A Study Guide for the l^2 Decoupling Theorem[*]

Jean BOURGAIN[1]   Ciprian DEMETER[2]

(*Dedicated to Haim Brezis on the occasion of his* 70*th birthday*)

Abstract This paper contains a detailed, self contained and more streamlined proof of the l^2 decoupling theorem for hypersurfaces from the paper of Bourgain and Demeter in 2015. The authors hope this will serve as a good warm up for the readers interested in understanding the proof of Vinogradov's mean value theorem from the paper of Bourgain, Demeter and Guth in 2015.

Keywords Restriction theorems, Multilinear Kakeya inequality, Decouplings
2000 MR Subject Classification 42B37

1 The Theorem

Consider the truncated (elliptic) paraboloid in $\mathbb{R}^n$,

$$\mathbb{P}^{n-1} := \{(\xi_1, \cdots, \xi_{n-1}, \xi_1^2 + \cdots + \xi_{n-1}^2) : \ 0 \le \xi_i \le 1\}.$$

For each cube Q in $[0,1]^{n-1}$ and $g : Q \to \mathbb{C}$, define the extension operator $E_Q^{(n)} = E_Q$ as follows:

$$E_Q g(x) = \int_Q g(\xi_1, \cdots, \xi_{n-1}) e(\xi_1 x_1 + \cdots + \xi_{n-1} x_{n-1} + (\xi_1^2 + \cdots + \xi_{n-1}^2) x_n) \mathrm{d}\xi,$$

where $e(z) = \mathrm{e}^{2\pi \mathrm{i} z}$,

$$\xi = (\xi_1, \cdots, \xi_{n-1})$$

and

$$x = (x_1, \cdots, x_n).$$

This can be interpreted as the Fourier transform $\widehat{gd\sigma}$, where the measure $d\sigma$ is the lift of the Lebesgue measure from $[0,1]^{n-1}$ to the paraboloid. When $Q = [0,1]^{n-1}$, we will sometimes write Eg for $E_{[0,1]^{n-1}} g$.

We will use the letters Q, q to denote cubes on the frequency side $[0,1]^{n-1}$. We will use the letters B, Δ to denote cubes on the spatial side $\mathbb{R}^n$. Throughout the whole paper we can and will implicitly assume that all cubes have side length in $2^{\mathbb{Z}}$. This in particular will place (harmless) restrictions on various parameters such as δ, σ, R, that we will not bother to write

down explicitly. Thanks to this assumption, we will be able to partition (rather than use finitely overlapping covers) large cubes into smaller cubes. Given a cube $Q \subset [0,1]^{n-1}$ with side length $l(Q) \in 2^{-\mathbb{N}}$ and $\alpha \in 2^{-\mathbb{N}}$ smaller than $l(Q)$, we will denote by $\mathrm{Part}_\alpha(Q)$ the (unique) partition of Q by using cubes Q_α of side length α. A similar notation will occasionally be used for spatial cubes B.

We will write $B = B(c_B, R)$ for the cube in $\mathbb{R}^n$ centered at c_B and with side length $l(B_R) = R$, and we will introduce the associated weight

$$w_B(x) = \frac{1}{\left(1 + \frac{|x - c_B|}{R}\right)^{100n}}.$$

The exponent $100n$ is chosen large enough to guarantee various integrability requirements. We will see that Theorem 1.1 remains true for any larger exponent $E \geq 100n$, and the implicit bounds will depend on E. This observation will allow us to run our induction argument, as explained in Section 3.

For a positive weight $v : \mathbb{R}^n \to [0, \infty)$ and for $f : \mathbb{R}^n \to \mathbb{C}$, we define the weighted integral

$$\|f\|_{L^p(v)} = \left(\int_{\mathbb{R}^n} |f(x)|^p v(x) \mathrm{d}x\right)^{\frac{1}{p}}.$$

For $2 \leq p \leq \infty$ and $\delta \in 4^{-\mathbb{N}}$, let $\mathrm{Dec}(\delta, p) = \mathrm{Dec}_n(\delta, p)$ be the smallest constant, such that the inequality

$$\|Eg\|_{L^p(w_B)} \leq \mathrm{Dec}(\delta, p) \Big(\sum_{Q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}([0,1]^{n-1})} \|E_Q g\|_{L^p(w_B)}^2 \Big)^{\frac{1}{2}}$$

holds for every cube $B \subset \mathbb{R}^n$ with side length δ^{-1} and every $g : [0,1]^{n-1} \to \mathbb{C}$.

The l^2 decoupling theorem proved in [3] reads as follows. We refer the reader to [3] for a few applications that motivate the theorem.

Theorem 1.1 *We have the following sharp (up to $\delta^{-\epsilon}$ losses) upper bound for $\mathrm{Dec}_n(\delta, p)$:*

$$\mathrm{Dec}_n(\delta, p) \lesssim_{\epsilon, p, n} \delta^{-\epsilon},$$

if $2 \leq p \leq \frac{2(n+1)}{n-1}$. The implicit constant depends on ϵ, p, n but not on δ.

We will present a rather detailed argument for this theorem. Essentially, we rewrite our original argument from [3] using a more streamlined approach. This approach has started to take shape in our subsequent papers on decouplings and has gotten to this final form in the joint work with Guth [4]. One new feature of our argument compared to [3] is that we avoid the special interpolation from [3], that relies on wave-packet decomposition. Another one is that we use the multilinear Kakeya inequality, rather than the multilinear restriction theorem. The argument we describe here also clarifies various technical aspects of the theory, such as the role of the weights w_B and the (essentially) locally constant behavior of Fourier transforms of measures supported on caps on the paraboloid.

We hope the argument will be accessible to experts outside the area of harmonic analysis. We believe this will serve as a warm up for the readers interested in understanding the proof of Vinogradov's mean value theorem from [4].

A brief summary of the argument is presented in Section 3. The most important sections are the last two. The details from the remaining sections may be skipped at the first reading.

2 More Notation

Throughout this paper we will write $A \lesssim_v B$ to denote the fact that $A \leq CB$, for a certain implicit constant C, that depends on the parameter v. Typically, this parameter is either ϵ, ν or K. The implicit constant will never depend on the scale δ, on the spatial cubes we integrate over, or on the function g. It will however most of the times depend on the degree n and on the Lebesgue index p. Since these can be thought of as being fixed parameters, we will in general not write $\lesssim_{p,n}$.

We will denote by B_R an arbitrary cube in $\mathbb{R}^n$ of side length $l(B_R) = R$. We use the following two notations for averaged integrals:

$$\fint_B F = \frac{1}{|B|} \int_B F,$$

$$\|F\|_{L_\sharp^p(w_B)} = \Big(\frac{1}{|B|} \int |F|^p w_B\Big)^{\frac{1}{p}}.$$

Given a function η on $\mathbb{R}^n$ and a cube $B = B(c, R)$ in $\mathbb{R}^n$, we will use the rescaled version

$$\eta_B(x) = \eta\Big(\frac{x - c}{R}\Big).$$

$|A|$ will refer to either the cardinality of A if A is finite, or to its Lebesgue measure if A has positive measure.

We will sometimes write $\langle f, g \rangle$ for the inner product $\int f \overline{g}$.

3 A Brief Description of the Argument

We use two types of mechanisms to decouple. One is the L^2 decoupling (see Section 6). This is very basic, it relies just on Hilbert space orthogonality, but it is nevertheless very efficient. It decouples into frequency cubes whose side length is as small as permitted by the uncertainty principle, namely equal to the reciprocal of the side length of the spatial cube. The second mechanism is a multilinear decoupling that relies on the multilinear Kakeya inequality (see Theorem 9.2). Combining these with multiple iterations leads to the multiscale inequality (10.7). This inequality has a very simple form when $2 \leq p \leq \frac{2n}{n-1}$, and a short warm up argument is presented in the end of Section 10 to prove Theorem 1.1 in this range.

For the general case, the argument will go as follows. We will introduce a family of constants $\mathrm{Dec}_n(\delta, p, \nu, m)$, and will show in Section 8 that they dominate $\mathrm{Dec}_n(\delta, p)$. On the other hand, in the last section, we use (10.7) to show that each $\mathrm{Dec}_n(\delta, p, \nu, m)$ can be controlled by a combination of powers of δ and some power of $\mathrm{Dec}_n(\delta, p)$ (see (11.10)). This inequality represents an improvement over the trivial estimate $\mathrm{Dec}_n(\delta, p, \nu, m) \lesssim \mathrm{Dec}_n(\delta, p)$. By playing the two bounds (see (11.10) and (8.3)) against each other, we arrive at the desired upper bound

$$\mathrm{Dec}_n(\delta, p) \lesssim_\epsilon \delta^{-\epsilon}.$$

An unfortunate technicality is the fact that we will need to work with the family of weights for a cube $B = B(c, R)$ in $\mathbb{R}^n$,

$$w_{B,E}(x) = \frac{1}{(1 + \frac{|x-c|}{R})^E}.$$

Here $E \geq 100n$. For each such exponent E, we will let, as before, $\mathrm{Dec}_n(\delta, p, E)$ denote the smallest constant that guarantees the following inequality for each $g, B = B_{\delta^{-1}}$:

$$\|Eg\|_{L^p(w_{B,E})} \leq \mathrm{Dec}_n(\delta, p, E) \Big(\sum_{Q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}([0,1]^{n-1})} \|E_Q g\|_{L^p(w_{B,E})}^2 \Big)^{\frac{1}{2}}.$$

All the quantities that will depend on weights will implicitly depend on E. This includes $\mathrm{Dec}_n(\delta, p, \nu, m)$, $D_t(q, B^r, g)$ and $A_p(q, B^r, s, g)$. Sometimes we will suppress the dependence on E, and will understand implicitly that the inequality is true for all $E \geq 100n$. The weight $w_{B,E}$ will always be the same on both sides of a given inequality. The implicit constants will depend on E but that is completely harmless.

We will prove Theorem 1.1 by using induction on the dimension n. We set a superficially stronger induction hypothesis, namely, we will assume that

$$\mathrm{Dec}_{n-1}(\delta, p, E) \lesssim_{\epsilon, E} \delta^{-\epsilon}$$

for each $2 \leq p \leq \frac{2n}{n-2}$ and each $E \geq 100(n-1)$. We will use this to prove

$$\mathrm{Dec}_n(\delta, p, E) \lesssim_{\epsilon, E} \delta^{-\epsilon}$$

for each $2 \leq p \leq \frac{2(n+1)}{n-1}$ and each $E \geq 100n$. The reason for such a hypothesis is coming from inequality (8.8), which essentially uses the lower dimensional constant $\mathrm{Dec}_{n-1}(\delta, p, F)$ to make a statement about $\mathrm{Dec}_n(\delta, p, E)$. Larger dimensions demand higher values of E due to integrability requirements.

4 A Few Useful Inequalities

One technical challenge in the proof of Theorem 1.1 is to preserve the exponent E for the weights w_B involved in various inequalities.

A key, easy to check property of the weights $w_B = w_{B,E}$ that will be used extensively, is the following inequality:

$$1_B \lesssim \sum_{\Delta \in \mathcal{B}} w_\Delta \lesssim w_B, \tag{4.1}$$

valid for all cubes B with $l(B) = R$ and all finitely overlapping covers $\mathcal{B}$ of B with cubes Δ of (fixed) side length $1 \leq R' \leq R$. The implicit constants in (4.1) will (harmlessly) depend on E, but crucially, they will be independent of R, R'.

We will find extremely useful the following simple result.

Lemma 4.1 *Let $\mathcal{W}$ be the collection of all weights, that is, positive, integrable functions on $\mathbb{R}^n$. Fix $R > 0$. Fix E. Let $O_1, O_2 : \mathcal{W} \to [0, \infty]$ have the following four properties:*

(W1) $O_1(1_B) \lesssim O_2(w_{B,E})$ *for all cubes $B \subset \mathbb{R}^n$ of side length R.*
(W2) $O_1(\alpha u + \beta v) \leq \alpha O_1(u) + \beta O_1(v)$ *for each $u, v \in \mathcal{W}$ and $\alpha, \beta > 0$.*
(W3) $O_2(\alpha u + \beta v) \geq \alpha O_2(u) + \beta O_2(v)$ *for each $u, v \in \mathcal{W}$ and $\alpha, \beta > 0$.*
(W4) *If $u \leq v$, then $O_i(u) \leq O_i(v)$.*
Then

$$O_1(w_{B,E}) \lesssim O_2(w_{B,E})$$

for each cube B with side length R. The implicit constant is independent of R, B and only depends on the implicit constant from (W1), on E and n.

