STABILITY OF WEIGHTED NORM INEQUALITIES

MICHEL ALEXIS, JOSE LUIS LUNA-GARCIA, ERIC T. SAWYER, AND IGNACIO URIARTE-TUERO

Abstract. We show that while individual Riesz transforms are two weight norm stable under biLipschitz change of variables on $A_\infty$ weights, they are two weight norm unstable under even rotational change of variables on doubling weights. More precisely, we show that individual Riesz transforms are unstable under a set of rotations having full measure, which includes rotations arbitrarily close to the identity. This provides an operator theoretic distinction between $A_\infty$ weights and doubling weights.

More generally, all iterated Riesz transforms of odd order are rotationally unstable on pairs of doubling weights, thus demonstrating the need for characterizations of iterated Riesz transform inequalities using testing conditions as in [AlSaUr], as opposed to the typically stable ‘bump’ conditions.

Contents

1. Introduction 2
1.1. Previous stability results 2
1.2. Description of results 3
1.3. BiLipschitz and rotational stability 3
1.4. History of stability 5
1.5. Proof of Stability 7
1.6. Open Problems 8
2. Preliminaries: grids, doubling, telescoping identities and dyadic testing 8
2.1. Notation for grids and cubes 8
2.2. Doubling 9
2.3. Telescoping identity 9
2.4. Dyadic testing 11
2.5. The Bellman construction of the dyadic weights 11
3. The supervisor and transplantation map 13
3.1. The informal description of the construction 13
3.2. The formal construction 13
4. Weak Convergence Properties of the Riesz transforms 16
4.1. Weak convergence properties of the Hilbert transform 17
4.2. From Hilbert to Riesz 19
5. Boundedness properties of the Riesz transforms 27
5.1. The formal argument in the plane 27
5.2. The Nazarov argument for Riesz transforms 28
5.3. Classical doubling, $A_2$ and dyadic testing in $\mathbb{R}^n$ 33
6. Iterated Riesz transforms 39
6.1. Rotations 41
6.2. Completion of proofs of main theorems 41
7. Appendix 42
7.1. BiLipschitz instability of the Hilbert transform for arbitrary weight pairs 42
7.2. Beyond biLipschitz maps for $A_2$ stability 42
7.3. Modification of Transplantation to get Classical Doubling and $A_2$ 43
References 46

E. Sawyer is partially supported by a grant from the National Research Council of Canada.
I. Uriarte-Tuero has been partially supported by grant MTM2015-65792-P (MINECO, Spain), and is partially supported by a grant from the National Research Council of Canada.
1. Introduction

We begin by describing two stability theorems for operator norms given three decades apart, that motivate the main results of this paper.

1.1. Previous stability results. Thirty-five years ago, Johnson and Neugebauer [JoNe] Theorem 2.10 (a), see also the preceding Remark 1] characterized the smooth homeomorphisms $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ that preserve Muckenhoupt’s $A_p(\mathbb{R}^n)$ condition on a weight $w$ under pushforward by $\Phi$, as precisely those quasiconformal maps $\Phi$ having their Jacobian $J = |\det D\Phi|$ in the intersection $\bigcap_{r > 1} A_r(\mathbb{R}^n)$ of the $A_r$ classes over $r > 1$. A variant of the one-dimensional case of this beautiful characterization, see [JoNe, Theorem 2.7 with $Hf$], can be reformulated in terms of stability of the ‘Muckenhoupt’ one weight norm inequality for the Hilbert transform under homeomorphisms of the real line.

**Theorem 1.** Suppose that $w : \mathbb{R} \to [0, \infty)$ is a nonnegative weight on the real line $\mathbb{R}$, that $\varphi : \mathbb{R} \to \mathbb{R}$ is an increasing homeomorphism with $\varphi$ and $\varphi^{-1}$ absolutely continuous, and that $H$ is the Hilbert transform, $Hf(x) = \text{pv} \int_{-\infty}^{\infty} \frac{f(y)}{x-y} \, dy$.

For $1 < p < \infty$, denote by $\mathcal{N}_{H,p}[w]$ the operator norm of the map $H : L^p(w) \to L^p(w)$, i.e. the best constant $C$ in the inequality

$$\int_{\mathbb{R}} |Hf(x)|^p \, w(x) \, dx \leq C \int_{\mathbb{R}} |f(x)|^p \, w(x) \, dx.$$ 

Then there is a positive constant $C_1$ such that

$$\mathcal{N}_{H,p}[(w \circ \varphi) \varphi'] \leq C_1 \mathcal{N}_{H,p}[w], \quad \text{for all weights } w,$$

if and only if $\varphi' \in \bigcap_{r > 1} A_r(\mathbb{R})$.

More recently, Tolsa [To] see abstract] characterized the ‘Ahlfors-David’ one weight inequality for the Cauchy transform, equivalently the 1-fractional vector Riesz transform $R^{1,2}$ in the plane $\mathbb{R}^2$ (see (1.2) below), in the case $p = 2$, namely

$$\int_{\mathbb{R}^2} |R^{1,2}(f \mu)(x)|^2 \, d\mu(x) \leq \mathcal{N}_{R^{1,2}}(\mu) \int_{\mathbb{R}^2} |f(x)|^2 \, d\mu(x),$$

in terms of a growth condition and Menger curvature. As a consequence, Tolsa obtained stability of the operator norm $\mathcal{N}_{R^{1,2}}(\mu)$ under biLipschitz pushforwards of the measure $\mu$. Even more recently, in papers by Dąbrowski and Tolsa and Tolsa, this result was extended to higher dimensions, and as a consequence they obtained stability of the operator norm $\mathcal{N}_{R^{1,n}}(\mu)$ of the 1-fractional vector Riesz transform $R^{1,n}$ under biLipschitz pushforwards of the measure $\mu$ in $\mathbb{R}^n$ [DaTo], [To2]. As an important application of norm stability, they obtain the stability of removable sets for Lipschitz harmonic functions under biLipschitz mappings [To2] see Corollary 1.6 and the discussion surrounding it.

Here we define the $\alpha$-fractional vector Riesz transform in $\mathbb{R}^n$ by

$$R_{j}^{\alpha,n}(f)(x) \equiv c_{\alpha,n} \text{pv} \int_{\mathbb{R}^n} \frac{x-y}{|x-y|^{n+1-\alpha}} f(y) \, dy, \quad x \in \mathbb{R}^n, 0 \leq \alpha < n.$$ 

Let $R_{j}^{\alpha,n} = (R_{1}^{\alpha,n}, \ldots, R_{2}^{\alpha,n})$ where we refer to the components $R_{j}^{\alpha,n}$ as individual $\alpha$-fractional Riesz transforms in $\mathbb{R}^n$. In the classical case $\alpha = 0$, we will usually drop the superscript $\alpha$ and simply write $R = (R_{1}, \ldots, R_{n})$ when the dimension $n$ is understood, and refer to the components $R_{j}$ as Riesz transforms.

In this paper we are mainly concerned with the fractional order $\alpha = 0$.

The main problem we consider in this paper is the extent to which the above theorems hold in the setting of two weight norm inequalities, and to include more general operators in higher dimensions. The complexities inherent in dealing with two weight norm inequalities - mainly that they are no longer characterized simply by $A_p$-like conditions or more generally by conditions of ‘positive nature’, but require testing conditions of ‘singular nature’ as well - suggests that we should limit ourselves to consideration of biLipschitz maps. Indeed, this much smaller class of maps is much more amenable to current two weight techniques, and allows for a rich theory where stability holds in certain ‘nice’ situations, while failing in small perturbations of these ‘nice’ situations.

Our analysis will be mainly restricted to the case $p = 2$ and iterated Riesz transforms of odd order in $\mathbb{R}^n$, where we show that stability of the two weight norm inequality is sensitive to the distinction between doubling and $A_{\infty}$ weights, even when the biLipschitz maps are restricted to rotations of $\mathbb{R}^n$. 

1.2. Description of results. The two weight norm inequality for an operator $T$ with a pair $(\sigma, \omega)$ of positive locally finite Borel measures on $\mathbb{R}^n$ and exponents $1 < p \leq q < \infty$ is informally, 

\[
(1.3) \quad \left( \int_{\mathbb{R}^n} |T (f \sigma)^{q} |^q \, d\omega \right)^{\frac{1}{q}} \leq \mathcal{N}_T \left( \int_{\mathbb{R}^n} |f|^p \, d\sigma \right)^{\frac{1}{p}}, \quad f \in L^p (\sigma).
\]

See e.g. [AlSaUr, Theorem 1 (2) and (3)] for two common definitions of what it means for \((1.3)\) to hold, and which are equivalent at least in the case of doubling measures. In the case $p = q = 2$, we first establish a distinction between weighted norm inequalities for positive operators $T$ in \((1.3)\), such as the maximal function and fractional integrals, on the one hand; and singular integral operators $T$ in \((1.3)\), such as the individual Riesz transforms and iterated Riesz transforms, on the other hand. Namely, that the former are two weight norm stable under biLipschitz change of variables for arbitrary locally finite positive Borel measures, while the latter are not in general, even on pairs of doubling measures.

Our main result, Theorem \[4\] shows that while individual Riesz transforms are two weight norm stable under biLipschitz change of variables on pairs of $A_\infty$ weights, they are two weight norm unstable under even a rotational change of variables on doubling weights. This provides an operator theoretic distinction between $A_\infty$ weights and doubling weight. \[4\]

We also show that all iterated Riesz transforms of odd order are rotationally unstable on pairs of doubling weights, thus demonstrating the need for characterizations of iterated Riesz transform inequalities using unstable conditions, such as the testing conditions in [AlSaUr], as opposed to the typically stable ‘bump’ conditions.

1.3. BiLipschitz and rotational stability.

Definition 2. Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be continuous and invertible.

(1) $\Phi$ is biLipschitz if

\[
\|\Phi\|_{\text{biLip}} \equiv \sup_{x, y \in \mathbb{R}^n} \frac{|\Phi (x) - \Phi (y)|}{|x - y|} + \sup_{x, y \in \mathbb{R}^n} \frac{\|\Phi^{-1} (x) - \Phi^{-1} (y)\|}{|x - y|} < \infty.
\]

(2) $\Phi$ is a rotation if $\Phi$ is linear and $\Phi^* = I$ and $\det \Phi = 1$.

Let $\mathcal{X}$ be a group of continuous invertible maps on $\mathbb{R}^n$, such as the group of biLipschitz or rotation transformations, which we denote by $\mathcal{X}_{\text{biLip}}$ and $\mathcal{X}_{\text{rot}}$ respectively. Denote by $\mathcal{M}$ the space of positive Borel measures on $\mathbb{R}^n$, and by $\Phi_*, \mu$ the pushforward of $\mu \in \mathcal{M}$ by a continuous map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$, i.e., $\Phi_*(\mu) (B) \equiv \mu (\Phi^{-1} (B))$. We say that a subclass $\mathcal{S} \subset \mathcal{M}$ of positive Borel measures is $\mathcal{X}$-invariant if $\Phi_* \mu \in \mathcal{S}$ for all $\mu \in \mathcal{S}$ and $\Phi \in \mathcal{X}$. Of course $\mathcal{M}$ itself is $\mathcal{X}$-invariant for the group $\mathcal{X}_{\text{cont inv}}$ of all continuous invertible maps, but less trivial examples of biLipschitz invariant classes include,

\[
(1.4) \quad \mathcal{S}_{A_p} \equiv \{ \mu \in \mathcal{M} : d\mu (x) = u (x) \, dx \text{ with } u \in A_p \}, \quad \text{for } 1 \leq p < \infty,
\]

\[
\mathcal{S}_{A_\infty} \equiv \{ \mu \in \mathcal{M} : d\mu (x) = u (x) \, dx \text{ with } u \in A_\infty \},
\]

\[
\mathcal{S}_{\text{doub}} \equiv \{ \mu \in \mathcal{M} : \mu \text{ is a doubling measure} \},
\]

\[
\mathcal{S}_{\text{Ahds}} \equiv \{ \mu \in \mathcal{M} : \mu \text{ is Ahlfors-David regular of degree } s \},
\]

\[
\mathcal{S}_{\text{fpB}} \equiv \{ \mu \in \mathcal{M} : \mu \text{ is a locally finite positive Borel measure} \}.
\]

To each of the above classes $\mathcal{S}$ we can associate a functional $\|\mu\|_{\mathcal{S}}$ for which $\mathcal{S} \equiv \{ \mu \in \mathcal{M} : \|\mu\|_{\mathcal{S}} < \infty \}$. For example we take

\[
(1.5) \quad \|\mu\|_{\mathcal{S}_{A_\infty}} = [\mu]_{A_\infty} \equiv \sup_Q \left( \frac{1}{|Q|} \int_Q \mu \right) \exp \left( \frac{1}{|Q|} \int_Q \ln \frac{1}{\mu} \right),
\]

and $\|\mu\|_{\mathcal{S}_{\text{doub}}} = C_{\text{doub}} (\mu)$ as in Definition \[9\]. In the case that $\mathcal{S} = \mathcal{S}_{\text{fpB}}$, there is no ‘natural’ choice of $\|\cdot\|_{\mathcal{S}}$ that measures the ‘size’ of the measure $\mu$ and so instead we define

\[
\|\mu\|_{\mathcal{S}_{\text{fpB}}} = \begin{cases} 
1 & \text{if } \mu \in \mathcal{S}_{\text{fpB}} \\
\infty & \text{otherwise}
\end{cases}.
\]

\[1\]In 1974, C. Fefferman and B. Muckenhoupt [FeMu] constructed an example of a doubling weight that was not $A_\infty$ using a self similar construction, on which many subsequent results have been based.
We also define
\[
\|\Phi\|_{X} = \begin{cases} 
\|\Phi\|_{biLip} & \text{if } X = X_{biLip} \\
1 & \text{if } X = X_{rot} \text{ and } \Phi \in X_{rot} \\
\infty & \text{if } X = X_{rot} \text{ and } \Phi \not\in X_{rot}.
\end{cases}
\]

Here is the main stability definition for a function \( F \) on measure pairs, a group \( X \in \{ X_{biLip}, X_{rot} \} \) and an \( X \)-invariant class \( S \) (or to be precise, for \( (S, \| \cdot \|_{S}) \)).

**Definition 3.** Let \( X \in \{ X_{biLip}, X_{rot} \} \), \( S \subset M \) be \( X \)-invariant, and let \( \mathcal{F} : S \times S \to [0, \infty] \) be a nonnegative extended real-valued function on the product set \( S \times S \). We say that the function \( \mathcal{F} \) is \( X \)-stable on \( S \) if there is a function \( \mathcal{G} : [0, \infty)^{4} \to [0, \infty) \) which maps bounded subsets of \([0, \infty)^{4}\) to bounded subsets of \([0, \infty)\), such that
\[
\mathcal{F}(\Phi, \sigma, \Phi, \omega) \leq \mathcal{G}(\|\Phi\|_{X}, \mathcal{F}(\sigma, \omega), \|\sigma\|_{S}, \|\omega\|_{S}),
\]
for all \( \sigma, \omega \in S \) such that \( \mathcal{F}(\sigma, \omega) < \infty \) and all \( \Phi \in X \).

Note that to check \( \mathcal{G} \) maps bounded sets to bounded sets, it suffices to show for instance that \( \mathcal{G} \) is continuous. Typically, we will take \( \mathcal{F} \) to be an operator norm on weighted spaces, though one may also take \( \mathcal{F} \) to be a common bump condition associated to the operator. Also note that in the case of \( S = Si_{PB} \), the fact that a function \( \mathcal{F} \) is \( X \)-stable means it is stable independent of any notion of 'size' of the measures \( \sigma \) and \( \omega \) being considered.

A simple example of a biLipschitz stable function on the class \( Si_{PB} \) is the classical two weight \( A_{2} \) characteristic for a pair of measures, namely
\[
\mathcal{F}(\sigma, \omega) = A_{2}(\sigma, \omega) = \sup_{\text{cubes } Q \text{ in } \mathbb{R}^{n}} \frac{|Q\sigma|}{|Q|} \frac{|Q\omega|}{|Q|}.
\]

Indeed,
\[
\frac{|Q\Phi, \sigma|}{|Q|} \frac{|Q\Phi, \omega|}{|Q|} = \frac{|\Phi^{-1}Q\sigma|}{|Q|} \frac{|\Phi^{-1}Q\omega|}{|Q|} \approx \frac{|\Phi^{-1}Q\sigma|}{|Q|} \frac{|\Phi^{-1}Q\omega|}{|Q|},
\]

since \( \Phi^{-1} \) is biLipschitz, and now observe that there is a cube \( P \) such that \( P \subset \Phi^{-1}Q \subset \rho P \) for some \( \rho > 1 \) by quasiconformality of \( \Phi \) \cite{AswMa}, Lemma 3.4.5, where \( \rho \) depends only on \( \Phi \in X_{biLip} \). Thus we have
\[
\frac{|Q\Phi, \sigma|}{|Q|} \frac{|Q\Phi, \omega|}{|Q|} \leq \frac{|\rho P\sigma|}{|\rho P|} \frac{|\rho P\omega|}{|\rho P|} \leq A_{2}(\sigma, \omega),
\]
and by taking supremums over cubes the reader can check that this gives
\[
A_{2}(\Phi, \sigma, \Phi, \omega) \leq \mathcal{G}(\|\Phi\|_{biLip}, A_{2}(\sigma, \omega), \|\sigma\|_{biLip}, \|\omega\|_{biLip}) = \mathcal{G}(\|\Phi\|_{biLip}, A_{2}(\sigma, \omega), 1, 1)
\]
for \( \mathcal{G}(w, x, y, z) = cw^{4n}x \), where \( c > 0 \) is independent of \( \Phi, \sigma \) and \( \omega \). The reader can also check that all of the usual 'bump' conditions that substitute for \( A_{2}(\sigma, \omega) \) are biLipschitz stable on any biLipschitz invariant subclass \( S \), e.g. Neugebauer’s condition,
\[
A_{2,r}(\sigma, \omega) = \sup_{\text{cubes } Q \text{ in } \mathbb{R}^{n}} \left( \frac{1}{|Q|} \int_{Q} \sigma(x)^{r} \, dx \right)^{\frac{1}{r}} \left( \frac{1}{|Q|} \int_{Q} \omega(x)^{r} \, dx \right)^{\frac{1}{r}},
\]
where \( 1 < r < \infty \) and \( \sigma, \omega \) are now absolutely continuous measures on \( \mathbb{R}^{n} \).

We mention in passing that the following form of the two weight \( A_{p} \) condition on the real line,
\[
\tilde{A}_{p}(v, w) \equiv \sup_{I \text{ an interval}} \left( \frac{1}{|I|} \int_{I} v \right)^{\frac{1}{p}} \left( \frac{1}{|I|} \int_{I} w \right)^{\frac{1}{p}} \left( \frac{1}{|I|} \int_{I} w^{p' - 1} \right)^{\frac{1}{p' - 1}},
\]
has been proved stable under an increasing homeomorphic change of variable \( \varphi \) (with both \( \varphi \) and \( \varphi^{-1} \) absolutely continuous) if and only if \( \varphi' \in A_{1}(\mathbb{R}) \), see \cite{LioNg} Corollary 4.4, but this condition is no longer equivalent to boundedness of the Hilbert transform for two weights, and moreover, the stability of \( A_{2}(v, w) \) is different from the stability of \( A_{2}(\sigma, \omega) \) considered above since composition and pushforward don’t commute, e.g. when \( p = 2, \Phi_{v} \neq (\Phi_{v}^{-1})^{-1} \) in general.
1.3.1. Main results. Our main result below on both stability and instability involves Riesz transforms and doubling measures, as well as Stein elliptic Calderón-Zygmund operators. Recall that if $K$ is a Calderón-Zygmund kernel, i.e. satisfies

\[ |K(x, y)| \leq C_{CZ} |x - y|^{-n}, \]

\[ |\nabla_x K(x, y)| + |\nabla_y K(x, y)| \lesssim C_{CZ} |x - y|^{-n-1}, \]

and if $T$ is a bounded linear operator on unweighted $L^2(\mathbb{R}^n)$, we say that $T$ is associated with the kernel $K$ if

\[ Tf(x) = \int K(x, y) f(y) dy, \quad \text{for all } x \in \mathbb{R}^n \setminus \text{supp } f, \]

and we refer to such operators as Calderón-Zygmund operators. Note in particular that a Calderón-Zygmund operator $T$ is bounded on unweighted $L^2(\mathbb{R}^n)$ by definition. Following [Ste2] (39) on page 210, we say that a Calderón-Zygmund operator $T$ is elliptic in the sense of Stein if there is a unit vector $u_0 \in \mathbb{R}^n$ and a constant $c > 0$ such that

\[ |K(x, x + tu_0)| \geq c |t|^{-n}, \quad \text{for all } t \in \mathbb{R}, \]

where $K(x, y)$ is the kernel of $T$.

**Theorem 4.** The two weight operator norms for individual Riesz transforms $R_i$, and more generally any Stein elliptic Calderón-Zygmund operator, are biLipschitz stable on $S_{\text{doub}}$. The individual Riesz transforms, as well as iterated Riesz transforms of odd order, are not even rotationally stable on $S_{\text{doub}}$, and even when the measures have doubling constants $C_{\text{doub}}$ arbitrarily close to $2^n$.

In fact, we can prove the following stronger rotational instability for iterated Riesz transforms of odd order, which in particular shows that instability can hold for rotations arbitrarily close to the identity.

**Theorem 5.** Iterated Riesz transforms of odd order are unstable on $S_{\text{doub}}$ under a set of rotations having full measure.

In contrast to the instability assertions in these theorems, most positive operators, such as maximal functions and fractional integral operators, are easily seen to be biLipschitz stable on $S_{A_p}, S_{A_{\infty}}, S_{\text{doub}}$ and $S_{\text{plB}}$.

For example, if $T = I_\alpha$ is the fractional integral of order $0 < \alpha < n$, and if $\Phi: \mathbb{R}^n \to \mathbb{R}^n$ is biLipschitz, then

\[
\|T_{\Phi, \sigma} f\|_{L^2(\Phi, \omega)}^2 = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |x - y|^{\alpha - n} f(y) d\Phi_\sigma(y) d\Phi_\omega(x)
\]

\[= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\Phi^{-1}x - \Phi^{-1}y|^{\alpha - n} f(\Phi^{-1}y) d\sigma(y) d\omega(x)
\]

\[\approx \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |x - y|^{\alpha - n} (f \circ \Phi^{-1})(y) d\sigma(y) d\omega(x) = \|T_\sigma (f \circ \Phi^{-1})\|_{L^2(\omega)}^2\]

and

\[\|f\|_{L^2(\Phi, \sigma)} = \int_{\mathbb{R}^n} |f(y)|^2 d\Phi_\sigma(y) = \int_{\mathbb{R}^n} |f(\Phi^{-1}y)|^2 d\sigma(y) = \|f \circ \Phi^{-1}\|_{L^2(\sigma)}^2.\]

A similar proof holds for the case when $T$ is a fractional maximal operator of order $0 \leq \alpha < n$.

1.4. History of stability. The class of Calderón-Zygmund kernels $K(x, y)$ satisfying (1.5) has long been known to be invariant under biLipschitz change of variable $x = \Phi(u)$. For example, if $K_\Phi(u, v) = K(\Phi(u), \Phi(v))$, then the chain rule gives

\[|\nabla_u K_\Phi(u, v)| = |D\Phi(u)(\nabla_x K)(u, v)| \lesssim \|D\Phi\|_{\text{biLip}} C_{CZ} |u - v|^{-n-1} \leq \|\Phi\|_{\text{biLip}} C_{CZ} |u - v|^{-n-1}.\]

It follows that if a Calderón-Zygmund operator $T$ associated with such a kernel $K$ satisfies the two weight norm inequality (1.3), then the pullback $T_\Phi$ with kernel $K_\Phi$ is also a Calderón-Zygmund operator (by a simple change of variables using that the Jacobian of $\Phi$ is bounded between two positive constants), and satisfies the inequality (1.3) with the pair of measures $(\sigma, \omega)$ replaced by the pair of pushforwards $(\Phi_\sigma, \Phi_\omega)$. This raises the question of when $T$ itself satisfies (1.3) with the pair of pushforwards $(\Phi_\sigma, \Phi_\omega)$ when $\Phi$ is
biLipschitz. Roughly speaking, our results show that the answer is yes if the measures $\sigma, \omega$ are $A_\infty$ weights, but no in general if the measures $\sigma, \omega$ are just doubling.

In $[\text{LaSaUr}]$, it was mentioned that the two weight norm inequality for the Hilbert transform is “unstable,” in the sense that for $\omega$ equal to the Cantor measure, and $\sigma$ an appropriate choice of weighted point masses in each removed third, the norm of the operator could go from finite to infinite with just arbitrarily small perturbations of the locations of the point masses, while the $A_2$ condition remained in force. In the appendix, we use this example to show that the Hilbert transform is two weight norm unstable under biLipschitz pushforwards of arbitrary measure pairs, and this instability extends to Riesz transforms in higher dimensions in a straightforward way. Thus the Riesz transforms in higher dimensions are biLipschitz unstable on arbitrary weight pairs, something which already shows that the more familiar bump-type conditions, e.g. $[\text{Neu}, \text{Theorem } 3]$, cannot characterize the two-weight problem for Riesz transforms alone.

On the other hand, we show below that Riesz transforms are biLipschitz stable under pairs of $A_\infty$ weights. So on one hand, for pairs of arbitrary measures we have instability, and on the other hand for pairs of $A_\infty$ weights, we have stability. This begs the question, what side-conditions on the weights in our weight pairs and only if they satisfy the $A_\infty$ condition, something which already shows that the more familiar bump-type conditions, e.g. $[\text{Neu}, \text{Theorem } 3]$, cannot characterize the two-weight problem for Riesz transforms alone.

The main result of this paper is that individual Riesz transforms are biLipschitz - and even rotationally - unstable for pairs of doubling weights. This provides an operator-theoretic means of distinguishing $A_\infty$ weights from doubling weights, sharpening the result of Fefferman and Muckenhoupt.

1.4.1. Our methods and their history. In 1976, Muckenhoupt and Wheeden showed in $[\text{MuWh}]$ that the two weight norm inequality for the maximal function $M$ implies the one-tailed $A_2$ condition, and conjectured that it was sufficient. Then in 1982, the third author disproved that conjecture in $[\text{Saw1}]$ by starting with a pair of simple radially decreasing weights $V, U$ constructed by Muckenhoupt in $[\text{Muc}]$, that were essentially constant on dyadic intervals $I_k = [2^{-k-1}, 2^{-k}]$ and failed the two weight inequality for $M$. Then the weights were disarranged into weights $v, u$, i.e. dilates and translates of the weights restricted to the dyadic intervals $I_k$ were essentially redistributed onto the unit interval $[0, 1]$ according to a self-similar “transplantation” rule. The resulting weights satisfied the one-tailed $A_2$ condition on $[0, 1]$ but failed the two-weight norm inequality for $M$. However, such weights cannot be doubling, because radially decreasing weights $w$ are doubling if and only if they satisfy the $A_1$ condition, $Mw \leq Cw$. This significant obstacle remained until the pioneering work of Nazarov $[\text{Naz}]$ and $[\text{NaVo}]$, to which we now turn.

Some years later, Treil and Volberg showed in $[\text{TrVo}]$ that the two-weight norm inequality for the Hilbert transform $H$ implies the two-tailed $A_2$ condition, and Sarason conjectured it was sufficient. $[\text{HaNi}, \text{s. } 7.9]$. Shortly after that, Nazarov disproved the conjecture in $[\text{Naz}]$ (which we were unable to locate till very recently, using the references in $[\text{KaTr}]$), even using doubling weights, in a beautiful proof involving the Bellman technique and a brilliant supervisor, or remodeling, argument - see also $[\text{NaVo}]$ for the details. This use of doubling weights here turns out to be crucial for our purposes. More specifically, Nazarov’s method consisted of first using the Bellman technique in a delicate argument to construct a weight pair $(V, U)$ on $\mathbb{T}$ that failed to satisfy the two weight inequality for the discrete Hilbert transform, but satisfied both dyadic doubling, with constant arbitrarily close to that of Lebesgue, and dyadic $A_2$. Then he transplanted highly oscillating functions according to a certain self-similar ‘supervisor’ rule having roots in $[\text{Bo}]$, that resulted in a pair of weights $(v, u)$ on $\mathbb{T}$ that satisfied the two-tailed $A_2$ condition, with doubling constant arbitrarily close to that of Lebesgue measure, and for which the testing condition was increasingly unbounded. Nazarov’s argument requires the clever use of highly oscillatory functions in order to deal with the singularity of the Hilbert transform, and the use of holomorphic function theory to prove weak convergence results for these increasingly oscillatory functions.

Very recently, it has come to our attention that Kakaroumpas and Treil extended Nazarov’s results to $p \neq 2$ using a non-Bellman and ‘remodeling’ construction $[\text{KaTr}]$. More precisely, Kakaroumpas and Treil first began with a pair of discretized weights with $A_p$ condition under control, a bilinear form involving the Haar shift having increasingly large norm, but doubling constant just as large. They then apply an

---

2 The reader can easily check that for a discretized version of these weights, the dyadic square function defined in Section 2 also has infinite two-weight norm.
iterative disarrangement of these weights to then construct new weights for which the $A_p$ condition and the norm of the bilinear form remain essentially unchanged, but the dyadic doubling constant of the weights is much closer to that of Lebesgue measure. This clever disarrangement is the innovative idea which replaces Nazarov’s Bellman construction, and provides weights for which one can compute explicit quantities. It is possible that our Riesz transform results can be proved using the Haar shift scheme of Kakaroumpas and Treil in place of the square function scheme of Nazarov, but we have not checked the details.

