ANALYSIS & PDE

Volume 19 No. 3 2026

POLONA DURCIK, LENKA SLAVÍKOVÁ AND CHRISTOPH THIELE

NORM-VARIATION OF TRIPLE ERGODIC AVERAGES FOR COMMUTING TRANSFORMATIONS

msp

ANALYSIS AND PDE
Vol. 19 (2026), No. 3, pp. 539–586
DOI: 10.2140/apde.2026.19.539

msp

NORM-VARIATION OF TRIPLE ERGODIC AVERAGES FOR COMMUTING TRANSFORMATIONS

POLONA DURCIK, LENKA SLAVÍKOVÁ AND CHRISTOPH THIELE

We prove an r-variation estimate, $r > 4$, in the norm for ergodic averages with respect to three commuting transformations. It is not known whether such estimates hold for all $r \geq 2$ as in the analogous cases for one or two commuting transformations, or whether such estimates hold for any $r < \infty$ for more than three commuting transformations.

1. Introduction

We prove the following norm variation bound for three commuting transformations.

Theorem 1.1. *For all $r > 4$, there exists a constant $C > 0$ such that the following holds. Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space, $T_0, T_1, T_2 \colon X \to X$ mutually commuting measure preserving transformations, and let J and $n_0 < n_1 < \cdots < n_J$ be positive integers. For any $f_0, f_1 \in L^8(X)$ and $f_2 \in L^4(X)$, each of respective norm 1, we have the bound*

$$\sum_{j=1}^{J} \big\| M_{n_j}(f_0, f_1, f_2) - M_{n_{j-1}}(f_0, f_1, f_2) \big\|_{L^2(X)}^{r} \leq C,$$

where we have defined for almost every $x \in X$

$$M_n(f_0, f_1, f_2)(x) := \frac{1}{n} \sum_{i=0}^{n-1} f_0(T_0^i x) f_1(T_1^i x) f_2(T_2^i x).$$

MSC2020: primary 42B20; secondary 37A30, 42B15.
Keywords: ergodic average, multilinear estimate, singular integral.

Norm variation bounds with $r \geq 2$ for one transformation were proven in [Jones et al. 1996] and for two commuting transformations in [Durcik et al. 2019a], following earlier work [Kovač 2016] in the finite group setting. Norm variation bounds with any $r < \infty$ for any number of commuting transformations were stated as an open problem in the closing section of [Avigad and Rute 2015]. Any such bounds remain unknown for more than three commuting transformations. It is natural to conjecture norm variation bounds for $r \geq 2$ for any number of commuting transformations. The passage from two to three commuting transformations is a critical transition as present techniques very clearly fail to address the sharp variation threshold $r \geq 2$.

Norm variation bounds for any r are strong quantitative forms of norm convergence. Qualitative norm convergence for three or more commuting transformations was proven by Tao [2008] by finitary methods. The case for two commuting transformations had been shown before using the tools from ergodic theory. Ergodic theoretic proofs of Tao's result were given in [Host 2009; Austin 2010], and a generalization to transformations generating a nilpotent group was proven in [Walsh 2012].

Norm convergence should be compared with the more difficult question of pointwise convergence almost everywhere. Such pointwise convergence is known by the classical Birkhoff theorem for a single transformation [Birkhoff 1931], with pointwise variational estimates proven in [Bourgain 1989]. Pointwise convergence almost everywhere remains a widely recognized open problem even in the case of two general commuting transformations. This contrasts with recent developments in the area concerning multiple ergodic averages with actions of polynomial powers $T^{p(n)}$, including a number of pointwise almost everywhere convergence results under the umbrella of the Furstenberg–Bergelson–Leibman conjecture such as the bilinear but not completely linear polynomial averages in [Krause et al. 2022] or the multiparameter polynomial averages in [Bourgain et al. 2023]. For further history on the ergodic means discussed in the present paper, we refer to the paper on two commuting transformations [Durcik et al. 2019a].

By a variant of the well-known Calderón transference principle, Theorem 1.1 follows from Theorem 1.2 below. We do not elaborate on the transference principle in the present paper but refer to the case of two commuting transformations in [Durcik et al. 2019a]. It reduces quantitative convergence results to analogous results on individual orbits of the action of the group spanned by the commuting transformations and parametrized by $\mathbb{Z}^3$. The further transfer from $\mathbb{Z}^3$ to $\mathbb{R}^3$ as in Theorem 1.2 is harmless and it can be made a part of the transference principle in our setting, unlike in the setting of actions $T^{p(n)}$ with polynomials of higher degree which face number theoretic complications.

Theorem 1.2. *For all $r > 4$, there exists a constant $C > 0$ such that the following holds. For any positive integer J and positive real numbers $t_0 < t_1 < \cdots < t_J$, any $f_0, f_1 \in L^8(\mathbb{R}^3)$ and $f_2 \in L^4(\mathbb{R}^3)$ with respective norm 1, we have*

$$\sum_{j=1}^{J} \left\| M_{t_j}(f_0, f_1, f_2) - M_{t_{j-1}}(f_0, f_1, f_2) \right\|_{L^2(\mathbb{R}^3)}^r \leq C, \tag{1-1}$$

where, with e_0, e_1, e_2 the standard unit vectors in $\mathbb{R}^3$, we have defined for almost every $x \in \mathbb{R}^3$

$$M_t(f_0, f_1, f_2)(x) := \frac{1}{t} \int_0^t f_0(x + \tau e_0) f_1(x + \tau e_1) f_2(x + \tau e_2)\, d\tau. \tag{1-2}$$

Only the choice of tuple of exponents (8, 8, 4) breaks the symmetry between the three functions in the above theorems. One therefore concludes the analogous estimates for permutations of these exponents. Interpolation gives further tuples of exponents, for example the symmetric tuple (6, 6, 6).

Theorem 1.2 is proven using the theory of singular Brascamp–Lieb forms. A singular Brascamp–Lieb datum $D = (n, S, \Pi, (\Pi_s)_{s\in S})$ is a tuple containing the dimension $n \geq 1$ of the domain of integration, the finite set S parametrizing the tuple of input functions, and linear maps Π and Π_s for $s \in S$ on the domain $\mathbb{R}^n$, where Π_s maps onto the domain of the input function with parameter s, typically of smaller dimension than n. Together with some singular integral kernel K on the range of Π, the singular Brascamp–Lieb form $\Lambda_{D,K}$ is defined as

$$\Lambda_{D,K}((f_s)_{s\in S}) = \int_{\mathbb{R}^n} K(\Pi x) \prod_{s\in S} f_s(\Pi_s x)\, dx,$$

where the integral is defined in some principal value sense or, if the kernel has additional qualitative regularity as is mostly the case in the present paper, in the Lebesgue integral sense. We also often talk about the multiplier m of the form, which is the Fourier transform of the kernel K. A singular Brascamp–Lieb inequality estimates this form by a constant times the product of Lebesgue norms $\prod_{s\in S} \|f_s\|_{p_s}$ for some tuple of exponents p_s.

Singular Brascamp–Lieb inequalities with the kind of data appearing in this paper are studied in [Durcik et al. 2019b; 2022; Durcik and Thiele 2021; Muscalu and Zhai 2022] when K is a classical Calderón–Zygmund kernel. Compared with this work, the novelty in the present paper is that the kernels K do not satisfy uniform Calderón–Zygmund bounds but rather multiparameter symbol estimates arising naturally in the investigation of variation norms. These symbol estimates no longer synchronize with a Whitney decomposition of frequency space but rather involve regions determined by an arbitrary sequence of jumps between scales, such as the red regions with arbitrary eccentricity in Figure 1.

Multiparameter singular Brascamp–Lieb forms of this type appear in more basic form already in the case of two commuting transformations [Durcik et al. 2019a]. Compared to two transformations, novel challenges for three commuting transformations arise from the absence of the cubical structure of the main singular Brascamp–Lieb form relevant to Theorem 1.2. For two commuting transformations the set S of the Brascamp–Lieb datum can be naturally identified with the corners of a square, but for three commuting transformations it cannot be identified with corners of a cube, but rather with vertices of a triangular prism. Cubical structure is important to allow a loss-free symmetrization of the form along the reflection symmetries of the cube. Lacking such cubical structure, the techniques available to us lead to an unavoidable loss analogous to the work on cancellation for the simplex Hilbert transform [Durcik et al. 2019b] and also [Durcik and Kovač 2022; Durcik and Stipčić 2025]. One novelty in the present paper is that this loss needs to be absorbed by a relaxation of the variation norm parameter r towards $r > 4$. In other words, we cannot allow a loss in difference between largest and smallest scale involved, i.e., in the total number of intermediate scales involved, but only a loss in the much smaller number of jumps between the scales. Thus we need to develop an analysis that carefully uses and preserves the particular structure of multipliers as depicted in Figure 1 throughout the argument.

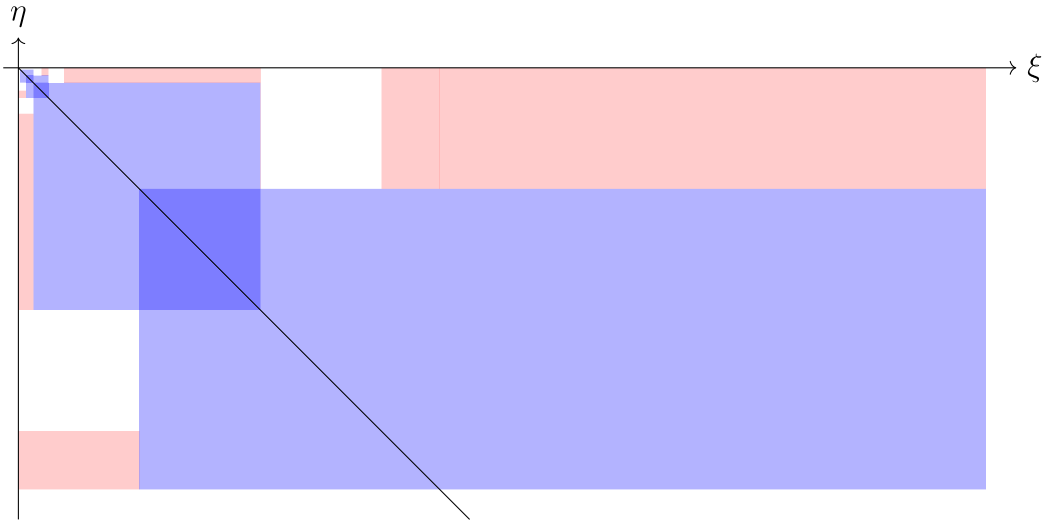

Figure 1. Structure of $m = \widehat{K}$ in Theorem 1.2.

We next provide an overview over the arguments of the present paper, which is structured into intermediate propositions and sections as in Figure 2.

Theorem 1.2 is deduced in Section 3 from (3-1), an estimate in terms of a fixed number J of jumps in the variation, which can be thought of as an endpoint estimate at $r = 4$ for (1-1). This endpoint estimate is reduced to two singular Brascamp–Lieb estimates, both with datum D_1 defined in (2-1), but with different two-dimensional kernels illustrated in Figure 1. For simplicity we focus on one quadrant in our discussion, as the other quadrants do not pose additional difficulties. The first singular Brascamp–Lieb estimate, Proposition 2.1, takes care of the so-called short variation with a multiplier that lives near the dark blue overlap regions of the light blue squares in Figure 1. The size of each dark blue square is comparable to its distance to the origin, a property we call Whitney. Proposition 2.3 takes care of the so-called long variation with multipliers living at the light blue squares themselves. Each light blue square includes potentially many scales and therefore is not in general Whitney. However, the piece of the multiplier associated with each light blue square has elementary tensor structure and telescopes into the difference between its largest and its smallest scale. For both of these propositions, it is important that the number of squares involved is controlled by J.

Proposition 2.3 is proved in Section 5. Multipliers vanishing on the diagonal in Figure 1 play a role as auxiliary objects. We use a positivity of multipliers symmetric across the diagonal to pass to a similar multiplier m_1 associated with light blue squares but constant on the diagonal. We then define two further multipliers m_2 and m_3 so that $m_1 + m_2 + m_3$ is constant in the entire plane. This constant multiplier allows a trivial bound, reducing the estimate for m_1 to estimates for m_2 and m_3. Multiplier m_2 is addressed in Proposition 2.6. It is supported near the red sticks in Figure 1. Each stick is away from the diagonal and has possibly many scales and is therefore not in general Whitney. However, the multiplier associated with each stick is an elementary tensor and as such telescopes into a small number of scales. The multiplier

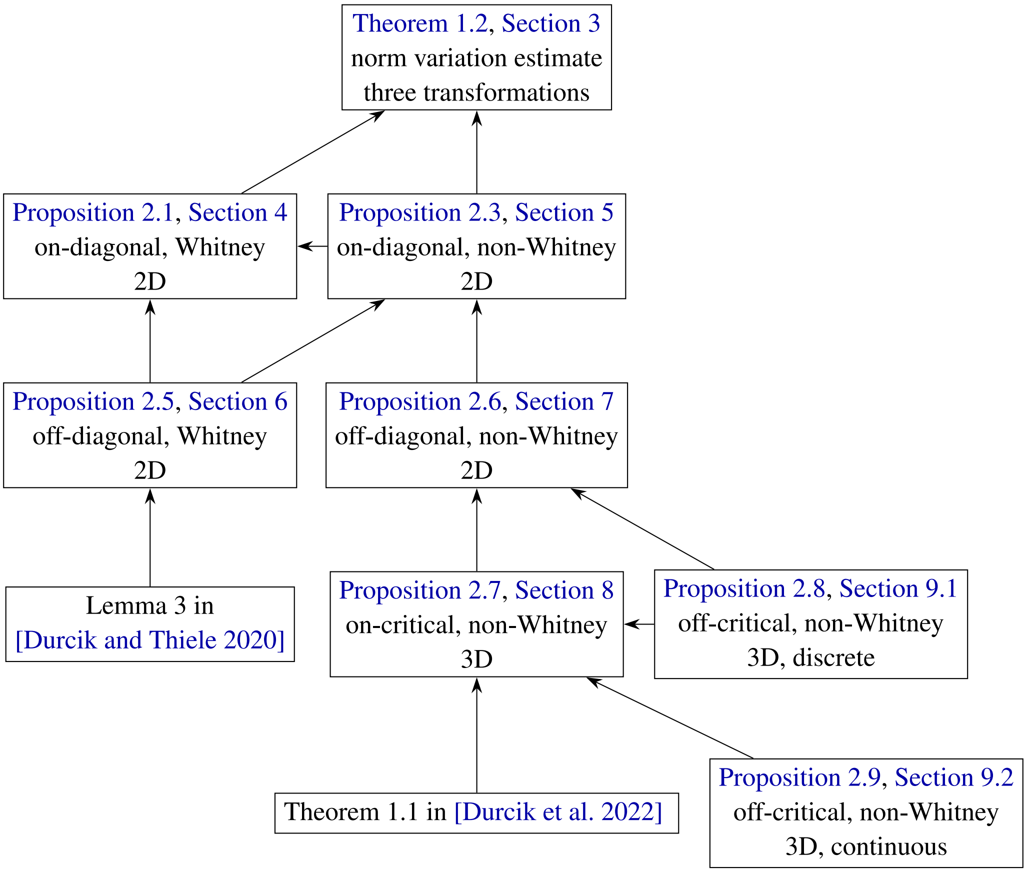

Figure 2. Structure of the proof of the main theorem.

$m_1 + m_2$ is constant both on the diagonal as well as on the white L-shaped regions in Figure 3. The multiplier m_3 is addressed in Proposition 2.5. It is supported in the at most J purple regions in between the white L-shaped regions and vanishes on the diagonal. Each purple region has a single scale and is Whitney. We decompose m_3 into a lacunary family of cones towards the diagonal shown in Figure 3. Each lacunary piece is estimated with Lemma 3 in [Durcik and Thiele 2020]. Thanks to vanishing on the diagonal, one has a geometric sum for the estimates in this family.

Proposition 2.1 is proved in Section 4. We combine the dark blue squares with a suitable family of light blue squares with tensor structure to obtain a multiplier vanishing on the diagonal. The light blue squares are estimated with Proposition 2.3, while the multiplier vanishing on the diagonal is estimated with Proposition 2.5.

Proposition 2.6 is proved in Section 7. Using the off-diagonal property of the multiplier to preserve crucial cancellation in the innermost integral, we apply the Cauchy–Schwarz inequality in the remaining integrals. We estimate one of the factors on the right-hand side of Cauchy–Schwarz using that the multiplier has J summands, which leads to the loss of $J^{1/2}$. The other factor we estimate loss-free thanks to the above mentioned cancellation. This loss-free estimate takes the form of a singular Brascamp–Lieb

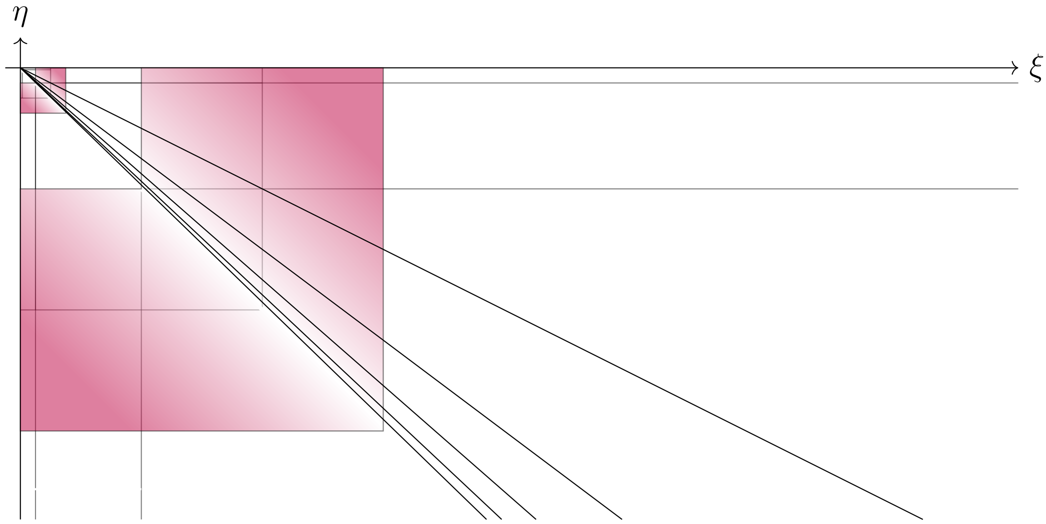

Figure 3. Lacunary cones.

form with datum D_2 defined in (2-10). The multiplier m is now three-dimensional, but consisting of pieces that are naturally of the form

$$\phi_1(\theta \cdot v_1)\phi_2(\theta \cdot v_2), \tag{1-3}$$

with two vectors $v_1 = (0, 0, 1)$ and $v_2 = (1, -1, 0)$ as shown in Figure 4. Typical behavior of the functions ϕ_1 and ϕ_2 is shown in Figure 4 on the planes perpendicular to v_1 and v_2. An important role is played by multipliers vanishing on the critical space spanned by v_1 and v_2. Such multipliers are estimated in Propositions 2.8 and 2.9 in the non-Whitney case and by Theorem 1.1 in [Durcik et al. 2022] in the Whitney case. The multiplier m does not vanish on the critical space. It is first modified using Proposition 2.8 towards a multiplier m' that also consists of pieces as in (1-3) but is more symmetric in the variables ξ, η. The multiplier m' is then estimated by Proposition 2.7.

Proposition 2.7 is proved in Section 8. The key to estimating non-Whitney multipliers such as m' is a new variant of a telescoping identity in this context that concerns three-dimensional multipliers with a two-dimensional flavor as in Figure 4 and (1-3). This identity telescopes a trivial multiplier, the product of the lightest blue squares minus the product of the darkest blue squares, into two sums consisting of products of a square in one plane with a difference of the corresponding square with a consecutive square in the other plane. One sum is arranged to allow a positivity argument using reflection symmetry across the diagonal in the ξ, η plane, the other sum is arranged similarly to allow a positivity argument with symmetry in the shown skew coordinates in the other plane. As the two positive sums add to a trivial multiplier, both are individually bounded. The given multiplier m' can be dominated by one of these constructed multipliers, using various modifications with Propositions 2.8 and 2.9 and Theorem 1.1 in [Durcik et al. 2022].

Propositions 2.8 and 2.9 are proved in Section 9. Vanishing of the multiplier on the critical space allows a lacunary decomposition away from the critical space and a further Cauchy–Schwarz. This leads to singular Brascamp–Lieb estimates with a standard cubical datum and three-fold telescoping identities

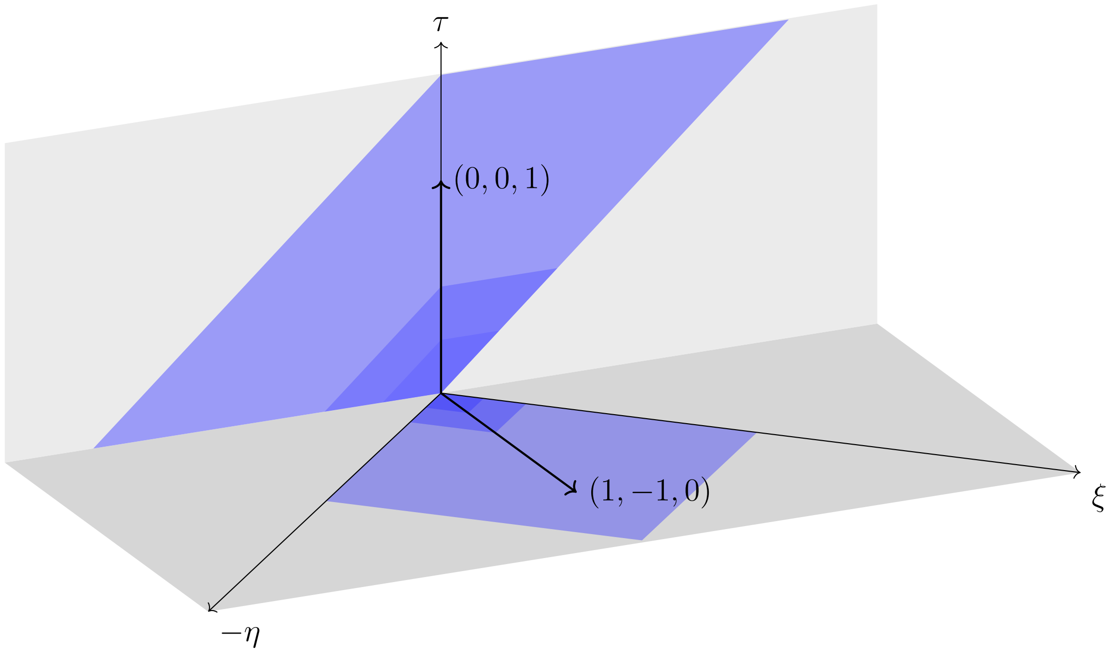

Figure 4. Three-dimensional multiplier.

for three-dimensional kernels. The non-Whitney property requires telescoping along the scales of the variation sequences. There is a mix of discrete telescoping and partial integration, with Proposition 2.8 more discrete and Proposition 2.9 more continuous. Analogous but simpler techniques appear in the Whitney case in [Durcik et al. 2022, Theorem 1.1].

We have kept the sections past Sections 1 and 2 independent of each other; each proves the theorem or one or two propositions and uses some of the other propositions or cited theorems as black boxes.