We will sometimes be able to check a stronger assumption than (W1), where $O_2(w_B)$ is replaced with $O_2(\eta_B)$ for some rapidly decreasing function η.

Proof Let $\mathcal{B}$ be a finitely overlapping cover of $\mathbb{R}^n$ with cubes $B' = B'(c_{B'}, R)$. It suffices to note that

$$w_B(x) \lesssim \sum_{B' \in \mathcal{B}} 1_{B'}(x) w_B(c_{B'})$$

and

$$\sum_{B' \in \mathcal{B}} w_{B'}(x) w_B(c_{B'}) \lesssim w_B(x).$$

Remark 4.1 It is rather immediate that for each f,

$$O_1(v) := \|f\|_{L^p(v)}^p$$

satisfies (W2) and (W4). Also, for fixed $p \geq 2$ and f_i, Minkowski's inequality in $l_{\frac{2}{p}}$ shows that

$$O_2(v) := \left(\sum_i \|f_i\|_{L^p(v)}^2 \right)^{\frac{p}{2}}$$

satisfies (W3) and (W4). Most applications of Lemma 4.1 will use this type of operators.

We complete this section with the following reverse Hölder inequality.

Corollary 4.1 *For each $q \geq p \geq 1$, each cube $Q \subset [0,1]^{n-1}$ with $l(Q) = \frac{1}{R}$ and each cube B in $\mathbb{R}^n$ with $l(B) = R$, we have*

$$\|E_Q g\|_{L^q_\sharp(w_{B,E})} \lesssim \|E_Q g\|_{L^p_\sharp(w_{B, \frac{Ep}{q}})} \tag{4.2}$$

with the implicit constant independent of R, Q, B and g.

Proof Let η be a positive smooth function on $\mathbb{R}^n$ satisfying $1_{B(0,1)} \leq \eta_{B(0,1)}$, such that the Fourier transform of $\eta^{\frac{1}{p}}$ is supported on the cube $B(0,1)$. We can thus write

$$\|E_Q g\|_{L^q(B)} \leq \|E_Q g\|_{L^q(\eta_B^{\frac{q}{p}})} = \|\eta_B^{\frac{1}{p}} E_Q g\|_{L^q(\mathbb{R}^n)}.$$

Let θ be a Schwartz function which equals to 1 on the cube $B(0, 10)$. Since the Fourier transform of $\eta_B^{\frac{1}{p}} E_Q g$ is supported in the cube $3Q$ (having the same center as Q and side length three times as large), we have

$$\eta_B^{\frac{1}{p}} E_Q g = (\eta_B^{\frac{1}{p}} E_Q g) * \widehat{\theta_Q},$$

and thus, by Young's inequality, we can write

$$\|\eta_B^{\frac{1}{p}} E_Q g\|_{L^q(\mathbb{R}^n)} \leq \|\eta_B^{\frac{1}{p}} E_Q g\|_{L^p(\mathbb{R}^n)} \|\widehat{\theta_Q}\|_{L^r(\mathbb{R}^n)} \lesssim R^{-\frac{n}{r'}} \|E_Q g\|_{L^p(\eta_B)}.$$

Here

$$\frac{1}{q} = \frac{1}{p} + \frac{1}{r} - 1 = \frac{1}{p} - \frac{1}{r'}.$$

Now, following the notation and ideas from the proof of Lemma 4.1, we may use the previous inequalities to write

$$\begin{aligned}
\int |E_Q g|^q w_{B,E} &\lesssim \sum_{B' \in \mathcal{B}} w_{B,E}(c_{B'}) \int_{B'} |E_Q g|^q \\
&\lesssim R^{-\frac{nq}{r'}} \sum_{B' \in \mathcal{B}} w_{B,E}(c_{B'}) \|E_Q g\|_{L^p(\eta_{B'})}^q \\
&\lesssim R^{-\frac{nq}{r'}} \Big(\sum_{B' \in \mathcal{B}} [w_{B,E}(c_{B'})]^{\frac{p}{q}} \|E_Q g\|_{L^p(\eta_{B'})}^p \Big)^{\frac{q}{p}} \\
&\lesssim R^{-\frac{nq}{r'}} \Big(\int |E_Q g|^p w_{B,\frac{Ep}{q}} \Big)^{\frac{q}{p}}.
\end{aligned}$$

Remark 4.2 Note that there is a loss in regularity in (4.2), as the weight exponent is $\frac{Ep}{q}$ on the right-hand side. A simple example shows that this exponent is optimal. This will later cause some minor technicalities. In particular, it will force us to use the smaller weight $w_{\Delta,10E}$ (as opposed to $w_{\Delta,E}$) in (8.1). This will in turn allow us to go from (11.8) to (11.9) by using (4.2) for indices p and 2.

5 An Equivalent Formulation

For $\delta < 1$ and $Q \subset [0,1]^{n-1}$, define the δ-neighborhood of $\mathbb{P}^{n-1}$ above Q to be

$$N_\delta(Q) = \{(\xi_1, \cdots, \xi_{n-1}, \xi_1^2 + \cdots + \xi_{n-1}^2 + t) : \xi_i \in Q \text{ and } 0 \leq t \leq \delta\}.$$

For each $f : \mathbb{R}^n \to \mathbb{C}$ and $R \subset \mathbb{R}^n$, denote by f_R the Fourier restriction of f to R

$$f_R(x) = \int_R \widehat{f}(\xi) e(x \cdot \xi) \mathrm{d}\xi.$$

In this section, we will make repeated use of the following inequalities, where B_R will refer to the cube centered at the origin in $\mathbb{R}^n$,

$$w_{B_R,E} * \Big(\frac{1}{(R')^n} w_{B_{R'},E} \Big) \lesssim w_{B_R,E}, \quad R' \leq R \tag{5.1}$$

and, when $n = 2$,

$$w_{B_R,E}(x_1, x_2) \leq \Big(\frac{1}{1 + \frac{|x_1|}{R}} \Big)^{E_1} \Big(\frac{1}{1 + \frac{|x_2|}{R}} \Big)^{E_2}, \quad E_1 + E_2 \leq E. \tag{5.2}$$

We will need the following alternate form of decoupling, when we will derive inequality (8.8).

Theorem 5.1 *For each $E \geq 100n$, the following statement is true for each $F \geq \Gamma_n(E)$, where $\Gamma_n(E)$ is a large enough constant depending on E and n. For $p \geq 2$, each $f : \mathbb{R}^n \to \mathbb{C}$ with Fourier transform supported in $N_{\frac{1}{R}}([0,1]^{n-1})$ and for each cube $B_R \subset \mathbb{R}^n$, we have*

$$\|f\|_{L^p(w_{B_R,E})} \lesssim \mathrm{Dec}_n(R^{-1},p,F)\Big(\sum_{Q \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1]^{n-1})} \|f_{N_{\frac{1}{R}}(Q)}\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}.$$

Proof To simplify notation, we will show the computations when $n = 2$. In this case, $\Gamma_2(E) = 2E + 2$ will suffice.

Using Remark 4.1, it will suffice to prove

$$\|f\|_{L^p(B_R)} \lesssim \mathrm{Dec}_2(R^{-1},p,F)\Big(\sum_{Q \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1])} \|f_{N_{\frac{1}{R}}(Q)}\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}.$$

Due to translation/modulation invariance, we may assume B_R to be centered at the origin.

A change of variables allows us to write

$$f(x_1, x_2) = \int_{N_{\frac{1}{R}}([0,1])} \widehat{f}(\xi)e(\xi \cdot x)\mathrm{d}\xi$$

$$= \sum_{Q \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1])} \int_{Q \times [0,\frac{1}{R}]} \widehat{f}(s, s^2 + t)e(sx_1 + s^2 x_2)e(tx_2)\mathrm{d}s\mathrm{d}t.$$

Next, combining this with the Taylor expansion

$$e(tx_2) = \sum_{j \geq 0} \frac{(2\pi)^j}{j!}\Big(\frac{2ix_2}{R}\Big)^j\Big(\frac{Rt}{2}\Big)^j,$$

we can write for $x \in B_R$,

$$|f(x)| \leq \sum_{j \geq 0} \frac{(4\pi)^j}{j!}\Big|\sum_{Q \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1])} E_Q g_j(x)\Big|, \tag{5.3}$$

where

$$g_j(s) = \int_0^{R^{-1}} \widehat{f}(s, s^2 + t)\Big(\frac{Rt}{2}\Big)^j \mathrm{d}t.$$

Obviously, (5.3) leads to the following inequality:

$$\|f\|_{L^p(B_R)} \leq \mathrm{Dec}_2(R^{-1},p,F)\sum_{j \geq 0} \frac{(4\pi)^j}{j!}\Big(\sum_{Q \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1])} \|E_Q g_j\|^2_{L^p(w_{B_R,F})}\Big)^{\frac{1}{2}}.$$

It remains to prove that (note that we have F on the left-hand side and E on the right-hand side)

$$\|E_Q g_j\|_{L^p(w_{B_R,F})} \lesssim \|f_{N_{1/R}(Q)}\|_{L^p(w_{B_R,E})},$$

uniformly over $j \geq 0$.

An easy computation allows us to assume $Q = [0, R^{-\frac{1}{2}}]$. Indeed, translating $[u, u + R^{-\frac{1}{2}}]$ to $[0, R^{-\frac{1}{2}}]$ on the frequency side will replace (x_1, x_2) with $(x_1 + 2ux_2, x_2)$ on the spatial side. Note that when $0 \le u \le 1$, these shear transformations affect the weights w_B only negligibly.

We start by writing

$$\|E_Q g_j\|_{L^p(w_{B_R}, F)}^p \sim \int \|E_Q g_j\|_{L^p_\sharp(B(y,R))}^p w_{B_R, F}(y) \mathrm{d}y.$$

Recall that

$$E_Q g_j(x) = \int_{N_{\frac{1}{R}}(Q)} \widehat{f}(\xi) \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)^j e((\xi_1^2 - \xi_2)x_2) e(\xi \cdot x) \mathrm{d}\xi.$$

For $x \in B(y, R)$, we write

$$e((\xi_1^2 - \xi_2)x_2) = e((\xi_1^2 - \xi_2)y_2) e((\xi_1^2 - \xi_2)(x_2 - y_2)),$$

and apply another Taylor expansion for $e((\xi_1^2 - \xi_2)(x_2 - y_2))$ to arrive at

$$|E_Q g_j(x)| \le \sum_{k \ge 0} \frac{(4\pi)^k}{k!} \left| \int_{N_{\frac{1}{R}}(Q)} \widehat{f}(\xi) \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)^{j+k} e((\xi_1^2 - \xi_2)y_2) e(\xi \cdot x) \mathrm{d}\xi \right|.$$

It now remains to prove

$$\int \left\| \int_{N_{\frac{1}{R}}(Q)} \widehat{f}(\xi) \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)^j e((\xi_1^2 - \xi_2)y_2) e(\xi \cdot x) \mathrm{d}\xi \right\|_{L^p_\sharp(B(y,R))}^p w_{B_R, F}(y) \mathrm{d}y$$

$$\lesssim \|f_{N_{\frac{1}{R}}(Q)}\|_{L^p(w_{B_R}, E)}^p,$$

uniformly over $j \ge 0$.

We write

$$\int_{N_{\frac{1}{R}}(Q)} \widehat{f}(\xi) \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)^j e((\xi_1^2 - \xi_2)y_2) e(\xi \cdot x) \mathrm{d}\xi = \int \widehat{F}(\xi) m_j(\xi) e(\xi_1 x_1 + \xi_2(x_2 - y_2)) \mathrm{d}\xi,$$

where

$$m_j(\xi) = m_{j, y_2}(\xi) = e(\xi_1^2 y_2) \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)^j 1_{[0, \frac{1}{2}]} \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right) \eta(R^{\frac{1}{2}} \xi_1) \eta(R\xi_2),$$

η is a Schwartz function equal to 1 on $[-2, 2]$ and supported in $[-3, 3]$, and

$$F = f_{N_{\frac{1}{R}}(Q)}.$$

Let $M_j(t)$ be a compactly supported Schwartz function which agrees with t^j on $[0, \frac{1}{2}]$ and satisfies the derivative bound

$$\left\| \frac{\mathrm{d}^k}{\mathrm{d}t^k} M_j \right\|_{L^\infty(\mathbb{R})} \lesssim_k 1, \tag{5.4}$$

uniformly over $j \ge 0$ for each $k \ge 0$.

