|I_1|} \otimes \cdots \otimes \frac{\mathbf{a}_{I_m}^{(m)}}{|I_m|} \otimes \frac{\mathbf{b}_{J_1}^{(1)}}{|J_1|} \otimes \cdots \otimes \frac{\mathbf{b}_{J_n}^{(n)}}{|J_n|} \right\rangle_{L^2(\mathbb{R}^{m+n})}$$

and  $\langle \cdot, \cdot \rangle_{L^2}$  denotes the standard inner product. Now we use the perfect cancellation condition. Since  $K$  is constant on cubes that do not intersect the diagonal, the corresponding coefficients (3.2) are zero, so only the terms with  $I_1 = \dots = I_m$  and  $J_1 = \dots = J_n$  remain. This leads to a representation of  $\Lambda_E$  as a finite sum of *entangled dyadic paraproducts*,

$$(3.3) \quad \Lambda_E = \sum_{\substack{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)} \in \{\mathbf{1}, \mathbf{h}\} \\ \mathbf{a}^{(i)}, \mathbf{b}^{(j)} \text{ are not all equal } \mathbf{1}}} \Theta_E^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}}.$$

These are defined explicitly by

$$(3.4) \quad \Theta((F_{i,j})_{(i,j) \in E}) = \Theta_E^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}}((F_{i,j})_{(i,j) \in E}) := \sum_{I \times J \in \mathcal{C}} \lambda_{I \times J} |I|^{2-m-n} \int_{\mathbb{R}^{m+n}} \prod_{(i,j) \in E} F_{i,j}(x_i, y_j) \times \mathbf{a}_I^{(1)}(x_1) \cdots \mathbf{a}_I^{(m)}(x_m) \mathbf{b}_J^{(1)}(y_1) \cdots \mathbf{b}_J^{(n)}(y_n) dx_1 \cdots dx_m dy_1 \cdots dy_n,$$

with coefficients

$$(3.5) \quad \lambda_{I \times J} = \lambda_{I \times J}^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}} := |I|^{m+n-2} \nu_{I \times \dots \times I \times J \times \dots \times J}^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}} = |I \times J|^{-1} \left\langle K, \mathbf{a}_I^{(1)} \otimes \cdots \otimes \mathbf{a}_I^{(m)} \otimes \mathbf{b}_J^{(1)} \otimes \cdots \otimes \mathbf{b}_J^{(n)} \right\rangle_{L^2(\mathbb{R}^{m+n})}.$$

Since we will be working with a fixed choice of  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}$  most of the time, we omit denoting that dependence in the following text and simply write  $\Theta$  and  $\lambda_{I \times J}$ . From Decomposition (3.3), we see that it is enough to prove bound (2.9) for each entangled dyadic paraproduct  $\Theta$ .

Let us introduce some convenient notation. The average value of a function  $f$  on an interval  $I$  will be denoted by  $[f(x)]_{x \in I}$ , while the same average of the function  $f \mathbf{h}_I$  will be denoted by  $\langle f(x) \rangle_{x \in I}$ . In other words,

$$[f(x)]_{x \in I} := \frac{1}{|I|} \int_{\mathbb{R}} f(x) \mathbf{1}_I(x) dx, \quad \langle f(x) \rangle_{x \in I} := \frac{1}{|I|} \int_{\mathbb{R}} f(x) \mathbf{h}_I(x) dx.$$

We will be dealing with multi-variable expressions, so it will be important to clarify in which variable the average is taken. To save space, we allow joining several averaging procedures into a single bracket notation, with all averaging variables listed in its subscript. If we set

$$S := \{i \in \{1, \dots, m\} : \mathbf{a}^{(i)} = \mathbf{h}\}, \quad T := \{j \in \{1, \dots, n\} : \mathbf{b}^{(j)} = \mathbf{h}\},$$

$$S^c := \{i \in \{1, \dots, m\} : \mathbf{a}^{(i)} = \mathbf{1}\}, \quad T^c := \{j \in \{1, \dots, n\} : \mathbf{b}^{(j)} = \mathbf{1}\},$$

then the paraproduct given in (3.4) can be written as

$$\Theta((F_{i,j})_{(i,j) \in E}) = \sum_{I \times J \in \mathcal{C}} \lambda_{I \times J} |I \times J| \mathcal{A}_{I \times J}((F_{i,j})_{(i,j) \in E}),$$

where

$$(3.6) \quad \mathcal{A}_{I \times J}((F_{i,j})_{(i,j) \in E}) := \left\langle \prod_{(i,j) \in E} F_{i,j}(x_i, y_j) \right\rangle_{\substack{x_i \in I \text{ for each } i \in S \\ y_j \in J \text{ for each } j \in T}} \Big|_{\substack{x_i \in I \text{ for each } i \in S^c \\ y_j \in J \text{ for each } j \in T^c}}.$$

We continue with some quite standard terminology. *Child* and *parent* relations on dyadic squares are defined as usually, so that each square has four children and exactly one parent. A *finite convex tree* is any finite collection  $\mathcal{T}$  of dyadic squares that is convex with respect to the set inclusion and that contains the largest square (i.e., a square that covers all others). It is denoted  $Q_{\mathcal{T}}$  and called the *tree-top* of  $\mathcal{T}$ . *Leaves* of  $\mathcal{T}$  are then defined to be dyadic squares that are not members of  $\mathcal{T}$  but their parents are. The family of leaves is denoted  $\mathcal{L}(\mathcal{T})$ . One can define a localized absolute version of each  $\Theta$  for a finite convex tree  $\mathcal{T}$  simply as

$$\Theta_{\mathcal{T}} := \sum_{Q \in \mathcal{T}} |\lambda_Q| |Q| \mathcal{A}_Q.$$

By a classical stopping time argument one can reduce the desired  $L^p$  estimate for  $\Theta$  to a “single tree estimate”,

$$(3.7) \quad \Theta_{\mathcal{T}}((F_{i,j})_{(i,j) \in E}) \lesssim |Q_{\mathcal{T}}| \prod_{(i,j) \in E} \max_{Q \in \mathcal{T} \cup \mathcal{L}(\mathcal{T})} [F_{i,j}^{d_{i,j}}]_Q^{1/d_{i,j}}$$

for each finite convex tree  $\mathcal{T}$  with tree-top  $Q_{\mathcal{T}}$ . Here  $[F]_Q$  denotes the average value of a function  $F$  on a square  $Q$ . This reduction can be found in [7, Section 4], which is tailored exactly to this situation, as all arguments given there are fairly general. We only remark that [7] does not explicitly mention any estimates involving  $L^\infty(\mathbb{R}^2)$ , but these are also easily derived. If  $p_{i,j} = \infty$  for some  $(i, j) \in E$ , then for those pairs  $(i, j)$  we simply control  $[F_{i,j}^{d_{i,j}}]_Q^{1/d_{i,j}}$  in (3.7) by  $\|F_{i,j}\|_{L^\infty(\mathbb{R}^2)}$  and apply the stopping time procedure to the remaining functions.

In order to establish (3.7) we need to distinguish two essentially different cases of entangled dyadic paraproducts. We call  $\Theta$  *non-cancellative* if there

exists an edge in  $G$  that is incident to all vertices from  $\{x_i : i \in S\} \cup \{y_j : j \in T\}$ . Otherwise we say that  $\Theta$  is *cancellative*. In order to clarify this concept, let us observe that cancellative paraproducts  $\Theta$  correspond to the cases

- (C1)  $\max\{|S|, |T|\} \geq 2$ ,
- (C2)  $S = \{i\}$ ,  $T = \{j\}$ , and  $(i, j) \notin E$ ,

while non-cancellative paraproducts appear in the cases

- (NC1)  $|S| + |T| = 1$ ,
- (NC2)  $S = \{i\}$ ,  $T = \{j\}$ , and  $(i, j) \in E$ .