While it is plausible that our approach can be upgraded to an iterative scheme that handles more than three commuting transformations, we decided to complete and circulate the argument in the case of three transformations. This case has a single transition step with the important new techniques and does not appear to involve all the complications that one expects for the general case.

2. A collection of propositions on singular Brascamp–Lieb forms

This section contains a number of propositions stating cancellation estimates for singular Brascamp–Lieb forms for some data and some class of kernels and with symmetric tuples of test functions.

The first four propositions and two corollaries share a common datum D_1, which, after suitable change of variables, arises directly out of the original problem in Theorem 1.2. Put coordinates

$$x = (x_0, x_1, x_2, x_3^0, x_3^1)$$

on $\mathbb{R}^5$. Define

$$D_1 := (5, S, \Pi, (\Pi_s)_{s\in S}) \tag{2-1}$$

with $S = \{0, 1, 2\} \times \{0, 1\}$, with Π mapping $\mathbb{R}^5$ to $\mathbb{R}^2$ as

$$\Pi(x) = (x_3^0 - x_0 - x_1 - x_2, x_3^1 - x_0 - x_1 - x_2),$$

and with Π_s for $s = (k, j)$ mapping $\mathbb{R}^5$ to $\mathbb{R}^3$ as

$$\Pi_{(k,j)}(x) = (x_0, x_1, x_2) - x_k e_k + x_3^j e_k.$$

Each of the three following propositions will have a constant C, a parameter J and formulate a class of kernels K such that the singular Brascamp–Lieb estimate

$$|\Lambda_{D_1,K}((f_s)_{s\in S})| \le C J^{\frac{1}{2}} \tag{2-2}$$

holds for all tuples of real-valued Schwartz functions $(f_s)_{s\in S}$ such that

$$f_{(k,0)} = f_{(k,1)} \tag{2-3}$$

for each $k \in \{0, 1, 2\}$ and

$$\|f_{(0,j)}\|_8 = \|f_{(1,j)}\|_8 = \|f_{(2,j)}\|_4 = 1 \tag{2-4}$$

for each $j \in \{0, 1\}$. We point out that the symmetry assumption (2-3) arises naturally when reducing Theorem 1.2 to the propositions stated below. While our arguments could be modified in order to prove these propositions without the extra assumption (2-3), we decided not to pursue this line of generalization.

Define for any function ϕ on $\mathbb{R}^d$ the L^1 normalized scaling

$$\phi_{(t)}(x) = t^{-d}\phi(t^{-1}x).$$

Define the Fourier transform $\widehat{\phi}$ of ϕ by integration against the kernel $(x, \xi) \mapsto e^{-2\pi i x\cdot\xi}$.

The kernels in the next proposition satisfy standard two-dimensional symbol estimates with bounds depending on the parameter k. They consist of pieces satisfying a positivity assumption. Such positivity assumption is used in the proof by adding further positive terms so as to achieve better behavior on some frequency diagonal. The complexity of these kernels is bounded by J.

Proposition 2.1 (on-diagonal, Whitney, 2D [proved in Section 4]). *Let $\lambda = \frac{3}{2}$. There exists a constant $C > 0$ such that the following holds for all $k \le 0$. Let J be a positive integer and $(k_j)_{j=1}^J$ a finite strictly monotone increasing sequence of integers. Let*

$$K = \sum_{j=1}^{J} \Phi_j,$$

where for each $1 \le j \le J$ we assume Φ_j is a real-valued function on $\mathbb{R}^2$, with symmetry

$$\Phi_j(u, v) = \Phi_j(v, u)$$

and positivity in the sense

$$\int_{\mathbb{R}^2} f(u)\overline{f(v)}\Phi_j(u, v)\, du\, dv \ge 0 \tag{2-5}$$

for all complex-valued f. We assume further

$$\operatorname{supp}(\widehat{\Phi}_j) \subset \big([-2^{-k_j+20}, -2^{-k_j-20}] \cup [2^{-k_j-20}, 2^{-k_j+20}]\big)^2$$

and, for all $(u, v) \in \mathbb{R}^2$,

$$|(\Phi_j)_{(2^{-k_j})}(u, v)| \le 2^{\lambda k}(1 + 2^k|u + v|)^{-10}(1 + |u - v|)^{-10}. \tag{2-6}$$

Then estimate (2-2) holds for any tuple as in (2-3), (2-4).

We note that the particular value $\lambda = \frac{3}{2}$ is not essential for the proof of Proposition 2.1. Evidently, the analogous statement of the proposition becomes stronger for smaller values of λ. Our proof can be pushed to $\lambda > 1$ at the expense of allowing the constant in (2-2) to depend on λ. On the other hand, the upper bound $\lambda < 2$ is needed to apply Proposition 2.1 to prove Theorem 1.2. There are also constants 10 and 20 chosen in this proposition which need to be large enough but also need to relate to similar other constants in other propositions to follow.

If, in the above proposition, each Φ_j is an elementary tensor of a suitable function ϕ_j with itself, then symmetry and positivity are automatic, and k is naturally chosen as 0. We formulate this as an immediate corollary.

Corollary 2.2. *There exists a constant $C > 0$ such that the following holds. Let J be a positive integer and $(k_j)_{j=1}^J$ a finite strictly monotone increasing sequence of integers. Let*

$$K = \sum_{j=1}^{J} \phi_j \otimes \phi_j,$$

where for each $1 \le j \le J$ we assume ϕ_j is a real-valued function on $\mathbb{R}$ with

$$\operatorname{supp}(\widehat{\phi}_j) \subset [-2^{-k_j+20}, -2^{-k_j-20}] \cup [2^{-k_j-20}, 2^{-k_j+20}]$$

and, for all $u \in \mathbb{R}$,

$$|(\phi_j)_{(2^{-k_j})}(u)| \le (1 + |u|)^{-20}.$$

Then estimate (2-2) holds for any tuple as in (2-3), (2-4).

We need the following technical notion of pairs in the next proposition. Let $N = 2^{18}$. This large number is necessitated by a somewhat inefficient referral in the proof of Proposition 2.7 to a theorem in [Durcik et al. 2022]. A more hands-on approach should be able to make this number much more moderate, but this is certainly not important for our argument. A *c-pair* is a pair (ϕ_0, ϕ_1) of two real-valued integrable even functions satisfying the following assumptions. Their Fourier transforms $\widehat{\phi}_0, \widehat{\phi}_1$ map $\mathbb{R}$ to $[0, 1]$, are supported on $[-1, 1]$ and constant 1 on $[-2^{-1}, 2^{-1}]$. They satisfy

$$(\widehat{\phi}_0)^2 + (1 - \widehat{\phi}_1)^2 = 1 \quad \text{and} \quad \|\widehat{\phi}_0^{(N+30)}\|_\infty, \|\widehat{\phi}_1^{(N+30)}\|_\infty \le c. \tag{2-7}$$

Here and in what follows, $\varphi^{(k)}$ stands for the k-th derivative of φ. Lemma 2.10 below shows that there is c such that a c-pair exists. When c is at most one million times the infimum of all positive numbers c' such that a c'-pair exists, then (ϕ_0, ϕ_1) is called a universal pair. A left window is a function ϕ such that there exists a function ψ such that (ϕ, ψ) is a universal pair. A right window is a function ϕ such that there exists a function ρ such that (ρ, ϕ) is a universal pair. Note that functions ϕ that are both left and right window may exist, but a notion of two sided windows needs caution as the corresponding functions ψ and ρ may not satisfy this notion.

The kernels of the next proposition do not satisfy two-dimensional symbol estimates, at least not uniformly in the choices of sequences k_j and l_j. They still consist of pieces with a positivity assumption and elementary tensor structure with only two different scales in it and have complexity controlled by J.

Proposition 2.3 (on-diagonal, non-Whitney, 2D [proved in Section 5]). *There exists $C > 0$ such that the following holds. Let J be a positive integer and $(k_j)_{j=1}^J$ and $(l_j)_{j=1}^J$ two finite sequences of integers that are interlaced in the sense that $k_j + 10 < l_j$ for $1 \le j \le J$ and $l_j < k_{j+1}$ for $1 \le j < J - 1$. Consider a kernel*

$$K = \sum_{j=1}^{J} (\phi_{0,j} - \phi_{1,j}) \otimes (\phi_{0,j} - \phi_{1,j}),$$

where, for each j, $(\phi_{0,j})_{(2^{-k_j})}$ is a left window and $(\phi_{1,j})_{(2^{-l_j})}$ is a right window.

Then estimate (2-2) holds for any tuple as in (2-3), (2-4).

Using Corollary 2.2, we have the following corollary of Proposition 2.3,

Corollary 2.4. *The variant of Proposition 2.3, where the assumption $k_j + 10 < l_j$ is replaced by the assumption $k_j < l_j$, holds.*

To see this corollary, we split the sequence into terms with $k_j + 10 \ge l_j$ and $k_j + 10 < l_j$. The former terms are estimated with Corollary 2.2, while the latter are estimated with Proposition 2.3.

In contrast to the last proposition, the kernel of the next proposition does not oscillate on the critical frequency diagonal $\xi + \eta = 0$. The complexity still is controlled by J. We no longer have the positivity assumptions, but we do satisfy standard symbol estimates, with bounds depending on the parameter k.

Proposition 2.5 (off-diagonal, Whitney, 2D [proved in Section 6]). *Let $\lambda = \frac{3}{2}$. There exists a constant $C > 0$ such that the following holds for all $k \le 0$. Let J be a positive integer and let $(k_j)_{j=1}^J$ be a finite strictly increasing sequence of integers. Let $(\Phi_j)_{j=1}^J$ be a finite sequence of real-valued functions on $\mathbb{R}^2$. Assume that*

$$\operatorname{supp}(\widehat{\Phi}_j) \subseteq \{(\xi, \eta) \in \mathbb{R}^2 : 2^{-k_j - 30} \le |(\xi, \eta)| \le 2^{-k_j + 30}\}.$$

Assume further that, for all $(u, v) \in \mathbb{R}^2$,

$$|(\Phi_j)_{(2^{-k_j})}(u, v)| \le 2^{\lambda k}(1 + 2^k|u + v|)^{-4}(1 + |u - v|)^{-4} + (1 + |u + v|)^{-4}(1 + |u - v|)^{-4}. \tag{2-8}$$

Let K be defined by

$$K = \sum_{j=1}^{J} \Phi_j,$$

and assume that $\widehat{K}$ vanishes on the diagonal $\{(\xi, \eta) \in \mathbb{R}^2 : \xi + \eta = 0\}$.

Then estimate (2-2) holds for any tuple as in (2-3), (2-4).

The kernel of the next proposition also vanishes on the critical diagonal. It does not satisfy standard two-dimensional symbol estimates uniformly in k_j. It has no positivity assumption, but similarly to some of the positive kernels above it is a sum of J tensors with few scales in it.

Proposition 2.6 (off-diagonal, non-Whitney, 2D [proved in Section 7]). *There exists a constant $C > 0$ such that the following holds. Let J be a positive integer and $(k_j)_{j=0}^J$ a finite increasing sequence of integers with $k_{j-1} + 10 \le k_j$ for $1 \le j \le J$. For $1 \le j \le J$, let $\phi_{0,j}, \phi_{1,j}, \phi_{2,j}$ be functions such that $(\phi_{0,j})_{(2^{-k_{j-1}})}$ is a left window, while $(\phi_{1,j})_{(2^{-k_j})}$ and $(\phi_{2,j})_{(2^{4-k_j})}$ are right windows. Define*

$$K = \sum_{j=1}^{J} (\phi_{0,j} - \phi_{2,j}) \otimes \phi_{1,j}. \tag{2-9}$$

Then estimate (2-2) holds for any tuple as in (2-3), (2-4).

The remaining propositions share a singular Brascamp–Lieb datum D_2. The datum D_2 arises as a reduction from D_1 after a Cauchy–Schwarz inequality. Put coordinates $x = (x_0, x_1, x_2^0, x_3^0, x_2^1, x_3^1)$ on $\mathbb{R}^6$. Define

$$D_2 := (6, S, \Pi, (\Pi_s)_{s \in S}) \tag{2-10}$$

with $S = \{0, 1\} \times \mathcal{C}$, where $\mathcal{C}$ is the set of functions $j : \{0, 1\} \to \{0, 1\}$, with Π mapping $\mathbb{R}^6$ to $\mathbb{R}^3$ as

$$\Pi(x) = \big(x_2^0 - x_0 - x_1 - x_3^0,\, x_2^1 - x_0 - x_1 - x_3^0,\, x_3^1 - x_3^0\big),$$

and with Π_s for $s = (k, j)$ mapping $\mathbb{R}^6$ to $\mathbb{R}^3$ as

$$\Pi_{(k,j)}(x) = (x_k, x_2^{j(0)}, x_3^{j(1)}).$$

For this datum D_2 and a kernel K, we are interested in a loss-free estimate

$$|\Lambda_{D_2,K}((f_s)_{s \in S})| \le C \tag{2-11}$$

for any tuple of real-valued Schwartz functions $(f_s)_{s \in S}$ with

$$f_{(k,j)} = f_{(k,j')} \tag{2-12}$$

for all $k \in \{0, 1\}$ and $j, j' \in \mathcal{C}$, and

$$\|f_s\|_8 = 1 \tag{2-13}$$

for all $s \in S$.

The next proposition is a variant of Proposition 2.3, adjusted to the datum D_2. The kernel has some positivity properties and pieces arising from suitable elementary tensor structure. The complexity J here is not relevant, as we obtain estimates independent of J.

We write g for the Gaussian $g(x) = e^{-\pi|x|^2}$, typically in one dimension but occasionally in more than one dimension. We have $\hat{g} = g$. We write h for the derivative of the Gaussian in one dimension, $h(x) = -2\pi x g(x)$. Recall $N = 2^{18}$.

Proposition 2.7 (on-critical, non-Whitney, 3D [proved in Section 8]). *There exists a constant $C > 0$ such that the following holds. Let $\alpha \geq 1$. Let J be a positive integer and $(k_j)_{j=0}^J$ a finite increasing sequence of integers with $k_{j-1} + 10 \leq k_j$ for $1 \leq j \leq J$. Let $(m_j)_{j=1}^J$ be a sequence of real numbers with $k_j - 1 \leq m_j \leq k_j$ for $1 \leq j \leq J$. For $0 \leq j \leq J$, let χ_j be a function such that $(\chi_j)_{(2^{2-k_j})}$ is a left window, and let ϕ_j be such that $\widehat{\phi}_j \geq 0$ and*

$$(\widehat{\phi}_j)^2 = (\widehat{\chi}_{j-1})^2 - (\widehat{\chi}_j)^2.$$

Let

$$K(u, v, z) = \alpha^{-N} \sum_{j=1}^{J} \int_{\mathbb{R}} g_{(\alpha 2^{m_j})}(u + p) g_{(\alpha 2^{m_j})}(v + p) \phi_j(z + p) \phi_j(p)\, dp.$$

Then estimate (2-11) holds for any tuple as in (2-12), (2-13).

Proposition 2.7 will be proven using the next two propositions. Both involve the datum D_2. Both exploit a vanishing of the function $\widehat{K}$ on the critical space $\xi + \eta = 0$.

Proposition 2.8 (off-critical, non-Whitney, 3D, discrete [proved in Section 9]). *There is a constant C such that the following holds. Let J be a positive integer. For $1 \leq i \leq 2$, let $(a_{i,j})_{j=1}^J$ be increasing sequences of positive real numbers.*

For $1 \leq j \leq J$, let $\rho_j : \mathbb{R}^4 \to \mathbb{R}$ be a continuous function satisfying

$$\int_{\mathbb{R}^2} |\rho_j|(u_1 + p, u_2 + p, u_3 + r, u_4 + r)\, dp\, dr \leq a_{1,j}^{-1}(1 + a_{1,j}^{-1}|u_1 - u_2|)^{-2} a_{2,j}^{-1}(1 + a_{2,j}^{-1}|u_3 - u_4|)^{-2} \quad (2\text{-}14)$$

for every $(u_1, u_2, u_3, u_4) \in \mathbb{R}^4$. Let $(c_j)_{j=0}^J$ be an increasing sequence of positive real numbers, well-separated in that $2c_{j-1} \leq c_j$ for $1 \leq j \leq J$. Let χ be a left window. For $1 \leq j \leq J$ let $\phi_j : \mathbb{R} \to \mathbb{R}$ be a continuous function, which exists due to the left window property of χ, satisfying $\widehat{\phi}_j \geq 0$ and

$$(\widehat{\phi}_j)^2 = (\widehat{\chi}_{(c_{j-1})})^2 - (\widehat{\chi}_{(c_j)})^2.$$

Let K be defined by

$$K(u, v, z) = \sum_{j=1}^{J} \int_{\mathbb{R}^3} \phi_j(p) \phi_j(q) \rho_j(u + p + q + r, v + p + q + r, z + r, r)\, dp\, dq\, dr. \quad (2\text{-}15)$$

Then estimate (2-11) holds for any tuple as in (2-12), (2-13).

The orthogonal complement $V^\perp$ of the subspace

$$V = \{(\xi, \eta, \tau, -(\xi + \eta + \tau), -(\xi + \eta), -(\xi + \eta)) : \xi, \eta, \tau \in \mathbb{R}\}$$

of $\mathbb{R}^6$ can be parametrized as

$$\{(p + q + r, p + q + r, r, r, p, q) : p, q, r \in \mathbb{R}\}.$$

As (2-15) is an integral over $V^\perp$ of a function F in $\mathbb{R}^6$, its Fourier transform is the restriction to V of the Fourier transform of $\widehat{F}$ to that subspace. Hence, for some universal constant C,

$$\widehat{K}(\xi,\eta,\tau) = C\sum_{j=1}^{J} \widehat{\phi}_j(\xi+\eta)^2 \hat{\rho}_j(\xi,\eta,\tau,-\xi-\eta-\tau). \tag{2-16}$$

This expression shows the vanishing of $\widehat{K}(\xi,\eta,\tau)$ on the hyperplane $\xi+\eta=0$.

Also in the following proposition, $\widehat{K}$ vanishes on $\xi+\eta$. It is made up by a very specific part in the variables ξ, η and a rather general part in the variables τ and $\tau+\xi+\eta$.

Proposition 2.9 (off-critical, non-Whitney, 3D, continuous [proved in Section 9]). *There is a constant C such that the following holds. Let J be a positive integer and $(a_j)_{j=0}^J$, $(b_j)_{j=1}^J$ be increasing sequences of positive real numbers. For $1\le j\le J$ let $\phi_j:\mathbb{R}^2\to\mathbb{R}$ be a continuous function satisfying*

$$|\phi_j(u_1,u_2)| \le (b_j)^{-2}(1+b_j^{-1}|(u_1,u_2)|)^{-4}. \tag{2-17}$$

Let K be a kernel such that

$$\widehat{K}(\xi,\eta,\tau) = \sum_{j=1}^{J}\int_{a_{j-1}}^{a_j} t^2(\xi+\eta)^2 g(t\xi)g(t\eta)\,\frac{dt}{t}\widehat{\phi}_j(\tau,-\xi-\eta-\tau). \tag{2-18}$$

Then estimate (2-11) holds for any tuple as in (2-12), (2-13).

We remark on a symmetry in the datum D_2. We do a change of variables in the kernel using the linear map

$$L(a,b,c) = (a+b-c,\, a-b,\, c).$$

Define

$$\widetilde{\Pi}(x) := L\circ\Pi(x) = \big(x_2^0+x_2^1-x_3^0-x_3^1-2(x_0+x_1),\, x_2^0-x_2^1,\, x_3^1-x_3^0\big).$$

Define $\widetilde{D}_2$ from D_2 by replacing Π by $\widetilde{\Pi}$, and choose $\widetilde{K}$ so that $\widetilde{K}\circ L=K$. We obtain

$$\Lambda_{D_2,K}((f_s)_{s\in S}) = \Lambda_{\widetilde{D}_2,\widetilde{K}}((f_s)_{s\in S}).$$

The map $\widetilde{\Pi}$ has a symmetry under interchanging the last two entries at the same time as precomposing with the involution

$$(x_0,x_1,x_2^0,x_3^0,x_2^1,x_3^1) \mapsto (x_0,x_1,-x_3^0,-x_2^0,-x_3^1,-x_2^1).$$

This involution can be seen as acting on the tuple of functions f_s, and hence we have the following consequence for the associated form. Define $\widetilde{K}^*(a,b,c)=\widetilde{K}(a,c,b)$. For $j\in\mathcal{C}$, define $j^*\in\mathcal{C}$ by $j^*(l)=j(1-l)$ and define $f^*_{(k,j)}(a,b,c)=f_{(k,j^*)}(a,-c,-b)$. Then

$$\Lambda_{\widetilde{D}_2,\widetilde{K}}((f_s)_{s\in S}) = \Lambda_{\widetilde{D}_2,\widetilde{K}^*}((f_s^*)_{s\in S}). \tag{2-19}$$

We finally introduce a further datum D_A, which is associated with a regular 3×3 matrix A and has $n=6$. Let S be the set of functions $S:\{0,1,2\}\to\{0,1\}$. We put coordinates $x=(x_1^0,x_2^0,x_3^0,x_1^1,x_2^1,x_3^1)$

on $\mathbb{R}^6$. We define

$$D_A := (6, S, \Pi, (\Pi_s)_{s\in S}), \tag{2-20}$$

where the projection $\Pi : \mathbb{R}^6 \to \mathbb{R}^3$ is given by

$$\Pi(x)^T = (I, A)x^T, \tag{2-21}$$

where I is the 3×3 identity matrix and (I, A) is a 3×6 block matrix. For $s \in S$, $\Pi_s : \mathbb{R}^6 \to \mathbb{R}^3$ is given by

$$\Pi_s(x) = (x_1^{s(0)}, x_2^{s(1)}, x_3^{s(2)}).$$

Note that after the relabelling of the coordinates, this datum has the same components as the datum D_2 except for the choice of the projection Π. We have used transposes in (2-21) as we usually write vectors as rows while the matrix equation (2-21) expects columns. The datum D_A will be used in the proofs of Propositions 2.5, 2.8, and 2.9. In the latter two cases we will only use it with $A = -I$.

We conclude this section with the previously announced existence result.