Note that the rotational stability problem is only significant in higher dimensions since in one dimension the only rotation is reflection about the origin, and that preserves the Hilbert transform. Our proof of rotational instability in higher dimensions begins by using the Bellman construction in \cite{NaVo}, and is then inspired by Nazarov and Volberg’s supervisor argument with highly oscillatory functions. In particular, we extend Nazarov’s supervisor/remodelling construction to higher dimensions, which we call “transplantation”, and which makes explicit where one extend Nazarov’s supervisor/remodelling construction to higher dimensions, which we call “transplantation”, and which makes explicit where one

Remark 6. In our construction, we show that a given iterated Riesz transform $T_0$ of order $N = 2m + 1$ fails one of the testing conditions, while all other iterated Riesz transforms $T$ of order $N = 2m + 1$ satisfy both testing conditions. Thus at this point, we have doubling measures satisfying the $A_2$ condition and both testing conditions for $T$. We now need to conclude that $T$ is two weight bounded. If the doubling constants are sufficiently small, then the $A_2$ condition implies the classical energy condition \cite{Gri}, and so one can apply either of the theorems in \cite{SaShUr7} and \cite{AlSaUr} (the main result in \cite{LaVi} can also be used for Riesz transforms of order $N = 1$). However, our construction can be slightly modified, as detailed in Remark 13, to yield pairs of doubling measures with arbitrarily large doubling constants that satisfy $A_2$ and both testing conditions for $T$, while failing the testing conditions for $T_0$. In order to show that instability continues to hold even when the doubling weights are permitted to have large doubling constant, we require the $T_1$ theorem in \cite{AlSaUr}.

1.5. Proof of Stability. We present here a simple proof of stability in Theorem 4 using a few classical facts on weights. Lemma 8 below is due to Hytonen and Lacey \cite{HyLa}, where they also prove a sharp dependence on the characteristics using much deeper tools. We begin with the following lemma of Neugebauer.

Lemma 7 (\cite[Theorem 3]{Neu}). Let $(u, v)$ be a pair of nonnegative functions. Then there exists $w \in A_p$ with $c_1 u \leq w \leq c_2 v$ if and only if there is $r > 1$ such that

$$\sup_Q \left( \frac{1}{|Q|} \int_Q u^r \right) \left( \frac{1}{|Q|} \int_Q v^{r(1-p')} \right)^{p-1} < \infty.$$  

Lemma 8 (\cite[see 1.2 Theorem]{HyLa}). Suppose that $T$ is a Stein elliptic Calderón-Zygmund operator, and that both $\omega$ and $\sigma$ are $A_\infty$ weights. Then $T$ satisfies the two weight norm inequality

$$\|T_p f\|_{L^2(\omega)}^2 \leq C \|f\|_{L^2(\sigma)}^2,$$

if and only if $A_2 (\sigma, \omega) < \infty.$

Proof. Since $T$ is Stein elliptic, a necessary condition for the two weight norm inequality is that $A_2 (\sigma, \omega) < \infty$ \cite[Lemma 10 for doubling measures]{AlSaUr}, so we now turn to proving sufficiency. But $\sigma, \omega \in A_\infty$ implies that both weights satisfy reverse Hölder conditions for some $r > 1$, and so

$$A_{2,r} (\sigma, \omega) \equiv \sup_Q \left( \frac{1}{|Q|} \int_Q \omega^r \right)^{1/r} \left( \frac{1}{|Q|} \int_Q \sigma^{r'} \right)^{1/r'} \leq \sup_Q \left( \frac{1}{|Q|} \int_Q \omega \right) \left( \frac{1}{|Q|} \int_Q \sigma \right) = A_2 (\sigma, \omega).$$

Now we apply Neugebauer’s lemma with $p = 2$ to the weight pair $(u, v) = (\omega, \sigma^{-1})$ to obtain that there exists $W \in A_2$ with $c_1 \omega (x) \leq W (x) \leq c_2 \sigma (x)^{-1}$. Then the extension of the weighted inequality of Coifman
and Fefferman [CoFe] for Calderón-Zygmund operators given in [Ste2] 6.13 on page 221, shows that

$$\|Tf\|_{L^2(\omega)}^2 \leq \|Tf\|_{L^2(W)}^2 \leq C \|f\|_{L^2(W)}^2 \leq C \|f\|_{L^2(\sigma^{-1})}^2,$$

i.e., $$\|T_\sigma f\|_{L^2(\omega)}^2 \leq C \|f\|_{L^2(\sigma)}^2$$

for all Calderón-Zygmund operators $T_\sigma$. \hfill \square

Proof of stability in Theorem 4. In particular, in the context above, suppose the norm inequality $$\|T_\sigma f\|_{L^2(\omega)}^2 \leq \mathcal{N}_T (\sigma, \omega) \|f\|_{L^2(\sigma)}^2$$ holds for a Calderón-Zygmund operator $T$ associated with a kernel $K$, and a pair of $A_\infty$ weights $(\sigma, \omega)$. Since (1.7) implies the biLipschitz stability of $A_2 (\sigma, \omega)$, and since the $A_\infty$-characteristics $[\sigma]_{A_\infty}$ and $[\omega]_{A_\infty}$ are easily seen to be biLipschitz stable as well (in fact they are stable under the more general class of quasiconformal change of variables [Uch, Theorem 2]), we conclude that the norm inequality also holds for the Calderón-Zygmund operator $T_\Phi$ with kernel

$$K_\Phi (x, y) \equiv K (\Phi (x), \Phi (y)).$$

As mentioned at the beginning of Subsection 1.4, $T_\Phi$ is a Calderón-Zygmund operator whenever $T$ is, i.e. satisfies (1.8) and is bounded on unweighted $L^2 (\mathbb{R}^n)$. Thus we conclude that $T$ is bounded on the weight pair $(\Phi \sigma, \Phi \omega)$.

We can also be more precise in our proof of stability, since [HNLa, Theorem 1.2] implies that the function

$$G(w, x, y, z) \equiv C w^{\alpha X} x (y^{\beta X} + z^{\beta X})$$

satisfies (1.6) for the functional $F = \mathcal{N}_T (\sigma, \omega)$, where $\alpha_X$ and $\beta_X$ are appropriately chosen constants. \hfill \square

The proof of instability in Theorem 4 is much more complicated, and treated throughout Section 5 with Sections 2-4 containing the necessary exposition and lemmas for the proof. In Section 6 we then extend our results to show that individual iterated Riesz transforms of odd order are rotationally unstable.

1.6. Open Problems. The question of stability of operator norms for singular integrals on weighted spaces is in general wide open. Here are four instances that might be more accessible.

1. (1) Theorem 4 shows biLipschitz stability of the norm inequality for iterated Riesz transforms of odd order over pairs of $A_\infty$ weights. This begs the question of characterizing the homeomorphisms for which the norm inequalities for iterated Riesz transforms, and even more general Calderón-Zygmund operators, are stable.

2. Only iterated Riesz transforms of odd order are treated in Theorem 4. Are Riesz transforms of even order, such as the real and imaginary parts of the Beurling transform, stable under rotations, or more generally biLipschitz pushforwards?

3. While the individual Riesz transforms $R_\ell$ are unstable under rotations of $\mathbb{R}^n$, the vector Riesz transform $R = (R_1, R_2, \ldots, R_n)$ is clearly rotationally stable since it is invariant under rotations. In fact, as mentioned at the beginning of the paper, Dąbrowski and Tolsa [DaTo] have demonstrated biLipschitz stability in the Ahlfors-David one weight setting for the $1$-fractional vector Riesz transform $R^{1-n}$. This motivates the question of whether or not the vector Riesz transform $R$ of fractional order $0$ is biLipschitz stable in the two weight setting.

2. Preliminaries: grids, doubling, telescoping identities and dyadic testing

We begin by introducing some notation, Haar bases and the telescoping identity. Then we recall the beautiful Bellman construction used in [NaVo] to obtain the dyadic weights $V, U$.

2.1. Notation for grids and cubes. We let $\mathcal{D} (J)$ denote the collection of dyadic subcubes of $J$, and for each $m \geq 0$ let $\mathcal{D}_m (J)$ denote the dyadic subcubes $I$ of $J$ satisfying $\ell(I) = 2^{-m} \ell(J)$. Let $\mathcal{P} (J)$ denote the collection of subcubes of $J$ with sides parallel to the coordinate axes, and $\mathcal{P}^0 \equiv \mathcal{P} ([0, 1]^n)$; unless otherwise specified, any cube mentioned in this paper is assumed to be axis-parallel. We also define $\mathcal{D}^0 \equiv \mathcal{D} ([0, 1]^n)$.

Given a cube $I \subset \mathbb{R}^n$, we will use the notational convention

$$I = I_1 \times I_2 \times \ldots \times I_n.$$ 

\footnote{see 6.13 on page 221 of [Ste2] for definitions}
Given an interval $I \subset \mathbb{R}$, let $I_{-}$ denote the left half and $I_{+}$ denote the right half. More generally, given a cube $I \subset \mathbb{R}^n$, let $I_{\pm} \equiv (I_{1})_{\pm} \times I_{2} \times \ldots \times I_{n}$.

Given a cube $I \subset \mathbb{R}^n$, we let $C^{(k)}(I)$ denote the $k$th generation dyadic grandchildren of $I$, and $C(I) \equiv C^{(1)}(I)$.

Given a dyadic grid $D$ and a cube $I$ in the grid, we let $\pi_D I$ denote the parent of $I$ in $D$. The same notation extends to arbitrary grids $K$, like in Section 3, where $\pi_K I$ denote the $K$-parent of $I$.

It will also be useful to keep track of the location of the children of $I$. In $\mathbb{R}^n$, let $\Theta$ denote the $2^n$ locations a dyadic child cube can be relative to its parent. For instance, when $n=2$ we can take $\Theta \equiv \{\text{NW, NE, SW, SE}\}$ the set of four locations of a dyadic square $Q$ within its $D$-parent $\pi_D Q$, where NW stands for Northwest, SE denotes Northeast, etc. Given a cube $I$ and $\theta \in \Theta$, we adopt the notation that $I_\theta$ denotes the dyadic child of $I$ at location $\theta$.

As usual we let $|J|_\mu \equiv \int_J d\mu$ for any positive Borel measure in $\mathbb{R}^n$, and $E_{J,\mu} \equiv \frac{1}{|J|} \int_J d\mu$. Given a locally integrable function $U$ in $\mathbb{R}^n$, we often abbreviate the absolutely continuous measure $U(x) \, dx$ by $U$ as well.

2.2. Doubling. We say that two distinct cubes $Q_1$ and $Q_2$ in $\mathbb{R}^n$ are adjacent if there exists a cube $Q$ for which $Q_1$ and $Q_2$ are dyadic children of $Q$.

**Definition 9.** Recall a measure $\mu$ on $\mathbb{R}^n$ is doubling if there exists a constant $C$ such that

$$\mu(2Q) \leq C \mu(Q) \quad \text{for all cubes } Q.$$  

The smallest such $C$ is called the doubling constant for $\mu$, denoted $C_{\text{doub}}$.

Equivalently, a doubling measure $\mu$ has doubling exponent $\theta > 0$ and a positive constant $c$ that satisfy the condition

$$2^{-j} |Q|_\mu \geq c 2^{-j\theta} |Q|_\mu, \quad \text{for all } j \in \mathbb{N}.$$  

The best such $\theta$ is denoted $\theta_{\text{doub}}$.

Equivalently, if $\mu$ is a doubling measure, then there exists $\lambda \geq 1$ such that for any two dyadic children $I_1$ and $I_2$ of an arbitrary cube $I$

$$\frac{E_{I_1,\mu}}{E_{I_2,\mu}} \in (\lambda^{-1}, \lambda).$$

The smallest such $\lambda$, denoted $\lambda_{\text{adj}}$, is referred to as the doubling ratio or adjacency constant of $\mu$, given by

$$\lambda_{\text{adj}}(\mu) \equiv \sup_{I, J \text{ adjacent}} \max \left\{ \frac{|I|_\mu}{|J|_\mu}, \frac{|J|_\mu}{|I|_\mu} \right\}.$$  

For a fixed cube $Q$ we define the relative doubling constant by

$$\lambda_{\text{adj};Q}(\mu) \equiv \sup_{I, J \text{ adjacent} \subset Q} \max \left\{ \frac{|I|_\mu}{|J|_\mu}, \frac{|J|_\mu}{|I|_\mu} \right\}.$$  

One may also consider the dyadic adjacency constant $\lambda_{\text{adj}}^{\text{dyad}}$ for a measure $\mu$, which is defined as above except that we only consider dyadic cubes $I_1, I_2$ with respect to a fixed grid $D$, which will be clear from context. Similarly for the relative constants $\lambda_{\text{adj};Q}^{\text{dyad}}$, where $I_1, I_2$ are taken to be dyadic subcubes of $Q$.

Given $\tau \in (0, 1)$, we say a doubling measure $\mu$ is $\tau$-flat if its adjacency constant $\lambda$ satisfies $\lambda, \lambda^{-1} \in (1-\tau, 1+\tau)$. One can make a similar definition in the dyadic setting.

For a doubling measure $\mu$ on $\mathbb{R}^n$, $C_{\text{doub}} \rightarrow 2^n$ if the doubling ratio of $\mu$ is $1 + o(1)$.

One can make similar definitions replacing $\mathbb{R}^n$ by an open subset, and modifying the definitions accordingly.

2.3. Telescoping identity. We begin by discussing the telescoping identity in the plane where matters can easily be made more explicit. For each square $Q$ in the plane define the 1-dimensional projection $E_Q$ by

$$E_Q f \equiv (E_Q f) 1_Q$$
where \( E_Q f \equiv \frac{1}{|Q|} \int_Q f \) is the average of \( f \) on \( Q \). Denote the four dyadic children of a square \( Q \) in the plane by \( Q_{NW}, Q_{NE}, Q_{SW}, Q_{SE} \) where NW stands for the northwest child, etc. Then define an orthonormal Haar basis \( \{ h_Q^{\text{horizontal}}, h_Q^{\text{vertical}}, h_Q^{\text{checkerboard}} \} \) associated with \( Q \) by

\[
\begin{align*}
\sqrt{|Q|} h_Q^{\text{horizontal}} &= +1_{Q_{NW}} - 1_{Q_{SW}} + 1_{Q_{SE}} \equiv s_Q^{\text{horizontal}}, \\
\sqrt{|Q|} h_Q^{\text{vertical}} &= -1_{Q_{NW}} + 1_{Q_{SE}} \equiv s_Q^{\text{vertical}}, \\
\sqrt{|Q|} h_Q^{\text{checkerboard}} &= +1_{Q_{NW}} - 1_{Q_{NE}} - 1_{Q_{SW}} + 1_{Q_{SE}} \equiv s_Q^{\text{checkerboard}},
\end{align*}
\]

where we associate the three matrices \([+ -], [- +], [+ +] \) with \( h_Q^{\text{horizontal}}, h_Q^{\text{vertical}}, h_Q^{\text{checkerboard}} \) respectively. We will also refer to these three matrices as the horizontal matrix, vertical matrix and checkerboard matrix respectively. Let \( \triangle_Q \) denote Haar projection onto the 3-dimensional space of functions that are constant on children of \( Q \), and that also have mean zero. Then we have the linear algebra formula,

\[
\triangle_Q f = (E_{Q_{NW}} f + E_{Q_{NE}} f + E_{Q_{SW}} f + E_{Q_{SE}} f) - E_Q f = \langle f, h_Q^{\text{horizontal}} \rangle_{h_Q^{\text{horizontal}}} + \langle f, h_Q^{\text{vertical}} \rangle_{h_Q^{\text{vertical}}} + \langle f, h_Q^{\text{checkerboard}} \rangle_{h_Q^{\text{checkerboard}}} h_Q^{\text{checkerboard}}
\]

\[
= \triangle_Q^{\text{horizontal}} f + \triangle_Q^{\text{vertical}} f + \triangle_Q^{\text{checkerboard}} f,
\]

where \( \triangle_Q^{\text{horizontal}} f \) is the rank one projection \( \langle f, h_Q^{\text{horizontal}} \rangle_{h_Q^{\text{horizontal}}} h_Q^{\text{horizontal}} \), etc.

Now given two cubes \( P \) and \( Q \) in \( D(P) \) with \( Q \subsetneq P \), define

\[
\{Q, P\} \equiv \{ I \in D(P) : Q \subsetneq I \subset P \}
\]

to be the tower of cubes from \( Q \) to \( P \) that includes \( P \) but not \( Q \). Similarly define the towers \( (Q, P), [Q, P], (Q, P) \).

Then, for \( (Q, P) \), we have the telescoping identity,

\[
(E_Q f - E_P f) 1_Q = \left( \sum_{I \in (Q, P)} \triangle_I f \right) 1_Q = \left( \sum_{I \in (Q, P)} \langle f, h_I^{\text{horizontal}} \rangle_{h_I^{\text{horizontal}}} \right) 1_Q + \left( \sum_{I \in (Q, P)} \langle f, h_I^{\text{vertical}} \rangle_{h_I^{\text{vertical}}} \right) 1_Q + \left( \sum_{I \in (Q, P)} \langle f, h_I^{\text{checkerboard}} \rangle_{h_I^{\text{checkerboard}}} \right) 1_Q
\]

\[
= \left( \sum_{I \in (Q, P)} \triangle_I^{\text{horizontal}} f \right) 1_Q + \left( \sum_{I \in (Q, P)} \triangle_I^{\text{vertical}} f \right) 1_Q + \left( \sum_{I \in (Q, P)} \triangle_I^{\text{checkerboard}} f \right) 1_Q.
\]

Turning now to dimension \( n \), we note that a similar telescoping identity holds in \( \mathbb{R}^n \). In particular, given a cube \( Q \subset \mathbb{R}^n \), if we let \( \triangle_Q \) denote the Haar projection onto the space of functions constant on the dyadic children of \( Q \) with mean 0, then

\[
\triangle_Q f = \sum_{j=1}^{d(n)} \langle f, h_Q^j \rangle h_Q^j \equiv \sum_{j=1}^{d(n)} \triangle_Q^j f,
\]

where \( \{ h_Q^j \}_{j=1}^{d(n)} \) is a choice of \( L^2(Q) \) orthonormal basis for the range of \( \triangle_Q \), and \( d(n) = 2^n - 1 \) is the dimension of this space. One of course has an analogue to the telescoping identity above. In our applications for \( n \geq 2 \), we will be interested in the case that \( h_Q^j = h_Q^{\text{horizontal}}^j \), where for \( Q = Q_1 \times \ldots \times Q_n \) we define

\[
\sqrt{|Q|} h_Q^{\text{horizontal}}(x) = \begin{cases} 
1 & \text{if } x \in Q_- \\
-1 & \text{if } x \in Q_+
\end{cases}
\]

We will not care about the choice of \( h_Q^2, h_Q^3, \ldots, h_Q^{d(n)} \) for each cube \( Q \), although we could simply take the \( h_Q^j \) to be products of one-dimensional Haar and indicator functions in each variable separately (leaving out the constant function).
2.4. **Dyadic testing.** Given weights \( V, U \) on a cube \( J \) define

\[
\gamma_{\text{horizontal}}(V, U; J) = \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} \| \Delta_{I, \text{horizontal}} V \|_{L^2(I)}^2 E_I U = \frac{1}{|J|} \sum_{I \in \mathcal{D}(J)} |\langle V, h_{I, \text{horizontal}}^1 \rangle|^2 E_I U .
\]

If \( \mathcal{D} \) is the dyadic grid, define the dyadic horizontal testing constant

\[
\mathfrak{T}_{\text{horizontal}}(V, U) \equiv \sup_{I \in \mathcal{D}} \frac{\gamma_{\text{horizontal}}(V, U; J)}{E_I V}.
\]

**Remark 10.** \( \mathfrak{T}_{\text{horizontal}}(V, U) \) is the \( L^2(\mathbb{V}) \to L^2(\mathbb{U}) \) testing condition for the ‘localized’ horizontal dyadic square function

\[
S_J^{\text{horizontal}} f(x) \equiv \sum_{I \in \mathcal{D}(J); \ x \in I} \frac{\| \Delta_{I, \text{horizontal}} f \|}{|I|}^2 = \sum_{I \in \mathcal{D}(J)} \frac{\| \Delta_{I, \text{horizontal}} f \|}{|I|}^2 \frac{1_I(x)}{|I|}.
\]

Indeed, we compute

\[
\int_J |S_J^{\text{horizontal}} (1_J V)(x)|^2 U(x) \, dx = \int_J \sum_{I \in \mathcal{D}(J)} \| \Delta_{I, \text{horizontal}} (1_J V) \|^2 U(x) \frac{1_I(x)}{|I|} \, dx = \sum_{I \in \mathcal{D}(J); I \subset J} \| \Delta_{I, \text{horizontal}} (1_J V) \|^2 E_I U,
\]

and so the norm squared of the dyadic testing condition for the localized horizontal square function is

\[
\sup_{j \in \mathcal{D}} \int_J |S_J^{\text{horizontal}} (1_J V)(x)|^2 U(x) \, dx = \sup_{j \in \mathcal{D}} \frac{\gamma_{\text{horizontal}}(V, U, J)}{E_J V}.
\]

Similarly when in the plane \( \mathbb{R}^2 \), we define \( \gamma_{\text{vertical}}(V, U; J), \gamma_{\text{checkerboard}}(V, U; J) \) and \( \mathfrak{T}_{\text{vertical}}(V, U), \mathfrak{T}_{\text{checkerboard}}(V, U) \). More generally in \( \mathbb{R}^n \), given a collection of Haar bases \( \{ (h_J^1, h_J^2, \ldots, h_J^{d_m}) \}_{j \in \mathcal{D}_0} \) (one \( L^2(J) \)-basis for each cube \( J \)), one can similarly define \( \gamma^1(V, U; J) \) and \( \mathfrak{T}^1(V, U) \). Of course these quantities are more meaningful for certain choices of \( \{ h_J^j \}_{j \in \mathcal{D}} \) than others.

2.5. **The Bellman construction of the dyadic weights.**

**Definition 11.** Given weights \( V, U \) on a cube \( J \) in \( \mathbb{R}^d \), we define the dyadic \( A_2^* \) constant relative to \( J \) by

\[
A_2^*(V, U; J) \equiv \sup_{I \in \mathcal{D}(J)} (E_I U) (E_I V).
\]

Following the Bellman construction used in [NaVo] gives the following key result.\(^4\)

**Theorem 12.** Given a cube \( J \) in \( \mathbb{R}^n \) and arbitrary constants \( \Gamma > 0, \tau \in (0, 1) \), there exist \( \tau \)-flat weights \( V, U \) on \( J \), with \( V, U \) constant on all cubes \( I \in \mathcal{D}(J) \) for some \( m > 0 \), such that

\[
A_2^*(V, U; J) \leq 1, \quad \gamma_{\text{horizontal}}(V, U; J) \geq \Gamma (E_J V),
\]

and for each \( I \in \mathcal{D}(J) \), if \( h_I \) is a function with mean 0 supported on \( I \), constant on the children of \( I \), and orthogonal to \( h_J^{\text{horizontal}} \) in \( L^2(I) \), then

\[
\langle U, h_I \rangle = \langle V, h_I \rangle = 0 .
\]

In particular when \( n = 2 \), the last conclusion implies

\[
\Delta_{I, \text{vertical}} V = 0, \quad \Delta_{I, \text{checkerboard}} V = 0.
\]

**Proof.** First note that the arbitrary dimension \( n \) case follows from the dimension \( n = 1 \) case: for instance, we show the \( n = 1 \) case implies \( n = 2 \) case, as the general case will be similar. Let \( J = J_1 \times J_2 \) be a square. So given parameters \( \Gamma \) and \( \tau \), suppose our 1-dimensional Theorem gives us weights \( (V_0, U_0) \) defined on \( J_1 \). Then define \( U \) by

\[
U(x_1, x_2) \equiv 1_{J_1}(x_2) U_0(x_1),
\]

and similarly for \( V \). Then note that

\[
E_I U = E_{I_1} U_0, \quad E_I V = E_{I_1} V_0, \quad \text{for } I \in \mathcal{D}(J).
\]

\(^4\)A simpler Bellman proof is provided in [Na2], one can also likely obtain the key result by using the disarrangement argument of [Ka12].
Since $U_0, V_0$ are $\tau$-flat and $A_2^{\text{dyadic}}(V_0, U_0; J_1) \leq 1$, then the above equation shows the same must be true of $V, U$ on $J$.

Then 2-dimensional testing is given by

$$
\gamma_{\text{horizontal}}(V, U; J) \approx \sum_{I \in D_k(J)} \frac{|I|}{|J|} (E_{I_{SW}} V + E_{I_{SE}} V - E_{I_{SE}} V - E_{I_{NE}} V)^2 E_I U
$$

$$
= \sum_{k=0}^{\infty} \sum_{J \in D_k(J)} 2^{-2k} (E_{I_{NW}} V + E_{I_{SW}} V - E_{I_{NE}} V - E_{I_{SE}} V)^2 E_I U
$$

$$
\approx \sum_{k=0}^{\infty} \sum_{K \in D_k(J_1)} \sum_{I \in D_k(J): I_i = K} 2^{-k} (E_{K-} V_0 - E_{K+} V_0)^2 E_K U_0
$$

$$
= \sum_{k=0}^{\infty} \sum_{K \in D_k(J_1)} |K| (E_{K-} V_0 - E_{K+} V_0)^2 E_K U_0
$$

$$
\approx \gamma_{\text{horizontal}}(V_0, U_0; J_1),
$$

which is at least $\Gamma(E_{J} V_0) = \Gamma(E_{J} V)$, which yields the first conclusions after relabeling $\Gamma$.

Now let $h_I$ be specified as in the theorem statement. Since $h_I$ is piecewise constant on the dyadic children of $I$, we may write

$$
\langle U, h_I \rangle = \int J U(x) h_I(x) dx
$$

$$
= E_{I_{SW}} U \int_{I_{NW}} h_I(x) dx + E_{I_{SW}} U \int_{I_{SW}} h_I(x) dx
$$

$$
+ E_{I_{SE}} U \int_{I_{SE}} h_I(x) dx + E_{I_{SE}} U \int_{I_{SE}} h_I(x) dx.
$$

Substituting averages of $U$ for averages of $U_0$, taking $a \equiv E_{(I_1)_-} U_0$ and $b \equiv E_{(I_1)_+} U_0$ for convenience, we get that this equals

$$
a \int_I h_I(x) dx + b \int_I h_I(x) dx = \frac{a + b}{2} \int_I h_I(x) dx + \frac{b - a}{2} \left( \int_I h_I(x) dx - \int_I h_I(x) dx \right).
$$

Since $h_I$ has mean 0, the first term on the right hand side of the above vanishes. Since $\langle h_I, h_I^{\text{horizontal}} \rangle = 0$, then the last term vanishes too, and thus $\langle U, h_I \rangle = 0$. Similarly for $V$. This completes the reduction to dimension 1.

The dimension $n = 1$ case follows from Nazarov’s Bellman argument in [Naz]. See also [NaVo] Section 3 for a stronger conclusion not used here. However, the argument there involves more difficult Hessian computations, and also requires an argument to show that their set of admissible weight pairs $\mathcal{F}_\epsilon$ is nonempty [AiLuSaUr], Lemma 12].

We will now adapt the supervisor argument of Nazarov to construct a pair of doubling weights $(v, u)$, first on a cube in $\mathbb{R}^n$ and eventually on the whole space $\mathbb{R}^n$, satisfying $A_2(v, u) \leq 1$ and such that the first Riesz transform $R_1$ has operator norm $\mathfrak{N}_{R_1}(v, u) > \Gamma$, while the other Riesz transforms $R_j$, $j \geq 2$, have operator norm $\mathfrak{N}_{R_j}(v, u) \leq 1$. Thus the individual Riesz transform $R_1$ is not stable under rotations of doubling weights in the plane. We will view the supervisor map more simply as a transplantation map, that readily exploits telescoping properties of projections.

Note that if $V$ and $U$ are $\tau$-flat for $\tau$ sufficiently small, then the classical pivotal condition holds [Gri], and we can apply the $T1$ theorem in [SaShUr7] in order to deduce $\mathfrak{N}_{R_1}(v, u) \leq 1$ from the testing conditions.
The same will apply to any more general operators in place of the Riesz transforms. Thus we need not use
our recent doubling theorem for arbitrary smooth operators in [AlSaUr].

Remark 13. A simple modification of our Bellman construction yields weights $V, U$ which have doubling
constants arbitrarily large. Indeed, given $2^\Gamma > 0$, define weights $V, U$ on $[0, 1]$ as follows: on $[0, \frac{1}{2}]$, take
$(V, U)$ as generated by Theorem 12 and take $b = E_{[0, \frac{1}{2}]}V$ and $a = E_{[0, \frac{1}{2}]}U$. Then on $[\frac{1}{2}, 1]$, define
$$U(x) = Ma, \quad V(x) = Mb.$$ Then by taking $M$ sufficiently large, we get $V, U$ will have dyadic adjacency ratio arbitrarily large, and hence
dyadic doubling constant arbitrarily large. In our subsequent arguments, this will yield weights with doubling
constant arbitrarily large while keeping the testing constant for $R_2$ bounded, with bound possibly depending on $M$,
while making the testing constant for $R_1$ be at least $\Gamma$, which can be chosen arbitrarily large independent of $M$.