Note that we can also write

$$\int_{N_{\frac{1}{R}}(Q)} \widehat{f}(\xi) \left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)^j e((\xi_1^2 - \xi_2)y_2) e(\xi \cdot x) \mathrm{d}\xi = \int \widehat{F}(\xi) \widetilde{m}_j(\xi) e(\xi_1 x_1 + \xi_2(x_2 - y_2)) \mathrm{d}\xi,$$

where

$$\widetilde{m}_{j,y_2}(\xi) = \widetilde{m}_j(\xi) = e(\xi_1^2 y_2) M_j\left(\frac{R(\xi_2 - \xi_1^2)}{2}\right)\eta(R^{\frac{1}{2}}\xi_1)\eta(R\xi_2).$$

Applying Hölder's inequality, we arrive at

$$\int \left\|\int \widehat{F}(\xi)\widetilde{m}_j(\xi)e(\xi_1 x_1 + \xi_2(x_2 - y_2))\mathrm{d}\xi\right\|_{L^p_\sharp(B(y,R))}^p w_{B_R,F}(y)\mathrm{d}y$$

$$\lesssim \iint |F|^p * |\widehat{\widetilde{m}_j}|(x)R^{-2}1_{B_R}(x_1 - y_1, x_2)w_{B_R,F}(y)\mathrm{d}x\mathrm{d}y$$

$$= \int |F|^p(x')\left[\iint |\widehat{\widetilde{m}_j}|(x - x')R^{-2}1_{B_R}(x_1 - y_1, x_2)w_{B_R,F}(y)\mathrm{d}x\mathrm{d}y\right]\mathrm{d}x'.$$

It remains to show that

$$\int |\widehat{\widetilde{m}_j}| * (R^{-2}1_{B_R})(y_1 - x_1, -x_2)w_{B_R,F}(y)\mathrm{d}y \lesssim w_{B_R,E}(x).$$

In fact, we will prove a slightly stronger inequality

$$\int |\widehat{\widetilde{m}_j}| * (R^{-2}1_{B_R})(y_1 - x_1, -x_2)w_{B_R,F}(y)\mathrm{d}y \lesssim \left(1 + \frac{|x_1|}{R}\right)^{-E}\left(1 + \frac{|x_2|}{R}\right)^{-E}. \qquad (5.5)$$

An easy computation using (5.4) shows that for each $s_1, s_2 \geq 0$,

$$\|\partial_{\xi_1}^{s_1}\partial_{\xi_2}^{s_2}\widetilde{m}_j\|_{L^\infty} \lesssim_{s_1,s_2} \left(R^{\frac{1}{2}} + \frac{|y_2|}{R^{\frac{1}{2}}}\right)^{s_1} R^{s_2}.$$

Combining this with the fact that $\widetilde{m}_j$ is compactly supported in $[-R^{-\frac{1}{2}}, R^{-\frac{1}{2}}] \times [-R^{-1}, R^{-1}]$ leads, via repeated integration by parts, to the following estimate for the Fourier transform:

$$|\widehat{\widetilde{m}_j}(x_1, x_2)| \leq \phi_1(x_1)\phi_2(x_2),$$

where

$$\phi_1(x_1) \lesssim_{s_1} \frac{1}{R^{\frac{1}{2}}}\left(\frac{1}{1 + \frac{|x_1|}{R^{\frac{1}{2}}+R^{-\frac{1}{2}}|y_2|}}\right)^{s_1} \qquad (5.6)$$

and

$$\phi_2(x_2) \lesssim_{s_2} \frac{1}{R}\left(\frac{1}{1 + \frac{|x_2|}{R}}\right)^{s_2}. \qquad (5.7)$$

Let $I_R = \left[-\frac{R}{2}, \frac{R}{2}\right]$ and recall that $B_R = I_R \times I_R$. Using (5.7) and (5.1) ($n = 1$), we may now write

$$\int (|\widehat{\widetilde{m}_j}| * (R^{-2}1_{B_R}))(y_1 - x_1, -x_2)w_{B_R,F}(y)\mathrm{d}y$$

$$\leq \left(\phi_2 * \left(\frac{1}{R}1_{I_R}\right)\right)(-x_2)\int\left(\phi_1 * \left(\frac{1}{R}I_R\right)\right)(y_1 - x_1)w_{B_R,F}(y)\mathrm{d}y$$

$$\lesssim \frac{1}{R}\left(1 + \frac{|x_2|}{R}\right)^{-E}\int\left(\phi_1 * \left(\frac{1}{R}I_R\right)\right)(y_1 - x_1)w_{B_R,F}(y)\mathrm{d}y.$$

Recalling (5.5), we are left with proving that

$$\int \left(\phi_1 * \left(\frac{1}{R} I_R \right) \right)(y_1 - x_1) w_{B_R, F}(y) \mathrm{d}y \lesssim R \left(1 + \frac{|x_1|}{R} \right)^{-E}. \tag{5.8}$$

We split the analysis into three cases. We will need $F \geq 2E + 2$.

(a) $|y_2| \leq R$. In this case,

$$\phi_1(x_1) \lesssim \frac{1}{R^{\frac{1}{2}}} \left(1 + \frac{|x_1|}{R^{\frac{1}{2}}} \right)^{-E}.$$

Using (5.1) with $n = 1$ twice (first $R' = R^{\frac{1}{2}}$ then $R' = R$) and (5.2) with $E_1 = E$, $E_2 = 2$, we get

$$\int_{|y_2| \lesssim R} \left(\phi_1 * \left(\frac{1}{R} I_R \right) \right)(y_1 - x_1) w_{B_R, F}(y) \mathrm{d}y$$

$$\lesssim \int \frac{1}{R} \left(1 + \frac{|x_1 - y_1|}{R} \right)^{-E} \left(1 + \frac{|y_1|}{R} \right)^{-E} \mathrm{d}y_1 \int \left(1 + \frac{|y_2|}{R} \right)^{-2} \mathrm{d}y_2$$

$$\lesssim R \left(1 + \frac{|x_1|}{R} \right)^{-E},$$

as needed.

(b) $|y_2| \sim KR$, with $K \in [1, R^{\frac{1}{2}}] \cap 2^{\mathbb{N}}$. In this case,

$$\phi_1(x_1) \lesssim \frac{1}{R^{\frac{1}{2}}} \left(1 + \frac{|x_1|}{KR^{\frac{1}{2}}} \right)^{-E}.$$

Using (5.1) twice (first $R' = KR^{\frac{1}{2}}$ then $R' = R$), and (5.2) with $E_1 = E$, $E_2 = 3$, we write

$$\int_{|y_2| \sim KR} \left(\phi_1 * \left(\frac{1}{R} I_R \right) \right)(y_1 - x_1) w_{B_R, F}(y) \mathrm{d}y$$

$$\lesssim K \int \frac{1}{R} \left(1 + \frac{|x_1 - y_1|}{R} \right)^{-E} \left(1 + \frac{|y_1|}{R} \right)^{-E} \mathrm{d}y_1 \int_{|y_2| \sim KR} \left(1 + \frac{|y_2|}{R} \right)^{-3} \mathrm{d}y_2$$

$$\lesssim K \left(1 + \frac{|x_1|}{R} \right)^{-E} \frac{KR}{K^3} = \frac{R}{K} \left(1 + \frac{|x_1|}{R} \right)^{-E}.$$

Note that summing over $K \in [1, R^{\frac{1}{2}}] \cap 2^{\mathbb{N}}$ leads to the desired estimate (5.8).

(c) $|y_2| \sim KR^{\frac{3}{2}}$ with $K \in [1, \infty) \cap 2^{\mathbb{N}}$. In this case,

$$\phi_1(x_1) \lesssim \frac{1}{R^{\frac{1}{2}}} \left(1 + \frac{|x_1|}{KR} \right)^{-E},$$

and so, by (5.1), we have

$$\left(\phi_1 * \left(\frac{1}{R} I_R \right) \right)(y_1 - x_1) \lesssim \frac{1}{R^{\frac{1}{2}}} \left(1 + \frac{|y_1 - x_1|}{KR} \right)^{-E}.$$

Next, combining this with (5.2) ($E_1 = E$, $E_2 = E + 2$) and then with (5.1), we get

$$\int_{|y_2| \sim KR^{\frac{3}{2}}} \left(\phi_1 * \left(\frac{1}{R} I_R \right) \right) (y_1 - x_1) w_{B_R, F}(y) \mathrm{d}y$$

$$\lesssim R^{\frac{1}{2}} \int \left(1 + \frac{|y_1 - x_1|}{KR} \right)^{-E} \frac{1}{R} \left(1 + \frac{|y_1|}{R} \right)^{-E} \mathrm{d}y_1 \int_{|y_2| \sim KR^{\frac{3}{2}}} \left(1 + \frac{|y_2|}{R} \right)^{-E-2} \mathrm{d}y_2$$

$$\lesssim R^{\frac{1}{2}} \left(1 + \frac{|x_1|}{KR} \right)^{-E} \frac{KR^{\frac{3}{2}}}{(KR^{\frac{1}{2}})^{E+2}} \lesssim R^{\frac{1}{2}} \left(1 + \frac{|x_1|}{KR} \right)^{-E} \frac{KR^{\frac{3}{2}}}{K^{E+2}R}$$

$$\leq \frac{R}{K} \left(1 + \frac{|x_1|}{R} \right)^{-E}.$$

Note that summing over $K \in [1, \infty) \cap 2^{\mathbb{N}}$ leads to the desired estimate (5.8).

6 L^2 Decoupling

We will use Lemma 4.1 to prove a very simple but efficient decoupling. This exploits L^2 orthogonality and will allow us to decouple to the smallest possible scale, equal to the inverse of the radius of the cube. This process is illustrated by the following simple result.

Proposition 6.1 (L^2 Decoupling) *Let Q be a cube with $l(Q) \geq R^{-1}$. Then for each cube $B_R \subset \mathbb{R}^n$ with side length R we have*

$$\|E_Q g\|_{L^2(w_{B_R})} \lesssim \left(\sum_{q \in \mathrm{Part}_{\frac{1}{R}}(Q)} \|E_q g\|_{L^2(w_{B_R})}^2 \right)^{\frac{1}{2}}.$$

Proof We will prove that

$$\|E_Q g\|_{L^2(w_{B_R})}^2 \lesssim \sum_{q \in \mathrm{Part}_{\frac{1}{R}}(Q)} \|E_q g\|_{L^2(w_{B_R})}^2. \tag{6.1}$$

Fix a positive Schwartz function η, such that the Fourier transform of $\sqrt{\eta}$ is supported in a small neighborhood of the origin, and such that $\eta \geq 1$ on $B(0, 1)$. By invoking Lemma 4.1, we see that inequality (6.1) will follow once we check that

$$\|E_Q g\|_{L^2(B')}^2 \lesssim \sum_{q \in \mathrm{Part}_{\frac{1}{R}}(Q)} \|E_q g\|_{L^2(\eta_{B'})}^2 \tag{6.2}$$

holds true for each cube B' with $l(B') = R$.

Note that the Fourier transform of $\sqrt{\eta_{B'}} E_q g$ will be supported inside the R^{-1}-neighborhood of the paraboloid above q, and that these neighborhoods are pairwise disjoint for two non-adjacent q. Since

$$\|E_Q g\|_{L^2(B')}^2 \lesssim \|E_Q g\|_{L^2(\eta_{B'})}^2 = \|\sqrt{\eta_{B'}} E_Q g\|_{L^2(\mathbb{R}^n)}^2,$$

(6.2) will now immediately follow from the L^2 orthogonality of the functions $\sqrt{\eta_{B'}} E_q g$.

7 Parabolic Rescaling

A nice property of the paraboloid $\mathbb{P}^{n-1}$ is the fact that each square-like cap on it can be stretched to the whole $\mathbb{P}^{n-1}$ via an affine transformation. Affine transformations interact well with the Fourier transform, and this facilitates a natural passage from the operator E_Q to $E_{[0,1]^{n-1}}$.

Proposition 7.1 *Let $0 < \delta \le \sigma < 1$ and $p \ge 2$. For each cube $Q \subset [0,1]^{n-1}$ with $l(Q) = \sigma^{\frac{1}{2}}$ and each cube $B \subset \mathbb{R}^n$ with $l(B) \ge \delta^{-1}$, we have*

$$\|E_Q g\|_{L^p(w_B)} \lesssim \mathrm{Dec}_p\Big(\frac{\delta}{\sigma}\Big)\Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(w_B)}^2\Big)^{\frac{1}{2}}.$$

The implicit constant is independent of δ, σ, Q, B.

