

Chapman University

Chapman University Digital Commons

Mathematics, Physics, and Computer Science
Faculty Articles and Research

Science and Technology Faculty Articles and
Research

9-27-2022

Local Bounds for Singular Brascamp–Lieb Forms with Cubical Structure

Polona Durcik

Lenka Slavíková

Chrisoph Thiele

Follow this and additional works at: https://digitalcommons.chapman.edu/scs_articles

Local Bounds for Singular Brascamp–Lieb Forms with Cubical Structure

Comments

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in *Mathematische Zeitschrift* in 2022 following peer review. The final publication may differ and is available at Springer via <https://doi.org/10.1007/s00209-022-03148-8>.

[A free-to-read copy of the final published article is available here.](#)

Copyright

Springer

LOCAL BOUNDS FOR SINGULAR BRASCAMP-LIEB FORMS WITH CUBICAL STRUCTURE

POLONA DURCIK, LENKA SLAVÍKOVÁ, AND CHRISTOPH THIELE

ABSTRACT. We prove a range of L^p bounds for singular Brascamp-Lieb forms with cubical structure. We pass through sparse and local bounds, the latter proved by an iteration of Fourier expansion, telescoping, and the Cauchy-Schwarz inequality. We allow $2^{m-1} < p \leq \infty$ with m the dimension of the cube, extending an earlier result that required $p = 2^m$. The threshold 2^{m-1} is sharp in our theorems.

1. INTRODUCTION

A singular Brascamp-Lieb form with cubical structure is a principal value integral

$$p.v. \int_{\mathbb{R}^{2m}} \prod_{j \in \mathcal{C}} F_j(\Pi_j x) K(\Pi x) dx. \quad (1.1)$$

It is called "Brascamp-Lieb", because the integrand is a product of several functions F_j and K , each composed with a linear surjection Π_j or Π . It is called singular, because K is a Calderón-Zygmund kernel.

The cubical structure of (1.1) lies in the choice of surjections. We let \mathcal{C} be the set of functions

$$j : \{1, \dots, m\} \rightarrow \{0, 1\}$$

and Π_j the projection from \mathbb{R}^{2m} to \mathbb{R}^m given by

$$\Pi_j(x_1^0, \dots, x_m^0, x_1^1, \dots, x_m^1)^T := (x_1^{j(1)}, \dots, x_m^{j(m)})^T, \quad (1.2)$$

while the surjection $\Pi : \mathbb{R}^{2m} \rightarrow \mathbb{R}^m$ is generic. Each index j can naturally be identified with the corner of an m -dimensional cube $[0, 1]^m$. The cubical structure allows for a symmetrization procedure, pioneered in [13], of the tuple of functions F_j along reflection symmetries of the cube. This is crucial for our understanding of L^p estimates for these forms. The paper [12] provides evidence that singular Brascamp-Lieb forms with cubical structure are key to understanding more general singular Brascamp-Lieb forms. Singular Brascamp-Lieb forms with cubical structure have found applications in enumerative combinatorics [7], [6], [14] and ergodic theory [8], [15].

We fix $m \geq 2$ throughout this paper and always denote by \mathcal{C} the cube and by Π_j the projections as in (1.2). We also fix the notation that components of vectors in \mathbb{R}^{2m} are written x_i^l as in (1.2) and we use matrix product notation for linear maps as in (1.2).

We assume in this paper, unless otherwise stated, that each F_j is a real-valued smooth compactly supported function on \mathbb{R}^m . The principal value in (1.1) means that the

Date: September 7, 2022.

2010 Mathematics Subject Classification. 42B20.

Schwartz function

$$\phi = \prod_{j \in \mathcal{C}} F_j \circ \Pi_j$$

is paired with the tempered distribution $K \circ \Pi$. A Calderón-Zygmund kernel K of degree n in \mathbb{R}^m is the Fourier transform of integration against a function \widehat{K} on $\mathbb{R}^m \setminus \{0\}$ which satisfies the symbol estimates

$$|\partial^\alpha \widehat{K}(\eta)| \leq |\eta|^{-|\alpha|}$$

for all multi-indices of order $|\alpha| \leq n$. Hence the principal value integral can be written as a standard integral

$$p.v. \int_{\mathbb{R}^{2m}} \phi(x) K(\Pi x) dx = \int_{\mathbb{R}^m} \widehat{\phi}(-\Pi^T \eta) \widehat{K}(\eta) d\eta.$$

An $m \times m$ matrix A is called Δ -regular, if

$$\max(|A|, |\det A|^{-1}) \leq \Delta,$$

where $|A|$ denotes the operator norm of A . Note that one can also control $|\det A|$ and $|A^{-1}|$ of a Δ -regular matrix by suitable Δ' depending on Δ .

The main purpose of this paper is to prove the following theorem, which generalizes a theorem in [10].

Theorem 1.1. *Let $\Delta > 1$. For each $j \in \mathcal{C}$ let*

$$2^{m-1} < p_j \leq \infty \tag{1.3}$$

and assume

$$\sum_{j \in \mathcal{C}} \frac{1}{p_j} = 1. \tag{1.4}$$

Then there is a constant C such that the following holds.

Let Π be a real $m \times 2m$ matrix such that for each $j \in \mathcal{C}$ the composition $\Pi_j \Pi^T$ is Δ -regular. For all Calderón-Zygmund kernels K in \mathbb{R}^m of order 2^{6m} and all tuples $(F_j)_{j \in \mathcal{C}}$ we have

$$\left| p.v. \int_{\mathbb{R}^{2m}} \prod_{j \in \mathcal{C}} F_j(\Pi_j x) K(\Pi x) dx \right| \leq C \prod_{j \in \mathcal{C}} \|F_j\|_{p_j}.$$

The result in [10] is the special case $p_j = 2^m$ for all $j \in \mathcal{C}$. There the necessity of regularity of $\Pi_j \Pi^T$ has been discussed in detail. The exponent 2^m is very particular and allows for a global argument in [10]. Any deviation from this exponent requires a localization of the arguments. The main point of the present paper is to develop such localization for these singular Brascamp-Lieb forms. A key difficulty arises from the general position of Π , which prevents the use of algebraic telescoping identities and instead requires cone decomposition (5.13) of certain multipliers. This leads to rectangles of arbitrary eccentricity, expressed via matrices D or E , and correspondingly a use of a mollified strong maximal function for example in (5.18), (5.19) and in (5.2). The localization procedure also leads to boundary terms associated with arbitrary eccentricities, which require some additive measure theoretic arguments in (5.28) and subsequent displays. The use of mollified strong maximal function, a superposition of anisotropically rescaled Hardy-Littlewood maximal functions as seen in Lemma 3.1, and additive measure theory in the context of singular Brascamp-Lieb inequalities seem a novelty in our paper.

Theorem 1.1 should be compared with [5], which proves a special case with $m = 2$ of the theorem, and with [22] in the simpler dyadic setting. Both of these references avoid general surjections Π .

On a formal level, we prove Theorem 1.1 with the technique of sparse bounds, which has seen prominent developments in the past decade and is a vehicle towards further results such as weighted and vector valued inequalities. Concretely, Theorem 1.1 follows from the sparse bounds formulated in the next Theorem 1.2.

A dyadic interval is one of the form $[2^k n, 2^k(n+1))$ with integers k, n . A dyadic cube Q in \mathbb{R}^m is of the form $I_1 \times \cdots \times I_m$, where the I_i are dyadic intervals of the same length. We denote by Q^* the cube with the same center as Q and three times the side length. A sparse ($\frac{1}{2}$ -sparse) collection \mathcal{S} of dyadic cubes is one that satisfies for all $Q \in \mathcal{S}$

$$\sum_{Q' \in \mathcal{S}: Q' \subseteq Q} |Q'| \leq 2|Q|,$$

with $|Q|$ denoting the volume of Q . For a measurable set $S \subseteq \mathbb{R}^m$, typically a cube or parallelepiped, and an integrable function F on \mathbb{R}^m , we denote

$$[F]_S := \left(\frac{1}{|S|} \int_S |F|^{2^{m-1}} \right)^{2^{1-m}}.$$

For a $2m \times m$ matrix J we denote the upper $m \times m$ block by J^0 and the lower $m \times m$ block by J^1 . Assuming that J^1 is regular, we write

$$\Pi^J := (I, -J^0(J^1)^{-1}),$$

which implies $\Pi^J \circ J = 0$. We will see in Section 2 that all maps Π in Theorem 1.1 may and will be assumed to be of the form Π^J .

Theorem 1.2. *Let $\Delta > 1$. For each $j \in \mathcal{C}$, let p_j be such that (1.3) and (1.4) hold. Then there is a constant C such that the following holds.*

Let J be a real $2m \times m$ matrix such that for each $j \in \mathcal{C}$ the matrix

$$A_j := \Pi_j J$$

is Δ -regular. For all Calderón-Zygmund kernels K in \mathbb{R}^m of order 2^{6m} and all tuples $(F_j)_{j \in \mathcal{C}}$ there is a sparse collection \mathcal{S} of dyadic cubes in \mathbb{R}^m such that

$$\left| p.v. \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j x) \right) K(\Pi^J x) dx \right| \leq C \sum_{Q \in \mathcal{S}} |Q| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*}. \quad (1.5)$$

The reduction of Theorem 1.1 to Theorem 1.2 will be presented in Section 2. It follows from a well known paradigm used in [3] in the context of the bilinear Hilbert transform, another prominent singular Brascamp–Lieb form. In our instance, the reduction is essentially a consequence of Theorem 1.11 in [23]. Indeed, one obtains in [23] more general bounds in weighted L^p spaces, with good dependence on the weight classes. Further, one can deduce vector valued bounds as in [20].

We highlight a further reduction of Theorem 1.2 to local bounds formulated in Theorem 1.3 below. The reduction follows some standard procedures and will be elaborated in Section 3. Note the appearance of the modified strong maximal function in Theorem 1.3, which will be key to the inductive approach to this theorem.

The quantity $K(\Pi^J x)$ is invariant under translation in direction of the kernel of Π^J , which is the range of J , and obeys strong regularity assumption invariant under isotropic scaling. It therefore can be naturally written as an integral over the group spanned by

dilations and by translations in direction of the range of J of forms defined by certain bump functions in x in place of $K(\Pi^J x)$. Following [10], using suitable windowed Fourier transforms on the multiplier \widehat{K} , we write (1.1) as superposition of forms

$$\int_0^\infty \int_{\mathbb{R}^m} \frac{c(t)}{1 + |u|^{4m}} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j x) \right) (\partial_i^0 \partial_i^1 g)_t(x - Jp + ut) dx dp \frac{dt}{t}, \quad (1.6)$$

where $1 \leq i \leq m$, $\|c\|_\infty \leq 1$, $u \in \mathbb{R}^{2m}$, $g(x) = e^{-\pi|x|^2}$ is the Gaussian in \mathbb{R}^{2m} , and for a function ϕ on \mathbb{R}^m we use the notation $\phi_t(x) = t^{-m}\phi(t^{-1}x)$. We have written ∂_i^0 and ∂_i^1 in analogy to the convention in (1.2) for vectors in \mathbb{R}^{2m} . It suffices to prove bounds of Theorem 1.2 with the left-hand side of (1.5) replaced by the absolute value of (1.6).

The variables (p, t) in (1.6) are in the upper half space $\mathbb{R}^m \times (0, \infty)$, the above mentioned symmetry group. We decompose the upper half space into boxes $Q \times (\ell(Q)/2, \ell(Q))$, where Q is a dyadic cube and $\ell(Q)$ denotes its side length. For a collection \mathcal{T} of dyadic cubes we define

$$\Omega_{\mathcal{T}} := \cup_{Q \in \mathcal{T}} Q \times (\ell(Q)/2, \ell(Q))$$

and we set

$$\Lambda_{\mathcal{T}}((F_j)_{j \in \mathcal{C}}) := \int_{\Omega_{\mathcal{T}}} \frac{c(t)}{1 + |u|^{4m}} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j x) \right) (\partial_i^0 \partial_i^1 g)_t(x - Jp + ut) dx dp \frac{dt}{t}. \quad (1.7)$$

A collection \mathcal{T} of dyadic cubes is called a convex tree or a stopping time, if there exists $Q_{\mathcal{T}} \in \mathcal{T}$ such that for all $Q \in \mathcal{T}$ we have $Q \subseteq Q_{\mathcal{T}}$ and $Q' \in \mathcal{T}$ whenever $Q \subseteq Q' \subseteq Q_{\mathcal{T}}$.

Let \mathcal{R} be the set of rectangular boxes in \mathbb{R}^m , that is boxes of the form $R = I_1 \times \cdots \times I_m$, with intervals I_1, \dots, I_m that are not necessarily dyadic and not necessarily of equal length. We denote by ε_R the reciprocal of the eccentricity of R , that is

$$\varepsilon_R = \left(\min_{1 \leq k \leq m} |I_k| \right) \left(\max_{1 \leq k \leq m} |I_k| \right)^{-1}.$$

We define for $\alpha \geq 0$ the modified strong maximal operator

$$M_\alpha(F)(x) := \sup_{R \in \mathcal{R}: x \in R} (\varepsilon_R^\alpha)^{2^{1-m}} [F]_R. \quad (1.8)$$

This is a weighted version, with weight depending on the eccentricity, of the classical strong $L^{2^{m-1}}$ maximal operator. We will typically see expressions of the type $M_\alpha(F)(Ax)$ for some $m \times m$ matrix A . For a dyadic cube Q we define

$$M_\alpha(F)(AQ) := \sup_{R \in \mathcal{R}: AQ \subseteq R} (\varepsilon_R^\alpha)^{2^{1-m}} [F]_R.$$

Here AS for some set S means the set of all Ax with $x \in S$. For a collection \mathcal{T} of dyadic cubes we define

$$M_\alpha(F)(A\mathcal{T}) := \sup_{Q \in \mathcal{T}} M_\alpha(F)(AQ).$$

Theorem 1.3. *Let $\Delta > 1$ and $\alpha < \frac{1}{m}$. There exists $C > 0$ such that the following holds. Let J, A_j be as in Theorem 1.2. For any convex tree \mathcal{T} , for all c, u, i satisfying conditions specified near (1.6) and all tuples $(F_j)_{j \in \mathcal{C}}$ we have*

$$|\Lambda_{\mathcal{T}}((F_j)_{j \in \mathcal{C}})| \leq C |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}).$$

The upper bound on the parameter α is not so important for the deduction of Theorem 1.2 from Theorem 1.3, here a classical 2^{m-1} -Hardy-Littlewood operator would do, which is a limit of M_α as α tends to ∞ . The upper bound is however needed for the proof of Theorem 1.3, except in particular cases such as $m = 2$ or $J^0 = J^1 = I$. While the classical strong 2^{m-1} -maximal operator, which is the case $\alpha = 0$ of M_α , does not satisfy a weak type 2^{m-1} bound, the operator M_α with $\alpha > 0$ does, as we will show in Lemma 3.1. Related to this we also recall that the classical strong maximal operator cannot be sparsely dominated [1], while the result of Theorem 1.3 with $\alpha > 0$ allows us to obtain the sparse bounds of Theorem 1.2.