3. The supervisor and transplantation map

We again begin our discussion in the plane where matters are more easily pictured. We will construct
our weight pair $(v, u)$ on a square $Q^0 \subset \mathbb{R}^2$ from the dyadic weight pair $(V, U)$ by adapting the supervisor
argument of Nazarov [NaVo] as follows. Let $\{k_1\}_{i=1}^\infty$ be an increasing sequence of positive integers to be
fixed later, and let $\mathcal{D}^0$ denote the collection of dyadic subsquares of $Q^0$. We denote by $\mathcal{K}_t = \mathcal{K}_t(Q^0)$
the collection of dyadic subsquares $Q$ of $Q^0$ in $\mathcal{D}^0$ whose side lengths satisfy $t(Q) = 2^{-k_1} \cdots 2^{-k_t}(Q^0)$, and then define
$$\mathcal{K} = \mathcal{K}(Q_0) = \bigcup_{t \in \mathbb{N}} \mathcal{K}_t(Q_0)$$
a subgrid of the dyadic grid $\mathcal{D}^0$. Recall we have $\Theta \equiv \{\text{NW}, \text{NE}, \text{SW}, \text{SE}\}$ the set of four locations of a dyadic
square $Q$ within its $D$-parent $\pi_D Q$.

3.1. The informal description of the construction. Here is an informal description of the transplantation
argument, that we will give more precisely later. Given a nonnegative integrable function $U \in L^1(Q^0)$, and $t \in \mathbb{N}$, we will define $u_t(x)$ to be a step function on $Q^0$ that is constant on each square in the $t$th level $\mathcal{K}_t$ of
and where the constants are among the expected values of $U$ on the squares in the $t$th level $\mathcal{D}_t^0$ of $\mathcal{D}^0$, but ‘scattered’ according to the following plan.

To each square $Q$ in $\mathcal{K}_t$, there is associated a unique descending ‘$K$-tower’ $T = (T_1, ..., T_t) \in \mathcal{K}^t = K \times \cdots \times K$
with $T_t = Q$, where the square $T_t$ is the unique square in $\mathcal{K}_t$ containing $Q$. To each component square $T_t$
of $T$ there is associated $\theta_t \in \Theta$, which describes the location of $T_t$ within its $D$-parent $\pi_D T_t$. Thus there
is a unique vector $\theta \in \Theta^t = \Theta \times \cdots \times \Theta$ of the locations $\theta_t$ associated to the squares $T_t$ in the tower $T$.
This vector $\theta$ then determines a unique square $S(Q)$ in $D_t$, with the property that the $D$-tower of $S(Q)$ has
the same location vector $\theta$. In the terminology of Nazarov [NaVo], $S(Q)$ is the supervisor of $Q$. We then
‘transplant’ the expected value $E_{S(Q)} U$ of $U$ on the supervisor to the cube $Q$ in $\mathcal{K}_t$ that is being supervised.
For example, when $k_t = 1$ for all $t$, this construction yields the identity
$$u_t = E_{S(Q)} U \equiv \sum_{Q \in \mathcal{D}^t(Q)} (E_Q U) 1_Q,$$ and when the $k_t$’s are bigger than 1, the values \( \frac{1}{|Q|}1_Q U \) are ‘scattered’ throughout $Q^0$. Now we give precise
details of the ‘scattering’ construction.

3.2. The formal construction. Throughout this section, we define the associated parent grid $\mathcal{P} \equiv \pi_D \mathcal{K}$
of $\mathcal{K}$ to consist of the $D$-parents of the squares in $\mathcal{K}$. Define a $K$-tower $T = (T_1, ..., T_t) \in \mathcal{K}^t \equiv \otimes^t K$ to satisfy
$T_t \in \mathcal{K}$ and $T_t \supset T_{t+1}$, and define the corresponding parent tower by
$$\mathcal{P} = (P_1, ..., P_t) \equiv \pi_D T \equiv (\pi_D T_1, ..., \pi_D T_t) \in \mathcal{P}^t \equiv \otimes^t \mathcal{P}.$$

\footnote{A simpler form of ‘disarranging’ a weight was used in [Saw1] to provide a counterexample to the conjecture of Muckenhoupt and Wheeden [MiWh, page 281] that a one-tailed $A_p$ condition was sufficient for the norm inequality for $M$, but the weights were not doubling.}
Given a \( t \)-vector of locations \( \theta = (\theta_1, ..., \theta_t) \in \Theta^t \equiv \otimes^t \Theta \), and a parent tower \( P = (P_1, ..., P_t) \in \mathcal{P}^t \), we say that \( P \) has structure \( \theta \), written \( P \in \mathcal{S}_\theta \) if

\[
\begin{align*}
P_1 &\in \mathfrak{c}_D^{(k_1-1)}(Q^0), \\
P_\ell &\in \mathfrak{c}_D^{(k_{\ell-1})}\left((P_{\ell-1})_{\theta_{\ell-1}}\right) \quad \text{for } \ell = 2, ..., t.
\end{align*}
\]

Finally, we set \( \mathcal{D}_t^0 \equiv \{ Q \in \mathcal{D}_t^0 : \ell(Q) = 2^{-t} \ell(Q^0) \} \) and \( \mathcal{K}_t \equiv \{ Q \in \mathcal{K} : \ell(Q) = 2^{-k_1-...-k_t} \ell(Q^0) \} \) to be the collection of squares at level \( t \) in each of the grids \( \mathcal{D} \) and \( \mathcal{K} \).

Let \( U \in L^1(Q^0) \) be a nonnegative integrable function, and let \( t \in \mathbb{N} \). To each square \( Q \) in \( \mathcal{K}_t \), which is the \( t^{th} \) level of \( \mathcal{K} \), there is associated a unique tower \( T = (T_1, ..., T_t) \in \mathcal{K}^t \) with \( T_t = Q \), where the square \( T_t \) is the unique square in \( \mathcal{K}_t \) containing \( Q \). To the parent tower \( P \) associated with \( T \), one can see that there is a unique vector \( \theta \in \Theta^t \) for which (3.1) is satisfied, in fact it is the vector \( \theta \) described in the informal argument above.

We then ‘transplant’ the expected value \( E_{\mathcal{S}(Q)} U \) of \( U \) on the supervisor \( \mathcal{S}(Q) \) to the cube \( Q \) in \( \mathcal{K}_t \) that is being supervised. Here are the precise formulas written out using the parent grid \( \mathcal{P} \), where for convenience we will use superscripts to track the level of a square in the grid \( \mathcal{D} \):

\[
u_0(x) = (E_{Q^0} U) \mathbf{1}_{Q^0}(x),
\]

and for \( t \geq 1,

\[
u_t(x) = \left( \sum_{\theta \in \Theta^t} \sum_{P \in \mathcal{S}_\theta} (E_{Q_{t_1, ..., t_t}^0} U) \mathbf{1}_{(P_{t_1})_{\theta_{t_1}}} \right)(x),
\]

where \( Q_{t_1, ..., t_t}^0 \) denotes the unique cube \( Q \in \mathcal{D}_t^0 \) such that the \( \mathcal{D} \)-tower of \( Q \) has location vector \( \theta = (\theta_1, ..., \theta_t) \), i.e. \( Q_{t_1, ..., t_t}^0 = \left( ... (Q^0)_{\theta_1} ... \right)_{\theta_t} \), using the notation introduced at the beginning of Section 2.

**Definition 14.** Given a cube \( Q = (P_t)_{\theta_t} \) with associated parent tower \( P \in \mathcal{S}_\theta \), where the structure vector \( \theta \in \Theta^{t+1} \), we define the supervisor \( \mathcal{S}(Q) \) of \( Q \) to be \( Q_{t_1, ..., t_t}^0 \).

Note that the supervisor map \( \mathcal{S} \) is many-to-one, indeed \( Q_{t_1, ..., t_t}^0 \) has \( C_{t-k_1-...-k_t} \) preimages under \( \mathcal{S} \). Furthermore we note that \( \mathcal{S}(\pi \mathcal{K} Q) = \pi \mathcal{D} \mathcal{S}(Q) \), i.e. \( \pi \) and \( \mathcal{S} \) commute. Recall the Haar projection \( \triangle_Q \) associated with \( Q \) satisfies

\[
\triangle_Q f \equiv \left( \sum_{Q' \in \mathcal{E}_D Q} \mathbb{E}_{Q'} f \right) - \mathbb{E}_Q f = \left( \sum_{Q' \in \mathcal{E}_D(Q)} (E_{Q'} f) \mathbf{1}_{Q'} \right) - (E_Q f) \mathbf{1}_Q.
\]

Given cubes \( Q, P \), let \( \phi_{P \rightarrow Q} \) denote the unique translation and dilation that takes \( P \) to \( Q \), and define

\[
h_Q^{\text{horizontal}}[P](x) \equiv h_Q^{\text{horizontal}}(\phi_{P \rightarrow Q}(x)).
\]

Note that this function does not have \( L^2(P) \) norm equal to 1. We can also make the same definition for \( h_Q^{\text{vertical}}[P], h_Q^{\text{checkerboard}}[P] \). Finally, define

\[
\triangle_Q [P] f(x) \equiv \left( \triangle_Q f \right) \phi_{P \rightarrow Q}(x)
\]

\[
= \langle f, h_Q^{\text{horizontal}}[P] \rangle_{h_Q^{\text{horizontal}}[P]} + \langle f, h_Q^{\text{vertical}}[P] \rangle_{h_Q^{\text{vertical}}[P]} + \langle f, h_Q^{\text{checkerboard}}[P] \rangle_{h_Q^{\text{checkerboard}}[P]}
\]

\[
\equiv \triangle_Q^{\text{horizontal}}[P] f(x) + \triangle_Q^{\text{vertical}}[P] f(x) + \triangle_Q^{\text{checkerboard}}[P] f(x).
\]
Then using (3.2) for \( t \geq 1 \), the first order differences of the weights \( u_t \) are given by

\[
u_{t+1}(x) - u_t(x)
= \sum_{\theta \in \Theta'} \sum_{P \in S_0} \left\{ \left( \sum_{\theta_{t+1} \in \Theta} \sum_{P_{t+1} \in \mathcal{E}_D^{(k_t+1)}((P_t)_{\theta_t})} \left( E_{Q_t}^0, \ldots, u_t \right) \mathbf{1}_{(P_{t+1})_{\theta_{t+1}}} (x) \right) - \left( E_{Q_t}^0, \ldots, u_t \right) \mathbf{1}_{(P_t)_{\theta_t}} (x) \right\}
= \sum_{\theta \in \Theta'} \sum_{P \in S_0} \left\{ \sum_{P_{t+1} \in \mathcal{E}_D^{(k_t+1)}((P_t)_{\theta_t})} \left( E_{Q_t}^0, \ldots, u_t \right) \mathbf{1}_{(P_{t+1})_{\theta_{t+1}}} (x) - \left( E_{Q_t}^0, \ldots, u_t \right) \mathbf{1}_{P_{t+1}} (x) \right\}
= \sum_{\theta \in \Theta'} \sum_{P \in S_0} \left\{ \sum_{P_{t+1} \in \mathcal{E}_D^{(k_t+1)}((P_t)_{\theta_t})} \left( E_{Q_t}^0, \ldots, u_t \right) \mathbf{1}_{(P_{t+1})_{\theta_{t+1}}} (x) - \left( E_{Q_t}^0, \ldots, u_t \right) \mathbf{1}_{P_{t+1}} (x) \right\}
= \sum_{\theta \in \Theta'} \sum_{P \in S_0} \left\{ \sum_{P_{t+1} \in \mathcal{E}_D^{(k_t+1)}((P_t)_{\theta_t})} \Delta_{Q_t, \ldots, \theta_{t}} [P_{t+1}] u_t (x) \right\}.
\]

Let \( \mathcal{B} \) denote a set indexing our choice of Haar basis: since we are working in dimension 2, we take

\[\mathcal{B} = \{\text{horizontal, vertical, checkerboard}\}.\]

For a square \( Q \) and an integer \( M \in \mathbb{N} \), we define three alternating functions, one for each pattern \( \in \mathcal{B} \):

\[s_{k}^{Q, \text{pattern}}(x) = \sum_{Q' \in \mathcal{E}^{(M-1)}(Q)} \sqrt{|Q'|} s_{k}^{Q', \text{pattern}}, \quad \text{pattern} \in \mathcal{B}.\]

Note that each of these three alternating functions is a constant \( \pm 1 \) on grandchildren \( P' \in \mathcal{E}^{(M)}(Q) \) of depth \( M \), and when restricted to a grandchild \( Q' \in \mathcal{E}^{(M-1)}(Q) \), each alternating function has the arrangement of \( \pm 1 \) given respectively by

\[
\begin{bmatrix}
+ & + & \ldots & + \\
+ & + & \ldots & + \\
\end{bmatrix}, \quad \begin{bmatrix}
- & - & \ldots & - \\
- & - & \ldots & - \\
\end{bmatrix}, \quad \begin{bmatrix}
+ & + & \ldots & + \\
- & - & \ldots & - \\
\end{bmatrix},
\]

which are the horizontal, vertical and checkerboard matrices respectively (horizontal, vertical and checkerboard refer to the direction of sign change). For instance, \( s_{k}^{Q, \text{horizontal}} \) is the function on \( Q \) consisting of \( \pm 1 \) arranged in the following fashion:

\[s_{k}^{Q, \text{horizontal}} \sim \text{the } 2^k \times 2^k \text{ matrix } \begin{bmatrix}
+ & + & \ldots & + \\
+ & + & \ldots & + \\
\vdots & \vdots & \ddots & \vdots \\
+ & + & \ldots & + \\
\end{bmatrix}, \]

and similarly

\[s_{k}^{Q, \text{vertical}} \sim \text{the } 2^k \times 2^k \text{ matrix } \begin{bmatrix}
+ & - & \ldots & - \\
+ & + & \ldots & - \\
\vdots & \vdots & \ddots & \vdots \\
+ & - & \ldots & - \\
\end{bmatrix} \quad \text{and} \quad s_{k}^{Q, \text{checkerboard}} \sim \text{the } 2^k \times 2^k \text{ matrix } \begin{bmatrix}
+ & - & \ldots & - \\
+ & - & \ldots & - \\
\vdots & \vdots & \ddots & \vdots \\
+ & - & \ldots & - \\
\end{bmatrix}.
\]

Remark 15. Notice that the matrix for \( s_{k}^{Q, \text{horizontal}} \) is given by transplanting \( 2^{2k-2} \) copies of the \( 2 \times 2 \) matrix \( \begin{bmatrix}
+ & - \\
+ & - \\
\end{bmatrix} \), which corresponds to the tensor product of a \( 1 \)-dimensional Haar function with matrix \( \begin{bmatrix}
+ & - \\
+ & - \\
\end{bmatrix} \), and an indicator function with matrix \( \begin{bmatrix}
+ & + \\
\end{bmatrix} \). This connects the alternating form of \( s_{k}^{Q, \text{horizontal}} \) with the testing condition for the ‘Haar square function’, as constructed in Section 3.
We now write the projections $\Delta Q U$ as a sum of the horizontal, vertical and checkerboard components,

$$
\Delta Q U = \Delta_{\text{horizontal}} Q U + \Delta_{\text{vertical}} Q U + \Delta_{\text{checkerboard}} Q U,
$$

where\(\Delta_{\text{horizontal}} Q U = \langle U, h^Q_{\text{horizontal}} \rangle h^Q_{\text{horizontal}}\) is nonnegative. Further,\(\Delta_{\text{vertical}} Q U = \langle U, h^Q_{\text{vertical}} \rangle h^Q_{\text{vertical}}\) and\(\Delta_{\text{checkerboard}} Q U = \langle U, h^Q_{\text{checkerboard}} \rangle h^Q_{\text{checkerboard}}\), to obtain for \(t \geq 1\),

$$
\begin{align*}
\sum_{\text{pattern}\in B} \sum_{\theta \in \Omega} \sum_{P \in S} \Delta_{\text{pattern}} \left\{ \sum_{P_{t+1} \in \epsilon_{D}^{(t+1-1)}((P_{t})_{\theta_{t}})} |P_{t+1}| \cdot U(x) \right\} \\
= \sum_{\text{pattern}\in B} \sum_{\theta \in \Omega} \sum_{P \in S} \left\{ \sum_{P_{t+1} \in \epsilon_{D}^{(t+1-1)}((P_{t})_{\theta_{t}})} \langle U, h^Q_{\text{pattern}} \rangle h^Q_{\text{pattern}} \left[ P_{t+1} \right] \right\} \\
= \sum_{\text{pattern}\in B} \sum_{\theta \in \Omega} \sum_{P \in S} \left\{ \sum_{P_{t+1} \in \epsilon_{D}^{(t+1-1)}((P_{t})_{\theta_{t}})} \langle U, h^Q_{\text{pattern}} \rangle h^Q_{\text{pattern}} \left[ P_{t+1} \right] \right\} \\
= \sum_{\text{pattern}\in B} \sum_{\theta \in \Omega} \sum_{P \in S} \left\{ \sum_{P_{t+1} \in \epsilon_{D}^{(t+1-1)}((P_{t})_{\theta_{t}})} \langle U, h^Q_{\text{pattern}} \rangle h^Q_{\text{pattern}} \left[ P_{t+1} \right] \right\}.
\end{align*}
$$

By reversing the roles of \(Q \in K_t\) and its supervisor \(S(Q) \in D_t\) for the squares \(P_{t+1} \in K_t\), we obtain

$$
\sum_{\text{pattern}\in B} \sum_{\theta \in \Omega} \sum_{P \in S} \left\{ \sum_{P_{t+1} \in \epsilon_{D}^{(t+1-1)}((P_{t})_{\theta_{t}})} \langle U, h^Q_{\text{pattern}} \rangle h^Q_{\text{pattern}} \left[ P_{t+1} \right] \right\}.
$$

Finally, we note that the weights \(u_{t+1}\) are nonnegative on \(Q^0\) since \(u_{t+1}\) is constant on each square \(Q\) in \(K_t\), and the value of this constant is the expectation \(E_{S(Q)} U\) on the supervisor \(S(Q)\), which is of course nonnegative.

In dimension \(n = 1\), the above transplantation construction reduces to the ‘supervisor and alternating function’ construction in Nazarov and Volberg [NaVo]. Furthermore, in the one dimensional setting, we may define \(s^Q_{\text{horizontal}}\) similarly, and since there is essentially only one choice of Haar basis in 1 dimension, \(\{\pm h^Q_{\text{horizontal}}\}\), we will use the simplified notation \(s^Q_{k} = s^Q_{k,\text{horizontal}}\) in dimension 1. We also note the following useful fact: \(|u_{t+1}|\) is bounded by a constant independent of the choice of \(\{k_{t}\}_{t \geq 0}\), namely \(\|u_{t+1}\|_{L^\infty} \leq \|U\|_{L^\infty}\) since the only values \(u_{t+1}\) can take on are precisely the expectations of \(U\) over the relevant cubes \(Q\).

Turning now to general dimension \(n\), we may define

$$
s^Q_{\text{horizontal}}(x) = \langle h_{x_{1}} \rangle 1_{Q_{2} \times \ldots \times Q_{n}}(x_{2}, \ldots, x_{n}),
$$

which is consistent with our initial definition in dimension \(n = 2\), and where the horizontal direction is the direction of sign change. All of the calculations above extend to dimension \(n\) using \(s^Q_{1,\text{horizontal}}\) as part of an otherwise arbitrarily chosen basis of Haar functions for the cube \(Q = Q_{1} \times \ldots \times Q_{n}\). For instance we could consider the ‘standard’ Haar basis \(\{g_{1} \otimes \ldots \otimes g_{n}\}\) consisting of all product functions \(g_{1}(x_{1}) \times \ldots \otimes g_{n}(x_{n})\) in which \(g_{j}\) is either the Haar function \(h_{x_{j}}\) on \(Q_{j}\), or the normalized indicator \(\frac{1}{|Q_{j}|} 1_{Q_{j}}\), and where the constant function on \(Q\) is discarded; then note that \(s^Q_{1,\text{horizontal}} = h_{x_{1}} \otimes \frac{1}{\sqrt{|Q_{2}|}} 1_{Q_{2}} \otimes \ldots \otimes \frac{1}{\sqrt{|Q_{n}|}} 1_{Q_{n}}\).

4. **Weak Convergence Properties of the Riesz Transforms**

We let \(H\) denote the Hilbert transform on \(\mathbb{R}\), i.e.

$$
Hf(x) = \frac{1}{\pi} \text{pv} \int_{\mathbb{R}} \frac{f(x - y)}{y} dy,
$$

and we let \(R_{j}\) denote the \(j\)th individual Riesz transform on \(\mathbb{R}^{n}\), i.e.

$$
(4.1) \quad R_{j}f(x) = c_{n} \text{pv} \int_{\mathbb{R}^{n}} \frac{y_{j}}{|y|^{n+1}} f(x - y) dy,
$$

where \(c_{n} = \frac{\Gamma \left( \frac{n+1}{2} \right)}{\pi^{\frac{n+1}{2}}}\).
Note that with these choices of constants, the symbols of the operators $H$ and $R_j$ are $-i \text{sgn} \xi$ and $-i \frac{\xi_j}{|\xi|}$ respectively. In what follows, all singular integrals are understood to be taken in the sense of principal values, even when we do not explicitly write $pv$ in front of the integral. If we apply the Riesz transform $R_j$ in the plane to the difference $u_{t+1} - u_t$ above we obtain

$$R_j \left( u_{t+1} - u_t \right) = \sum_{\text{pattern}\in \mathbb{B}} \sum_{Q \in K_t} \left< U, h_{Q, \text{pattern}}^\text{pattern} \right> \frac{1}{\sqrt{|S(Q)|}} R_j s_Q^Q, \text{pattern}$$

and in particular, if $\Delta_p \text{vertical} U$, $\Delta_p \text{vertical} V$, $\Delta_p \text{checkerboard} U$ and $\Delta_p \text{checkerboard} V$ vanish for all $P$, then we have both

$$R_j \left( u_{t+1} - u_t \right) = \sum_{Q \in K_t} \left< U, h_{Q, \text{horizontal}}^\text{horizontal} \right> \frac{1}{\sqrt{|S(Q)|}} R_j s_Q^Q, \text{horizontal},$$

$$R_j \left( v_{t+1} - v_t \right) = \sum_{Q \in K_t} \left< V, h_{Q, \text{horizontal}}^\text{horizontal} \right> \frac{1}{\sqrt{|S(Q)|}} R_j s_Q^Q, \text{horizontal}.$$

We now wish to establish three key testing estimates: for an arbitrarily large $\Gamma$,

1. $\sup_{Q \ni |u|} \int_Q |R_1 1_Q v|^2 u \geq \Gamma$,
2. $\sup_{Q \ni |u|} \int_Q |R_2 1_Q v|^2 u \leq 1$,
3. $\sup_{Q \ni |u|} \int_Q |R_2 1_Q u|^2 v \leq 1$.

As in [NaVo], this is accomplished by inductively choosing the rapidly increasing sequence $\{k_t\}_{t=1}^m$ of positive integers so that at each stage of the construction labelled by $t$, the discrepancy $\int |R_j \left( v_{t+1} \right)|^2 u_{t+1} - \int |R_j \left( v_t \right)|^2 u_t$ looks like $\sum_{l=1}^N \left< U, E^2\sum_{k=1}^t \Delta_k^\text{horizontal} V \right> E^2 U$, whose sum over $t$ exceeds $\Gamma$. These considerations also extend to higher dimensions. But a considerable amount of preparation is needed for this, and we begin with a discussion of the notion of weak convergence, which we use in connection with the alternating functions introduced in Section 3.

**Definition 16.** Let $1 < p < \infty$. We say $f_i \rightarrow 0$ weakly in $L^p(\mathbb{R}^n)$ if

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}^n} f_i(x) b(x) \, dx = 0,$$

for all functions $b \in L^p(\mathbb{R}^n)$.

Note that if $f_i \rightarrow 0$ weakly in $L^p(\mathbb{R}^n)$, then so does $Tf_i$ for any Calderón-Zygmund operator $T$ bounded on unweighted $L^2$. Note also that if (1.2) holds only for $b$ belonging to some dense subset $X$ of $L^p(\mathbb{R}^n)$ and if $\{f_i\}$ is uniformly bounded in $L^p(\mathbb{R}^n)$, then a density argument shows (1.2) holds for all $b \in L^p(\mathbb{R}^n)$. In particular, we will use the cases where $X$ is the space of step functions with compact support on $\mathbb{R}^n$, those that are constant on dyadic subintervals of $\mathbb{R}^n$ of fixed size, or where $X$ is the larger space $L^\infty(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$.

We now turn to some lemmas in dimension $n = 1$ that we will use for establishing the three key testing estimates listed above.

**4.1. Weak convergence properties of the Hilbert transform.** In Nazarov’s supervisor argument in [NaVo], the weak limits appearing in Lemma 18 below, for the alternating functions $s_t^1$, were proved using holomorphic function theory. The key observation was the following lemma.

**Lemma 17.** Suppose $p \in (1, \infty)$. Consider a bounded sequence $\{f_k\}$ in $L^p(\mathbb{R})$. Then $f_k \rightarrow 0$ weakly in $L^p(\mathbb{R})$ if and only if $\lim_{k \rightarrow \infty} \mathbb{P} f_k(z) = 0$ for all $z \in \mathbb{R}^2_+$, where $\mathbb{P} f_k$ is the Poisson extension of $f_k$.

**Proof.** Since $f_k \in L^p(\mathbb{R})$, then for every $z \in \mathbb{R}^2_+$ the Poisson extension formula yields

$$\mathbb{P} f_k(z) = \mathbb{P} f_k(x + iy) \equiv \int_\mathbb{R} f_k(t) \, P_{x+iy}(t) \, dt,$$

where $P_{x+iy}(t) = \frac{y}{(x-t)^2 + y^2}$ is the Poisson kernel.

If $f_k \rightarrow 0$ weakly in $L^p(\mathbb{R})$, then by the formula above we clearly have $\lim_{k \rightarrow \infty} \mathbb{P} f_k(z) = 0$ for all $z \in \mathbb{R}^2_+$. 

Proof. Let us first show (\[NaVo\) Section 4)} that mixed terms \(s_k\) weakly in \(L^p(R)\), since \(z\) for all \(0 \rightarrow 0\) goes to which gives Lemma 19.\(^6\)

\[
\text{Proof. Since } \lim_{k \rightarrow \infty} \int_R s_k^1(t) g(t) dt = 0 \text{ for all step functions } g \text{ on } R, \text{ and since finite linear combinations of step functions are dense in } L^p(R), \text{ we conclude that } s_k^1 \rightarrow 0 \text{ weakly in } L^p(R). \text{ Since } H \text{ is bounded on } L^p(R), \text{ we also have } Hs_k^1 \rightarrow 0 \text{ weakly in } L^p(R). \text{ Let } f_k^1 = s_k^1 + iHs_k^1 \in H^p(R). \text{ By an application of Lemma 17 using } f_k^1 \rightarrow 0 \text{ weakly in } L^p(R), \text{ followed by the fact that } (\overline{P f_k^1})^2 \text{ is holomorphic, and then finally writing } (f_k^1)^2 \text{ in terms of its real and imaginary parts, we get}
\]

\[
0 = \left[ \lim_{k \rightarrow \infty} \mathbb{P} f_k^1(z) \right]^2 = \lim_{k \rightarrow \infty} \mathbb{P} \left[ (f_k^1)^2 \right](z) = \lim_{k \rightarrow \infty} \mathbb{P} \left[ (s_k^1)^2 - (Hs_k^1)^2 + i2s_k^1Hs_k^1 \right](z)
\]

for all \(z \in \mathbb{R}^2\). By Lemma 17 again,

\[
s_k^1Hs_k^1 \rightarrow 0 \text{ weakly in } L^p(R), \quad 1_I - (Hs_k^1)^2 = (s_k^1)^2 - (Hs_k^1)^2 \rightarrow 0 \text{ weakly in } L^p(R),
\]

since \((s_k^1)^2 = 1_I\). Similarly, we see that the real part of \((f_k^1)^3 \rightarrow 0 \text{ weakly in } L^p(R), \text{ i.e.}
\]

\[
(s_k^1)^3 - 3(s_k^1)(Hs_k^1)^2 \rightarrow 0 \text{ weakly in } L^p(R),
\]

which gives \(s_k^1(Hs_k^1)^2 \rightarrow 0 \text{ weakly in } L^p(R)\) since \((s_k^1)^2 = 1_I\) and \(s_k^1 \rightarrow 0 \text{ weakly in } L^p(R)\).

The more general statement involving powers \(a\) and \(b\) follows similar arguments which we leave for the reader, noting that they are not used in this paper. \(\square\)

To carry out Nazarov’s supervisor argument in \(\[NaVo\), one also needs to understand the weak convergence of mixed terms \(s_k^1(Hs_k^1)(Hs_k^1)\), where \(I, J, K\) are dyadic intervals of same side length. We will often make use of the trivial observation that if \(I_1, I_2, \ldots, I_N\) are pairwise disjoint sets, and functions \(a_{I_j}^j\) are supported on \(I_j\), then \(\sum_{j=1}^N a_{I_j}^j \rightarrow 0 \text{ weakly in } L^p(R)\) if and only if \(a_{I_j}^j \rightarrow 0 \text{ weakly in } L^p(R)\) for each \(j = 1, 2, \ldots, N\).

