Composition of Dyadic Paraproducts

Pott, Sandra; Reguera, Maria Carmen; Sawyer, Eric T.; Wick, Brett D.

DOI:
10.1016/j.aim.2016.04.024

COMPOSITION OF DYADIC PARAPRODUCTS

SANDRA POTT, MARIA CARMEN REGUERA[1], ERIC T. SAWYER[2], AND BRETT D. WICK[3]

Abstract. We obtain necessary and sufficient conditions to characterize the boundedness of the composition of dyadic paraproduct operators.

Contents

1. Introduction

Recall that a Toeplitz operator on the Hardy space of analytic functions $H^2(\mathbb{D})$ is defined by

$$T_\varphi : H^2(\mathbb{D}) \to H^2(\mathbb{D}) \text{ where } T_\varphi f = \mathbb{P}_{H^2}(\varphi f).$$

It is well known that this operator is bounded if and only if $\varphi \in L^\infty(\mathbb{T})$. Equivalently, the Toeplitz operator T_φ is bounded if and only if $\sup_{\lambda\in\mathbb{D}} \|T_\varphi k_\lambda\|_{H^2} < \infty$ where $k_\lambda(z) = \frac{1}{1-\bar{\lambda}z}$ is the reproducing kernel for $H^2(\mathbb{D})$. An infamous conjecture of Sarason, [8], states that the composition of two (potentially unbounded) Toeplitz operators is bounded, i.e. $T_\varphi T_{\overline{\psi}}$ is a bounded operator, if and only if a certain relatively simple testing condition on the symbols φ and ψ hold, see [10]. However, even though this conjecture seems quite reasonable, a beautiful counterexample was constructed by F. Nazarov in [2] disproving this simple testing condition.

In this paper we are interested in a discrete dyadic analogue of the Sarason conjecture. This discrete problem is already very challenging and captures much of the difficulty associated with Sarason's original conjecture but is more amenable to study because of the dyadic nature of the problem. In particular, we are concerned with dyadic Haar paraproducts, and obtaining necessary and sufficient conditions for the boundedness of the composition of

1. Research supported by grants 2009SGR-000420 (Generalitat de Catalunya) and MTM-2010-16232 (Spain).

2. Research supported in part by a NSERC Grant.

3. Research supported in part by National Science Foundation DMS grants # 1001098 and # 955432.

The authors would like to thank the Banff International Research Station for the Banff–PIMS Research in Teams support for the project: The Sarason Conjecture and the Composition of Paraproducts.

two such paraproducts. The conditions characterizing the boundedness will be much more general than just those characterizing boundedness for each individual paraproduct - just as the condition $\|bd\|_\infty < \infty$ that characterizes boundedness of the composition $M_b \circ M_d$ of pointwise multipliers is much more general than the conditions $\|b\|_\infty < \infty$ and $\|d\|_\infty < \infty$ that characterize individual boundedness of the pointwise multipliers.

Let $\mathcal{D}$ denote the usual dyadic grid of intervals on the real line. We consider sequences $b = \{b_I\}_{I\in\mathcal{D}}$ of complex numbers on $\mathcal{D}$, which we often refer to as <u>symbols</u>. Define the Haar function h_I^0 and averaging function h_I^1 by

$$h_I^0 \equiv h_I \equiv \frac{1}{\sqrt{|I|}}\left(-\mathbf{1}_{I_-} + \mathbf{1}_{I_+}\right) \text{ and } h_I^1 \equiv \frac{1}{|I|}\mathbf{1}_I \ , \qquad I \in \mathcal{D}.$$

The operators considered in this paper are the following dyadic paraproducts.

Definition 1.1. *Given a symbol $b = \{b_I\}_{I\in\mathcal{D}}$ and a pair $(\alpha,\beta) \in \{0,1\}\times\{0,1\}$, define the <u>dyadic paraproduct</u> acting on a function f by*

$$\mathsf{P}_b^{(\alpha,\beta)} f \equiv \sum_{I\in\mathcal{D}} b_I \left\langle f, h_I^\beta \right\rangle_{L^2(\mathbb{R})} h_I^\alpha,$$

where h_I^0 is the Haar function associated with I, and h_I^1 is the average function associated with I. The index (α,β) is referred to as the <u>type</u> of $\mathsf{P}_b^{(\alpha,\beta)}$.