Lemma 2.10. *There exists a $c > 0$ and a c-pair as defined near (2-7).*

Proof. Let $\psi : [0, \infty) \to \mathbb{R}$ be a smooth monotone decreasing function with

$$\psi(x) = \tfrac{1}{2} \quad \text{for } x \in \left[0, \tfrac{5}{6}\right], \qquad \psi(x) = 0 \quad \text{for } x \in [1, \infty).$$

Let $\rho : [0, \infty) \to \mathbb{R}$ be a smooth monotone increasing function with

$$\rho(x) = 0 \quad \text{for } x \in \left[0, \tfrac{1}{2}\right], \qquad \rho(x) = 3^{\frac{1}{2}} \quad \text{for } x \in \left[\tfrac{4}{6}, \infty\right).$$

There exists a smooth even function ϕ_0 on $\mathbb{R}$ such that its Fourier transform is nonnegative and satisfies on $[0, \infty)$

$$(\widehat{\phi_0})^2 = (4 - \rho^2)\psi^2,$$

because the right-hand side equals ψ^2 on $\left[\tfrac{4}{6}, \infty\right)$ and is bounded below by $\tfrac{1}{4}$ on $\left[0, \tfrac{5}{6}\right]$ and constant 1 on $\left[0, \tfrac{1}{2}\right)$. There exists a smooth even function ϕ_1 on $\mathbb{R}$ such that its Fourier transform is nonnegative and fulfills on the interval $[0, \infty)$

$$(1 - \widehat{\phi_1})^2 = 1 - (4 - \rho^2)\psi^2,$$

because the right-hand side equals $\tfrac{1}{4}\rho^2$ on $\left[0, \tfrac{5}{6}\right]$ and is bounded below by $\tfrac{3}{4}$ on $\left[\tfrac{4}{6}, \infty\right)$ and constant on $\left[0, \tfrac{1}{2}\right]$. The pair (ϕ_0, ϕ_1) then satisfies the assumptions for a c-pair with

$$c = \max(\|\widehat{\phi}_0^{(N+30)}\|_\infty, \|\widehat{\phi}_1^{(N+30)}\|_\infty). \qquad \square$$

We write $A \lesssim B$ if there exists a constant $C > 0$ such that $|A| \le C|B|$ uniformly over all values of parameters appearing in the expressions A and B.

3. Proof of Theorem 1.2 from Proposition 2.1 and Corollaries 2.2 and 2.4

This section follows the corresponding argument in [Durcik et al. 2019a] for two commuting transformations with minor modifications. We summarize and streamline the argument.

Let J be given, without loss of generality we may assume $J > 2$. Let also positive real numbers $t_0 < t_1 < \cdots < t_J$ be given. Let f_0, f_1, f_2 be real-valued measurable functions on $\mathbb{R}^3$, normalized as

$$\|f_0\|_4 = \|f_1\|_8 = \|f_2\|_8 = 1.$$

We will prove a weak-type endpoint estimate at $r = 4$, namely for any $f_0 \in L^4(\mathbb{R}^3)$ and $f_1, f_2 \in L^8(\mathbb{R}^3)$ with respective norm 1,

$$\sum_{j=1}^{J} \|M_{t_j}(f_0, f_1, f_2) - M_{t_{j-1}}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}. \tag{3-1}$$

We call (3-1) an endpoint estimate as it would follow from the hypothetical inequality (1-1) with $r = 4$ by the Cauchy–Schwarz inequality, and conversely (3-1) implies (1-1) for parameters $r > 4$. Namely, (3-1) allows by Chebyshev's inequality to estimate the number of λ-jumps of the norm by $O(\lambda^{-4})$, which then allows to deduce (1-1) by a layer cake representation of the r-variation. Theorem 1.2 will thus follow as soon as we prove (3-1).

We decompose the characteristic function $\mathbb{1}_{[0,1)}$ into smoother functions. Let χ be a left window and define

$$\theta := \chi - \chi_{(2)}.$$

Then $\hat{\theta}$ is supported in $[-1, -2^{-2}] \cup [2^{-2}, 1]$ and, as detailed in [Durcik et al. 2019a, Section 2.4],

$$\mathbb{1}_{[0,1)} = \mathbb{1}_{[0,1)} * \chi + \sum_{k=-\infty}^{-1} \mathbb{1}_{[0,\infty)} * \theta_{(2^k)} - \sum_{k=-\infty}^{-1} \mathbb{1}_{[1,\infty)} * \theta_{(2^k)} =: \varphi + \sum_{k=-\infty}^{-1} \varphi_{0,k} + \sum_{k=-\infty}^{-1} \varphi_{1,k}. \tag{3-2}$$

For $\vartheta \in L^1(\mathbb{R})$ we define in analogy with (1-2) for $x \in \mathbb{R}^3$

$$M_t^{\vartheta}(f_0, f_1, f_2)(x) := \int_{\mathbb{R}} f_0(x + ue_0) f_1(x + ue_1) f_2(x + ue_2) \vartheta_{(t)}(u)\, du.$$

Using (3-2) and the triangle inequality on the sum in k, it suffices to show in place of (1-1) for every $k \leq -1$,

$$\sum_{j=1}^{J} \|M_{t_j}^{\varphi}(f_0, f_1, f_2) - M_{t_{j-1}}^{\varphi}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}, \tag{3-3}$$

$$\sum_{j=1}^{J} \|M_{t_j}^{\varphi_{0,k}}(f_0, f_1, f_2) - M_{t_{j-1}}^{\varphi_{0,k}}(f_0, f_1, f_2)\|_2^2 \lesssim 2^{2k} J^{\frac{1}{2}}, \tag{3-4}$$

$$\sum_{j=1}^{J} \|M_{t_j}^{\varphi_{1,k}}(f_0, f_1, f_2) - M_{t_{j-1}}^{\varphi_{1,k}}(f_0, f_1, f_2)\|_2^2 \lesssim 2^{\gamma k} J^{\frac{1}{2}}, \tag{3-5}$$

where $\gamma = \frac{1}{2}$. In fact, it will follow from our argument that inequality (3-5) continues to hold with any $\gamma < 1$, at the expense of allowing the constant in that inequality to depend on γ. The estimate (3-3) is acceptable and the estimates (3-4) and (3-5) give a geometric series over $k \le -1$ and are thus acceptable as well.

We first prove (3-3). We reduce further (3-3) to the analogous estimate but with the bump function φ replaced by one whose Fourier transform is constant near the origin. We write

$$\varphi = \chi + (\varphi - \chi) = \chi + \sum_{l=-2}^{\infty} (\varphi - \chi) * \theta_{(2^l)} =: \chi + \sum_{l=-2}^{\infty} \varphi_{2,l}.$$

It then suffices to show

$$\sum_{j=1}^{J} \| M_{t_j}^{\chi}(f_0, f_1, f_2) - M_{t_{j-1}}^{\chi}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}, \tag{3-6}$$

$$\sum_{j=1}^{J} \| M_{t_j}^{\varphi_{2,l}}(f_0, f_1, f_2) - M_{t_{j-1}}^{\varphi_{2,l}}(f_0, f_1, f_2)\|_2^2 \lesssim 2^{-2l} J^{\frac{1}{2}}. \tag{3-7}$$

We first prove (3-6). We split into long and short variation as in [Jones et al. 2008]. Enlarging the sequence t_j if necessary while at most doubling the number of terms and retaining at least a quarter of the left-hand side of (3-6), we may assume that for each t_j there is a t_i which is an integer power of 2 with $t_i \le t_j < 2t_i$. Let $(k_i)_{i=0}^{I}$ be the increasing sequence of all k_i such that the power 2^{k_i} occurs in the sequence $(t_j)_{j=1}^{J}$. We have $I \le J$. It then suffices to show the short and long variation bounds,

$$\sum_{i=0}^{I} \sum_{j:\, 2^{k_i} < t_j \le 2^{k_i+1}} \| M_{t_j}^{\chi}(f_0, f_1, f_2) - M_{t_{j-1}}^{\chi}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}, \tag{3-8}$$

$$\sum_{i=1}^{I} \| M_{2^{k_i}}^{\chi}(f_0, f_1, f_2) - M_{2^{k_{i-1}}}^{\chi}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}. \tag{3-9}$$

We first discuss the short variation (3-8). We define $T\chi(s) := (s\chi(s))'$, so that

$$(T\chi)_{(t)}(s) = -t\partial_t(\chi_{(t)}(s)),$$

and we will use T throughout the section. By the fundamental theorem of calculus and the Cauchy–Schwarz inequality, we have for $x \in \mathbb{R}^3$ and every $1 \le i \le I$,

$$\sum_{j:\, 2^{k_i} < t_j \le 2^{k_i+1}} \left| M_{t_j}^{\chi}(f_0, f_1, f_2)(x) - M_{t_{j-1}}^{\chi}(f_0, f_1, f_2)(x) \right|^2 \le \int_1^2 \left(M_{2^{k_i}t}^{T\chi}(f_0, f_1, f_2)(x) \right)^2 \frac{dt}{t}.$$

It then suffices to show

$$\sum_{i=0}^{I} \int_{\mathbb{R}^3} \int_1^2 \left(M_{2^{k_i}t}^{T\chi}(f_0, f_1, f_2)(x) \right)^2 \frac{dt}{t}\, dx \lesssim J^{\frac{1}{2}}.$$

Expanding the square and moving the integral in t outside, the left-hand side becomes

$$\int_1^2 \sum_{i=0}^{I} \int_{\mathbb{R}^5} \left[\prod_{n=0}^{2} f_n(x + ue_n) f_n(x + ve_n) \right] (T\chi)_{(2^{k_i}t)}(u) (T\chi)_{(2^{k_i}t)}(v)\, dx\, du\, dv\, \frac{dt}{t}. \tag{3-10}$$

The expression (3-10) takes the form

$$\int_1^2 \Lambda_{D_1,K_t}((f_s)_{s\in S})\,\frac{dt}{t}, \tag{3-11}$$

where for $s=(k,j)\in\{0,1,2\}\times\{0,1\}$ and $y=(y_0,y_1,y_2)$ we have set

$$f_s(y)=f_k\big(y-(y_0+y_1+y_2-y_k)e_k\big) \tag{3-12}$$

and

$$K_t(u,v):=\sum_{i=0}^{I}(T\chi)_{(2^{k_i}t)}(u)(T\chi)_{(2^{k_i}t)}(v). \tag{3-13}$$

Indeed, writing $x=(x_0,x_1,x_2)$ and changing variables

$$u=x_3^0-x_0-x_1-x_2,\qquad v=x_3^1-x_0-x_1-x_2, \tag{3-14}$$

we obtain with the projections Π_s of the datum D_1

$$f_{(k,0)}(\Pi_{(k,0)}(x,x_3^0,x_3^1))=f_k(x+ue_k),\qquad f_{(k,1)}(\Pi_{(k,1)}(x,x_3^0,x_3^1))=f_k(x+ve_k).$$

It suffices to prove bounds uniformly for fixed $t\in[1,2]$ on the integrand of (3-11). For this we apply Corollary 2.2 with the sequence $(k_i)_{i=0}^{I}$ and ϕ_i suitable multiples of $(T\chi)_{(2^{k_i}t)}$ and use

$$\operatorname{supp}(\widehat{T}\chi)\subset[-1,-2^{-1}]\cup[2^{-1},1],\qquad |T\chi(u)|\lesssim(1+|u|)^{-20}. \tag{3-15}$$

This proves (3-8).

Next, we prove the long variation bound (3-9). Recalling the universal pair (χ,ϕ), by the triangle inequality, it suffices to show

$$\sum_{i=1}^{I}\|M^{\chi}_{2^{k_{i-1}}}(f_0,f_1,f_2)-M^{\phi}_{2^{k_i}}(f_0,f_1,f_2)\|_2^2\lesssim J^{\frac12}, \tag{3-16}$$

$$\sum_{i=1}^{I}\|M^{\chi}_{2^{k_i}}(f_0,f_1,f_2)-M^{\phi}_{2^{k_i}}(f_0,f_1,f_2)\|_2^2\lesssim J^{\frac12}. \tag{3-17}$$

We first prove (3-16). We expand out the square of the L^2 norm to reduce matters to estimating

$$\sum_{i=1}^{I}\int_{\mathbb{R}^5}\Bigg[\prod_{n=0}^{2}f_n(x+ue_n)f_n(x+ve_n)\Bigg](\chi_{(2^{k_{i-1}})}-\phi_{(2^{k_i})})(u)(\chi_{(2^{k_{i-1}})}-\phi_{(2^{k_i})})(v)\,dx\,du\,dv. \tag{3-18}$$

Performing the same change of variables in the Brascamp–Lieb datum as in (3-10), we rewrite it as

$$\Lambda_{D_1,K}((f_s)_{s\in S}), \tag{3-19}$$

with

$$K(u,v)=\sum_{i=1}^{I}(\chi_{(2^{k_{i-1}})}-\phi_{(2^{k_i})})(u)(\chi_{(2^{k_{i-1}})}-\phi_{(2^{k_i})})(v).$$

We estimate this with Corollary 2.4 of Propositions 2.1 and 2.3, using that $\widehat{\chi}$ is a left window and $\widehat{\phi}$ is a right window, and after splitting the sum into even and odd indices j to assure spacing of the sequences k_j and l_j. This completes the discussion of (3-16). Similarly, estimating (3-17) reduces to estimating a form (3-19) with kernel

$$K(u,v)=\sum_{i=1}^{I}(\chi_{(2^{k_i})}-\phi_{(2^{k_i})})(u)(\chi_{(2^{k_i})}-\phi_{(2^{k_i})})(v).$$

This is done with Corollary 2.2. This completes the discussion of (3-17) and thus the discussion of (3-6).

Next, we consider the decaying lacunary pieces near the origin (3-7). We define

$$\varphi_{3,l}(x):=2^l(\varphi_{2,l})_{(2^{-l})}(x)$$

and we replace t_j by $2^l t_j$, using that the sequence t_j was arbitrary, to turn (3-7) into

$$\sum_{j=1}^{J}\|M_{t_j}^{\varphi_{3,l}}(f_0,f_1,f_2)-M_{t_{j-1}}^{\varphi_{3,l}}(f_0,f_1,f_2)\|_2^2\lesssim J^{\frac{1}{2}}. \tag{3-20}$$

Analogously to our discussion of (3-6), we pass to short and long variation. The short variation we estimate analogously using in place of (3-15)

$$\operatorname{supp}(\widehat{T}\varphi_{3,l})\subset[-1,-2^{-1}]\cup[2^{-1},1],\qquad |T\varphi_{3,l}(u)|\lesssim(1+|u|)^{-20}, \tag{3-21}$$

which follows because $\widehat{\varphi}-\widehat{\chi}$ vanishes at the origin. This completes the estimate for the short variation.

The long variation we expand similarly as (3-18) above into

$$\Lambda(f_0,f_1,f_2):=\sum_{i=1}^{I}\int_{\mathbb{R}^5}\Big[\prod_{n=0}^{2}f_n(x+ue_n)f_n(x+ve_n)\Big] \times\big((\varphi_{3,l})_{(2^{k_{i-1}})}-(\varphi_{3,l})_{(2^{k_i})}\big)(u)\big((\varphi_{3,l})_{(2^{k_{i-1}})}-(\varphi_{3,l})_{(2^{k_i})}\big)(v)\,dx\,du\,dv. \tag{3-22}$$

By the distributive law, (3-22) is the difference of the two terms of the form

$$\sum_{i=1}^{I}\int_{\mathbb{R}^5}\Big[\prod_{n=0}^{2}f_n(x+ue_n)f_n(x+ve_n)\Big](\varphi_{3,l})_{(2^{m_i})}(u)\big((\varphi_{3,l})_{(2^{k_{i-1}})}-(\varphi_{3,l})_{(2^{k_i})}\big)(v)\,dx\,du\,dv, \tag{3-23}$$

with $m_i=k_i$ and with $m_i=k_{i-1}$, respectively. We write (3-23) as

$$\sum_{i=1}^{I}\int_{\mathbb{R}^3}\Big[\int_{\mathbb{R}}\prod_{n=0}^{2}f_n(x+ue_n)(\varphi_{3,l})_{(2^{m_i})}(u)\,du\Big]\Big[\int_{\mathbb{R}}\prod_{n=0}^{2}f_n(x+ve_n)\big((\varphi_{3,l})_{(2^{k_{i-1}})}-(\varphi_{3,l})_{(2^{k_i})}\big)(v)\,dv\Big]dx$$

and apply the Cauchy–Schwarz inequality in x and in the summation. This gives

$$\Lambda(f_0,f_1,f_2)\le\tilde{\Lambda}(f_0,f_1,f_2)^{\frac{1}{2}}\Lambda(f_0,f_1,f_2)^{\frac{1}{2}}$$

with

$$\tilde{\Lambda}(f_0,f_1,f_2)=\Big[\sum_{i=1}^{I}\int_{\mathbb{R}^5}\Big[\prod_{n=0}^{2}f_n(x+ue_n)f_n(x+ve_n)\Big](\varphi_{3,l})_{(2^{m_i})}(u)(\varphi_{3,l})_{(2^{m_i})}(v)\,dx\,du\,dv\Big]^{\frac{1}{2}}. \tag{3-24}$$

By bootstrapping, it suffices to prove a bound on $\tilde{\Lambda}(f_0,f_1,f_2)$ in place of $\Lambda(f_0,f_1,f_2)$. This should be compared with the integrand in (3-10) for fixed t. By the same change of variables as there, (3-24) equals $\Lambda_{D_1,K}((f_s)_{s\in S})$ with

$$K(u,v)=\sum_{i=1}^{I}(\varphi_{3,l})_{(2^{m_i})}(u)(\varphi_{3,l})_{(2^{m_i})}(v). \tag{3-25}$$

Applying Corollary 2.2 of Proposition 2.1 yields a bound for this term and finishes the proof of (3-7). The assumptions of Corollary 2.2 are satisfied, which can be verified similarly as inequalities (3-21) observed earlier. This completes the proof of the estimate (3-3).

Now we prove (3-4). We write

$$\varphi_{0,k} = \mathbb{1}_{(-\infty,0)} * \theta_{(2^k)} = 2^k \tilde{\theta}_{(2^k)},$$

where $\tilde{\theta} := \mathbb{1}_{(-\infty,0)} * \theta$ is the primitive of θ. It has high-order decay since θ has integral zero. By rescaling, it suffices to show

$$\sum_{j=1}^{J} \|M_{t_j}^{\tilde{\theta}}(f_0, f_1, f_2) - M_{t_{j-1}}^{\tilde{\theta}}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}.$$

This now follows in the same way as (3-20), using

$$\operatorname{supp}(\hat{\tilde{\theta}}) \subset [-1, -2^{-2}] \cup [2^{-2}, 1]$$

and high-order decay of $\tilde{\theta}$. This completes the proof of (3-4).

It remains to prove (3-5). Define

$$\varphi_{4,k}(u) := 2^k \tilde{\theta}(u - 2^{-k}).$$

We have

$$\varphi_{1,k}(u) = (2^k \tilde{\theta}(u - 2^{-k}))_{(2^k)} = (\varphi_{4,k})_{(2^k)}(u).$$

By rescaling, it suffices to show

$$2^{-\gamma k} \sum_{j=1}^{J} \|M_{t_j}^{\varphi_{4,k}}(f_0, f_1, f_2) - M_{t_{j-1}}^{\varphi_{4,k}}(f_0, f_1, f_2)\|_2^2 \lesssim J^{\frac{1}{2}}.$$

We split into long and short variation as in (3-20). To estimate the short variation, we use the fundamental theorem of calculus and the Cauchy–Schwarz inequality, which yields the bound

$$\begin{aligned} &\sum_{i=0}^{I} \sum_{j:\, 2^{k_i} < t_j \le 2^{k_i+1}} 2^{-\gamma k} \|M_{t_j}^{\varphi_{4,k}}(f_0, f_1, f_2) - M_{t_{j-1}}^{\varphi_{4,k}}(f_0, f_1, f_2)\|_2^2 \\ &\lesssim \biggl[\biggl[\sum_{i=0}^{I} 2^{-(\gamma+1)k} \int_{\mathbb{R}^3} \int_1^2 \bigl(M_{2^{k_i} t}^{\varphi_{4,k}}(f_0, f_1, f_2)(x)\bigr)^2 \frac{dt}{t}\, dx\biggr] \\ &\qquad \times \biggl[\sum_{i=0}^{I} 2^{(1-\gamma)k} \int_{\mathbb{R}^3} \int_1^2 \bigl(M_{2^{k_i} t}^{T\varphi_{4,k}}(f_0, f_1, f_2)(x)\bigr)^2 \frac{dt}{t}\, dx\biggr]\biggr]^{\frac{1}{2}}. \end{aligned} \tag{3-26}$$

We are going to estimate each factor in the square brackets as $\lesssim J^{\frac{1}{2}}$. We begin with the first factor, that we expand as

$$\sum_{i=0}^{I} \int_{\mathbb{R}^5} \biggl[\prod_{n=0}^{2} f_n(x + ue_n) f_n(x + ve_n)\biggr] \biggl[2^{-(\gamma+1)k} \int_1^2 (\varphi_{4,k})_{(2^{k_i} t)}(u) (\varphi_{4,k})_{(2^{k_i} t)}(v) \frac{dt}{t}\biggr] dx\, du\, dv.$$

Similarly as (3-11) and (3-13), this takes the form

$$\Lambda_{D_1,K}((f_s)_{s\in S})$$

with

$$K(u,v)=\sum_{i=0}^{I}\biggl[\int_1^2 2^{-(\gamma+1)k}(\varphi_{4,k})_{(t)}(u)(\varphi_{4,k})_{(t)}(v)\,\frac{dt}{t}\biggr]_{(2^{k_i})}=:\sum_{i=0}^{I}\Phi_{(2^{k_i})}(u,v).$$

We apply Proposition 2.1 with $\lambda=2-\gamma>1$, using that Φ is symmetric and positive as a superposition of positive terms, and using

$$\operatorname{supp}(\widehat{\Phi})\subseteq([-1,-2^{-3}]\cup[2^{-3},1])^2$$

and the bound

$$\begin{aligned}|\Phi(u,v)|&\le 2^{-(\gamma+1)k}\int_1^2|\varphi_{4,k}(t^{-1}u)\varphi_{4,k}(t^{-1}v)|\,dt\\&\le 2^{(1-\gamma)k}\int_1^2|\tilde{\theta}(t^{-1}(u-t2^{-k}))\tilde{\theta}(t^{-1}(v-t2^{-k}))|\,dt\\&\lesssim 2^{(1-\gamma)k}\int_1^2(1+|u+v-t2^{1-k}|)^{-10}(1+|u-v|)^{-10}\,dt\\&\lesssim 2^{(2-\gamma)k}\int_{2^{1-k}}^{2^{2-k}}(1+|u+v-t|)^{-10}(1+|u-v|)^{-10}\,dt\\&\lesssim 2^{(2-\gamma)k}(1+2^k|u+v|)^{-10}(1+|u-v|)^{-10}.\end{aligned}\tag{3-27}$$

Here we estimated the integral for $|u+v|<2^{3-k}$ by the integral over $\mathbb{R}$ and for $|u+v|>2^{3-k}$ we estimated the integrand by its supremum norm. We used along the way decay estimates of $\tilde{\theta}$ it inherits from the window χ.