Recall that  $S = T = \emptyset$  is not possible because then all  $\mathbf{a}^{(i)}$  and  $\mathbf{b}^{(j)}$  would be  $\mathbf{1}$ .

**3.2. Cancellative paraproducts.** Types (C1) and (C2) will be resolved by the following proposition.

PROPOSITION 3. *Suppose that  $\Theta$  is a cancellative entangled dyadic paraproduct defined by (3.4) with coefficients (3.5).*

(a) *If condition (2.7) holds, then the coefficients  $\lambda = (\lambda_Q)_{Q \in \mathcal{C}}$  satisfy*

$$\|\lambda\|_{\ell^\infty} := \sup_{Q \in \mathcal{C}} |\lambda_Q| \lesssim 1.$$

(b) *Each localized form  $\Theta_{\mathcal{T}}$  for a finite convex tree  $\mathcal{T}$  satisfies*

$$\Theta_{\mathcal{T}}((F_{i,j})_{(i,j) \in E}) \lesssim \|\lambda\|_{\ell^\infty} |Q_{\mathcal{T}}| \prod_{(i,j) \in E} \max_{Q \in \mathcal{T} \cup \mathcal{L}(\mathcal{T})} [F_{i,j}^{d_{i,j}}]_Q^{1/d_{i,j}}.$$

*Proof.* (a) Take an arbitrary dyadic square  $I \times J \in \mathcal{C}$ . Let us begin by substituting  $F_{i,j} = \mathbf{1}_I \otimes \mathbf{1}_J$  for each  $(i, j) \in E$  into (2.2). Since there are no isolated vertices in  $G$ , the characteristic function  $\mathbf{1}_I(x_i)$  is attached to each variable  $x_i$  (possibly several times) and the same holds for the variables  $y_j$ . Thus, Equation (2.2) becomes

$$\Lambda_E(\underbrace{\mathbf{1}_I \otimes \mathbf{1}_J, \dots, \mathbf{1}_I \otimes \mathbf{1}_J}_{|E|}) = \langle K, \underbrace{\mathbf{1}_I \otimes \dots \otimes \mathbf{1}_I}_m \otimes \underbrace{\mathbf{1}_J \otimes \dots \otimes \mathbf{1}_J}_n \rangle_{L^2(\mathbb{R}^{m+n})},$$

so inequality (2.7) gives

$$(3.8) \quad \left| \langle K, \mathbf{1}_I \otimes \dots \otimes \mathbf{1}_I \otimes \mathbf{1}_J \otimes \dots \otimes \mathbf{1}_J \rangle_{L^2(\mathbb{R}^{m+n})} \right| \lesssim |I \times J|.$$

We can split each of  $\mathbf{a}_I^{(1)}, \dots, \mathbf{a}_I^{(m)}$  as  $\mathbf{1}_{I_{\text{left}}} \pm \mathbf{1}_{I_{\text{right}}}$  and each of  $\mathbf{b}_J^{(1)}, \dots, \mathbf{b}_J^{(n)}$  as  $\mathbf{1}_{J_{\text{left}}} \pm \mathbf{1}_{J_{\text{right}}}$ . That way we can estimate  $|\lambda_{I \times J}|$  by

$$(3.9) \quad |I \times J|^{-1} \sum_{\substack{I^{(1)}, \dots, I^{(m)} \in \{I_{\text{left}}, I_{\text{right}}\} \\ J^{(1)}, \dots, J^{(n)} \in \{J_{\text{left}}, J_{\text{right}}\}}} \left| \langle K, \mathbf{1}_{I^{(1)}} \otimes \dots \otimes \mathbf{1}_{I^{(m)}} \otimes \mathbf{1}_{J^{(1)}} \otimes \dots \otimes \mathbf{1}_{J^{(n)}} \rangle_{L^2(\mathbb{R}^{m+n})} \right|.$$

The terms with  $I^{(1)} = \dots = I^{(m)}$ ,  $J^{(1)} = \dots = J^{(n)}$  can be controlled using (3.8) applied to each of the four subsquares

$$I_{\text{left}} \times J_{\text{left}}, \quad I_{\text{left}} \times J_{\text{right}}, \quad I_{\text{right}} \times J_{\text{left}}, \quad I_{\text{right}} \times J_{\text{right}}$$

to yield

$$|I \times J|^{-1} \sum_{\substack{I' \in \{I_{\text{left}}, I_{\text{right}}\} \\ J' \in \{J_{\text{left}}, J_{\text{right}}\}}} |\langle K, \mathbf{1}_{I'} \otimes \dots \otimes \mathbf{1}_{I'} \otimes \mathbf{1}_{J'} \otimes \dots \otimes \mathbf{1}_{J'} \rangle_{L^2(\mathbb{R}^{m+n})}| \lesssim 1.$$

Each of the  $2^{m+n} - 4$  remaining terms in (3.9) will be estimated individually. Without loss of generality suppose  $m \geq 2$ ,  $I^{(1)} = I_{\text{left}}$ ,  $I^{(2)} = I_{\text{right}}$  and denote the left endpoint of  $I_{\text{right}}$  by  $x_0$ . If  $x_1 \in I_{\text{left}}$ ,  $x_2 \in I_{\text{right}}$ ,  $x_3, \dots, x_m \in I$ ,  $y_1, \dots, y_n \in J$ , then

$$\begin{aligned} |x_1 - x_2| &= |x_1 - x_0| + |x_2 - x_0| \\ |x_i - x_1| + |x_i - x_2| &\geq |x_i - x_0|, \quad i = 3, \dots, m \end{aligned}$$

and consequently also

$$\begin{aligned} &\sum_{1 \leq i_1 < i_2 \leq m} |x_{i_1} - x_{i_2}| + \sum_{1 \leq j_1 < j_2 \leq n} |y_{j_1} - y_{j_2}| \\ &\geq \sum_{i=1}^m |x_i - x_0| + \sum_{j=1}^{n-1} |y_j - y_n| \geq \left( \sum_{i=1}^m (x_i - x_0)^2 + \sum_{j=1}^{n-1} (y_j - y_n)^2 \right)^{1/2}. \end{aligned}$$

Passing to the spherical coordinates in  $\mathbb{R}^{m+n-1}$  and using the kernel size estimate (2.1) we obtain

$$\begin{aligned} &|I \times J|^{-1} |\langle K, \mathbf{1}_{I_{\text{left}}} \otimes \mathbf{1}_{I_{\text{right}}} \otimes \mathbf{1}_{I^{(3)}} \otimes \dots \otimes \mathbf{1}_{I^{(m)}} \otimes \mathbf{1}_{J^{(1)}} \otimes \dots \otimes \mathbf{1}_{J^{(n)}} \rangle_{L^2(\mathbb{R}^{m+n})}| \\ &\leq |I|^{-2} \int_J \left( \int_{\text{Ball}((x_0, y_n), (m+n)|I|)} |K(x_1, \dots, x_m, \right. \\ &\quad \left. y_1, \dots, y_{n-1}, y_n)| dx_1 \cdots dx_m dy_1 \cdots dy_{n-1} \right) dy_n \\ &\leq |I|^{-2} \int_J \left( \int_0^{(m+n)|I|} r^{2-m-n} r^{m+n-2} dr \right) dy_n \lesssim 1. \end{aligned}$$

Thus,  $|\lambda_{I \times J}| \lesssim 1$ . The implicit constant certainly depends on  $m, n$  and the constants from (2.1), (2.7), but these dependencies are understood.