We highlight a further reduction of Theorem 1.3 to estimates in Theorem 1.4 for a form that has localization of the x variables rather than localization in the (p, t) variables. The localization in x is somewhat more precise, because it truncates $2m$ variables rather than m variables. Moreover, it facilitates Fourier transform techniques in the variable p . A truncation in x is however not useful in the deduction of Theorem 1.2, as this truncation requires global information on the tree and therefore poorly commutes with the tree selection algorithm in the proof of sparse domination.

With notation near (1.7), define $\tilde{\Lambda}_{\mathcal{T}}((F_j)_{j \in \mathcal{C}})$ as

$$\sum_{k \in \mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{\mathbb{R}^m} \frac{c(t)}{1 + |u|^{4m}} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} F_j 1_{A_j T_k}(\Pi_j x) \right) (\partial_i^0 \partial_i^1 g)_t(x - Jp + ut) dx dp \frac{dt}{t},$$

where for a tree \mathcal{T} we have written

$$T_k := \cup \{Q \in \mathcal{T} : l(Q) = 2^k\}.$$

Theorem 1.4. *Let $\Delta > 1$ and $\alpha < \frac{1}{m}$. There exists $C > 0$ such that the following holds. Let J, A_j be as in Theorem 1.2. For any convex tree \mathcal{T} , for all c, u, i satisfying conditions specified near (1.6), and all tuples $(F_j)_{j \in \mathcal{C}}$ we have*

$$|\tilde{\Lambda}_{\mathcal{T}}((F_j)_{j \in \mathcal{C}})| \leq C |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}).$$

The following theorem shows sharpness of one assumption in Theorem 1.1:

Theorem 1.5. *The variant of Theorem 1.1, where the threshold 2^{m-1} in (1.3) is replaced by a smaller number, is false.*

The examples used to prove Theorem 1.5 are similar to the ones in [13] and [9]. Theorem 1.5 also implies that similarly the lower bound 2^{m-1} in Theorem 1.2 can not be lowered. We stress that we do not know the sharp range of tuples of exponents $(p_j)_j$ for the conclusion of Theorem 1.1 to hold. It may be possible to lower some of the exponents p_j below 2^{m-1} at the expense of demanding additional constraints on the other exponents.

The results in this paper are continuous analogues of some of the estimates in [22] in the dyadic setting. It would be desirable to have a more complete understanding of the analogs of results in [22] in the continuous setting. In the broader context, there are a number of similar results in the dyadic setting which await a transfer into the continuous setting, such as for example [19], [21], [12].

In Section 2, we reduce Theorem 1.1 to Theorem 1.2. In Section 3, we reduce Theorem 1.2 to Theorem 1.3. In Section 4, we reduce Theorem 1.3 to Theorem 1.4, this section should be compared with [5, Lemma 6 and Lemma 7]. In Section 5, we prove Theorem

1.4 by an induction on the number of symmetries of the form and the tuple $(F_j)_{j \in \mathcal{C}}$ as in [10]. The elaborate induction statement splits the proof into three natural steps, which are done in three separate subsections. In Section 6, we prove Theorem 1.5.

A survey on singular Brascamp-Lieb forms is found in [11]. We like to point the reader also at some more recent developments since the survey such as [2], [17]. Comparing for example the results in [2] with our result, note that there the argument of the kernel K has much larger dimension than the arguments of the functions, while in our result the argument of the kernel has same dimension as the arguments of the functions. Also [2] has more general exponents than Hölder. It might be interesting to study Brascamp-Lieb data in the vicinity of our data that lead to non-Hölder scaling. We have not endeavoured in this direction yet.

Acknowledgements. The first author is supported by NSF DMS-2154356. The second author is supported by the Primus research programme PRIMUS/21/SCI/002 of Charles University. The third author acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2047/1 - 390685813 as well as SFB 1060. Part of the research was carried out during a delightful workshop on Real and Harmonic Analysis at the Oberwolfach Research Institute for Mathematics. The authors thank the anonymous referee for their careful reading of the paper and valuable suggestions.

2. PROOF OF THEOREM 1.1

In this section we reduce Theorem 1.1 to Theorem 1.2. Let $\Delta > 1$. Let $(p_j)_{j \in \mathcal{C}}$ satisfy (1.3) and (1.4). Let K be a Calderón-Zygmund kernel in \mathbb{R}^m of order 2^{6m} and let $(F_j)_{j \in \mathcal{C}}$ be a tuple of smooth compactly supported functions. Let Π be a real $m \times 2m$ matrix such that $\Pi_j \Pi^T$ is Δ -regular for each $j \in \mathcal{C}$.

The Δ -regularity of $\Pi_0 \Pi^T$ and $\Pi_1 \Pi^T$ implies that the left and right $m \times m$ blocks of Π are Δ -regular. Here we denote by 0 and 1 the respective constant functions in \mathcal{C} . Thus, the matrix Π is of the form $((J^0)^{-1}, -(J^1)^{-1})$ for some $m \times m$ Δ -regular matrices J^0, J^1 . We write

$$K(\Pi x) = K(J_0^{-1} J_0 \Pi x) = \tilde{K}(J_0 \Pi x) = \tilde{K}(\Pi^J x)$$

with $\tilde{K} = K \circ J_0^{-1}$ and $J = ((J^0)^T, (J^1)^T)^T$. Then \tilde{K} is up to a constant multiple depending on Δ also Calderón-Zygmund, and we will apply Theorem 1.2 with the kernel \tilde{K} and the matrix J .

We verify that the matrices $A_j = \Pi_j J$ are Δ' regular for all j and a Δ' depending on Δ . The upper bound on the norm of A_j follows from upper bounds on the norm of J , and it then suffices to prove a lower bound on the absolute value of the determinant of A_j . By Δ -regularity of J_1 , it suffices to prove lower bounds on the determinant of

$$B_j := A_j((J^1)^{-1}) = \Pi_j((J^0(J^1)^{-1})^T, I)^T.$$

The matrix B_j arises from $V := J^0(J^1)^{-1}$ by replacing each row with an index i with $j(i) = 1$ by the corresponding i -th row of the identity matrix. The determinant of B_j is therefore equal to the determinant of the submatrix W of V obtained by erasing all rows and columns with index i such that $j(i) = 1$. Similarly, the determinant of $-W$ is equal to the determinant of $\Pi_{\bar{j}}(\Pi^J)^T$, where \bar{j} denotes the corner of \mathcal{C} opposite to j . The latter determinant is however bounded below by the assumed Δ -regularity of $\Pi_{\bar{j}}(\Pi^J)^T$. This proves the desired lower bound on the determinant of A_j .

Theorem 1.2 yields a sparse collection \mathcal{S} of dyadic cubes in \mathbb{R}^m such that the analogue of (1.5) for \bar{K} holds. To deduce Theorem 1.1 from this bound, we first estimate the right-hand side of (1.5) by a sum of averages over cubes in dyadic grids. Here, by a dyadic grid we mean any collection \mathcal{D} of measurable subsets of \mathbb{R}^m such that any $Q, Q' \in \mathcal{D}$ are either disjoint or one is contained in the other. It follows from Definition 2.1 and Theorem 3.1 in [18] that there exist 3^m dyadic grids \mathcal{D}_i , $1 \leq i \leq 3^m$, such that each Q^* associated with a cube $Q \in \mathcal{S}$ belongs to \mathcal{D}_i for some i . Thus,

$$\sum_{Q \in \mathcal{S}} |Q| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*} \leq \sum_{i=1}^{3^m} \sum_{Q^* \in \mathcal{S}_i} |Q^*| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*}, \quad (2.1)$$

where

$$\mathcal{S}_i = \{Q^* \in \mathcal{D}_i : Q \in \mathcal{S}\}.$$

The collections \mathcal{S}_i are $\frac{1}{2}3^{-m}$ -sparse because

$$\sum_{(Q')^* \in \mathcal{S}_i : (Q')^* \subseteq Q^*} |Q'^*| \leq 3^m \sum_{Q' \in \mathcal{S} : (Q')^* \subseteq Q^*} |Q'| \leq 2 \cdot 3^m |Q^*|,$$

where in the last inequality we used that $(Q')^* \subseteq Q^*$ implies that Q' is contained in one of the 3^m cubes of the same size as Q that are contained in Q^* and we applied the assumption of sparsity of \mathcal{S} on each of these cubes. Applying Theorem 1.11 in [23] with $\rho_i := 0$, $r_i := 1$, $t_i := p_i$, and $w_i := 1$ to (2.1) for each $1 \leq i \leq 3^m$ yields Theorem 1.1.

Alternatively, the reduction of Theorem 1.1 to Theorem 1.2 might possibly be accomplished with estimates for sparse operators rather than sparse forms as in [16]. We did not try to pursue this approach but refer to [18] for a survey. In the formalism of outer L^p spaces of [4], the reduction is an estimate of the right-hand side of (1.5) with a suitable outer Hölder inequality by

$$\leq C \|1_{\mathcal{S}}\|_{L^\infty(\ell^1)} \prod_{j \in \mathcal{C}} \|\mathcal{F}_j F_j\|_{L^{p_j}(\ell^\infty)}$$

for suitable embeddings \mathcal{F}_j . Sparsity of \mathcal{S} is tautological to a bound on $\|1_{\mathcal{S}}\|_{L^\infty(\ell^1)}$, and embedding theorems give the estimates $\|\mathcal{F}_j F_j\|_{L^{p_j}(\ell^\infty)} \leq C \|F_j\|_{p_j}$.

3. PROOF OF THEOREM 1.2

We present the proof of Theorem 1.2, using Theorem 1.3 and some lemmas that will be stated and proved at the end of this section. Let Δ , p_j be as in Theorem 1.2. We need to find a constant C such that the sparse bound of Theorem 1.2 holds for all J , A_j as in Theorem 1.2, all tuples $(F_j)_{j \in \mathcal{C}}$ of compactly supported smooth functions and all Calderón-Zygmund kernels K in \mathbb{R}^m . We will find constants for various related inequalities and by abuse of notation will use the letter C for each of them. As a consequence, the meaning of C may change from line to line.

As discussed in the introduction and elaborated in detail in [10], by superposition it suffices to show that the sparse bound holds for the specific class of Calderón-Zygmund kernels associated with forms Λ of the type (1.6). The integral in the (t, p) variables in (1.6) is absolutely integrable, a typical theme in Calderón Zygmund theory that we elaborate in Lemma 3.3. It therefore suffices to show the sparse bound for the form $\Lambda_{\mathcal{Q}'}$ as in (1.7) with arbitrary finite collection \mathcal{Q}' of dyadic cubes. Let thus c, u, i be given as detailed near (1.6) and let \mathcal{Q}' be given. By making \mathcal{Q}' larger if needed, we may assume that \mathcal{Q}' is convex in the sense that if $Q \subseteq Q' \subseteq Q''$ and $Q, Q'' \in \mathcal{Q}'$, then also $Q' \in \mathcal{Q}'$.

We may also assume that each maximal cube Q in \mathcal{Q}' has the property that the support of each of the functions $F_j \circ A_j$ is contained in Q^* . Here and in what follows, maximality of cubes is with respect to the partial order by set inclusion.

We construct a sparse collection. For each dyadic cube Q let $\mathcal{I}(Q)$ be the set of maximal dyadic cubes $Q' \subseteq Q$ satisfying

$$\sup_{R \in \mathcal{R}: A_j Q' \subseteq R} (\varepsilon_R^\alpha)^{2^{1-m}} [F_j 1_{A_j Q^*}]_R > c [F_j \circ A_j]_{Q^*}$$

for some $j \in \mathcal{C}$, where c is a certain positive constant determined a few lines below. For each $Q' \in \mathcal{I}(Q)$, there is $j \in \mathcal{C}$ such that

$$M_\alpha(F_j 1_{A_j Q^*})(A_j x) > c [F_j \circ A_j]_{Q^*}$$

for all $x \in Q'$. With the weak bound for M_α in Lemma 3.1 below, we conclude

$$\begin{aligned} \sum_{Q' \in \mathcal{I}(Q)} |Q'| &\leq C \sum_{j \in \mathcal{C}} |\{x \in \mathbb{R}^m : M_\alpha(F_j 1_{A_j Q^*})(A_j x) > c [F_j \circ A_j]_{Q^*}\}| \\ &\leq C \sum_{j \in \mathcal{C}} (c [F_j \circ A_j]_{Q^*})^{-2^{m-1}} \int_{A_j Q^*} |F_j|^{2^{m-1}} = C c^{-2^{m-1}} \sum_{j \in \mathcal{C}} |A_j Q^*| \leq \frac{1}{2} |Q|, \end{aligned} \quad (3.1)$$

where c is chosen large enough depending on Δ such that the last inequality holds.

Let \mathcal{S}_0 be the collection of maximal dyadic cubes in \mathcal{Q}' . It is a finite collection of disjoint cubes. For $n \geq 1$ define

$$\mathcal{S}_n := \bigcup_{Q \in \mathcal{S}_{n-1}} \mathcal{I}(Q)$$

and let \mathcal{S} be the union of the sets \mathcal{S}_n over all natural numbers n . Using (3.1) and summing a geometric series, we deduce that the collection \mathcal{S} is sparse.

We aim to show

$$|\Lambda_{\mathcal{Q}'}((F_j)_{j \in \mathcal{C}})| \leq C \sum_{Q \in \mathcal{S}} |Q| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*}. \quad (3.2)$$

Given a dyadic cube Q , denote by $\mathcal{T}_{\leq Q}$ the set of all dyadic cubes Q' in \mathcal{Q}' with $Q' \subseteq Q$. We observe

$$\Lambda_{\mathcal{Q}'}((F_j)_{j \in \mathcal{C}}) = \sum_{Q \in \mathcal{S}_0} \Lambda_{\mathcal{T}_{\leq Q}}((F_j)_{j \in \mathcal{C}}) = \sum_{Q \in \mathcal{S}_0} \Lambda_{\mathcal{T}_{\leq Q}}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}}). \quad (3.3)$$

The first equality follows from the definition of \mathcal{S}_0 and $\mathcal{T}_{\leq Q}$. The second equality follows by the assumption on the maximal cubes Q in \mathcal{Q}' .