We turn to the second factor in (3-26). We proceed as above; in place of (3-27) we compute

$$2^{(1-\gamma)k}\biggl|\int_1^2 t\partial_t((\varphi_{4,k})_{(t)}(u))t\partial_t((\varphi_{4,k})_{(t)}(v))\,\frac{dt}{t}\biggr|$$
$$=2^{(3-\gamma)k}\biggl|\int_1^2 t\partial_t\bigl(t^{-1}\tilde{\theta}(t^{-1}(u-t2^{-k}))\bigr)t\partial_t\bigl(t^{-1}\tilde{\theta}(t^{-1}(v-t2^{-k}))\bigr)\,\frac{dt}{t}\biggr|.$$

Applying Leibniz and chain rules, most terms will be analogous to the above. However, when a derivative falls on $t2^{-k}$, we obtain a factor 2^{-k}. The worst term is the one where both derivatives fall on the $t2^{-k}$. Thus we get the estimate

$$\lesssim 2^{(1-\gamma)k}\int_1^2(1+|u+v-t2^{1-k}|)^{-10}(1+|u-v|)^{-10}\,\frac{dt}{t}.$$

As above, this is estimated by

$$\lesssim 2^{(2-\gamma)k}(1+2^k|u+v|)^{-10}(1+|u-v|)^{-10}.$$

To treat the long variation, we proceed as for (3-22), where after a bootstrapping estimate we are led to estimate, analogously to (3-25), $\Lambda_{D_1,K}((f_s)_{s\in S})$ with

$$K(u,v) = 2^{-\gamma k}\sum_{i=1}^{I}(\varphi_{4,k})_{(2^{m_i})}(u)(\varphi_{4,k})_{(2^{m_i})}(v).$$

Similarly as in (3-27) we estimate

$$\begin{aligned}2^{-\gamma k}|\varphi_{4,k}(u)\varphi_{4,k}(v)| &= 2^{(2-\gamma)k}|\tilde\theta(u-2^{-k})\tilde\theta(v-2^{-k})|\\ &\lesssim 2^{(2-\gamma)k}(1+|u+v-2^{1-k}|)^{-10}(1+|u-v|)^{-10}\\ &\lesssim 2^{(2-\gamma)k}(1+2^k|u+v|)^{-10}(1+|u-v|)^{-10}.\end{aligned}$$

Applying Proposition 2.1 again completes the proof of (3-5).

4. Proof of Proposition 2.1 using Propositions 2.3 and 2.5

Let $\lambda=\frac{3}{2}$. Let $k\le 0$, let J be a positive integer and $(k_j)_{j=1}^J$ a strictly increasing sequence of integers. By splitting into a hundred subsequences, using the triangle inequality to separate these sequences, we may assume $k_j+100\le k_{j+1}$ for $1\le j<J$.

Let Φ_j for $1\le j\le J$ be as in Proposition 2.1. In particular, $\widehat{\Phi}_j(\xi,-\xi)$ is continuous and even in ξ by the symmetry assumption on the kernel Φ_j. Furthermore, we claim that $\widehat{\Phi}_j(\xi,-\xi)$ is positive for all $\xi\in\mathbb{R}$. To see this, first apply Plancherel to the positivity assumption (2-5) in Proposition 2.1 to conclude

$$0\le\int_{\mathbb{R}^2}\hat f(-\xi)\overline{\hat f(\eta)}\widehat{\Phi}_j(\xi,\eta)\,d\xi\,d\eta$$

for all Schwartz functions f. Now we see the claim by using testing functions $\hat f$ which approximate the Dirac delta at ξ.

As $\|\Phi_j\|_1$ has a universal bound, for suitable universal constant c we have

$$\widehat{\Phi}_j(\xi,-\xi)\le c(\widehat{\phi}_{0,j}-\widehat{\phi}_{1,j})(\xi)^2$$

with even real functions $\phi_{0,j}$ and $\phi_{1,j}$, such that $(\phi_{0,j})_{2^{-k_j+25}}$ is a left window and $(\phi_{1,j})_{2^{-k_j-25}}$ is a right window. Moreover, there exists a real even function ψ_j such that

$$\widehat{\psi_j}(\xi)^2 := 2c(\widehat{\phi}_{0,j}-\widehat{\phi}_{1,j})(\xi)^2-\ \widehat{\Phi}_j(\xi,-\xi). \tag{4-1}$$

Namely, outside the support of $\xi\mapsto\widehat{\Phi}_j(\xi,-\xi)$, the function $\widehat{\psi_j}$ can be chosen to equal $\sqrt{2c}(\widehat{\phi}_{0,j}-\widehat{\phi}_{1,j})$, while on a neighborhood of this support, the function on the right-hand side is at least c and thus has square root. The function $(\widehat{\psi_j})_{(2^{-k_j+25})}$ has support in $[-1,1]$. To understand derivative bounds for this function, let $F(\xi)=(\widehat{\Phi}_j)_{(2^{-k_j})}(\xi,-\xi)$. Then we have, for $0\le a\le 8$,

$$|F^{(a)}(\xi)| = \left|(-2\pi i)^a\int_{\mathbb{R}^2}(\Phi_j)_{(2^{-k_j})}(u,v)(u-v)^a e^{-2\pi i\xi(u-v)}\,du\,dv\right|\lesssim 2^{(\lambda-1)k},$$

by (2-6). Thus,

$$|((\widehat{\psi_j})_{(2^{-k_j})})^{(a)}|\lesssim 1, \tag{4-2}$$

as one can see outside the support of $(\widehat{\Phi}_j)_{(2^{-k_j})}$ from bounds for derivatives of the windows and on the support using a lower bound on the right-hand side of (4-1) and upper bounds on the derivative of the right-hand side of (4-1).

To show a bound on $\Lambda_{D_1,K}((f_s)_{s\in S})$ with $K=\sum_{j=1}^J \Phi_j$, which is positive, it suffices to show a bound on $\Lambda_{D_1,K_0}((f_s)_{s\in S})$ with

$$K_0=\sum_{j=1}^J \Phi_j+\psi_j\otimes\psi_j$$

because the form associated with the datum D_1 and the difference K_0-K is positive as well.

By Proposition 2.3, the form $\Lambda_{D_1,K_1}((f_s)_{s\in S})$ is bounded, where

$$K_1=2c\sum_{j=1}^J(\phi_{0,j}-\phi_{1,j})\otimes(\phi_{0,j}-\phi_{1,j}).$$

Hence it suffices to prove a bound on $\Lambda_{D_1,K_3}((f_s)_{s\in S})$, where $K_3=K_0-K_1$.

This is done by an application of Proposition 2.5. Note that we have on the diagonal

$$\widehat{K}_3(\xi,-\xi)=\sum_{j=1}^J \widehat{\Phi}_j(\xi,-\xi)+\widehat{\psi}_j(\xi)^2-2c(\widehat{\phi}_{0,j}-\widehat{\phi}_{1,j})(\xi)^2=0.$$

We verify the remaining assumptions of Proposition 2.5 for

$$\Psi_j:=\Phi_j+\psi_j\otimes\psi_j-2c(\phi_{0,j}-\phi_{1,j})\otimes(\phi_{0,j}-\phi_{1,j}).$$

We have

$$\begin{aligned}\operatorname{supp}(\widehat{\Phi}_j)&\subseteq\big([-2^{-k_j+20},-2^{-k_j-20}]\cup[2^{-k_j-20},2^{-k_j+20}]\big)^2,\\ \operatorname{supp}\big((\widehat{\phi}_{0,j}-\widehat{\phi}_{1,j})\otimes(\widehat{\phi}_{0,j}-\widehat{\phi}_{1,j})\big)&\subseteq\big([-2^{-k_j+25},-2^{-k_j-26}]\cup[2^{-k_j-26},2^{-k_j+25}]\big)^2,\\ \operatorname{supp}(\widehat{\psi}_j\otimes\widehat{\psi}_j)&\subseteq\big([-2^{-k_j+25},-2^{-k_j-26}]\cup[2^{-k_j-26},2^{-k_j+25}]\big)^2.\end{aligned}$$

Thus,

$$\operatorname{supp}(\widehat{\Psi}_j)\subseteq\big\{(\xi,\eta)\in\mathbb{R}^2:2^{-k_j-30}<|(\xi,\eta)|\le 2^{-k_j+30}\big\}.$$

Note also that, using in particular (4-2),

$$\begin{aligned}|(\Phi_j)_{(2^{-k_j})}(u,v)|&\lesssim 2^{\lambda k}(1+2^k|u+v|)^{-10}(1+|u-v|)^{-10},\\ \big|((\phi_{0,j}-\phi_{1,j})\otimes(\phi_{0,j}-\phi_{1,j}))_{(2^{-k_j})}(u,v)\big|&\lesssim(1+|u+v|)^{-4}(1+|u-v|)^{-4},\\ |(\psi_j\otimes\psi_j)_{(2^{-k_j})}(u,v)|&\lesssim(1+|u+v|)^{-4}(1+|u-v|)^{-4}.\end{aligned}$$

Hence

$$|(\Psi_j)_{(2^{-k_j})}(u,v)|\lesssim 2^{\lambda k}(1+2^k|u+v|)^{-4}(1+|u-v|)^{-4}+(1+|u+v|)^{-4}(1+|u-v|)^{-4}.$$

The final claim now follows from Proposition 2.5.

5. Proof of Proposition 2.3 using Propositions 2.5 and 2.6

Let J be a positive integer and $(k_j)_{j=1}^J$ and $(l_j)_{j=1}^J$ two finite sequences of integers with $k_j + 10 < l_j$ for $1 \le j \le J$ and $l_j < k_{j+1}$ for $1 \le j < J-1$. By splitting the sequence into subsequences of even and odd j if necessary, we may assume without loss of generality that $l_j + 10 < k_{j+1}$ for each $1 \le j < J$. Assume a tuple $(f_s)_{s\in S}$ as in (2-3) and (2-4) is given.

Assume we are given $\phi_{0,j}$ and $\phi_{1,j}$ for each j such that $(\phi_{0,j})_{(2^{-k_j})}$ is a left window and $(\phi_{1,j})_{(2^{-l_j})}$ is a right window. Pick corresponding functions $\psi_{0,j}$ and $\psi_{1,j}$ so that the rescaled functions give universal pairs, and hence

$$(1-\widehat{\phi}_{1,j})^2 + (\widehat{\psi}_{0,j})^2 = 1, \tag{5-1}$$

$$(\widehat{\phi}_{0,j})^2 + (1-\widehat{\psi}_{1,j})^2 = 1. \tag{5-2}$$

Then

$$(1-\widehat{\psi}_{1,1})^2 + \sum_{j=1}^{J} (\widehat{\phi}_{0,j} - \widehat{\phi}_{1,j})^2 + \sum_{j=1}^{J-1} (\widehat{\psi}_{0,j} - \widehat{\psi}_{1,j+1})^2 + (\widehat{\psi}_{0,J})^2 = 1. \tag{5-3}$$

To see this, note that at every point at most one of the functions $\widehat{\phi}_{0,j}$, $\widehat{\phi}_{1,j}$, $1 \le j \le J$, is neither 0 nor 1, and the functions $\widehat{\psi}_{0,j}$, $\widehat{\psi}_{1,j}$ are neither 0 nor 1 precisely when the respective function $\widehat{\phi}_{1,j}$, $\widehat{\phi}_{0,j}$ is not 0 or 1. Therefore, at any point at most one pair $(\widehat{\psi}_{0,j}, \widehat{\phi}_{1,j})$ or $(\widehat{\psi}_{1,j}, \widehat{\phi}_{0,j})$ takes values other than 0 and 1, and we can apply (5-1) or (5-2), respectively.

As $\Lambda_{D_1,K}((f_s)_{s\in S})$ in Proposition 2.3 is positive, it suffices to estimate its sum with another positive term, and thus it suffices to estimate $\Lambda_{D_1,K_1}((f_s)_{s\in S})$ with

$$\begin{aligned}\widehat{K}_1(\xi,\eta) = (1-\widehat{\psi}_{1,1})(\xi)(1-\widehat{\psi}_{1,1})(\eta) + \sum_{j=1}^{J} (\widehat{\phi}_{0,j} - \widehat{\phi}_{1,j})(\xi)(\widehat{\phi}_{0,j} - \widehat{\phi}_{1,j})(\eta) \\ + \sum_{j=1}^{J-1} (\widehat{\psi}_{0,j} - \widehat{\psi}_{1,j+1})(\xi)(\widehat{\psi}_{0,j} - \widehat{\psi}_{1,j+1})(\eta) + (\widehat{\psi}_{0,J})(\xi)(\widehat{\psi}_{0,J})(\eta).\end{aligned}$$

This can be rewritten in a more compressed form as

$$\widehat{K}_1 = \sum_{j=0}^{2J} (\widehat{\varphi}_{0,j} - \widehat{\varphi}_{1,j}) \otimes (\widehat{\varphi}_{0,j} - \widehat{\varphi}_{1,j}),$$

where $\widehat{\varphi}_{0,0} = 1$, for $1 \le j \le J$,

$$\varphi_{0,2j-1} = \phi_{0,j}, \quad \varphi_{0,2j} = \psi_{0,j}, \quad \varphi_{1,2j-1} = \phi_{1,j}, \quad \varphi_{1,2j-2} = \psi_{1,j},$$

and $\varphi_{1,2J} = 0$. Define for $1 \le j \le J$

$$m_{2j-2} = k_j \quad \text{and} \quad m_{2j-1} = l_j.$$

Observe that for each $1 \le j \le 2J-1$ we have that $(\varphi_{0,j})_{(2^{-m_{j-1}})}$ is a left window and $(\varphi_{1,j})_{(2^{-m_j})}$ is a right window.

In order to apply Proposition 2.6, we introduce for $0 \le j \le 2J$ the functions

$$\varphi_{2,j} = (\varphi_{1,j})_{(2^{-4})}.$$

Observe that $(\varphi_{2,j})_{2^{4-m_j}}$ is a right window whenever $0 \le j \le 2J-1$. We write for $\widehat{K}_1$

$$-\sum_{j=0}^{2J} (\widehat{\varphi}_{0,j} - \widehat{\varphi}_{2,j}) \otimes \widehat{\varphi}_{1,j} + \widehat{\varphi}_{1,j} \otimes (\widehat{\varphi}_{0,j} - \widehat{\varphi}_{2,j}) \tag{5-4}$$

$$-\sum_{j=0}^{2J} (\widehat{\varphi}_{2,j} - \widehat{\varphi}_{1,j}) \otimes \widehat{\varphi}_{1,j} + \widehat{\varphi}_{1,j} \otimes (\widehat{\varphi}_{2,j} - \widehat{\varphi}_{1,j}) \tag{5-5}$$

$$+\sum_{j=0}^{2J} \widehat{\varphi}_{0,j} \otimes \widehat{\varphi}_{0,j} - \widehat{\varphi}_{1,j} \otimes \widehat{\varphi}_{1,j}. \tag{5-6}$$

In (5-4), the bound for the sum of these terms over $1 \le j \le 2J-1$ follows from Proposition 2.6, applied to the sequence $(m_j)_{j=0}^{2J-1}$ and the rescaled windows $\varphi_{0,j}$, $\varphi_{1,j}$, $\varphi_{2,j}$ for $1 \le j \le 2J-1$. The term for $j = 2J$ in (5-4) vanishes. To deal with the term for $j = 0$ in (5-4), we use $\widehat{\varphi}_{0,0} = 1$ and rewrite this term as

$$-\widehat{\varphi}_{1,0}(\eta) + \widehat{\varphi}_{2,0}(\xi)\widehat{\varphi}_{1,0}(\eta) + \widehat{\varphi}_{1,0}(\xi) - \widehat{\varphi}_{1,0}(\xi)\widehat{\varphi}_{2,0}(\eta).$$

Denoting by f_k, $k = 0, 1, 2$, the functions defined via (3-12) and using the change of variables as in (3-14), we estimate

$$\left|\Lambda_{D_1,\varphi_{1,0}\otimes\delta}((f_s)_{s\in S})\right| = \left|\int_{\mathbb{R}^4} \left[\prod_{k=0}^{2} f_k(x+ue_k) f_k(x)\right] \varphi_{1,0}(u)\,dx\,du\right| \le \|\varphi_{1,0}\|_1 \lesssim 1,$$

where for a fixed u we used Hölder's inequality in x and δ denotes the Dirac delta at the origin. Similarly,

$$\left|\Lambda_{D_1,\varphi_{1,0}\otimes\varphi_{2,0}}((f_s)_{s\in S})\right| = \left|\int_{\mathbb{R}^5} \left[\prod_{k=0}^{2} f_k(x+ue_k) f_k(x+ve_k)\right] \varphi_{1,0}(u)\varphi_{2,0}(v)\,dx\,du\,dv\right|$$
$$\le \|\varphi_{1,0}\|_1 \|\varphi_{2,0}\|_1 \lesssim 1.$$

By symmetry, this bounds the form associated with the $j = 0$ summand in (5-4).

It remains to estimate the form associated with K_2 where $\widehat{K}_2$ is the sum of (5-5), (5-6). As $\widehat{K}_1$ is constant 1 on the diagonal $\xi + \eta = 0$ by (5-3) and the stick terms (5-4) vanish on this diagonal, the function $\widehat{K}_2$ is still constant 1 on this diagonal.

We define K_3 by $\widehat{K}_3 := \widehat{K}_2 - 1$. It suffices to prove bounds for the form associated with K_3, because $K_2 - K_3$ is the Dirac delta and

$$\left|\Lambda_{D_1,K_2-K_3}((f_s)_{s\in S})\right| = \left|\int_{\mathbb{R}^3} \prod_{k=0}^{2} f_k^2(x)\,dx\right| \le 1,$$

where the functions f_k are as in (3-12). We rewrite $\widehat{K}_3$ as

$$-\sum_{j=0}^{2J}(\widehat{\varphi}_{2,j}-\widehat{\varphi}_{1,j})\otimes\widehat{\varphi}_{1,j}+\widehat{\varphi}_{1,j}\otimes(\widehat{\varphi}_{2,j}-\widehat{\varphi}_{1,j}) \tag{5-7}$$

$$+\sum_{j=1}^{2J}\widehat{\varphi}_{0,j}\otimes\widehat{\varphi}_{0,j}-\widehat{\varphi}_{1,j-1}\otimes\widehat{\varphi}_{1,j-1}, \tag{5-8}$$

where we have reshuffled (5-6) and used $\widehat{\phi}_{0,0}=1$ and $\widehat{\phi}_{1,2J}=0$. Bounds for the sum of (5-7) and (5-8) follow from Proposition 2.5. Indeed, for each $0\le j\le 2J$,

$$\operatorname{supp}((\widehat{\varphi}_{2,j}-\widehat{\varphi}_{1,j})\otimes\widehat{\varphi}_{1,j})\subseteq\big([-2^{-m_j+4},-2^{-m_j-1}]\cup[2^{-m_j-1},2^{-m_j+4}]\big)\times[-2^{-m_j},2^{-m_j}].$$

By symmetry, the j-th summand in (5-7) is supported in

$$\big\{(\xi,\eta)\in\mathbb{R}^2:2^{-m_j-30}<|(\xi,\eta)|\le 2^{-m_j+30}\big\}=:A.$$

The j-th summand also satisfies a bound by

$$\big|(\varphi_{2,j}-\varphi_{1,j})\otimes\varphi_{1,j}+\varphi_{1,j}\otimes(\varphi_{2,j}-\varphi_{1,j})\big|_{(2^{-m_j})}(u,v)\lesssim(1+|u+v|)^{-4}(1+|u-v|)^{-4}$$

due to the functions being windows.

Similarly, for $0\le j\le 2J-1$ we have

$$\operatorname{supp}\big(\widehat{\varphi}_{0,j+1}\otimes\widehat{\varphi}_{0,j+1}-\widehat{\varphi}_{1,j}\otimes\widehat{\varphi}_{1,j}\big)\subseteq[-2^{-m_j},2^{-m_j}]^2\setminus[-2^{-m_j-1},2^{-m_j-1}]^2\subseteq A$$

and the decay

$$\big|\varphi_{0,j+1}\otimes\varphi_{0,j+1}-\varphi_{1,j}\otimes\varphi_{1,j}\big|_{(2^{-m_j})}(u,v)\lesssim(1+|u+v|)^{-4}(1+|u-v|)^{-4}.$$

Thus, bounds for Λ_{D_1,K_3} follow from Proposition 2.5.

6. Proof of Proposition 2.5 using Lemma 3 in [Durcik and Thiele 2020]

Given a regular 3×3 matrix A, let D_A be the datum defined in (2-20). We recall the following lemma, which is a special instance of a more general result proved in [Durcik and Thiele 2020].

Lemma 6.1 [Durcik and Thiele 2020, Lemma 3]. *For all $0<\varepsilon<1$, there exists a constant C such that the following holds.*

Let A be a regular 3×3 matrix which differs from $-I$ by at most one row and satisfies

$$|\det A|>\varepsilon \quad \textit{and} \quad \|A\|_{\mathrm{HS}}\le\varepsilon^{-1}, \tag{6-1}$$

where $\|A\|_{\mathrm{HS}}$ stands for the Hilbert–Schmidt norm of A. With S as in the datum D_A, let $(f_s)_{s\in S}$ be a tuple of real-valued Schwartz functions such that $\|f_s\|_8=1$ for all $s\in S$. Let $i=1,2,3$, and let K be the kernel satisfying

$$K(\Pi x)=\int_0^\infty\int_{\mathbb{R}^3}(\partial_i\partial_{i+3}g)_{(t)}\big(x+((-Ap^T)^T,p)\big)\,dp\,\frac{dt}{t}. \tag{6-2}$$

Then

$$|\Lambda_{D_A,K}((f_s)_{s\in S})|\le C.$$

Proof of Proposition 2.5. Let $\lambda = \frac{3}{2}$. Let $k \le 0$ be given. Let an integer $J \ge 1$ and a strictly increasing sequence $(k_j)_{j=1}^J$ of integers be given. Let $(\Phi_j)_{j=1}^J$ and K be given as in the proposition. Let $(f_s)_{s\in S}$ be given as in (2-3) and (2-4). Set $f_k := f_{(k,0)} = f_{(k,1)}$ for each $k = 0, 1, 2$.

Let $\theta : \mathbb{R} \to \mathbb{R}$ be a function whose Fourier transform is supported in $\left[-2, -\frac{1}{2}\right] \cup \left[\frac{1}{2}, 2\right]$ and whose derivatives up to order 8 are $\lesssim 1$. Assume further that

$$\int_0^\infty \hat{\theta}(r\xi)\,\frac{dr}{r} = 1$$

for all $\xi \ne 0$. We do the two-parameter lacunary decomposition of $\widehat{K}$ in directions $\xi + \eta$ and $\xi - \eta$ and collect these pieces into lacunary cones away from the line $\xi + \eta = 0$ centered at the origin. In detail, we write

$$\widehat{K}(\xi, \eta) = \int_0^\infty \widehat{K}^{(z)}(\xi, \eta)\,\frac{dz}{z} \tag{6-3}$$

with

$$\widehat{K}^{(z)}(\xi, \eta) = \int_0^\infty \widehat{K}(\xi, \eta)\hat{\theta}(t(\xi - \eta))\hat{\theta}(z^{-1}t(\xi + \eta))\,\frac{dt}{t}. \tag{6-4}$$

We break the integral in (6-3) into the integrals over the domains $(0, 1)$ and $(1, \infty)$ and do the estimates for these integrals separately. We begin with the case $z \in (0, 1)$. Here we do an estimate for each z separately and show for all $z < 1$ that

$$|\Lambda_{D_1, K^{(z)}}((f_s)_{s\in S})| \lesssim z^{(\lambda-1)^2/(2\lambda)} J^{\frac{1}{2}}, \tag{6-5}$$

which is an integrable upper bound with respect to the measure dz/z. Fix $z \in (0, 1)$.