(b) We begin by estimating  $|\lambda_Q| \leq \|\lambda\|_{\ell^\infty}$ , so we can indeed normalize  $\|\lambda\|_{\ell^\infty} = 1$  and then abandon the coefficients by estimating them from above by 1. Since the expression  $\mathcal{A}_Q$  is scale-invariant, one can even re-scale the tree  $\mathcal{T}$  to achieve  $|Q_{\mathcal{T}}| = 1$  and then (using homogeneity) also normalize the functions by

$$(3.10) \quad \max_{Q \in \mathcal{T} \cup \mathcal{L}(\mathcal{T})} [F_{i,j}^{d_{i,j}}]_Q^{1/d_{i,j}} = 1$$

for each  $(i, j) \in E$ . Let us realize the structural splitting with respect to the graph components. We denote

$$\mathcal{A}_{I \times J}^{(\ell)} := \left[ \left\langle \prod_{(i,j) \in E_\ell} F_{i,j}(x_i, y_j) \right\rangle_{\substack{x_i \in I \text{ for each } i \in S \cap \mathcal{X}_\ell \\ y_j \in J \text{ for each } j \in T \cap \mathcal{Y}_\ell}} \right]_{\substack{x_i \in I \text{ for each } i \in S^c \cap \mathcal{X}_\ell \\ y_j \in J \text{ for each } j \in T^c \cap \mathcal{Y}_\ell}},$$

so that  $\mathcal{A}_Q = \prod_{\ell=1}^k \mathcal{A}_Q^{(\ell)}$  for each dyadic square  $Q = I \times J$ .

The paper [7] was devoted exactly to entangled dyadic paraproducts that fall into class (C1), although this terminology was not used there. In case (C1) estimate (3.7) is covered precisely by Proposition 3.1 in [7]. However, it is important to notice that there is even an improvement over our claim: it establishes (3.7) when  $d_{i,j}$  is simply replaced with  $\max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\}$  for every  $(i, j) \in E_\ell$ , that is, there is no need to add 1 to it. It is worth observing that quantities  $[F_{i,j}^d]_Q^{1/d}$  are increasing in  $d$  by Jensen’s inequality.

In what follows we show how to handle entangled paraproducts from class (C2), that is, those that have  $S = \{i_0\}$ ,  $T = \{j_0\}$ ,  $(i_0, j_0) \notin E$ . Such  $i_0$  and  $j_0$  will be fixed for the whole succeeding discussion. We further split into two subcases.

*Case 1.* The vertices  $x_{i_0}$  and  $y_{j_0}$  of  $G$  belong to the same component.

Without loss of generality we can assume  $i_0 \in \mathcal{X}_1$ ,  $j_0 \in \mathcal{Y}_1$ . After Normalization (3.10) and multiple applications of Hölder’s inequality or Lemma 3.2 from [7], we can bound  $|\mathcal{A}_Q^{(\ell)}| \leq 1$  for each  $Q \in \mathcal{T}$  and  $\ell \neq 1$ . This in turn gives  $|\mathcal{A}_Q| \leq |\mathcal{A}_Q^{(1)}|$ , so we can work with the component  $G_1$  only. Compare this reduction with [7, Section 3.2].

Exploiting the fact  $(i_0, j_0) \notin E$  we can rewrite  $\mathcal{A}_Q^{(1)}$  as

$$\begin{aligned} \mathcal{A}_Q^{(1)} &= \left[ \left\langle \prod_{j:(i_0,j) \in E_1} F_{i_0,j}(x_{i_0}, y_j) \right\rangle_{x_{i_0} \in I} \right. \\ &\quad \times \left. \left\langle \prod_{i:(i,j_0) \in E_1} F_{i,j_0}(x_i, y_{j_0}) \right\rangle_{y_{j_0} \in J} \prod_{\substack{(i,j) \in E_1 \\ i \neq i_0, j \neq j_0}} F_{i,j}(x_i, y_j) \right]_{\substack{x_i \in I \text{ for } i \in \mathcal{X}_1 \setminus \{i_0\} \\ y_j \in J \text{ for } j \in \mathcal{Y}_1 \setminus \{j_0\}}} \end{aligned}$$

and then estimate  $|\mathcal{A}_Q^{(1)}| \leq \frac{1}{2} \tilde{\mathcal{A}}_Q + \frac{1}{2} \bar{\mathcal{A}}_Q$ , where

$$\begin{aligned} \tilde{\mathcal{A}}_Q &:= \left[ \left\langle \prod_{j:(i_0,j) \in E_1} F_{i_0,j}(x_{i_0}, y_j) \right\rangle_{x_{i_0} \in I}^2 \prod_{\substack{(i,j) \in E_1 \\ i \neq i_0, j \neq j_0}} F_{i,j}(x_i, y_j) \right]_{\substack{x_i \in I \text{ for } i \in \mathcal{X}_1 \setminus \{i_0\} \\ y_j \in J \text{ for } j \in \mathcal{Y}_1 \setminus \{j_0\}}}, \\ \bar{\mathcal{A}}_Q &:= \left[ \left\langle \prod_{i:(i,j_0) \in E_1} F_{i,j_0}(x_i, y_{j_0}) \right\rangle_{y_{j_0} \in J}^2 \prod_{\substack{(i,j) \in E_1 \\ i \neq i_0, j \neq j_0}} F_{i,j}(x_i, y_j) \right]_{\substack{x_i \in I \text{ for } i \in \mathcal{X}_1 \setminus \{i_0\} \\ y_j \in J \text{ for } j \in \mathcal{Y}_1 \setminus \{j_0\}}}. \end{aligned}$$

Consider the new entangled paraproducts  $\tilde{\Theta}$  and  $\bar{\Theta}$  with local versions defined by

$$\tilde{\Theta}_{\mathcal{T}} := \sum_{Q \in \mathcal{T}} |Q| \tilde{\mathcal{A}}_Q \quad \text{and} \quad \bar{\Theta}_{\mathcal{T}} := \sum_{Q \in \mathcal{T}} |Q| \bar{\mathcal{A}}_Q.$$

Observe that  $\tilde{\Theta}$  corresponds to the graph obtained from  $G_1$  by “doubling” the vertex  $x_{i_0}$  and deleting the vertex  $y_{j_0}$ . This way we obtain a (connected or disconnected) graph with at most  $\max\{|\mathcal{X}_1| + 1, |\mathcal{Y}_1| - 1\}$  vertices in each bipartition. It belongs to the class (C1), so we can invoke [7, Proposition 3.1] again. Analogously, we proceed with  $\bar{\Theta}_{\mathcal{T}}$ , which leads us to

$$\Theta_{\mathcal{T}}((F_{i,j})_{(i,j) \in E}) \leq \frac{1}{2} \tilde{\Theta}_{\mathcal{T}}((F_{i,j})_{(i,j) \in E}) + \frac{1}{2} \bar{\Theta}_{\mathcal{T}}((F_{i,j})_{(i,j) \in E}) \lesssim 1.$$

Note that in [7, Proposition 3.1] the thresholds  $d_{i,j}$  are defined to be  $\max\{|\mathcal{X}_1|, |\mathcal{Y}_1|\}$ , but since we apply this proposition to a modified connected component with component size changed by 1, we only obtain the desired estimate with the threshold  $\max\{|\mathcal{X}_1|, |\mathcal{Y}_1|\} + 1$  with respect to the original graph. However, this is all we claimed when  $\max\{|\mathcal{X}_1|, |\mathcal{Y}_1|\} \geq 3$  and the component in question is not a complete graph, and thus proves the desired bound for  $\Theta_{\mathcal{T}}$  in this case.