Proof Let us first assume $l(B) = \delta^{-1}$. We will apply Lemma 4.1 to

$$O_1(v) = \|E_Q g\|_{L^p(v)}^p,$$

$$O_2(v) = \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(v)}^2\Big)^{\frac{p}{2}}$$

(see Remark 4.1). It thus suffices to prove that

$$\|E_Q g\|_{L^p(B)} \lesssim \mathrm{Dec}_p\Big(\frac{\delta}{\sigma}\Big)\Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(w_B)}^2\Big)^{\frac{1}{2}}.$$

Assume $Q = a + [0, \sigma^{\frac{1}{2}}]^{n-1}$ with $a = (a_1, \cdots, a_{n-1})$. We will perform a parabolic rescaling via the affine transformation $L = L_Q$,

$$L_Q(\xi_1, \cdots, \xi_{n-1}) = (\xi_1', \cdots, \xi_{n-1}') = \Big(\frac{\xi_1 - a_1}{\sigma^{\frac{1}{2}}}, \cdots, \frac{\xi_{n-1} - a_{n-1}}{\sigma^{\frac{1}{2}}}\Big).$$

A simple computation shows that for each cube $\widetilde{Q}$, we have

$$|E_{\widetilde{Q}} g(x_1, \cdots, x_{n-1}, x_n)| = \sigma^{\frac{n-1}{2}} |E_{\widetilde{Q}_L} g_L((x_1 + 2a_1 x_n)\sigma^{\frac{1}{2}}, \cdots, (x_{n-1} + 2a_{n-1} x_n)\sigma^{\frac{1}{2}}, x_n \sigma)|,$$

where $\widetilde{Q}_L = L(\widetilde{Q})$, $g_L = g \circ L$. The image S of B under the affine transformation $T = T_Q$

$$T_Q(x_1, \cdots, x_{n-1}, x_n) = ((x_1 + 2a_1 x_n)\sigma^{\frac{1}{2}}, \cdots, (x_{n-1} + 2a_{n-1} x_n)\sigma^{\frac{1}{2}}, x_n \sigma)$$

can be covered with a family $\mathcal{F}$ of pairwise disjoint cubes Δ with side length $\delta^{-1}\sigma$, such that we have the following double inequality, in the same spirit as (4.1):

$$1_S(x) \lesssim \sum_{B' \in \mathcal{F}} w_{B'}(x) \lesssim w_B(T^{-1}x). \tag{7.1}$$

The second inequality is very easy to guarantee for a proper covering, as $l(B') \leq l(B)$. After a change of variables on the spatial side, we get (since $Q_L = [0,1]^{n-1}$)

$$\|E_Q g\|_{L^p(B)} = \sigma^{\frac{n-1}{2}} \sigma^{-\frac{n+1}{2p}} \|E g_L\|_{L^p(S)}$$

$$\lesssim \sigma^{\frac{n-1}{2} - \frac{n+1}{2p}} \Big(\sum_{B' \in \mathcal{F}} \|E g_L\|_{L^p(w_{B'})}^p \Big)^{\frac{1}{p}}$$

$$\leq \sigma^{\frac{n-1}{2} - \frac{n+1}{2p}} \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big[\sum_{B' \in \mathcal{F}} \Big(\sum_{q' \in \mathrm{Part}_{(\frac{\delta}{\sigma})^{\frac{1}{2}}}([0,1]^{n-1})} \|E_{q'} g_L\|_{L^p(w_{B'})}^2 \Big)^{\frac{p}{2}} \Big]^{\frac{1}{p}}$$

$$= \sigma^{\frac{n-1}{2} - \frac{n+1}{2p}} \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big[\sum_{B' \in \mathcal{F}} \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_{q_L} g_L\|_{L^p(w_{B'})}^2 \Big)^{\frac{p}{2}} \Big]^{\frac{1}{p}}.$$

Using Minkowski's inequality followed by (7.1), this is dominated by

$$\sigma^{\frac{n-1}{2} - \frac{n+1}{2p}} \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_{q_L} g_L\|_{L^p(\sum w_{B'})}^2 \Big)^{\frac{1}{2}}$$

$$\lesssim \sigma^{\frac{n-1}{2} - \frac{n+1}{2p}} \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_{q_L} g_L\|_{L^p(w_B \circ T^{-1})}^2 \Big)^{\frac{1}{2}}.$$

By changing back to the original variables, this is easily seen to be the same as

$$\mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(w_B)}^2 \Big)^{\frac{1}{2}}.$$

This finishes the proof in the case $l(B) = \delta^{-1}$.

Let us next assume $l(B) \geq \delta^{-1}$. By invoking again Lemma 4.1 (see Remark 4.1), it suffices to prove

$$\|E_Q g\|_{L^p(B)} \lesssim \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(w_B)}^2 \Big)^{\frac{1}{2}}.$$

Using (4.1) and Minkowski's inequality, we may complete the argument as follows:

$$\|E_Q g\|_{L^p(B)} \lesssim \Big(\sum_{\Delta \in \mathrm{Part}_{\delta^{-1}}(B)} \|E_Q g\|_{L^p(w_\Delta)}^p \Big)^{\frac{1}{p}}$$

$$\lesssim \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big(\sum_{\Delta \in \mathrm{Part}_{\delta^{-1}}(B)} \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(w_\Delta)}^2 \Big)^{\frac{p}{2}} \Big)^{\frac{1}{p}}.$$

$$\lesssim \mathrm{Dec}_p \Big(\frac{\delta}{\sigma} \Big) \Big(\sum_{q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q)} \|E_q g\|_{L^p(w_B)}^2 \Big)^{\frac{1}{2}}.$$

8 Linear Versus Multilinear Decoupling

Let $\pi : \mathbb{P}^{n-1} \to [0,1]^{n-1}$ be the projection map.

Definition 8.1 *We say that the cubes* $Q_1, \cdots, Q_n \subset [0,1]^{n-1}$ *are* ν-*transverse if the volume of the parallelepiped spanned by unit normals* $n(P_i)$ *is greater than* ν, *for each choice of* $P_i \in \mathbb{P}^{n-1}$ *with* $\pi(P_i) \in Q_i$.

For $E \geq 100n$, $2 \leq p \leq \infty$, $m \in \mathbb{N}$ and $0 < \nu < 1$, we let $\mathrm{Dec}(\delta, p, \nu, m, E) = \mathrm{Dec}_n(\delta, p, \nu, m, E)$ be the smallest constant, such that the inequality

$$\left[\sum_{\Delta \in \mathrm{Part}_{\mu^{-1}}(B)} \left(\prod_{i=1}^{n} \|E_{Q_i} g\|_{L^p(w_{\Delta, 10E})}^p \right)^{\frac{1}{n}} \right]^{\frac{1}{p}}$$

$$\leq \mathrm{Dec}(\delta, p, \nu, m, E) \left[\prod_{i=1}^{n} \sum_{q_i \in \mathrm{Part}_{\delta^{\frac{1}{2}}}(Q_i)} \|E_{q_i} g\|_{L^p(w_{B,E})}^2 \right]^{\frac{1}{2n}} \qquad (8.1)$$

holds for each cube $B \subset \mathbb{R}^n$ with $l(B) = \delta^{-1}$, each $g : [0,1]^{n-1} \to \mathbb{C}$ and each ν-transverse cubes Q_i with equal side lengths μ satisfying $\mu \geq \delta^{2^{-m}}$. Recall that $\mathrm{Part}_{\mu^{-1}}(B)$ is the partition of B by using cubes Δ with $l(\Delta) = \mu^{-1}$. The lower bound which we impose on the size of μ is a bit more severe than the minimal lower bound $\mu \geq \delta^{\frac{1}{2}}$ needed in order to make sense of the quantity $\mathrm{Part}_{\delta^{\frac{1}{2}}}(Q_i)$. This restriction can be ignored now and should only be paid attention to in the final argument from the last section, when we dominate (11.7) by (11.8).

Note that we use $w_{\Delta, 10E}$ rather than $w_{\Delta, E}$ in (8.1). This is done for purely technical reasons, as explained in Remark 4.2.

Since $|E_{Q_i} g|$ can be thought of as being essentially constant on each Δ, the quantity

$$\left[\sum_{\Delta \in \mathrm{Part}_{\mu^{-1}}(B)} \left(\prod_{i=1}^{n} \|E_{Q_i} g\|_{L^p(w_{\Delta, 10E})}^p \right)^{\frac{1}{n}} \right]^{\frac{1}{p}}$$

can be viewed as being comparable to

$$\left\| \prod_{i=1}^{n} |E_{Q_i} g|^{\frac{1}{n}} \right\|_{L^p(w_{B, 10E})}.$$

The former will be a preferred substitute for the latter due to purely technical reasons.

Several applications of Hölder's inequality combined with (4.1) show that for each ν, m,

$$\mathrm{Dec}(\delta, p, \nu, m, E) \lesssim \mathrm{Dec}(\delta, p, E). \qquad (8.2)$$

This inequality is too basic and will never be used. We will instead derive a stronger form of it in the last section (see (11.10)), which dominates $\mathrm{Dec}(\delta, p, \nu, m, E)$ by using a combination of powers of δ and some power of $\mathrm{Dec}(\delta, p)$.

We will now prove and later use the following approximate reverse inequality. Recall the definition of $\Gamma_n(E)$ from Theorem 5.1.

Theorem 8.1 *Let* $E \geq 100n$. *Assume that one of the following holds:*
(i) $n = 2$.
(ii) $n \geq 3$ *and* $\mathrm{Dec}_{n-1}(\delta, p, \Gamma_{n-1}(10E)) \lesssim_{\epsilon, E} \delta^{-\epsilon}$.

Then for each $0 < \nu \leq 1$, there is $\epsilon(\nu) = \epsilon(\nu, p, E)$ with $\lim_{\nu \to 0} \epsilon(\nu) = 0$ and $C_{\nu,m}$, such that for each $m \geq 1$, we have

$$\mathrm{Dec}_n(R^{-1}, p, E) \leq C_{\nu,m} R^{\epsilon(\nu)} \Big(1 + \sup_{1 \leq R' \leq R} \mathrm{Dec}_n(R'^{-1}, p, \nu, m, E) \Big) \tag{8.3}$$

for each $R \gtrsim_{\nu,m} 1$.

We next prove the case $n = 3$ of the theorem and will then indicate the modifications needed for $n \geq 4$. The argument will also show how to deal with the case $n = 2$.

Remark 8.1 If $P_1, P_2, P_3 \in \mathbb{P}^2$, the volume of the parallelepiped spanned by the unit normals $n(P_i)$ is comparable to the area of the triangle with vertices $\pi(P_i)$.

The key step in the proof of Theorem 8.1 for $n = 3$ is the following result.

Proposition 8.1 *Assume* $\mathrm{Dec}_2(\delta, p, \Gamma_2(10E)) \lesssim_\epsilon \delta^{-\epsilon}$. *Then there are constants* C, C_ϵ, *such that for each* $m \geq 1$ *and each* $R \geq K^{2^m}$,

$$\|Eg\|_{L^p(w_{B_R,E})} \leq C_\epsilon K^\epsilon \Bigg[\Big(\sum_{\alpha \in \mathrm{Part}_{K^{-1}}([0,1]^2)} \|E_\alpha g\|^2_{L^p(w_{B_R,E})} \Big)^{\frac{1}{2}}$$

$$+ \Big(\sum_{\beta \in \mathrm{Part}_{K^{-\frac{1}{2}}}([0,1]^2)} \|E_\beta g\|^2_{L^p(w_{B_R,E})} \Big)^{\frac{1}{2}} \Bigg]$$

$$+ K^C \mathrm{Dec}_3(R^{-1}, p, K^{-2}, m, E) \Big(\sum_{\Delta \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1]^2)} \|E_\Delta g\|^2_{L^p(w_{B_R,E})} \Big)^{\frac{1}{2}}.$$

Proof Using Lemma 4.1 (see Remark 4.1), it suffices to prove the inequality with the unweighted quantity $\|Eg\|_{L^p(B_R)}$ on the left-hand side. Cover B_R with a family $\mathrm{Part}_K(B_R)$ of cubes $B_K \subset \mathbb{R}^3$ with side length K.

Let $\psi : \mathbb{R}^3 \to \mathbb{C}$ be a Schwartz function with Fourier transform equal to 1 on $B(0, 10)$. For each $\alpha \in \mathrm{Part}_{K^{-1}}([0, 1]^2)$, define

$$c_\alpha(B_K) = \Big(\frac{1}{|B_K|} \int |E_\alpha g|^p w_{B_K,10E} \Big)^{\frac{1}{p}}.$$

Note that since $E_\alpha g = (E_\alpha g) * \psi_K$ for an appropriate modulation/dilation ψ_K of ψ, we have

$$\sup_{x \in B_K} |E_\alpha g(x)| \lesssim c_\alpha(B_K).$$

This is a manifestation of the uncertainty principle that asserts that $|E_\alpha g|$ is essentially constant at scale K. Let $\alpha^* = \alpha^*(K) \in \mathrm{Part}_{K^{-1}}([0, 1]^2)$ be a square that maximizes $c_\alpha(B_K)$. Define also

$$S_{\mathrm{big}} = \{ \alpha : c_\alpha(B_K) \geq K^{-C} c_{\alpha^*}(B_K) \}.$$

The number C will change its value from one line to the next one, but crucially, it will always be independent of K.