Let g be the one-dimensional Gaussian and let $h = g'$. Set $\widehat{\omega} = (\hat{h})^{-1}\hat{\theta}$. The function $\widehat{\omega}$ satisfies similar support and derivative estimates as $\hat{\theta}$ since $\hat{h}$ and its derivatives are essentially constant on the support of $\hat{\theta}$. In addition, let $\widehat{\phi}$ be a function supported in the annulus $\frac{1}{16} \le |(\xi, \eta)| \le 16$ such that its derivatives up to order 8 are $\lesssim 1$ and $\widehat{\phi}(\xi, \eta)\hat{g}(\xi)\hat{g}(\eta) = 1$ if $\frac{1}{8} \le |(\xi, \eta)| \le 8$. Then, for all $\xi, \eta \in \mathbb{R}$,

$$\hat{\theta}(\xi - \eta)\hat{\theta}(z^{-1}(\xi + \eta)) = \hat{\theta}(\xi - \eta)\widehat{\omega}(z^{-1}(\xi + \eta))\hat{h}(z^{-1}(\xi + \eta))\widehat{\phi}(\xi, \eta)\hat{g}(\xi)\hat{g}(\eta). \tag{6-6}$$

Note that this equality holds since the left-hand side is supported in the set where

$$\tfrac{1}{8} \le |(\xi, \eta)| \le 8. \tag{6-7}$$

Indeed, on the support of the left-hand side of (6-6) we have $|\xi + \eta| \le 2z \le 2$ and $\frac{1}{2} \le |\xi - \eta| \le 2$. This yields (6-7).

For $z \in (0, 1)$ and $t > 0$ we define the function $w^{z,t}$ via

$$\widehat{w}^{z,t}(\xi, \eta) = \widehat{K}(t^{-1}(\xi, \eta))\widehat{\phi}(\xi, \eta)\hat{\theta}(\xi - \eta)\widehat{\omega}(z^{-1}(\xi + \eta)). \tag{6-8}$$

Let Π be the projection associated with the datum D_1. Using the Fourier inversion formula and equations (6-4), (6-6) and (6-8), we write $K^{(z)}(\Pi x)$ as

$$\int_0^\infty\!\!\int_{\mathbb{R}^2} \widehat{w}^{z,t}(t(\xi, \eta))\hat{h}(z^{-1}t(\xi+\eta))\hat{g}(t\xi)\hat{g}(t\eta)\exp\bigl(2\pi i\bigl(\xi(x_3^0-x_0-x_1-x_2)+\eta(x_3^1-x_0-x_1-x_2)\bigr)\bigr)\,d\xi\,d\eta\,\frac{dt}{t}. \tag{6-9}$$

Since the Fourier transform of $w^{z,t}$ is supported in the set where $\frac{1}{8} \le |(\xi, \eta)| \le 8$, we observe that $w^{z,t}$ vanishes unless t is in the set

$$M := \bigcup_{j=1}^{J} [2^{k_j - 33}, 2^{k_j + 33}].$$

We may thus restrict the region of t-integration in (6-9) to M. Further, we may interpret the inner integral in (6-9) as the integral of the Fourier transform of the function

$$(y_0, y_1, y_2, y_3, y_4) \mapsto w^{z,t}_{(t)}(y_0 + x_0 + x_1, y_1 + x_0 + x_1) h_{(z^{-1}t)}(y_2 + x_2) g_{(t)}(y_3 + x_3^0) g_{(t)}(y_4 + x_3^1)$$

over the hyperplane

$$\{(-\xi, -\eta, -\xi - \eta, \xi, \eta) : \xi \in \mathbb{R},\ \eta \in \mathbb{R}\}.$$

It is therefore up to universal multiplicative constant equal to the integral of the function itself over the orthogonal complement

$$\{(p + q - r, q - r, r, p + q, q) : p, q, r \in \mathbb{R}\}.$$

The form $\Lambda_{D_1, K^{(z)}}((f_s)_{s \in S})$ can then be rewritten as

$$\int_M \int_{\mathbb{R}^8} \Big[\prod_{s \in S} f_s(\Pi_s x) \Big] w^{z,t}_{(t)}(x_0 + x_1 + p + q - r, x_0 + x_1 + q - r) \\ \times h_{(z^{-1}t)}(x_2 + r) g_{(t)}(x_3^0 + p + q) g_{(t)}(x_3^1 + q)\, dx\, dp\, dq\, dr\, \frac{dt}{t}. \tag{6-10}$$

We write the integral in x_2 as the innermost and use the Cauchy–Schwarz inequality in the remaining variables. This bounds (6-10) by the geometric mean of

$$\int_M \int_{\mathbb{R}^7} \Big[\prod_{i=0,1} |f_2(x_0, x_1, x_3^i)|^2 \Big] \big| w^{z,t}_{(t)}(x_0 + x_1 + p + q - r, x_0 + x_1 + q - r) \big| \\ \times g_{(t)}(x_3^0 + p + q) g_{(t)}(x_3^1 + q)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp\, dq\, dr\, \frac{dt}{t} \tag{6-11}$$

and

$$\int_0^\infty \int_{\mathbb{R}^7} \Big[\int_{\mathbb{R}} \Big[\prod_{i=0,1} f_0(x_3^i, x_1, x_2) f_1(x_0, x_3^i, x_2) \Big] h_{(z^{-1}t)}(x_2 + r)\, dx_2 \Big]^2 \\ \times \big| w^{z,t}_{(t)}(x_0 + x_1 + p + q - r, x_0 + x_1 + q - r) \big| \\ \times g_{(t)}(x_3^0 + p + q) g_{(t)}(x_3^1 + q)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp\, dq\, dr\, \frac{dt}{t}. \tag{6-12}$$

In order to bound (6-11) and (6-12), we prove a pointwise estimate for $w^{z,t}$. We first claim

$$|w^{z,t}(u, v)| \lesssim z^{\lambda}. \tag{6-13}$$

To verify the claim, we observe that since $\widehat{K}$ vanishes on the diagonal $\xi+\eta=0$, the function $\widehat{K}_{(t^{-1})}*\phi$ has the same property. Therefore

$$\begin{aligned}
|\widehat{K}_{(t^{-1})}*\phi(\xi,\eta)| &= \bigl|\widehat{K}_{(t^{-1})}*\phi(\xi,\eta)-\widehat{K}_{(t^{-1})}*\phi\bigl((\xi-\eta)/2,-(\xi-\eta)/2\bigr)\bigr| \\
&= \left|\int_{\mathbb{R}^2} K_{(t^{-1})}*\phi(u,v)e^{-\pi i(\xi-\eta)(u-v)}(e^{-\pi i(\xi+\eta)(u+v)}-1)\,du\,dv\right| \\
&\lesssim \int_{\mathbb{R}^2}\bigl|K_{(t^{-1})}*\phi(u,v)\bigr|\min\{|\xi+\eta|\,|u+v|,1\}\,du\,dv \\
&\le |\xi+\eta|^{\lambda-1}\int_{\mathbb{R}^2}\bigl|K_{(t^{-1})}*\phi(u,v)\bigr||u+v|^{\lambda-1}\,du\,dv, \qquad (6\text{-}14)
\end{aligned}$$

as $\lambda-1\in(0,1)$. We observe that

$$|K_{(t^{-1})}*\phi(u,v)|\lesssim 2^{\lambda k}(1+2^k|u+v|)^{-4}(1+|u-v|)^{-4}+(1+|u+v|)^{-4}(1+|u-v|)^{-4}, \qquad (6\text{-}15)$$

thanks to the derivative estimates on ϕ, to the support properties of $\widehat{\phi}$ and $\widehat{\Phi}_j$ and to (2-8). Therefore,

$$\begin{aligned}
&\int_{\mathbb{R}^2}\bigl|K_{(t^{-1})}*\phi(u,v)\bigr||u+v|^{\lambda-1}\,du\,dv \\
&\quad\lesssim 2^k\int_{\mathbb{R}^2}(1+2^k|u+v|)^{\lambda-5}(1+|u-v|)^{-4}\,du\,dv+\int_{\mathbb{R}^2}(1+|u+v|)^{\lambda-5}(1+|u-v|)^{-4}\,du\,dv\lesssim 1.
\end{aligned}$$

Combining this with (6-14) and passing to $w^{z,t}$, we thus obtain

$$|\widehat{w}^{z,t}(\xi,\eta)|\lesssim z^{\lambda-1}.$$

Estimating the Fourier inversion formula by $L^1\to L^\infty$ bounds, inequality (6-13) follows.

We note that the right-hand side of (6-13) has the desired decay as z tends to 0, however, it does not have a good behavior with respect to (u,v). We therefore derive a yet another estimate for $w^{z,t}$ in which the right-hand side possesses merely L^1 scaling in z but decays sufficiently fast as $|(u,v)|$ tends to infinity. We set

$$F(u,v)=\omega_{(z^{-1})}((u+v)/2)\theta((u-v)/2).$$

By (6-8), we have $w^{z,t}=K_{(t^{-1})}*\phi*F$. Recall that the functions $\widehat{\omega}$ and $\hat{\theta}$ are supported in $[-2,2]$ and have derivatives up to order 8 bounded by $\lesssim 1$. Using (6-15), we therefore obtain

$$|w^{z,t}(u,v)|\lesssim 2^{\lambda k}(1+2^k|u+v|)^{-4}(1+|u-v|)^{-4}+z(1+z|u+v|)^{-4}(1+|u-v|)^{-4}\quad\text{if } 2^k\le z \qquad (6\text{-}16)$$

and

$$|w^{z,t}(u,v)|\lesssim z(1+z|u+v|)^{-4}(1+|u-v|)^{-4}\quad\text{if } z\le 2^k. \qquad (6\text{-}17)$$

Finally, we write $|w^{z,t}|=|w^{z,t}|^{(\lambda-1)/(2\lambda)}|w^{z,t}|^{(\lambda+1)/(2\lambda)}$ and use the estimate (6-13) for the first factor and the estimates (6-16) and (6-17) for the second factor. This yields the desired bounds

$$|w^{z,t}(u,v)|\lesssim z^{\frac{(\lambda-1)^2}{2\lambda}}[z(1+z|u+v|)^{-2-\frac{2}{\lambda}}+2^k(1+2^k|u+v|)^{-2-\frac{2}{\lambda}}](1+|u-v|)^{-2-\frac{2}{\lambda}} \qquad (6\text{-}18)$$

if $2^k \le z$, and

$$|w^{z,t}(u,v)| \lesssim z^{\frac{(\lambda-1)^2}{2\lambda}} z(1+z|u+v|)^{-2-\frac{2}{\lambda}}(1+|u-v|)^{-2-\frac{2}{\lambda}} \quad \text{if } z \le 2^k. \tag{6-19}$$

Having inequalities (6-18) and (6-19) at our disposal, we proceed to bound the term (6-11). We observe that this term can be written as

$$\int_M \int_{\mathbb{R}^5} \Big[\prod_{i=0,1} |f_2(x_0,x_1,x_3^i)|^2 \Big] \times [|w^{z,t}| * (g \otimes g)]_{(t)}\big(x_3^0 - x_0 - x_1 + r, x_3^1 - x_0 - x_1 + r\big)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dr\, \frac{dt}{t}. \tag{6-20}$$

Applying the Cauchy–Schwarz inequality, we bound (6-20) with

$$v_{z,t} := [|w^{z,t}| * (g \otimes g)]_{(t)}$$

by

$$\int_M \prod_{i=0,1} \Big[\int_{\mathbb{R}^5} |f_2(x_0,x_1,x_3^i)|^4 v_{z,t}(x_3^0 - x_0 - x_1 + r, x_3^1 - x_0 - x_1 + r)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dr \Big]^{\frac{1}{2}} \frac{dt}{t}.$$

The product of the square roots of the integrals for $i = 0, 1$ equals

$$\|f_2\|_4^4 \|v_{z,t}\|_1 \lesssim z^{\frac{(\lambda-1)^2}{2\lambda}}.$$

The last identity can be seen by integrating first in x_3^{1-i} and then in r to obtain the L^1 norm of $v_{z,t}$. What remains is then the L^4 norm of f_2 raised to the fourth power. Using that $\int_M dt/t \lesssim J$, we deduce that (6-11) is bounded by a multiple of

$$z^{\frac{(\lambda-1)^2}{2\lambda}} J.$$

We next focus on the term (6-12). Using the estimates (6-18) and (6-19), bounding the form (6-12) reduces to estimating

$$\int_0^\infty \int_{\mathbb{R}^7} \Big[\int_{\mathbb{R}} \Big[\prod_{i=0,1} f_0(x_3^i, x_1, x_2) f_1(x_0, x_3^i, x_2) \Big] h_{(z^{-1}t)}(x_2 + r)\, dx_2 \Big]^2 \times t^{-1}\gamma\big(1 + t^{-1}\gamma|x_0 + x_1 + p/2 + q - r|\big)^{-2-\frac{2}{\lambda}} t^{-1}(1+t^{-1}|p|)^{-2-\frac{2}{\lambda}} \times g_{(t)}(x_3^0 + p + q) g_{(t)}(x_3^1 + q)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp\, dq\, dr\, \frac{dt}{t},$$

where $\gamma = z$, or $\gamma = 2^k$ if $2^k \le z$. We will prove a bound independent of z and k, which will bound (6-12) by $\lesssim z^{(\lambda-1)^2/(2\lambda)}$ thanks to the extra factor $z^{(\lambda-1)^2/(2\lambda)}$ in (6-18) and (6-19).

We dominate

$$\begin{aligned} t^{-1}\gamma\big(1 + t^{-1}\gamma|x_0 + x_1 + p/2 + q - r|\big)^{-2-\frac{2}{\lambda}} t^{-1}(1+t^{-1}|p|)^{-2-\frac{2}{\lambda}} \\ \lesssim t^{-2}\gamma\big(1 + t^{-1}\big|(\gamma(x_0 + x_1 + p/2 + q - r), 2p)\big|\big)^{-2-\frac{2}{\lambda}} \\ \lesssim \int_2^\infty g_{(\alpha\gamma^{-1}t)}(x_0 + x_1 + p/2 + q - r) g_{(\alpha t)}(2p)\, \frac{d\alpha}{\alpha^{1+2/\lambda}}. \end{aligned}$$

It thus suffices to estimate the form

$$\int_0^\infty\int_{\mathbb{R}^7}\biggl[\int_{\mathbb{R}}\biggl[\prod_{i=0,1} f_0(x_3^i,x_1,x_2)f_1(x_0,x_3^i,x_2)\biggr]h_{(z^{-1}t)}(x_2+r)\,dx_2\biggr]^2$$
$$\times g_{(\alpha\gamma^{-1}t)}(x_0+x_1+p/2+q-r)g_{(\alpha t)}(2p)g_{(t)}(x_3^0+p+q)g_{(t)}(x_3^1+q)\,dx_0\,dx_1\,dx_3^0\,dx_3^1\,dp\,dq\,dr\,\frac{dt}{t} \quad (6\text{-}21)$$

with a bound $\lesssim\alpha$ and then integrate over α, using that $2/\lambda>1$. We claim that

$$\int_{\mathbb{R}} g_{(\alpha\gamma^{-1}t)}(x_0+x_1+p/2+q-r)g_{(\alpha t)}(2p)g_{(t)}(x_3^0+p+q)\,dp$$
$$\lesssim g_{(2^{1/2}\alpha\gamma^{-1}t)}(x_0+x_1+q-r)g_{(\alpha t)}(x_3^0+q). \quad (6\text{-}22)$$

Indeed, we have

$$g_{(\alpha t)}(2p)\lesssim e^{-\pi\gamma^2\alpha^{-2}t^{-2}(p/2)^2}g_{(2^{-1/2}\alpha t)}(p).$$

The elementary inequality $e^{-2(a+b)^2}e^{-2b^2}\le e^{-a^2}$ yields

$$g_{(\alpha\gamma^{-1}t)}(x_0+x_1+p/2+q-r)e^{-\pi\gamma^2\alpha^{-2}t^{-2}(p/2)^2}\lesssim g_{(2^{1/2}\alpha\gamma^{-1}t)}(x_0+x_1+q-r).$$

Thus, the left-hand side of (6-22) is bounded by

$$g_{(2^{1/2}\alpha\gamma^{-1}t)}(x_0+x_1+q-r)(g_{(2^{-1/2}\alpha t)}*g_{(t)})(x_3^0+q)\lesssim g_{(2^{1/2}\alpha\gamma^{-1}t)}(x_0+x_1+q-r)g_{(\alpha t)}(x_3^0+q),$$

as desired.

Expressing further $g_{(2^{1/2}\alpha\gamma^{-1}t)}(x_0+x_1+q-r)$ as a convolution of two Gaussians and using the evenness of the Gaussian, (6-21) is bounded by

$$\alpha\int_0^\infty\int_{\mathbb{R}^7}\biggl[\int_{\mathbb{R}}\biggl[\prod_{i=0,1} f_0(x_3^i,x_1,x_2)f_1(x_0,x_3^i,x_2)\biggr]h_{(z^{-1}t)}(x_2+r)\,dx_2\biggr]^2$$
$$\times g_{(\alpha\gamma^{-1}t)}(x_0+p)g_{(\alpha\gamma^{-1}t)}(x_1-p+q-r)g_{(\alpha t)}(x_3^0+q)g_{(\alpha t)}(x_3^1+q)\,dx_0\,dx_1\,dx_3^0\,dx_3^1\,dp\,dq\,dr\,\frac{dt}{t}. \quad (6\text{-}23)$$

After renaming of variables, naming the variable x_2 that is twice an integration variable once as x_2^0 and once as x_2^1, then renaming the variables $x_0,x_1,x_2^0,x_2^1,x_3^0,x_3^1$ in this order as $x_1^1,x_1^0,x_3^0,x_3^1,x_2^0,x_2^1$, and finally introducing functions $\tilde f_0(a,b,c)=f_0(b,a,c)$ and $\tilde f_1=f_1$, we write (6-23) as

$$\alpha\int_0^\infty\int_{\mathbb{R}^7}\biggl[\prod_{i=0,1}\int_{\mathbb{R}}\tilde f_0(x_1^0,x_2^0,x_3^i)\tilde f_1(x_1^1,x_2^0,x_3^i)\tilde f_0(x_1^0,x_2^1,x_3^i)\tilde f_1(x_1^1,x_2^1,x_3^i)h_{(z^{-1}t)}(x_3^i+r)\,dx_3^i\biggr]$$
$$\times g_{(\alpha\gamma^{-1}t)}(x_1^0-p+q-r)g_{(\alpha\gamma^{-1}t)}(x_1^1+p)g_{(\alpha t)}(x_2^0+q)g_{(\alpha t)}(x_2^1+q)\,dx_1^0\,dx_1^1\,dx_2^0\,dx_2^1\,dp\,dq\,dr\,\frac{dt}{t}.$$

Let S and $(\Pi_s)_{s\in S}$ be as in the datum D_A. Introducing $f_s=\tilde f_{s(1)}$ for $s\in S$, we may write the last display as

$$\alpha\int_0^\infty\int_{\mathbb{R}^9}\biggl[\prod_{s\in S} f_s(\Pi_s x)\biggr]g_{(\alpha\gamma^{-1}t)}(x_1^0-p+q-r)g_{(\alpha\gamma^{-1}t)}(x_1^1+p)$$
$$\times g_{(\alpha t)}(x_2^0+q)g_{(\alpha t)}(x_2^1+q)h_{(z^{-1}t)}(x_3^0+r)h_{(z^{-1}t)}(x_3^1+r)\,dx\,dp\,dq\,dr\,\frac{dt}{t}.$$

Let $V:\mathbb{R}^3\to\mathbb{R}^3$ be a mapping given by $V(v_0,v_1,v_2)=(\alpha\gamma^{-1}v_0,\alpha v_1,z^{-1}v_2)$. We perform the change of variables with respect to this mapping for each of the triples (p,q,r), (x_1^0,x_2^0,x_3^0) and (x_1^1,x_2^1,x_3^1). After this transformation, the above form becomes

$$\alpha\int_0^\infty\int_{\mathbb{R}^9}\Bigl[\prod_{s\in S}\alpha^{\frac14}\gamma^{-\frac18}z^{-\frac18}f_s(V\Pi_s x)\Bigr]g_{(t)}(x_1^0-p+\gamma q-\gamma z^{-1}\alpha^{-1}r)g_{(t)}(x_1^1+p)$$
$$\times g_{(t)}(x_2^0+q)g_{(t)}(x_2^1+q)h_{(t)}(x_3^0+r)h_{(t)}(x_3^1+r)\,dx\,dp\,dq\,dr\,\frac{dt}{t}. \tag{6-24}$$

This can be recognized as an α multiple of

$$\Lambda_{D_A,K}((\alpha^{\frac14}\gamma^{-\frac18}z^{-\frac18}f_s\circ V)_{s\in S}),$$

where K has the form (6-2) with $i=3$ and

$$A=\begin{pmatrix}1&-\gamma&\gamma z^{-1}\alpha^{-1}\\0&-1&0\\0&0&-1\end{pmatrix}.$$

Since $0<\gamma\le z\le 1\le\alpha$, the matrix A satisfies the assumption (6-1) with $\varepsilon=5^{-1/2}$. Observing further that the function $\alpha^{1/4}\gamma^{-1/8}z^{-1/8}f_s\circ V$ has the same L^8 norm as f_s, we deduce from Lemma 6.1 that (6-24) is bounded by $\lesssim\alpha$. This yields the desired bound for (6-12).

Combining the estimates for (6-11) and (6-12), we obtain (6-5).

It remains to consider the part of the integral in (6-3) where $z\in(1,\infty)$. Let φ be the function defined via its Fourier transform by

$$\widehat{\varphi}(\xi)=\int_1^\infty\hat\theta(z\xi)\,\frac{dz}{z}.$$

Then we can write

$$\int_1^\infty\widehat{K}^{(z)}(\xi,\eta)\,\frac{dz}{z}=\int_0^\infty\widehat{K}(\xi,\eta)\widehat{\varphi}(t(\xi-\eta))\hat\theta(t(\xi+\eta))\,\frac{dt}{t}. \tag{6-25}$$

Formally, this expression has the same form as (6-4) when $z=1$, except that the function $\hat\theta$ is at one occurrence replaced by $\widehat{\varphi}$. Due to this similarity, we will denote (6-25) by $\widehat{K}^{(1)}(\xi,\eta)$. Note that $\hat\theta$ is supported in $\bigl[-2,-\frac12\bigr]\cup\bigl[\frac12,2\bigr]$, $\widehat{\varphi}$ is supported in $[-2,2]$ and the support properties of $\hat\theta$ and $\widehat{\varphi}$ ensure that $\widehat{\varphi}(\xi-\eta)\hat\theta(\xi+\eta)$ is supported in the set where $\frac18\le|(\xi,\eta)|\le 8$. We may therefore apply an argument analogous to the case $z\in(0,1)$, arriving at the estimate

$$|\Lambda_{D_1,K^{(1)}}((f_s)_{s\in S})|\lesssim J^{\frac12}.$$

Combining this with (6-5) yields the conclusion of the proposition. □

7. Proof of Proposition 2.6 using Propositions 2.7 and 2.8

Let $(k_j)_{j=0}^J$ be a finite increasing sequence of integers with $k_{j-1}+10\le k_j$. For $1\le j\le J$, let $\phi_{0,j}$, $\phi_{1,j}$, $\phi_{2,j}$ be rescaled respective left or right windows as in the proposition, and define K as in (2-9). Let $(f_s)_{s\in S}$ be a tuple of functions as in (2-3) and (2-4). Set $f_k:=f_{(k,0)}=f_{(k,1)}$ for each $k=0,1,2$.