Under the assumption of case 1 the graph  $G_1$  is never complete. It remains to consider the case when  $G_1$  satisfies  $\max\{|\mathcal{X}_1|, |\mathcal{Y}_1|\} \leq 2$ . After relabeling if necessary, this implies under the assumption of case 1 that  $E_1 = \{(1, 1), (1, 2), (2, 1)\}$ ,  $S = \{2\}$ ,  $T = \{2\}$ . This graph was already depicted in the left half of Figure 1. Using the Cauchy–Schwarz inequality,  $Q = I \times J \in \mathcal{T}$ , and  $[F_{1,1}^2]_Q^{1/2} \leq 1$ ,

$$\begin{aligned} |\mathcal{A}_Q^{(1)}| &= \left| \left[ \langle F_{2,1}(x_2, y_1) \rangle_{x_2 \in I} \langle F_{1,2}(x_1, y_2) \rangle_{y_2 \in J} F_{1,1}(x_1, y_1) \right]_{x_1 \in I, y_1 \in J} \right| \\ &\leq \left[ \langle F_{2,1}(x_2, y_1) \rangle_{x_2 \in I}^2 \langle F_{1,2}(x_1, y_2) \rangle_{y_2 \in J}^2 \right]_{x_1 \in I, y_1 \in J}^{1/2} \\ &\quad \times [F_{1,1}(x_1, y_1)^2]_{x_1 \in I, y_1 \in J}^{1/2} \\ &= \left[ \langle F_{2,1}(x_2, y_1) \rangle_{x_2 \in I}^2 \right]_{y_1 \in J}^{1/2} \left[ \langle F_{1,2}(x_1, y_2) \rangle_{y_2 \in J}^2 \right]_{x_1 \in I}^{1/2} \\ &\quad \times [F_{1,1}(x_1, y_1)^2]_{x_1 \in I, y_1 \in J}^{1/2} \\ &\leq \frac{1}{2} \left[ \langle F_{2,1}(x_2, y_1) \rangle_{x_2 \in I}^2 \right]_{y_1 \in J} + \frac{1}{2} \left[ \langle F_{1,2}(x_1, y_2) \rangle_{y_2 \in J}^2 \right]_{x_1 \in I}. \end{aligned}$$

This time we ended up with paraproducts

$$\begin{aligned} \tilde{\Theta}_{\mathcal{T}} &:= \sum_{I \times J \in \mathcal{T}} |I \times J| \left[ \langle F_{2,1}(x, y) \rangle_{x \in I} \right]_{y \in J}, \\ \bar{\Theta}_{\mathcal{T}} &:= \sum_{I \times J \in \mathcal{T}} |I \times J| \left[ \langle F_{1,2}(x, y) \rangle_{y \in J} \right]_{x \in I}, \end{aligned}$$

corresponding to two graphs on only three vertices. They are both in the class (C1), which completes the discussion of case 1.

*Case 2.* The vertices  $x_{i_0}$  and  $y_{j_0}$  of  $G$  belong to different components.

Without loss of generality, we can assume that  $i_0 \in \mathcal{X}_1$  and  $j_0 \in \mathcal{Y}_2$ . By [7, Lemma 3.2] again we can estimate

$$|\mathcal{A}_Q| = \prod_{\ell=1}^k |\mathcal{A}_Q^{(\ell)}| \leq |\mathcal{A}_Q^{(1)}| |\mathcal{A}_Q^{(2)}| \leq \frac{1}{2} (\mathcal{A}_Q^{(1)})^2 + \frac{1}{2} (\mathcal{A}_Q^{(2)})^2$$

for each  $Q \in \mathcal{T}$ . This in turn immediately reduces bound (3.7) for  $\Theta_{\mathcal{T}}$  to the same bound for

$$\tilde{\Theta}_{\mathcal{T}} := \sum_{Q \in \mathcal{T}} |Q| (\mathcal{A}_Q^{(1)})^2 \quad \text{and} \quad \bar{\Theta}_{\mathcal{T}} := \sum_{Q \in \mathcal{T}} |Q| (\mathcal{A}_Q^{(2)})^2.$$

Both of these are indeed localized versions of entangled paraproductions falling into the class (C1). More precisely,  $\tilde{\Theta}$  corresponds to a graph with two components, each being a copy of  $G_1$ , so Proposition 3.1 from [7] can be applied once again and it yields (3.7) with all exponents equal to  $\max\{|\mathcal{X}_1|, |\mathcal{Y}_1|\}$ . Similarly we deal with  $\bar{\Theta}$ . Note that we have proved a sharper variant of (3.7) again, that is, the one when  $d_{i,j}$  is replaced with  $\max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\}$  for  $(i, j) \in E_\ell$ .  $\square$

**3.3. Non-cancellative paraproductions.** Now we handle paraproduction types (NC1) and (NC2).

**PROPOSITION 4.** *Suppose that  $\Theta$  is a non-cancellative entangled dyadic paraproduction defined by (3.4) with coefficients (3.5).*

(a) *If conditions (2.8) hold, then the coefficients  $\lambda = (\lambda_Q)_{Q \in \mathcal{C}}$  satisfy*

$$\|\lambda\|_{\text{bmo}} := \sup_{Q_0 \in \mathcal{C}} \left( \frac{1}{|Q_0|} \sum_{Q \in \mathcal{C}: Q \subseteq Q_0} |Q| |\lambda_Q|^2 \right)^{1/2} \lesssim 1.$$

(b) *Each localized form  $\Theta_{\mathcal{T}}$  for a finite convex tree  $\mathcal{T}$  satisfies*

$$\Theta_{\mathcal{T}}((F_{i,j})_{(i,j) \in E}) \lesssim \|\lambda\|_{\text{bmo}} |Q_{\mathcal{T}}| \prod_{(i,j) \in E} \max_{Q \in \mathcal{T} \cup \mathcal{L}(\mathcal{T})} [F_{i,j}^{d_{i,j}}]_Q^{1/d_{i,j}}.$$

*Proof.* (a) Let us begin by fixing  $(u, v) \in E$  and computing  $T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2})$ . We achieve this by substituting  $F_{i,j} = \mathbf{1}_{\mathbb{R}^2}$  for  $(i, j) \neq (u, v)$  in  $\Lambda_E((F_{i,j})_{(i,j) \in E})$  and using representation (3.3). In this case,

$$\begin{aligned} & \Theta_E^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}}((F_{i,j})_{(i,j) \in E}) \\ &= \sum_{I \times J \in \mathcal{C}} \lambda_{I \times J}^{\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)}, \mathbf{b}^{(1)}, \dots, \mathbf{b}^{(n)}} |I|^{2-m-n} \\ & \quad \times \left( \prod_{i \neq u} \int_{\mathbb{R}} \mathbf{a}_I^{(i)} \right) \left( \prod_{j \neq v} \int_{\mathbb{R}} \mathbf{b}_J^{(j)} \right) \int_{\mathbb{R}^2} F_{u,v}(x, y) \mathbf{a}_I^{(u)}(x) \mathbf{b}_J^{(v)}(y) dx dy, \end{aligned}$$



which can only be non-zero when  $S \subseteq \{u\}$  and  $T \subseteq \{v\}$ . Let us rather denote the coefficients by  $\lambda_{I \times J}^{S,T}$ , so that

$$\Lambda_E((F_{i,j})_{(i,j) \in E}) = \sum_{I \times J \in \mathcal{C}} \int_{\mathbb{R}^2} F_{u,v}(x,y) (\lambda_{I \times J}^{\emptyset, \{v\}} \mathbf{1}_I(x) \mathbf{h}_J(y) + \lambda_{I \times J}^{\{u\}, \emptyset} \mathbf{h}_I(x) \mathbf{1}_J(y) + \lambda_{I \times J}^{\{u\}, \{v\}} \mathbf{h}_I(x) \mathbf{h}_J(y)) dx dy.$$