**Lemma 19.** Suppose \(p \in (1, \infty)\). Let \(I, J, K\) be dyadic intervals all having the same side length. Then

\[
s_k^1 (Hs_k^1) \rightarrow 0 \text{ weakly in } L^p(R) \text{ as } k \rightarrow \infty, \quad (Hs_k^1)(Hs_k^1) \rightarrow 0 \text{ weakly in } L^p(R) \text{ as } k \rightarrow \infty \text{ if } I \neq J, \quad s_k^1 (Hs_k^1)(Hs_k^1) \rightarrow 0 \text{ weakly in } L^p(R) \text{ as } k \rightarrow \infty.
\]

**Proof.** Let us first show \(s_k^1Hs_k^1 \rightarrow 0 \text{ weakly in } L^p(R)\). If \(I = J\), this follows by Lemma 18 So assume \(I \neq J\).
Write \(f_k^1 = s_k^1 + iHs_k^1\), and similarly for \(J\). Since \(f_k^1f_k^1 \in H^p(R)\) (because \(H\) is bounded on \(L^2p(R)\)), the method of proof of Lemma 18 combined with Lemma 17 implies that the real and imaginary parts of \(f_k^1f_k^1\) go to 0 weakly in \(L^p(R)\). In particular since \(s_k^1s_k^1 = 0\) because of their disjoint support, we get

\[
-(Hs_k^1)(Hs_k^1) \rightarrow 0 \text{ weakly in } L^p(R), \quad s_k^1Hs_k^1 + s_k^1Hs_k^1 \rightarrow 0 \text{ weakly in } L^p(R),
\]

\(^6\text{Hint: Consider the unit circle } \mathbb{T} = \{0, 2\pi\}. \text{ Let } f \in C(\mathbb{T}) \text{ and } \epsilon > 0. \text{ For } r < 1 \text{ sufficiently close to } 1, \text{ and for } n \text{ sufficiently large depending on } r, \text{ we have}
\]

\[
\left| P_r * f(x) - \sum_{k=0}^{n-1} \left( \int_{2\pi \frac{k+1}{n}}^{2\pi \frac{k+1}{n}} f \right) P_r \left( x - \frac{2\pi k}{n} \right) \right| \leq \epsilon.
\]
which immediately proves the second line in the lemma. As for the first line, since $I, J$ are disjoint, it follows that $s_k^I H s_k^J \to 0$ weakly in $L^p(\mathbb{R})$ and $s_k^J H s_k^I \to 0$ weakly in $L^p(\mathbb{R})$. At this point we have proved the first two lines in the lemma.

Now let us show that $s_k^I (H s_k^J) (H s_k^J) \to 0$ weakly in $L^p(\mathbb{R})$. Define $f_k^I, f_k^J, f_k^K$ as above. We will expand $f_k^I f_k^J f_k^K$ into its real and imaginary parts, which by Lemma 19 go to 0 weakly in $L^p(\mathbb{R})$. We will consider various cases involving the dyadic intervals $I, J, K$.

**Case 1:** $I = J = K$. Then $s_k^I (H s_k^J) (H s_k^J) = s_k^I (H s_k^J)^2 \to 0$ weakly in $L^p(\mathbb{R})$ by Lemma 18.

**Case 2:** $I \neq J = K$. Then using that $|s_k^I|^2 = 1$, and similarly for $J$, we compute the real part

$$\text{Re} (f_k^I f_k^J f_k^K) = \text{Re} \left( f_k^I (f_k^J)^2 \right) = \text{Re} \left( (s_k^I + i H s_k^J) (s_k^I + i H s_k^J)^2 \right) = -2 (H s_k^J) s_k^I (H s_k^J) - s_k^I (H s_k^J)^2 .$$

Since the real part is the sum of two functions with disjoint support, by Lemma 17, $s_k^I (H s_k^J)^2 \to 0$ weakly in $L^p(\mathbb{R})$.

**Case 3:** $I = J \neq K$ or $I = K \neq J$. Assume without loss of generality that $I = J \neq K$. Using that $s_k^I s_k^K = 0$ because they have disjoint supports, we get $f_k^I f_k^J f_k^K$ has real part

$$-2 s_k^I (H s_k^J) (H s_k^J)^2 - s_k^K (H s_k^J)^2 \to 0$$

weakly in $L^p(\mathbb{R})$ by Lemma 17. But the two terms have disjoint support $J$ and $K$, so each goes to 0 weakly in $L^p(\mathbb{R})$.

**Case 4:** $I, J, K$ are pairwise disjoint. We compute the real part of $f_k^I f_k^J f_k^K$ equals

$$-s_k^I (H s_k^J) (H s_k^J) - (H s_k^I) s_k^K (H s_k^J) - s_k^K (H s_k^J) \to 0$$

weakly in $L^p(\mathbb{R})$, by Lemma 17. Since the three terms have pairwise disjoint support, then each individual term goes to 0 weakly in $L^p(\mathbb{R})$. \hfill □

### 4.2. From Hilbert to Riesz.

In analogy with $(H s_k^I)^2 \to 1_I$ weakly in $L^2$, we want to show that $(R_1 s_k^{P, \text{horizontal}})^2 \to c \text{I}P$ weakly in $L^2$ for some positive constant $c$, and also that $R_2 s_k^{P, \text{horizontal}} \to 0$ strongly in $L^2$, even $L^p$, as $k \to \infty$. Using real variable techniques, we will calculate matters in such a way that our claim for $R_1$ reduces to that of the Hilbert transform $H$, where the holomorphic methods used by Nazarov are available, while the claim for $R_2$ does not need reduction to $H$.

The following notation will also be useful.

**Notation 20.** Given a sequence $\{f_k\}_{k=1}^\infty$ of functions in $L^2(\mathbb{R}^n)$, we write

$$f_k = o_k^{\text{weakly}}(1)$$

if

$$\lim_{k \to \infty} \int_{\mathbb{R}^n} f_k(t) g(t) dt = 0 \text{ for all } g \in L^2(\mathbb{R}^n),$$

and we write $f_k = o_k^{\text{strongly}}(1)$ if

$$\lim_{k \to \infty} \int_{\mathbb{R}^n} |f_k(t)|^p dt = 0 \text{ for all } p \in (1, \infty).$$

We first need an elementary consequence of the alternating series test.

**Lemma 21.** If $b$ is a bounded function on $[0, 1]$ and there exists a partition $\{z_0 \equiv 0 < z_1 < \ldots < z_{N-1} < z_N \equiv 1\}$ such that $b$ is monotone, and of one sign, on each subinterval $(z_j, z_{j+1})$, then

$$\left| \int b(x) s_k^{[0,1]}(x) dx \right| \leq CN 2^{-k} \|b\|_{\infty} .$$

**Proof.** If $b$ is monotone on $[0, 1]$, and if say $b(0) > b(1) \geq 0$, then

$$\left| \int b(x) s_k^{[0,1]}(x) dx \right| = \sum_{j=0}^{2^k} (-1)^j \int_{\frac{z_j}{2^k}}^{\frac{z_{j+1}}{2^k}} b(x) dx \leq \int_0^{\frac{z_1}{2^k}} |b(x)| dx \leq 2^{-k} \|b\|_{\infty} ,$$

by the alternating series test. More generally, we can apply this argument to the subinterval $[z_{m-1}, z_m]$ if the endpoints lie in $\{j 2^{-k}\}_{j=0}^{2^k}$, the points of change in sign of $s_k^{[0,1]}$. In the general case, note that if we
denote by \( \frac{x_m-1}{m} \) or \( \frac{x_m}{m} \) the leftmost (or rightmost) point of the form \( \frac{x}{m} \) in \([z_{m-1}, z_m]\), then the integrals in each one of the intervals \([z_{m-1}, \frac{x_m-1}{m}]\), \([\frac{x_m-1}{m}, \frac{x_m}{m}]\), and \([\frac{x_m}{m}, z_m]\) all satisfy the same bound as \(13\). □

We will use Lemma 24 to prove the following lemmas, which encompass the technical details for the estimates in this section. We first need to establish some notation.

**Definition 22.** We say that a function \( g \) on \([a, b]\) is \( M \)-piecewise monotone if there is a partition \( \{a = t_1 < t_2 < \ldots < t_M = b\} \) such that \( g \) is monotone and of one sign on each subinterval \((t_k, t_{k+1})\), \( 1 \leq k < M \).

**Notation 23.** For \( x \in \mathbb{R}^n \) and \( P = P_1 \times \ldots \times P_n \) a cube in \( \mathbb{R}^n \), we write
\[
    x = (x_1, \ldots, x_n) = (x_1, x') = (x_1, x_2, x'') = (\bar{x}, x_n) = (x_1, \bar{x}, x_n),
\]
\[
    P = P_1 \times \ldots \times P_n = P_1 \times P' = P_1 \times P_2 \times P'' = \tilde{P} \times P_n = P_1 \times \tilde{P} \times P_n.
\]

**Definition 24.** The common definition of the \( \delta \)-halo of a cube \( P \) is given by
\[
    H_{\delta}^P \equiv \{ x \in \mathbb{R}^n : \text{dist}(x, \partial P) < \delta \}.
\]

Given a cube \( Q \supset P \) we define the \( Q \)-extended halo of \( P \) by
\[
    H_{\delta}^{Q, P} \equiv \{ x \in Q : \text{dist}(x, \partial P) < \delta \} \text{ for some } 1 \leq j \leq n\}.
\]

We also write \( s_k \) in place of \( s_k^{-1,1} \).

**Lemma 25.** Let \( p \in (1, \infty) \) and \( M \geq 1 \). Let \( P = P_1 \times P' \) be a subcube of a cube \( Q = Q_1 \times Q' \subset \mathbb{R} \times \mathbb{R}^{n-1} \). Furthermore suppose that
\[
    F : Q \times P_1 \times \tilde{P} \to \mathbb{R}
\]
satisfies the following three properties:
\[
    (4.5)
\]
(i) \( y_1 \to F(x, y_1, \bar{y}) \) is \( M \)-piecewise monotone for each \((x, \bar{y}) \in (Q \setminus H_{\delta}^{Q, P}) \times \tilde{P}\) for all \( 0 < \delta < \frac{1}{2} \),

(ii) \[
    \sup_{(x, y_1, \bar{y}) \in (Q \setminus H_{\delta}^{Q, P}) \times P_1 \times \tilde{P}} |F(x, y_1, \bar{y})| \leq C_{\delta} < \infty \text{ for all } 0 < \delta < \frac{1}{2},
\]

(iii) \[
    1_{H_{\delta}^{Q, P}}(x) \int_{P} \int_{P_1} |F(x, y_1, \bar{y})| \ dy_1 d\bar{y} \to 0 \text{ strongly in } L^p(Q) \text{ as } \delta \to 0.
\]

Then
\[
    \int_{P} \int_{P_1} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 d\bar{y} \to 0 \text{ strongly in } L^p(Q) \text{ as } k \to \infty.
\]

**Proof.** Write
\[
    \int_{P} \int_{P_1} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 d\bar{y} = \left\{ 1_{H_{\delta}^{Q, P}}(x) + 1_{Q \setminus H_{\delta}^{Q, P}}(x) \right\} \int_{P} \int_{P_1} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 d\bar{y}.
\]

For the first term use
\[
    1_{H_{\delta}^{Q, P}}(x) \int_{P} \int_{P_1} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 d\bar{y} \leq 1_{H_{\delta}^{Q, P}}(x) \int_{P} \int_{P_1} |F(x, y_1, \bar{y})| \ dy_1 d\bar{y}
\]

and the third assumption in \(4.5\).

For the second term we will use the alternating series test Lemma 21 adapted to the interval \( P_1 \) on the integral \( \int_{P_1} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 \) together with the first and second assumptions in \(4.5\). Indeed, by the first assumption and Lemma 21 we have that for \((x, \bar{y}) \in (Q \setminus H_{\delta}^{Q, P}) \times \tilde{P} \), there exists a partition \( \{t_0, t_1, \ldots, t_M\} \) of \( P_1 \) depending on \((x, \bar{y})\), but with \( M \) independent of \((x, \bar{y})\), such that
\[
    \left| \int_{P_1} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 \right| \leq \sum_{j=0}^{M-1} \int_{t_j}^{t_{j+1}} F(x, y_1, \bar{y}) \ s_k(y_1) \ dy_1 \leq CM \alpha 2^{-k},
\]

where the final inequality follows from the second assumption. Thus away from the halo we have uniform convergence to zero, and altogether we obtain the desired conclusion. □
We will also need a version of the previous lemma in which some of the \( y \) variables have been integrated out.

**Lemma 26.** Let \( p \in (1, \infty) \) and \( M \geq 1 \). Let \( P = P_1 \times P' \) be a subcube of a cube \( Q = Q_1 \times Q' \subset \mathbb{R} \times \mathbb{R}^{n-1} \). Furthermore suppose that
\[
F : Q \times P_1 \to \mathbb{R}
\]
satisfies the following three properties:

\[
\begin{align*}
(\text{i}) \quad & y_1 \to F(x, y_1) \text{ is } M\text{-piecewise monotone for each } x \in Q, \\
(\text{ii}) \quad & \sup_{(x,y_1) \in (Q \setminus H^{p,\infty}_s)^*} |F(x, y_1)| \leq C_\delta < \infty \text{ for all } 0 < \delta < \frac{1}{2}, \\
(\text{iii}) \quad & 1_{H^{p,\infty}_s}(x) \int_{P_1} |F(x, y_1)| dy_1 \to 0 \text{ strongly in } L^p(Q) \text{ as } \delta \to 0.
\end{align*}
\]

Then
\[
\int_{P_1} F(x, y_1) s_k(y_1) dy_1 \to 0 \text{ strongly in } L^p(Q) \text{ as } k \to \infty.
\]

**Proof.** Apply Lemma 25 to the function \((x, y_1, \bar{y}) \mapsto F(x, y_1)\). \(\square\)

Here is our main reduction of the action of Riesz transforms on \( s_k^{P,\text{horizontal}}(x) \) to that of the Hilbert transform \( H \) on \( s_k^{P_1}(x_1) \).

**Lemma 27.** Given \( n \geq 1 \), a cube \( P \subset \mathbb{R}^{n} \) and \( p \in (1, \infty) \), we have for \( x = (x_1, x') \in \mathbb{R}^1 \times \mathbb{R}^{n-1} \),
\[
R_1 s_k^{P,\text{horizontal}}(x) = B_n H s_k^{P_1}(x_1) 1_{P'}(x') + E_k^P(x),
\]

where
\[
B_n = c_n A_n A_{n-1} \ldots A_1, \quad c_n \text{ is as in (4.7)}, \quad A_n = \int_{\mathbb{R}} \frac{1}{(1 + z^2)^{n+1}} dz > 0,
\]
and \( E_k^P \) tends to 0 strongly in \( L^p(\mathbb{R}^n) \), i.e.
\[
\lim_{k \to \infty} \left\| E_k^P \right\|_{L^p(\mathbb{R}^n)} = 0.
\]

**Proof.** We prove the lemma by induction on the dimension \( n \geq 1 \). Since \( B_1 = 1 \), the case \( n = 1 \) is a tautology (with the understanding that \( R_1 = H \) on \( \mathbb{R} \), note that the constants in front of the integrals match) and so we now suppose that \( n \geq 2 \), and assume the conclusion of the lemma holds with \( n - 1 \) in place of \( n \).

Let \( \varepsilon > 0 \). For every \( M > 1 \), we have
\[
R_1 s_k^{P,\text{horizontal}}(x) = 1_{\mathbb{R}^n \setminus MP}(x) R_1 s_k^{P,\text{horizontal}}(x) + 1_{\mathbb{R}^n \setminus MP}(x) R_1 s_k^{P,\text{horizontal}}(x).
\]

We note that the second term \( 1_{\mathbb{R}^n \setminus MP}(x) R_1 s_k^{P,\text{horizontal}}(x) \) goes to 0 strongly in \( L^p(\mathbb{R}^n) \) as \( M \to \infty \), since
\[
\int_{\mathbb{R}^n \setminus MP} \left| R_1 s_k^{P,\text{horizontal}}(x) \right|^p dx \leq C \int_{\mathbb{R}^n \setminus MP} \left( \int_{\mathbb{R}^n \setminus MP} \left| \frac{1}{P} \right| dx' \right)^p dx' \leq C \int_{\mathbb{R}^n \setminus MP} \left( \frac{|P|}{|\text{dist}(x, P)|^{n+1}} \right)^p dx,
\]
which goes to 0 as \( M \to \infty \) uniformly in \( k \); in particular choose \( M \) such that \( \int_{\mathbb{R}^n \setminus MP} \left| R_1 s_k^{P,\text{horizontal}}(x) \right|^p dx < \frac{\varepsilon}{2} \) for all \( k \geq 0 \). With \( Q = MP \), it will suffice to show that \( \lim_{k \to \infty} \left\| E_k^P \right\|_{L^p(Q)} < \frac{\varepsilon}{2} \) for \( k \) sufficiently large, where \( E_k^P \) is implicitly defined as in the statement of the lemma.

Without loss of generality we suppose that \( P = [-1,1]^n \). Recalling that \( \hat{x} = (x_1, \ldots, x_{n-1}) \), \( \hat{y} = (y_1, \ldots, y_{n-1}) \), we write
\[
R_1 s_k^{P,\text{horizontal}}(x) = c_n \int_{-1}^{1} \int_{[-1,1]^{n-1}} \frac{(x_1 - y_1) s_k^{[-1,1]}(y_1)}{[(x_1 - y_1)^2 + |x' - y'|^2]^{n/2}} dy_1 \ldots dy_{n-1} dy_n.
\]
where, by the change of variables $z = \frac{x-y}{|x-y|}$, we have

$$
\Psi(\tilde{x}, x_n, \tilde{y}) = c_n \int_{-1}^{1} \frac{x_1 - y_1}{|\tilde{x} - \tilde{y}|^2 + |x_n - y_n|^2} dy_n = \frac{c_n}{c_{n-1}} K_1^{[n-1]}(\tilde{x} - \tilde{y}) \int \frac{z_{n-1} + 1}{(1 + z^2)^{n-\frac{1}{2}}} dz,
$$

and where $K_1^{[m]}$ is the kernel of the first individual Riesz transform $R_1^{[m]}$ in $m$ dimensions. Note

$$
\Phi^{n-1}(\tilde{x}, x_n, \tilde{y}) \equiv \int \frac{z_{n-1} + 1}{(1 + z^2)^{n-\frac{1}{2}}} dz
$$

is a bounded function of $(\tilde{x}, x_n, \tilde{y})$ with

$$
\|\Phi^{n-1}\|_\infty \leq \int \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz = A_n > 0.
$$

With $l_n(x, \tilde{y}) \equiv \frac{x_n + 1}{|x-n|}$ and $u_n(x, \tilde{y}) \equiv \frac{x_n + 1}{|x-n|}$ we may further decompose $\Phi^{n-1}(\tilde{x}, x_n, \tilde{y})$ as

$$
\left\{ \int_{l_n}^0 + \int_{u_n}^0 \right\} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz = \left\{ -\text{sgn}(x_n - 1) \int_{0}^{l_n} + \text{sgn}(x_n + 1) \int_{u_n}^{0} \right\} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz
$$

$$
= A_n \mathbf{1}_{P_n}(x_n) + \left\{ -\text{sgn}(x_n - 1) \int_{0}^{l_n} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz - A_n \right\} + \text{sgn}(x_n + 1) \int_{u_n}^{0} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz - A_n
$$

$$
= A_n \mathbf{1}_{P_n}(x_n) - \text{sgn}(x_n - 1) L_1^1(x, \tilde{y}) + \text{sgn}(x_n + 1) L_2^2(x, \tilde{y}).
$$

Relating the above computations to $R_1^{[n-1]}$ and $R_1^{[n]}$, we obtain

$$
R_1^{[n]} s_k^{\text{horizontal}}(x) = \frac{c_n}{c_{n-1}} A_n R_1^{[n-1]} \left( s_k^{[-1,1]} \otimes \mathbf{1}_{P_n} \right)(\tilde{x}) \mathbf{1}_{P_n}(x_n)
$$

$$
- \frac{c_n}{c_{n-1}} \text{sgn}(x_n - 1) \int_{[-1,1]} \frac{x_1 - y_1}{|\tilde{x} - \tilde{y}|} L_1^1(x, \tilde{y}) s_k^{[-1,1]}(y_1) dy_1 + \frac{c_n}{c_{n-1}} \text{sgn}(x_n + 1) \int_{[-1,1]} \frac{x_1 - y_1}{|\tilde{x} - \tilde{y}|} L_2^2(x, \tilde{y}) s_k^{[-1,1]}(y_1) dy_1
$$

$$
\equiv \frac{c_n}{c_{n-1}} A_n R_1^{[n-1]} \left( s_k^{[-1,1]} \otimes \mathbf{1}_{\tilde{P}} \right)(\tilde{x}) \mathbf{1}_{P_n}(x_n) + E_1^1(x) + E_2^1(x).
$$

We now apply our induction hypothesis to the term $R_1^{[n-1]} \left( s_k^{[-1,1]} \otimes \mathbf{1}_{\tilde{P}} \right)(\tilde{x})$ to obtain

$$
\frac{c_n}{c_{n-1}} A_n R_1^{[n-1]} \left( s_k^{[-1,1]} \otimes \mathbf{1}_{\tilde{P}} \right)(\tilde{x}) \mathbf{1}_{P_n}(x_n) = B_n H s_k^{P_n}(x_1) \mathbf{1}_{\tilde{P}}(x_2, ..., x_n) + \frac{c_n}{c_{n-1}} A_n E_1^\tilde{P}(\tilde{x}) \mathbf{1}_{P_n}(x_n)
$$

where $E_2^1(\tilde{x}) \mathbf{1}_{P_n}(x_n)$ tends to 0 strongly in $L^p(Q)$ by the induction hypothesis.

So it remains only to show that both $E_1^1(x)$ and $E_2^1(x)$ go to 0 strongly in $L^p(Q)$, and by symmetry it suffices to consider just $E_2^1(x)$. We have

$$
L_2^2(x, \tilde{y}) = \int_{0}^{\|u_n\|} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz - \frac{A_n}{2} = - \int_{\|u_n\|}^{\infty} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz,
$$

where we recall that $|u_n| = \frac{x_n + 1}{|x_n - y|}$.

We now see that it suffices to verify (i), (ii) and (iii) of Lemma 25 for the cube $Q$ and the function

$$
F(x, y_1, \tilde{y}) = \frac{x_1 - y_1}{|\tilde{x} - \tilde{y}|} \int_{\|u_n\|}^{\infty} \frac{1}{(1 + z^2)^{n-\frac{1}{2}}} dz.
$$

We first turn to verifying property (i), and since we only require estimates at this point, we will ignore absolute constants. The case $n = 2$ turns out to be rather special and easily handled so we dispose of that case first. We have when $n = 2$ that

$$
F(x, y_1) = \frac{K(x, y_1)}{x_1 - y_1} = \frac{1}{x_1 - y_1} \int_{|y_2(x, y_1)|}^{\infty} \frac{1}{(1 + z^2)^2} dz.
$$
where \( |u_2(x, y_1)| = \frac{|x_2 + 1|}{\sqrt{x_1 - y_1}} \). For any fixed \( x \), \( |u_2(x, y_1)| \) is monotone as a function of \( |x_1 - y_1| \). We now claim that the function \( F(x, y_1) \) is \( M \)-piecewise monotone for \( M = 7 \) as a function of \( y_1 \). Since \( F(x, y_1) \) only changes sign once, to see this it suffices to show that with \( s = |u_2(x, y_1)| \) the function

\[
H_\beta(s) = s \int_s^\infty (1 + t^2)^{-\beta} \, dt, \quad \text{for } s \in (-\infty, \infty), \quad \beta > \frac{1}{2},
\]

has 3 changes in monotonicity on \( (-\infty, \infty) \). We compute

\[
H''_\beta(s) = 2 \{(\beta - 1) s^2 - 1\} (1 + s^2)^{-\beta - 1}
\]

has at most 2 zeroes in \( (-\infty, \infty) \), hence \( H'_\beta(s) \) has at most 3 zeroes, which proves our claim.

Now we turn to the more complicated case \( n \geq 3 \). Let \( t = x - y \). Then we may write

\[
F(x, y_1, \bar{y}) = \frac{t_1}{(t_1^2 + |\bar{y}|^2)^\frac{n}{2}} V_n \left( \frac{|x_1 + 1|}{(t_1^2 + |\bar{y}|^2)^\frac{n}{2}} \right),
\]

where \( V_n(w) \equiv \int_w^\infty \frac{1}{(1 + z^2)^\frac{n}{2}} \, dz \). Note that the antiderivative

(4.8)

\[
\int \frac{1}{(1 + z^2)^\frac{n}{2}} \, dz = \int \frac{1}{(1 + \tan^2 \theta)^\frac{n+1}{2}} \, d \tan \theta = \int \frac{\sec^2 \theta}{(\sec^2 \theta)^\frac{n+2}{2}} \, d \theta = \int \cos^{n-1} \theta \, d \theta = C_n \theta + R_n \left( z, \sqrt{1 + z^2} \right), \quad z = \tan \theta,
\]

where \( R_n \) is a rational function of \( z = \tan \theta \) and \( \sqrt{1 + z^2} = \sec \theta \), and \( C_n = 0 \) when \( n \) is even. Indeed, one can use the well known recursion

\[
\int \cos^m \theta \, d \theta = \frac{1}{m} \cos^{m-1} \theta \, \sin \theta + \frac{m-1}{m} \int \cos^{m-2} \theta \, d \theta = \frac{1}{m} \sec^m \theta \, \tan \theta + \frac{m-1}{m} \int \cos^{m-2} \theta \, d \theta.
\]

Then setting

\[
z = \tan \theta = \frac{|x_1 + 1|}{(t_1^2 + |\bar{y}|^2)^\frac{n}{2}}, \quad E_0 \equiv \left( \frac{|x_1 + 1|}{|\bar{y}|} \right)^2, \quad E_1 \equiv \frac{|\bar{y}|}{|x_1 + 1|},
\]

and using (4.8), we may write (4.7) as

\[
F(x, y_1, \bar{y}) = \frac{t_1}{(t_1^2 + |\bar{y}|^2)^\frac{n}{2}} \left\{ R_n \left( z, \sqrt{1 + z^2} \right) + C_n \theta + C \right\} = E_1 \tan^{n-1} \theta \sqrt{E_0 - \tan^2 \theta} \left\{ R_n \left( z, \sqrt{1 + z^2} \right) + C_n \theta + C \right\} = D_{x, \bar{y}}(\theta).
\]

At this point, we employ the convention that \( R_n, T_n, U_n \) are rational functions which may change line to line, or instance to instance, but their degree will be bounded a constant depending only on the dimension \( n \). Similarly, we will take \( M \) to be an integer which may change line to line or instance to instance, but will only depend on the dimension \( n \). We also recall the fact that the function \( R_n \left( z, \sqrt{1 + z^2} \right) \) can equal 0 or \( \infty \) at most \( M \) times: indeed, \( R_n \) is a rational function of \( z \) and \( \sqrt{1 + z^2} \), which is in turn a nontrivial rational function of \( \sin \theta \) and \( \cos \theta \), with degree depending only on \( n \). Thus the number of zeros or poles it possesses is at most a constant depending only the degree, i.e. a constant which only depends \( n \).

Now fix \( x \) and \( \bar{y} \), or equivalently \( x \) and \( t \), and let us only consider when \( t_1 = x_1 - y_1 > 0 \), as the case \( t_1 < 0 \) will be similar. Then since \( t_1 \mapsto \theta(t_1) \) is a decreasing injective map from \( \mathbb{R}_+ \to (0, \frac{\pi}{2}) \), then \( y_1 \mapsto F(x, y_1, \bar{y}) \) is \( M \)-piecewise monotone on \( \{ y_1 \in \mathbb{R} : y_1 < x_1 \} \) if \( \theta \mapsto D_{x, \bar{y}}(\theta) \) is \( M \)-piecewise monotone on \( (0, \frac{\pi}{2}) \). Since \( t_1 > 0 \), then \( F \) is positive and so is \( D_{x, \bar{y}}(\theta) \) when \( \theta > 0 \), since both functions possess the same sign. Since \( u \mapsto u^2 \) is increasing for \( u > 0 \), then \( D_{x, \bar{y}}(\theta)^2 \) is \( M \)-piecewise monotone, which we will now show below.
In the reasoning that follows, we assume all rational functions we consider below are non-constant; in the case one of them is constant or even identically 0, the proof of $M$-piecewise monotonicity is even simpler than the proof below, the details of which we leave to the reader. We have
\[
D_{x,\theta}(\theta)^2 = E_1^2 \left[ E_0 - \tan^2 \theta \right] R_n \left( z, \sqrt{1+z^2} \right) \tan^{n-1} \theta + (C_n \theta + C) \tan^{n-1} \theta)^2
= R_n \left( z, \sqrt{1+z^2} \right) \theta^2 + T_n \left( z, \sqrt{1+z^2} \right) \theta + U_n \left( z, \sqrt{1+z^2} \right).
\]
To check $D_{x,\theta}(\theta)^2$ is $M$ monotone, it suffices to show $D_{x,\theta}(\theta)^2$ has at most $M$ critical points. For this we compute
\[
\frac{d}{d\theta} D_{x,\theta}(\theta)^2 = R_n \left( z, \sqrt{1+z^2} \right) \theta^2 + T_n \left( z, \sqrt{1+z^2} \right) \theta + U_n \left( z, \sqrt{1+z^2} \right)
= R_n \left( z, \sqrt{1+z^2} \right) \left\{ \theta^2 + T_n \left( z, \sqrt{1+z^2} \right) \theta + U_n \left( z, \sqrt{1+z^2} \right) \right\}
\]
which equals 0 or $\infty$ if
\[
R_n \left( z, \sqrt{1+z^2} \right) = 0 \text{ or } \infty, \text{ or } \theta^2 + \theta R_n \left( z, \sqrt{1+z^2} \right) + T_n \left( z, \sqrt{1+z^2} \right) = 0 \text{ or } \infty.
\]
The first equality can clearly only hold for at most $M$ values of $\theta$. To show the function
\[
\theta^2 + \theta R_n \left( z, \sqrt{1+z^2} \right) + T_n \left( z, \sqrt{1+z^2} \right)
\]
can equal 0 or $\infty$ at most $M$ times, it suffices to show that this function also has at most $M$ critical points.
Its derivative is of the form
\[
R_n \left( z, \sqrt{1+z^2} \right) \left( \theta + T_n \left( z, \sqrt{1+z^2} \right) \right),
\]
which we claim equals 0 or $\infty$ at most $M$ times. Indeed, $R_n$ equals 0 or $\infty$ at most $M$ times, and the function
\[
\theta + T_n \left( z, \sqrt{1+z^2} \right)
\]
equals 0 or $\infty$ at most $M$ times because its derivative is given by
\[
1 + T_n \left( z, \sqrt{1+z^2} \right),
\]
which in turn equals 0 or $\infty$ at most $M$ times.
Thus $y_1 \mapsto F(x, y_1, \tilde{y})$ is $M$-piecewise monotone for some $M$ depending only on $n$, and not on the additional parameters $x$ and $y_2, ..., y_n$. This completes the verification of property (i) in Lemma 25.