We will show that for each $B_K \in \mathrm{Part}_K(B_R)$ there exists a line $L = L(B_K)$ in the (ξ_1, ξ_2) plane, such that if

$$S_L = \Big\{ (\xi_1, \xi_2) : \mathrm{dist}((\xi_1, \xi_2), L) \leq \frac{C}{K} \Big\},$$

then for $x \in B_K$,

$$|Eg(x)|$$

$$\leq Cc_{\alpha^*}(B_K) \tag{8.4}$$

$$+ K^4 \max_{\substack{\alpha_1,\alpha_2,\alpha_3 \\ K^{-2}-\text{transverse}}} \Big(\prod_{i=1}^3 c_{\alpha_i}(B_K) \Big)^{\frac{1}{3}} \tag{8.5}$$

$$+ \Big| \sum_{\alpha \subset S_L} E_\alpha g(x) \Big|. \tag{8.6}$$

To see this, we distinguish three scenarios. First, if there is no $\alpha \in S_{\text{big}}$ with $\text{dist}(\alpha, \alpha^*) \geq \frac{10}{K}$, then (8.4) suffices as

$$|Eg(x)| \leq \sum_\alpha c_\alpha(B_K) \leq Cc_{\alpha^*}(B_K).$$

Otherwise, there is $\alpha^{**} \in S_{\text{big}}$ with $\text{dist}(\alpha^{**}, \alpha^*) \geq \frac{10}{K}$. The line L is determined by the centers of α^1, α^2, which are chosen to be furthest apart among all possible pairs in S_{big}. Note that the distance between these centers is at least $\frac{10}{K}$.

Second, if there is $\alpha^3 \in S_{\text{big}}$, such that α^3 intersects the complement of S_L, then (8.5) suffices. To see this, note first that α^3 is forced to intersect the strip between α^1 and α^2 perpendicular to L. Thus, a triangle determined by any three points in α^i has area $\gg K^{-2}$. Combining this with Remark 8.1 shows that $\alpha^1, \alpha^2, \alpha^3$ are K^{-2} transverse for C large enough.

Third, if all $\alpha \in S_{\text{big}}$ are inside S_L, the sum of (8.4) and (8.6) will obviously suffice.

We now claim that (8.4)–(8.6) imply the following:

$$\|Eg\|_{L^p(B_K)}$$

$$\leq C_\epsilon K^\epsilon \Big[\Big(\sum_{\alpha \in \text{Part}_{K^{-1}}([0,1]^2)} \|E_\alpha g\|^2_{L^p(w_{B_K},10E)} \Big)^{\frac{1}{2}} + \Big(\sum_{\beta \in \text{Part}_{K^{-\frac{1}{2}}}([0,1]^2)} \|E_\beta g\|^2_{L^p(w_{B_K},10E)} \Big)^{\frac{1}{2}} \Big]$$

$$+ K^C \max_{\substack{\alpha_1,\alpha_2,\alpha_3 \\ K^{-2}-\text{transverse}}} \Big(\prod_{i=1}^3 \|E_{\alpha_i} g\|_{L^p(w_{B_K},10E)} \Big)^{\frac{1}{3}}. \tag{8.7}$$

Only the third scenario above needs an explanation. Cover S_L by pairwise disjoint rectangles U of dimensions K^{-1} and $K^{-\frac{1}{2}}$, with the long side parallel to L. To simplify notation, assume the equation of L is $\eta = 1$ and that $B_K = [0, K]^3$. For each fixed y the Fourier transform of $(x, z) \mapsto E_{S_L} g(x, y, z)$ is supported in the $O(K^{-1})$ neighborhood of the parabola $\eta = \xi^2 + 1$. Using Theorem 5.1 and our hypothesis $\text{Dec}_2(K^{-1}, p, \Gamma_2(10E)) \lesssim_\epsilon K^\epsilon$, we can write

$$\|E_{S_L} g(x, y, z)\|_{L^p_{x,z}([0,K]^2)} \lesssim \text{Dec}_2(K^{-1}, p, 10E) \Big(\sum_U \|E_U g(x, y, z)\|^2_{L^p_{x,z}(w_{[0,K]^2},10E)} \Big)^{\frac{1}{2}}$$

$$\lesssim_\epsilon \delta^{-\epsilon} \Big(\sum_U \|E_U g(x, y, z)\|^2_{L^p_{x,z}(w_{[0,K]^2},10E)} \Big)^{\frac{1}{2}}. \tag{8.8}$$

Next raise this inequality to the power p, integrate over $y \in [0, K]$ and use

$$w_{[0,K]^2,10E}(x, z) 1_{[0,K]}(y) \lesssim w_{B_K,10E}(x, y, z) \tag{8.9}$$

and Minkowski's inequality to write

$$\Big\| \sum_{\alpha:\alpha\subset S_L} E_\alpha g \Big\|_{L^p(B_K)} \lesssim_\epsilon K^\epsilon \Big(\sum_U \|E_U g\|^2_{L^p(w_{B_K,10E})} \Big)^{\frac{1}{2}}.$$

Note however that since we are dealing with the third scenario, the contribution of $E_{[0,1]^2\setminus S_L}g$ is small

$$\|E_{[0,1]^2\setminus S_L}g\|_{L^p(w_{B_K,10E})} \leq \sum_{\alpha\notin S_{\mathrm{big}}} \|E_\alpha g\|_{L^p(w_{B_K,10E})} \leq Cc_{\alpha^*}(B_K)|B_K|^{\frac{1}{p}}.$$

Using the triangle inequality, we get

$$\Big(\sum_U \|E_U g\|^2_{L^p(w_{B_K,10E})} \Big)^{\frac{1}{2}} \leq \Big(\sum_{\beta\in\mathrm{Part}_{K^{-\frac{1}{2}}}([0,1]^2)} \|E_\beta g\|^2_{L^p(w_{B_K,10E})} \Big)^{\frac{1}{2}} + C\|E_{\alpha^*}g\|_{L^p(w_{B_K,10E})}.$$

We conclude that (8.7) holds under the third scenario. The first two scenarios are quite immediate.

Using $w_{B_K,10E} \leq w_{B_K,E}$, (8.7) further implies that

$$\|Eg\|_{L^p(B_K)}$$
$$\leq C_\epsilon K^\epsilon \Big[\Big(\sum_{\alpha\in\mathrm{Part}_{K^{-1}}([0,1]^2)} \|E_\alpha g\|^2_{L^p(w_{B_K,E})} \Big)^{\frac{1}{2}} + \Big(\sum_{\beta\in\mathrm{Part}_{K^{-\frac{1}{2}}}([0,1]^2)} \|E_\beta g\|^2_{L^p(w_{B_K,E})} \Big)^{\frac{1}{2}} \Big]$$
$$+ K^C \max_{\substack{\alpha_1,\alpha_2,\alpha_3 \\ K^{-2}-\mathrm{transverse}}} \Big(\prod_{i=1}^3 \|E_{\alpha_i}g\|_{L^p(w_{B_K,10E})} \Big)^{\frac{1}{3}}. \tag{8.10}$$

Finally, we raise (8.10) to the power p and sum over all $B_K \in \mathrm{Part}_K(B_R)$, by invoking Minkowski's inequality and (4.1), to get

$$\|Eg\|_{L^p(B_R)}$$
$$\leq C_\epsilon K^\epsilon \Big[\Big(\sum_{\alpha\in\mathrm{Part}_{K^{-1}}([0,1]^2)} \|E_\alpha g\|^2_{L^p(w_{B_R,E})} \Big)^{\frac{1}{2}} + \Big(\sum_{\beta\in\mathrm{Part}_{K^{-\frac{1}{2}}}([0,1]^2)} \|E_\beta g\|^2_{L^p(w_{B_R,E})} \Big)^{\frac{1}{2}} \Big]$$
$$+ K^C \mathrm{Dec}_3(R^{-1},p,K^{-2},m,E) \Big(\sum_{\Delta\in\mathrm{Part}_{R^{-\frac{1}{2}}}([0,1]^2)} \|E_\Delta g\|^2_{L^p(w_{B_R,E})} \Big)^{\frac{1}{2}}.$$

An application of Lemma 4.1 finishes the proof.

Parabolic rescaling as in the proof of Proposition 7.1 leads to the following. The details are left to the reader.

Proposition 8.2 *Let $\tau \subset [0,1]^2$ be a square with side length $\delta \geq R^{-\frac{1}{2}}K^{2^{m-1}}$. Assume*

$$\mathrm{Dec}_2(\delta',p,\Gamma_2(10E)) \lesssim_\epsilon \delta'^{-\epsilon}$$

for all $\delta' < 1$. Then if $R \geq K^{2^m}$, we have

$$\|E_\tau g\|_{L^p(w_{B_R,E})} \leq C_\epsilon K^\epsilon \Big[\Big(\sum_{\alpha \in \mathrm{Part}_{\delta K^{-1}}(\tau)} \|E_\alpha g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}$$

$$+ \Big(\sum_{\beta \in \mathrm{Part}_{\delta K^{-\frac{1}{2}}}(\tau)} \|E_\beta g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}\Big]$$

$$+ K^C \mathrm{Dec}_3((R\delta^2)^{-1}, p, K^{-2}, m, E)\Big(\sum_{\Delta \in \mathrm{Part}_{R^{-\frac{1}{2}}}(\tau)} \|E_\Delta g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}.$$

The constants C_ϵ, C are independent of δ, R, τ, K.

We are now ready to prove Theorem 8.1 for $n = 3$. Let $K = \nu^{-\frac{1}{2}}$. Let also $R \geq K^{2^m} = \nu^{-2^{m-1}}$. Iterate Proposition 8.2 starting with scale $\delta = 1$, until we reach scale $\delta = R^{-\frac{1}{2}}K^{2^{m-1}}$. Each iteration lowers the scale of the square from δ to at least $\frac{\delta}{K^{\frac{1}{2}}}$. Thus we have to iterate $O(\log_K R)$ times. We use the following immediate consequence of Hölder's inequality:

$$\Big(\sum_{\beta \in \mathrm{Part}_{R^{-\frac{1}{2}}K^{2^{m-1}}}([0,1]^2)} \|E_\beta g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}$$

$$\lesssim K^{O(1)}\Big(\sum_{\Delta \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1]^2)} \|E_\Delta g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}. \tag{8.11}$$

Since

$$\mathrm{Dec}_3((\delta^2 R)^{-1}, p, \nu, m, E) \leq \sup_{1 \leq R' \leq R} \mathrm{Dec}_3(R'^{-1}, p, \nu, m, E),$$

we get

$$\|E_{[0,1]^2} g\|_{L^p(w_{B_R,E})}$$

$$\leq (CC_\epsilon K^\epsilon)^{O(\log_K R)} K^{O(1)}\Big(1 + \sup_{1 \leq R' \leq R} \mathrm{Dec}_3(R'^{-1}, p, \nu, m, E)\Big)$$

$$\cdot \Big(\sum_{\Delta \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1]^2)} \|E_\Delta g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}$$

$$= R^{-O(1)\log_\nu(CC_\epsilon)+\epsilon} \nu^{-O(1)}\Big(1 + \sup_{1 \leq R' \leq R} \mathrm{Dec}_3(R'^{-1}, p, \nu, m, E)\Big)$$

$$\cdot \Big(\sum_{\Delta \in \mathrm{Part}_{R^{-\frac{1}{2}}}([0,1]^2)} \|E_\Delta g\|^2_{L^p(w_{B_R,E})}\Big)^{\frac{1}{2}}.$$

The result in Theorem 8.1 now follows since C, C_ϵ do not depend on ν.

To summarize, the proof of Theorem 8.1 for $n = 3$ relied on the hypothesis that the contribution coming from squares β living near a line is controlled by the negligible lower dimensional quantity $\mathrm{Dec}_2(\delta, p) = O(\delta^{-\epsilon})$. When $n \geq 4$, the contribution from the cubes near a hyperplane H in $[0,1]^{n-1}$ will be similarly controlled by $\mathrm{Dec}_{n-1}(\delta, p)$. That is because $\pi^{-1}(H)$ is a lower dimensional elliptic paraboloid whose principal curvatures are ~ 1, uniformly over H. This paraboloid is an affine image of $\mathbb{P}^{n-2}$, and can be analyzed by using parabolic rescaling. When $n = 2$, there is no such lower dimensional contribution.