Since  $F_{u,v}$  can be chosen arbitrarily, using (2.4) we have obtained

$$T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2}) = \sum_{I \times J \in \mathcal{C}} (\lambda_{I \times J}^{\emptyset, \{v\}} \mathbf{1}_I \otimes \mathbf{h}_J + \lambda_{I \times J}^{\{u\}, \emptyset} \mathbf{h}_I \otimes \mathbf{1}_J + \lambda_{I \times J}^{\{u\}, \{v\}} \mathbf{h}_I \otimes \mathbf{h}_J).$$

From (2.5), one easily concludes

$$\begin{aligned} (3.11) \quad & \|T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2})\|_{\text{BMO}(\mathbb{R}^2)} \\ &= \sup_{Q_0 \in \mathcal{C}} \left( \frac{1}{|Q_0|} \sum_{Q \in \mathcal{C}: Q \subseteq Q_0} |Q| (|\lambda_Q^{\emptyset, \{v\}}|^2 + |\lambda_Q^{\{u\}, \emptyset}|^2 + |\lambda_Q^{\{u\}, \{v\}}|^2) \right)^{1/2}. \end{aligned}$$

Combining condition (2.8) and equality (3.11), we obtain

$$\|\lambda^{\emptyset, \{v\}}\|_{\text{bmo}}, \|\lambda^{\{u\}, \emptyset}\|_{\text{bmo}}, \|\lambda^{\{u\}, \{v\}}\|_{\text{bmo}} \lesssim 1.$$

It remains to observe that for each non-cancellative paraproduct  $\Theta$  there exists  $(u, v) \in E$  such that  $S \subseteq \{u\}$  and  $T \subseteq \{v\}$ , simply by definition. Therefore, its coefficients satisfy the desired bound.

(b) Again we normalize to achieve  $|Q_{\mathcal{T}}| = 1$  and (3.10). Proceed by applying the Cauchy–Schwarz inequality to get

$$\Theta_{\mathcal{T}} \leq \left( \sum_{Q \in \mathcal{T}} |Q| |\lambda_Q|^2 \right)^{1/2} \left( \sum_{Q \in \mathcal{T}} |Q| (\mathcal{A}_Q)^2 \right)^{1/2}.$$

On the one hand,

$$\sum_{Q \in \mathcal{T}} |Q| |\lambda_Q|^2 \leq \sum_{Q \in \mathcal{C}: Q \subseteq Q_{\mathcal{T}}} |Q| |\lambda_Q|^2 \leq \|\lambda\|_{\text{bmo}}^2 |Q_{\mathcal{T}}| = \|\lambda\|_{\text{bmo}}^2.$$

On the other hand,

$$\tilde{\Theta}_{\mathcal{T}} := \sum_{Q \in \mathcal{T}} |Q| (\mathcal{A}_Q)^2$$

is an example of a local version of an entangled paraproduct falling into the class (C1). Its graph is obtained by simply doubling all connected components of  $G$ . From Proposition 3.1 of [7], we get  $\tilde{\Theta}_{\mathcal{T}} \lesssim 1$ , so  $\Theta_{\mathcal{T}} \lesssim \|\lambda\|_{\text{bmo}}$ , which completes the proof.  $\square$

Let us comment once again that Propositions 3 and 4 together prove estimate (3.7) for each entangled paraprodut  $\Theta$ , which in turn has to be combined with a stopping time argument and decomposition (3.3) to establish estimate (2.9).

**4. Discussion of the exponent range and the exceptional cases**

The purpose of this section is to ensure that the results of Theorems 1 and 2 are not void for any non-trivial graph  $G$ , that is, whenever  $m, n \geq 2$ . We have already established part (a) of Theorem 1 for “most” of the graphs and now we have to identify and resolve the several exceptional cases. Begin by observing that exponents  $p_{i,j}$  such that  $\sum_{(i,j) \in E} p_{i,j}^{-1} = 1$  and  $d_{i,j} < p_{i,j} < \infty$  will certainly exist as soon as (2.6) is satisfied.

**4.1. Non-emptiness of the range.** Each component  $G_\ell$  has at least  $|\mathcal{X}_\ell| + |\mathcal{Y}_\ell| - 1$  edges, as otherwise it would not be a connected graph. If  $\max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\} \leq 2$ , then the component contributes to the sum in (2.6) with at least

$$\frac{|\mathcal{X}_\ell| + |\mathcal{Y}_\ell| - 1}{\max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\}} \geq 1,$$

while if  $\max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\} \geq 3$ , then it adds at least

$$\frac{|\mathcal{X}_\ell| + |\mathcal{Y}_\ell| - 1}{\max\{|\mathcal{X}_\ell|, |\mathcal{Y}_\ell|\} + 1} \geq \frac{3}{4}.$$

We conclude that (2.6) surely holds when  $G$  has two or more components.

If  $G$  is connected, then by  $m, n \geq 2$  we have

$$\frac{m + n - 1}{\max\{m, n\}} > 1 \quad \text{and} \quad \frac{m + n - 1}{\max\{m, n\} + 1} = 1 + \frac{\min\{m, n\} - 2}{\max\{m, n\} + 1},$$

so (2.6) can fail only when  $\min\{m, n\} = 2$  and  $\max\{m, n\} \geq 3$ . Moreover, if (2.6) is false, then

$$1 \leq \frac{|E|}{m + n - 1} = \frac{|E|}{\max\{m, n\} + 1} = \sum_{(i,j) \in E} \frac{1}{d_{i,j}} \leq 1,$$

which gives  $|E| = m + n - 1$ . We see that the number of edges in  $G$  is by 1 smaller than the number of vertices, which implies that  $G$  is a tree (i.e., it has no cycles), by a well-known result from graph theory.