Denote for any cube $Q \in \mathcal{S}$ by \mathcal{T}_Q the convex tree

$$\mathcal{T}_Q := \mathcal{T}_{\leq Q} \setminus \bigcup_{Q' \in \mathcal{I}(Q)} \mathcal{T}_{\leq Q'}.$$

We have for $Q \in \mathcal{S}$

$$\Lambda_{\mathcal{T}_{\leq Q}}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}}) = \Lambda_{\mathcal{T}_Q}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}}) + \sum_{Q' \in \mathcal{I}(Q)} \Lambda_{\mathcal{T}_{\leq Q'}}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}}). \quad (3.4)$$

By Theorem 1.3 and the definition of $\mathcal{I}(Q)$, the first term on the right-hand side of (3.4) can be estimated as

$$|\Lambda_{\mathcal{T}_Q}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}})| \leq C |Q| \prod_{j \in \mathcal{C}} M_\alpha(F_j 1_{A_j Q^*})(A_j \mathcal{T}_Q) \leq C |Q| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*}.$$

To estimate the second term on the right-hand side of (3.4), fix $Q' \in \mathcal{I}(Q)$ and split

$$F_j 1_{A_j Q^*} = F_j 1_{A_j(Q')^*} + F_j 1_{A_j Q^* \setminus A_j(Q')^*}$$

for each $j \in \mathcal{C}$. Multilinearity of the form $\Lambda_{\mathcal{T}_{\leq Q'}}$ gives

$$\Lambda_{\mathcal{T}_{\leq Q'}}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}}) = \Lambda_{\mathcal{T}_{\leq Q'}}((F_j 1_{A_j(Q')^*})_{j \in \mathcal{C}}) + \sum_{(E_j)_{j \in \mathcal{C}} \in \mathcal{E}} \Lambda_{\mathcal{T}_{\leq Q'}}((F_j E_j)_{j \in \mathcal{C}}),$$

where \mathcal{E} is the set of tuples with entries $1_{A_j(Q')^*}$ or $1_{A_j Q^* \setminus A_j(Q')^*}$ such that at least one j satisfies $E_j = 1_{A_j Q^* \setminus A_j(Q')^*}$. The first term on the right hand side will be estimated by a recursive procedure, while the second term is estimated by Lemma 3.5 and by the choice of $\mathcal{I}(Q)$. We obtain for the second term

$$\sum_{(E_j)_{j \in \mathcal{C}} \in \mathcal{E}} \Lambda_{\mathcal{T}_{\leq Q'}}((F_j E_j)_{j \in \mathcal{C}}) \leq C|Q'| \prod_{j \in \mathcal{C}} M_\alpha(F_j 1_{A_j Q^*})(A_j Q') \leq C|Q'| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*}.$$

Combining (3.3) and (3.4) and using the bounds we have just discussed, we have

$$|\Lambda_{Q'}((F_j)_{j \in \mathcal{C}})| \leq C \sum_{Q \in \mathcal{S}_0} |Q| \prod_{j \in \mathcal{C}} [F_j \circ A_j]_{Q^*} + \sum_{Q \in \mathcal{S}_1} |\Lambda_{\mathcal{T}_{\leq Q}}((F_j 1_{A_j Q^*})_{j \in \mathcal{C}})|.$$

Iterating this argument over $\mathcal{S}_1, \mathcal{S}_2$, and so on, and using that $\mathcal{S}_n \cap Q'$ is empty for large n by finiteness of Q' , we obtain (3.2). This completes the proof of Theorem 1.2.

Lemma 3.1 (Weak type 2^{m-1} bound for M_α). *For $\alpha > 0$ and $m \geq 2$, there is a $C > 0$ such that the following holds. For a measurable function F on \mathbb{R}^m and $\lambda > 0$, we have with M_α as in (1.8),*

$$|\{x : M_\alpha(F)(x) > \lambda\}| \leq C \lambda^{-2^{m-1}} \|F\|_{2^{m-1}}^{2^{m-1}}.$$

Proof. We denote by \mathcal{Z} the set consisting of all $k = (k_1, \dots, k_m) \in \mathbb{Z}^m$ with $k_m = 0$. Given $k \in \mathcal{Z}$, we define \mathcal{R}_k to be the set of all $I_1 \times \dots \times I_m \in \mathcal{R}$ such that for each $1 \leq i \leq m-1$,

$$2^{k_{i-1}} < \frac{|I_i|}{|I_m|} \leq 2^{k_i}.$$

We introduce the maximal function

$$M_k F(x) := \sup_{R \in \mathcal{R}_k : x \in R} [F]_R.$$

We claim the weak type bound

$$|\{x : M_k(F)(x) > \lambda\}| \leq C \lambda^{-2^{m-1}} \|F\|_{2^{m-1}}^{2^{m-1}}.$$

By a change of variables $y = Dx$ if necessary, with diagonal matrix D with entries $2^{k_i - k_m}$, it suffices to show this claim for $k_1 = \dots = k_{m-1} = 0$. As each rectangular box in $\mathcal{R}_{(0, \dots, 0)}$ has eccentricity very close to 1, it is contained in a cubical box of volume no larger than 2^m times the volume of the rectangular box. The claim follows then from the weak type 1 bound for the standard Hardy-Littlewood maximal operator, applied to the function $|F|^{2^{m-1}}$.

Define for $k \in \mathcal{Z}$

$$\varepsilon_k := \frac{\min_{1 \leq i \leq m} 2^{k_i}}{\max_{1 \leq i \leq m} 2^{k_i}}$$

and note that for $R \in \mathcal{R}_k$ we have

$$2^{-1} \varepsilon_k \leq \varepsilon_R \leq 2 \varepsilon_k.$$

We also have

$$\sum_{k \in \mathcal{Z}} \varepsilon_k^\alpha < C,$$

because the exponential decay of the summand in the length of the vector k beats the polynomial growth of the volume of the balls in \mathbb{Z}^{m-1} . We obtain

$$\begin{aligned} |\{y : M_\alpha(F)(y) > \lambda\}| &\leq |\{y : \sup_{k \in \mathcal{Z}} (\varepsilon_k^\alpha)^{2^{1-m}} M_k(F)(y) > C\lambda\}| \\ &\leq \sum_{k \in \mathcal{Z}} |\{y : (\varepsilon_k^\alpha)^{2^{1-m}} M_k(F)(y) > C\lambda\}| \\ &\leq \sum_{k \in \mathcal{Z}} C \varepsilon_k^\alpha \lambda^{-2^{m-1}} \|F\|_{2^{m-1}}^{2^{m-1}} \leq C \lambda^{-2^{m-1}} \|F\|_{2^{m-1}}^{2^{m-1}}. \end{aligned}$$

□

Lemma 3.2 (Brascamp-Lieb inequality for Π_j). *For any measurable functions F_j on \mathbb{R}^m we have*

$$\left| \int_{\mathbb{R}^{2m}} \prod_{j \in \mathcal{C}} F_j(\Pi_j x) dx \right| \leq \prod_{j \in \mathcal{C}} \|F_j\|_{2^{m-1}}.$$

Proof. We induct on $m \geq 1$. By Fubini, we have

$$\int_{\mathbb{R}^2} |F_0(x_1^0) F_1(x_1^1)| dx = \|F_0\|_1 \|F_1\|_1,$$

which gives the desired inequality for $m = 1$. For $m > 1$, we proceed by Cauchy-Schwarz:

$$\begin{aligned} \int_{\mathbb{R}^{2m}} \prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| dx &= \int_{\mathbb{R}^{2m}} \prod_{\sigma=0}^1 \prod_{j(m)=\sigma} |F_j(\Pi_j x)| dx \\ &\leq \int_{\mathbb{R}^2} \prod_{\sigma=0}^1 \left(\int_{\mathbb{R}^{2m-2}} \prod_{j(m)=\sigma} |F_j(\Pi_j x)|^2 dx_1^0 \dots dx_{m-1}^0 dx_1^1 \dots dx_{m-1}^1 \right)^{1/2} dx_m^0 dx_m^1. \end{aligned}$$

Note that the integrals in x_m^0 and x_m^1 split. By induction, we obtain an upper bound by

$$\prod_{\sigma=0}^1 \int_{\mathbb{R}} \prod_{j(m)=\sigma} \|F_j^2\|_{L^{2^{m-2}}(x_1, \dots, x_{m-1})}^{1/2} dx_m \leq C \prod_{\sigma=0}^1 \prod_{j(m)=\sigma} \|F_j\|_{2^{m-1}} = C \prod_{j \in \mathcal{C}} \|F_j\|_{2^{m-1}},$$

where we have used the identity $\|F^2\|_{2^{m-2}} = \|F\|_{2^{m-1}}^2$ and Hölder's inequality in the integral over x_m . □

Lemma 3.3 (Integrability over upper half space). *The integral in the (t, p) variables in (1.6) is absolutely integrable for compactly supported smooth functions F_j .*

Proof. For fixed t , the integral

$$\int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j x) \right) (\partial_i^0 \partial_i^1 g)_t(x - Jp + ut) dx \right| dp \quad (3.5)$$

can be estimated with Fubini and Lemma 3.2 by

$$\leq C \left(\prod_{j \in \mathcal{C}} \|F_j\|_{2^{m-1}} \right) \sup_{x \in \mathbb{R}^{2m}} \int_{\mathbb{R}^m} |(\partial_i^0 \partial_i^1 g)_t(x - Jp + ut)| dp.$$

The supremum in x is bounded by Ct^{-m} , which makes (3.5) absolutely integrable over the half line $t > 1$ with measure dt/t . For $t < 1$, a partial integration on the x integral estimates (3.5) by

$$Ct^2 \int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} |\partial_i^0 \partial_i^1 \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j x) \right)| g_t(x - Jp + ut) dx dp.$$

We split \mathbb{R}^{2m} as the direct sum of the kernel V of Π^J and its complement. We then estimate (3.5) with Fubini by

$$\begin{aligned} & t^2 \int_V \int_{V^\perp} |\partial_i^0 \partial_i^1 \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j(y+z)) \right)| \int_{\mathbb{R}^m} g_t(y+z - Jp + ut) dp dy dz \\ & \leq t^2 \int_V \sup_{y \in V^\perp} |\partial_i^0 \partial_i^1 \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j(y+z)) \right)| \left(\int_{y \in V^\perp} \int_{\mathbb{R}^m} g_t(y+z - Jp + ut) dp dy \right) dz. \end{aligned}$$

The Gaussian integral is bounded by a constant independent of t , and the remaining integral is finite since V^\perp is orthogonal to $J\mathbb{R}^m$. Therefore, the whole last display is bounded by Ct^2 . This can be integrated over $t < 1$ against the measure dt/t . \square

Lemma 3.4 (Gaussian estimate). *With notation as in Theorem 1.3, for every $x \in \mathbb{R}^{2m}$, $t > 0$ and $p, p' \in \mathbb{R}^m$ with $|p' - p| < 2^{m+3}t$, we have*

$$\left| \frac{1}{1 + |u|^{4m}} (\partial_i^0 \partial_i^1 g)_t(x - Jp + ut) \right| \leq Ct^{-2m} \sum_{n \geq 0} 2^{-4nm} \chi(|x - Jp'| \leq 2^n t) \quad (3.6)$$

with the indicator $\chi(\mathcal{A})$ that is 1 if \mathcal{A} is satisfied and 0 otherwise.

Proof. Assume first the case

$$t^{-1}|x - Jp'| \leq 8 + 2^{m+5}|J| \quad (3.7)$$

with the operator norm $|J|$ of J . We estimate the left-hand side of (3.6) by the quantity $t^{-2m} \|\partial_i^0 \partial_i^1 g\|_\infty$, which for suitable C is bounded by the single term on the right-hand side with the smallest index n , for which the right-hand side of (3.7) is less than 2^n .

It remains to consider the case that (3.7) is not satisfied. Pick $n \geq 4$ such that

$$2^{n-1} < t^{-1}|x - Jp'| \leq 2^n.$$

Considering only the n -th term on the right-hand side of (3.6), it suffices to show

$$\left| \frac{1}{1 + |u|^{4m}} \partial_i^0 \partial_i^1 g(t^{-1}(x - Jp) + u) \right| \leq C2^{-4nm}. \quad (3.8)$$

If $|u| > 2^{n-3}$, this follows from a bound for $\|\partial_i^0 \partial_i^1 g\|_\infty$. If $|u| \leq 2^{n-3}$, we have

$$|t^{-1}(x - Jp) + u| \geq |t^{-1}(x - Jp')| - 2^{m+3}|J| - 2^{n-3} \geq 2^{n-3},$$

where we used the assumed bound on $p - p'$ and estimated $|J|$ by the assumed reverse inequality to (3.7). Estimate (3.8) then follows from the Schwartzian bound,

$$|\partial_i^0 \partial_i^1 g(y)| \leq C|y|^{-4m}$$

for $|y| \geq 1$. \square

Lemma 3.5 (Off-diagonal estimate for a tree). *Let $\Delta > 1$. There is a constant $C > 0$ such that for J, A_j as in Theorem 1.2, for any convex tree \mathcal{T} and any tuple $(F_j)_{j \in \mathcal{C}}$ of compactly supported bounded measurable functions on \mathbb{R}^m such that for some $j_0 \in \mathcal{C}$, the support of $F_{j_0} \circ A_{j_0}$ is disjoint from $Q_{\mathcal{T}}^*$, we have*

$$|\Lambda_{\mathcal{T}}((F_j)_{j \in \mathcal{C}})| \leq C|Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} \sup_{Q \in \mathcal{Q}: Q_{\mathcal{T}}^* \subseteq Q^*} [F_j \circ A_j]_{Q^*}.$$

Here \mathcal{Q} denotes the set of all dyadic cubes in \mathbb{R}^m .