Let (χ, ϕ) be a universal pair and define $\chi_j := \chi_{(2^{k_j-2})}$ and $\phi_j := \phi_{(2^{k_j-2})}$. Define

$$\phi_{3,j} := \chi_{j-1} - \phi_j \tag{7-1}$$

and consequently,

$$(\widehat{\phi}_{3,j})^2 = (\widehat{\chi}_{j-1})^2 - (\widehat{\chi}_j)^2.$$

If (ξ, η) is in the support of $(\widehat{\phi}_{0,j} - \widehat{\phi}_{2,j}) \otimes \widehat{\phi}_{1,j}$, then

$$2^{-k_j+3} - 2^{-k_j} \le |\xi + \eta| \le 2^{-k_{j-1}} + 2^{-k_j}.$$

In this range, $\widehat{\chi}_{j-1}(\xi + \eta)$ is constant 1 and $\widehat{\chi}_j(\xi + \eta)$ is constant zero. We can therefore introduce artificial factors $\widehat{\phi}_{3,j}$ in $\widehat{K}$ as follows:

$$\widehat{K}(\xi, \eta) = \sum_{j=1}^{J} (\widehat{\phi}_{0,j} - \widehat{\phi}_{2,j})(\xi)\widehat{\phi}_{1,j}(\eta) = \sum_{j=1}^{J} (\widehat{\phi}_{0,j} - \widehat{\phi}_{2,j})(\xi)\widehat{\phi}_{1,j}(\eta)\widehat{\phi}_{3,j}(-\xi - \eta).$$

Taking the Fourier transform, we obtain for some universal constant C,

$$\begin{aligned} K(u, v) &= \int_{\mathbb{R}^2} \sum_{j=1}^{J} (\widehat{\phi}_{0,j} - \widehat{\phi}_{2,j})(\xi) e^{2\pi i u\xi} \widehat{\phi}_{1,j}(\eta) e^{2\pi i v\eta} \widehat{\phi}_{3,j}(-\xi - \eta)\, d\xi\, d\eta \\ &= C \sum_{j=1}^{J} \int_{\mathbb{R}} (\phi_{0,j} - \phi_{2,j})(u + p)\phi_{1,j}(v + p)\phi_{3,j}(p)\, dp, \end{aligned} \tag{7-2}$$

where we used that the integral of a function in $\mathbb{R}^3$ over the diagonal $\{(p, p, p) : p \in \mathbb{R}\}$ equals the integral of its Fourier transform over the orthogonal complement of the diagonal, suitably normalized.

Therefore, with S and Π_s as in the datum D_1, and doing a variable transformation $p \to x_2 + p$,

$$\begin{aligned} \Lambda_{D_1,K}((f_s)_{s\in S}) = \sum_{j=1}^{J} \int_{\mathbb{R}^6} \Big[\prod_{s\in S} f_s(\Pi_s x)\Big] \\ \times (\phi_{0,j} - \phi_{2,j})(x_3^0 - x_0 - x_1 + p)\phi_{1,j}(x_3^1 - x_0 - x_1 + p)\phi_{3,j}(x_2 + p)\, dx\, dp. \end{aligned} \tag{7-3}$$

We apply Fubini in (7-3) to have the integral in x_2 as the innermost and then apply the Cauchy–Schwarz inequality in $x_0, x_1, x_3^0, x_3^1, p$, which bounds $|\Lambda_{D_1,K}((f_s)_{s\in S})|$ up to a constant by the geometric mean of

$$\sum_{j=1}^{J} \int_{\mathbb{R}^5} \Big[\prod_{i=0,1} |f_2(x_0, x_1, x_3^i)|^2\Big] \mu_j(x_3^0 - x_0 - x_1 + p, x_3^1 - x_0 - x_1 + p)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp \tag{7-4}$$

and

$$\begin{aligned} \sum_{j=1}^{J} \int_{\mathbb{R}^5} \Big[\int_{\mathbb{R}} \Big[\prod_{i=0,1} f_0(x_3^i, x_1, x_2) f_1(x_0, x_3^i, x_2)\Big] \phi_{3,j}(x_2 + p)\, dx_2\Big]^2 \\ \times \mu_j\big(x_3^0 - x_0 - x_1 + p, x_3^1 - x_0 - x_1 + p\big)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp, \end{aligned} \tag{7-5}$$

where we have introduced the weight μ_j defined by

$$\mu_j(u,v) = |\phi_{0,j} - \phi_{2,j}|(u)|\phi_{1,j}|(v).$$

We will estimate (7-4) as $\lesssim J$ and (7-5) as $\lesssim 1$, thereby proving Proposition 2.6.

We begin with (7-4). Applying the Cauchy–Schwarz inequality in the remaining integration variables, we bound (7-4) by

$$\sum_{j=1}^{J} \prod_{i=0,1} \bigg[\int_{\mathbb{R}^5} |f_2(x_0, x_1, x_3^i)|^4 \mu_j\big(x_3^0 - x_0 - x_1 + p, x_3^1 - x_0 - x_1 + p\big)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp \bigg]^{\frac{1}{2}} = \sum_{j=1}^{J} \|f_2\|_4^4 \|\mu_j\|_1 \lesssim J.$$

Here the identity is seen by integrating first in x_3^{1-i} then in p to obtain the L^1 norm of μ_j.

It remains to estimate (7-5). We use decay of μ_j thanks to control of derivatives of Fourier transform of windows and the superposition estimate

$$(1 + |(u,v)|)^{-N-20} \lesssim \int_1^\infty g_{(\alpha)}(u) g_{(\alpha)}(v) \,\frac{d\alpha}{\alpha^{N+10}},$$

which we scale isotropically and anisotropically, to dominate

$$\mu_j(u,v) \lesssim \int_1^\infty g_{(\alpha 2^{k_j})}(u) g_{(\alpha 2^{k_j})}(v) \,\frac{d\alpha}{\alpha^{N+10}} + \int_1^\infty g_{(\alpha 2^{k_{j-1}})}(u) g_{(\alpha 2^{k_j})}(v) \,\frac{d\alpha}{\alpha^{N+10}}.$$

By superposition of positive terms, it suffices to estimate as $\lesssim \alpha^N$ the variant of (7-5) with $\mu_j(u,v)$ replaced by

$$g_{(\alpha 2^{l_j})}(u) g_{(\alpha 2^{k_j})}(v)$$

and for each of the sequences $l_j = k_j$ and $l_j = k_{j-1}$. Define the sequence of real numbers $(m_j)_{j=1}^J$ by

$$2^{2m_j} + 2^{2m_j} = 2^{2k_j} + 2^{2l_j}.$$

Note that $k_j - 1 \le m_j \le k_j$, because $l_j \le k_j$. Adding and subtracting terms, it suffices to estimate as $\lesssim \alpha^N$ the variants of (7-5) with $\mu_j(u,v)$ replaced by

$$g_{(\alpha 2^{m_j})}(u) g_{(\alpha 2^{m_j})}(v) \tag{7-6}$$

and by $(\nu_j)_{(\alpha)}(u,v)$, where

$$\nu_j(u,v) := g_{(2^{l_j})}(u) g_{(2^{k_j})}(v) - g_{(2^{m_j})}(u) g_{(2^{m_j})}(v). \tag{7-7}$$

We begin with (7-6). We need to estimate

$$\sum_{j=1}^{J} \int_{\mathbb{R}^5} \bigg[\int_{\mathbb{R}} \bigg[\prod_{i=0,1} f_0(x_3^i, x_1, x_2) f_1(x_0, x_3^i, x_2) \bigg] \phi_{3,j}(x_2 + p)\, dx_2 \bigg]^2 \times g_{(\alpha 2^{m_j})}(x_3^0 - x_0 - x_1 + p) g_{(\alpha 2^{m_j})}(x_3^1 - x_0 - x_1 + p)\, dx_0\, dx_1\, dx_3^0\, dx_3^1\, dp. \tag{7-8}$$

A renaming of variables, naming the variable x_2 that is twice an integration variable once as x_2^0 and once as x_2^1, then renaming the variables $x_0, x_1, x_2^0, x_2^1, x_3^0, x_3^1$ in this order as $x_1, x_0, x_3^0, x_3^1, x_2^0, x_2^1$, and finally introducing functions $\tilde{f}_0(a,b,c) = f_0(b,a,c)$ and $\tilde{f}_1 = f_1$, we write (7-8) as

$$\sum_{j=1}^{J} \int_{\mathbb{R}^5} \Bigl[\prod_{i=0}^{1} \int_{\mathbb{R}} \tilde{f}_0(x_0, x_2^0, x_3^i) \tilde{f}_1(x_1, x_2^0, x_3^i) \tilde{f}_0(x_0, x_2^1, x_3^i) \tilde{f}_1(x_1, x_2^1, x_3^i) \phi_{3,j}(x_3^i + p)\, dx_3^i\Bigr] \\ \times g_{(\alpha 2^{m_j})}(x_2^0 - x_0 - x_1 + p) g_{(\alpha 2^{m_j})}(x_2^1 - x_0 - x_1 + p)\, dx_0\, dx_1\, dx_2^0\, dx_2^1\, dp. \quad (7\text{-}9)$$

Introducing for the datum D_2 the tuple $f_{(k,j)} = \tilde{f}_k$ for $k = 0, 1$ and $j \in \mathcal{C}$, we may write (7-9) as $\Lambda_{D_2,K_1}((f_s)_{s\in S})$, with

$$K_1(u,v,z) = \sum_{j=1}^{J} \int_{\mathbb{R}} g_{(\alpha 2^{m_j})}(u+p) g_{(\alpha 2^{m_j})}(v+p) \phi_{3,j}(z+p) \phi_{3,j}(p)\, dp.$$

Proposition 2.7 implies $\Lambda_{D_2,K_1}((f_s)_{s\in S}) \lesssim \alpha^N$.

It remains to estimate the term with (7-7). We may assume $l_j = k_{j-1}$, because (7-7) vanishes in the case $k_j = l_j$. With similar transformations as for term (7-6), we write the form associated with (7-7) as $\Lambda_{D_2,K_2}((f_s)_{s\in S})$ with

$$K_2(u,v,z) = \sum_{j=1}^{J} \int_{\mathbb{R}} (\nu_j)_{(\alpha)}(u+p, v+p) \phi_{3,j}(z+p) \phi_{3,j}(p)\, dp.$$

We decompose $\nu_j = \sum_{n\in\mathbb{Z}} \nu_{j,n}$, where

$$\hat{\nu}_{j,0}(\xi,\eta) = \hat{\nu}_j(\xi,\eta)((\widehat{\chi}_{(2^{k_{j-1}})})^2 - (\widehat{\chi}_{(2^{k_j})})^2)(\xi+\eta),$$

and for $n < 0$,

$$\hat{\nu}_{j,n}(\xi,\eta) = \hat{\nu}_j(\xi,\eta)((\widehat{\chi}_{(2^{k_{j-1}+n})})^2 - (\widehat{\chi}_{(2^{k_{j-1}+n+1})})^2)(\xi+\eta) \quad (7\text{-}10)$$

and for $n > 0$,

$$\hat{\nu}_{j,n}(\xi,\eta) = \hat{\nu}_j(\xi,\eta)((\widehat{\chi}_{(2^{k_j+n-1})})^2 - (\widehat{\chi}_{(2^{k_j+n})})^2)(\xi+\eta). \quad (7\text{-}11)$$

We split $K_2 = \sum_{n\in\mathbb{Z}} K_{2,n}$ accordingly and estimate for each n

$$\Lambda_{D_2,K_{2,n}}((f_s)_{s\in S}) \lesssim 2^{-|n|}.$$

Upon summing over n, we obtain the desired bound for (7-7).

We begin with $n = 0$. We have, similarly as in (7-2), for some universal constant C,

$$K_2(u,v,z) = C \sum_{j=1}^{J} \int_{\mathbb{R}^3} (\hat{\nu}_j)_{(\alpha)}(\xi,\eta) e^{2\pi i(u\xi+v\eta)} \widehat{\phi}_{3,j}(\tau) e^{2\pi i z\tau} \widehat{\phi}_{3,j}(-\tau-\xi-\eta)\, d\xi\, d\eta\, d\tau,$$

and thus

$$\widehat{K}_{2,0}(\xi,\eta,\tau) = C \sum_{j=1}^{J} ((\widehat{\chi}_{(\alpha 2^{k_{j-1}})})^2 - (\widehat{\chi}_{(\alpha 2^{k_j})})^2)(\xi+\eta)(\hat{\nu}_j)_{(\alpha)}(\xi,\eta) \widehat{\phi}_{3,j}(\tau) \widehat{\phi}_{3,j}(-\tau-\xi-\eta).$$

Preparing to apply Proposition 2.8, we note that $K_{2,0}$ is of the form (2-15) with ρ_j defined by

$$\rho_j(u_1, u_2, u_3, u_4) := (\nu_j)_{(\alpha)}(u_1, u_2)\phi_{3,j}(u_3)\phi_{3,j}(u_4),$$

as can be seen from the Fourier transform side (2-16). We do not attempt to show that ρ_j itself satisfies the assumptions of Proposition 2.8, but we split into eight pieces by the distributive law, splitting ν_j into two pieces as in its definition (7-7) and each $\phi_{3,j}$ into two as in its definition (7-1). A typical piece is

$$g_{(\alpha 2^{l_j})}(u_1) g_{(\alpha 2^{k_j})}(u_2)\chi_{j-1}(u_3)\phi_j(u_4),$$

which satisfies the assumptions of Proposition 2.8, because

$$\int_{\mathbb{R}^2} g_{(\alpha 2^{l_j})}(u_1+p) g_{(\alpha 2^{k_j})}(u_2+p)\chi_{j-1}(u_3+r)\phi_j(u_4+r)\,dp\,dr \\ \lesssim (g*g)_{(\alpha 2^{1+k_j})}(u_1-u_2) 2^{-k_j}(1+2^{-k_j}|u_3-u_4|)^{-2}.$$

This along with similar estimates for the other seven pieces completes the bound for $\Lambda_{D_2,K_{2,0}}((f_s)_{s\in S})$ by Proposition 2.8.

We turn to $n>0$. We introduce artificial factors that are constant 1 where relevant, using that the sequence k_j is well separated, and write

$$\widehat{K}_{2,n}(\xi,\eta,\tau) = C\sum_{j=1}^{J}((\widehat{\chi}_{(\alpha 2^{k_{j-1}+n+1})})^2 - (\widehat{\chi}_{(\alpha 2^{k_j+n+1})})^2)(\xi+\eta) \\ \times ((\widehat{\chi}_{(\alpha 2^{k_j+n-1})})^2 - (\widehat{\chi}_{(\alpha 2^{k_j+n})})^2)(\xi+\eta)(\hat{\nu}_j)_{(\alpha)}(\xi,\eta)\widehat{\phi}_{3,j}(\tau)\widehat{\phi}_{3,j}(-\tau-\xi-\eta).$$

This kernel is of the form (2-15) with

$$\hat{\rho}_j(\xi,\eta,\tau,\sigma) = (\hat{\nu}_{j,n})_{(\alpha)}(\xi,\eta)\widehat{\phi}_{3,j}(\tau)\widehat{\phi}_{3,j}(\sigma),$$

with $\nu_{j,n}$ defined in (7-11). We break both functions $\widehat{\phi}_{3,j}$ into pieces as above. All pieces are done similarly, we discuss a typical piece of ρ_j given by

$$\hat{\varrho}_j(\xi,\eta,\tau,\sigma) = (\hat{\nu}_{j,n})_{(\alpha)}(\xi,\eta)\widehat{\chi}_{j-1}(\tau)\widehat{\phi}_j(\sigma).$$

Using that χ_j and ϕ_j are even, we have

$$\int_{\mathbb{R}} \chi_{j-1}(u_3+r)\phi_j(u_4+r)\,dr = (\chi_{j-1}*\phi_j)(u_3-u_4) \lesssim 2^{-k_j}(1+2^{-k_j}|u_3-u_4|)^{-2}.$$

With Lemma 7.1 below, we obtain

$$\int_{\mathbb{R}^2} |\varrho_j|(u_1+p, u_2+p, u_3+r, u_4+r)\,dp\,dr \\ \lesssim 2^{-n}\alpha^{-1}2^{-k_j}(1+\alpha^{-1}2^{-k_j}|u_1-u_2|)^{-2}2^{-k_j}(1+2^{-k_j}|u_3-u_4|)^{-2}.$$

Proposition 2.8 gives $\Lambda_{D_2,K_{2,n}}((f_s)_{s\in S}) \lesssim 2^{-n}$, as desired.

Lemma 7.1. *We have for every $1\le j\le J$ and every $x, y\in\mathbb{R}$ the estimate*

$$|\nu_{j,n}(x,y)| \lesssim 2^{-n}2^{-k_j}(1+2^{-k_j}|x-y|)^{-4}2^{-n-k_j}(1+2^{-n-k_j}|x+y|)^{-4}.$$

Proof of Lemma 7.1. Scaling by a factor 2^{k_j} allows us to assume $k_j = 0$ and $-1 \le m_j \le 0$ and $l_j \le 0$. We fix j and omit the index j. We thus have to prove

$$|\nu_n(x, y)| \lesssim 2^{-2n}(1 + |x - y|)^{-4}(1 + 2^{-n}|x + y|)^{-4}, \tag{7-12}$$

where

$$\hat{\nu}_n(\xi, \eta) = ((\widehat{\chi}_{(2^{-1})})^2 - (\widehat{\chi})^2)(2^n(\xi + \eta))\hat{\nu}(\xi, \eta) \tag{7-13}$$

with

$$\hat{\nu}(\xi, \eta) = g(2^l \xi)g(\eta) - g(2^m \xi)g(2^m \eta). \tag{7-14}$$

We claim that for $0 \le \alpha \le 4$ and $0 \le \beta \le 4$ and $|\xi + \eta| \le 1$

$$|\partial^\alpha_{(1,-1)}\partial^\beta_{(1,1)}\hat{\nu}(\xi, \eta)| \lesssim |\xi + \eta|^{\max\{1-\beta,0\}}(1 + |\xi - \eta|)^{-2}.$$

For $\beta > 0$, this follows by deriving (7-14) and using the decay of Gaussians and their derivatives for those Gaussians whose argument contains m or k, because $-1 \le m \le 0$ and $k = 0$. Here we also use the fact that whenever $|\xi + \eta| \le 1$ and $|\xi - \eta| \ge 1$, then the three quantities $|\xi - \eta|$, $|\xi|$ and $|\eta|$ are comparable.

We next estimate the term with $\beta = 0$. By the choice of m, the function $\hat{\nu}$ vanishes on the diagonal $\xi + \eta = 0$, and thus the same property holds also for $\partial^\alpha_{(1,-1)}\hat{\nu}$. Therefore,

$$\begin{aligned}
|\partial^\alpha_{(1,-1)}\hat{\nu}(\xi, \eta)| &= \big|\partial^\alpha_{(1,-1)}\hat{\nu}(\xi, \eta) - \partial^\alpha_{(1,-1)}\hat{\nu}\big((\xi - \eta)/2, -(\xi - \eta)/2\big)\big| \\
&\lesssim |\xi + \eta|\left|\int_0^1 \partial^\alpha_{(1,-1)}\partial_{(1,1)}\hat{\nu}\big((\xi - \eta)/2 + r(\xi + \eta)/2, -(\xi - \eta)/2 + r(\xi + \eta)/2\big)\, dr\right| \\
&\lesssim |\xi + \eta| \sup_{0 \le r \le 1} \big|\partial^\alpha_{(1,-1)}\partial_{(1,1)}\hat{\nu}\big((\xi - \eta)/2 + r(\xi + \eta)/2, -(\xi - \eta)/2 + r(\xi + \eta)/2\big)\big| \\
&\lesssim |\xi + \eta|(1 + |\xi - \eta|)^{-2}.
\end{aligned}$$

Turning to $\hat{\nu}_n$ as in (7-13), using that $\widehat{\chi}^2_{(2^{-1})} - \widehat{\chi}^2$ is supported in $[-2, 2]$, we obtain by differentiating

$$\begin{aligned}
|\hat{\nu}_n(\xi, \eta)| &\lesssim 2^{-n} 1_{|2^n(\xi+\eta)|<1}(1 + |\xi - \eta|)^{-2}, \\
|\partial^4_{(1,-1)}\hat{\nu}_n(\xi, \eta)| &\lesssim 2^{-n} 1_{|2^n(\xi+\eta)|<1}(1 + |\xi - \eta|)^{-2}, \\
|\partial^4_{(1,1)}\hat{\nu}_n(\xi, \eta)| &\lesssim 2^{3n} 1_{|2^n(\xi+\eta)|<1}(1 + |\xi - \eta|)^{-2}, \\
|\partial^4_{(1,-1)}\partial^4_{(1,1)}\hat{\nu}_n(\xi, \eta)| &\lesssim 2^{3n} 1_{|2^n(\xi+\eta)|<1}(1 + |\xi - \eta|)^{-2}.
\end{aligned}$$

Hence, estimating the Fourier inversion formula crudely by $L^1 \to L^\infty$ bounds,

$$|\nu_n(x, y)| \lesssim 2^{-2n}, \quad |x-y|^4|\nu_n(x, y)| \lesssim 2^{-2n}, \quad |x+y|^4|\nu_n(x, y)| \lesssim 2^{2n}, \quad |x-y|^4|x+y|^4|\nu_n(x, y)| \lesssim 2^{2n}.$$

We can summarize these findings into (7-12), as can be seen by splitting into four cases depending on whether $2^n \le |x + y|$ or $2^n > |x + y|$ and depending on whether $1 \le |x - y|$ or $1 > |x - y|$. This proves the lemma. □

We finally turn to $n < 0$. As in the previous case, we introduce an artificial factor and write

$$\widehat{K}_{2,n}(\xi,\eta,\tau) = C\sum_{j=1}^{J}((\widehat{\chi}_{(\alpha 2^{k_{j-1}+n-1})})^2 - (\widehat{\chi}_{(\alpha 2^{k_j+n-1})})^2)(\xi+\eta) \times((\widehat{\chi}_{(\alpha 2^{k_{j-1}+n})})^2 - (\widehat{\chi}_{(\alpha 2^{k_{j-1}+n+1})})^2)(\xi+\eta)(\hat{\nu}_j)_{(\alpha)}(\xi,\eta)\widehat{\phi}_{3,j}(\tau)\widehat{\phi}_{3,j}(-\tau-\xi-\eta).$$

This kernel is of the form (2-15) with

$$\hat{\rho}_j(\xi,\eta,\tau,\sigma) = (\hat{\nu}_{j,n})_{(\alpha)}(\xi,\eta)\widehat{\phi}_{3,j}(\tau)\widehat{\phi}_{3,j}(\sigma)$$

with $\nu_{j,n}$ as in (7-10).

We break both functions $\widehat{\phi}_{3,j}$ into pieces as above. All pieces are done similarly, we discuss a typical piece of ρ_j given by

$$\hat{\varrho}_j(\xi,\eta,\tau,\sigma) = (\hat{\nu}_{j,n})_{(\alpha)}(\xi,\eta)\widehat{\chi}_{j-1}(\tau)\widehat{\phi}_j(\sigma).$$

With Lemma 7.2, we obtain

$$\int_{\mathbb{R}^2} |\varrho_j|(u_1+p,u_2+p,u_3+r,u_4+r)\,dp\,dr \lesssim 2^n\alpha^{-1}2^{-k_j}(1+\alpha^{-1}2^{-k_j}|u_1-u_2|)^{-4}2^{-k_j}(1+2^{-k_j}|u_3-u_4|)^{-2}.$$

Proposition 2.8 gives $\Lambda_{D_2,K_{2,n}}((f_s)_{s\in S}) \lesssim 2^n$, as desired.