9 From Multilinear Kakeya to Multilinear Decouplings

We start by recalling the following multilinear Kakeya inequality due to Bennett, Carbery and Tao [1]. We refer the reader to [6] for a short proof.

Theorem 9.1 *Let $0 < \nu < 1$. Consider n families $\mathcal{P}_j$ consisting of tiles (rectangular boxes) P in $\mathbb{R}^n$ having the following properties:*

(i) Each P has $n-1$ side lengths equal to $R^{\frac{1}{2}}$ and one side length equal to R which points in the direction of the unit vector v_P.

(ii) $v_{P_1} \wedge \cdots \wedge v_{P_n} \geq \nu$ for each $P_i \in \mathcal{P}_i$.

(iii) All tiles are subsets of a fixed cube B_{4R} with side length $4R$.

Then we have the following inequality:

$$\fint_{B_{4R}} \Big| \prod_{j=1}^n F_j \Big|^{\frac{1}{n-1}} \lesssim_{\epsilon, \nu} R^\epsilon \prod_{j=1}^n \Big| \fint_{B_{4R}} F_j \Big|^{\frac{1}{n-1}} \tag{9.1}$$

for all functions F_j of the form

$$F_j = \sum_{P \in \mathcal{P}_j} c_P 1_P.$$

The implicit constant will not depend on $R, c_P, \mathcal{P}_j$.

We use this to prove the following key result.

Theorem 9.2 *Let $p \geq \frac{2n}{n-1}$ and $\delta < 1$. Consider n ν-transverse cubes $Q_1, \cdots, Q_n \subset [0,1]^{n-1}$. Let B be an arbitrary cube in $\mathbb{R}^n$ with side length δ^{-2}, and let $\mathcal{B}$ be the (unique) partition of B into cubes Δ of side length δ^{-1}. Then for each $g : [0,1] \to \mathbb{C}$ we have*

$$\frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} \Big[\prod_{i=1}^n \Big(\sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} \|E_{Q_{i,1}} g\|^2_{L_\sharp^{\frac{p(n-1)}{n}}(w_\Delta)} \Big)^{\frac{1}{2}} \Big]^{\frac{p}{n}} \tag{9.2}$$

$$\lesssim_{\epsilon, \nu} \delta^{-\epsilon} \Big[\prod_{i=1}^n \Big(\sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} \|E_{Q_{i,1}} g\|^2_{L_\sharp^{\frac{p(n-1)}{n}}(w_B)} \Big)^{\frac{1}{2}} \Big]^{\frac{p}{n}}. \tag{9.3}$$

Moreover, the implicit constant is independent of g, δ, B.

Remark 9.1 This result is part of a two-stage process. Note that, strictly speaking, this inequality is not a decoupling, since the size on the frequency cubes $Q_{i,1}$ remains unchanged. However, the size of the spatial cube increases from δ^{-1} to δ^{-2}, which will facilitate a subsequent decoupling, as we shall later see in Proposition 10.1.

Proof Since we can afford logarithmic losses in δ, it suffices to prove the inequality with the summation on both sides restricted to families of $Q_{i,1}$ for which $\|E_{Q_{i,1}} g\|_{L_\sharp^{\frac{p(n-1)}{n}}(w_B)}$ have comparable size (within a multiplicative factor of 2) for each i. Indeed, the cubes $Q'_{i,1}$ satisfying (for some large enough $C = O(1)$)

$$\|E_{Q'_{i,1}} g\|_{L_\sharp^{\frac{p(n-1)}{n}}(w_B)} \leq \delta^C \max_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} \|E_{Q_{i,1}} g\|_{L_\sharp^{\frac{p(n-1)}{n}}(w_B)}$$

can be easily dealt with by using the triangle inequality, since we automatically have

$$\max_{\Delta \in \mathcal{B}} \|E_{Q'_{i,1}} g\|_{L_{\sharp}^{\frac{p(n-1)}{n}}(w_{\Delta})} \leq \delta^{C} \max_{Q_{i,1} \in \mathrm{Part}_{\delta}(Q_i)} \|E_{Q_{i,1}} g\|_{L_{\sharp}^{\frac{p(n-1)}{n}}(w_B)}.$$

This leaves only $\log_2(\delta^{-O(1)})$ sizes to consider.

Let us now assume that we have N_i cubes $Q_{i,1}$, with $\|E_{Q_{i,1}} g\|_{L_{\sharp}^{\frac{p(n-1)}{n}}(w_B)}$ of comparable size. Since $p \geq \frac{2n}{n-1}$, by Hölder's inequality, (9.2) is at most

$$\Big(\prod_{i=1}^{n} N_i^{\frac{1}{2} - \frac{n}{p(n-1)}} \Big)^{\frac{p}{n}} \frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} \Big[\prod_{i=1}^{n} \Big(\sum_{Q_{i,1}} \|E_{Q_{i,1}} g\|_{L_{\sharp}^{\frac{p(n-1)}{n}}(w_{\Delta})}^{\frac{p(n-1)}{n}} \Big) \Big]^{\frac{1}{n-1}}. \qquad (9.4)$$

For each cube $Q = Q_{i,1}$ centered at c_Q, we cover B with a family $\mathcal{F}_Q$ of pairwise disjoint, mutually parallel tiles T_Q. They have $n-1$ short sides of length δ^{-1} and one longer side of length δ^{-2}, pointing in the direction of the normal $N(c_Q)$ to the paraboloid $\mathbb{P}^{n-1}$ at c_Q. Moreover, we can assume these tiles to be inside the cube $4B$. We let $T_Q(x)$ be the tile containing x, and we let $2T_Q$ be the dilation of T_Q by a factor of 2 around its center.

Let us use q to abbreviate $\frac{p(n-1)}{n}$. Our goal is to control the expression

$$\frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} \prod_i \Big(\sum_{Q_{i,1}} \|E_{Q_{i,1}} g\|_{L_{\sharp}^q(w_{\Delta})}^q \Big)^{\frac{1}{n-1}}.$$

We now define F_Q for $x \in \cup_{T_Q \in \mathcal{F}_Q} T_Q$ by

$$F_Q(x) := \sup_{y \in 2T_Q(x)} \|E_Q g\|_{L_{\sharp}^q(w_{B(y, \delta^{-1})})}.$$

For any point $x \in \Delta$, we have $\Delta \subset 2T_Q(x)$, and so we also have

$$\|E_Q g\|_{L_{\sharp}^q(w_{\Delta})} \leq F_Q(x).$$

Therefore,

$$\frac{1}{|\mathcal{B}|} \sum_{\Delta \in \mathcal{B}} \prod_i \Big(\sum_{Q_{i,1}} \|E_{Q_{i,1}} g\|_{L_{\sharp}^q(w_{\Delta})}^q \Big)^{\frac{1}{n-1}} \lesssim \fint_{4B} \prod_i \Big(\sum_{Q_{i,1}} F_{Q_{i,1}}^q \Big)^{\frac{1}{n-1}}.$$

Moreover, the function F_Q^q is constant on each tile $T_Q \in \mathcal{F}_Q$. Applying Theorem 9.1, we get the bound

$$\fint_{4B} \prod_i \Big(\sum_{Q_{i,1}} F_{Q_{i,1}}^q \Big)^{\frac{1}{n-1}} \lesssim_{\epsilon, \nu} \delta^{-\epsilon} \prod_i \Big(\sum_{Q_{i,1}} \fint_{4B} F_{Q_{i,1}}^q \Big)^{\frac{1}{n-1}}.$$

It remains to check that for each $Q = Q_{i,1}$,

$$\|F_Q\|_{L_{\sharp}^q(4B)} \lesssim \|E_Q g\|_{L_{\sharp}^q(w_B)}. \qquad (9.5)$$

Once this is established, it follows that (9.4) is dominated by

$$\delta^{-\epsilon} \Big(\prod_{i=1}^{n} N_i^{\frac{1}{2} - \frac{n}{p(n-1)}} \Big)^{\frac{p}{n}} \prod_{i=1}^{n} \Big(\sum_{Q_{i,1}} \|E_{Q_{i,1}} g\|_{L_{\sharp}^{\frac{p(n-1)}{n}}(w_B)}^{\frac{p(n-1)}{n}} \Big)^{\frac{1}{n-1}}. \qquad (9.6)$$

Recalling the restriction, we have made on $Q_{i,1}$, (9.6) is comparable to

$$\delta^{-\epsilon}\Big[\prod_{i=1}^{n}\Big(\sum_{Q_{i,1}}\|E_{Q_{i,1}}g\|^2_{L_\sharp^{\frac{p(n-1)}{n}}(w_B)}\Big)^{\frac{1}{2}}\Big]^{\frac{p}{n}},$$

as desired.

To prove (9.5), we may assume $Q = \big[-\frac{\delta}{2},\frac{\delta}{2}\big]^{n-1}$, and thus $\widehat{E_Q g}$ will be supported in $[-\delta,\delta]^{n-1}\times[-\delta^2,\delta^2]$. Fix $x=(x_1,\cdots,x_n)$ with $T_Q(x)\in\mathcal{F}_Q$ and let $y\in 2T_Q(x)$. Note that $T_Q(x)$ has sides parallel to the coordinate axes. In particular, $y=x+y'$ with $|y'_j|<4\delta^{-1}$ for $1\le j\le n-1$ and $|y'_n|<4\delta^{-2}$. Then

$$\|E_Q g\|^q_{L^q(w_{B(y,\delta^{-1})})}$$
$$\lesssim \int |E_Q g(x_1+u_1,\cdots,x_{n-1}+u_{n-1},x_n+u_n+y'_n)|^q w_{B(0,\delta^{-1})}(u)\mathrm{d}u. \qquad (9.7)$$

Now, using Taylor series, we can write

$$|E_Q g(x_1+u_1,\cdots,x_{n-1}+u_{n-1},x_n+u_n+y'_n)|$$
$$=\Big|\int \widehat{E_Q g}(\lambda)e(\lambda\cdot(x+u))e(\lambda_n y'_n)\mathrm{d}\lambda\Big|$$
$$\le \sum_{s_n=0}^{\infty}\frac{1000^{s_n}}{s_n!}\Big|\int \widehat{E_Q g}(\lambda)e(\lambda\cdot(x+u))\Big(\frac{\lambda_n}{2\delta^2}\Big)^{s_n}\mathrm{d}\lambda\Big|$$
$$=\sum_{s_n=0}^{\infty}\frac{1000^{s_n}}{s_n!}|M_{s_n}(E_Q g)(x+u)|.$$

Here M_{s_n} is the operator with Fourier multiplier $1_{\mathbb{R}^{n-1}}(\lambda_1,\cdots,\lambda_{n-1})m_{s_n}\big(\frac{\lambda_n}{2\delta^2}\big)$, where

$$m_{s_n}(\lambda_n)=(\lambda_n)^{s_n}1_{[-\frac{1}{2},\frac{1}{2}]}(\lambda_n).$$

We are able to insert the cutoff because of our initial restriction on the Fourier support of $E_Q g$.

Plugging this estimate into (9.7), we obtain

$$\|E_Q g\|^q_{L_\sharp^q(w_{B(y,\delta^{-1})})}\lesssim \sum_{s_n=0}^{\infty}\frac{1000^{s_n}}{s_n!}\|M_{s_n}(E_Q g)\|^q_{L_\sharp^q(w_{B(x,\delta^{-1})})}.$$

Recalling the definition of F_Q and the fact that

$$\delta^n\int_{4B}w_{B(x,\delta^{-1})}(z)\mathrm{d}x\lesssim w_B(z),\quad z\in\mathbb{R}^n,$$

we conclude that

$$\|F_Q\|^q_{L_\sharp^q(4B)}\lesssim \sum_{s_n=0}^{\infty}\frac{1000^{s_n}}{s_n!}\|M_{s_n}(E_Q g)\|^q_{L_\sharp^q(w_B)}. \qquad (9.8)$$

Note that $t\mapsto t^{s_n}1_{[-\frac{1}{2},\frac{1}{2}]}(t)$ agrees on $[-\frac{1}{2},\frac{1}{2}]$ with a compactly supported smooth function $m^*_{s_n}$ defined on $\mathbb{R}$, with derivatives of any given order uniformly bounded over s_n. It follows that

$$|\widehat{m^*_{s_n}}(x_n)|\lesssim \xi(x_n)$$

with implicit constant independent of s_n, where

$$\xi(x_n) \lesssim_M (1 + |x_n|)^{-M}$$

for all $M > 0$. Let $M^*_{s_n}$ denote the operator with multiplier $1_{\mathbb{R}^{n-1}}(\lambda_1, \cdots, \lambda_{n-1}) m^*_{s_n}(\frac{\lambda_n}{2\delta^2})$. We can now write

$$|M_{s_n}(E_Q g)(x)| = |M^*_{s_n}(E_Q g)(x)| \lesssim |E_Q g| \odot \xi_{\delta^2}(x),$$

where $\odot$ denotes the convolution with respect to the last variable x_n, and

$$\xi_{\delta^2}(x_n) = \delta^2 \xi(\delta^2 x_n).$$

Using this, one can easily check that

$$\|M_{s_n}(E_Q g)\|^q_{L^q(w_B)} \lesssim \langle |E_Q g|^q \odot \xi_{\delta^2}, w_B \rangle = \langle |E_Q g|^q, \xi_{\delta^2} \odot w_B \rangle \lesssim \langle |E_Q g|^q, w_B \rangle.$$

Combining this with (9.8) leads to the proof of (9.5)

$$\|F_Q\|^q_{L^q_\sharp(4B)} \lesssim \sum_{s_n=0}^{\infty} \frac{1000^{s_n}}{s_n!} \|E_Q g\|^q_{L^q_\sharp(w_B)} \lesssim \|E_Q g\|^q_{L^q_\sharp(w_B)}.$$

The argument is now complete.