Proof. Let $Q' \in \mathcal{T}$. Using Lemmas 3.4 and 3.2 we estimate

$$\begin{aligned} & \left| \int_{l(Q')/2}^{l(Q')} \int_{Q'} \frac{c(t)}{1+|u|^{4m}} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} F_j(\Pi_j x) \right) (\partial_i^0 \partial_i^1 g)_t(x - Jp + ut) dx dp \frac{dt}{t} \right| \quad (3.9) \\ & \leq C|Q'| \sup_{l(Q')/2 \leq t \leq l(Q')} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| \right) t^{-2m} \sum_{n \geq 0} 2^{-4nm} \chi(|x - J\mathbf{c}(Q')| \leq 2^n t) dx \\ & \leq C \sum_{n \geq 0} 2^{-4nm} |Q'|^{-1} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| \chi(|\Pi_j x - A_j \mathbf{c}(Q')| \leq 2^n l(Q')) \right) dx \\ & \leq C \sum_{n \geq 0} 2^{-4nm} |Q'|^{-1} \prod_{j \in \mathcal{C}} \|F_j(y) \chi(|y - A_j \mathbf{c}(Q')| \leq 2^n l(Q'))\|_{L^{2^{m-1}}(y)}. \end{aligned}$$

Here $\mathbf{c}(Q)$ denotes the center of the cube Q .

For the summation indices n with $|A_{j_0}^{-1}|2^n l(Q') < l(Q_{\mathcal{T}})$, we see that the factor with F_{j_0} vanishes because $F_{j_0} \circ A_{j_0}$ is supported outside $Q_{\mathcal{T}}^*$ while $\mathbf{c}(Q')$ is inside $Q_{\mathcal{T}}$. Thus, it remains to sum over n with $|A_{j_0}^{-1}|2^n l(Q') \geq l(Q_{\mathcal{T}})$. We estimate

$$\begin{aligned} & \|F_j(y) \chi(|y - A_j \mathbf{c}(Q')| \leq 2^n l(Q'))\|_{L^{2^{m-1}}(y)} \\ & \leq C 2^{nm} 2^{1-m} |Q'|^{2^{1-m}} \sup_{Q \in \mathcal{Q}: Q_{\mathcal{T}}^* \subseteq Q^*} [F_j \circ A_j]_{Q^*}. \end{aligned}$$

In the last inequality we used that $\mathbf{c}(Q') \in Q_{\mathcal{T}}$ and that averages are over boxes of volume comparable or larger in volume than $Q_{\mathcal{T}}^*$.

We continue the estimate of (3.9) with

$$\begin{aligned} & \leq C \sum_{|A_{j_0}^{-1}|2^n l(Q') \geq l(Q_{\mathcal{T}})} 2^{-4nm} 2^{2nm} |Q'| \prod_{j \in \mathcal{C}} \sup_{Q \in \mathcal{Q}: Q_{\mathcal{T}}^* \subseteq Q^*} [F_j \circ A_j]_{Q^*} \\ & \leq C|Q'| \frac{|Q'|}{|Q_{\mathcal{T}}|} \prod_{j \in \mathcal{C}} \sup_{Q \in \mathcal{Q}: Q_{\mathcal{T}}^* \subseteq Q^*} [F_j \circ A_j]_{Q^*}. \end{aligned}$$

Summing (3.9) over all $Q' \in \mathcal{T}$ and using

$$\sum_{Q' \in \mathcal{T}} |Q'| \left(\frac{|Q'|}{|Q_{\mathcal{T}}|} \right) = \sum_{k \leq 0} 2^{mk} \sum_{Q' \in \mathcal{T}: l(Q') = 2^k l(Q_{\mathcal{T}})} |Q'| \leq \sum_{k \leq 0} 2^{mk} |Q_{\mathcal{T}}| \leq C|Q_{\mathcal{T}}|$$

proves the lemma. \square

4. PROOF OF THEOREM 1.3

We present the proof of Theorem 1.3 using Theorem 1.4, several lemmas from the previous section and one from the end of this section. With notation as in Theorems 1.3 and 1.4, we need to prove

$$|\widetilde{\Lambda}_{\mathcal{T}}((F_j)_{j \in \mathcal{C}}) - \Lambda_{\mathcal{T}}((F_j)_{j \in \mathcal{C}})| \leq C|Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_{\alpha}(F_j)(A_j \mathcal{T}). \quad (4.1)$$

We estimate the left-hand side of (4.1) with Lemma 3.4 by

$$\sum_{n \geq 0} 2^{-4mn} \sum_{k \in \mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{T_k} \int_{S_k^c} \left(\prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| \right) t^{-2m} \chi(|x - Jp| \leq 2^n t) dx dp \frac{dt}{t} \quad (4.2)$$

$$+ \sum_{n \geq 0} 2^{-4mn} \sum_{k \in \mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{T_k^c} \int_{S_k} \left(\prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| \right) t^{-2m} \chi(|x - Jp| \leq 2^n t) dx dp \frac{dt}{t}, \quad (4.3)$$

where S_k is the set of all $x \in \mathbb{R}^{2m}$ such that $\Pi_j x \in A_j T_k$ for all $j \in \mathcal{C}$.

Fix n and k and first estimate the corresponding summand in (4.2) by

$$C2^{-2mk} \int_{T_k} \int_{S_k^c} \left(\prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| \chi(|\Pi_j x - A_j p| \leq 2^n 2^k) \right) dx dp.$$

Let E be the set of $p \in T_k$ such that the inner integral of the last display is not zero. Then we estimate the last display with the Brascamp-Lieb inequality from Lemma 3.2 by

$$\begin{aligned} &\leq C2^{-2mk} \int_E \prod_{j \in \mathcal{C}} \|F_j(y) \chi(|y - A_j p| \leq 2^n 2^k)\|_{L^{2^{m-1}}(y)} dp \\ &\leq C2^{2mn} |E| \prod_{j \in \mathcal{C}} M_{\alpha}(F_j)(A_j \mathcal{T}). \end{aligned} \quad (4.4)$$

Here the last inequality followed by identifying the $L^{2^{m-1}}$ norm of the truncated function F_j as an average over a cube of side length $2^n 2^k$ about the point $A_j p$, which is contained in a set $A_j Q^*$ of comparable volume for a Q containing a cube $Q' \in \mathcal{T}$.

We estimate $|E|$. If $p \in E$, then there is $x \in S_k^c$ such that for all $j \in \mathcal{C}$ we have

$$|\Pi_j x - A_j p| \leq 2^n 2^k.$$

By definition of S_k , there is a $j_0 \in \mathcal{C}$ such that $\Pi_{j_0} x \notin A_{j_0} T_k$. Let Q_p be a dyadic cube of side length 2^k containing p and let Q_x be a dyadic cube of side length 2^k such that $\Pi_{j_0} x \in A_{j_0} Q_x$. Then $Q_p \subseteq T_k$ and $Q_x \not\subseteq T_k$. But both $A_{j_0} Q_p$ and $A_{j_0} Q_x$ are contained in the ball B of radius $C2^n 2^k$ about $A_{j_0} p$ for sufficiently large $C > 1$ depending on Δ . Hence there is

$$y \in \partial T_k \cap (2^k \mathbb{Z})^m$$

such that $y \in A_{j_0}^{-1} B$. But then p is contained in the ball of radius $C2^n 2^k$ about y . Hence

$$|E| \leq C2^{nm} 2^{km} \#(\partial T_k \cap (2^k \mathbb{Z})^m).$$

Using this estimate, (4.4), and summing the geometric series in n , we obtain for (4.2) the upper bound

$$\leq C \left(\sum_{k \in \mathbb{Z}} 2^{km} \#(\partial T_k \cap (2^k \mathbb{Z})^m) \right) \prod_{j \in \mathcal{C}} M_{\alpha}(F_j)(A_j \mathcal{T}).$$

With Lemma 4.1 we obtain the desired bound for (4.2).

It remains to estimate (4.3). Fix again n and k and estimate the corresponding summand by

$$C2^{-2mk} \int_{T_k^c} \int_{S_k} \left(\prod_{j \in \mathcal{C}} |F_j(\Pi_j x)| \chi(|\Pi_j x - A_j p| \leq 2^n 2^k) \right) dx dp.$$

Let E be the set of $p \in T_k^c$ such that the inner integral of the last display is not zero. Then we estimate the last display with the Brascamp-Lieb inequality from Lemma 3.2 by

$$\leq C2^{-2mk} \int_E \prod_{j \in \mathcal{C}} \|F_j(y) \chi(|y - A_j p| \leq 2^n 2^k)\|_{L^{2^{m-1}}(y)} dp. \quad (4.5)$$

If $p \in E$, then there is $x \in S_k$ such that for all $j \in \mathcal{C}$ we have

$$|\Pi_j x - A_j p| \leq 2^n 2^k.$$

By definition of S_k , for every $j \in \mathcal{C}$ there is a $p_j \in T_k$ such that $\Pi_j x = A_j p_j$. Using the triangle inequality, we may estimate (4.5) by

$$\begin{aligned} &\leq C2^{-2mk} \int_E \prod_{j \in \mathcal{C}} \|F_j(y) \chi(|y - A_j p_j| \leq 2^{n+1} 2^k)\|_{L^{2^{m-1}}(y)} dp \\ &\leq C2^{2mn} |E| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}). \end{aligned}$$

Here we may argue in the last inequality as for (4.4), because $p_j \in T_k$. We may also argue as before that the ball of radius $C2^{n+1}2^k$ about p contains p_j and we may argue as before that it also contains a point in $\partial T_k \cap (2^k \mathbb{Z})^m$. We may estimate as before

$$|E| \leq C2^{nm} 2^{km} \#(\partial T_k \cap (2^k \mathbb{Z})^m)$$

and conclude the desired bound for (4.3). This concludes the proof of Theorem 1.3.

Lemma 4.1. *There is a constant $C > 0$ such that for any convex tree \mathcal{T} in \mathbb{R}^m ,*

$$\sum_{k \in \mathbb{Z}} 2^{mk} \#(\partial T_k \cap (2^k \mathbb{Z})^m) \leq C|Q_{\mathcal{T}}|.$$

Proof. Let \mathcal{Q} be the collection of maximal cubes $Q \subseteq Q_{\mathcal{T}}^*$ such that $Q \notin \mathcal{T}$. As these cubes are pairwise disjoint and contained in $Q_{\mathcal{T}}^*$, we have

$$\sum_{Q \in \mathcal{Q}} |Q| \leq C|Q_{\mathcal{T}}|.$$

For each $Q \in \mathcal{Q}$ let \mathcal{Q}_Q be the set of all cubes $Q' \subseteq Q$ such that the boundary of Q' intersects the boundary of Q . With k such that $l(Q) = 2^k$, we have

$$\sum_{Q' \in \mathcal{Q}_Q} |Q'| \leq \sum_{k' \leq k} \sum_{Q' \in \mathcal{Q}_Q: l(Q')=2^{k'}} |Q'| \leq C \sum_{k' \leq k} 2^{k'-k} |Q| \leq C|Q|,$$

where we counted the cubes Q' of fixed side length by using that they are contained in a thin layer around the boundary of the cube Q .

For each point $p \in \partial T_k \cap (2^k \mathbb{Z})^m$, there is a cube $Q_{p,k}$ of side length 2^k which is adjacent to p and element of \mathcal{Q}_Q for some $Q \in \mathcal{Q}$. Conversely, each such cube $Q_{p,k}$ is equal to

some other cube $Q_{p',k'}$ only for $k' = k$ and boundedly many points $p' \in \partial T_k \cap (2^k \mathbb{Z})^m$. We obtain

$$\sum_{k \in \mathbb{Z}} 2^{mk} \#(\partial T_k \cap (2^k \mathbb{Z})^m) \leq C \sum_{Q \in \mathcal{Q}} \sum_{Q' \in \mathcal{Q}_Q} |Q'| \leq C \sum_{Q \in \mathcal{Q}} |Q| \leq C |Q_{\mathcal{T}}|.$$

□

5. PROOF OF THEOREM 1.4

The cube \mathcal{C} has natural reflection symmetries. We formulate in Proposition 5.1 a statement, which can be proven by downward induction on a parameter l counting the reflection symmetries that preserve the value of $\Lambda((F_j)_{j \in \mathcal{C}})$ for a given form Λ and given functions F_j . This proposition, applied to the case of no symmetry assumption, implies Theorem 1.4.

The induction consists of three by now well established key arguments. Symmetrization of the form is achieved by the Cauchy-Schwarz inequality as from (5.36) to (5.37). One breaks the cube into two halves along a reflection symmetry, and Cauchy-Schwarz symmetrizes the two pieces.

It is important that the precious cancellation of the Calderón-Zygmund kernel K is preserved through this Cauchy-Schwarz by being moved to bump functions associated with the two faces that are not cut by the symmetry plane. This is second crucial ingredient, often done by partial integration or algebra, moving the cancellation around the faces. Here, due to the general projection Π , one has to work with cone decompositions (5.13) in addition to algebra (5.23) and these decompositions require careful control of the geometry and dependence of various parameters, done mostly in Section 5.1.

The third ingredient, to avoid that pieces already symmetrized will reappear in later steps, is a positivity argument. All the previously and the newly symmetrized pieces appear as positive terms (5.42) of a sum and it suffices to estimate the sum rather than the individual pieces.

Define for a convex tree \mathcal{T} , for an injection $J : \mathbb{R}^m \rightarrow \mathbb{R}^{2m}$, and for a function

$$K : (0, \infty) \times \mathbb{R}^{2m} \rightarrow \mathbb{C}$$

that is measurable in the first variable $t \in (0, \infty)$ such that $x \rightarrow t^m K(t, tx)$ is Schwartz with integral zero and Schwartz norm bounds uniform in t , the form

$$\Lambda(\mathcal{T}, J, K)((F_j)_{j \in \mathcal{C}}) = \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} (F_j 1_{(A_j T_k)})(\Pi_j x) \right) \int_{2^{k-1}}^{2^k} K(t, x) \frac{dt}{t} dx. \quad (5.1)$$

If $1 \leq i \leq m$ and $j \in \mathcal{C}$, we define the reflected corner $i * j \in \mathcal{C}$ by determining that $(i * j)(k)$ coincides with $j(k)$ at all positions k except at the position $k = i$, where it differs. Define for a function $\varphi : \mathbb{R}^{2m} \rightarrow \mathbb{C}$ and a regular $2m \times 2m$ matrix D

$$\varphi_D(x) = (\det D)^{-1} \varphi(D^{-1}x).$$

We write J_h^σ for the h -th row of the matrix J^σ .