(ii) For any $x \in Q$ we have from (4.17) and $|u_n| = \frac{|1+x_n|}{|x-y|}$ that
\[
|F(x, y_1, \tilde{y})| \leq \frac{|x_1-y_1|}{|x-y|} \int_{u_n}^{\infty} \frac{1}{u_n^2 (1+z^2)^{n+1}} \frac{1}{z^2} dz = \frac{|x_1-y_1|}{|1+x_n|^n} \int_{u_n}^{\infty} \frac{1}{u_n^n (1+z^2)^{n+1}} \frac{1}{z^2} dz
\]
We claim that $|u_n|^n \int_{u_n}^{\infty} \frac{1}{u_n^n (1+z^2)^{n+1}} \frac{1}{z^2} dz \leq C_n$. Indeed, when $|u_n| \leq 1$, this follows from integrability of the integrand, and when $|u_n| \geq 1$ this follows from a direct computation using the fact that $(1+z^2)^{n+1} \approx z^{n+1}$.
Thus $|F(x, y_1, \tilde{y})| \leq C_n \frac{|x_1-y_1|}{|1+x_n|^n} \leq C_{n, Q, \delta}$ when $y \in P$, $x \in Q \setminus H^{P_Q}$.

(iii) To show $1_{H^{P_Q}}(x) \int_{-1}^{1} \int_{[-1,1]^{n-2}} |F(x, y_1, \tilde{y})| d\tilde{y} dy_1 \to 0$ strongly in $L^p(\mathbb{R}^n)$ as $\delta \to 0$, we split
\[
1_{H^{P_Q}}(x) \int_{-1}^{1} \int_{[-1,1]^{n-2}} |F(x, y_1, \tilde{y})| d\tilde{y} dy_1 \leq 1_{H^{P_Q}}(x) \int_{\tilde{y} \in [-1,1]^{n-1} : |\tilde{x}-\tilde{y}| > |1+x_n|} |F(x, y_1, \tilde{y})| d\tilde{y}
+ 1_{H^{P_Q}}(x) \int_{\tilde{y} \in [-1,1]^{n-1} : |\tilde{x}-\tilde{y}| \leq |1+x_n|} |F(x, y_1, \tilde{y})| d\tilde{y}.
\]
To bound the first term, we use the estimate
\[
|F(x, y_1, \tilde{y})| \leq \frac{|x_1-y_1|}{|x-y|} \int_{u_n}^{\infty} \frac{1}{u_n^2 (1+z^2)^{n+1}} \frac{1}{z^2} dz \leq \frac{1}{|x-y|^{n-1}} \int_{u_n}^{\infty} \frac{1}{u_n} \frac{1}{(1+z^2)^{n+1}} \frac{1}{z^2} dz \leq C_n \frac{1}{|x-y|^{n-1}},
\]
and polar coordinates to get
\[
\int_{\{\tilde{y}\in[-1,1]^{-1}:|\tilde{x}-\tilde{y}|>|1+x_n|\}} |F(x,y_1,\tilde{y})| \, d\tilde{y} \leq C_n \int_{\{\tilde{y}\in[-1,1]^{-1}:|\tilde{x}-\tilde{y}|>|1+x_n|\}} \frac{1}{|\tilde{x}-\tilde{y}|^{n-1}} \, d\tilde{y}
\]
\[
\leq C_n \int_{\mathbb{R}^n-2} \int_{|1+x_n|}^{cQ} \frac{1}{r} \, dr \, d\theta \leq C_n \ln \frac{1}{\text{dist}(x_n, \partial P)}
\]
where we have used the fact that $|\tilde{x}-\tilde{y}| \leq cQ$. Thus
\[
1_{H^p_Q}(x) \int_{\{\tilde{y}\in[-1,1]^{-1}:|\tilde{x}-\tilde{y}|>|1+x_n|\}} |F(x,y_1,\tilde{y})| \, d\tilde{y} \to 0
\]
strongly in $L^p(Q)$ as $\delta \to 0$.

As for the second term, for $|u_n| \geq 1$ we estimate
\[
|F(x,y_1,\tilde{y})| \leq \frac{|x_1-y_1|}{|\tilde{x}-\tilde{y}|^{n}} \int_{|u_n|}^{\infty} \frac{1}{(1+z^2)^{\frac{n+1}{2}}} \, dz \leq \frac{1}{|\tilde{x}-\tilde{y}|^{n-1}} \int_{|u_n|}^{\infty} \frac{1}{(1+z^2)^{\frac{n+1}{2}}} \, dz
\]
and so
\[
\int_{\{\tilde{y}\in[-1,1]^{-1}:|\tilde{x}-\tilde{y}|\leq|1+x_n|\}} |F(x,y_1,\tilde{y})| \, d\tilde{y} \leq C_n \int_{\{\tilde{y}\in[-1,1]^{-1}:|\tilde{x}-\tilde{y}|\leq|1+x_n|\}} \frac{1}{|1+x_n|^{n-1}} \, d\tilde{y} \leq C_n.
\]

Thus
\[
1_{H^p_Q}(x) \int_{\{\tilde{y}\in[-1,1]^{-1}:|\tilde{x}-\tilde{y}|\leq|1+x_n|\}} |F(x,y_1,\tilde{y})| \, d\tilde{y} \to 0
\]
strongly in $L^p(Q)$ as $\delta \to 0$.

The next lemma is an extension of the one-dimensional lemma of Nazarov in [NaVo].

**Lemma 28.** Suppose $p \in (1, \infty)$. Let $a$ and $b$ be nonnegative integers, not both zero. Given a cube $P = P_1 \times P_2 \times \ldots \times P_n \subset \mathbb{R}^n$, we have

1. \( \lim_{k \to \infty} \int_{\mathbb{R}^n} \left( s_k^{P,\text{horizontal}}(x) \right)^a \left( R_1 s_k^{P,\text{horizontal}}(x) \right)^b f(x) \, dx = 0 \) for all functions $f \in L^p(\mathbb{R}^n)$ when $a$ or $b$ is odd.
2. \( \lim_{k \to \infty} \int_{\mathbb{R}^n} \left( s_k^{P,\text{horizontal}}(x) \right)^a \left( R_1 s_k^{P,\text{horizontal}}(x) \right)^b f(x) \, dx = C_{a,b} \int_{\mathbb{R}^n} f(x) \, dx \) for all functions $f \in L^p(\mathbb{R}^n)$ when both $a$ and $b$ are even, and $C_{a,b} > 0$ and $C_{a,0} = B_2^2$.
3. For all $k \geq 1$ we have
   \[
   \left| R_{2k}^{P,\text{horizontal}}(x) \right| \leq C \ln \frac{1}{\text{dist}(x_2, \partial P_2)} 1_{\{\text{dist}(x_2, \partial P_2) < \delta\}}(x) + C_4 2^{-k} \left( 1_{\{\text{dist}(x_2, \partial P_2) > \delta\}}(x) \right),
   \]
   so that $R_{2k}^{P,\text{horizontal}}(x)$ tends to 0 strongly in $L^p(\mathbb{R}^n)$ for all $p \in (1, \infty)$.

**Proof.** (1) and (2): By Lemma 27, we may write
\[
\left( s_k^{P,\text{horizontal}}(x) \right)^a \left( R_1 s_k^{P,\text{horizontal}}(x) \right)^b f(x) = B_n^{b} \left( s_k^{P,\text{horizontal}}(x) \right)^a \left( H s_k^{P,\text{horizontal}}(x) \right)^b f(x) \mathbf{1}_{P'}(x') + E_k^{P,\text{horizontal}}(x),
\]
where $E_k^{P,\text{horizontal}}(x)$ goes to 0 strongly in $L^1(Q)$, and $P' = P_2 \times \ldots \times P_n$ and $x = (x_2, \ldots, x_n)$. Thus integrating over $\mathbb{R}^n$ and using Lemma 13 yields the conclusions sought.

(3): Let $\varepsilon > 0$. Arguing as in the proof of Lemma 27 for every $M > 1$, we have
\[
R_{2k}^{P,\text{horizontal}}(x) = R_{2k}^{P,\text{horizontal}}(x) 1_{\mathbb{R}^n \setminus M P}(x) + R_{2k}^{P,\text{horizontal}}(x) 1_{\mathbb{R}^n \setminus M P}(x).
\]
We note that the second term $R_{2k}^{P,\text{horizontal}}(x) 1_{\mathbb{R}^n \setminus M P}(x)$ goes to 0 strongly in $L^p(\mathbb{R}^n)$ as $M \to \infty$, since
\[
\int_{\mathbb{R}^n \setminus M P} \left| R_{2k}^{P,\text{horizontal}}(x) \right|^p \, dx \leq C \int_{P \setminus M P} \left( \frac{1}{\text{dist}(x, P)} \right)^p \, dx \leq C \int_{P \setminus M P} \left( \frac{|P|}{\text{dist}(x, P)} \right)^p \, dx,
\]
which goes to 0 as \( M \to \infty \). So choose \( M \) such that \( \int_{\mathbb{R}^d \setminus M} |R^P_{2k}\|_{L^p}^p \) \( dx < \frac{\varepsilon}{2} \). Thus with \( Q = MP \), it remains to show that \( \|R^P_{2k}\|_{L^p(Q)} < \frac{\varepsilon}{2} \) for \( k \) sufficiently large.

Again we may assume that \( P = [-1, 1]^n \). We have

\[
R^P_{2k}(x) = \int_{-1}^{1} \int_{-1}^{1} \cdots \int_{-1}^{1} c \frac{(x_2 - y_2) s_{\frac{-1}{1}}^{k-1}(y_1)}{[\sqrt{x - y_2} + |x'' - y''|]} \, dy_1 \, dy' = \int_{-1}^{1} F(x, y_1) s_{\frac{-1}{1}}^{k-1}(y_1) \, dy_1,
\]

and it suffices to show that

\[
F(x, y_1) \equiv c_n \int_{[-1, 1]^{n-1}} \frac{x_2 - y_2}{[\sqrt{x - y_2} + |x'' - y''|]} \, dy'
\]

satisfies conditions (i)-(iii) of Lemma 26. Before proceeding, we note that by oddness of the kernel in \( y_2 \), and the change of variable \( t = x_2 - y_2 \), we have

\[
F(x, y_1) = -c_n \int_{[-1, 1]^{n-2}} \left\{ \int_{-1}^{x_2-1} \frac{t}{[\sqrt{x - y_2} + |x'' - y''|]} \, dt \right\} \, dy''
\]

\[
= -c_n \int_{[-1, 1]^{n-2}} \left\{ \int_{x_2+1}^{1} \frac{t}{[\sqrt{x - y_2} + |x'' - y''|]} \, dt \right\} \, dy''
\]

\[
= c_n \zeta(x) \int_{[-1, 1]^{n-2}} \left\{ \int_{\max\{|x_2+1|,|x_2-1|\}}^{\min\{|x_2+1|,|x_2-1|\}} \frac{t}{[\sqrt{x - y_2} + |x'' - y''|]} \, dt \right\} \, dy'',
\]

where \( \zeta(x) \) is a unimodular function.

(i) We note that \( F(x, y_1) \zeta(x) \) is an integral of concave down functions in \( y_1 \) with critical point at \( x_1 \), which by interchanging differentiation and integration, must itself also be a concave down function in \( y_1 \) with critical point at \( x_1 \), and hence \( F \) is 3-piecelwise monotone.

(ii) By the form of \( F \) we computed using oddness of the kernel above, we have

\[
|F(x, y_1)| \leq c_n \int_{[-1, 1]^{n-2}} \left\{ \int_{\max\{|x_2+1|,|x_2-1|\}}^{\min\{|x_2+1|,|x_2-1|\}} \frac{t}{[\sqrt{x - y_2} + |x'' - y''|]} \, dt \right\} \, dy''
\]

\[
\leq c_n \int_{[-1, 1]^{n-2}} \left\{ \int_{\max\{|x_2+1|,|x_2-1|\}}^{\min\{|x_2+1|,|x_2-1|\}} \frac{t}{\delta^{n+1}} \, dt \right\} \, dy'',
\]

since if \( x \in Q \setminus H^P_{\delta} \) then \( t > \delta \) by separation. Thus \( \left| \chi_{Q \setminus H^P_{\delta}}(x) F(x, y_1) \right| \leq C^{-\frac{1}{3}} \).

(iii) Let

\[
A \equiv \{(y_1, y'') \in [-1, 1]^{n-1} : |(x_1 - y_1, x'' - y'')| > |1 - x_2|\},
\]

\[
B \equiv \{(y_1, y'') \in [-1, 1]^{n-1} : |(x_1 - y_1, x'' - y'')| < |1 - x_2|\},
\]

we have

\[
\int_{\mathbb{R}^d \setminus M} |R^P_{2k}|^p + \int_{\mathbb{R}^d \setminus M} |R^P_{2k}|^p \leq c_n \int_{[-1, 1]^{n-2}} \left\{ \int_{\max\{|x_2+1|,|x_2-1|\}}^{\min\{|x_2+1|,|x_2-1|\}} \frac{t}{[\sqrt{x - y_2} + |x'' - y''|]} \, dt \right\} \, dy''.
\]
and assume without loss of generality that $|x_2 - 1| \leq |x_2 + 1|$. For $x \in H^{p,q}_s$, we have

$$
\int_{-1}^{1} |F(x, y_1)| dy_1 \leq \int_{-1}^{1} \int_{[-1,1]^{n-2}} \left\{ \int_{\min\{|x_2+1,|x_2-1|\}}^{\infty} \frac{t}{(x_1 - y_1)^2 + t^2 + |x' - y'|^2} \frac{dt}{\sqrt{\frac{2\pi}{x_2}}} \right\} dy'' dy_1
$$

$$
= \frac{1}{n-1} \int_{-1}^{1} \int_{[-1,1]^{n-2}} \left\{ \frac{1}{(x_1 - y_1)^2 + (1 - x_2)^2 + |x' - y'|^2} \right\} dy'' dy_1
$$

$$
\leq \left\{ \int_{A} + \int_{B} \right\} \left\{ \frac{1}{(x_1 - y_1)^2 + (1 - x_2)^2 + |x' - y'|^2} \right\} dy'' dy_1
$$

$$
\leq \int_{A} \frac{1}{(x_1 - y_1)^2 + |x' - y'|^2} \frac{dy''}{\sqrt{\frac{2\pi}{x_2}}} + \int_{B} \frac{1}{|1 - x_2|^{n-1}} dy'' dy_1.
$$

By a crude estimate the second integral is bounded by

$$
\int_{B} \frac{1}{|1 - x_2|^{n-1}} dy'' \leq C_n |B| \frac{1}{|1 - x_2|^{n-1}} \leq C_n.
$$

As for the first integral, integration using polar coordinates yields the upper bound

$$
\int_{1-\delta}^{1-\epsilon} r^{n-2} dr = \ln \frac{c_n}{|1 - x_2|^{n-1}} \in L^p(Q).
$$

Similar estimates hold when $x_2 < 0$ and $x \in H^{p,q}_s$. Thus $1_{H^{p,q}_s}(x) \int_{-1}^{1} |F(x, y_1)| dy_1$ goes to 0 strongly in $L^p(Q)$ as $\delta \to 0$. \qed

**Theorem 29.** The conclusions of Lemma 19 hold if one replaces $H$ by $R_1$ and $s^{k}_h$ by $s^{l}_{k \text{ horizontal}}$, and similarly for $J, K$.

**Proof.** One argues as previously in the proofs of Lemma 28 parts (1) and (2), in particular using Lemmas 27 and Lemma 19. \qed

5. Boundedness properties of the Riesz transforms

We now are equipped with the convergence results we need to complete the proof of the main theorem by following the supervisor argument of Nazarov in [NaVo]. We begin with a short formal argument in the plane, and then complete the proof by adapting Nazarov’s supervisor argument for the Hilbert transform to the transplantation of Riesz transforms.

5.1. The formal argument in the plane. We now take $Q^0 = [0, 1]^2$ to be the unit square in the plane, and let $V, U$ be as arising from Theorem 12. We apply the transplantation argument of Section 3 to $V, U$ to obtain weights $v_t, u_t$ for all $1 \leq t \leq m$, with $u = u_m, v = v_m$. We will compute the testing conditions for $(u, v)$ by first estimating them for the pair $(v_{t+1} - v_t, u_t)$. Since both vertical $\triangle_{P \text{ vertical}}$ and checkerboard $\triangle_{P \text{ checkerboard}}$ components of $V, U$ vanish for all $P$, then by the estimates of Section 3 we obtain that in the limit only the diagonal terms in $|R_1(v_{t+1} - v_t)|^2$ survive the integration with $U$. Indeed, recall that

$$
R_j (v_{t+1} - v_t) = \sum_{Q \in K_t} \left\langle V, s^{h_{S(Q)}}_{Q, \text{horizontal}} \right\rangle \frac{1}{\sqrt{\|S(Q)\|}} R_j s^{Q, \text{horizontal}}_{k+1},
$$

and the vanishing weak convergence results of Section 3 yield for $k_{t+1} >> k_t$

$$
\int R_1 s^{Q, \text{horizontal}}_{k_{t+1}} R_1 s^{Q', \text{horizontal}}_{k+1} u_t \to \begin{cases} 0 & \text{if } Q \neq Q' \text{ on } [0, 1]^2, \\ (B_2)^2 \int u & \text{if } Q = Q' \end{cases}
$$
where $B_2$ is the constant appearing in Lemma 27, and so using once again the vanishing weak convergence results of Section 4 we get for $k_{t+1} \gg k_t$

$$
\int [R_1 (u_{t+1} - u_t)]^2 u_t = \int \left[ \sum_{Q \in K_t} \langle V, h_{S(Q)}^\text{horizontal} \rangle \frac{1}{\sqrt{|S(Q)|}} R_{f, S(Q)}^{k_{t+1}} \right]^2 u_t
$$

$$
= \sum_{Q \in K_t} \int \langle V, h_{S(Q)}^\text{horizontal} \rangle^2 [R_1 (u_{t+1} - u_t)]^2 \frac{1}{|S(Q)|} u_t + \text{offdiagonal} \rightarrow (B_2)^2 \sum_{Q \in K_t} \langle V, h_{S(Q)}^\text{horizontal} \rangle^2 \frac{1}{|S(Q)|} \int_Q u_t,
$$

and if we now add these results in $t$, pigeonhole cubes $Q$ based on their supervisor $S$, use the fact that $E_Q u_t = E_S U$, and finally $\sum_{Q \in K_t} \frac{|Q|}{|S|} = 1$, we obtain

$$
\int \left[ \sum_{t=1}^m (u_{t+1} - u_t) \right]^2 u_t \approx \sum_{t=1}^m \int [R_1 (u_{t+1} - u_t)]^2 u_t
$$

$$
\approx (B_2)^2 \sum_{t=1}^m \sum_{Q \in K_t} \langle V, h_{S(Q)}^\text{horizontal} \rangle^2 \frac{1}{|S(Q)|} \int_Q u_t = (B_2)^2 \sum_{t=1}^m \sum_{S \in D_t} \sum_{Q \in K_t, S(Q) = S} \langle V, h_S^\text{horizontal} \rangle^2 E_S U \frac{|Q|}{|S|}
$$

$$
= (B_2)^2 \sum_{t=1}^m \sum_{S \in D_t} \langle V, h_S^\text{horizontal} \rangle^2 E_S U > (B_2)^2 \Gamma (E_{[0,1]} v V),
$$

which shows that testing for $R_1$ blows up. On the other hand, when we test with $R_2$ we get the bound

$$
\left( \sum_{P \in D : \ell (P) \geq 2^{-m}} \langle V, h_P^\text{horizontal} \rangle^2 \frac{1}{|P|} \int_P U \right)^2
$$

which is small for $k_1, k_2, \ldots, k_{t+1}$ all chosen large enough in an inductive fashion.

Now we consider the dual inequalities for both $R_1$ and $R_2$. Note however, that we needn’t bother with the dual testing condition for $R_1$ since we already know the forward testing condition fails to hold, and so $R_1$ is unbounded in any event. Thus we must estimate the term $\int [R_2 (u_{t+1} - u_t)]^2 v$. But since the vertical and checkerboard components both vanish, interchanging the roles of $u$ and $v$ yields that the dual testing condition for $R_2$ is bounded.

To make this formal argument precise in the next subsection, we follow the corresponding argument in \cite{NaVo}.

### 5.2. The Nazarov argument for Riesz transforms

We now continue to carry out our adaptation of Nazarov’s supervisor argument to the higher dimensional setting of the supervisor and transplantation map. Equipped with the supervisor and transplantation map, and the weak convergence results above, this remaining argument is now virtually verbatim the corresponding argument in \cite{NaVo}.

Recall that $\{k_t\}_{t=0}^\infty$ is a strictly increasing sequence of nonnegative integers $k_t \in \mathbb{Z}_+$ with $k_0 = 0$, and whose members will be chosen sufficiently large in the arguments below. We define $\mathcal{K} = \bigcup_{t=0}^\infty \mathcal{K}_t$ where $\mathcal{K}_0 = \{Q^0\} = \{[0,1]^2\}$ and

$$
\mathcal{K}_t = \{Q \in \mathcal{D} : \ell (Q) = 2^{-k_1 - k_2 - \ldots - k_t} \}, \quad t \geq 1.
$$
Proposition 30 (Nazarov [NaVo] in the case of the Hilbert transform). For every $\Gamma > 1$ and $0 < \tau < 1$ there exist positive weights $u, v$ on the unit cube $Q^0 \equiv [0,1]^n$ satisfying

$$\int_{[0,1]^n} |R_1 v(x)|^2 u(x) \, dx \geq \Gamma \int_{[0,1]^n} v(x) \, dx,$$

$$\int_I |R_2 1 v(x)|^2 u(x) \, dx \leq \int_I v(x) \, dx, \quad \text{for all dyadic cubes } I \in \mathcal{D}^0,$$

$$\int_I |R_2 1 u(x)|^2 v(x) \, dx \leq \int_I u(x) \, dx, \quad \text{for all dyadic cubes } I \in \mathcal{D}^0,$$

$$\left(\frac{1}{|I|} \int_I u(x) \, dx\right) \left(\frac{1}{|I|} \int_I v(x) \, dx\right) \leq 1, \quad \text{for all cubes } I \in \mathcal{P}^0,$$

$$1 - \tau < \frac{E_{jk} u_{jk}}{E_{kk} v_{jk}} < 1 + \tau, \quad \text{for arbitrary adjacent cubes } J, K \in \mathcal{P}^0.$$

Proof: We begin by considering just dyadic subcubes of $Q^0$ and later discuss the extension to cubes in $\mathcal{P}^0$, as defined in Section 2.

Let $V, U$ be as arising from Theorem 12 with $\frac{2(V, U, Q^0)}{E_{kk} v} > \Gamma'$ sufficiently large. We apply the transplantation argument of Section 3 to $V, U$ to obtain weights $v_t, u_t$ with $1 \leq t \leq m$, with $u = u_m, \ v = v_m$, where $m$ is as in Theorem 12. For convenience we give the details in the plane, and leave the easy extension to higher dimensions for the reader. In particular then we have

$$u \equiv 1_{Q^0} E_{Q} u + \sum_{t=0}^{m-1} \sum_{Q \in \mathcal{K}_t} \left\langle U, h_{S(Q)} \right\rangle \frac{1}{\sqrt{|S(Q)|}} s_{k_{t+1}}^{Q, \text{horizontal}},$$

$$v \equiv 1_{Q^0} E_{Q} v + \sum_{t=0}^{m-1} \sum_{Q \in \mathcal{K}_t} \left\langle V, h_{S(Q)} \right\rangle \frac{1}{\sqrt{|S(Q)|}} s_{k_{t+1}}^{Q, \text{horizontal}},$$

Recall from the construction above that

$$v_{t+1}(x) - v_t(x) = \sum_{Q \in \mathcal{K}_t} \left\langle V, h_{S(Q)} \right\rangle \frac{1}{\sqrt{|S(Q)|}} s_{k_{t+1}}^{Q, \text{horizontal}}(x),$$

$$v_t(x) = v_m(x),$$

and similarly for $u(x)$. We recall that both $u$ and $v$ are positive on $[0,1]^2$. It will be convenient to denote the differences

$$\eta_{t+1} \equiv u_{t+1} - u_t,$$

$$\delta_{t+1} \equiv v_{t+1} - v_t,$$

respectively. Note that $\eta_t, \delta_t$ are sums of the form

$$\sum_{Q \in \mathcal{K}_t} c_Q \frac{1}{\sqrt{|S(Q)|}} s_{k_{t+1}}^{Q, \text{horizontal}} = \alpha_{k_{t+1} \to \infty}^{\text{weakly}}(1),$$

where the sum is $\alpha_{k_{t+1} \to \infty}^{\text{weakly}}(1)$ because the constants $c_Q$ depend only on the levels $1$ through $t$ of the construction and the number of terms in the sum only depends on $k_1, \ldots, k_t$. Let $R_j$ denote the Riesz transform with convolution kernel $K_j(z) = \frac{2}{|z|^j}$.

$$R_j f(x) = \int \frac{y_j - x_j}{|y - x|^j} f(y) \, dy.$$

We will now focus on the ‘testing’ constants $\sum_{Q \in \mathcal{K}_t} \frac{1}{|Q|} \int_{[0,1]^2} |R_1 v(x)|^2 u(x) \, dx$, and $\sup_{Q \in \mathcal{D}(Q^0)} \frac{1}{|Q|} \int_{Q} |R_1 v|^2 u \, dx$, and show that the first is large, and second and third are small, provided we take the integers $k_t$ sufficiently large in an inductive fashion. Define the discrepancy and dual discrepancy
for $R_j$ on $Q$ by

$$\text{Disc}^{v \to u}_{2;Q}(t) = \int_Q (R_j^1 Q u_{t+1}(x))^2 u_{t+1}(x) \, dx - \int_Q (R_j^1 Q v_t(x))^2 u_t(x) \, dx,$$

$$\text{Disc}^{v \to u}_{3;Q}(t) = \int_Q (R_j^1 Q u_{t+1}(x))^2 v_{t+1}(x) \, dx - \int_Q (R_j^1 Q u_t(x))^2 v_t(x) \, dx.$$  

We begin with Nazarov’s identity,

$$\text{Disc}^{v \to u}_{1;Q}(t) = \int_Q (R_j^1 Q \delta_{t+1} + R_j^1 Q v_t)^2 u_{t+1} - \int_Q (R_j^1 Q v_t)^2 u_t$$

$$= \int_Q (R_j^1 Q \delta_{t+1})^2 u_{t+1} + \int_Q \{2 (R_j^1 Q \delta_{t+1}) (R_j^1 Q v_t)\} (u_t + \eta_{t+1}) + \int_Q (R_j^1 Q v_t)^2 (u_{t+1} - u_t)$$

$$= \langle (R_j^1 Q \delta_{t+1})^2, u_{t+1} \rangle_{L^2(Q)} + 2 \langle (R_j^1 Q \delta_{t+1}) (R_j^1 Q v_t), u_t \rangle_{L^2(Q)}$$

$$+ 2 \langle (R_j^1 Q v_t), \eta_{t+1} \rangle_{L^2(Q)} + \langle (R_j^1 Q v_t)^2, \eta_{t+1} \rangle_{L^2(Q)}$$

$$= A_{1;Q} + B_{1;Q} + C_{1;Q} + D_{1;Q}.$$  

We first claim that with $\text{Disc}^{v \to u}_{1;Q}(t) \equiv \text{Disc}^{v \to u}_{1;[0,1]^2}(t)$ and $A_1 \equiv A_{1;[0,1]^2}$, etc...,

$$\text{Disc}^{v \to u}_{1}(t) = A_1 + B_1 + C_1 + D_1 = (B_2)^2 \sum_{I \in D: \ell(I) = 2^{-t}} \langle \Delta_{I} \text{horizontal} V, (E_I U) \rangle + \sum_{r=0}^{t} o_{k+1 \to \infty} (1).$$

We will see in a moment that $A_1$ is the main term. Using that $u_t, v_t$ and $\delta_{t+1}, \eta_{t+1}$ are supported in $[0,1]^2$,

$$B_1 = 2 \langle (R_1 v_t) u_t, R_1 \delta_{t+1} \rangle_{L^2([0,1]^2)} = -2 \langle R_1 [(R_1 v_t) u_t], \delta_{t+1} \rangle_{L^2([0,1]^2)} = o_{k+1 \to \infty} (1)$$

since the function $R_1 [(R_1 v_t) u_t] \in L^p (\mathbb{R}^2)$ for all $p \in (1, \infty)$, and in particular belongs to $L^2 (\mathbb{R}^2)$, and is independent of $k_{t+1}$, and finally that $\delta_{t+1} \sim o_{k_{t+1} \to \infty} (1)$. Similarly, since $R_1 v_t \in L^4 (\mathbb{R}^2)$, we have

$$D_1 = \langle (R_1 v_t)^2, \eta_{t+1} \rangle_{L^2([0,1]^2)} = o_{k_{t+1} \to \infty} (1).$$

For term $C_1$ we have

$$C_1 = 2 \langle (R_1 \delta_{t+1}) (R_1 v_t), \eta_{t+1} \rangle_{L^2([0,1]^2)}$$

$$= 2 \int_{[0,1]^2} \left( \sum_{Q \in \mathcal{K}_t} \left\langle V, h_{\text{horizontal}}^{Q, \text{horizontal}} \frac{1}{\sqrt{|S(Q)|}} \frac{1}{s_{k_{t+1}}} \right\rangle \left( R_1 \right)_{Q, \text{horizontal}} \left( \sum_{Q' \in \mathcal{K}_t} \left\langle U, h_{\text{horizontal}}^{Q', \text{horizontal}} \frac{1}{\sqrt{|S(Q')|}} \frac{1}{s_{k_{t+1}}} \right\rangle \left( R_1 \right)_{Q', \text{horizontal}} \right) \right)$$

$$= 2 \sum_{Q, Q' \in \mathcal{K}_t} \left\langle V, h_{\text{horizontal}}^{Q, \text{horizontal}} \frac{1}{\sqrt{|S(Q)|}} \frac{1}{s_{k_{t+1}}} \right\rangle \left( R_1 \right)_{Q, \text{horizontal}} \left( \left\langle U, h_{\text{horizontal}}^{Q', \text{horizontal}} \frac{1}{\sqrt{|S(Q')|}} \frac{1}{s_{k_{t+1}}} \right\rangle \left( R_1 \right)_{Q', \text{horizontal}} \right)$$

by Theorem [29] since $k_t$ and $R_1 v_t$ are both independent of $k_{t+1}$, while $\left( R_1 s_{k_{t+1}}^{Q, \text{horizontal}} \right) s_{k_{t+1}}^{Q', \text{horizontal}} \to 0$ weakly in $L^2(\mathbb{R}^2)$.