The purpose of this paper is to characterize boundedness on $L^2(\mathbb{R})$ of the compositions $\mathsf{P}_b^{(\alpha,\beta)} \circ \mathsf{P}_d^{(\gamma,\delta)}$. We denote the composition $\mathsf{P}_b^{(\alpha,\beta)} \circ \mathsf{P}_d^{(\gamma,\delta)}$ by $\mathsf{P}_{b,d}^{(\alpha,\beta,\gamma,\delta)}$, and refer to the index $(\alpha,\beta,\gamma,\delta)$ as the <u>type</u> of the product $\mathsf{P}_b^{(\alpha,\beta)} \circ \mathsf{P}_d^{(\gamma,\delta)}$. The dual $\left(\mathsf{P}_b^{(\alpha,\beta)}\right)^* = \mathsf{P}_b^{(\beta,\alpha)}$ of the operator $\mathsf{P}_b^{(\beta,\alpha)}$ is obtained by exchanging exponents, which then reduces the total number of products to be investigated. We are able to give reasonable characterizations of the operator norm $\left\|\mathsf{P}_{b,d}^{(\alpha,\beta,\gamma,\delta)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}$ in two special cases, namely when the type of $\mathsf{P}_{b,d}^{(\alpha,\beta,\gamma,\delta)}$ is of the form $(\alpha,0,0,\delta)$ or $(0,\beta,\gamma,0)$.

In the first case, the product $\mathsf{P}_{b,d}^{(\alpha,0,0,\delta)}$ reduces to a single paraproduct $\mathsf{P}_{b\circ d}^{(\alpha,\delta)}$ whose symbol $b\circ d$ is built from the sequences in a very simple manner.

In the second case, the compositions are not as easy since there is less cancellation. However, we are able to transplant the problem, first to an operator on the discrete Bergman space on $\mathcal{D}$, and then to a two weight norm inequality for a positive or singular operator on $L^2(\mathcal{H})$. The positive operator inequality reduces to the tree inequality in [1], while the singular operator inequality is solved by an extension of a two weight theorem in [3]. The transplantation idea seems to be a novel element in our approach to paraproducts, and should find application elsewhere.

Our main results are then the following theorems that characterize the compositions in certain cases. To state them requires some additional notation. For a sequence $a = \{a_I\}_{I\in\mathcal{D}}$ define.

$$\begin{aligned} \|a\|_{\ell^\infty} &\equiv \sup_{I\in\mathcal{D}} |a_I| \ ; \\ \|a\|_{CM} &\equiv \sqrt{\sup_{I\in\mathcal{D}} \frac{1}{|I|}\sum_{J\subset I} |a_J|^2} \ . \end{aligned}$$

Given two sequences $b = \{b_I\}_{I\in\mathcal{D}}$ and $d = \{d_I\}_{I\in\mathcal{D}}$ let $b \circ d$ denote the Schur product of the sequences, i.e.

$$b \circ d \equiv \{b_I d_I\}_{I\in\mathcal{D}} \ .$$

In the case of a composition of type $(0,0,0,1)$, $(1,0,0,0)$ or $(0,0,0,0)$ we have the following characterization.

Theorem 1.1. *The composition $\mathsf{P}_b^{(0,0)} \circ \mathsf{P}_d^{(0,1)}$ and $\mathsf{P}_b^{(1,0)} \circ \mathsf{P}_d^{(0,0)}$ is bounded on $L^2(\mathbb{R})$ if and only if $\|b \circ d\|_{CM} < \infty$. Moreover, the operator norm of the composition satisfies*

$$\left\|\mathsf{P}_b^{(0,0)} \circ \mathsf{P}_d^{(0,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} = \left\|\mathsf{P}_b^{(1,0)} \circ \mathsf{P}_d^{(0,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \|b \circ d\|_{CM} \ .$$

The composition $\mathsf{P}_b^{(0,0)} \circ \mathsf{P}_d^{(0,0)}$ is bounded on $L^2(\mathbb{R})$ if and only if $\|b \circ d\|_{\ell^\infty} < \infty$ Moreover, the operator norm of the composition satisfies

$$\left\|\mathsf{P}_b^{(0,0)} \circ \mathsf{P}_d^{(0,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \|b \circ d\|_{\ell^\infty} \ .$$