Proposition 5.1. *Let $\Delta > 1$ and $\alpha < \frac{1}{m}$. There exists $C > 0$ such that the following holds. Let $0 \leq l \leq m$ be an integer. Let J, A_j be as in Theorem 1.2 and write $\Pi := \Pi^J$. Assume that for each $1 \leq h \leq l$ we have $J_h^0 = J_h^1$. Let \mathcal{T} be any convex tree. Let D be a*

$2m \times 2m$ diagonal matrix whose diagonal entries d_h^σ are positive, satisfy $d_h^0 = d_h^1 =: d_h$, and, in case $l < m$, for all $1 \leq h \leq l < i \leq m$

$$d_h \leq d_{l+1} = d_i. \quad (5.2)$$

Then, for every tuple $(F_j)_{j \in \mathcal{C}}$ satisfying $F_j = F_{i_*j}$ for all $j \in \mathcal{C}$ and $1 \leq i \leq l$,

$$|\Lambda(\mathcal{T}, J, K)((F_j)_{j \in \mathcal{C}})| \leq C \left(\prod_{i=1}^m \max\{d_i^{2\alpha}, d_i^{-2}\} \right) |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}) \quad (5.3)$$

holds for the following three types of kernels K . If $l > 0$, then (5.3) holds for the kernel

$$K(t, x) = \frac{c(t)}{1 + |u|^{4m}} \int_{\mathbb{R}^m} (\partial_i^0 \partial_i^1 g)_{tD}(x + Jp + tDu) dp \quad (5.4)$$

with c, u, i as in (1.7). If $l > 0$, then (5.3) holds for the kernel

$$K(t, x) = \int_{\mathbb{R}^m} |tD\Pi^T \tilde{\Theta} \xi|^2 g(tD\Pi^T \xi) e^{2\pi i \xi^T \Pi x} d\xi, \quad (5.5)$$

and for the kernel

$$K(t, x) = \sum_{i=1}^l \int_{\mathbb{R}^m} (\partial_i^0 \partial_i^1 g)_{tD}(x + Jp) dp. \quad (5.6)$$

Here $\tilde{\Theta}$ is the diagonal $m \times m$ matrix with i -th diagonal entry $\tilde{\Theta}_i = 0$ if $i \leq l$ and $\tilde{\Theta}_i = 1$ if $i > l$.

The bound (5.3) for (5.4) in case $l = 0$ and D the identity matrix gives Theorem 1.4. Note a sign change in front of Jp , which is harmless as we integrate over \mathbb{R}^m . The bound (5.3) for (5.4) in case $l = m$ is void as there is no index $l < i \leq m$, and hence the statement is vacuously true for $l = m$.

In the next three sections, we will show that bounds for (5.5) for some l follow from bounds for (5.4) for the same l , that bounds for (5.6) for some l follow from bounds for (5.5) for the same l , and that bounds for (5.4) for some l follow from bounds (5.6) for $l + 1$. By induction, this concludes the proof of Theorem 1.4. As the induction consists of finitely many steps, it is permissible that the constants C which may depend on Δ , m , α but not on D, \mathcal{T}, c, u, i may change each inductive step between hypothesis and conclusion.

5.1. Expansion: Bounds for the first kernel imply bounds for the second kernel. Let $J, \alpha, 0 < l \leq m, D$ be given, set $\Pi = \Pi^J$, and let K be the kernel (5.5). If $l = m$, then $\tilde{\Theta} = 0$ and the bounds (5.3) are trivial. Thus, we assume $l < m$. Let $d := d_{l+1}$ denote the maximal entry of D . Let $\Theta : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ be the diagonal matrix consisting of two copies of $\tilde{\Theta}$ on the diagonal.

Lemma 5.2. *For each $j \in \mathcal{C}$, the map Π^T maps $(1 - \tilde{\Theta})\mathbb{R}^m$ to $(1 - \Theta)\mathbb{R}^{2m}$ and the map $\Pi_j \Pi^T$ maps $(1 - \tilde{\Theta})\mathbb{R}^m$ onto itself. Finally, we have $\Theta \Pi^T \tilde{\Theta} = \Theta \Pi^T$.*

Proof. By assumption we have $(1 - \tilde{\Theta})J^0 = (1 - \tilde{\Theta})J^1$ and hence

$$(1 - \tilde{\Theta})(-J^0(J^1)^{-1}) = \tilde{\Theta} - 1.$$

We obtain

$$(1 - \tilde{\Theta})\Pi = (1 - \tilde{\Theta})(I, -J^0(J^1)^{-1}) = (1 - \tilde{\Theta})(I, -I).$$

This proves the first statement by transposition. Moreover,

$$(1 - \tilde{\Theta})\Pi\Pi_j^T = (1 - \tilde{\Theta})E_j \quad (5.7)$$

for the diagonal matrix E_j with h -th diagonal entry $(-1)^{j(h)}$. Hence $\Pi_j\Pi^T$ maps $(1 - \tilde{\Theta})\mathbb{R}^m$ into itself. As $\Pi_j\Pi^T$ is regular, it maps $(1 - \tilde{\Theta})\mathbb{R}^m$ onto itself. This proves the second statement. The last statement follows from

$$\Theta\Pi^T(1 - \tilde{\Theta}) = 0,$$

which is clear from the first statement. \square

Lemma 5.3. *There is a constant C depending on Δ such that for all $j \in \mathcal{C}$ and all $\xi \in \mathbb{R}^m$ we have*

$$|\Pi^T\tilde{\Theta}\xi| \leq C|\tilde{\Theta}\Pi_j\Pi^T\tilde{\Theta}\xi| \leq C|\Theta\Pi^T\tilde{\Theta}\xi|. \quad (5.8)$$

Proof. The second inequality follows from $\tilde{\Theta}\Pi_j = \Pi_j\Theta$ and the fact that Π_j is a projection. To prove the first inequality, by homogeneity, we may assume $|\tilde{\Theta}\xi| = 1$. As $\Pi_j\Pi^T$ maps $(1 - \tilde{\Theta})\mathbb{R}^m$ onto itself, there is an $\eta \in (1 - \tilde{\Theta})\mathbb{R}^m$ such that

$$\tilde{\Theta}\Pi_j\Pi^T\tilde{\Theta}\xi = \Pi_j\Pi^T(\tilde{\Theta}\xi + \eta).$$

It follows by Δ -regularity of $\Pi_j\Pi^T$ that with C depending on Δ

$$|\tilde{\Theta}\Pi_j\Pi^T\tilde{\Theta}\xi| \geq C|\tilde{\Theta}\xi + \eta| \geq C|\tilde{\Theta}\xi|. \quad (5.9)$$

The first inequality in (5.8) follows by Δ -regularity of both blocks of Π^T . \square

Define $W = D\Pi^T\tilde{\Theta}\mathbb{R}^m$.

Lemma 5.4. *There exists $\delta > 0$ depending on Δ such that for each $\eta \in W$ with $1 \leq |\eta|$ there is an $l < i \leq m$ such that for all $\zeta \in W$ with $|\zeta - \eta| < \delta$ we have $\delta \leq |\zeta_i^0|$ and $\delta \leq |\zeta_i^1|$.*

Proof. With C from Lemma 5.3, we set $\delta = (2Cm)^{-1}$. We then have for any ξ and j

$$|D\Pi^T\tilde{\Theta}\xi| \leq d|\Pi^T\tilde{\Theta}\xi| \leq Cd|\tilde{\Theta}\Pi_j\Pi^T\tilde{\Theta}\xi| = C|\tilde{\Theta}\Pi_jD\Pi^T\tilde{\Theta}\xi|, \quad (5.10)$$

where we used $\tilde{\Theta}\Pi_j = \Pi_j\Theta$ and $\Theta D = d\Theta$ from (5.2). Now fix $\eta \in W$ with $1 \leq |\eta|$. Then there is $\xi \in \mathbb{R}^m$ with $\eta = D\Pi^T\tilde{\Theta}\xi$. Pick j so that for $l < i \leq m$ we have

$$|\eta_i^{j(i)}| = \min(|\eta_i^0|, |\eta_i^1|).$$

Then, with (5.10),

$$1 \leq |\eta| \leq C|\tilde{\Theta}\Pi_j\eta| \leq Cm \max_{l < i \leq m} \min(|\eta_i^0|, |\eta_i^1|).$$

Pick $l < i \leq m$ so that the maximum is attained in the last display. For each $\zeta \in W$ with $|\zeta - \eta| < \delta$ we have

$$(Cm)^{-1} \leq \min(|\eta_i^0|, |\eta_i^1|) \leq \delta + \min(|\zeta_i^0|, |\zeta_i^1|).$$

Therefore, with the definition of δ ,

$$\delta \leq \min(|\zeta_i^0|, |\zeta_i^1|).$$

\square

We need a partition of unity of $W \setminus \{0\}$, which we find as follows. Consider the shell

$$S := \{\eta \in W : 1 \leq |\eta| \leq 4\}.$$

Let Γ be a maximal set of $\delta/4$ -separated vectors in S . The cardinality of Γ depends on Δ . For each $\gamma \in \Gamma$, let B_γ be the open ball in W of radius δ about γ , and let ρ_γ be a smooth function supported on the closure of B_γ and strictly positive on B_γ . The collection of balls B_γ covers S , while the collection of all $2^s S$ with $s \in \mathbb{Z}$ covers $W \setminus \{0\}$. Hence

$$\psi_\gamma(\eta) := \rho_\gamma(\eta) \left(\sum_{s \in \mathbb{Z}} \sum_{\gamma \in \Gamma} \rho_\gamma(2^s \eta) \right)^{-1}$$

is well defined, supported on the closure of B_γ , and smooth for each $\gamma \in \Gamma$. We have

$$\sum_{s \in \mathbb{Z}} \sum_{\gamma \in \Gamma} \psi_\gamma(2^s \eta) = 1$$

for all $\eta \in W \setminus \{0\}$. For each $\gamma \in \Gamma$, by construction of the ball B_γ and Lemma 5.4, there is an $i = i_\gamma$ with $l < i \leq m$ and a smooth compactly supported function q_γ with support and derivative bounds independent of D , so that for η in B_γ we have

$$|\eta|^2 = \eta_i^0 \eta_i^1 q_\gamma(\eta). \quad (5.11)$$

We fix such a choice of i_γ and q_γ .

We further need an annular partition of unity of $\mathbb{R}^{2m} \setminus \{0\}$. Let ρ be a non-negative smooth function on \mathbb{R}^{2m} supported on the closed annulus $1 \leq |\tau| \leq 4$ and positive on the open annulus and define for $r > 0$

$$\phi_r(\tau) = \rho(2^{-r} \tau) \left(\sum_{k \in \mathbb{Z}} \rho(2^{-k} \tau) \right)^{-1}$$

and

$$\phi_0(\tau) = \left(\sum_{r \leq 0} \rho(2^{-r} \tau) \right) \left(\sum_{k \in \mathbb{Z}} \rho(2^{-k} \tau) \right)^{-1}.$$

Then we have for all $\tau \in \mathbb{R}^{2m} \setminus \{0\}$,

$$\sum_{r \geq 0} \phi_r(\tau) = 1.$$

We now introduce a linear map L that will be needed in the subsequent computations. With the definition of W and Lemma 5.2 we have

$$\Theta W = \Theta D \Pi^T \tilde{\Theta} \mathbb{R}^m = \Theta D \Pi^T \mathbb{R}^m.$$

Recall from (5.9) that $\tilde{\Theta} \Pi_j \Pi^T$ is injective on $\tilde{\Theta} \mathbb{R}^m$ for each j . Hence $\Theta \Pi^T$ is injective on $\tilde{\Theta} \mathbb{R}^m$, and so is $\Theta D \Pi^T$. Therefore, the dimension of

$$\Theta W = \Theta D \Pi^T \tilde{\Theta} \mathbb{R}^m$$

is $m - l$, the same as W . Hence there is a regular linear map $L : \Theta W \rightarrow W$ such that for all $\xi \in \mathbb{R}^m$

$$D \Pi^T \tilde{\Theta} \xi = L \Theta D \Pi^T \xi.$$

We note that ΘL is the identity on ΘW , because for all $\xi \in \mathbb{R}^m$, with Lemma 5.2,

$$\Theta L \Theta D \Pi^T \xi = \Theta D \Pi^T \tilde{\Theta} \xi = \Theta D \Pi^T \xi. \quad (5.12)$$

Observe for $\xi \in \mathbb{R}^m$

$$|L \Theta D \Pi^T \xi| = |D \Pi^T \tilde{\Theta} \xi| \leq d |\Pi^T \tilde{\Theta} \xi| \leq C d |\Theta \Pi^T \tilde{\Theta} \xi| = C |\Theta D \Pi^T \xi|,$$

where we used Lemma 5.2, Lemma 5.3, and (5.2). Hence $\|L\| \leq C$ with a constant depending on Δ but not on D .

Using L , we write for the kernel in (5.5)

$$\begin{aligned} K(t, x) &= \int_{\mathbb{R}^m} |tL\Theta D\Pi^T \xi|^2 g(tD\Pi^T \xi) e^{2\pi i \xi^T \Pi x} d\xi \\ &= C \int_V |tL\Theta D\eta|^2 g(tD\eta) e^{2\pi i \eta^T x} d\eta, \end{aligned}$$

with normalized Lebesgue measure on the space $V = \Pi^T \mathbb{R}^m$ and a constant C that is controlled by Δ .

We use the above partitions of unity to decompose

$$\Lambda(\mathcal{T}, \Pi, K) = \sum_{r \geq 0} \sum_{s \in \mathbb{Z}} \sum_{\gamma \in \Gamma} \Lambda(\mathcal{T}, \Pi, K_{r,s,\gamma}) \quad (5.13)$$

with

$$K_{r,s,\gamma}(t, x) := \int_V |tL\Theta D\eta|^2 g(tD\eta) \phi_r(tD\eta) \psi_\gamma(2^{s-r} tL\Theta D\eta) e^{2\pi i \eta^T x} d\eta. \quad (5.14)$$

As L is bounded, there is an $s_0 \in \mathbb{Z}$ depending on Δ so that for $s < s_0$ we have

$$|2^{s-r} tL\Theta D\eta| \leq 4^{-1} |2^{-r} tD\eta|.$$

If $\phi_r(tD\eta)$ is non-zero, then the right-hand side of the last display is less than 1. But then $\psi_\gamma(2^{s-r} tL\Theta D\eta)$ is zero. Hence we can restrict the sum in (5.13) to $s \geq s_0$.