Finally, for term $A_1$ we have

$$A_1 = \langle (R_1 \delta_{t+1})^2, u_{t+1} \rangle_{L^2([0,1]^2)} = \left\langle \left( \sum_{I \in \mathcal{K}_t} \left\langle V, h_{\text{horizontal}}^{I, \text{horizontal}} \frac{1}{\sqrt{|S(I)|}} \frac{1}{s_{k_{t+1}}} \right\rangle \left( R_1 \right)_{I, \text{horizontal}} \right)^2, u_{t+1} \right\rangle_{L^2([0,1]^2)}.$$
We first note that if the sum is taken outside the square, so that we consider only the ‘diagonal’ terms, we have
\[
\left\langle \sum_{I \in K_t} \left( V, h_{S(I)}^\text{horizontal} \right) R_1 \frac{1}{\sqrt{|S(I)|}} s_{k_{t+1}}^\text{horizontal}, u_{t+1} \rightangle
\]
\[
= \sum_{I \in K_t} \frac{1}{|S(I)|} \left\langle V, h_{S(I)}^\text{horizontal} \rightangle^2 \left\langle \frac{|I|}{|S(I)|} E_S(I) U \right\rangle + \sum_{I \in K_t} \frac{1}{|S(I)|} \left\langle V, h_{S(I)}^\text{horizontal} \rightangle^2 \left\langle \frac{|I|}{|S(I)|} E_S(I) U \right\rangle + o_{k_{t+1} \to \infty} (1)
\]
\[
\equiv F + G + o_{k_{t+1} \to \infty} (1)
\]
by Lemma 28 part (2) for \( k_{t+1} \) sufficiently large, and since \( \frac{1}{|I|} \int_I u_t = E_{S(I)} U \).

Turning now to the sum of the off diagonal terms,
\[
\sum_{I \neq I'} \frac{1}{|S(I)|} \frac{1}{|S(I')|} \left\langle R_1 \left[ \left( V, h_{S(I)}^\text{horizontal} \right) s_{k_{t+1}}^\text{horizontal} \right], R_1 \left[ \left( V, h_{S(I')}^\text{horizontal} \right) s_{k_{t+1}}^\text{horizontal} \right], \eta_{t+1} \right\rangle,
\]
we see that they all tend to 0 weakly as \( k_{t+1} \to \infty \) by Theorem 29. Thus we can choose the components of the sequence \( \{ k_t \}_{t=1}^m \) sufficiently large that
\[
\int_{[0,1]^2} |R_1 v(x)|^2 u(x) dx \geq (1' - CA_2^\text{dyadic}(V, U, [0,1]^2)) \int_{[0,1]^2} v(x) dx,
\]
since we also have
\[
\int_{[0,1]^2} |R_1 v_0(x)|^2 v_0(x) dx = \int_{[0,1]^2} |R_1 1_{[0,1]^2} E_{[0,1]^2} V|^2 1_{[0,1]^2} E_{[0,1]^2} U dx
\]
\[
= \left( E_{[0,1]^2} V \right)^2 \left( E_{[0,1]^2} U \right) \int_{[0,1]^2} |R_1 1_{[0,1]^2}|^2 dx = C \left( E_{[0,1]^2} V \right)^2 \left( E_{[0,1]^2} U \right) \leq CA_2(V, U, [0,1]^2) E_{[0,1]^2} V.
\]

Our next task is to show that for each \( t \geq 1 \), we have
\[
\text{Disc}_{2;Q}^{u \to v} (t) = O(1) \quad \text{and} \quad \text{Disc}_{2;Q}^{u \to v} (t) = O(1),
\]
for all \( Q \in D^0 \) such that \( \ell(Q) \geq 2^{-k_1-k_2-\ldots-k_t} \).

The two discrepancies above are symmetric, and so it suffices to show only the first assertion. However, arguing using part (3) of Lemma 28 and Theorem 29, we obtain \( \text{Disc}_{2;Q}^{u \to v} (t) = o_{k_{t+1} \to \infty} (1) \).

Let \( t = \ell(Q) \), if it exists, be such that \( 2^{-k_1-k_2-\ldots-k_t} \leq \ell(Q) < 2^{-k_1-k_2-\ldots-k_{t-1}} \). We will deal with the remaining cubes \( Q \) later. At each stage \( t \), there are only finitely many cubes \( Q \in D^0 \) such that \( \ell(Q) \geq 2^{-k_1-k_2-\ldots-k_t} \), and hence only finitely many terms which are \( o_{k_{t+1} \to \infty} (1) \) in (5.4).
weights $u_{t-1}$ and $v_{t-1}$ are constant on such cubes $Q$, and $u = u_{t-1} + \eta_t + \sum_{s=t+1}^{m-1} \eta_s$ and $v = v_{t-1} + \delta_t + \sum_{s=t}^{m-1} \delta_{s+1}$, we can choose the components of the sequence $\{k_t\}_{t=1}^m$ sufficiently large that
\[
\frac{1}{|Q|} \int_Q |R_21Qv(x)|^2 u(x) \, dx = \frac{1}{|Q|} \int_Q |R_21Q(u_{t-1} + \delta_t)(x)|^2 u(x) \, dx + o(1), \quad \text{for all } Q \in \mathcal{D}^0,
\]
\[
\frac{1}{|Q|} \int_Q |R_21Qu(x)|^2 v(x) \, dx = \frac{1}{|Q|} \int_Q |R_21Q(u_{t-1} + \eta_t)(x)|^2 v(x) \, dx + o(1), \quad \text{for all } Q \in \mathcal{D}^0.
\]
Indeed, we use $u(x) \leq \|U\|_\infty$ independent of the choice of $k_1, \ldots, k_m$, which gives using part (3) of Lemma 28,
\[
\frac{1}{|Q|} \int_Q |R_21Qv_{\text{top}}(x)|^2 u(x) \, dx \leq \frac{1}{|Q|} \int_Q |R_21Q\left(\sum_{s=t}^{m-1} \delta_s\right)(x)|^2 u(x) \, dx \rightarrow 0 \text{ weakly as } k_{t+j} \rightarrow \infty, \quad j = 1, 2, \ldots, m-t,
\]
where we recall that $t = t(Q)$, and similarly for the second line.
Thus it remains to show that
\[
\frac{1}{|Q|} \int_Q |R_21Q(u_{t-1} + \delta_t)(x)|^2 u(x) \, dx = O(1) \quad \text{and} \quad \frac{1}{|Q|} \int_Q |R_21Q(u_{t-1} + \eta_t)(x)|^2 v(x) \, dx = O(1).
\]
By symmetry it suffices to prove the first assertion. We define
\[
u_{\text{top}} \equiv u_{t-1} + \eta_t, \quad \nu_{\text{below}} \equiv u - u_{\text{top}}, \quad t = t(Q),
\]
and similarly for $v$.

The left hand side of the first assertion can be written as
\[
\int_Q (R_21Q\nu_{\text{top}})^2 u \, dx = \int_Q (R_21Q\nu_{\text{top}})^2 u_{\text{top}} \, dx + \int_Q (R_21Q\nu_{\text{top}})^2 u_{\text{below}} \, dx.
\]
The second term can be made arbitrarily small by choosing $k_{t+1}$ sufficiently large, and using the weak convergence of $u_{\text{below}} \rightarrow 0$ in $L^p(\mathbb{R}^n)$, as $R_21Q\nu_{\text{top}}$ is independent of $k_{t+1}$.

So we are left with estimating $\int_Q (R_21Q\nu_{\text{top}})^2 u_{\text{top}} \, dx$. Note now that $E_Q\nu_{\text{top}} = E_Qv = E_{S(Q^*)}V$, where $Q^*$ is the unique cube in $K_t$ containing $Q$. Note as well that $\nu_{\text{top}}$ is constant on each $I \in K_{t+1}$, and satisfies the pointwise estimate $\nu_{\text{top}}(x) \leq (E_{S(Q^*)}V) (1 + \tau)$, since $\nu_{\text{top}}$ inherits dyadic $\tau$-flatness from $V$; similarly for $u$. Then applying the pointwise estimate to $\nu_{\text{top}}$, followed by using $\|R_21Q\nu_{\text{top}}\|_{L^2} \leq \|1Q\nu_{\text{top}}\|_{L^2}$ by boundedness of $R_2$, and then the pointwise estimate applied to $\nu_{\text{top}}$, we get
\[
\int_Q (R_21Q\nu_{\text{top}})^2 u_{\text{top}} \, dx \leq (1 + \tau)^3 (E_{S(Q^*)}U) \int_Q (R_21Q\nu_{\text{top}})^2 \, dx \leq (1 + \tau)^3 \|E_{S(Q^*)}V\|^2 |Q|.
\]
Since $A_2\text{dyadic}(V;U;Q^0) \leq 1$, the above is controlled by
\[
(1 + \tau)^3(E_QV)|Q| = (1 + \tau)^3(E_QV)|Q| = (1 + \tau)^3 \int_Q v.
\]
Finally we consider $\ell(Q) < 2^{-k_1-k_2-\ldots-k_m}$. Then $v, u$ are constant on $Q$ with $E_Qv = E_{S(Q^*)}V$, $E_Qu = E_{S(Q^*)}U$, and so
\[
\int_Q (R_21Qv)^2 u = (E_{S(Q^*)}V)^2 (E_{S(Q^*)}U) \int_Q (R_21Q)^2 \leq (E_{S(Q^*)}V) |Q| = \int_Q v,
\]
where in the inequality we used that $(E_{S(Q^*)}V) (E_{S(Q^*)}U) \leq 1$ and $\|R_2\|_{L^2 \rightarrow L^2} = 1$.

Since $\tau \in (0, 1)$, we obtain that the dual testing constant for $R_2$ on dyadic squares is bounded; similarly for the testing constant on dyadic squares.

Finally, to remove the restriction of dyadic from the $A_2$ and doubling conditions, one can modify the transplantation argument following [NaVo], as described in Appendix [NaVo]. However, complete proofs were not provided in [NaVo] and we invite the reader to consult Appendix [NaVo] for missing details.
Finally, by multiplying \( v, u \) by an appropriate positive constant, we obtain the statements in the theorem with the required constants.

**Remark 31.** The weights \( u(x), v(x) \) in \([0, 1]^n\) that are constructed in the proof of Proposition 30 depend only on the first variable \( x_1 \) of \( x \).

**Remark 32.** A careful reading of the proof shows that our weights \( v, u \) satisfy the \( L^p \)-testing and dual \( L^p \)-testing conditions for the operator \( R_2 \) when \( p \in (1, \infty) \). Thus if there was a \( T1 \) theorem for \( L^p \) with doubling weights, our results regarding \( R_2 \) would extend to \( L^p \).

In order to complete the proof of Theorem 4, we need to extend our doubling conclusions to classical doubling, and remove the restriction to dyadic cubes in our testing conditions for the weight pair \((v, u)\) in Proposition 30.

### 5.3 Classical doubling, \( A_2 \) and dyadic testing in \( \mathbb{R}^n \)

Recall that a measure \( \mu \) has doubling exponent \( \theta_{\text{doub}}(\mu) \to n \) if the adjacency constant \( \lambda_{\text{adj}}(\mu) \to 1 \).

By Proposition 30, we have constructed a pair of \( \tau \)-flat weights \((v, u)\) on \( Q^0 = [0, 1]^n \), which we relabel here as \((\sigma, \omega)\), that satisfy the \( A_2(\sigma, \omega; [0, 1]^n) \) condition as well as the dyadic testing conditions,

\[
\int_{Q^0} |R_2(1_{Q^0})|^2 d\sigma \leq \left( \sum_{Q \in D^0} |\sigma_Q| \right)^2 |Q|_{\sigma}, \quad Q \in D^0,
\]

where we have included the superscript \( D^0 \) in the testing constants to indicate that the cubes \( Q \) are restricted to the dyadic grid \( D^0 \).

We now extend these measures to the entire space by reflecting in each coordinate separately to obtain an extension to \([0, 2]^n\), and then by adding translates \([0, 2]^n + 2(\alpha_1, \alpha_2, ..., \alpha_n), \alpha \in \mathbb{Z}^n\), so as to be periodic of period two on the entire space \( \mathbb{R}^n \). Most importantly, after this reflection process the pair \((\sigma, \omega)\) satisfies and \( A_2 \) condition, as well as the dyadic testing conditions for all \( D \)-dyadic cubes of side length at most 1. Furthermore note that adjacent cubes from neighbouring dyadic cubes of side length 1 also satisfy the adjacent doubling condition, and with constant 1 since they have equal measures by the reflection extension process, and so for any adjacent dyadic cubes \( I_1 \) and \( I_2 \), we have \( \frac{\mu_{I_1}^{E_2}(\sigma)}{\mu_{I_2}^{E_1}(\sigma)} \in (1 - \tau, 1 + \tau) \), and similarly for \( \omega \).

We now set

\[
Q_{\alpha} \equiv [0, 1]^n + (\alpha_1, \alpha_2, ..., \alpha_n), \quad \text{for all } \alpha \in \mathbb{Z}^n.
\]

Let \( \tau \in (0, 1) \) be as in Proposition 30 and multiply each of these measures by the factor

\[
\varphi_\tau(x) = \sum_\alpha a_\alpha 1_{Q_\alpha}(x),
\]

where \( a_\alpha = \frac{1}{|Q_\alpha|} \int_{Q_\alpha} d\mu_\tau \) and \( d\mu_\tau(x) = \frac{dx}{(1 + |x|)^\tau} \), and consider the measure pairs \((\sigma_\tau, \omega_\tau)\) with \( \sigma_\tau = \varphi_\tau(x) d\sigma(x) \) and \( \omega_\tau = \varphi_\tau(x) d\omega(x) \). We set \( A = ||Q_\alpha||_\sigma \) and \( B = ||Q_\alpha||_\omega \) for all \( \alpha \in \mathbb{Z}^n \), and \( AB \leq A_2(\sigma, \omega) \).

**Lemma 33.** The measures \((\sigma_\tau, \omega_\tau)\) are \( o(1) \)-flat, or equivalently the adjacent doubling constant tends to 1 as \( \tau \searrow 0 \).

**Proof.** Indeed, if \( Q_\alpha \) and \( Q_{\alpha'} \) are two adjacent cubes of the form \( Q_\alpha \equiv [0, 1]^n + (\alpha_1, \alpha_2, ..., \alpha_n) \), then

\[
\frac{\int_{Q_\alpha} \sigma_\tau}{\int_{Q_{\alpha'}} \sigma_\tau} = \frac{a_\alpha}{a_{\alpha'}} \frac{\int_{Q_\alpha} \sigma}{\int_{Q_{\alpha'}} \sigma} \quad \text{and} \quad \frac{\mu_\tau}{\mu_{\tau'}} = \frac{\int_{Q_\alpha} d\mu_\tau}{\int_{Q_{\alpha'}} d\mu_{\tau}},
\]

tends to 1 as \( \tau \searrow 0 \) independent of the pair \((Q_\alpha, Q_{\alpha'})\) since \( \mu_\tau \) is a doubling weight on \( \mathbb{R}^n \) with adjacent doubling constant roughly \( 1 + O_{\tau \to 0}(\tau) \). If instead we consider adjacent cubes \( P \) and \( P' \) that are each a
union of cubes $Q_\alpha$, then
\[
\frac{\int_{P} \sigma_{\tau}}{\int_{P'} \sigma_{\tau}} = \frac{\sum_{\alpha:Q_\alpha \subset P} a_\alpha |Q_\alpha|_{\sigma}}{\sum_{\alpha:Q_\alpha \subset P'} a_\alpha |Q_\alpha|_{\sigma}} = \frac{\sum_{\alpha:Q_\alpha \subset P} \int_{Q_\alpha} d\mu_{\tau}}{\sum_{\alpha:Q_\alpha \subset P'} \int_{Q_\alpha} d\mu_{\tau}} = \frac{\int_{P} d\mu_{\tau}}{\int_{P'} d\mu_{\tau}},
\]
which again tends to 1 as $\tau \searrow 0$ independent of the pair $(P, P')$. Thus for any adjacent dyadic cubes $I_1$ and $I_2$, we have $E_{I_1} \sigma_{\tau} \leq E_{I_2} \sigma_{\tau} \in (1 - \tau, 1 + \tau)$. A standard argument shows that $\sigma_{\tau}$ has adjacent doubling constant equal to $1 + o(1)$ as $\tau \searrow 0$, and similarly for $\omega_{\tau}$.

Next we turn to the final task of establishing the testing conditions for $R_2$ on the doubling measure pair $(\sigma_{\tau}, \omega_{\tau})$ uniformly for any $\tau \in (0, 1)$, which then leads to boundedness of $R_2$ via the main result in \[AlSaUr\] (or in the case $\tau > 0$ sufficiently small, one can use either \[SaShUr7\] or \[LaWi\]). Of course, testing fails for $R_1$. For this we will need the definition of a weighted norm inequality as used in \[AlSaUr\].

We follow the approach in \[SaShUr9\], see page 314]. So we suppose that $K^\alpha$ is a standard smooth $\alpha$-fractional Calderón-Zygmund kernel, and $\sigma, \omega$ are locally finite positive Borel measures on $\mathbb{R}^n$, and we introduce a family \(\{\eta^\alpha_{\delta,R}\}_{0 < \delta < R < \infty}\) of nonnegative functions on $[0, \infty)$ so that the truncated kernels
\[
K^\alpha_{\delta,R}(x,y) = \eta^\alpha_{\delta,R}|x-y||K^\alpha(x,y)
\]
are bounded with compact support for fixed $x$ or $y$, and uniformly satisfy the smooth Calderón-Zygmund kernel estimates. Then the truncated operators
\[
T^\alpha_{\sigma,\delta,R} f(x) \equiv \int_{\mathbb{R}^n} K^\alpha_{\delta,R}(x,y) f(y) d\sigma(y), \quad x \in \mathbb{R}^n,
\]
are pointwise well-defined when $f$ is bounded with compact support, and we will refer to the pair $(K^\alpha, \{\eta^\alpha_{\delta,R}\}_{0 < \delta < R < \infty})$ as an $\alpha$-fractional singular integral operator, which we typically denote by $T^\alpha$, suppressing the dependence on the truncations. In the event that $\alpha = 0$ and $T^0$ is bounded on unweighted $L^2(\mathbb{R}^n)$, we say that $T^0$ is a Calderón-Zygmund operator.

**Definition 34.** We say that an $\alpha$-fractional singular integral operator $T^\alpha = (K^\alpha, \{\eta^\alpha_{\delta,R}\}_{0 < \delta < R < \infty})$ satisfies the norm inequality
\[
\|T^\alpha f\|_{L^2(\omega)} \leq \mathcal{N}_{T^\alpha} \|f\|_{L^2(\sigma)}, \quad f \in L^2(\sigma),
\]
provided
\[
\|T^\alpha_{\sigma,\delta,R} f\|_{L^2(\omega)} \leq \mathcal{N}_{T^\alpha_{\sigma,\delta,R}} (\sigma, \omega) \|f\|_{L^2(\sigma)}, \quad f \in L^2(\sigma), 0 < \delta < R < \infty.
\]

**Independence of Truncations:** In the presence of the classical Muckenhoupt condition $A^2_\infty$, the norm inequality (5.5) is independent of the choice of truncations used, including nonsmooth truncations as well - see \[LaSaShUr3\].

In Section 2 we introduced dyadic testing conditions. Here we introduce their continuous counterparts.

**Definition 35.** We say that an $\alpha$-fractional singular integral operator $T^\alpha = (K^\alpha, \{\eta^\alpha_{\delta,R}\}_{0 < \delta < R < \infty})$ satisfies the testing conditions if
\[
\Xi_T^{\alpha,\sigma,\omega}(\sigma, \omega)^2 \equiv \sup_{Q \in \mathcal{P}_n} \frac{1}{|Q|_{\sigma}} \int_Q |T^\alpha_{\sigma}(1_Q \sigma)|^2 d\omega < \infty, \quad \Xi_{T^{\alpha,\sigma,\omega}}(\omega, \sigma)^2 \equiv \sup_{Q \in \mathcal{P}_n} \frac{1}{|Q|_{\omega}} \int_Q |T^{\alpha,\sigma,\omega}(1_Q \omega)|^2 d\sigma < \infty.
\]

When we say that testing conditions hold for singular integrals, we will also mean that they hold uniformly over all admissible truncations.

Finally, we record an estimate from \[Saw0\] that will be used in proving the next lemma.

**Lemma 36 (\[Saw6\]).** If $\mu$ is a doubling measure and $P$ is a cube, then for every $\delta \in (0, 1)$ we have
\[
|\{x \in P : \text{dist}(x, \partial P) < \delta \ell(P)\}|_\mu \lesssim \frac{1}{\ln \frac{1}{\delta}}.
\]

**Lemma 37.** For all $\tau > 0$, the second Riesz transform $R_2$ satisfies the norm inequality for the measure pair $(\sigma_{\tau}, \omega_{\tau})$.
Proof. First fix a dyadic cube \( Q \in \mathcal{D} \). If \( Q \) has side length at most 1, then \( Q \) is contained in one of the cubes \( Q_\alpha \), where we have already shown that the testing conditions for \((\sigma, \omega)\) hold in Proposition [30]. In particular we have the following inequality that will be used repeatedly below,

\[
(5.6) \quad \int_{Q_\alpha} |R_2 (1_{Q_\alpha} \sigma_t)|^2 \, d\omega_t = a_\alpha^3 \int_{Q_\alpha} |R_2 (1_{Q_\alpha} \sigma)|^2 \, d\omega \leq C_* a_\alpha |Q_\alpha|_\sigma = C_* |Q_\alpha|_{\sigma_t}, \quad \alpha \in \mathbb{Z}^n.
\]

So suppose \( Q \) has side length \( 2^k \) with \( k \geq 1 \) for some \( k \in \mathbb{N} \). Then \( Q \) is a finite pairwise disjoint union of cubes \( Q_\beta \), say \( Q = \bigcup_{|\beta| \leq 2^k} Q_\beta \), where \( |\beta| = \max \{ \beta_1, \beta_2, \ldots, \beta_n \} \). We will suppose that \( Q = [0, 2^k]^n \) as the general case follows the same argument. Finally we note that \( a_\alpha \approx \frac{1}{(1+|\alpha_t|)^{r}} \). Now we write

\[
(5.7) \quad \int_Q |R_2 (1_Q \sigma_t)|^2 \, d\omega_t = \sum_{|\alpha_1|, |\alpha_2|, |\alpha_3| = 0}^{2^k} \int_{Q_{\alpha_1}} R_2 (1_{Q_{\alpha_2}} \sigma_t) \cdot R_2 (1_{Q_{\alpha_3}} \sigma_t) \, d\omega_t
\]

\[
\leq \sum_{|\alpha_1|, |\alpha_2|, |\alpha_3| = 0}^{2^k} \left[ \frac{1 + |\alpha_2 - \alpha_1|}{(1 + |\alpha_2|)^{r}} \right] \left[ \frac{1 + |\alpha_3 - \alpha_1|}{(1 + |\alpha_3|)^{r}} \right] \frac{|Q_{\alpha_2}|_\sigma |Q_{\alpha_3}|_\sigma |Q_{\alpha_1}|_\omega}{(1 + |\alpha_1|)^{r}},
\]

and consider several cases separately.

First we assume that \(|\alpha_2 - \alpha_1| \geq 2\) and \(|\alpha_3 - \alpha_1| \geq 2\), so that what we need to bound is

\[
\sum_{|\alpha_1|, |\alpha_2|, |\alpha_3| = 0}^{2^k} \left( \frac{1 + |\alpha_2 - \alpha_1|}{(1 + |\alpha_2|)^{r}} \right) \left( \frac{1 + |\alpha_3 - \alpha_1|}{(1 + |\alpha_3|)^{r}} \right) \frac{|Q_{\alpha_2}|_\sigma |Q_{\alpha_3}|_\sigma |Q_{\alpha_1}|_\omega}{(1 + |\alpha_1|)^{r}},
\]

where we suppress the specified conditions \(|\alpha_2 - \alpha_1| \geq 2\) and \(|\alpha_3 - \alpha_1| \geq 2\) in the sum. Summing first over \(\alpha_3\) and using \(|Q_{\alpha_3}|_\sigma = A\) we see that the above term is dominated by

\[
\sum_{|\alpha_1|, |\alpha_2| = 0}^{2^k} \left( \frac{1 + |\alpha_2 - \alpha_1|}{(1 + |\alpha_2|)^{r}} \right) \left( \frac{1 + |\alpha_3 - \alpha_1|}{(1 + |\alpha_3|)^{r}} \right) \frac{A |Q_{\alpha_2}|_\sigma |Q_{\alpha_1}|_\omega}{(1 + |\alpha_1|)^{r}}
\]

\[
\leq A \sum_{|\alpha_1|, |\alpha_2| = 0}^{2^k} \left[ \sum_{|\alpha_3| = 0}^{2^k} \left( \frac{1 + |\alpha_3 - \alpha_1|}{(1 + |\alpha_3|)^{r}} \right) \frac{A |Q_{\alpha_2}|_\sigma |Q_{\alpha_1}|_\omega}{(1 + |\alpha_1|)^{r}} \right] \frac{1 + |\alpha_2 - \alpha_1|}{(1 + |\alpha_2|)^{r}}
\]

\[
\leq A \sum_{|\alpha_1|, |\alpha_2| = 0}^{2^k} \left[ \sum_{|\alpha_3| = 0}^{2^k} \frac{1 + |\alpha_3 - \alpha_1|}{(1 + |\alpha_3|)^{r}} \frac{A |Q_{\alpha_2}|_\sigma |Q_{\alpha_1}|_\omega}{(1 + |\alpha_1|)^{r}} \right] \frac{1 + |\alpha_2 - \alpha_1|}{(1 + |\alpha_2|)^{r}}
\]

Now summing over \(\alpha_1\), using that using \(|Q_{\alpha_1}|_\omega = B\) and that \(AB \leq A_2(\sigma, \omega)\), we obtain in a similar way that the final line above is at most a constant times

\[
A_2(\sigma, \omega) \sum_{|\alpha_2| = 0}^{2^k} \left[ \frac{\ln (2 + |\alpha_2|)}{(1 + |\alpha_2|)^{3r}} \right] A_2(\sigma, \omega) \sum_{|\alpha_2| = 0}^{2^k} \left[ \frac{\ln (2 + |\alpha_2|)}{(1 + |\alpha_2|)^{2r}} \right] A_2(\sigma, \omega)
\]

\[
\leq CA_2(\sigma, \omega) \sum_{|\alpha_2| = 0}^{2^k} |Q_{\alpha_2}|_{\sigma_t} = CA_2(\sigma, \omega) |Q|_{\sigma_t},
\]

where we used that \(AB \leq A_2(\sigma_t, \omega_t) = A_2\).
The relatively simple case we just proved is case (6) in the following exhaustive list of cases, which we
delineate based on the relationship of the indices $\alpha_2$ and $\alpha_3$ to the distinguished index $\alpha_1$:

1. $\alpha_1 = \alpha_2 = \alpha_3$.
2. $\alpha_1 = \alpha_2$ and $Q_{\alpha_1}, Q_{\alpha_3}$ are separated,
3. $\alpha_1 = \alpha_3$ and $Q_{\alpha_1}, Q_{\alpha_2}$ are separated,
4. $Q_{\alpha_1}, Q_{\alpha_2}$ are adjacent and $Q_{\alpha_1}, Q_{\alpha_3}$ are separated,
5. $Q_{\alpha_1}, Q_{\alpha_3}$ are adjacent and $Q_{\alpha_1}, Q_{\alpha_2}$ are separated,
6. $Q_{\alpha_1}, Q_{\alpha_2}$ are separated and $Q_{\alpha_1}, Q_{\alpha_3}$ are separated

(7) \[
\begin{aligned}
&\alpha_1 = \alpha_2 \quad \text{and} \quad Q_{\alpha_1}, Q_{\alpha_3} \text{ are adjacent} \\
&\alpha_1 = \alpha_3 \quad \text{and} \quad Q_{\alpha_1}, Q_{\alpha_2} \text{ are adjacent} \\
&Q_{\alpha_1}, Q_{\alpha_2} \text{ are adjacent} \quad \text{and} \quad Q_{\alpha_1}, Q_{\alpha_3} \text{ are adjacent}
\end{aligned}
\]

where we say that $Q_{\alpha_1}, Q_{\alpha_2}$ are \textit{separated} if $|\alpha_1 - \alpha_2| \geq 2$, and of course $Q_{\alpha_1}, Q_{\alpha_2}$ are adjacent if and only if $|\alpha_1 - \alpha_2| = 1$.