For compositions of type $(1,0,0,1)$ we also have a characterization, but again require some additional notation. Given a symbol $a = \{a_I\}_{I\in\mathcal{D}}$, we define the *sweep*, $\widehat{S}(a)$, of a by

$$\widehat{S}(a) \equiv \left\{\left\langle \sum_{J\in\mathcal{D}} a_J h_J^1, h_I \right\rangle_{L^2(\mathbb{R})}\right\}_{I\in\mathcal{D}} = \left\{\sum_{J\subsetneq I} a_J \widehat{h_J^1}(I)\right\}_{I\in\mathcal{D}} , \tag{1.1}$$

and also the sequence $E(a)$ by

$$E(a) \equiv \left\{\frac{1}{|J|}\sum_{I\subset J} a_I\right\}_{J\in\mathcal{D}} . \tag{1.2}$$

The characterization is then given by the following theorem.

Theorem 1.2. *The composition $\mathsf{P}_b^{(1,0)} \circ \mathsf{P}_d^{(0,1)}$ is bounded on $L^2(\mathbb{R})$ if and only if $\left\|\widehat{S}(b \circ d)\right\|_{CM} < \infty$ and $\|E(b \circ d)\|_{\ell^\infty} < \infty$. Moreover, the operator norm of the composition $\mathsf{P}_b^{(1,0)} \circ \mathsf{P}_d^{(0,1)}$ on $L^2(\mathbb{R})$ satisfies*

$$\left\|\mathsf{P}_b^{(1,0)} \circ \mathsf{P}_d^{(0,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} = \left\|\mathsf{P}_{b\circ d}^{(1,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \left\|\widehat{S}(b \circ d)\right\|_{CM} + \|E(b \circ d)\|_{\ell^\infty} \ .$$

In the case of the composition of type $(0,1,1,0)$ we obtain the following theorem. To state the characterization again requires slightly more notation. Given a function $f \in L^2(\mathbb{R})$ and an interval $I \in \mathcal{D}$ we let

$$\mathsf{Q}_I f \equiv \sum_{J\subset I} \langle f, h_J\rangle_{L^2(\mathbb{R})} h_J \tag{1.3}$$

denote the projection of the function f onto the span of the Haar functions supported within the interval I. When applied to a sequence $a = \{a_I\}_{I\in\mathcal{D}}$ the operator Q_I takes the following form:

$$\mathsf{Q}_I a \equiv \sum_{J\subset I} a_J h_J. \tag{1.4}$$

Notice that this definition encompasses the definition when applied to functions since we can always identify a function with its sequence of Haar coefficients.

Our characterization is then the following theorem.

Theorem 1.3. *The composition* $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ *is bounded on* $L^2(\mathbb{R})$ *if and only if both*

$$\begin{aligned} \left\| \mathsf{Q}_I \mathsf{P}_b^{(0,1)} \mathsf{P}_d^{(1,0)} \left(\mathsf{Q}_I \overline{d} \right) \right\|_{L^2(\mathbb{R})}^2 &\leq C_1^2 \left\| \mathsf{Q}_I d \right\|_{L^2(\mathbb{R})}^2 ; \\ \left\| \mathsf{Q}_I \mathsf{P}_d^{(0,1)} \mathsf{P}_b^{(1,0)} \left(\mathsf{Q}_I \overline{b} \right) \right\|_{L^2(\mathbb{R})}^2 &\leq C_2^2 \left\| \mathsf{Q}_I b \right\|_{L^2(\mathbb{R})}^2 \end{aligned}$$

for all $I \in \mathcal{D}$*; i.e. for all* $I \in \mathcal{D}$ *the following inequalities are true*

$$\begin{aligned} \sum_{J \subset I} |b_J|^2 \frac{1}{|J|^2} \left(\sum_{L \subset J} |d_L|^2 \right)^2 &\leq C_1^2 \sum_{L \subset I} |d_L|^2 ; \\ \sum_{J \subset I} |d_J|^2 \frac{1}{|J|^2} \left(\sum_{L \subset J} |b_L|^2 \right)^2 &\leq C_2^2 \sum_{L \subset I} |b_L|^2 . \end{aligned}$$

Moreover, the norm of $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ *on* $L^2(\mathbb{R})$ *satisfies*

$$\left\| \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)} \right\|_{L^2(\mathbb{R}) \to L^2(\mathbb{R})} \approx C_1 + C_2$$

where C_1 *and* C_2 *are the best constants in appearing above.*

In the case of composition of type $(0,1,0,0)$, and by duality and symmetry the type $(0,0,1,0)$, we have the following characterization of the composition of Haar paraproducts.