We estimate each $\Lambda(\mathcal{T}, \Pi, K_{r,s,\gamma})$ separately and then sum the estimates by the triangle inequality. Let r, s, γ be arbitrary as in the index set of this sum and $s \geq s_0$.

We apply (5.11) with $\xi = 2^{s-r} tL\Theta D\eta$. Using $\Theta L\Theta = \Theta$ on DV due to (5.12) and homogeneity of the quadratic and linear factors, we obtain for (5.14)

$$\int_V (tD\eta)_i^0 (tD\eta)_i^1 q_\gamma(2^{s-r} tL\Theta D\eta) g(tD\eta) \phi_r(tD\eta) \psi_\gamma(2^{s-r} tL\Theta D\eta) e^{2\pi i \eta^T x} d\eta. \quad (5.15)$$

Define E to be the $2m \times 2m$ diagonal matrix

$$E = 2^{s-r} \Theta D + 2^{-r} (1 - \Theta) D.$$

Define $Y := E\Pi^T \mathbb{R}^m$. Define $c_{r,s,\gamma}(u)$ for $u \in Y$ to be

$$(1 + |u|^{8m}) \int_Y g(\zeta)^{-1} g(DE^{-1}\zeta) \phi_r(DE^{-1}\zeta) \psi_\gamma(L\Theta\zeta) q_\gamma(L\Theta\zeta) e^{-2\pi i \zeta^T u} d\zeta, \quad (5.16)$$

where the integral is with respect to normalized Lebesgue measure on Y .

By the Fourier inversion formula, we have for $\zeta \in Y$

$$g(\zeta)^{-1} g(DE^{-1}\zeta) \phi_r(DE^{-1}\zeta) \psi_\gamma(L\Theta\zeta) q_\gamma(L\Theta\zeta) = \int_Y \frac{c_{r,s,\gamma}(u)}{1 + |u|^{8m}} e^{2\pi i u^T \zeta} du$$

and note that $c_{r,s,\gamma}(u)$ is Schwartz in u . Applying this with $\zeta = tE\eta$ gives for (5.15)

$$\int_Y \frac{2^{2r-2s} c_{r,s,\gamma}(u)}{1 + |u|^{8m}} \int_V e^{2\pi i \eta^T x} (tE\eta)_i^0 (tE\eta)_i^1 g(tE\eta) e^{2\pi i u^T tE\eta} d\eta du.$$

We identify the inner integral as an integral over the range of Π^T of the Fourier transform of F given by

$$F(p) = (\partial_i^0 \partial_i^1 g)_{tE}(x + p + tEu),$$

as then

$$\widehat{F}(\eta) = 4\pi^2 (tE\eta)_i^0 (tE\eta)_i^1 g(tE\eta) e^{2\pi i \eta^T tEu} e^{2\pi i \eta^T x}.$$

We may replace the inner integral by an integral of F itself over the orthogonal complement of the range of Π^T . Thus we obtain with a constant C for (5.15)

$$\int_Y \frac{2^{2r-2s} c_{r,s,\gamma}(u)}{1 + |u|^{8m}} \int_{\mathbb{R}^m} (\partial_i^0 \partial_i^1 g)_{tE}(x + Jp + tEu) dp du. \quad (5.17)$$

Using the assumed bound (5.3) for (5.4) and integrability of $(1 + |u|^{4m})^{-1}$ over Y , we obtain the bound

$$\begin{aligned} & |\Lambda(\mathcal{T}, \Pi, K_{r,s,\gamma})((F_j)_{j \in \mathcal{C}})| \quad (5.18) \\ & \leq C 2^{2r-2s} \|c_{r,s,\gamma}\|_\infty \left(\prod_{i=1}^m \max\{e_i^{2\alpha}, e_i^{-2}\} \right) |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}) \\ & \leq C 2^{(2+2m)r + (2m\alpha - 2)s} \|c_{r,s,\gamma}\|_\infty \left(\prod_{i=1}^m \max\{d_i^{2\alpha}, d_i^{-2}\} \right) |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}). \end{aligned}$$

Here we used that $r \geq 0$ and $s \geq s_0$ and waived the potential gain from any powers of 2 by negative multiples of r or s .

To prove the desired bound (5.3) for (5.5), it remains to show

$$\sum_{r \geq 0} \sum_{s \geq s_0} \sum_{\gamma \in \Gamma} 2^{(2+2m)r + (2m\alpha - 2)s} \|c_{r,s,\gamma}\|_\infty \leq C. \quad (5.19)$$

To see (5.19), it suffices to show for each r, s, γ

$$\|c_{r,s,\gamma}\|_\infty \leq C 2^{-2r}.$$

Decompose ζ in Y into $\Theta\zeta$ and $(1 - \Theta)\zeta$ and note

$$DE^{-1}\zeta = 2^{r-s}\Theta\zeta + 2^r(1 - \Theta)\zeta.$$

For $\psi_\gamma(L\Theta\zeta)$ to be non-zero, we need $\Theta\zeta$ to be bounded by some constant depending on Δ . For $\phi_r(DE^{-1}\zeta)$ to be non-zero, we need $(1 - \Theta)\zeta$ to be bounded by C . Hence the integrand in (5.16) is supported in a ball of radius C . It will then suffice to show that the partial derivatives of the function

$$\zeta \mapsto g(\zeta)^{-1} g(DE^{-1}\zeta) \phi_r(DE^{-1}\zeta) \psi_\gamma(L\Theta\zeta) q_\gamma(L\Theta\zeta)$$

satisfy a bound $C 2^{-2r}$ with constant depending on Δ and the order of derivative. The smallness ultimately comes from the smallness of the Gaussian $g(DE^{-1}\zeta)$ on the support of the function ϕ_r . More precisely, as all factors are smooth, it suffices to assume $r > 2$. The first factor and last two factors do not depend on r and are compactly supported and smooth, it therefore suffices to show the analogous function value and derivative bound on

$$\zeta \mapsto g(DE^{-1}\zeta) \phi_r(DE^{-1}\zeta).$$

However,

$$\phi_r(DE^{-1}\zeta)$$

is nonzero only on an annulus of radius 2^r , where the Gaussian is very small. The desired bound then follows from elementary calculus. This completes the reduction of bounds for (5.5) to bounds for (5.4).

We conclude this section with remarks on the cases $J = (I, I)^T$, i.e., $\Pi = (I, -I)$, or $m = 2$, in which the above proof can be simplified. If $\Pi = (I, -I)$ then we have

$$|D\Pi^T\tilde{\Theta}\xi|^2 = 2 \sum_{l < i \leq m} d_i^2 \xi_i^2,$$

while

$$\sum_{l < i \leq m} (D\Pi^T\xi)_i^0 (D\Pi^T\xi)_i^1 = - \sum_{i=l+1}^m d_i^2 \xi_i^2.$$

Using the expression for the kernel (5.4) on the frequency side as between (5.15) and (5.17), the kernel (5.5) is trivially a linear combination of kernels (5.4) with $u = 0$ and $c(t) = 1$. Similarly, if $l = m - 1$, which is the only case that needs to be discussed if $m = 2$, then both $|D\Pi^T\tilde{\Theta}\xi|^2$ and $(D\Pi^T\xi)_m^0 (D\Pi^T\xi)_m^1$ are non-zero multiples of $d_m^2 \xi_m^2$, which simplifies the proof.

However, in general,

$$|D\Pi^T\tilde{\Theta}\xi|^2$$

is not a linear combination of terms of the form $(D\Pi^T\xi)_i^0 (D\Pi^T\xi)_i^1$, where $l < i \leq m$. To see this, take $m = 3$, $l = 1$, $D = I$ and $\Pi = (I, A)$ with

$$A = \begin{pmatrix} -1 & 0 & 0 \\ a & -1 & b \\ c & d & -1 \end{pmatrix}.$$

Then

$$\begin{aligned} |D\Pi^T\tilde{\Theta}\xi|^2 &= |(I, A)^T(0, \xi_2, \xi_3)|^2 = \xi_2^2 + \xi_3^2 + (a\xi_2 + c\xi_3)^2 + (-\xi_2 + d\xi_3)^2 + (b\xi_2 - \xi_3)^2 \\ &= (2 + a^2 + b^2)\xi_2^2 + (2 + c^2 + d^2)\xi_3^2 + 2(ac - b - d)\xi_2\xi_3. \end{aligned}$$

Further,

$$(\Pi^T\xi)_2^0 (\Pi^T\xi)_2^1 = \xi_2(A^T\xi)_2 = -\xi_2^2 + d\xi_2\xi_3$$

and

$$(\Pi^T\xi)_3^0 (\Pi^T\xi)_3^1 = \xi_3(A^T\xi)_3 = -\xi_3^2 + b\xi_2\xi_3.$$

Thus, comparing the coefficients at ξ_2^2 and ξ_3^2 , we see that in order to have

$$|D\Pi^T\tilde{\Theta}\xi|^2 = \alpha(\Pi^T\xi)_2^0 (\Pi^T\xi)_2^1 + \beta(\Pi^T\xi)_3^0 (\Pi^T\xi)_3^1 \quad (5.20)$$

one necessarily needs $\alpha = -(2 + a^2 + b^2)$ and $\beta = -(2 + c^2 + d^2)$. With this choice of α, β , equation (5.20) is fulfilled if and only if the corresponding coefficients at $\xi_2\xi_3$ on both sides of (5.20) coincide, which amounts to

$$a^2d + b^2d + bc^2 + bd^2 + 2ac = 0.$$

The last equality is not satisfied in general.

5.2. Telescoping: Bounds for the second kernel imply bounds for the third kernel. Let $\Delta, J, A_j, \alpha, l, D, F_j$ be as in Proposition 5.1. Assume $0 < l$.

Let K be the kernel (5.6), it is a translation invariant integral of the function

$$y \mapsto \sum_{1 \leq i \leq l} (\partial_i^0 \partial_i^1 g)_{tD}(x + y)$$

over the kernel of Π . This is equal to a translation invariant integral of the Fourier transform of the function over the range Π^T . Therefore, with a constant C controlled by Δ ,

$$\begin{aligned} K(t, x) &= C \sum_{1 \leq i \leq l} \int_{\mathbb{R}^m} (tD\Pi^T \xi)_i^0 (tD\Pi^T \xi)_i^1 g(tD\Pi^T \xi) e^{2\pi i x^T \Pi^T \xi} d\xi \quad (5.21) \\ &= C \int_{\mathbb{R}^m} \langle (1 - \tilde{\Theta})\Pi_0(tD\Pi^T \xi), (1 - \tilde{\Theta})\Pi_1(tD\Pi^T \xi) \rangle g(tD\Pi^T \xi) e^{2\pi i x^T \Pi^T \xi} d\xi. \end{aligned}$$

Note that for $\eta \in \mathbb{R}^{2m}$ we have

$$|\eta^2| + 2 \langle (1 - \tilde{\Theta})\Pi_0 \eta, (1 - \tilde{\Theta})\Pi_1 \eta \rangle \quad (5.22)$$

$$= |\tilde{\Theta}\Pi_0 \eta|^2 + |\tilde{\Theta}\Pi_1 \eta|^2 + |(1 - \tilde{\Theta})(\Pi_0 + \Pi_1)\eta|^2. \quad (5.23)$$

The expression (5.23) takes the same value for the vectors $\eta = D\Pi^T \xi$ and $\eta = D\Pi^T \tilde{\Theta} \xi$ with $\xi \in \mathbb{R}^m$. This follows individually for the first two terms in (5.23) from Lemma 5.2 and for the third term from (5.7) in the form

$$(\Pi_0 + \Pi_1)\Pi^T(1 - \tilde{\Theta}) = 0.$$

Hence the expression (5.22) also takes the same value for these vectors and we obtain

$$|D\Pi^T \xi^2| + 2 \langle (1 - \tilde{\Theta})\Pi_0 D\Pi^T \xi, (1 - \tilde{\Theta})\Pi_1 D\Pi^T \xi \rangle = |D\Pi^T \tilde{\Theta} \xi|^2,$$

where we have dropped the inner product term on the right hand side because it vanishes because $\Pi_0 \Pi^T$ is the identity and $(1 - \tilde{\Theta})\Pi_0 \Pi^T \tilde{\Theta} = 0$. We remark that the vanishing of this last term is due to our choice of normalization of $\Pi = \Pi^J$ with $\Pi^J \circ J = 0$. If one worked with other choices, one would use the full quadratic form (5.22) for $\eta = D\Pi^T \tilde{\Theta} \xi$ in the definition of the kernel (5.5).

Then the kernel $2K(t, x)$ is a difference of

$$\tilde{K}(t, x) := C \int_{\mathbb{R}^m} |tD\Pi^T \tilde{\Theta} \xi|^2 g(tD\Pi^T \xi) e^{2\pi i x^T \Pi^T \xi} d\xi$$

and

$$\tilde{K}(t, x) := C \int_{\mathbb{R}^m} |tD\Pi^T \xi|^2 g(tD\Pi^T \xi) e^{2\pi i x^T \Pi^T \xi} d\xi.$$

By the induction hypothesis, $\Lambda(\mathcal{T}, J, \tilde{K})$ satisfies the analogue of (5.3).

Thus, to show the desired bound (5.3) for $\Lambda(\mathcal{T}, J, K)$, it suffices to show the analogue of (5.3) for $\Lambda(\mathcal{T}, J, \tilde{K})$. Analogous Fourier transform arguments as the ones leading to (5.21), put in reverse, give

$$\tilde{K}(t, x) = \int_{\mathbb{R}^m} (\Delta g)_{tD}(x + Jp) dp.$$

As is well known, the gaussians give fundamental solutions to the heat equation, more precisely, we have

$$(\Delta g)_{tD}(x) = t^{-2m}(\det D)^{-1}(-4\pi m + 4\pi^2|(tD)^{-1}x|^2)e^{-\pi|(tD)^{-1}x|^2} = 2\pi t\partial_t(g_{tD}).$$

Let $k_{\mathcal{T}}$ be defined by $l(Q_{\mathcal{T}}) = 2^{k_{\mathcal{T}}}$. Using the heat equation in the kernel of \tilde{K} we obtain

$$(2\pi)^{-1}\Lambda(\mathcal{T}, J, \tilde{K}) = \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} (F_j 1_{A_j T_k})(\Pi_j x) \right) \int_{2^{k-1}}^{2^k} t \partial_t(g_{tD})(x + Jp) \frac{dt}{t} dx dp.$$

By the fundamental theorem of calculus in t , this equals

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} (F_j 1_{A_j T_k})(\Pi_j x) \right) (g_{2^k D} - g_{2^{k-1} D})(x + Jp) dx dp \\ &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} (F_j 1_{A_j Q_{\mathcal{T}}})(\Pi_j x) \right) g_{2^{k_{\mathcal{T}}} D}(x + Jp) dx dp \end{aligned} \quad (5.24)$$

$$+ \sum_{k < k_{\mathcal{T}}} \int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} (F_j 1_{A_j T_k})(\Pi_j x) - \prod_{j \in \mathcal{C}} (F_j 1_{A_j T_{k+1}})(\Pi_j x) \right) g_{2^k D}(x + Jp) dx dp. \quad (5.25)$$

We estimate the two terms (5.24) and (5.25) separately.