In the first of these seven cases, the right hand side of (5.7) satisfies

\[
\sum_{|\alpha| = 0}^{2^k} \int_{Q_{\alpha}} |R_2 (1, \sigma_{\tau})|^2 \, d\omega_{\tau} \leq C_{\ast} \sum_{|\alpha| = 1}^{2^k} |Q_{\alpha}|_{\sigma_{\tau}} = C_{\ast} |Q|_{\sigma_{\tau}},
\]

independent of $\tau \in (0, 1)$ by (5.6).

In the second of these cases, we will use the separation between $Q_{\alpha_1}$ and $Q_{\alpha_3}$, as well as the fact that

\[
(5.8) \quad \left| \int_{Q_{\alpha_1}} R_2 (1, \sigma_{\tau}) \, d\omega_{\tau} \right| \leq \left( \int_{Q_{\alpha_1}} |R_2 (1, \sigma_{\tau})|^2 \, d\omega_{\tau} \right)^{\frac{1}{2}} \sqrt{|Q_{\alpha_1}|_{\sigma_{\tau}}}
\leq \sqrt{C_{\ast}} \sqrt{|Q_{\alpha_1}|_{\sigma_{\tau}}} \sqrt{|Q_{\alpha_1}|_{\omega_{\tau}}} \approx \sqrt{C_{\ast}} \frac{AB}{(1 + |\alpha_1|)^{\tau}},
\]

where the second inequality follows from reasoning using (5.6), similar to the previous display. Thus recalling
that $AB \leq A_2 (\sigma, \omega)$, we dominate the right hand side of (5.7) using (5.8) by

\[
\sum_{|\alpha_1|, |\alpha_3| = 0}^{2^k} \int_{Q_{\alpha_1}} \frac{|R_2 (1, \sigma_{\tau})|}{(1 + |\alpha_3 - \alpha_1|)^{\tau}} \frac{|Q_{\alpha_3}|_{\sigma}}{(1 + |\alpha_1|)^{\tau}} \, d\omega \leq A_2 \sqrt{C_{\ast}} \sum_{|\alpha_3| = 0}^{2^k} \frac{|Q_{\alpha_3}|_{\sigma}}{(1 + |\alpha_3|)^{\tau}} \sum_{|\alpha_1| = 0}^{2^k} \frac{(1 + |\alpha_3 - \alpha_1|)^{\tau}}{(1 + |\alpha_1|)^{\tau}} \leq A_2 \sqrt{C_{\ast}} \sum_{|\alpha_3| = 0}^{2^k} \frac{|Q_{\alpha_3}|_{\sigma} \ln (2 + |\alpha_3|)}{(1 + |\alpha_3|)^{2\tau}} \leq CA_2 \sqrt{C_{\ast}} |Q|_{\sigma_{\tau}}.
\]

To handle the cases where $Q_{\alpha_1}$ is adjacent to one of the cubes $Q_{\alpha_2}$ or $Q_{\alpha_3}$ or both, we use the fact in
Lemma 36 that doubling measures charge halos with reciprocal log control. Indeed, in the fourth case above, namely $|\alpha_1 - \alpha_2| = 1$ and $|\alpha_1 - \alpha_3| \geq 2$, we follow the same argument just used except that in place of the
testing condition in (5.8), we use

\[
\int_{Q_{\alpha_1}} R_2 (1, \sigma_{\tau}) \, d\omega_{\tau} = I + II.
\]

We control the first term $I$ by $\delta$-separation between $(1 - \delta) Q_{\alpha_1}$ and $Q_{\alpha_2}$:

\[
|I| \leq \int_{(1 - \delta)Q_{\alpha_1}} C \frac{1}{\delta n} |Q_{\alpha_2}|_{\sigma_{\tau}} \, d\omega_{\tau} \leq C \frac{1}{\delta n} |Q_{\alpha_2}|_{\sigma_{\tau}} |Q_{\alpha_1}|_{\omega_{\tau}} = C \frac{1}{\delta n} \frac{AB}{(1 + |\alpha_1|)^{\tau} (1 + |\alpha_2|)^{\tau}} \leq C \frac{1}{\delta n} \frac{A_2}{(1 + |\alpha_1|)^{\tau} (1 + |\alpha_2|)^{\tau}}.
\]
We control the second term $II$ by using Lemma 36:

$$
|II| \leq \int_{Q_{\alpha_1} \setminus (1-\delta) Q_{\alpha_1}} |R_2(1_{Q_{\alpha_2}} \sigma_\tau)| \, d\omega_\tau \leq N_\delta R_2(\sigma_\tau, \omega_\tau) \sqrt{|Q_{\alpha_2}|_{\sigma_\tau} |Q_{\alpha_1} \setminus (1-\delta) Q_{\alpha_1}|_{\omega_\tau}} \\
\leq \frac{C}{\sqrt{\ln \frac{1}{\delta}}} N_\delta R_2(\sigma_\tau, \omega_\tau) \frac{\sqrt{A_2}}{(1 + |\alpha_1|)^{\frac{3}{2}}} \leq \frac{C}{\sqrt{\ln \frac{1}{\delta}}} N_\delta R_2(\sigma_\tau, \omega_\tau) \frac{\sqrt{A_2}}{(1 + |\alpha_1|)^{\frac{3}{2}}}.
$$

Altogether, our replacement for (5.8) is

(5.9) \quad \left| \int_{Q_{\alpha_1}} R_2(1_{Q_{\alpha_2}} \sigma_\tau) \, d\omega_\tau \right| \leq \left( C_1 \sqrt{A_2} + \frac{C}{\sqrt{\ln \frac{1}{\delta}}} N_\delta R_2(\sigma_\tau, \omega_\tau) \right) \frac{\sqrt{A_2}}{(1 + |\alpha_1|)}.

since $|\alpha_1 - \alpha_2| = 1$. Now the previous argument can continue using (5.9) in place of (5.8), which proves the fourth case since there are just $3^n - 1$ points $\alpha_2$ for each fixed point $\alpha_1$. Indeed, we have

$$
\sum_{|\alpha_1|, |\alpha_3|=0}^{2^k} \int_{Q_{\alpha_1}} \frac{|R_2(1_{Q_{\alpha_2}} \sigma_\tau)| (1 + |\alpha_3 - \alpha_1|)^{-n} |Q_{\alpha_1}|_{\sigma}}{(1 + |\alpha_1|)^{\frac{3}{2}}} \frac{|Q_{\alpha_2}|_{\sigma}}{(1 + |\alpha_1|)^{\frac{3}{2}}} \, d\omega \\
\leq \left( C_1 \sqrt{A_2} + \frac{C}{\sqrt{\ln \frac{1}{\delta}}} N_\delta R_2(\sigma_\tau, \omega_\tau) \right) \sum_{|\alpha_3|=0}^{2^k} \frac{|Q_{\alpha_3}|_{\sigma}}{(1 + |\alpha_3|)^{\frac{3}{2}}} \sum_{|\alpha_1|=0}^{2^k} \frac{(1 + |\alpha_3 - \alpha_1|)^{-n}}{(1 + |\alpha_1|)^{\frac{3}{2}}}.
$$

The third and fifth cases are symmetric to those just handled. So it remains to consider the remaining seventh case, where one of the following three subcases holds:

$$
\begin{cases}
\alpha_1 = \alpha_2 & \text{and } |\alpha_1 - \alpha_3| = 1 \\
\alpha_1 = \alpha_3 & \text{and } |\alpha_1 - \alpha_2| = 1 \\
|\alpha_1 - \alpha_2| = 1 & \text{and } |\alpha_1 - \alpha_3| = 1.
\end{cases}
$$

In all three of these subcases, there is essentially only the sum over $\alpha_1$ since for each fixed $\alpha_1$, there are at most $3^{2n}$ pairs $(\alpha_2, \alpha_3)$ satisfying one of the three subcases. If both $Q_{\alpha_2}$ and $Q_{\alpha_3}$ are adjacent to $Q_{\alpha_1}$, we write

$$
\int_{Q_{\alpha_1}} R_2(1_{Q_{\alpha_2}} \sigma_\tau) R_2(1_{Q_{\alpha_3}} \sigma_\tau) \, d\omega_\tau = \int_{Q_{\alpha_1}} R_2(1_{(1-\delta)Q_{\alpha_2}} \sigma_\tau) R_2(1_{(1-\delta)Q_{\alpha_3}} \sigma_\tau) \, d\omega_\tau \\
+ \int_{Q_{\alpha_1}} R_2(1_{(1-\delta)Q_{\alpha_2}} \sigma_\tau) R_2(1_{Q_{\alpha_3} \setminus (1-\delta)Q_{\alpha_2}} \sigma_\tau) \, d\omega_\tau + \int_{Q_{\alpha_1}} R_2(1_{Q_{\alpha_2} \setminus (1-\delta)Q_{\alpha_2}} \sigma_\tau) R_2(1_{Q_{\alpha_3}} \sigma_\tau) \, d\omega_\tau.
$$

The first term of the right-hand side is handled by the $\delta$-separation between $Q_{\alpha_1}$ and $(1-\delta) Q_{\alpha_2}$, as well as between $Q_{\alpha_1}$ and $(1-\delta) Q_{\alpha_3}$, together with the $A_2$ condition $AB \leq 1$ to obtain

$$
\left| \int_{Q_{\alpha_1}} R_2(1_{(1-\delta)Q_{\alpha_2}} \sigma_\tau) R_2(1_{(1-\delta)Q_{\alpha_3}} \sigma_\tau) \, d\omega_\tau \right| \leq \frac{1}{\delta^{2n}} \int_{Q_{\alpha_1}} |Q_{\alpha_2}|_{\sigma_\tau} |Q_{\alpha_3}|_{\sigma_\tau} \, d\omega_\tau \leq \frac{1}{\delta^{2n}} |Q_{\alpha_3}|_{\sigma_\tau} A_2 \leq \frac{1}{\delta^{3n}} |Q_{\alpha_3}|_{\sigma_\tau} A_2
$$

and since for each fixed $\alpha_3$, there are at most $3^{2n}$ pairs $(\alpha_1, \alpha_2)$, we can sum to obtain the bound $C \frac{1}{\delta^{3n}} |Q|_{\sigma_\tau}$.
To handle the terms involving a halo $Q_{\alpha_j} \setminus (1 - \delta) Q_{\alpha_j}$ we use Lemma 35 together with the norm constant $\mathfrak{N}_{R_2} = \mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau)$. For example,

$$\left| \int_{Q_{\alpha_1}} R_2 \left(1_{Q_{\alpha_2} \setminus (1 - \delta)Q_{\alpha_2}} \sigma_\tau \right) R_2 \left(1_{Q_{\alpha_2}} \sigma_\tau \right) d\omega_\tau \right|$$

$$\leq \left( \left| \int_{Q_{\alpha_1}} R_2 \left(1_{Q_{\alpha_2} \setminus (1 - \delta)Q_{\alpha_2}} \sigma_\tau \right)^2 d\omega_\tau \right| \right)^{\frac{1}{2}} \left( \left| \int_{Q_{\alpha_1}} R_2 \left(1_{Q_{\alpha_2}} \sigma_\tau \right)^2 d\omega_\tau \right| \right)^{\frac{1}{2}}$$

$$\leq \mathfrak{N}_{R_2} \sqrt{|Q_{\alpha_2} \setminus (1 - \delta) Q_{\alpha_2}|_{\sigma_\tau}} \mathfrak{N}_{R_2} \left( |Q_{\alpha_2}|_{\sigma_\tau} \right)^{\frac{1}{2}} = \left( \mathfrak{N}_{R_2} \right)^2 \frac{C}{\sqrt{\ln \frac{1}{\delta}}} \sqrt{|Q_{\alpha_2}|_{\sigma_\tau}}$$

and again we can sum to obtain the bound $(\mathfrak{N}_{R_2})^2 \frac{C}{\ln \frac{1}{\delta}} |Q|_{\sigma_\tau}$ because the indices $\alpha_j$ are distance one from each other. The other terms are handled similarly and we thus obtain in this seventh case that

$$\sigma_\tau$$

interpretation of uniformly over all admissible truncations of the independence of truncations mentioned above, the above arguments actually prove that (5.10) holds for any choice of $\delta_1 \in (0, 1)$, where the constant $C_*$ arises in (5.6).

Now we turn to the case of a general cube $Q$. In this case we first fix $M \in \mathbb{N}$ large to be chosen later, and write $Q$ as a union of roughly $2^M \ln n$ dyadic subcubes $\{Q_s \}_{s}$ of side length $\delta_2 \equiv \delta(\bar{Q}) > 0$, in such a way that the remaining portion of $Q$ is contained in the $5\delta_2$-halo of $Q$. Then the above argument shows that the testing condition holds except for the terms that arise from the halo. But by Lemma 35 these leftover terms in $\left( \int_Q |R_2(1_Q \sigma_\tau)|^2 d\omega_\tau \right)^{\frac{1}{2}}$ are dominated by $C \frac{1}{\sqrt{\ln \frac{1}{\delta}}} \mathfrak{N}_{R_2} \sqrt{|Q|_{\sigma_\tau}}$, so that altogether we obtain that the continuous testing constant satisfies

$$\mathfrak{T}_{R_2}(\sigma_\tau, \omega_\tau) \leq C_2 \mathfrak{T}_{R_2}(\sigma_\tau, \omega_\tau) + C \frac{1}{\sqrt{\ln \frac{1}{\delta}}} \mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau)$$

$$+ C \delta_2 \left( C_\star + C_* C_\tau + \frac{1}{\delta_2^2} A_2 + \frac{\left( \mathfrak{N}_{R_2} \right)^2}{\sqrt{\ln \frac{1}{\delta}}} \right)^{\frac{1}{2}} + C \frac{\mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau)}{\sqrt{\ln \frac{1}{\delta}}}$$

$$\leq C_2 \sigma_\tau \sqrt{C_\star} + C \delta_2 \frac{1}{\delta_1^{\frac{3}{2}}} \sqrt{A_2} + \left( \frac{C_\delta_2}{\sqrt{\ln \frac{1}{\delta}}} + \frac{C}{\sqrt{\ln \frac{1}{\delta}}} \right) \mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau).$$

Note both weights $\sigma_\tau, \omega_\tau$ are bounded step functions, and so by the boundedness of the principal value interpretation of $R_2$ on Lebesgue spaces, we have

$$\mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau) \leq \|\sigma_\tau\|_{\infty} \|\omega_\tau\|_{\infty}.$$  

Thus by boundedness of maximal truncations (see e.g. [Ste2, Proposition 1 page 31]) together with the independence of truncations mentioned above, the above arguments actually prove that (5.10) holds uniformly over all admissible truncations of $R_2$, which is the hypothesis used in [AlSaUr]. Thus noting Definition 24 we can apply the main theorem in [AlSaUr] to obtain

$$\mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau) \leq C \sqrt{A_2(\sigma_\tau, \omega_\tau)} + C \mathfrak{T}_{R_2}(\sigma_\tau, \omega_\tau) + C \mathfrak{T}_{R_2}(\omega_\tau, \sigma_\tau)$$

$$\leq C \sqrt{A_2(\sigma_\tau, \omega_\tau)} + 2 \left( C \delta_2 \sigma_\tau \sqrt{C_\tau} + C \delta_2 \frac{1}{\delta_1^{\frac{3}{2}}} \sqrt{A_2(\sigma_\tau, \omega_\tau)} + \left( \frac{C_\delta_2}{\sqrt{\ln \frac{1}{\delta}}} + \frac{C}{\sqrt{\ln \frac{1}{\delta}}} \right) \mathfrak{N}_{R_2}(\sigma_\tau, \omega_\tau) \right).$$
for any admissible truncation of \( R_2 \). Thus with \( \delta_2 > 0 \) chosen sufficiently small that \( \frac{C_2}{\sqrt{\ln \frac{1}{\delta_2}}} < \frac{1}{4} \), and then \( \delta_1 > 0 \) chosen sufficiently small that \( \frac{C_1}{\sqrt{\ln \frac{1}{\delta_1}}} < \frac{1}{4} \), an absorption completes the proof that the norm inequality for \( R_2 \) holds (recall that truncations of \( R_2 \) are a priori bounded).

We have thus proved the following special case of Theorem 4 for the individual Riesz transforms \( R_1 \) and \( R_2 \).

**Proposition 38.** For every \( \Gamma > 1 \) and \( 0 < \tau < 1 \), there is a pair of positive weights \((\sigma, \omega)\) in \( \mathbb{R}^n \) satisfying

\[
\int_{\mathbb{R}^n} \left| R_1 \left( 1_{[0,1]^n} \sigma \right)(x) \right|^2 (x) \, d\omega(x) \geq \Gamma \int_{[0,1]^n} d\sigma(x),
\]

\[
\int_I |R_2 1_I \sigma(x)|^2 \, d\omega(x) \leq \int_I d\sigma(x), \quad \text{for all cubes } I \in \mathcal{P}^n,
\]

\[
\int_I |R_2 1_I \omega(x)|^2 \, d\sigma(x) \leq \int_I d\omega(x), \quad \text{for all cubes } I \in \mathcal{P}^n,
\]

\[
\left( \frac{1}{|I|} \int_I d\sigma \right) \left( \frac{1}{|I|} \int_I d\omega \right) \leq 1, \quad \text{for all cubes } I \in \mathcal{P}^n,
\]

\[
1 - \tau < \frac{E_{1I} \sigma}{E_{1J} \sigma} \frac{E_{1I} \omega}{E_{1K} \omega} < 1 + \tau, \quad \text{for arbitrary adjacent cubes } J, K \text{ in } \mathbb{R}^n.
\]

The argument used in proving this proposition also shows that in the main theorem in \( \text{[AlSaUr]} \), the testing may be carried out over only cubes in any fixed dyadic grid \( \mathcal{D} \), and here is one possible formulation of this improvement.

**Theorem 39.** Suppose \( 0 \leq \alpha < n \), and let \( T^\alpha \) be an \( \alpha \)-fractional Calderón-Zygmund singular integral operator on \( \mathbb{R}^n \) with a smooth \( \alpha \)-fractional kernel \( K^\alpha \). Assume that \( \sigma \) and \( \omega \) are doubling measures on \( \mathbb{R}^n \). Set \( T^\alpha f = T^\alpha (f \sigma) \) for any smooth truncation of \( T^\alpha \). Finally fix a dyadic grid \( \mathcal{D} \) on \( \mathbb{R}^n \).

Then the best constant \( \mathcal{N}_{T^\alpha} \) in the weighted norm inequality

\[
\| T^\alpha f \|_{L^2(\omega)} \leq C \| f \|_{L^2(\sigma)},
\]

satisfies

\[
(5.11) \quad \mathcal{N}_{T^\alpha} \leq C_{\alpha,n} \left( \sqrt{A^2_{T^\alpha}} + \mathcal{T}^P_{T^\alpha} + \mathcal{T}^D_{T^\alpha} \right),
\]

where the constant \( C_{\alpha,n} \) also depends on the Calderón-Zygmund kernel, \( A^2_{T^\alpha} \) is the classical Muckenhoupt constant, and \( \mathcal{T}^P_{T^\alpha}, \mathcal{T}^D_{T^\alpha} \) are the dual \( \mathcal{D} \)-dyadic testing constants.

In order to complete the proof of Theorem 3 we need to consider iterated Riesz transforms.

### 6. Iterated Riesz Transforms

Throughout Section 4 and 5 we considered Riesz transforms of order 1. However our results extend to arbitrary iterated Riesz transforms of odd order in \( \mathbb{R}^n \). We will extend the results of Section 4 to their appropriate analogues to make the reasoning of Section 5 hold for the appropriate iterated Riesz transforms, and we begin by establishing the following theorem.

**Theorem 40.** The odd order pure iterated Riesz transforms \( R_1^{2m+1} \) are unstable on \( \mathbb{R}^n \) for pairs of doubling measures under 90° rotations in any coordinate plane. In fact, there exists a measure pair of doubling measures on which \( R_1^{2m+1} \) is unbounded, and all iterated Riesz transforms of order \( 2m+1 \) that are not a pure power of \( R_1 \), are bounded.

**Proof.** Recall the notation \( T_\alpha f = T(f \sigma) \). We begin first by considering Riesz transforms of arbitrary order, even or odd. Using the identity

\[
R_1^2 + \ldots + R_n^2 = -I,
\]

and \( N \geq 2 \) we have for an arbitrary positive measure \( \sigma \) that

\[
(R_1^N)_\sigma = (R_1^{N-2} R_1^2)_\sigma = -(R_1^{N-2})_\sigma - \sum_{j=2}^n (R_1^{N-2} R_j^2)_\sigma.
\]
Iteration then yields for \( N \geq 1 \),
\[
(R_1^N)_{\sigma} = \begin{cases} 
\pm I_{\sigma} + \sum_{k=0}^m \left[ \pm \sum_{j=2}^n (R_1^{N-2k} R_j^2)_{\sigma} \right] & \text{if } N = 2m \text{ is even } \\
\pm (R_1)_{\sigma} + \sum_{k=0}^m \left[ \pm \sum_{j=2}^n (R_1^{N-2k} R_j^2)_{\sigma} \right] & \text{if } N = 2m + 1 \text{ is odd } 
\end{cases}.
\]

For the weight pairs \((\sigma_\tau, \omega_\tau)\) constructed in Section 5 and with \( N = 2m + 1 \) odd, the second line in (6.2) yields
\[
\left\| (R_1^N)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} \geq \left\| (R_1)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} - \sum_{k=0}^m \sum_{j=2}^n \left\| (R_1^{N-2k} R_j^2)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} \geq \Gamma - \sum_{k=0}^m \sum_{j=2}^n \left\| (R_1^{N-2k} R_j^2)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)},
\]
where \( \Gamma \) is the constant in the construction of the weight pair \((\sigma_\tau, \omega_\tau)\). Note that the operator norm dominates the testing constant, which was shown to exceed \( \Gamma \).

We now claim that the sum of the operator norms on the right hand side is bounded independently of \( \Gamma \), i.e.
\[
\sum_{k=0}^m \sum_{j=2}^n \left\| (R_1^{N-2k} R_j^2)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} = O(1).
\]

In fact if \( j \geq 2 \) and \( R^{\alpha} = R_1^{\alpha_1} R_2^{\alpha_2} \ldots R_n^{\alpha_n} \) with \( \alpha_j > 0 \), then by Lemma 28 part (3),
\[
\limsup_{k \to \infty} \int \left| R_j R^{\alpha} s_k(\sigma_\tau, \omega_\tau) \right|^2 \, dx = \limsup_{k \to \infty} \int \left| R_j s_k \right|^2 \left( R_j R^{\alpha} s_k(\sigma_\tau, \omega_\tau) \right) \, dx \leq \limsup_{k \to \infty} \int \left| R_j s_k \right|^2 \left( R_j R^{\alpha} \right) \, dx = O(1),
\]
for all \( N \in \mathbb{N} \).

The reasoning in Proposition 30 and Lemma 37 shows that iterated Riesz transforms of order \( N \) which are not pure powers of \( R_1 \) have dyadic testing constants on the weight pairs \((\sigma_\tau, \omega_\tau)\) that are \( O(1) \). Then Theorem 39 shows that the operator norms of such operators, including \( R_1^{N-2k} R_j^2 \), are \( O(1) \), which proves our claim, and completes the proof of the second assertion of the theorem. The first assertion regarding \( R_1^{2m+1} \) now follows from the fact that a rotation in the \((x_1, x_2)\)-plane interchanges \( R_1^{2m+1} \) and \( R_j^{2m+1} \).

\[\square\]

The key to our proof of Theorem 40 is the construction of weight pairs \((\sigma_\tau, \omega_\tau)\) satisfying the inequality
\[
\left\| (R_1^N)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} \geq \Gamma \text{ for } \Gamma \text{ arbitrarily large,}
\]
when \( N \) is odd. In fact, the inequality (6.3) actually fails for the weight pairs we construct when \( N \) is even. Indeed, from the first line in (6.2), and the fact that the proof of Theorem 40 shows that
\[
\sum_{k=0}^m \sum_{j=2}^n \left\| (R_1^{N-2k} R_j^2)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} = O(1),
\]
we get
\[
\left\| (R_1^N)_{\sigma_\tau} \right\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} \leq \|I_{\sigma}\|_{L^2(\sigma_\tau) \to L^2(\omega_\tau)} + O(1).
\]
The right hand side of the above display is bounded since the operator norm of \( I_{\sigma} \) is bounded by \( A_2(\sigma_\tau, \omega_\tau) \): indeed, when \( \sigma \) and \( \omega \) are weights (where a locally finite measure \( u(x) \, dx \) is called a weight if \( 0 < u(x) < \infty \) for all \( x \in \mathbb{R}^n \)), we have \( \|\sigma \omega\|_{\infty} \leq A_2(\sigma, \omega) \) by the Lebesgue differentiation theorem, and so
\[
\|I_{\sigma}f\|_{L^2(\omega)}^2 = \int_{\mathbb{R}^n} f^2 \sigma^2 \omega \leq A_2(\sigma, \omega) \int_{\mathbb{R}^n} f^2 \sigma = \|f\|_{L^2(\sigma)}^2.
\]
Moreover, it is easily shown that \( \|I_\sigma\|_{L^2(\sigma) \to L^2(\omega)} = A_2(\sigma, \omega) \) for arbitrary weights \( \sigma \) and \( \omega \). Thus \( R_1^N \) must then satisfy the testing conditions for the measure pair \( (\sigma, \omega) \).

In the next subsection we show that every odd order iterated Riesz transform \( R^\beta = R_1^{\beta_1}R_2^{\beta_2}...R_n^{\beta_n} \) is unstable under rotations, by showing that \( R_1^{\beta_1}R_2^{\beta_2}...R_n^{\beta_n} \) is some rotation of \( R^{(\beta,0,...,0)} \) whenever \( \beta \neq |\beta|e_k \) for some \( k \). When \( \beta = |\beta|e_k \) some \( k \), then we may assume without loss of generality that \( k = 2 \).

### 6.1. Rotations

Let \( \beta \) be a multiindex of length \( |\beta| = N \). The symbol of the iterated Riesz transform \( R^\beta = R_1^{\beta_1}R_2^{\beta_2}...R_n^{\beta_n} \) is \( i^N \sum_{|\beta|=N} \frac{\partial^\beta}{\partial \xi^\beta} \). We already know that \( R^{(N,0,...,0)} \) is unstable, and the following lemma will be used to show all \( R^\beta \) are unstable.

**Lemma 41.** If \( P(\xi) \) is a nontrivial homogeneous polynomial of degree \( N \) that doesn’t contain the monomial \( \xi_1^N \), then there is a set of rotations of full-measure \( \Lambda \), and for any rotation \( \Theta \in \Lambda \), we have \( \xi = \Theta \eta \) is such that \( P(\Theta \eta) \) contains the monomial \( \eta_1^N \).