Theorem 1.4. *The composition* $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)}$ *is bounded on* $L^2(\mathbb{R})$ *if and only if both*

$$\begin{aligned} |d_I| \left\| \mathsf{P}_b^{(0,1)} h_I \right\|_{L^2(\mathbb{R})} &\leq C_1; \\ \left\| \mathsf{Q}_I \mathsf{P}_d^{(0,0)} \mathsf{P}_b^{(1,0)} \mathsf{Q}_I \overline{b} \right\|_{L^2(\mathbb{R})} &\leq C_2 \left\| \mathsf{Q}_I b \right\|_{L^2(\mathbb{R})} \end{aligned}$$

for all $I \in \mathcal{D}$*.; i.e. for all* $I \in \mathcal{D}$ *the following inequalities are true*

$$\begin{aligned} |d_I| \left(\frac{1}{|I|} \sum_{L \subsetneq I} |b_L|^2 \right)^{\frac{1}{2}} &\leq C_1; \\ \left(\sum_{J \subset I} \frac{|d_J|^2}{|J|} \left(\sum_{K \subset J_+} |b_K|^2 - \sum_{K \subset J_-} |b_K|^2 \right)^2 \right)^{\frac{1}{2}} &\leq C_2 \left(\sum_{L \subset I} |b_L|^2 \right)^{\frac{1}{2}} . \end{aligned}$$

Moreover, the norm of $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ *on* $L^2(\mathbb{R})$ *satisfies*

$$\left\| \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)} \right\|_{L^2(\mathbb{R}) \to L^2(\mathbb{R})} \approx C_1 + C_2$$

where C_1 *and* C_2 *are the best constants in appearing above.*

The outline of the paper is as follows. In Section 2 we carry out the proof of Theorems 1.1 and 1.2. These essentially reduce to the characterizations of when a single paraproduct is bounded. In Section 3 we give the proof of Theorems 1.3 and 1.4. These characterizations are more difficult, but can be studied via techniques used to obtain two-weight inequalities for positive and well-localized operators.

As an application of these results it is possible to provide a new proof of the following:

Theorem 1.5 (Petermichl, [4]). *Let $w \in A_2$. Then*

$$\|H\|_{L^2(w)\to L^2(w)} \lesssim [w]_{A_2}.$$

Above, the A_2 characteristic of the function w is the quantity:

$$[w]_{A_2} \equiv \sup_{I\in\mathcal{D}} \langle w\rangle_I \left\langle w^{-1}\right\rangle_I;$$

while the Hilbert transform is defined by

$$H(f)(x) \equiv \int_{\mathbb{R}} \frac{f(y)}{y-x}dy,$$

with the integral taken in the principle value sense. One first notes

$$H : L^2(w) \to L^2(w) \iff M_{w^{\frac{1}{2}}} H M_{w^{-\frac{1}{2}}} : L^2 \to L^2.$$

The second reduction is to note that since the Hilbert transform can be recovered by averaging the Haar shifts, and since we are only after an upper bound, it will be sufficient to study the following dyadic model operator

$$M_{w^{\frac{1}{2}}} \mathcal{S} M_{w^{-\frac{1}{2}}} : L^2 \to L^2$$

where $\mathcal{S}$ is a shift operator defined on the Haar basis by $\mathcal{S}h_I \equiv h_{I_-} - h_{I_+}$. Because of linearity, it suffices to consider just "half" of the shift operator $\mathcal{S}$ defined by the operator $h_{I_-} \otimes h_I$. This averaging of shifts to recover H is the key observation made by Petermichl in [5] and played a decisive role in her proof of Theorem 1.5. Then note that

$$M_{w^{\pm\frac{1}{2}}} = \mathsf{P}^{(0,1)}_{\widehat{w^{\pm\frac{1}{2}}}} + \mathsf{P}^{(1,0)}_{\widehat{w^{\pm\frac{1}{2}}}} + \mathsf{P}^{(0,0)}_{\langle w^{\pm\frac{1}{2}}\rangle}$$