We begin with (5.25). We estimate

$$g \leq C \sum_{n \geq 0} e^{-2^n} \chi_{2^n}, \quad (5.26)$$

where χ is the characteristic function of the square $[-1, 1]^{2m}$ and C is a constant that depends only on m . We insert (5.26) into (5.25). We fix $k < k_{\mathcal{T}}$ and $n \geq 0$ and consider

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} (F_j 1_{A_j T_k})(\Pi_j x) - \prod_{j \in \mathcal{C}} (F_j 1_{A_j T_{k+1}})(\Pi_j x) \right) \chi_{2^{n+k} D}(x + Jp) dx dp.$$

Using the distributive law and $T_k \subseteq T_{k+1}$, we have

$$\begin{aligned} & \left| \prod_{j \in \mathcal{C}} (F_j 1_{A_j T_k})(\Pi_j x) - \prod_{j \in \mathcal{C}} (F_j 1_{A_j T_{k+1}})(\Pi_j x) \right| \\ & \leq \sum_{j_0 \in \mathcal{C}} |F_{j_0}| 1_{A_{j_0} T_{k+1} \setminus A_{j_0} T_k}(\Pi_{j_0} x) \prod_{j \neq j_0} |F_j| 1_{A_j T_{k+1}}(\Pi_j x). \end{aligned}$$

Fix j_0 . For simplicity of notation consider $j_0 = 0$, other values of j_0 will be analogous.

Let Q be a cube of side length 2^k contained in $T_{k+1} \setminus T_k$. We consider

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} |F_0| 1_{A_0 Q}(\Pi_0 x) \left(\prod_{j \neq 0} |F_j| 1_{A_j T_{k+1}}(\Pi_j x) \right) \chi_{2^{n+k} D}(x + Jp) dx dp. \quad (5.27)$$

Let E be the $2m \times 2m$ diagonal matrix such that for all $1 \leq i \leq m$ and $\sigma = 0, 1$,

$$e_i^\sigma = \max(d_i^\sigma, 1).$$

Denote by \tilde{E} the upper left $m \times m$ block of E . If the integrand in (5.27) is not zero, then

$$\Pi_0 x \in A_0 Q, \quad (5.28)$$

$$x + Jp \in 2^{n+k} D[-1, 1]^{2m} \subseteq 2^{n+k} E[-1, 1]^{2m}. \quad (5.29)$$

Let us denote

$$S := 2^{n+k} \tilde{E}[-1, 1]^m.$$

Applying Π_0 and Π_1 to (5.29), we obtain

$$\Pi_0 x + A_0 p \in S, \quad (5.30)$$

$$\Pi_1 x + A_1 p \in S. \quad (5.31)$$

Combining (5.28) and (5.30), we obtain

$$A_0 p \in S - A_0 Q. \quad (5.32)$$

Applying $A_1 A_0^{-1}$, we obtain

$$A_1 p \in A_1 A_0^{-1} S - A_1 Q. \quad (5.33)$$

Combining (5.31) with (5.33), we obtain

$$\Pi_1 x \in S - A_1 A_0^{-1} S + A_1 Q. \quad (5.34)$$

We cover the set in (5.34) by pairwise disjoint boxes

$$P = P_1 \times \cdots \times P_m,$$

for example with corners in a suitable lattice, such that $|P_h| = 2^n 2^k \tilde{e}_h$. Let \mathcal{P} denote the set of boxes in this cover. We estimate (5.27) using the cover \mathcal{P} and then integrating p by

$$\sum_{P \in \mathcal{P}} \frac{|S_P|}{2^{2m(k+n)} \det D} \int_{\mathbb{R}^{2m}} |F_0| 1_{A_0 Q}(\Pi_0 x) |F_1| 1_{A_1 T_{k+1} \cap P}(\Pi_1 x) \left(\prod_{j \neq 0,1} |F_j| 1_{A_j T_{k+1}}(\Pi_j x) \right) dx \quad (5.35)$$

where

$$S_P := (A_1^{-1}(S - P)) \cap (A_0^{-1}S - Q)$$

contains the essential support of p by (5.31), (5.32), and $\Pi_1 x \in P$.

We cover $A_0 Q$ by a rectangular box

$$A_0 Q \subseteq I_1^0 \times \cdots \times I_m^0$$

such that for each $1 \leq h \leq m$ we have $|I_h^0| \leq C 2^k$ with C depending on Δ . We fix one box $P \in \mathcal{P}$ and write I_h^1 for P_h . Write

$$I_j = I_1^{j(1)} \times \cdots \times I_m^{j(m)}.$$

Then we estimate the summand in (5.35) for this box P by

$$\frac{|S_P|}{2^{2m(k+n)} \det D} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} |F_j| 1_{A_j T_{k+1} \cap I_j}(\Pi_j x) \right) dx.$$

Here we used that $Q \subseteq T_{k+1}$ and that $\Pi_0 x \in I_0$ together with $\Pi_1 x \in I_1$ implies $\Pi_j x \in I_j$.

Applying the Brascamp-Lieb inequality in Lemma 3.2 estimates the last display by

$$\begin{aligned} &\leq C \frac{|S_P|}{2^{2m(k+n)} \det D} \prod_{j \in \mathcal{C}} |I_j|^{2^{1-m}} [F_j 1_{A_j T_{k+1}}]_{I_j} \\ &\leq C \frac{|S_P| \det \tilde{E}}{2^{mn} \det D} (1 + |\tilde{E}|)^{2\alpha} \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}). \end{aligned}$$

Here we estimated averages by the modified maximal functions using that if $A_j T_{k+1}$ has a nonempty intersection with I_j then the box CI_j for suitable C depending on Δ covers

the set $A_j Q'$ for some $Q' \in \mathcal{T}$. To estimate the product of $|I_j|^{2^{1-m}}$ we used that each I_h^σ appears in the product 2^{m-1} times.

The parallelepipeds $A_1^{-1}(P - S)$ have bounded overlap depending on m as P runs through \mathcal{P} , because the box $P - S$ is contained in $3P$, which intersects $3P'$ only for near neighbors P' of P . Hence

$$\sum_{P \in \mathcal{P}} |S_P| \leq C |A_0^{-1}S - Q| \leq C 2^{m(n+k)} \det \tilde{E}.$$

Here, the last inequality is obtained as follows. The set S can be covered by $C 2^{mn} \det \tilde{E}$ dyadic boxes of side length 2^k . Then the set $A_0^{-1}S$ can also be covered by $C 2^{mn} \det \tilde{E}$ dyadic boxes of side length 2^k with a constant depending on Δ . Finally, the set $A_0^{-1}S - Q$ can be covered by $C 2^{mn} \det \tilde{E}$ dyadic boxes of side length 2^{k+2} . This gives the above volume estimate.

Summing over the boxes in \mathcal{P} estimates (5.27) by

$$C \frac{\det \tilde{E}^2}{\det D} (1 + |\tilde{E}|)^{2\alpha} 2^{km} \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}).$$

Summing over the various $j_0 \in \mathcal{C}$, all estimated analogously to $j_0 = 0$, and then summing over the disjoint cubes Q of volume 2^{km} in $T_{k+1} \setminus T_k$ and summing over $k < k_{\mathcal{T}}$, using that the $T_{k+1} \setminus T_k$ are disjoint in $Q_{\mathcal{T}}$, and summing over n estimates (5.25) by

$$\begin{aligned} & C \frac{\det \tilde{E}^2}{\det D} (1 + |\tilde{E}|)^{2\alpha} |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}) \\ &= C \left(\max_{i=1, \dots, m} (1 + d_i)^{2\alpha} \prod_{i=1}^m \max\{d_i^{-2}, 1\} \right) |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_\alpha(F_j)(A_j \mathcal{T}). \end{aligned}$$

It remains to estimate (5.24), which is done similarly as (5.25) but simpler. Estimating the gaussian by a superposition of characteristic functions of boxes we consider

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^{2m}} \left(\prod_{j \in \mathcal{C}} |F_j| 1_{A_j Q_{\mathcal{T}}}(\Pi_j x) \right) \chi_{2^{n+k}D}(x + Jp) dx dp,$$

which then is estimated analogously to (5.27).

5.3. Cauchy-Schwarz: Bounds for the third kernel imply bounds for the first kernel one level down. Let $\Delta, A_j, \alpha, l, D, F_j$ be as in Proposition 5.1. Let c, u, i be as in (1.6) and assume $i > l$. Without loss of generality we assume $i = l + 1$, as we may interchange the indices i and $l + 1$ if necessary in the desired estimate and all assumptions remain intact.

Consider the quantity

$$(1 + |u|^{4m}) \Lambda(\mathcal{T}, J, K)((F_j)_{j \in \mathcal{C}})$$

for the kernel K defined by (5.4). This quantity is equal to

$$\int_{\mathbb{R}^{2m}} \int_0^\infty \int_{\mathbb{R}^m} c(t) \left(\prod_{j \in \mathcal{C}} (F_j 1_{A_j T_i})(\Pi_j x) \right) (\partial_i^0 \partial_i^1 g)_{tD}(x + Jp + tDu) dp \frac{dt}{t} dx, \quad (5.36)$$

where we have set $T_t := T_k$ if $2^{k-1} \leq t < 2^k$. We change the order of integration, so that the innermost integration variables are x_i^0 and x_i^1 . These two integrals separate. Hence we obtain for (5.36)

$$\int_{\mathbb{R}^{2m-2}} \int_0^\infty \int_{\mathbb{R}^m} c(t) \prod_{\sigma=0}^1 \left(\int_{\mathbb{R}} \left(\prod_{j(i)=\sigma} (F_j 1_{A_j T_t})(\Pi_j x) \right) (g')_{td_i}((x + Jp + tDu)_i^\sigma) dx_i^\sigma \right) d\mu$$

with

$$d\mu = \prod_{1 \leq h \leq m, h \neq i} \left(\prod_{\sigma=0}^1 g_{td_h}((x + Jp + tDu)_h^\sigma) dx_h^\sigma \right) dp \frac{dt}{t},$$

where g denotes the one-dimensional Gaussian.

We estimate $|c(t)|$ by 1 and apply Cauchy-Schwarz in the Hilbert space $L^2(d\mu)$ to estimate the modulus square of (5.36) by

$$\prod_{\sigma=0}^1 \int_{\mathbb{R}^{2m-2}} \int_0^\infty \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}} \left(\prod_{j(i)=\sigma} (F_j 1_{A_j T_t})(\Pi_j x) \right) (g')_{td_i}((x + Jp + tDu)_i^\sigma) dx_i^\sigma \right)^2 d\mu. \quad (5.37)$$

Fix $\sigma \in \{0, 1\}$. The matrix J^σ is invertible. Define $q \in \mathbb{R}^m$ by $q := -(J^\sigma)^{-1}v$, where $v_i = u_i^\sigma$ and $v_h = 0$ for $h \neq i$. Then we have $Jq = \Theta Jq$, as one can see from

$$(J^0 q)_h = (J^1 q)_h = -v_h = 0$$

for $1 \leq h \leq i-1$ as the first $i-1$ rows of J^0 and J^1 coincide. As a consequence, using that $d_h = \underline{d}_i$ for $h \geq i$, we have

$$DJq = Jqd_i.$$

Moreover,

$$|q| \leq C|u|$$

with a constant C depending on Δ . For fixed t , we use an affine change of variables to substitute p by

$$p + tqd_i$$

and we then obtain for the σ -th factor in (5.37)

$$\int_{\mathbb{R}^{2m-2}} \int_0^\infty \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}} \left(\prod_{j(i)=\sigma} (F_j 1_{A_j T_t})(\Pi_j x) \right) (g')_{td_i}((x + Jp)_i^\sigma) dx_i^\sigma \right)^2 d\tilde{\mu}, \quad (5.38)$$

where

$$d\tilde{\mu} := \prod_{1 \leq h \leq m, h \neq i} \prod_{\lambda=0}^1 g_{td_h}((x + Jp + tD(Jq + u))_h^\lambda) dp \frac{dt}{t}.$$

The square in (5.38) is positive. We therefore may estimate (5.38) by replacing the measure $d\tilde{\mu}$ by the larger measure given by

$$C(1 + |u|)^{2m-2} \prod_{1 \leq h \leq m, h \neq i} \prod_{\lambda=0}^1 g_{2td_h(1+|u|)}((x + Jp)_h^\lambda) dp \frac{dt}{t} \quad (5.39)$$

with a constant C depending on Δ . Here we estimated for each h

$$(td_h)^{-1}g((td_h)^{-1}(x + Jp)_h^\lambda + (Jq + u)_h^\lambda) \leq C(td_h)^{-1}g((2 + 2|u|)^{-1}(td_h)^{-1}(x + Jp)_h^\lambda)$$

using the elementary inequality

$$g(a+b) \leq Cg((2+2d)^{-1}a)$$

for real a, b and positive c, d with $|b| \leq cd$ and a constant C depending only on c . The elementary inequality is proven by considering separately the case $|a| \leq 2|b|$, in which the right-hand side is larger than 1 for sufficiently large C , and $|a| > 2|b|$, in which the inequality holds with constant $C = 1$ by monotonicity of g on either half axis.