Therefore, the choice (3.1) guarantees non-trivial  $L^p$  estimates in Theorems 1 and 2 in all but the following exceptional case:  $\min\{m, n\} = 2$ ,  $\max\{m, n\} \geq 3$ , and  $G$  is a tree. By symmetry, we suppose  $m = 2 < n$  and then vertices  $x_1$  and  $x_2$  must have precisely one common neighbor among the  $y_j$ 's, in order for the graph to stay connected and to avoid cycles. By relabeling the vertices, we can suppose that there is  $r \in \{1, \dots, n\}$ ,  $r \geq \lceil \frac{n+1}{2} \rceil \geq 2$  such that

$$E = \{(1, 1), (1, 2), \dots, (1, r), (2, r), \dots, (2, n)\}.$$

For  $r \neq n$ , we can modify the proof of Proposition 3 to obtain the single tree estimate with  $d_{i,j} := n$  for each  $(i, j) \in E$ . The only cancellative para-products that are not covered by [7, Proposition 3.1] are the ones from class (C2). Without loss of generality,  $S = \{1\}$  and  $T = \{n\}$ . Then we can control the para-product term as

$$\begin{aligned}
 |\mathcal{A}_Q| &= \left| \left\langle \prod_{j=1}^r F_{1,j}(x_1, y_j) \right\rangle_{x_1 \in I} \times \left\langle F_{2,n}(x_2, y_n) \right\rangle_{y_n \in J} \prod_{j=r}^{n-1} F_{2,j}(x_2, y_j) \right|_{x_2 \in I, y_1, \dots, y_{n-1} \in J} \\
 (4.1) \quad &\leq \frac{1}{2} \left[ \left\langle \prod_{j=1}^r F_{1,j}(x_1, y_j) \right\rangle_{x_1 \in I}^2 \prod_{j=r}^{n-1} F_{2,j}(x_2, y_j) \right]_{x_2 \in I, y_1, \dots, y_{n-1} \in J} \\
 (4.2) \quad &+ \frac{1}{2} \left[ \left\langle F_{2,n}(x_2, y_n) \right\rangle_{y_n \in J}^2 \prod_{j=r}^{n-1} F_{2,j}(x_2, y_j) \right]_{x_2 \in I, y_1, \dots, y_{n-1} \in J}.
 \end{aligned}$$

Note that (4.1) and (4.2) lead to (possibly disconnected) graphs having at most

$$\max\{3, n - 1, n - r + 2, r + 1\} \leq n$$

vertices in each of their bipartition vertex-sets, so [7, Proposition 3.1] can be applied once again. On the other hand, the proof of Proposition 4, which handles non-cancellative para-products, can be left unchanged. This in turn establishes (2.9) whenever conditions (2.7) and (2.8) are satisfied.

Finally, the only case we left out is when  $r = n$ . An example of such graph is shown in Figure 2. This time we rather take

$$(4.3) \quad d_{1,j} := n \quad \text{for } j = 1, \dots, n - 1, \quad d_{1,n} := 2n - 2, \quad d_{2,n} := n.$$

Cancellative para-products of type (C1) are again resolved by [7, Proposition 3.1], even in a slightly larger range. All para-products in class (C2) are

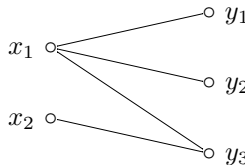


FIGURE 2. Example of an exceptional bipartite graph.

essentially the same, so choose  $S = \{2\}$ ,  $T = \{1\}$  and begin by estimating

$$\begin{aligned}
 |\mathcal{A}_Q| &= \left| \left[ \left\langle F_{1,1}(x_1, y_1) \right\rangle_{y_1 \in J} \left\langle F_{2,n}(x_2, y_n) \right\rangle_{x_2 \in I} \prod_{j=2}^n F_{1,j}(x_1, y_j) \right]_{x_1 \in I, y_2, \dots, y_n \in J} \right| \\
 &\leq \left[ \left\langle F_{1,1}(x_1, y_1) \right\rangle_{y_1 \in J}^2 \right]_{x_1 \in I}^{1/2} \left[ \left\langle F_{2,n}(x_2, y_n) \right\rangle_{x_2 \in I}^2 \right]_{y_n \in J}^{1/2} \\
 &\quad \times \left[ \prod_{j=2}^n F_{1,j}(x_1, y_j)^2 \right]_{x_1 \in I, y_2, \dots, y_n \in J}^{1/2}.
 \end{aligned}$$

From (3.10) and

$$\left[ \prod_{j=2}^n F_{1,j}(x_1, y_j)^2 \right]_{x_1 \in I, y_2, \dots, y_n \in J}^{1/2} \leq \prod_{j=2}^n [F_{1,j}(x_1, y_j)^{2n-2}]_{x_1 \in I, y_j \in J}^{1/(2n-2)}$$

we obtain

$$|\mathcal{A}_Q| \leq \frac{1}{2} \left[ \left\langle F_{1,1}(x_1, y_1) \right\rangle_{y_1 \in J}^2 \right]_{x_1 \in I} + \frac{1}{2} \left[ \left\langle F_{2,n}(x_2, y_n) \right\rangle_{x_2 \in I}^2 \right]_{y_n \in J},$$

which this time yields the single tree estimate (3.7) with an unusual choice  $\tilde{d}_{1,j} = 2n - 2$  for each  $1 \leq j \leq n$  and  $\tilde{d}_{2,n} = 2$ . Consequently, bound (2.9) holds for  $p_{i,j} > \tilde{d}_{i,j}$ . We are going to apply the fiberwise Calderón–Zygmund decomposition of Bernicot [3] to expand the exponent region so that it contains the one determined by (4.3). This is important because all entangled para-products corresponding to the same multilinear form should satisfy the same  $L^p$  estimate in order for this estimate to hold for the form  $\Lambda_E$  itself.

We only give a short comment on Bernicot’s argument, as we apply it in almost exactly the same way as in [8, Section 5]. By choosing  $(u, v) = (1, n)$  and by duality, the problem becomes to extend estimates

$$\|T_{1,n}(F_{2,n}, F_{1,1}, \dots, F_{1,n-1})\|_{L^p(\mathbb{R}^2)} \lesssim \|F_{2,n}\|_{L^{p_{2,n}}(\mathbb{R}^2)} \prod_{j=1}^{n-1} \|F_{1,j}\|_{L^{p_{1,j}}(\mathbb{R}^2)}$$

for the operator

$$\begin{aligned}
 &T_{1,n}(F_{2,n}, F_{1,1}, \dots, F_{1,n-1})(x, y) \\
 &= \sum_{I \times J \in \mathcal{C}} \lambda_{I \times J} |I|^{-n} \left( \int_{\mathbb{R}} F_{2,n}(x', y) \mathbf{a}_I^{(2)}(x') dx' \right) \\
 &\quad \times \prod_{j=1}^{n-1} \left( \int_{\mathbb{R}} F_{1,j}(x, y') \mathbf{b}_J^{(j)}(y') dy' \right) \mathbf{a}_I^{(1)}(x) \mathbf{b}_J^{(n)}(y)
 \end{aligned}$$

from the range

$$\frac{1}{p_{2,n}} < \frac{1}{2}, \quad \frac{1}{p_{1,1}}, \dots, \frac{1}{p_{1,n-1}} < \frac{1}{2n-2}, \quad \frac{1}{p} = \frac{1}{p_{2,n}} + \sum_{j=1}^{n-1} \frac{1}{p_{1,j}} > 1 - \frac{1}{2n-2}$$

to a range that includes

$$(4.4) \quad \frac{1}{p_{2,n}}, \frac{1}{p_{1,1}}, \dots, \frac{1}{p_{1,n-1}} < \frac{1}{n}, \quad \frac{1}{p} = \frac{1}{p_{2,n}} + \sum_{j=1}^{n-1} \frac{1}{p_{1,j}} > 1 - \frac{1}{2n-2}.$$

We do that by performing the Calderón–Zygmund decomposition in each fiber  $F_{1,j}(x, \cdot)$  for  $j = 1, \dots, n - 1$  respectively. In  $j$ th step that extends the range of weak  $L^p$  estimates by raising  $p_{1,j}^{-1}$  all the way to 1. The intersection of the new range with (4.4) lies inside the Banach simplex, so real multilinear interpolation actually provides strong estimates there. In  $n - 1$  steps we recover the whole region (4.4).