where, we can recognize the operators above as paraproducts by setting

$$\begin{aligned} \widehat{b}(I) &= \left\langle b, h_I^0\right\rangle_{L^2(\mathbb{R})} \\ \langle b\rangle_I &= \left\langle b, h_I^1\right\rangle_{L^2(\mathbb{R})}, \end{aligned}$$

for $I \in \mathcal{D}$. Then writing

$$\left(\mathsf{P}^{(0,1)}_{\widehat{w^{\frac{1}{2}}}} + \mathsf{P}^{(1,0)}_{\widehat{w^{\frac{1}{2}}}} + \mathsf{P}^{(0,0)}_{\langle w^{\frac{1}{2}}\rangle}\right) \mathcal{S} \left(\mathsf{P}^{(0,1)}_{\widehat{w^{-\frac{1}{2}}}} + \mathsf{P}^{(1,0)}_{\widehat{w^{-\frac{1}{2}}}} + \mathsf{P}^{(0,0)}_{\langle w^{-\frac{1}{2}}\rangle}\right) \tag{1.5}$$

we can recognize these terms as composition of paraproducts. One then can apply the above theorems characterizing the composition of paraproducts, and then verify that the testing conditions that appear can all be controlled by a linear power of the A_2 characteristic. In particular this strengthens the linear bound for $\mathcal{S}$ on $L^2(w)$ by obtaining such a bound for each of the nine operators arising in the canonical decomposition (1.5). A simpler proof of this, but again using the strategy outlined above, appears in [6].

We thank the referee for helpful comments about the current version of this paper.

2. Reduction to Single Paraproducts

Given two sequences $b = \{b_I\}_{I\in\mathcal{D}}$ and $d = \{d_I\}_{I\in\mathcal{D}}$ let $b \circ d$ denote the Schur product of the sequences, i.e.

$$b \circ d \equiv \{b_I d_I\}_{I\in\mathcal{D}}.$$

The composition $\mathsf{P}_b^{(\alpha,0)} \circ \mathsf{P}_d^{(0,\beta)}$ is given by

$$\begin{aligned}
\left(\mathsf{P}_b^{(\alpha,0)} \circ \mathsf{P}_d^{(0,\beta)}\right) f &= \mathsf{P}_b^{(\alpha,0)}\left(\mathsf{P}_d^{(0,\beta)} f\right) = \sum_{I\in\mathcal{D}} b_I \left\langle \mathsf{P}_d^{(0,\beta)} f, h_I^0 \right\rangle_{L^2(\mathbb{R})} h_I^\alpha \\
&= \sum_{I\in\mathcal{D}} b_I \left\langle \sum_{J\in\mathcal{D}} d_J \left\langle f, h_J^\beta \right\rangle_{L^2(\mathbb{R})} h_J^0, h_I^0 \right\rangle_{L^2(\mathbb{R})} h_I^\alpha \\
&= \sum_{I\in\mathcal{D}} b_I d_I \left\langle f, h_I^\beta \right\rangle_{L^2(\mathbb{R})} h_I^\alpha \\
&= \mathsf{P}_{b\circ d}^{(\alpha,\beta)} f \ . \qquad (2.1)
\end{aligned}$$

Thus the boundedness of the product $\mathsf{P}_b^{(\alpha,0)} \circ \mathsf{P}_d^{(0,\beta)}$ reduces to that of a single paraproduct $\mathsf{P}_{b\circ d}^{(\alpha,\beta)}$. There are three cases in which a single paraproduct is easily characterized, namely $\mathsf{P}_a^{(0,0)}$, $\mathsf{P}_a^{(0,1)}$ and $\mathsf{P}_a^{(1,0)}$.