Lemma 7.2. *We have for every $1 \le j \le J$ and every $x, y \in \mathbb{R}$ the estimate*

$$|\nu_{j,n}(x,y)| \lesssim 2^n 2^{-k_{j-1}}(1+2^{-k_{j-1}}|x|)^{-4}2^{-k_j}(1+2^{-k_j}|y|)^{-4}.$$

Proof. We split the function

$$\hat{\nu}_j(\xi,\eta) = g(2^{l_j}\xi)g(2^{k_j}\eta) - g(2^{m_j}\xi)g(2^{m_j}\eta)$$

into its two summands and consider the summands separately. Consider the term

$$g(2^{l_j}\xi)g(2^{k_j}\eta).$$

Scaling by the factor 2^{l_j} in ξ and 2^{k_j} in η reduces the matter to proving

$$|\mu_n(x,y)| \lesssim 2^n(1+|x|)^{-4}(1+|y|)^{-4}, \tag{7-15}$$

where

$$\widehat{\mu}_n(\xi,\eta) = (\widehat{\chi}^2 - \widehat{\chi}^2_{(2)})(2^n\xi + 2^{n+l}\eta)g(\xi)g(\eta)$$

with $l \le 0$. On the support of the function

$$(\widehat{\chi}^2 - \widehat{\chi}^2_{(2)})(2^n\xi + 2^{n+l}\eta),$$

we have

$$|\partial_\xi^\alpha \partial_\eta^\beta g(\xi)g(\eta)| \lesssim 2^n(1+|\xi|)^{-4}(1+|\eta|)^{-4}$$

for all $0 \le \alpha, \beta \le 4$. By the Leibniz rule, analogous bounds hold for $\widehat{\mu}_n$. The function μ_n then satisfies the bound (7-15). This is the desired estimate for the term $g(2^{l_j}\xi)g(2^{k_j}\eta)$.

To estimate the term $g(2^{m_j}\xi)g(2^{m_j}\eta)$, we rescale by 2^{m_j} in both variables and claim

$$|\mu_n(x,y)| \lesssim 2^n 2^{5l}(1+|x|)^{-4}(1+|y|)^{-4},$$

where

$$\widehat{\mu}_n(\xi,\eta) = (\widehat{\chi}^2 - \widehat{\chi}^2_{(2)})(2^{n+l}(\xi+\eta))g(\xi)g(\eta)$$

and $l = k_{j-1} - m_j \le 0$. This follows similarly as before, using the decay of the Gaussians. As

$$2^{5l}(1+|x|)^{-4} \lesssim 2^{-l}(1+2^{-l}|x|)^{-4},$$

this completes the proof of the lemma. □

8. Proof of Proposition 2.7 using Propositions 2.8, 2.9, and Theorem 1.1 in [Durcik et al. 2022]

Let $\alpha \ge 1$. Let J be a positive integer and $(k_j)_{j=0}^J$ a finite increasing sequence of integers with $k_{j-1}+10 \le k_j$ for $1 \le j \le J$, let $(m_j)_{j=1}^J$ be a sequence of real numbers with $k_j - 1 \le m_j \le k_j$. For $0 \le j \le J$, let χ_j be a function such that $(\chi_j)_{(2^{2-k_j})}$ is a left window and let ϕ_j be as in the statement of the proposition, i.e.,

$$(\widehat{\phi}_j)^2 = (\widehat{\chi}_{j-1})^2 - (\widehat{\chi}_j)^2.$$

Let a tuple $(f_s)_{s\in S}$ be given as in (2-12), (2-13) and write $f_{(0,j)} = f_0,\ f_{(1,j)} = f_1$ for any $j \in \mathcal{C}$.

Taking the Fourier transform, the kernel K of the proposition reads as

$$\widehat{K}(\xi,\eta,\tau) = \alpha^{-N} \sum_{j=1}^{J} \hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\widehat{\phi}_j(\tau)\widehat{\phi}_j(-\xi-\eta-\tau).$$

Define the kernel K_1 by

$$\widehat{K}_1(\xi,\eta,\tau) := \alpha^{-N} \sum_{j=1}^{J} \hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\big(\widehat{\chi}_{j-1}(\tau)\widehat{\chi}_{j-1}(-\tau-\xi-\eta) - \widehat{\chi}_j(\tau)\widehat{\chi}_j(-\tau-\xi-\eta)\big).$$

Therefore, on the critical space $\xi+\eta=0$, the kernels are equal, i.e., for all ξ,τ we have

$$\widehat{K}(\xi,-\xi,\tau) = \widehat{K}_1(\xi,-\xi,\tau).$$

By the triangle inequality, it suffices to estimate $\Lambda_{D_2,K-K_1}$ and Λ_{D_2,K_1}.

We begin with the latter. Since $\alpha \ge 1$, we observe that it in fact suffices to prove the (stronger) bound $|\Lambda_{D_2,\alpha^N K_1}((f_s)_{s\in S})| \lesssim 1$. Define the kernel K_2 by

$$\widehat{K}_2(\xi,\eta,\tau) := \sum_{j=1}^{J} \big(\hat{g}_{(\alpha 2^{m_{j-1}})}(\xi)\hat{g}_{(\alpha 2^{m_{j-1}})}(\eta) - \hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\big)\widehat{\chi}_{j-1}(\tau)\widehat{\chi}_{j-1}(-\tau-\xi-\eta)$$

and define

$$\widehat{\sigma}_j(\xi,\eta,\tau) := \hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\widehat{\chi}_j(\tau)\widehat{\chi}_j(-\tau-\xi-\eta).$$

Here, we formally set $m_0 = k_0$. By telescoping, we have

$$\alpha^N K_1 + K_2 = \sigma_0 - \sigma_J.$$

For each j, $\Lambda_{D_2,\sigma_j}((f_s)_{s\in S})$ equals

$$\int_{\mathbb{R}^7}\Big[\prod_{s\in S} f_s(\Pi_s x)\Big] g_{(\alpha 2^{m_j})}(x_2^0-x_0-x_1+p)g_{(\alpha 2^{m_j})}(x_2^1-x_0-x_1+p)\chi_j(x_3^0+p)\chi_j(x_3^1+p)\,dx\,dp,$$

where $x=(x_0,x_1,x_2^0,x_2^1,x_3^0,x_3^1)$. This can be estimated using a classical Brascamp–Lieb inequality as

$$|\Lambda_{D_2,\sigma_j}((f_s)_{s\in S})|\lesssim \|g_{(\alpha 2^{m_j})}\|_1^2\|\chi_j\|_1^2\prod_{s\in S}\|f_s\|_8\lesssim 1. \tag{8-1}$$

One can verify this Brascamp–Lieb inequality by interpolation between estimates that put one of the functions f_s in L^1 and all others in L^∞.

The estimate of $\Lambda_{D_2,\alpha^N K_1}$ is thus reduced to an estimate of Λ_{D_2,K_2}, which we now proceed to do. We use the fundamental theorem of calculus to split up a difference of Gaussians with parameters a,b as

$$\begin{aligned}g(a\xi)g(a\eta)-g(b\xi)g(b\eta)&=\int_a^b -t\partial_t(g(t\xi)g(t\eta))\,\frac{dt}{t}=2\pi\int_a^b t^2(\xi^2+\eta^2)g(t\xi)g(t\eta)\,\frac{dt}{t}\\&=2\pi\int_a^b t^2(\xi+\eta)^2g(t\xi)g(t\eta)\,\frac{dt}{t}-4\pi\int_a^b t^2\xi\eta g(t\xi)g(t\eta)\,\frac{dt}{t}.\end{aligned} \tag{8-2}$$

Using this splitting, in place of Λ_{D_2,K_2} we may estimate Λ_{D_2,K_3} and Λ_{D_2,K_4} with

$$\widehat{K}_3(\xi,\eta,\tau):=\sum_{j=1}^J\int_{\alpha 2^{m_{j-1}}}^{\alpha 2^{m_j}} t^2(\xi+\eta)^2g(t\xi)g(t\eta)\,\frac{dt}{t}\,\widehat{\chi}_{j-1}(\tau)\widehat{\chi}_{j-1}(-\tau-\xi-\eta)$$

and, using $h=g'$ and that $\hat{h}(\xi)$ is a constant multiple of $\xi\hat{g}(\xi)$,

$$\widehat{K}_4(\xi,\eta,\tau):=\sum_{j=1}^J\int_{\alpha 2^{m_{j-1}}}^{\alpha 2^{m_j}} \hat{h}(t\xi)\hat{h}(t\eta)\,\frac{dt}{t}\,\widehat{\chi}_{j-1}(\tau)\widehat{\chi}_{j-1}(-\tau-\xi-\eta).$$

Proposition 2.9 gives

$$|\Lambda_{D_2,K_3}((f_s)_{s\in S})|\lesssim 1.$$

We turn to Λ_{D_2,K_4}, which we write on the spatial side as

$$\begin{aligned}\sum_{j=1}^J\int_{\alpha 2^{m_{j-1}}}^{\alpha 2^{m_j}}\int_{\mathbb{R}^5}\Big[\int_{\mathbb{R}}\Big[\prod_{i=0,1} f_0(x_0,x_2^i,x_3)f_1(x_1,x_2^i,x_3)\Big]h_{(t)}(x_3+p)\,dx_3\Big]^2&\\ \times\chi_{j-1}(x_2^0-x_0-x_1+p)\chi_{j-1}(x_2^1-x_0-x_1+p)\,dx_0\,dx_1\,dx_2^0\,dx_2^1\,dp\,\frac{dt}{t}&.\end{aligned}$$

Using positivity of the square in this expression, we may dominate

$$|\chi_{j-1}(u)\chi_{j-1}(v)|\le\int_1^\infty g_{(\beta 2^{m_{j-1}})}(u)g_{(\beta 2^{m_{j-1}})}(v)\beta^{-N}\,\frac{d\beta}{\beta}.$$

Then it suffices to estimate for fixed $\beta\ge 1$ the form Λ_{D_2,K_5}, where

$$\widehat{K}_5(\xi,\eta,\tau):=\sum_{j=1}^J\int_{\alpha 2^{m_{j-1}}}^{\alpha 2^{m_j}} \hat{h}_{(t)}(\xi)\hat{h}_{(t)}(\eta)\,\frac{dt}{t}\,\hat{g}_{(\beta 2^{m_{j-1}})}(\tau)\hat{g}_{(\beta 2^{m_{j-1}})}(-\tau-\xi-\eta).$$

We introduce a new kernel

$$\widehat{K}_6(\xi,\eta,\tau)=\sum_{j=1}^{J}\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\int_{\beta 2^{m_{j-1}}}^{\beta 2^{m_j}}\hat{h}_{(t)}(\tau)\hat{h}_{(t)}(-\tau-\xi-\eta)\,\frac{dt}{t}.$$

The kernel K_6 is symmetric to K_5 under the symmetry (2-19). We note that, for some M, which is even in all variables and symmetric under switching the first two variables or switching the second two variables,

$$K_5(x,y,z)=\int_{\mathbb{R}}M(x+p,y+p,z+p,p)\,dp.$$

With $\widetilde{K}_5$ as defined near (2-19), we have

$$\begin{aligned}\widetilde{K}_5(x,y,z)&=\int_{\mathbb{R}}M\big(\tfrac{1}{2}(x+y+z)+p,\tfrac{1}{2}(x-y+z)+p,z+p,p\big)\,dp\\&=\int_{\mathbb{R}}M\big(-p,-y-p,-\tfrac{1}{2}(x-z+y)-p,-\tfrac{1}{2}(x+z+y)-p\big)\,dp,\end{aligned}$$

where we obtained the last identity by the substitution of p by $-p-\frac{1}{2}(x+y+z)$. For $\widetilde{K}_5^*$ as defined near (2-19), we obtain

$$\widetilde{K}_5^*(x,y,z)=\int_{\mathbb{R}}M\big(-p,-z-p,-\tfrac{1}{2}(x-y+z)-p,-\tfrac{1}{2}(x+y+z)-p\big)\,dp.$$

Using that M is an even function and that it is invariant under interchanging the first two entries or the second two entries, we obtain

$$\widetilde{K}_5^*(x,y,z)=\int_{\mathbb{R}}M\big(z+p,p,\tfrac{1}{2}(x+y+z)+p,\tfrac{1}{2}(x-y+z)+p\big)\,dp.$$

Inverting the tilde operation, we identify the kernel

$$K_6^*(x,y,z)=\int_{\mathbb{R}}M(z+p,p,x+p,y+p)\,dp.$$

Hence, the star symmetry acts on M by interchanging the first two variables with the second two variables in M.

As $\Lambda_{D_2,K_5}((f_s)_{s\in S})$ is positive by the above construction, it follows by symmetry that Λ_{D_2,K_6} is positive as well and it suffices to estimate the sum Λ_{D_2,K_5+K_6}.

We reverse the arguments leading from K_2 to K_4, with a Gaussian in place of χ_{j-1}, and apply these arguments both to K_5 and symmetrically to K_6.

In place of Λ_{D_2,K_3}, we obtain the corresponding forms Λ_{D_2,K_7} and Λ_{D_2,K_8} with

$$\widehat{K}_7(\xi,\eta,\tau):=\sum_{j=1}^{J}\int_{\alpha 2^{m_{j-1}}}^{\alpha 2^{m_j}}t^2(\xi+\eta)^2g(t\xi)g(t\eta)\,\frac{dt}{t}\,\hat{g}_{(\beta 2^{m_{j-1}})}(\tau)\hat{g}_{(\beta 2^{m_{j-1}})}(-\tau-\xi-\eta),$$

$$\widehat{K}_8(\xi,\eta,\tau):=\sum_{j=1}^{J}\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\int_{\beta 2^{m_{j-1}}}^{\beta 2^{m_j}}t^2(\xi+\eta)^2g(t\tau)g(t(-\tau-\xi-\eta))\,\frac{dt}{t}.$$

Note that to arrive at K_8, in place of symmetry arguments, we may also use in place of (8-2) the identity

$$(\xi+\eta)^2+2\tau(\tau+\xi+\eta)=\tau^2+(\tau+\xi+\eta)^2.$$

The forms Λ_{D_2,K_7} and symmetrically Λ_{D_2,K_8} are estimated analogously to Λ_{D_2,K_3} using Proposition 2.9.

Having thus reverted the above steps and having arrived at the analogue of Λ_{D_2,K_2}, we have reduced the bound of Λ_{D_2,K_5+K_6} to a bound on Λ_{D_2,K_9} with

$$\begin{aligned}
&\widehat{K}_9(\xi,\eta,\tau)\\
&=\sum_{j=1}^{J}\big[\hat{g}_{(\alpha 2^{m_{j-1}})}(\xi)\hat{g}_{(\alpha 2^{m_{j-1}})}(\eta)-\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\big]\hat{g}_{(\beta 2^{m_{j-1}})}(\tau)\hat{g}_{(\beta 2^{m_{j-1}})}(-\tau-\xi-\eta)\\
&\qquad+\sum_{j=1}^{J}\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\big[\hat{g}_{(\beta 2^{m_{j-1}})}(\tau)\hat{g}_{(\beta 2^{m_{j-1}})}(-\tau-\xi-\eta)-\hat{g}_{(\beta 2^{m_j})}(\tau)\hat{g}_{(\beta 2^{m_j})}(-\tau-\xi-\eta)\big]\\
&=\hat{g}_{(\alpha 2^{m_0})}(\xi)\hat{g}_{(\alpha 2^{m_0})}(\eta)\hat{g}_{(\beta 2^{m_0})}(\tau)\hat{g}_{(\beta 2^{m_0})}(-\tau-\xi-\eta)\\
&\qquad\qquad-\hat{g}_{(\alpha 2^{m_J})}(\xi)\hat{g}_{(\alpha 2^{m_J})}(\eta)\hat{g}_{(\beta 2^{m_J})}(\tau)\hat{g}_{(\beta 2^{m_J})}(-\tau-\xi-\eta),
\end{aligned}$$

where in the last identity we have telescoped the sum. We then obtain

$$|\Lambda_{D_2,K_9}((f_s)_{s\in S})|\lesssim 1$$

by a standard Brascamp–Lieb inequality analogously to the bound (8-1). This completes the bound for Λ_{D_2,K_1}.

It remains to estimate $\Lambda_{D_2,K-K_1}$. We have

$$(\widehat{K}-\widehat{K}_1)(\xi,\eta,\tau)=\alpha^{-N}\sum_{j=1}^{J}(\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\psi_j(\tau,-\tau-\xi-\eta)$$

with

$$\psi_j=\phi_j\otimes\phi_j-(\chi_{j-1}\otimes\chi_{j-1}-\chi_j\otimes\chi_j).$$

Define

$$\vartheta_{1,j}=\phi_j-\chi_{j-1},\quad \vartheta_{2,j}=\chi_{j-1}-(\chi_j)_{(2^{-4})},\quad \varrho_j=\psi_j-\vartheta_{1,j}\otimes\vartheta_{2,j}-\vartheta_{2,j}\otimes\vartheta_{1,j},$$

$$\widehat{K}_{10}(\xi,\eta,\tau)=\sum_{j=1}^{J}\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\hat{\vartheta}_{1,j}(\tau)\hat{\vartheta}_{2,j}(-\tau-\xi-\eta),$$

$$\widehat{K}_{11}(\xi,\eta,\tau)=\sum_{j=1}^{J}\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\hat{\vartheta}_{2,j}(\tau)\hat{\vartheta}_{1,j}(-\tau-\xi-\eta),$$

$$\widehat{K}_{12}(\xi,\eta,\tau)=\alpha^{-N}\sum_{j=1}^{J}\hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\hat{\varrho}_j(\tau,-\tau-\xi-\eta).$$

By the triangle inequality, it remains to estimate $\Lambda_{D_2,K_{10}}$, $\Lambda_{D_2,K_{11}}$, $\Lambda_{D_2,K_{12}}$, separately.

We begin with $\Lambda_{D_2,K_{10}}$. Recall that $(\chi_j)_{(2^{2-k_j})}$ is a left window. If $(\tau, -\tau-\xi-\eta)$ is in the support of $\hat{\vartheta}_{1,j} \otimes \hat{\vartheta}_{2,j}$, then

$$|\tau| \le 2^{-k_j+2}, \quad 2^{-k_j+5} \le |\tau+\xi+\eta| \le 2^{-k_{j-1}+2}, \quad 2^{-k_j+4} < |\xi+\eta| < 2^{-k_{j-1}+3}.$$

Defining

$$\vartheta_{3,j} := (\phi_j)_{(2^{-2})}$$

we have

$$\widehat{K}_{10}(\xi,\eta,\tau) = \sum_{j=1}^{J} \hat{g}_{(\alpha 2^{m_j})}(\xi)\hat{g}_{(\alpha 2^{m_j})}(\eta)\hat{\vartheta}_{1,j}(\tau)\hat{\vartheta}_{2,j}(-\tau-\xi-\eta)\hat{\vartheta}_{3,j}(\xi+\eta)^2$$

because the additional factor involving $\hat{\vartheta}_{3,j}$ is constant 1 on the support of the original summand in the definition of K_{10}. The bound

$$|\Lambda_{D_2,K_{10}}((f_s)_{s\in S})| \lesssim 1$$

then follows from Proposition 2.8 applied with

$$\rho_j := \hat{g}_{(\alpha 2^{m_j})} \otimes \hat{g}_{(\alpha 2^{m_j})} \otimes \hat{\vartheta}_{1,j} \otimes \hat{\vartheta}_{2,j}.$$

The form $\Lambda_{D_2,K_{11}}$ is estimated analogously to the form $\Lambda_{D_2,K_{10}}$. It remains to estimate $\Lambda_{D_2,K_{12}}$. This form is a more standard singular Brascamp–Lieb form with a kernel associated with a Hörmander–Mikhlin multiplier and we will apply Theorem 1.1 in [Durcik et al. 2022], which was the reason to set $N = 2^{18}$.

That theorem will give

$$|\Lambda_{D_2,K_{12}}((f_s)_{s\in S})| \lesssim 1$$

provided

$$|\partial^\gamma \widehat{K}_{12}(\xi,\eta,\tau)| \lesssim |(\xi,\eta,\tau)|^{-|\gamma|} \tag{8-3}$$

for all multiindices γ of order $0 \le |\gamma| \le N$. The assumption of that theorem that $\Pi_s \Pi^T$ is regular for the present datum D_2 is satisfied. It thus remains to show (8-3).

By definition of ψ_j and $\vartheta_{1,j}$, we obtain

$$\psi_j = \chi_{j-1} \otimes \vartheta_{1,j} + \vartheta_{1,j} \otimes \chi_{j-1} + \vartheta_{1,j} \otimes \vartheta_{1,j} + \chi_j \otimes \chi_j.$$

Using further the definition of ϱ_j and $\vartheta_{2,j}$, we obtain

$$\varrho_j = (\chi_j)_{(2^{-4})} \otimes \vartheta_{1,j} + \vartheta_{1,j} \otimes (\chi_j)_{(2^{-4})} + \vartheta_{1,j} \otimes \vartheta_{1,j} + \chi_j \otimes \chi_j. \tag{8-4}$$

Note that $\hat{\vartheta}_{1,j}$ vanishes outside

$$[-2^{-k_j+2}, 2^{-k_j+2}].$$

Hence $\hat{\varrho}_j$ is supported on the ball of radius 2^{10-k_j} around the origin. In addition, $\hat{\vartheta}_{1,j}$ coincides with -1 on $[-2^{-k_j+1}, 2^{-k_j+1}]$. Using that $(\chi_j)_{(2^{2-k_j})}$ is a left window, we then see that the Fourier transform of the first two terms on the right-hand side of (8-4) is equal to -1 on $[-2^{-k_j+1}, 2^{-k_j+1}]^2$ while the Fourier transform of the last two terms coincides with 1 on the same set. Therefore, $\hat{\varrho}_j$ vanishes inside the ball of radius 2^{-k_j} around the origin. The support properties of $\hat{\varrho}_j$ together with the estimates $|\hat{\varrho}_j| \lesssim 1$ and $g \lesssim 1$ yield that $|\widehat{K}_{12}| \lesssim 1$.