**4.2. A counterexample.** Finally, let us show that the degenerate case  $m = 1$  gives objects that need not be bounded just by assuming all other conditions of Theorem 1. Since  $G$  can have no isolated vertices, there must be an edge between  $x_1$  and each of the vertices  $y_j$ , so  $E = \{(1, 1), \dots, (1, n)\}$ .

If  $m = n = 1$ , there can be no open range of estimates, so suppose  $m = 1 < n$ . The form can be written as

$$\begin{aligned} \Lambda_E(F_{1,1}, \dots, F_{1,n}) &= \int_{\mathbb{R}} \left( \int_{\mathbb{R}^n} K(x_1, y_1, \dots, y_n) \prod_{j=1}^n F_{1,j}(x_1, y_j) dy_1 \cdots dy_n \right) dx_1. \end{aligned}$$

One can notice that the inner integral has the symmetry of a dyadic one-dimensional multilinear Calderón–Zygmund form for each fixed  $x_1$ . Still, we claim that the two-dimensional conditions (2.7) and (2.8) are not enough to have any  $L^p$  bounds for  $\Lambda_E$ . Instead one should impose one-dimensional testing conditions in each fiber of  $K$ .

To present the counterexample, we take the kernel

$$K(x_1, y_1, \dots, y_n) := \sum_{I \times J \in \mathcal{C}} \lambda_{I \times J} |I|^{1-n} \mathbf{h}_I(x_1) \mathbf{1}_J(y_1) \cdots \mathbf{1}_J(y_n),$$

where the coefficients  $\lambda = (\lambda_Q)_{Q \in \mathcal{C}}$  are given by

$$\lambda_{I \times J} := \begin{cases} 1 & \text{if } I = [0, 2^{-k}) \text{ and } J \subseteq [0, 1) \text{ for some } k \in \{0, 1, \dots, r - 1\}, \\ 0 & \text{otherwise} \end{cases}$$

and  $r$  is a positive integer. Using only  $|\lambda_{I \times J}| \leq 1$ , we get

$$\begin{aligned} |K(x_1, y_1, \dots, y_n)| &\leq \sum_{\substack{I \times J \in \mathcal{C} \\ x_1 \in I, y_1, \dots, y_n \in J}} |J|^{1-n} \lesssim \left( \min_{\substack{J \in \mathcal{D} \\ y_1, \dots, y_n \in J}} |J| \right)^{1-n} \\ &\leq \left( \max_{1 \leq j_1 < j_2 \leq n} |y_{j_1} - y_{j_2}| \right)^{1-n}, \end{aligned}$$

under conventions  $\min \emptyset = \infty$  and  $\frac{1}{\infty} = 0$ , which confirms the size estimate (2.1) with  $m = 1$ . It is also immediate to verify

$$|\Lambda_E(\mathbf{1}_Q, \dots, \mathbf{1}_Q)| = \left| \sum_{Q' \in \mathcal{C}: Q' \supseteq Q} \lambda_{Q'} |Q'|^{(1-n)/2} |Q|^{(n+1)/2} \right| \leq \frac{|Q| (\frac{1}{2})^{n-1}}{1 - (\frac{1}{2})^{n-1}} \leq |Q|$$

for each  $Q \in \mathcal{C}$  and

$$\begin{aligned} \|T_{1,j}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2})\|_{\text{BMO}(\mathbb{R}^2)} &= \left\| \sum_{I \times J \in \mathcal{C}} \lambda_{I \times J} \mathbf{h}_I \otimes \mathbf{1}_J \right\|_{\text{BMO}(\mathbb{R}^2)} \\ &= \|\lambda\|_{\text{bmo}} = \left( \sum_{k=0}^{r-1} 2^k (2^{-k})^2 \right)^{1/2} \leq \sqrt{2} \end{aligned}$$

for  $1 \leq j \leq n$ . Thus, the conditions of Theorem 1 are fulfilled.

On the other hand, for each  $j$  we define

$$F_{1,j}(x, y) := \begin{cases} \left(\frac{2^\ell}{\ell(\ell+1)}\right)^{1/n} & \text{for } x \in [2^{-\ell}, 2^{-\ell+1}), y \in [0, 1), \ell \in \{1, \dots, r\}, \\ 0 & \text{otherwise,} \end{cases}$$

and substitute into (2.2) to obtain

$$\begin{aligned} \Lambda_E(F_{1,1}, \dots, F_{1,n}) &= \sum_{k=0}^{r-1} 2^k (2^{-k})^{1-n} (2^{-k})^n \left( \sum_{\ell=k+2}^r \frac{1}{\ell(\ell+1)} - \frac{1}{(k+1)(k+2)} \right) \\ &= \sum_{k=0}^{r-1} \frac{1}{k+2} - \frac{r-1}{r+1} - 1 + \frac{1}{r+2} \rightarrow \infty \quad \text{as } r \rightarrow \infty. \end{aligned}$$

Since

$$\|F_{1,j}\|_{L^n} = \left( \sum_{\ell=1}^r 2^{-\ell} \frac{2^\ell}{\ell(\ell+1)} \right)^{1/n} = \left( 1 - \frac{1}{r+1} \right)^{1/n} \leq 1,$$

we see that bound (2.9) does not hold when  $p_{1,1} = \dots = p_{1,n} = n$ . Indeed, the estimate cannot hold for any choice of the exponents, because symmetry and interpolation would then recover the above ‘‘central point’’ bound.

### 5. Necessity of the testing conditions

We begin this section by reducing Theorem 2 to part (a) of Theorem 1.

*Proof of Theorem 2.* Assume that conditions (2.12) are satisfied. Let  $r > 0$  be large enough so that  $K$  is supported in  $[-r, r]^{m+n}$ .

Fix  $(u, v) \in E$  and take a square  $Q = I \times J \in \mathcal{C}$ . Let us temporarily denote

$$\mathcal{C}(Q) := \{Q' \in \mathcal{C} : |Q'| = |Q|, Q' \text{ intersects } [-r, r] \times [-r, r]\},$$

which is obviously a finite collection. We start by decomposing

$$\begin{aligned}
 & T_{u,v}((\mathbf{1}_{\mathbb{R}^2})_{(i,j) \in E \setminus \{(u,v)\}})(x_u, y_v) \mathbf{1}_Q(x_u, y_v) \\
 &= \sum_{\substack{Q_{i,j} \in \mathcal{C}(Q) \\ \text{for each } (i,j) \in E \setminus \{(u,v)\}}} T_{u,v}((\mathbf{1}_{Q_{i,j}})_{(i,j) \in E \setminus \{(u,v)\}})(x_u, y_v) \\
 &\quad \times \mathbf{1}_Q(x_u, y_v) \\
 &= \sum_{\substack{Q_{i,j} \in \mathcal{C}(Q) \\ \text{for each } (i,j) \in E \setminus \{(u,v)\}}} \int_{\mathbb{R}^{m+n-2}} K(x_1, \dots, x_m, y_1, \dots, y_n) \\
 &\quad \times \prod_{(i,j) \in E} \mathbf{1}_{Q_{i,j}}(x_i, y_j) \prod_{i \neq u} dx_i \prod_{j \neq v} dy_j,
 \end{aligned}$$

where for convenience we write  $Q_{u,v} := Q$ . Due to the term  $\prod_{(i,j)} \mathbf{1}_{Q_{i,j}}(x_i, y_j)$  the above summand can be non-zero only when it is of the form