10 The Iteration Scheme

Let $0 < \nu < 1$. Throughout this section, we fix some $0 < \delta < 1$ and also n ν-transverse cubes $Q_1, \cdots, Q_n \subset [0,1]^{n-1}$ with side length at least δ.

For a positive integer s, B^s will refer to cubes in $\mathbb{R}^n$ with side length $l(B^s) = \delta^{-s}$ and arbitrary centers. We will only encounter cubes $B \subset \mathbb{R}^n$ with side length $l(B) \in 2^{\mathbb{N}}$. This will allow us to perform decompositions by using cubes of smaller size in $2^{\mathbb{N}}$.

The implicit constants will be independent of δ, g and the spatial cubes Q_i.

Let $t, p \geq 1$ and consider the positive integers $q \leq s \leq r$. We define

$$D_t(q, B^r, g) = \Big[\prod_{i=1}^{n} \Big(\sum_{Q_{i,q} \in \text{Part}_{\delta q}(Q_i)} \|E_{Q_{i,q}} g\|^2_{L^t_\sharp(w_{B^r})} \Big)^{\frac{1}{2}} \Big]^{\frac{1}{n}}.$$

To simplify notation, we will denote by $\mathcal{B}_s(B^r) = \text{Part}_{\delta^{-s}}(B^r)$ the (unique) cover of B^r with cubes B^s of side length δ^{-s}. Define

$$A_p(q, B^r, s, g) = \Big(\frac{1}{|\mathcal{B}_s(B^r)|} \sum_{B^s \in \mathcal{B}_s(B^r)} D_2(q, B^s, g)^p \Big)^{\frac{1}{p}}.$$

The letter A will remind us that we have an average. Note that when $r = s$,

$$A_p(q, B^r, r, g) = D_2(q, B^r, g).$$

For $\frac{2n}{n-1} \leq p$, let $0 \leq \kappa_p \leq 1$ satisfy

$$\frac{n}{p(n-1)} = \frac{1 - \kappa_p}{2} + \frac{\kappa_p}{p}.$$

In other words,

$$\kappa_p = \frac{pn - p - 2n}{(p-2)(n-1)}.$$

Set also $\kappa_p = 0$ for $2 \leq p \leq \frac{2n}{n-1}$.

The next proposition will combine our main two decoupling devices, Theorem 9.2 and the L^2 decoupling. The result is a partial decoupling. Indeed, note that the term $A_p(1, B^2, 1, g)$ in (10.1) involves frequency cubes of size δ, while the term $A_p(2, B^2, 2, g)$ involves frequency cubes of smaller size δ^2. Inequality (10.1) is only a partial decoupling in the range $p > \frac{2n}{n-1}$, since the weight κ_p of the term $D_p(1, B^2, g)$ is nonzero. But this weight is zero when $p \leq \frac{2n}{n-1}$. For these values of p, inequality (10.1) has the very simple form

$$A_p(1, B^2, 1, g) \lesssim_{\epsilon,\nu} \delta^{-\epsilon} A_p(2, B^2, 2, g).$$

This can be easily iterated and leads to a simpler proof of Theorem 1.1 in the range $2 \leq p \leq \frac{2n}{n-1}$. See the discussion at the end of this section.

Proposition 10.1 *We have for each B^2 and $p \geq 2$,*

$$A_p(1, B^2, 1, g) \lesssim_{\epsilon,\nu} \delta^{-\epsilon} A_p(2, B^2, 2, g)^{1-\kappa_p} D_p(1, B^2, g)^{\kappa_p}. \tag{10.1}$$

Proof Assume first that $p \geq \frac{2n}{n-1}$. By Hölder's inequality,

$$\|E_{Q_{i,1}} g\|_{L^2_\sharp(w_{B^1})} \lesssim \|E_{Q_{i,1}} g\|_{L^{\frac{p(n-1)}{n}}_\sharp (w_{B^1})}.$$

Using this and Theorem 9.2, we can write

$$A_p(1, B^2, 1, g) \lesssim_{\epsilon,\nu} \delta^{-\epsilon} \Big(\prod_{i=1}^n \sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} \|E_{Q_{i,1}} g\|^2_{L^{\frac{p(n-1)}{n}}_\sharp (w_{B^2})} \Big)^{\frac{1}{2n}}. \tag{10.2}$$

Using Hölder's inequality, we can dominate this by

$$\leq \Big(\prod_{i=1}^n \sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} \|E_{Q_{i,1}} g\|^2_{L^2_\sharp(w_{B^2})} \Big)^{\frac{1-\kappa_p}{2n}} \Big(\prod_{i=1}^n \sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} \|E_{Q_{i,1}} g\|^2_{L^p_\sharp(w_{B^2})} \Big)^{\frac{\kappa_p}{2n}}. \tag{10.3}$$

It suffices now to apply L^2 decoupling (Proposition 6.1) to the first term in (10.3).

We have "interpolated" between L^2 and L^p. We have used L^2 because (as explained in Section 6) this space facilitates the most efficient decoupling. Indeed, note that the term $A_p(2, B^2, 2, g)$ on the right-hand side of (10.1) has cubes of side length δ^2, which is as small as one can hope, given the size of the spatial cube B^2.

If $p < \frac{2n}{n-1}$, using (10.1) with $p = \frac{2n}{n-1}$, we can write

$$A_p(1, B^2, 1, g) \leq A_{\frac{2n}{n-1}}(1, B^2, 1, g) \lesssim_{\epsilon,\nu} \delta^{-\epsilon} A_{\frac{2n}{n-1}}(2, B^2, 2, g) = \delta^{-\epsilon} A_p(2, B^2, 2, g).$$

Inequality (10.1) is easily seen to be true with κ_p replaced with 1, by simply invoking (4.1) and the fact that $D_2 \lesssim D_p$. Consequently, it will be true for each exponent in the interval $[\kappa_p, 1]$. The example $g = 1_Q$ with $l(Q) = \delta$ shows that one can not consider exponents smaller

than κ_p. The relevant thing about κ_p that will be used in the final section is the fact that $\kappa_p < \frac{1}{2}$ precisely in the subcritical range $p < \frac{2(n+1)}{n-1}$.

The following sequence of propositions will allow us to rewrite (10.1) in a form that is more suitable for iteration.

Proposition 10.2 *We have for each cube B^M with $M \geq 2$ and $p \geq 2$,*

$$A_p(1, B^M, 1, g) \lesssim_{\epsilon,\nu} \delta^{-\epsilon} A_p(2, B^M, 2, g)^{1-\kappa_p} D_p(1, B^M, g)^{\kappa_p}. \tag{10.4}$$

The implicit constant is independent of M.

Proof Raising (10.1) to the power p, summing over all cubes $B^2 \in \mathcal{B}_2(B^M)$ and using Hölder's inequality, we have

$$\|(a_j b_j)_j\|_{l^1} \leq \|(a_j)_j\|_{l^{\frac{1}{1-\kappa_p}}} \|(b_j)_j\|_{l^{\frac{1}{\kappa_p}}}.$$

The only thing that needs to be verified is the inequality

$$\sum_{B^2 \in \mathcal{B}_2(B^M)} D_p(1, B^2, g)^p \lesssim D_p(1, B^M, g)^p. \tag{10.5}$$

This however immediately follows from Minkowski's inequality (recall $p \geq 2$) and the fact that

$$\sum_{B^2 \in \mathcal{B}_2(B^M)} w_{B^2} \lesssim w_{B^M}.$$

Proposition 10.3 *Let $l, m \in \mathbb{N}$ with $l + 1 \leq m$. We have for each cube B^{2^m} and $p \geq 2$,*

$$A_p(2^l, B^{2^m}, 2^l, g) \lesssim_{\epsilon,\nu} \delta^{-2^l \epsilon} A_p(2^{l+1}, B^{2^m}, 2^{l+1}, g)^{1-\kappa_p} D_p(2^l, B^{2^m}, g)^{\kappa_p}. \tag{10.6}$$

The implicit constant is independent of l, m.

Proof Apply (10.4) with δ replaced by δ^{2^l} and $M = 2^{m-l}$.

We can now iterate Proposition 10.3 to get the following immediate conclusion.

Proposition 10.4 *If $m \geq 1$ and $p \geq 2$,*

$$A_p(1, B^{2^m}, 1, g)$$
$$\lesssim_{\epsilon,\nu,m} \delta^{-\epsilon} A_p(2^{m-1}, B^{2^m}, 2^{m-1}, g)^{(1-\kappa_p)^{m-1}} \prod_{l=0}^{m-2} D_p(2^l, B^{2^m}, g)^{\kappa_p(1-\kappa_p)^l}. \tag{10.7}$$

The implicit constant is now allowed to depend on m, but this dependence will prove to be completely harmless.

We complete this section with a quick proof of

$$\mathrm{Dec}_n(\delta, p) \lesssim_\epsilon \delta^{-\epsilon}$$

for $2 \leq p \leq \frac{2n}{n-1}$. This fact was first proved in [2]. In this range $\kappa_p = 0$ and (10.7) becomes a very satisfactory inequality

$$A_p(1, B^{2^m}, 1, g) \lesssim_{\epsilon,\nu,m} \delta^{-\epsilon} A_p(2^{m-1}, B^{2^m}, 2^{m-1}, g).$$

Combining this with (11.5), we may write

$$\Big\| \prod_{i=1}^{n} \Big[\sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} |E_{Q_{i,1}}g|^2 \Big]^{\frac{1}{2}n} \Big\|_{L^p(B^{2^m})} \lesssim_{\epsilon,\nu,m} \delta^{-\epsilon} D_p(2^{m-1}, B^{2^m}).$$

By invoking Cauchy–Schwarz, we can afford a rather trivial decoupling

$$\Big\| \Big| \prod_{i=1}^{n} E_{Q_i}g \Big|^{\frac{1}{n}} \Big\|_{L^p(B^{2^m})} \leq \delta^{-\frac{n-1}{2}} \Big\| \prod_{i=1}^{n} \Big[\sum_{Q_{i,1} \in \mathrm{Part}_\delta(Q_i)} |E_{Q_{i,1}}g|^2 \Big]^{\frac{1}{2}n} \Big\|_{L^p(B^{2^m})}.$$

Combining these two and substituting $\delta^{2^m} \mapsto \delta$ we can write

$$\Big\| \Big| \prod_{i=1}^{n} E_{Q_i}g \Big|^{\frac{1}{n}} \Big\|_{L^p(B^1)} \lesssim_{\epsilon,\nu,m} \delta^{-\epsilon} \delta^{-(n-1)2^{-m-1}} D_p\Big(\frac{1}{2}, B^1\Big).$$

Now choose m as large as desired to argue that

$$\mathrm{Dec}_n(\delta, p, \nu) \lesssim_{\epsilon,\nu} \delta^{-\epsilon}.$$

Finally, combine this with Theorem 8.1 by using induction on n to argue that

$$\mathrm{Dec}_n(\delta, p) \lesssim_\epsilon \delta^{-\epsilon}.$$

Now back to the case $p > \frac{2n}{n-1}$. As mentioned earlier, (10.7) is only a partial decoupling in this range. The argument for this case presented in the next section will go as follows. Assume the linear decoupling constant satisfies $\mathrm{Dec}_n(\delta, p) \sim \delta^{-\eta_p}$. We will first apply parabolic rescaling to majorize the terms D_p in (10.7) by some powers of $\delta^{-\eta_p}$. Then we will combine (10.7) with a trivial decoupling (Cauchy–Schwarz) to derive an upper bound on the multilinear constant $\mathrm{Dec}_n(\delta, p, \nu)$ in terms of $\delta^{-\eta_p}$. We play this against Theorem 8.1, which produced a lower bound for $\mathrm{Dec}_n(\delta, p, \nu)$ involving $\delta^{-\eta_p}$. These will force η_p to be zero.