Lemma 2.1. *We have the characterizations*

$$\left\|\mathsf{P}_a^{(0,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} = \|a\|_{\ell^\infty} \ ; \qquad (2.2)$$

$$\left\|\mathsf{P}_a^{(0,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} = \left\|\mathsf{P}_a^{(1,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \|a\|_{CM} \ . \qquad (2.3)$$

Proof. With the notation

$$\widehat{f}(K) = \langle f, h_K\rangle_{L^2(\mathbb{R})} \ ,$$

the identities

$$\begin{aligned}
\left\|\mathsf{P}_a^{(0,0)} f\right\|_{L^2(\mathbb{R})}^2 &= \sum_{I,I'\in\mathcal{D}} a_I \overline{a_{I'}} \langle f, h_I\rangle_{L^2(\mathbb{R})} \overline{\langle f, h_{I'}\rangle_{L^2(\mathbb{R})}} \langle h_I, h_{I'}\rangle_{L^2(\mathbb{R})} \\
&= \sum_{I\in\mathcal{D}} |a_I|^2 \left|\widehat{f}(I)\right|^2 , \\
\|f\|_{L^2(\mathbb{R})}^2 &= \sum_{I\in\mathcal{D}} \left|\widehat{f}(I)\right|^2 ,
\end{aligned}$$

immediately gives (2.2). Then the Carleson Embedding Theorem gives

$$\begin{aligned}
\left\|\mathsf{P}_a^{(0,1)} f\right\|_{L^2(\mathbb{R})}^2 &= \sum_{I,I'\in\mathcal{D}} a_I \overline{a_{I'}} \left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})} \overline{\left\langle f, h_{I'}^1\right\rangle_{L^2(\mathbb{R})}} \langle h_I, h_{I'}\rangle_{L^2(\mathbb{R})} \\
&= \sum_{I\in\mathcal{D}} |a_I|^2 \left|\left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})}\right|^2 \lesssim \left\{\sup_{I\in\mathcal{D}} \frac{1}{|I|} \sum_{J\subset I} |a_J|^2\right\} \|f\|_{L^2(\mathbb{R})}^2 .
\end{aligned}$$

So we have that

$$\left\|\mathsf{P}_a^{(0,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \lesssim \|a\|_{CM} .$$

To see that the other inequality holds, simply test on a Haar function. Indeed, let $\hat{I}$ denote the parent of I, and then we have

$$\left\|\mathsf{P}_a^{(0,1)} h_{\hat{I}}\right\|_{L^2(\mathbb{R})}^2 \leq \left\|\mathsf{P}_a^{(0,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}^2 .$$

However, a computation shows that

$$\begin{aligned}\left\|\mathsf{P}_a^{(0,1)}h_{\hat{I}}\right\|_{L^2(\mathbb{R})}^2 &= \left\|\sum_{J\in\mathcal{D}} a_J \left\langle h_{\hat{I}}, h_J^1\right\rangle h_J\right\|_{L^2(\mathbb{R})}^2 \\ &= \sum_{J\in\mathcal{D}} |a_J|^2 \left|\left\langle h_{\hat{I}}, h_J^1\right\rangle_{L^2(\mathbb{R})}\right|^2 \\ &\gtrsim \frac{1}{|I|}\sum_{J\subset I} |a_J|^2 ,\end{aligned}$$

which proves (2.3). ∎

It is clear that Lemma 2.1 coupled with the computations above prove Theorem 1.1.

2.1. **The Pott-Smith Identity.** We now recall a useful identity obtained by Pott and Smith in [7, Proposition 2.3] related to the composition of certain types of paraproducts. We will use the identity to obtain necessary and sufficient conditions for $\mathsf{P}_a^{(1,1)}$ to be bounded on $L^2(\mathbb{R})$.

We remind the reader that for a symbol $a = \{a_I\}_{I\in\mathcal{D}}$, in equation (1.1) the sweep $\widehat{S}(a)$ of a was defined by

$$\widehat{S}(a) \equiv \left\{\left\langle \sum_{J\in\mathcal{D}} a_J h_J^1, h_I\right\rangle_{L^2(\mathbb{R})}\right\}_{I\in\mathcal{D}} = \left\{\sum_{J\subsetneq I} a_J \widehat{h_J^1}(I)\right\}_{I\in\mathcal{D}} ,$$

and equation (1.2) we defined the sequence $E(a)$ by

$$E(a) \equiv \left\{\frac{1}{|J|}\sum_{I\subset J} a_I\right\}_{J\in\mathcal{D}} .$$