**Proof.** In dimension \( n = 2 \), we have
\[
P(\xi_1, \xi_2) = \sum_{m=1}^{N} c_m \xi_1^m \xi_2^{N-m},
\]
where not all \( c_m = 0 \), and the restriction of this polynomial to the unit circle cannot vanish identically (otherwise \( P \) itself would vanish identically by homogeneity, a contradiction). Thus there is \( \theta \in [0,2\pi) \) such that
\[
0 \neq P(\cos \theta, \sin \theta) = \sum_{m=1}^{N} c_m \cos^m \theta \sin^{N-m} \theta.
\]
However, if we make the rotational change of variable, i.e.
\[
\begin{pmatrix}
\xi_1 \\
\xi_2
\end{pmatrix}
= \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
\eta_1 \\
\eta_2
\end{pmatrix},
\]
then
\[
P(\xi_1, \xi_2) = \sum_{m=1}^{N} c_m \xi_1^m \xi_2^{N-m} = \sum_{m=1}^{N} c_m (\eta_1 \cos \theta - \eta_2 \sin \theta)^m (\eta_1 \sin \theta + \eta_2 \cos \theta)^{N-m}
= \eta_1^N \sum_{m=1}^{N} c_m \cos^m \theta \sin^{N-m} \theta + \sum_{\beta \neq e_1: |\beta| = N} \eta^\beta f_\beta(\theta)
\]
where \( \sum_{m=1}^{N} c_m \cos^m \theta \sin^{N-m} \theta \neq 0 \). The case \( n \geq 3 \) is similar. \( \square \)

### 6.2. Completion of proofs of main theorems

To complete the proof of Theorem 4 we use the above Lemma, together with Proposition 58, and we see that any iterated Riesz transform \( R^\beta \) of odd order \( N = |\beta| \) with \( \beta \neq (N,0,...,0) \), is bounded on the higher dimensional analogue of the weight pair \( (\sigma, \omega) \) constructed in Proposition 58, and can be rotated into a sum \( S \) of iterated Riesz transforms that includes \( R^{(N,0,...,0)} \), and hence \( S \) is unbounded on the weight pair \( (\sigma, \omega) \). Since stability under rotational change of variables is unaffected by rotation of the operator, this completes our proof that all iterated Riesz transforms \( R^\beta \) of odd order are unstable under rotational changes of variable, even when the measures are doubling with exponents \( \theta_{\text{domb}} \) arbitrarily close to the exponent \( n \) of Lebesgue measure, i.e. \( \lambda_{\text{adj}} \) is arbitrarily close to 1. This completes the proof of the main Theorem 4.

To prove Theorem 5, suppose \( R^\beta \) is an odd order iterated Riesz transform; without loss of generality, assume that \( R^\beta \neq R_1^{\beta_1} \). Then by Lemma 41 there is a set \( \Lambda \) of rotations of full measure such that for each \( \Theta \in \Lambda \), \( \Theta \) rotates \( R^\beta \) to \( c(\Theta) R_1^{\beta_1} \) plus mixed iterated Riesz transforms, where \( c(\Theta) \neq 0 \). Then our construction yields a weight pair \( (\sigma, \omega) \) for which the norm inequality for \( R^\beta \) is bounded, but the norm inequality for the rotated operator can be made arbitrarily large.
7. Appendix

We begin by using the counterexamples in [LaSaUr] to show that the Hilbert transform is two weight norm biLipschitz unstable on $S_{plb}$. Then we demonstrate that the notion of stability that is maximal for preserving the classical $A_2$ condition, is that of biLipschitz stability. Finally, we give the details for arguments surrounding classical doubling and classical $A_2$ which were omitted from [NaVo].

7.1. BiLipschitz instability of the Hilbert transform for arbitrary weight pairs. Here we show that the Hilbert transform $H$ is two weight norm unstable under biLipschitz transformations. We consider the measure pairs $(\sigma, \omega)$ and $(\tilde{\sigma}, \omega)$ constructed in [LaSaUr], where $(\sigma, \omega)$ satisfies the two weight norm inequality for $H$, while $(\tilde{\sigma}, \omega)$ does not, although it continues to satisfy the two-tailed Muckenhoupt $A_2$ condition. The measure $\omega$ is the standard Cantor measure on $[0,1]$ supported in the middle-third Cantor set $E$. The measures $\sigma = \sum_{k,j} s_j^k \delta_{z_j^k}$ and $\tilde{\sigma} = \sum_{k,j} s_j^k \delta_{\tilde{z}_j^k}$ are sums of weighted point masses located at positions $z_j^k$ and $\tilde{z}_j^k$ within the component $G_j^k$ removed at the $k^{th}$ stage of the construction of $E$, and satisfy

$$0 < c_1 \frac{\text{dist}(z_j^k, \partial G_j^k)}{|G_j^k|}, \frac{\text{dist}(\tilde{z}_j^k, \partial G_j^k)}{|G_j^k|} < c_2 < 1,$$

independent of $k,j$. See [LaSaUr] for notation and proofs.

It remains to construct a biLipschitz map $\Phi : R \to R$ such that $(\tilde{\sigma}, \omega) = (\Phi_* \sigma, \Phi_* \omega)$. For this, we first define biLipschitz maps $\Phi : G_j^k \to \tilde{G}_j^k$ so that $\Phi$ fixes the endpoints of $G_j^k$ and $\tilde{z}_j^k = \Phi(z_j^k)$, and note that this can be done with bounds independent of $k,j$ by (7.1). Now we extend the definition of $\Phi$ to all of $R$ by the identity map, and it is evident that $\Phi$ is biLipschitz and pushes $(\sigma, \omega)$ forward to $(\tilde{\sigma}, \omega)$.

7.2. Beyond biLipschitz maps for $A_2$ stability. Here we initiate an investigation of how general a map can be, and still preserve the two weight $A_2$ condition for all pairs of measures $(\sigma, \omega)$. We begin by defining some of the terminology we will use in this subsection.

Definition 42. Let $\mu$ be a locally finite positive Borel measure on $\mathbb{R}^n$, and let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be a Borel measurable function. We define the pushforward of the measure $\mu$ by the map $\Phi$ as the unique measure $\Phi_* \mu$ such that

$$\int_E \Phi_* \mu = \int_{\Phi^{-1}(E)} \mu, \quad \text{for all Borel sets } E \subset \mathbb{R}^n.$$

In the case $d\mu(x) = w(x) \, dx$ is absolutely continuous, its pushforward for $\Phi$ sufficiently smooth is given by

$$(\Phi_* \mu)(y) \equiv w(\Phi(y)) \left| \det \frac{\partial \Phi}{\partial x}(y) \right|.$$

Definition 43. A map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is $A_2$-stable, if there exists a constant $C > 0$ such that for every pair of locally finite positive Borel measures $\sigma, \omega$ we have

$$A_2(\Phi_* \sigma, \Phi_* \omega) \leq CA_2(\sigma, \omega).$$

Definition 44. A map $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ (not necessarily invertible) is shape-preserving if there exists $K \geq 1$ such that for every cube $Q \subset \mathbb{R}^n$ we can find cubes $Q_{\text{small}}$ and $Q_{\text{big}}$ with the properties,

$$Q_{\text{small}} \subset \Phi^{-1}(Q) \subset Q_{\text{big}} \quad \frac{\ell(Q_{\text{big}})}{\ell(Q_{\text{small}})} \leq K.$$

Note that homeomorphisms on the real line are automatically shape-preserving, as are quasiconformal maps in $\mathbb{R}^n$ [AslwMa Lemmam 3.4.5].

Theorem 45. Let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ be shape-preserving and Borel-measurable. Then the following two conditions are equivalent:

1. There exists a constant $C_1 > 0$ such that $|\Phi^{-1}(Q)| \leq C_1 |Q|$ for every cube $Q$.
2. $\Phi$ is $A_2$-stable.

Remark 46. If $\Phi$ is sufficiently regular that the usual change of variables formula holds, e.g. $\Phi^{-1}$ is locally Lipschitz, then condition (1) becomes $|\det D\Phi^{-1}| \lesssim 1$. 


Proof. Assume condition (1) holds where $\Phi$ is shape-preserving with constant $K$, and let $Q$ be an arbitrary cube in $\mathbb{R}^n$. Then

$$A_2 (\Phi \ast \sigma, \Phi \ast \omega) = \sup_Q \left( \frac{\int_Q d\Phi \ast \sigma}{|Q|} \right) \left( \frac{\int_Q d\Phi \ast \omega}{|Q|} \right) = \sup_Q \left( \frac{\int_{\Phi^{-1}(Q)} d\sigma}{|Q|} \right) \left( \frac{\int_{\Phi^{-1}(Q)} d\omega}{|Q|} \right) \leq C_2^2 \sup_Q \left( \frac{\int_{\Phi^{-1}(Q)} d\sigma}{|Q|} \right) \left( \frac{\int_{\Phi^{-1}(Q)} d\omega}{|Q|} \right) \leq C_2^2 K^2 A_2 (\sigma, \omega).$$

Conversely, if condition (2) holds, then with both measures $\sigma$ and $\omega$ equal to Lebesgue measure, and for any cube $Q$, we have,

$$\left( \frac{\Phi^{-1}(Q)}{|Q|} \right)^2 = \left( \frac{\int_{\Phi^{-1}(Q)} dx}{|Q|} \right) \left( \frac{\int_{\Phi^{-1}(Q)} dx}{|Q|} \right) \leq C.$$

Remark 47. If the pair $(\Phi \ast \sigma, \Phi \ast \omega)$ is in $A_2$ for the single choice of weights $d\sigma (x) = d\omega (x) = dx$, then the above proof shows that $\Phi$ preserves all $A_2$ pairs under the side assumption of shape-preservation.

Corollary 48. Assume $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is a shape-preserving invertible Lipschitz map with $\|D\Phi\|_{\infty} \leq 1$. Then $\Phi$ is $A_2$-stable if and only if $\Phi$ is biLipschitz.

Proof. By Theorem 15 and Remark 40 we see that $\Phi$ is $A_2$-stable if and only if $|\det D\Phi| \geq 1$. But then $1 \leq C |\det D\Phi| \leq C' |D\Phi|^n$, together with $\|D\Phi\|_{\infty} \leq 1$, shows that $\Phi$ is $A_2$-stable if and only if $\Phi$ is biLipschitz.

7.3. Modification of Transplantation to get Classical Doubling and $A_2$. In Section 3 we constructed functions $v, u$ on a cube $Q^0$ such that both $v, u$ are dyadically $\tau$-flat on $Q_0$ and $A_2^\text{dyadic} (v, u; Q^0) \leq 1$. However, dyadic doubling and dyadic $A_2$ do not imply continuous doubling or classical $A_2$ in $Q^0$. As such, we will need to modify the transplantation argument to smooth out $v, u$ into weights $v', u'$ which are classically doubling and satisfy the classical $A_2$ condition, as done in [NaVo].

We will describe how to attain $u'$ from $u$, as the process for $v'$ and $v$ is identical. Recall in Proposition 30 we define $u$ by

$$u = (E_{Q^0; U}) 1_{Q^0} + \sum_{t=0}^{m} \sum_{Q \in K_t} \langle U, h^\text{horizontal} \rangle \frac{1}{\sqrt{|S(Q)|}} s_{k_{t+1}}^{Q, \text{horizontal}},$$

where $s_{k_{t+1}}$ is constant on cubes in $K_{t+1}$.

Define the grid $\hat{K}$ from $K$ inductively as follows. First set $\hat{K}_0 \equiv K_0$. Now given $Q \in \hat{K}_t$, a cube $R \in K_{t+1}$ is called a transition cube for $Q$ if $Q = \pi_K R$ and $(\partial \pi_D R) \cap \partial Q)$ is non-empty; as such define $\hat{K}_{t+1}$ to consist of the cubes $P \in K_{t+1}$ such that $\pi_K P \in \hat{K}_t$ and $P$ is not a transition cube. Finally, set $\hat{K} \equiv \bigcup_t \hat{K}_t$.

One can see that $\hat{K}$ consists of the cubes in $K$ not contained in a transition cube. This implies that if $R$ is a transition cube, then $\pi_K R \in K$. It also implies that no two transition cubes have overlapping interiors. Visually, the union of the transition cubes for a cube $Q$ forms a “halo” for $Q$. Recall that two distinct dyadic cubes in $\hat{D}$ of the same size are adjacent if their boundaries intersect, even if only at a point. Note that two cubes in $\hat{K}$ are adjacent, then they must have the same $K$-parent, while adjacent transition cubes must be close to each other in the tree distance of $K$. The proof of the following lemma is left to the reader, who is encouraged to draw a picture. It helps to note that in $\mathbb{R}$, if two transition intervals $R_1$ and $R_2$ are at levels $s$ and $s + 2$, then there must be a transition interval $R$ at level $s + 1$ such that $R$ lies between $R_1$ and $R_2$.

Lemma 49. Let $R_1 \in K_t$ be a transition cube.

1. If $R_2 \in K_t$ is a transition cube such that the interiors of $R_2$ and $R_1$ are disjoint, but not their closures, then $t \in \{s - 1, s, s + 1\}$.

2. If $K \in K_t$ is such that the interiors of $K$ and $R_1$ are disjoint, but not their closures, then $t \in \{s - 1, s\}$. And if $t = s$, then $\pi_K K = \pi_K R_1$. 


With this in mind, given \( Q \in \hat{K}_t \), define

\[
u_{\ell, k_t+1}^Q(x) = \begin{cases} s^Q_{\ell, k_t+1}(x) & \text{if } x \text{ is not contained in a transition cube for } Q \text{.} \\ 0 & \text{otherwise} \end{cases}
\]

Now we may define

\[
u'_\ell \equiv (E_{Q^0}U)1_{Q^0} + \sum_{t=0}^{\ell-1} \sum_{Q \in \hat{K}_t} (U, h_{S(Q)}^{\text{horizontal}}) \frac{1}{|S(Q)|} \nu_{\ell, k_t+1}^Q(x), \quad 0 \leq \ell \leq m,
\]

\[
u' \equiv \nu'_m \text{ and } \nu'' \equiv \nu'_m.
\]

Given \( x \in Q^0 \) and \( \ell < m \), if we define

\[
t(x) \equiv \begin{cases} t & \text{if } x \text{ is contained in a transition cube belonging to } K_t \text{ for some } t < \ell \\ \ell & \text{otherwise} \end{cases}
\]

then pointwise we have

\[
u'_\ell(x) = (E_{Q^0}U)1_{Q^0} + \sum_{t=0}^{t(x)-1} \sum_{Q \in \hat{K}_t} (U, h_{S(Q)}^{\text{horizontal}}) \frac{1}{|S(Q)|} s_{\ell, k_t+1}^Q(x), \quad 0 \leq \ell \leq m.
\]

The function \( u' \) is nearly a transplantation of \( U \), as exhibited by the following lemma, whose proof we leave to the reader. The reader should note that for each cube contained in a transition cube, the value of \( u'_\ell \) is equal to its average on the transition cube containing it.

**Lemma 50.** Let \( K \) be as above.

1. If \( P \in K \) is not contained in a transition cube, then \( E_P u'_\ell = E_{S(P)} U \).
2. If \( P \in K \) is contained in a transition cube \( R \), then \( E_P u'_\ell = E_{S(\pi_K R)} U \).
3. If \( P \in D \) is a cube for which \( K_{t+1} \subseteq P < K_t \) where \( K_{t+1} \subseteq K_t \) and \( K_t \subseteq K \), then \( E_P u'_\ell = E_{K_t} u'_t \).

**7.3.1. Classical Doubling.**

**Lemma 51.** If \( P_1, P_2 \) are adjacent dyadic subcubes of \( Q^0 \), then \( \frac{E_{P_1} u'_\ell}{E_{P_2} u'_\ell} \in (1 - \tau, 1 + \tau) \). Similarly for \( \nu'' \).

**Proof of Lemma 51.** Let \( P_1, P_2 \) be adjacent dyadic subcubes of \( Q^0 \). By Lemma 50 part (3), it suffices to check the case when \( P_1, P_2 \in \hat{K} \). We consider various cases.

Case 1: neither \( P_1 \) nor \( P_2 \) is contained in a transition cube, i.e. both belong to \( \hat{K} \). Then \( P_1 \) and \( P_2 \) must have common \( K \)-parent, meaning

\[ \pi_D S(P_1) = S(\pi_K P_1) = S(\pi_K P_2) = S(P_2) \]

and so \( S(P_1) \) and \( S(P_2) \) must be equal or dyadic siblings. By the first formula of Lemma 50 we get

\[ \frac{E_{P_1} u'_\ell}{E_{P_2} u'_\ell} \in (1 - \tau, 1 + \tau). \]

Case 2: exactly one of the cubes, say \( P_1 \), is contained in a transition cube \( R_1 \). Since \( P_2 \) is not in a transition cube, then the only way for \( P_1, P_2 \) to be adjacent is for both to have the same \( K \)-parent. And since \( P_2 \) is not contained in a transition cube, then \( R_1 \) must in fact equal \( P_1 \), i.e. \( P_1 \) is a transition cube: indeed, if \( P_1 \) were a level below \( R_1 \) in the grid \( K \), then the only way \( P_2 \) can be adjacent to \( P_1 \) is by being in a transition cube adjacent to \( R_1 \) or in \( R_1 \) itself, but the latter can’t happen by assumption on \( P_2 \).

Altogether, the above yields that \( S(\pi_K P_1) = S(\pi_K P_2) = \pi_D S(P_2) \). Thus by Lemma 50 parts (1) and (2), dyadic \( \tau \)-flatness of \( U \), and the fact that \( P_1 \) is a transition cube, we have

\[ \frac{E_{P_1} u'_\ell}{E_{P_2} u'_\ell} = \frac{E_{S(\pi_K P_1)} U}{E_{S(P_2)} U} = \frac{E_{\pi_D S(P_2)} U}{E_{S(P_2)} U} \in (1 - \tau, 1 + \tau). \]

Case 3: both \( P_1 \) and \( P_2 \) are contained within transition cubes, say \( R_1 \) and \( R_2 \) respectively. Using Lemma 50 it suffices to show the ratio

\[ \frac{E_{P_1} u'_\ell}{E_{P_2} u'_\ell} = \frac{E_{S(\pi_K R_1)} U}{E_{S(\pi_K R_2)} U} \]

lies between \( 1 - \tau \) and \( 1 + \tau \). Note adjacency of \( P_1, P_2 \) implies \( R_1 \) and \( R_2 \) have disjoint interiors, but not closures, or are equal.
Case 3a: $R_1 = R_2$. Then we get $\frac{E_{P,u'}}{E_{P,u}} = 1$.

Case 3b: $R_1$ and $R_2$ are of the same sidelength, but $R_1 \neq R_2$. Then both $R_1$ and $R_2$ are adjacent, and so $S(\pi_K R_1)$ and $S(\pi_K R_2)$ must be equal or dyadic siblings. In either case, by the formula above $\frac{E_{P,u'}}{E_{P,u}} \in (1 - \tau, 1 + \tau)$.

Case 3c: $R_1$ and $R_2$ are of different sidelengths, say $\ell(R_1) > \ell(R_2)$. Since $P_1, P_2$ are adjacent then $R_1$ and $R_2$ have disjoint interiors, but not closures. It follows that if $R_1 \in K_t$, then $R_2 \in K_{t+1}$ by Lemma 49.

Thus $R_1$ is adjacent to $\pi_K R_2$. In fact, since $R_1$ is a transition cube but $\pi_K R_2$ is not, then by Lemma 49 (2) we have $\pi_K R_1 = \pi_K^{(2)} R_2$ and so

$$S(\pi_K R_1) = S(\pi_K^{(2)} R_2) = \pi_P S(\pi_K R_2).$$

Thus

$$\frac{E_{P,u'}}{E_{P,u}} = \frac{E_S(\pi_K R_1) U}{E_S(\pi_K R_2) U} = \frac{E_P S(\pi_K R_2) U}{E_S(\pi_K R_2) U} \in (1 - \tau, 1 + \tau).$$

This completes the proof.

Showing $u'$ has relative adjacency constant $1 + o(1)$ as $\tau \to 0$ on $Q^0$ follows from Lemma 51 and a standard argument, and similarly for $v'$.

7.3.2. Classical $A_2$. We will use reasoning as in the previous subsection on continuous doubling, and will use of the same notation as that subsection. Just as for $v$, it is easy to check that $A_2^{dyadic}(u', u'; Q^0) \leq 1$.

Lemma 52. Suppose a cube $I \subset \mathbb{R}^n$ is the union of $2^n$ pairwise adjacent dyadic subcubes $\{I_k\}_{k=1}^{2^n}$ of $Q^0$, each of equal sidelength. Then

$$(E_I u')(E_I v') \leq 81.$$

Proof. Write $\{I_k\}_{k=1}^{2^n} = \{I_k^{\text{rigid}}\}_{k=1}^{a} \cup \{I_k^{\text{transition}}\}_{k=1}^{b}$, where the first collection consists of all the $I_k$’s which are not contained in a transition cube, and the second collection consists of those cubes contained in a transition cube.

Regarding the collection $\{I_k^{\text{rigid}}\}_{k=1}^{a}$, let $t$ be the largest integer such that for each $1 \leq k \leq a$, there exists $K_k \in K_t$ such that $I_k^{\text{rigid}} \subset K_k$. Note the choice of $K_k$ is unique for each $k$. Regarding the second collection $\{I_k^{\text{transition}}\}_{k=1}^{b}$, let $R_k$ be the transition cube containing $I_k^{\text{transition}}$. Since $\{I_k^{\text{transition}}\}_{k=1}^{b}$ consists of pairwise adjacent cubes, then any two cubes in $\{R_k\}_{k=1}^{b}$ have disjoint interiors, but not closures. By Lemma 49 there exists an $s$ such that for all $1 \leq k \leq b$, we have $R_k$ belongs to $K_s$ or $K_{s+1}$. As such write

$$\{R_k\}_{k=1}^{b} = \{R_k\}_{k=1}^{s+1} \cup \{R_k^{s+1}\}_{k=1}^{d},$$

where $R_k^j \in K_j$ for $j = s, s + 1$. Similarly, write

$$\{I_k^{\text{transition}}\}_{k=1}^{b} = \{I_k^{\text{transition}}\}_{k=1}^{s+1} \cup \{I_k^{s+1}\}_{k=1}^{d},$$

so that $I_k^{\text{transition}, j} \subset R_k^j$.

We begin by computing

$$\int_I u'dx = \sum_{k=1}^{a} \int_{I_k^{\text{rigid}}} u'dx + \sum_{k=1}^{c} \int_{I_k^{\text{transition}}} u'dx + \sum_{k=1}^{d} \int_{I_k^{\text{transition}, s+1}} u'dx$$

$$= 2^{-n} |I| \left( \sum_{k=1}^{a} E_{I_k^{\text{rigid}}} u' + \sum_{k=1}^{c} E_{I_k^{\text{transition}}} u' + \sum_{k=1}^{d} E_{I_k^{\text{transition}, s+1}} u' \right),$$

where we used the fact that all $I_k$’s satisfy $|I_k| = 2^{-n} |I|$. By applying Lemma 50 part (3) to the first sum, the above becomes

$$2^{-n} |I| \left( \sum_{k=1}^{a} E_{K_k} u' + \sum_{k=1}^{c} E_{I_k^{\text{transition}}} u' + \sum_{k=1}^{d} E_{I_k^{\text{transition}, s+1}} u' \right),$$

for certain $E_{K_k}$.
and by applying Lemma [50] parts (1) and (2), this then becomes

\[ 2^{-n}|I| \left( \sum_{k=1}^{a} E_{S(K_k)} U + \sum_{k=1}^{c} E_{S(\pi_{K_k} R_k)} U + \sum_{k=1}^{d} E_{S(\pi_{K_k} R_k')} U \right). \]

Since all the $I_k$'s are adjacent, then so are the $K_k$'s. Since the $K_k$'s are all pairwise adjacent in the grid $\widehat{K}$, then their supervisors $\{S(K_k)\}_k$ must all have a common $D$-parent, say $L$. Similarly, $\{S(\pi_{K_k} R_k)\}_k$ all have common $D$-parent which we'll call $T_j$ for $j = s, s+1$. Then the above term may be written as

\[ 2^{-n}|I| \left( \frac{1}{2^{-n}|L|} \sum_{k=1}^{a} \int_{S(K_k)} U dx + \frac{1}{2^{-n}|T_s|} \sum_{k=1}^{c} \int_{S(\pi_{K_k} R_k)} U dx + \frac{1}{2^{-n}|T_{s+1}|} \sum_{k=1}^{d} \int_{S(\pi_{K_k} R_k')} U dx \right). \]

Putting everything together, we get

\[ (7.2) \quad E_{Iu'} \leq \frac{1}{|L|} \sum_{k=1}^{a} \int_{S(K_k)} U dx + \frac{1}{|T_s|} \sum_{k=1}^{c} \int_{S(\pi_{K_k} R_k)} U dx + \frac{1}{|T_{s+1}|} \sum_{k=1}^{d} \int_{S(\pi_{K_k} R_k')} U dx \]

\[ \leq 1_{a \neq 0} \left( \frac{1}{|L|} \sum_{k=1}^{a} \int_{S(K_k)} U dx \right) + 1_{c \neq 0} \left( \frac{1}{|T_s|} \sum_{k=1}^{c} \int_{S(\pi_{K_k} R_k)} U dx \right) \]

\[ + 1_{d \neq 0} \left( \frac{1}{|T_{s+1}|} \sum_{k=1}^{d} \int_{S(\pi_{K_k} R_k')} U dx \right) \]

where $1_{a \neq 0} \equiv \begin{cases} 1 & \text{if } a \neq 0, \\ 0 & \text{if } a = 0. \end{cases}$ A similar inequality holds for $v'$ and $V$.

Since $a + c + d = 2^n$, then $a$, $c$, or $d$ must be nonzero, say without loss of generality $a$ (the other cases are similar). Now note the following: since the $I_j$'s were all adjacent, then any $K_k$ and $R_k'$ have disjoint interiors, but not closures. The $K_k \in K$ are not transition cubes, so by Lemma [49] the only way that $K_k$ and $R_k'$ can have disjoint interiors, but not closures, is if $|t - s| \leq 1$. In turn this means that $L$ and $T_s$ must be separated by at most two dyadic levels in $D$ if $c \neq 0$. Similarly for $L$ and $T_{s+1}$ if $d \neq 0$.

Thus by dyadic $\tau$-flatness of $U$, we get

\[ E_{T_k} U \leq (1 + \tau)^2 E_{L} U \quad \text{if } c \neq 0; \quad \text{and } E_{T_{s+1}} U \leq (1 + \tau)^2 E_{L} U \quad \text{if } d \neq 0. \]

Thus (7.2) yields $E_{Iu'} \leq 9 E_{L} U$ since $\tau \in (0, 1)$. Similarity for $v'$ and $V$. Altogether, we get

\[ (E_{Iu'}) (E_{Iv'}) \leq 81 (E_{L} u') (E_{L} v') \leq 81 A^2_{dyadic} (V, U; Q^0) \leq 81. \]

\[ \square \]

Obtaining classical $A_2$ on $Q_0$ from Lemma [52] is a standard exercise.

References

[AlLuSaWe] M. Alexis, J. L. Luna-Garcia, E.T. Sawyer, and I. Uriarte-Tuero, Stability of Weighted Norm Inequalities, \arxiv{2208.08400v2}.

[AlSaWe] M. Alexis, E. T. Sawyer and I. Uriarte-Tuero, A T1 theorem for general Calderón-Zygmund operators with doubling weights, and optimal cancellation conditions, II, \arxiv{2111.06277}.

[AlSaWe2] M. Alexis, E. T. Sawyer and I. Uriarte-Tuero, Tops of dyadic grids and T1 theorems, \arxiv{2201.02897}.

[AeIwMa] Kari Astala, Tadeusz Iwaniec and Gaven Martin, Elliptic Partial Differential Equations and Quasiconformal Mappings in the Plane, Princeton University Press, Princeton, New Jersey 2009.

[Bou] J. Bourgain, Some remarks on Banach spaces in which martingale difference sequences are unconditional, Ark. Mat. 21 (1–2) (1983) 163–168.

[CoFe] R. R. Coifman and C. L. Fefferman, Weighted norm inequalities for maximal functions and singular integrals, Studia Math. 51 (1974), 241-250.

[DaTo] Damian Dąbrowski and Xavier Tolsa, The measures with $L^2$ bounded Riesz transform satisfying a subcritical Wolff-type energy condition, \arxiv{2106.00303}.

[FeMu] C. L. Fefferman and B. Muckenhoupt, Two nonequivalent conditions for weight functions, Proc. A.M.S. 45 (1974), 99-104.

[Gra] Grafakos, L., Classical Fourier Analysis. New York: Springer, (2008).

[Gri] Grigoris, C., Necessary/sufficient conditions in weighted theory, \arxiv{2009.12091}.

[HaNi] V.P. Haviv, N.K. Nikol’ski, Linear and Complex Analysis. Problem Book 3. Part I, Lecture Notes in Mathematics, vol. 1573, Springer-Verlag, Berlin, 1994.
A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).

E. Sawyer, A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).

E. Sawyer, A T1 theorem for general Calderón-Zygmund operators with comparable doubling weights and optimal cancellation conditions, arXiv:1906.05602v10, to appear in Journal d'Analyse.

Sawyer, Eric T., Shen, Chun-Yen, Uriarte-Tuero, Ignacio, A two weight theorem for α-fractional singular integrals with an energy side condition, Journal of Functional Analysis, Volume 263 (2012), Issue 2, 305-363.

Sawyer, Eric T., Wick, Brett D., Two weight inequalities for Riesz transforms: uniformly full dimension weights, arXiv:1312.6163v3.

B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, Trans. A.M.S. 165 (1972), 207-226.

B. Muckenhoupt and R. L. Wheeden, Two weight function norm inequalities for the Hardy-Littlewood maximal function and the Hilbert transform, Studia Math. 55 (1976), 279-294.

F. Nazarov, S. Treil and A. Volberg, The Bellman function, the two weight Hilbert transform, and the embeddings of the model space $K_p$, J. d'Analyse Math. 87 (2002), 385-414.

C. J. Neugebauer, Inserting $A_p$ weights, Proceedings of the American Mathematical Society, vol. 87, no. 4, 1983, pp. 644-48.

E. Sawyer, A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).

E. Sawyer, A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).

E. Sawyer, A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).

E. Sawyer, A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).

E. Sawyer, A characterization of a two-weight norm inequality for maximal operators, Studia Math. 75 (1982), 1-11, MR(676801 (84i:42032)).