11 The Final Argument

In this section, we present the details for the proof of Theorem 1.1. Let $E \geq 100n$. By combining the triangle and Cauchy-Schwarz inequalities, we find that $\mathrm{Dec}_n(\delta, p, E) \lesssim \delta^{-C_p}$ for some C_p large enough. For $p \geq 2$, let $\eta_{p,n,E} = \eta_{p,E} \geq 0$ be the unique (finite) number, such that

$$\lim_{\delta \to 0} \mathrm{Dec}_n(\delta, p, E) \delta^{\eta_{p,E}+\sigma} = 0 \quad \text{for each } \sigma > 0 \tag{11.1}$$

and

$$\limsup_{\delta \to 0} \mathrm{Dec}_n(\delta, p, E) \delta^{\eta_{p,E}-\sigma} = \infty \quad \text{for each } \sigma > 0. \tag{11.2}$$

We will use induction on n, as described at the end of Section 3. Assume either $n = 2$ or $n \geq 3$, and in addition, assume that we have

$$\mathrm{Dec}_{n-1}(\delta, p, E) \lesssim_\epsilon \delta^{-\epsilon}$$

for $E \geq 100(n-1)$ and $2 \leq p \leq \frac{2n}{n-2}$. We need to prove $\eta_{p,n,E} = 0$ for $E \geq 100n$ and $2 \leq p \leq \frac{2(n+1)}{n-1}$. Note that for such p, we automatically have that p is smaller than $\frac{2n}{n-2}$, the critical index for decouplings in $\mathbb{R}^{n-1}$. In particular, if $n \geq 3$, our induction hypothesis guarantees that

$$\mathrm{Dec}_{n-1}(\delta, p, E) \lesssim_\epsilon \delta^{-\epsilon} \tag{11.3}$$

for each $E \geq 100n$ and $2 \leq p \leq \frac{2(n+1)}{n-1}$.

Fix $2 \leq p < \frac{2(n+1)}{n-1}$ for the rest of the proof (so in particular we have $p \leq 20$). Fix also $E \geq 100n$. All quantities A_p and D_p will be implicitly assumed to be relative to this E. The case $p = \frac{2(n+1)}{n-1}$ will follow via a standard limiting argument explained in the end of the section. Note that for $2 \leq p < \frac{2(n+1)}{n-1}$, we have

$$2(1 - \kappa_p) > 1. \tag{11.4}$$

We start with the following rather immediate consequence of Proposition 10.4.

Theorem 11.1 *Consider n ν-transverse cubes $Q_1, \cdots, Q_n \subset [0,1]^{n-1}$ with side length at least δ. Then for $m \geq 1$ and $p \geq 2$, we have*

$$A_p(1, B^{2^m}, 1, g) \lesssim_{\epsilon,\nu,m} \delta^{-(\eta_p + \epsilon)\left(2^m - \frac{2\kappa_p}{2\kappa_p - 1} + \frac{(2(1-\kappa_p))^m}{2\kappa_p - 1}\right)} D_p(2^{m-1}, B^{2^m}, g)$$

with the implicit constant independent of Q_i.

Proof This will follow from Proposition 10.4, once we make a few observations.

First,

$$A_p(2^{m-1}, B^{2^m}, 2^{m-1}, g) \lesssim D_p(2^{m-1}, B^{2^m}, g). \tag{11.5}$$

This is a consequence of Hölder's inequality, Minkowski's inequality in $l^{\frac{p}{2}}$ and (4.1) (like (10.5) very much).

Second, an application of Proposition 7.1 shows that

$$D_p(2^l, B^{2^m}, g) \lesssim \mathrm{Dec}_n(\delta^{2^m - 2^{l+1}}, p) D_p(2^{m-1}, B^{2^m}, g).$$

Finally, combine these with (11.1) and Proposition 10.4.

By replacing δ^{2^m} with δ, we prefer to write the inequality in Theorem 11.1 as follows:

$$A_p(2^{-m}, B^1, 2^{-m}, g) \lesssim_{\epsilon,\nu,m} \delta^{-(\eta_p + \epsilon)\left(1 - 2^{-m}\frac{2\kappa_p}{2\kappa_p - 1} + \frac{(1-\kappa_p)^m}{2\kappa_p - 1}\right)} D_p\left(\frac{1}{2}, B^1, g\right) \tag{11.6}$$

with the implicit constant independent of the cubes Q_i. Here the assumption is $l(Q_i) \geq \delta^{2^{-m}}$.

Let $B = B^1$ be a cube in $\mathbb{R}^n$ with $l(B) = \delta^{-1}$. Consider n ν-transverse cubes $Q_1, \cdots, Q_n \subset [0,1]^{n-1}$ with side length $\mu \geq \delta^{2^{-m}}$. Let, as before, $\mathrm{Part}_{\mu^{-1}}(B)$ denote the partition of B by using cubes Δ with $l(\Delta) = \mu^{-1}$. Denote also by $\mathcal{B}_m(B)$ the partition of B by using cubes Δ_m with $l(\Delta_m) = \delta^{-2^{-m}}$.

We may write, first by combining Cauchy–Schwarz and (4.1),

$$\left[\frac{1}{|\mathcal{B}_{\mu^{-1}}(B)|}\sum_{\Delta\in\mathrm{Part}_{\mu^{-1}}(B)}\left(\prod_{i=1}^n\|E_{Q_i}g\|_{L^p_\sharp(w_{\Delta,10E})}^p\right)^{\frac{1}{n}}\right]^{\frac{1}{p}} \tag{11.7}$$

$$\lesssim \delta^{-(n-1)2^{-m-1}}\left[\frac{1}{|\mathcal{B}_m(B)|}\sum_{\Delta_m\in\mathcal{B}_m(B)}\left(\prod_{i=1}^n\left\|\left(\sum_{q_i\in\mathrm{Part}_{\delta^{2-m}}(Q_i)}|E_{q_i}g|^2\right)^{\frac{1}{2}}\right\|_{L^p_\sharp(w_{\Delta_m,10E})}^p\right)^{\frac{1}{n}}\right]^{\frac{1}{p}}, \tag{11.8}$$

and then by using Minkowski's inequality and (4.2) (recall that $p\le 20$), we have

$$\le \delta^{-(n-1)2^{-m-1}}\left[\frac{1}{|\mathcal{B}_m(B)|}\sum_{\Delta_m\in\mathcal{B}_m(B)}\left(\prod_{i=1}^n\sum_{q_i\in\mathrm{Part}_{\delta^{2-m}}(Q_i)}\|E_{q_i}g\|_{L^p_\sharp(w_{\Delta_m,10E})}^2\right)^{\frac{p}{2n}}\right]^{\frac{1}{p}} \tag{11.9}$$

$$\le \delta^{-(n-1)2^{-m-1}}\left[\frac{1}{|\mathcal{B}_m(B)|}\sum_{\Delta_m\in\mathcal{B}_m(B)}\left(\prod_{i=1}^n\sum_{q_i\in\mathrm{Part}_{\delta^{2-m}}(Q_i)}\|E_{q_i}g\|_{L^2_\sharp(w_{\Delta_m,E})}^2\right)^{\frac{p}{2n}}\right]^{\frac{1}{p}}$$

$$= A_p(2^{-m}, B^1, 2^{-m}, g).$$

Invoking (11.6) and removing the normalization, we conclude that

$$\left[\sum_{\Delta\in\mathrm{Part}_{\mu^{-1}}(B)}\left(\prod_{i=1}^n\|E_{Q_i}g\|_{L^p(w_\Delta,10E)}^p\right)^{\frac{1}{n}}\right]^{\frac{1}{p}}$$

$$\lesssim_{\epsilon,\nu,m} \delta^{-(\eta_p+\epsilon)\left(1-2^{-m}\frac{2\kappa_p}{2\kappa_p-1}+\frac{(1-\kappa_p)^m}{2\kappa_p-1}\right)}\delta^{-(n-1)2^{-m-1}}\left[\prod_{i=1}^n\sum_{q_i\in\mathrm{Part}_{\delta^{\frac{1}{2}}}(Q_i)}\|E_{q_i}g\|_{L^p(w_{B,E})}^2\right]^{\frac{1}{2n}}.$$

By taking a supremum over all Q_i, B, g as above, we deduce the following inequality, which is a stronger substitute for (8.2):

$$\mathrm{Dec}_n(\delta, p, \nu, m, E) \lesssim_{\epsilon,\nu,m} \delta^{-(\eta_p+\epsilon)\left(1-2^{-m}\frac{2\kappa_p}{2\kappa_p-1}+\frac{(1-\kappa_p)^m}{2\kappa_p-1}\right)}\delta^{-(n-1)2^{-m-1}}. \tag{11.10}$$

Combining this with Theorem 8.1 (use (11.3)) and (11.2), we may now write

$$\delta_l^{-\eta_p+\epsilon+\epsilon(\nu)} \lesssim_{\epsilon,\nu,m} \delta_l^{-(\eta_p+\epsilon)\left(1-2^{-m}\frac{2\kappa_p}{2\kappa_p-1}+\frac{(1-\kappa_p)^m}{2\kappa_p-1}\right)}\delta_l^{-(n-1)2^{-m-1}}$$

for some sequence δ_l converging to zero. This in turn forces

$$-\eta_p+\epsilon+\epsilon(\nu) \ge -(\eta_p+\epsilon)\left(1-2^{-m}\frac{2\kappa_p}{2\kappa_p-1}+\frac{(1-\kappa_p)^m}{2\kappa_p-1}\right) - (n-1)2^{-m-1}$$

for each $\epsilon,\nu>0$. Thus, letting $\epsilon,\nu\to 0$, we get

$$-\eta_p \ge -\eta_p\left(1-2^{-m}\frac{2\kappa_p}{2\kappa_p-1}+\frac{(1-\kappa_p)^m}{2\kappa_p-1}\right) - (n-1)2^{-m-1},$$

and by rearranging terms, we have

$$(n-1)2^{-1} \ge \eta_p\frac{[2(1-\kappa_p)]^m - 2\kappa_p}{1-2\kappa_p}. \tag{11.11}$$

As this holds for each $m \geq 1$, (11.4) will immediately force $\eta_p = 0$.

Let us now show that $\eta_{p_n} = 0$ for $p_n = \frac{2(n+1)}{n-1}$. Let $B \subset \mathbb{R}^n$ be a cube with $l(B) = \delta^{-1}$. Using inequality (4.2), for $p < p_n$, we can write

$$\|E_{[0,1]^{n-1}}g\|_{L^{p_n}(B)} \lesssim \|E_{[0,1]^{n-1}}g\|_{L^p(w_B)}.$$

Combining this with Hölder's inequality, we get

$$\|E_{[0,1]^{n-1}}g\|_{L^{p_n}(B)} \lesssim \mathrm{Dec}_n(\delta, p)\Big(\sum_{Q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}([0,1]^{n-1})} \|E_Q g\|^2_{L^p(w_B)} \Big)^{\frac{1}{2}}$$

$$\lesssim_\epsilon \delta^{-\epsilon}\|\mathbf{1}\|_{L^{\frac{q}{q-1}}(w_B)} \Big(\sum_{Q \in \mathrm{Part}_{\delta^{\frac{1}{2}}}([0,1]^{n-1})} \|E_Q g\|^2_{L^{p_n}(w_B)} \Big)^{\frac{1}{2}}.$$

It suffices to note that $q \to 1$ as $p \to p_n$.

References

[1] Bennett, J., Carbery, A. and Tao, T., On the multilinear restriction and Kakeya conjectures, *Acta Math.*, **196**(2), 2006, 261–302.

[2] Bourgain, J., Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces, *Israel J. Math.*, **193**(1), 2013, 441–458.

[3] Bourgain, J. and Demeter, C., The proof of the l^2 decoupling conjecture, *Annals of Math.*, **182**(1), 2015, 351–389.

[4] Bourgain, J., Demeter, C. and Guth, L., Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three, *Ann. of Math.*, **184**(2), 2016, 633–682.

[5] Bourgain, J. and Guth, L., Bounds on oscillatory integral operators based on multilinear estimates, *Geom. Funct. Anal.*, **21**(6), 2011, 1239–1295.

[6] Guth, L., A short proof of the multilinear Kakeya inequality, *Math. Proc. Cambridge Philos. Soc.*, **158**(1), 2015, 147–153.