We now decompose the paraproduct $\mathsf{P}_a^{(1,1)}$ into paraproducts with simpler types, each having at least one 0 in the index. The most natural idea is to expand the averaging functions in a Haar series: $h_I^1 = \sum_{J\supsetneq I} \widehat{h_I^1}(J)\, h_J$, and then to split the resulting double sum over intervals into diagonal, upper and lower parts. Carrying out this strategy we obtain:

$$\begin{aligned}\mathsf{P}_a^{(1,1)} f &= \sum_{I\in\mathcal{D}} a_I \left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})} h_I^1 \\ &= \sum_{I\in\mathcal{D}} a_I \left\langle f, \left(\sum_{J\supsetneq I} \widehat{h_I^1}(J)\, h_J\right)\right\rangle_{L^2(\mathbb{R})} \left(\sum_{K\supsetneq I} \widehat{h_I^1}(K)\, h_K\right) \\ &= \left\{\sum_{J\subsetneq K} + \sum_{K\subsetneq J} + \sum_{J=K}\right\} \sum_{I\subset J\cap K} a_I \widehat{h_I^1}(J)\, \widehat{f}(J)\, \widehat{h_I^1}(K)\, h_K \\ &\equiv T^{(1,0)} f + T^{(0,1)} f + T^{(0,0)} f .\end{aligned}$$

Now we have

$$\begin{aligned} T^{(0,0)}f &= \sum_{J\in\mathcal{D}}\sum_{I\subset J} a_I \widehat{h_I^1}(J)\,\widehat{f}(J)\,\widehat{h_I^1}(J)\,h_J = \sum_{J\in\mathcal{D}}\sum_{I\subset J} a_I \widehat{h_I^1}(J)^2\,\widehat{f}(J)\,h_J \\ &= \sum_{J\in\mathcal{D}}\left(\frac{1}{|J|}\sum_{I\subset J} a_I\right)\widehat{f}(J)\,h_J = \mathsf{P}^{(0,0)}_{E(a)}f \end{aligned}$$

where we have use the definition of $E(a)$ in (1.2). We also have with

$$\left\{\sum_{I\subset J} a_I \widehat{h_I^1}(J)\right\}_{J\in\mathcal{D}} = \left\{\left\langle \sum_{I\subsetneq J} a_I h_I^1, h_J\right\rangle_{L^2(\mathbb{R})}\right\}_{J\in\mathcal{D}} = \left\{\widehat{\sum_{I\in\mathcal{D}} a_I h_I^1}(J)\right\}_{J\in\mathcal{D}}$$

that

$$\begin{aligned} T^{(1,0)}f &= \sum_{J\subsetneq K}\sum_{I\subset J} a_I \widehat{h_I^1}(J)\,\widehat{f}(J)\,\widehat{h_I^1}(K)\,h_K = \sum_{J\in\mathcal{D}}\sum_{I\subset J}\sum_{K\supsetneq J} a_I \widehat{h_I^1}(J)\,\widehat{f}(J)\,\widehat{h_I^1}(K)\,h_K \\ &= \sum_{J\in\mathcal{D}}\left(\sum_{I\subset J} a_I \widehat{h_I^1}(J)\right)\widehat{f}(J)\left(\sum_{K\supsetneq J}\widehat{h_J^1}(K)\,h_K\right) \\ &= \sum_{J\in\mathcal{D}}\widehat{\sum_{I\subset J} a_I h_I^1}(J)\;\;\widehat{f}(J)\;\;h_J^1 = \mathsf{P}^{(1,0)}_{\widehat{S}(a)}f. \end{aligned}$$

Similarly,

$$T^{(0,1)}f = \mathsf{P}^{(0,1)}_{\widehat{S}(a)}f,$$

and altogether we have the desired decomposition

$$\mathsf{P}^{(1,1)}_a = \mathsf{P}^{(1,0)}_{\widehat{S}(a)} + \mathsf{P}^{(0,1)}_{\widehat{S}(a)} + \mathsf{P}^{(0,0)}_{E(a)}\,. \tag{2.4}$$

Thus we see that the single paraproduct $\mathsf{P}^{(1,1)}_a$ not covered by (2.2) reduces to types already characterized. Using this we can then obtain the characterization of the paraproduct $\mathsf{P}^{(1,1)}_a$.