Assume next that β is a multiindex with $1 \le |\beta| \le N$. Then $\hat{\varrho}_j$ satisfies symbol estimates adapted to the ball of radius 2^{11-k_j} around the origin, namely

$$|\partial^\beta \hat{\varrho}_j(\tau,\sigma)| \lesssim 2^{k_j|\beta|} 1_{|(\tau,\sigma)| \le 2^{11-k_j}}.$$

Now assume first $|\xi-\eta| \le |(\xi+\eta,\tau)|$. Then, using that all derivatives of g up to order N are $\lesssim 1$, and using that $|m_j - k_j| \le 1$ and $\alpha \ge 1$,

$$|\partial^\beta \widehat{K}_{12}(\xi,\eta,\tau)| \lesssim \alpha^{-N} \sum_{j=1}^{J} (\alpha 2^{k_j})^{|\beta|} 1_{|(\tau,\tau+\xi+\eta)| \le 2^{11-k_j}}.$$

Using further that $\alpha \ge 1$ and $|\beta| \le N$ we estimate the last display by

$$\lesssim |(\tau,\xi+\eta)|^{-|\beta|} \lesssim |(\tau,\xi,\eta)|^{-|\beta|},$$

where in the last inequality we have used $|\xi-\eta| \le |(\tau,\xi+\eta)|$. Now assume to the contrary that $|\xi-\eta| \ge |(\xi+\eta,\tau)|$. Then we use that $|\partial^\beta g(\xi)| \lesssim e^{-|\xi|}$ for all $|\beta| \le N$. Then

$$\begin{aligned} |\partial^\beta \widehat{K}_{12}(\xi,\eta,\tau)| \lesssim \alpha^{-N} \sum_{j=1}^{J} (\alpha 2^{k_j})^{|\beta|} e^{-\alpha 2^{k_j}|\xi-\eta|} &\lesssim |\xi-\eta|^{-|\beta|} \sum_{j=1}^{J} (2^{k_j}|\xi-\eta|)^{|\beta|} e^{-2^{k_j}|\xi-\eta|} \\ &\lesssim |\xi-\eta|^{-|\beta|} \sum_{n\in\mathbb{Z}} 2^{n|\beta|} e^{-2^n} \lesssim |(\xi,\eta,\tau)|^{-|\beta|}. \end{aligned}$$

9. Proof of Propositions 2.8 and 2.9

The proofs of these propositions have some similarities, so we put them into one section and do the second proof analogously to the first.

9.1. ***Proof of Proposition 2.8.*** For $1 \le i \le 2$ let $(a_{i,j})_{j=1}^{J}$ be increasing sequences of positive real numbers, we choose $a_{i,0} > 0$ so that $(a_{i,j})_{j=0}^{J}$ is still increasing. For $1 \le j \le J$ let $\rho_j : \mathbb{R}^4 \to \mathbb{R}$ be a continuous function satisfying (2-14), and pick a further such function $\rho_0 : \mathbb{R}^4 \to \mathbb{R}$. Let $(c_j)_{j=0}^{J}$ be a well separated increasing sequence of positive numbers. Let χ be a left window, and let ϕ_j be a function on $\mathbb{R}$ which for $1 \le j \le J$ satisfy $\widehat{\phi}_j \ge 0$ and

$$(\widehat{\phi}_j)^2 = (\widehat{\chi}_{(c_{j-1})})^2 - (\widehat{\chi}_{(c_j)})^2.$$

Let K be defined by (2-15) and let a tuple $(f_s)_{s\in S}$ be given as in (2-12), (2-13).

The integrand of the integral expressing $\Lambda_{D_2,K}((f_s)_{s\in S})$ factors into functions depending on x_0 and functions depending on x_1. We write the integrals in x_0 and x_1 innermost and separate these. With $p_0 := p$

and $p_1 := q$ we obtain

$$\Lambda_{D_2,K}((f_s)_{s\in S}) = \sum_{j=1}^{J} \int_{\mathbb{R}^7} \Bigl[\prod_{i=0,1} \int_{\mathbb{R}} \Bigl[\prod_{q\in\mathcal{C}} f_{(i,q)}(\Pi_{(i,q)}x) \Bigr] \phi_j(x_i + p_i)\, dx_i \Bigr] \\ \times \rho_j\bigl(x_2^0 + p_0 + p_1 + r, x_2^1 + p_0 + p_1 + r, x_3^0 + r, x_3^1 + r\bigr)\, dx_2^0\, dx_2^1\, dx_3^0\, dx_3^1\, dp_0\, dp_1\, dr. \quad (9\text{-}1)$$

Applying the Cauchy–Schwarz inequality in the seven exterior variables bounds the last display by the geometric mean of two forms, parametrized by $i = 0, 1$, which with the change of variables $p_{1-i} \to p_{1-i} - p_i - r$ we write as

$$\sum_{j=1}^{J} \int_{\mathbb{R}^7} \Bigl[\int_{\mathbb{R}} \Bigl[\prod_{q\in\mathcal{C}} f_{(i,q)}(\Pi_{(i,q)}x) \Bigr] \phi_j(x_i + p_i)\, dx_i \Bigr]^2 \\ \times |\rho_j|\bigl(x_2^0 + p_{1-i}, x_2^1 + p_{1-i}, x_3^0 + r, x_3^1 + r\bigr)\, dx_2^0\, dx_2^1\, dx_3^0\, dx_3^1\, dp_0\, dp_1\, dr. \quad (9\text{-}2)$$

Fix i and write f for $f_{(i,j)}$, which thanks to (2-12) does not depend on j.

Using the decay (2-14) for ρ_j, we dominate

$$\int_{\mathbb{R}^2} |\rho_j|\bigl(x_2^0 + p_{1-i}, x_2^1 + p_{1-i}, x_3^0 + r, x_3^1 + r\bigr)\, dp_{1-i}\, dr \\ \lesssim \int_1^\infty \int_1^\infty (g*g)_{(\alpha a_{1,j})}(x_2^0 - x_2^1)(g*g)_{(\beta a_{2,j})}(x_3^0 - x_3^1)\, \frac{d\alpha}{\alpha^2}\, \frac{d\beta}{\beta^2}. \quad (9\text{-}3)$$

It suffices to consider fixed α and β, and prove uniform bounds in α and β for (9-2) with (9-3) replaced by

$$(g*g)_{(\alpha a_{1,j})}(x_2^0 - x_2^1)(g*g)_{(\beta a_{2,j})}(x_3^0 - x_3^1).$$

Modifying the sequences $a_{i,j}$ if necessary, we may assume $\alpha = \beta = 1$.

Expanding the square in (9-2) and integrating in p_i, our task becomes to show

$$\sum_{j=1}^{J} \int_{\mathbb{R}^6} \Bigl[\prod_{s\in S} f(\Pi_s x) \Bigr] (\phi_j * \phi_j)(x_1^0 - x_1^1)(g*g)_{a_{1,j}}(x_2^0 - x_2^1)(g*g)_{a_{2,j}}(x_3^0 - x_3^1)\, dx \lesssim 1, \quad (9\text{-}4)$$

where S and $(\Pi_s)_{s\in S}$ are as in the datum D_{-I}, which is the datum defined in (2-20) in the case $A = -I$.

Define the kernels

$$K_1 := \sum_{j=1}^{J} \bigl((\chi*\chi)_{(c_{j-1})} - (\chi*\chi)_{(c_j)}\bigr) \otimes (g*g)_{(a_{1,j})} \otimes (g*g)_{(a_{2,j})},$$

$$K_2 := \sum_{j=1}^{J} (\chi*\chi)_{(c_{j-1})} \otimes \bigl((g*g)_{(a_{1,j-1})} - (g*g)_{(a_{1,j})}\bigr) \otimes (g*g)_{(a_{2,j})},$$

$$K_3 := \sum_{j=1}^{J} (\chi*\chi)_{(c_{j-1})} \otimes (g*g)_{(a_{1,j-1})} \otimes \bigl((g*g)_{(a_{2,j-1})} - (g*g)_{(a_{2,j})}\bigr),$$

and for $0 \le j \le J$ also

$$\sigma_j = (\chi*\chi)_{(c_j)} \otimes (g*g)_{(a_{1,j})} \otimes (g*g)_{(a_{2,j})}.$$

We have the telescoping identity

$$K_1 + K_2 + K_3 = \sigma_0 - \sigma_J. \tag{9-5}$$

The form (9-4) to be estimated becomes $\Lambda_{D_{-I},K_1}((f)_{s\in S})$. For each $0 \le j \le J$ one has by a standard Brascamp–Lieb inequality

$$|\Lambda_{D_{-I},\sigma_j}((f)_{s\in S})| \lesssim 1.$$

It then suffices to estimate the forms associated with K_2 and K_3 instead. By symmetry, we will only elaborate on $\Lambda_{D_{-I},K_2}((f)_{s\in S})$.

Next we would like to dominate $|(\chi * \chi)_{(c_{j-1})}|$ in these two forms by superposition of Gaussians in such a way that the cancellation is preserved. To do that, we will use the identity

$$(g*g)_{(a_{1,j-1})} - (g*g)_{(a_{1,j})} = -\frac{1}{\pi}\int_{a_{1,j-1}}^{a_{1,j}} (h*h)_{(t)}\,\frac{dt}{t}, \tag{9-6}$$

which follows by taking the Fourier transform of the identity

$$g(a\xi)^2 - g(b\xi)^2 = -\int_a^b \partial_t g(t\xi)^2\,dt = \frac{1}{\pi}\int_a^b (2\pi t\xi g(t\xi))^2\,\frac{dt}{t} = -\frac{1}{\pi}\int_a^b (\hat{h}(t\xi))^2\,\frac{dt}{t}$$

for any $a, b > 0$. Using further that h is odd and thus

$$-h*h(x-y) = \int_{\mathbb{R}} h(x+p)h(y+p)\,dp,$$

we obtain

$$\Lambda_{D_{-I},K_2}((f)_{s\in S}) = \frac{1}{\pi}\sum_{j=1}^{J}\int_{a_{1,j-1}}^{a_{1,j}}\int_{\mathbb{R}^5}\Bigl[\prod_{i=0,1}\int_{\mathbb{R}}\Bigl[\prod_{s(1)=i} f(\Pi_s x)\Bigr]h_{(t)}(x_2^i+p)\,dx_2^i\Bigr]$$
$$\times(\chi*\chi)_{(c_{j-1})}(x_1^0-x_1^1)(g*g)_{(a_{2,j})}(x_3^0-x_3^1)\,dx_1^0\,dx_1^1\,dx_3^0\,dx_3^1\,dp\,\frac{dt}{t}. \tag{9-7}$$

The product over $i = 0, 1$ has two identical factors and thus is nonnegative. We may therefore estimate the last display by dominating

$$|(\chi*\chi)_{(c_{j-1})}| \lesssim \int_1^\infty (g*g)_{(\beta c_{j-1})}\beta^{-5}\,d\beta.$$

It suffices to prove bounds of (9-7) with $(\chi*\chi)_{(c_{j-1})}$ replaced by $(g*g)_{(\beta c_{j-1})}$ uniformly in β. Fix β. By changing c_j if necessary, we may assume $\beta = 1$. Define again kernels

$$K_4 := \sum_{j=1}^{J}\bigl((g*g)_{(c_{j-1})} - (g*g)_{(c_j)}\bigr)\otimes(g*g)_{(a_{1,j})}\otimes(g*g)_{(a_{2,j})},$$

$$K_5 := \sum_{j=1}^{J}(g*g)_{(c_{j-1})}\otimes\bigl((g*g)_{(a_{1,j-1})} - (g*g)_{(a_{1,j})}\bigr)\otimes(g*g)_{(a_{2,j})},$$

$$K_6 := \sum_{j=1}^{J}(g*g)_{(c_{j-1})}\otimes(g*g)_{(a_{1,j-1})}\otimes\bigl((g*g)_{(a_{2,j-1})} - (g*g)_{(a_{2,j})}\bigr).$$

Similarly as near (9-5),

$$\Lambda_{D_{-I},K_4}((f)_{s\in S}) + \Lambda_{D_{-I},K_5}((f)_{s\in S}) + \Lambda_{D_{-I},K_6}((f)_{s\in S}) \tag{9-8}$$

telescopes into a form that is $\lesssim 1$ by a standard Brascamp–Lieb inequality. We have seen above that $\Lambda_{D_{-I},K_5}((f)_{s\in S})$ is positive. By symmetric arguments, the other summands in (9-8) are also positive. Hence each summand is $\lesssim 1$. This completes the proof of Proposition 2.8.

9.2. *Proof of Proposition 2.9*. Let a positive integer J be given as well as increasing sequences of positive real numbers $(a_j)_{j=0}^J$, $(b_j)_{j=1}^J$. Pick $b_0 > 0$ so that $(b_j)_{j=0}^J$ is an increasing sequence. For $1 \le j \le J$ let ϕ_j be given as in (2-17). Let K be defined by (2-18). Let a tuple $(f_s)_{s\in S}$ be given as in (2-12), (2-13).

We write

$$\begin{aligned} t^2(\xi+\eta)^2 g(t\xi)g(t\eta) &= t^2(\xi+\eta)^2 g(2^{-3/2}t(\xi+\eta))^2 g(2^{-1}t(\xi-\eta))g(2^{-1/2}t\xi)g(2^{-1/2}t\eta) \\ &= -\hat{h}(2^{-3/2}t(\xi+\eta))^2\hat{\rho}(2^{-3/2}t(\xi,\eta)) \end{aligned}$$

with

$$\hat{\rho}(2^{-3/2}(\xi,\eta)) := \frac{2}{\pi^2} g(2^{-1}(\xi-\eta))g(2^{-1/2}\xi)g(2^{-1/2}\eta).$$

Hence passing to the spatial side as near (2-16), replacing the arbitrary sequence a_j by $2^{-3/2}a_j$ to avoid the cumbersome factors $2^{-3/2}$,

$$K(u,v,z) = \sum_{j=1}^J \int_{\mathbb{R}^3}\int_{a_{j-1}}^{a_j} h_{(t)}(p)h_{(t)}(q)\rho_{(t)}(u+p+q+r, v+p+q+r)\,\frac{dt}{t}\phi_j(z+r,r)\,dp\,dq\,dr.$$

We thus have analogously to (9-1),

$$\begin{aligned} \Lambda_{D_2,K}((f_s)_{s\in S}) = \sum_{j=1}^J \int_{\mathbb{R}^7}\int_{a_{j-1}}^{a_j} \biggl[\prod_{i=0,1}\int_{\mathbb{R}}\biggl[\prod_{q\in\mathcal{C}} f_{(i,q)}(\Pi_{(i,q)}x)\biggr]h_{(t)}(x_i+p_i)\,dx_i\biggr] \\ \times \rho_{(t)}(x_2^0+p_0+p_1+r, x_2^1+p_0+p_1+r)\phi_j(x_3^0+r, x_3^1+r)\,\frac{dt}{t}\,dx_2^0\,dx_2^1\,dx_3^0\,dx_3^1\,dp_0\,dp_1\,dr. \end{aligned}$$

Applying the Cauchy–Schwarz inequality as in (9-2), we need to estimate for $i = 0, 1$,

$$\begin{aligned} \sum_{j=1}^J \int_{\mathbb{R}^7}\int_{a_{j-1}}^{a_j} \biggl[\int_{\mathbb{R}}\biggl[\prod_{q\in\mathcal{C}} f_{(i,q)}(\Pi_{(i,q)}x)\biggr]h_{(t)}(x_i+p_i)\,dx_i\biggr]^2 \\ \times |\rho_{(t)}|(x_2^0+p_{1-i}, x_2^1+p_{1-i})|\phi_j|(x_3^0+r, x_3^1+r)\,\frac{dt}{t}\,dx_2^0\,dx_2^1\,dx_3^0\,dx_3^1\,dp_0\,dp_1\,dr. \end{aligned}$$

Thanks to the square, the above integrand is positive and we dominate

$$|\rho_{(t)}| \lesssim g_{(4t)}\otimes g_{(4t)} \quad\text{and}\quad |\phi_j| \lesssim \int_1^\infty g_{(\beta b_j)}\otimes g_{(\beta b_j)}\beta^{-3}\,d\beta.$$

It suffices to prove bounds with $g_{(\beta b_j)}\otimes g_{(\beta b_j)}$ in place of $|\phi_j|$ uniformly in β. Fix β; we may assume $\beta = 1$ by modifying the otherwise arbitrary sequence b_j.

Performing the analogous steps as leading to (9-4) we end up having to estimate $\Lambda_{D_{-I},K_1}((f)_{s\in S})$, where now

$$K_1 := -\sum_{j=1}^J \int_{a_{j-1}}^{a_j} (h*h)_{(t)}\otimes(g*g)_{(4t)}\,\frac{dt}{t}\otimes(g*g)_{(b_j)}.$$

Define

$$K_2 := -\sum_{j=1}^{J} \int_{a_{j-1}}^{a_j} (g*g)_{(t)} \otimes (h*h)_{(4t)} \frac{dt}{t} \otimes (g*g)_{(b_j)},$$

$$K_3 := -\sum_{j=1}^{J} (g*g)_{(a_{j-1})} \otimes (g*g)_{(4a_{j-1})} \otimes \int_{b_{j-1}}^{b_j} (h*h)_{(t)} \frac{dt}{t},$$

and for $0 \le j \le J$ also

$$\sigma_j = (g*g)_{(a_j)} \otimes (g*g)_{(4a_j)} \otimes (g*g)_{(b_j)}.$$

Then we have the telescoping identity

$$K_1 + K_2 + K_3 = \pi(\sigma_0 - \sigma_J). \tag{9-9}$$

Indeed, this follows with (9-6), which gives

$$K_1 + K_2 = \pi \sum_{j=1}^{J} \int_{a_{j-1}}^{a_j} -t\partial_t\big((g*g)_{(t)} \otimes (g*g)_{(4t)}\big) \frac{dt}{t} \otimes (g*g)_{(b_j)}$$

$$= \pi \sum_{j=1}^{J} \big((g*g)_{(a_{j-1})} \otimes (g*g)_{(4a_{j-1})} - (g*g)_{(a_j)} \otimes (g*g)_{(4a_j)}\big) \otimes (g*g)_{(b_j)},$$

$$K_3 = \pi \sum_{j=1}^{J} (g*g)_{(a_{j-1})} \otimes (g*g)_{(4a_{j-1})} \otimes \big((g*g)_{(b_{j-1})} - (g*g)_{(b_j)}\big).$$

By the identity (9-9),

$$\Lambda_{D_{-I},K_1}((f)_{s\in S}) + \Lambda_{D_{-I},K_2}((f)_{s\in S}) + \Lambda_{D_{-I},K_3}((f)_{s\in S}) \lesssim 1.$$

All quantities on the left-hand side are nonnegative. For $\Lambda_{D_{-I},K_1}((f)_{s\in S})$, this can be seen as it resulted after an application of the Cauchy–Schwarz inequality, while for $\Lambda_{D_{-I},K_2}((f)_{s\in S})$ and $\Lambda_{D_{-I},K_3}((f)_{s\in S})$ it follows by symmetry. This gives the desired upper bound

$$\Lambda_{D_{-I},K_1}((f)_{s\in S}) \lesssim 1.$$

Acknowledgements

The authors thank Terence Tao for an exchange on the introduction of an earlier version of the paper. Durcik was partially supported by the NSF grant DMS-2154356. Slavíková was supported by the Primus research programme PRIMUS/21/SCI/002 of Charles University and by Charles University Research program UNCE/SCI/023. Thiele was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy — EXC-2047/1–390685813, as well as SFB 1060. Durcik and Thiele were also supported by the NSF grant DMS-1929284 while the authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the inspiring Harmonic Analysis and Convexity program. The authors thank the referees for their suggestions.

References

[Austin 2010] T. Austin, "On the norm convergence of non-conventional ergodic averages", *Ergodic Theory Dynam. Systems* **30**:2 (2010), 321–338. MR

[Avigad and Rute 2015] J. Avigad and J. Rute, "Oscillation and the mean ergodic theorem for uniformly convex Banach spaces", *Ergodic Theory Dynam. Systems* **35**:4 (2015), 1009–1027. MR

[Birkhoff 1931] G. D. Birkhoff, "Proof of the ergodic theorem", *Proc. Nat. Acad. Sci.* **17**:12 (1931), 656–660.

[Bourgain 1989] J. Bourgain, "Pointwise ergodic theorems for arithmetic sets", *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), 5–45. MR

[Bourgain et al. 2023] J. Bourgain, M. Mirek, E. M. Stein, and J. Wright, "On a multi-parameter variant of the Bellow–Furstenberg problem", *Forum Math. Pi* **11** (2023), art. id. e23. MR

[Durcik and Kovač 2022] P. Durcik and V. Kovač, "A Szemerédi-type theorem for subsets of the unit cube", *Anal. PDE* **15**:2 (2022), 507–549. MR

[Durcik and Stipčić 2025] P. Durcik and M. Stipčić, "Quantitative bounds for products of simplices in subsets of the unit cube", *Israel J. Math.* **267**:1 (2025), 135–156. MR

[Durcik and Thiele 2020] P. Durcik and C. Thiele, "Singular Brascamp–Lieb inequalities with cubical structure", *Bull. Lond. Math. Soc.* **52**:2 (2020), 283–298. MR

[Durcik and Thiele 2021] P. Durcik and C. Thiele, "Singular Brascamp–Lieb: a survey", pp. 321–349 in *Geometric aspects of harmonic analysis*, edited by P. Ciatti and A. Martini, Springer INdAM Ser. **45**, Springer, 2021. MR

[Durcik et al. 2019a] 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**:3 (2019), 658–688. MR

[Durcik et al. 2019b] P. Durcik, V. Kovač, and C. Thiele, "Power-type cancellation for the simplex Hilbert transform", *J. Anal. Math.* **139**:1 (2019), 67–82. MR

[Durcik et al. 2022] P. Durcik, L. Slavíková, and C. Thiele, "Local bounds for singular Brascamp–Lieb forms with cubical structure", *Math. Z.* **302**:4 (2022), 2375–2405. MR

[Host 2009] B. Host, "Ergodic seminorms for commuting transformations and applications", *Studia Math.* **195**:1 (2009), 31–49. MR

[Jones et al. 1996] R. L. Jones, I. V. Ostrovskii, and J. M. Rosenblatt, "Square functions in ergodic theory", *Ergodic Theory Dynam. Systems* **16**:2 (1996), 267–305. MR

[Jones et al. 2008] R. L. Jones, A. Seeger, and J. Wright, "Strong variational and jump inequalities in harmonic analysis", *Trans. Amer. Math. Soc.* **360**:12 (2008), 6711–6742. MR

[Kovač 2016] V. Kovač, "Quantitative norm convergence of double ergodic averages associated with two commuting group actions", *Ergodic Theory Dynam. Systems* **36**:3 (2016), 860–874. MR

[Krause et al. 2022] B. Krause, M. Mirek, and T. Tao, "Pointwise ergodic theorems for non-conventional bilinear polynomial averages", *Ann. of Math.* (2) **195**:3 (2022), 997–1109. MR

[Muscalu and Zhai 2022] C. Muscalu and Y. Zhai, "Five-linear singular integral estimates of Brascamp–Lieb-type", *Anal. PDE* **15**:4 (2022), 1011–1095. MR

[Tao 2008] T. Tao, "Norm convergence of multiple ergodic averages for commuting transformations", *Ergodic Theory Dynam. Systems* **28**:2 (2008), 657–688. MR

[Walsh 2012] M. N. Walsh, "Norm convergence of nilpotent ergodic averages", *Ann. of Math.* (2) **175**:3 (2012), 1667–1688. MR

Received 30 May 2024. Revised 7 Mar 2025. Accepted 28 Apr 2025.

POLONA DURCIK: durcik@chapman.edu
Schmid College of Science and Technology, Chapman University, Orange, CA, United States

LENKA SLAVÍKOVÁ: slavikova@karlin.mff.cuni.cz
Department of Mathematical Analysis, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

CHRISTOPH THIELE: thiele@math.uni-bonn.de
Mathematisches Institut, Universität Bonn, Bonn, Germany

mathematical sciences publishers msp

# Analysis & PDE

msp.org/apde

Cover image: Eric J. Heller: "Linear Ramp"

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2026 is US $500/year for the electronic version, and $780/year (+$75, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 19 No. 3 2026