Performing the above for each $\sigma = 0, 1$ in (5.37) we estimate

$$\begin{aligned} & (1+|u|)^{4m} \Lambda(\mathcal{T}, J, K)((F_j)_{j \in \mathcal{C}})^2 \\ & \leq C \prod_{\sigma=0}^1 \int_{\mathbb{R}^{2m-2}} \int_0^\infty \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}} \left(\prod_{j(i)=\sigma} (F_j 1_{A_j T_i})(\Pi_j x) \right) (g')_{td_i}((x+Jp)_i^\sigma) dx_i^\sigma \right)^2 \\ & \quad \times \prod_{1 \leq h \leq m, h \neq i} \prod_{\lambda=0}^1 g_{2td_h(1+|u|)}((x+Jp)_h^\lambda) dp \frac{dt}{t}. \end{aligned} \quad (5.40)$$

We aim to interpret the two factors with $\sigma = 0, 1$ as forms Λ as in (5.1) for some modification of J . For $\sigma = 0, 1$ we define the matrix ${}_\sigma J$ as follows. For each $1 \leq h \leq m$ and each $\lambda = 0, 1$ we set

$${}_\sigma J_h^\lambda := J_h^\lambda$$

if $\lambda = \sigma$ or $h \neq i$ and

$${}_\sigma J_h^\lambda := J_h^{1-\lambda}$$

if $\lambda \neq \sigma$ and $h = i$. For $\sigma = 0, 1$ we define ${}_\sigma F_j := F_j$ if $j(i) = \sigma$ and ${}_\sigma F_j := F_{i^*j}$ if $j(i) \neq \sigma$. We also define ${}_\sigma A_j := \Pi_j({}_\sigma J)$ so that if $j(i) = \sigma$ we have $A_j = {}_\sigma A_j$ and if $j(i) \neq \sigma$ then we have $A_{i^*j} = {}_\sigma A_j$. In particular, we see that the matrices ${}_\sigma A_j$ are Δ -regular.

Let E be a $2m \times 2m$ diagonal matrix with two identical $m \times m$ diagonal blocks, where the first m diagonal elements e_1, \dots, e_m of E are defined by $e_h = 2(1+|u|)d_h$ if $h \neq i$ and $e_i = d_i$. The matrix E satisfies the analogue of (5.2) with l replaced by i . Then we expand out the square in (5.40) and regroup the terms to write (5.40) as

$$\begin{aligned} & \prod_{\sigma=0}^1 \int_{\mathbb{R}^{2m}} \int_0^\infty \int_{\mathbb{R}^m} \left(\prod_{j \in \mathcal{C}} ({}_\sigma F_j 1_{{}_\sigma A_j T_i})(\Pi_j x) \right) (\partial_{l+1}^0 \partial_{l+1}^1 g)_{tE}(x + \sigma Jp) dp \frac{dt}{t} dx \\ & = \prod_{\sigma=0}^1 \Lambda(\mathcal{T}, {}_\sigma J, {}_\sigma K_{l+1})(({}_\sigma F_j)_{j \in \mathcal{C}}), \end{aligned} \quad (5.41)$$

where we used $i = l+1$ and for $1 \leq \kappa \leq l+1$ we defined

$${}_\sigma K_\kappa(t, x) := \int_{\mathbb{R}^m} (\partial_\kappa^0 \partial_\kappa^1 g)_{tE}(x + \sigma Jp) dp.$$

We observe that for $1 \leq \kappa \leq l+1$

$$\Lambda(\mathcal{T}, {}_\sigma J, {}_\sigma K_\kappa)(({}_\sigma F_j)_{j \in \mathcal{C}}) \geq 0. \quad (5.42)$$

For $\kappa = l+1$ this follows from evident positivity of (5.40). Similarly we can use the symmetries of ${}_\sigma F_j$ and ${}_\sigma J$ to write

$$\Lambda(\mathcal{T}, {}_\sigma J, {}_\sigma K_\kappa)(({}_\sigma F_j)_{j \in \mathcal{C}})$$

for $\kappa < l + 1$ in analogous form to (5.40). Thus, to estimate (5.41) it suffices to estimate

$$\prod_{\sigma=0}^1 \Lambda(\mathcal{T}, \sigma J, \sum_{\kappa=1}^{l+1} \sigma K_{\kappa})((\sigma F_j)_{j \in \mathcal{C}}).$$

We identify two kernels as in (5.6). Using (5.3) for $l + 1$ we estimate the last display by

$$\begin{aligned} &\leq C \prod_{\sigma=0}^1 \left(\prod_{i=1}^m \max\{e_i^{2\alpha}, e_i^{-2}\} \right) |Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_{\alpha}(\sigma F_j)(\sigma A_j \mathcal{T}) \\ &= C \left(\prod_{i=1}^m \max\{e_i^{2\alpha}, e_i^{-2}\} \right)^2 \left(|Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_{\alpha}(F_j)(A_j \mathcal{T}) \right)^2. \end{aligned} \quad (5.43)$$

Here we reshuffled the products using

$$M_{\alpha}(\lambda F_j)(\lambda A_j \mathcal{T}) = M_{\alpha}(F_j)(A_j \mathcal{T})$$

if $j(i) = \lambda$ and

$$M_{\alpha}(\lambda F_j)(\lambda A_j \mathcal{T}) = M_{\alpha}(F_{i^*j})(A_{i^*j} \mathcal{T})$$

if $j(i) \neq \lambda$. We use the definition of e_i to estimate (5.43) as

$$\leq (1 + |u|)^{4m} C \left(\prod_{i=1}^m \max\{d_i^{2\alpha}, d_i^{-2}\} \right)^2 \left(|Q_{\mathcal{T}}| \prod_{j \in \mathcal{C}} M_{\alpha}(F_j)(A_j \mathcal{T}) \right)^2.$$

This proves the desired estimate.

6. PROOF OF THEOREM 1.5

Assume to get a contradiction that there is $q < 2^{m-1}$ so that the analog of Theorem 1.1 holds with the lower threshold 2^{m-1} in (1.3) replaced by q .

Set $\Pi = (I, I)$ so that $\Pi_j \Pi^T = I$ for all $j \in \mathcal{C}$. Let K be the Riesz kernel

$$K(x) = \frac{x_1}{|x|^{m+1}}, \quad x \in \mathbb{R}^m \setminus \{0\}.$$

For $j \in \mathcal{C}$ with $j(1) = 0$ let

$$p_j = 2^{m-1}$$

and for $j \in \mathcal{C}$ with $j(1) = 1$ let

$$p_j = \infty.$$

These values p_j satisfy the assumptions of the analog of Theorem 1.1.

Let N be a large positive number. For $j \in \mathcal{C}$ with $j(1) = 0$, let $F_j : \mathbb{R}^m \rightarrow [0, 1]$ be smooth, supported in the box

$$[0, 1] \times [-2N, 2N]^{m-1}$$

and taking value 1 on the smaller box

$$[1/4, 1/2] \times [-N, N]^{m-1}.$$

Then

$$\|F_j\|_{2^{m-1}} \leq (4N)^{(m-1)2^{1-m}}.$$

Let $F_1 : \mathbb{R}^m \rightarrow [0, 1]$ be smooth, supported in the box

$$[1/2, 2N] \times [-2N, 2N]^{m-1}$$

and taking value 1 on the smaller box

$$[1, N] \times [-N, N]^{m-1}.$$

Finally, for $j \in \mathcal{C}$ with $j(1) = 1$ but $j \neq 1$, let $F_j : \mathbb{R}^m \rightarrow [0, 1]$ be smooth and compactly supported, taking value 1 on the box $[-N, N]^m$.

With $\|F_j\|_\infty = 1$ for all $j \in \mathcal{C}$, we obtain by assumption a constant C independent of N so that

$$\left| p.v. \int_{\mathbb{R}^{2m}} K(\Pi x) \prod_{j \in \mathcal{C}} F_j(\Pi_j x) dx \right| \leq C \prod_{j \in \mathcal{C}} \|F_j\|_{p_j} \leq C(4N)^{m-1}. \quad (6.1)$$

On the support of the product of the functions $F_j \circ \Pi_j$, one has $x_1^0 \geq 0$ and $x_1^1 \geq 1/2$. Hence $K(\Pi x)$ is positive on this support. We restrict the integral to a domain where all functions F_j take value 1 and thus estimate the left-hand side of (6.1) from below by

$$\int_{[1/4, 1/2] \times [-N, N]^{m-1} \times [1, N] \times [-N, N]^{m-1}} \frac{x_1^0 + x_1^1}{|x^0 + x^1|^{m+1}} dx.$$

For $2 \leq h \leq m$, we introduce variables $y_h^+ = x_h^0 + x_h^1$ and $y_h^- = x_h^0 - x_h^1$ and combine these variables to a vector $y = (y^+, y^-)$. Shrinking the domain of integration from line to line as convenient, we estimate the last display from below by

$$\begin{aligned} & 2^{2-2m} \int_{[1/4, 1/2]} \int_{[1, N]} \int_{[-N/2, N/2]^{2m-2}} \frac{x_1^0 + x_1^1}{|(x_1^0 + x_1^1, y^+)|^{m+1}} dy dx_1^1 dx_1^0 \\ & \geq 2^{-2m} N^{m-1} \int_{[1, N]} \int_{[-N/2, N/2]^{m-1}} \frac{x_1^1}{|(2x_1^1, y^+)|^{m+1}} dy^+ dx_1^1 \\ & \geq 2^{-2m} N^{m-1} \int_{[1, N/4]} \int_{[x_1^1, 2x_1^1]^{m-1}} \frac{x_1^1}{(2mx_1^1)^{m+1}} dy^+ dx_1^1 \\ & \geq (8m)^{-m-1} N^{m-1} \int_{[1, N/4]} \frac{1}{x_1^1} dx_1^1 = (8m)^{-m-1} N^{m-1} \ln(N/4). \end{aligned}$$

The right-hand side grows faster in N than the right-hand side of (6.1), which gives a contradiction for N large enough. This completes the proof of Theorem 1.5.

REFERENCES

- [1] A. Barron, J.M. Conde-Alonso, Y. Ou, and G. Rey, *Sparse domination and the strong maximal function*, Adv. Math. **345** (2019), 1–26.
- [2] C. Benea and C. Muscalu, *Mixed-norm estimates via the helicoidal method*, preprint (2020), arXiv:2007.01080.
- [3] A. Culiuc, F. Di Plinio, and Y. Ou, *Domination of multilinear singular integrals by positive sparse forms*, J. Lond. Math. Soc. (2) **98** (2018), no. 2, 369–392.
- [4] Y. Do, and C. Thiele, *L^p theory for outer measures and two themes of Lennart Carleson united*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 2, 249–296.
- [5] P. Durcik, *L^p estimates for a singular entangled quadrilinear form*, Trans. Amer. Math. Soc. **369** (2017), no. 10, 6935–6951.
- [6] P. Durcik, V. Kovač, and L. Rimanić, *On side lengths of corners in positive density subsets of the Euclidean space*, Int. Math. Res. Not., (2018), no. 22, 6844–6869.
- [7] P. Durcik and V. Kovač, *A Szemerédi type theorem for subsets of the unit cube*, Anal. PDE. **15** (2022), no. 2, 507–549.
- [8] P. Durcik, V. Kovač, K.A. Škreb, and C. Thiele, *Norm variation of ergodic averages with respect to two commuting transformations*, Ergodic Theory Dynam. Systems **39** (2019), no. 3, 658–688.

- [9] P. Durcik and J. Roos, *Averages of simplex Hilbert transforms*, Proc. Amer. Math. Soc. **149** (2021), no. 2, 633–647.
- [10] P. Durcik and C. Thiele, *Singular Brascamp-Lieb inequalities with cubical structure*, Bull. Lond. Math. Soc. **52** (2020), no. 2, 283–298.
- [11] P. Durcik and C. Thiele, *Singular Brascamp-Lieb: a survey*, Geometric Aspects of Harmonic Analysis, Vol. 45, Springer INdAM Series, Springer International Publishing, 2021.
- [12] V. Kovač, C. Thiele, and P. Zorin-Kranich, *Dyadic triangular Hilbert transform of two general functions and one not too general function*, Forum Math. Sigma **3** (2015), Paper No. e25, 27.
- [13] V. Kovač, *Boundedness of the twisted paraproduct*, Rev. Mat. Iberoam. **28** (2012), no. 4, 1143–1164.
- [14] V. Kovač, *Density theorems for anisotropic point configurations*, Canad. J. Math., to appear, arXiv:2008.01060.
- [15] V. Kovač and M. Stipčić, *Convergence of ergodic-martingale paraproducts*, Statist. Probab. Lett. **164** (2020), 108826, 6.
- [16] K. Li, K. Moen, and W. Sun, *The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators*, J. Fourier Anal. Appl. **20** (2014), no. 4, 751–765.
- [17] C. Muscalu and Y. Zhai, *Five-linear singular integral estimates of Brascamp-Lieb type*, preprint (2020), arXiv:2001.09064.
- [18] A. K. Lerner and F. Nazarov, *Intuitive dyadic calculus: the basics*, Expositiones Mathematicae **37** (2019), no. 3, 225–265.
- [19] M. Mirek and C. Thiele, *A local $T(b)$ theorem for perfect multilinear Calderón-Zygmund operators*, Proc. Lond. Math. Soc. (3) **114** (2017), no.1, 35–59.
- [20] Z. Nieraeth, *Quantitative estimates and extrapolation for multilinear weight classes*, Math. Ann. **375** (2019), no. 1-2, 453–507.
- [21] R. Oberlin and C. Thiele, *New uniform bounds for a Walsh model of the bilinear Hilbert transform*, Indiana Univ. Math. J. **60** (2011), no. 5, 1693–1712.
- [22] M. Stipčić, *$T(1)$ theorem for dyadic singular integral forms associated with hypergraphs*, J. Math. Anal. Appl. **481** (2020), no. 2, 123496, 27.
- [23] P. Zorin-Kranich, *A_p - A_∞ estimates for multilinear maximal and sparse operators*, J. Anal. Math., **138** (2019), no. 2, 871–889.

SCHMID COLLEGE OF SCIENCE AND TECHNOLOGY, CHAPMAN UNIVERSITY, ONE UNIVERSITY DRIVE, ORANGE, CA 92866, USA

Email address: durcik@chapman.edu

DEPARTMENT OF MATHEMATICAL ANALYSIS, FACULTY OF MATHEMATICS AND PHYSICS, CHARLES UNIVERSITY, SOKOLOVSKÁ 83, 186 75 PRAHA 8, CZECH REPUBLIC

Email address: slavikova@karlin.mff.cuni.cz

MATHEMATISCHES INSTITUT, UNIVERSITÄT BONN, ENDENICHER ALLEE 60, 53115 BONN, GERMANY

Email address: thiele@math.uni-bonn.de