$$\begin{aligned}
 (5.1) \quad & \int_{\mathbb{R}^{m+n-2}} K(x_1, \dots, x_m, y_1, \dots, y_n) \\
 & \times \prod_{i=1}^m \mathbf{1}_{I^{(i)}}(x_i) \prod_{j=1}^n \mathbf{1}_{J^{(j)}}(y_j) \prod_{i \neq u} dx_i \prod_{j \neq v} dy_j
 \end{aligned}$$

for some dyadic intervals  $I^{(1)}, \dots, I^{(m)}, J^{(1)}, \dots, J^{(n)}$  of the same length and such that  $I^{(u)} \times J^{(v)} = Q$ . Recall that  $K$  is constant on all dyadic cubes in  $\mathbb{R}^{m+n}$  that are disjoint from the diagonal  $D$ . Thus, the expression in (5.1) is constant for all  $(x_u, y_v) \in Q$ , except possibly when  $I^{(i)} = I$  for each  $i$  and  $J^{(j)} = J$  for each  $j$ , in which case (5.1) becomes simply

$$\begin{aligned}
 & \int_{\mathbb{R}^{m+n-2}} K(x_1, \dots, x_m, y_1, \dots, y_n) \prod_{(i,j) \in E} \mathbf{1}_Q(x_i, y_j) \prod_{i \neq u} dx_i \prod_{j \neq v} dy_j \\
 &= T_{u,v}((\mathbf{1}_Q)_{(i,j) \in E \setminus \{(u,v)\}})(x_u, y_v) \mathbf{1}_Q(x_u, y_v).
 \end{aligned}$$

This discussion leads us to

$$\begin{aligned}
 (5.2) \quad & T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2})(x, y) - \frac{1}{|Q|} \int_Q T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2}) \\
 &= T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)(x, y) - \frac{1}{|Q|} \int_Q T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)
 \end{aligned}$$

for every  $(x, y) \in Q$ . Combining this equality with conditions (2.12) we obtain

$$\begin{aligned}
 & \left| \frac{1}{|Q|} \int_Q T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2}) - \frac{1}{|Q|} \int_Q T_{u,v}(\mathbf{1}_{\mathbb{R}^2}, \dots, \mathbf{1}_{\mathbb{R}^2}) \right| \\
 & \leq \frac{2}{|Q|} \int_Q |T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)| \lesssim 1.
 \end{aligned}$$

It is well known (as an easy consequence of the John–Nirenberg inequality) that the quantity

$$\sup_{Q \in \mathcal{C}} \frac{1}{|Q|} \int_Q \left| F - \frac{1}{|Q|} \int_Q F \right|$$

is comparable with  $\|F\|_{\text{BMO}(\mathbb{R}^2)}$ , see [1, Section 3]. That way we have established condition (2.8).

Verification of condition (2.7) using (2.12) is easy, as for some  $(u, v) \in E$  and an arbitrary  $Q \in \mathcal{C}$ ,

$$\begin{aligned} |\Lambda_E(\mathbf{1}_Q, \dots, \mathbf{1}_Q)| &= \left| \int_{\mathbb{R}^2} T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q) \mathbf{1}_Q \right| \\ &\leq \|T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)\|_{L^1(Q)} \lesssim |Q|. \end{aligned}$$

Therefore, Theorem 1(a) can be applied and this completes the proof.  $\square$

In order to prove part (b) of Theorem 1, it is now enough to reduce its hypotheses to conditions (2.12), which imply (2.7) and (2.8), as we have just shown.

*Proof of Theorem 1(b).* Suppose that estimate (2.9) holds with some choice of exponents  $p_{i,j}$ . Fix  $(u, v) \in E$ , take  $Q \in \mathcal{C}$ , choose  $F_{i,j} = \mathbf{1}_Q$  for  $(i, j) \in E \setminus \{(u, v)\}$ , and leave  $F_{u,v}$  arbitrary. Let  $p'_{u,v}$  denote the conjugated exponent of  $p_{u,v}$ . By (2.4) and (2.9),

$$\begin{aligned} \left| \int_{\mathbb{R}^2} T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q) F_{u,v} \right| \\ \lesssim \|F_{u,v}\|_{L^{p_{u,v}}} \prod_{(i,j) \neq (u,v)} \|\mathbf{1}_Q\|_{L^{p_{i,j}}} = \|F_{u,v}\|_{L^{p_{u,v}}} |Q|^{1/p'_{u,v}}, \end{aligned}$$

so duality implies

$$\|T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)\|_{L^{p'_{u,v}}(Q)} \lesssim |Q|^{1/p'_{u,v}}.$$

Furthermore, Jensen's inequality gives

$$\frac{1}{|Q|} \int_Q |T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)| \leq \left( \frac{1}{|Q|} \int_Q |T_{u,v}(\mathbf{1}_Q, \dots, \mathbf{1}_Q)|^{p'_{u,v}} \right)^{1/p'_{u,v}} \lesssim 1,$$

which is exactly (2.12).  $\square$

Let us conclude with a comment that the short proofs given in this section are adaptations of classical arguments from [9]. If the perfect cancellation of  $K$  was replaced with the standard Hölder continuity condition, the difference of the two sides in (5.2) would only be a bounded function, which would still be enough to follow the reasoning of the corresponding part of the proof.



## REFERENCES

- [1] P. Auscher, S. Hofmann, C. Muscalu, T. Tao and C. Thiele, *Carleson measures, trees, extrapolation, and  $T(b)$  theorems*, Publ. Mat. **46** (2002), no. 2, 257–325. MR 1934198
- [2] A. Bényi, C. Demeter, A. R. Nahmod, C. Thiele, R. H. Torres and P. Villarroya, *Modulation invariant bilinear  $T(1)$  theorem*, J. Anal. Math. **109** (2009), 279–352. MR 2585397
- [3] F. Bernicot, *Fiber-wise Calderón–Zygmund decomposition and application to a bi-dimensional paraproduct*, Illinois J. Math. **56** (2012) 415–422. MR 3161332
- [4] G. David and J.-L. Journé, *A boundedness criterion for generalized Calderón–Zygmund operators*, Ann. of Math. (2) **120** (1984), no. 2, 371–397. MR 0763911
- [5] C. Demeter and C. Thiele, *On the two-dimensional bilinear Hilbert transform*, Amer. J. Math. **132** (2010), no. 1, 201–256. MR 2597511
- [6] L. Grafakos and R. H. Torres, *Multilinear Calderón–Zygmund theory*, Adv. Math. **165** (2002), no. 1, 124–164. MR 1880324
- [7] V. Kovač, *Bellman function technique for multilinear estimates and an application to generalized paraproducts*, Indiana Univ. Math. J. **60** (2011), no. 3, 813–846. MR 2985857
- [8] V. Kovač, *Boundedness of the twisted paraproduct*, Rev. Mat. Iberoam. **28** (2012), no. 4, 1143–1164. MR 2990138
- [9] E. M. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, Princeton Math. Ser., vol. 43, Princeton Univ. Press, Princeton, NJ, 1993. MR 1232192

VJEKOSLAV KOVAČ, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ZAGREB, BIJENIČKA CESTA 30, 10000 ZAGREB, CROATIA

*E-mail address:* [vjekovac@math.hr](mailto:vjekovac@math.hr)

CHRISTOPH THIELE, UNIVERSITY OF CALIFORNIA AT LOS ANGELES AND UNIVERSITÄT BONN, ENDENICHER ALLEE 60, 53115 BONN, GERMANY

*E-mail address:* [thiele@math.uni-bonn.de](mailto:thiele@math.uni-bonn.de)