Corollary 2.2. *The operator norm* $\left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}$ *of* $\mathsf{P}^{(1,1)}_a$ *on* $L^2(\mathbb{R})$ *satisfies*

$$\left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \left\|\widehat{S}(a)\right\|_{CM} + \|E(a)\|_{\ell^\infty}\,. \tag{2.5}$$

Proof. From (2.4) by applying Lemma 2.1 we have the following estimate

$$\left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \lesssim \left\|\widehat{S}(a)\right\|_{CM} + \|E(a)\|_{\ell^\infty}\,. \tag{2.6}$$

We now turn to showing that inequality (2.6) can be reversed. Suppose that $\left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}$ is finite. Then an easy computation shows that

$$\begin{aligned} \left\langle \mathsf{P}^{(1,1)}_a h_I, h_I\right\rangle_{L^2(\mathbb{R})} &= \left\langle \mathsf{P}^{(1,0)}_{\widehat{S}(a)} + \mathsf{P}^{(0,1)}_{\widehat{S}(a)} + \mathsf{P}^{(0,0)}_{E(a)} h_I, h_I\right\rangle_{L^2(\mathbb{R})} \\ &= \left\langle \mathsf{P}^{(1,0)}_{\widehat{S}(a)} h_I, h_I\right\rangle_{L^2(\mathbb{R})} + \left\langle \mathsf{P}^{(0,1)}_{\widehat{S}(a)} h_I, h_I\right\rangle_{L^2(\mathbb{R})} + \left\langle \mathsf{P}^{(0,0)}_{E(a)} h_I, h_I\right\rangle_{L^2(\mathbb{R})} \\ &= E(a)_I \end{aligned}$$

since

$$\left\langle \mathsf{P}^{(1,0)}_{\widehat{S}(a)} h_I, h_I \right\rangle_{L^2(\mathbb{R})} = \widehat{S}\,(a)_I \left\langle h_I^1, h_I \right\rangle_{L^2(\mathbb{R})} = 0.$$

A similar computation demonstrates that $\left\langle \mathsf{P}^{(0,1)}_{\widehat{S}(a)} h_I, h_I \right\rangle_{L^2(\mathbb{R})} = 0$ as well. Thus, we have

$$\|E(a)\|_{\ell^\infty} = \sup_{I\in\mathcal{D}} |E(a)_I| \leq \sup_{I\in\mathcal{D}} \left| \left\langle \mathsf{P}^{(1,1)}_a h_I, h_I \right\rangle_{L^2(\mathbb{R})} \right| \leq \left\| \mathsf{P}^{(1,1)}_a \right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}. \tag{2.7}$$

Again, let $\hat{I}$ denote the parent of the dyadic interval I. Now set

$$F_I \equiv \sum_{J\subset I} \overline{\widehat{S}(a)_J} h_J.$$

Then simple straightforward computations demonstrate that

$$\begin{aligned} \|F_I\|^2_{L^2(\mathbb{R})} &= \sum_{J\subset I} \left| \widehat{S}(a)_J \right|^2, \\ \left\langle F_I, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} &= 0, \\ \left\langle F_I, h^1_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} &= 0. \end{aligned}$$

First, observe that

$$\left| \left\langle \mathsf{P}^{(1,1)}_a F_I, h_{\hat{I}} \right\rangle \right|_{L^2(\mathbb{R})} \leq \left\| \mathsf{P}^{(1,1)}_a \right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \left(\sum_{J\subset I} \left| \widehat{S}(a)_J \right|^2 \right)^{\frac{1}{2}}. \tag{2.8}$$

Next, observe that the computations above involving F_I give that

$$\begin{aligned} \left\langle P^{(0,0)}_{E(a)} F_I, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} &= \sum_{K\in\mathcal{D}} E(a)_K \left\langle F_I, h_K \right\rangle_{L^2(\mathbb{R})} \left\langle h_K, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} \\ &= E(a)_{\hat{I}} \left\langle F_I, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} = 0; \end{aligned}$$

and

$$\begin{aligned} \left\langle P^{(0,1)}_{\widehat{S}(a)} F_I, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} &= \sum_{K\in\mathcal{D}} \widehat{S}(a)_K \left\langle F_I, h^1_K \right\rangle_{L^2(\mathbb{R})} \left\langle h_K, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} \\ &= \widehat{S}(a)_{\hat{I}} \left\langle F_I, h^1_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} = 0. \end{aligned}$$

Thus, using (2.4), (2.8), and the computations above we have that

$$\left| \left\langle \mathsf{P}^{(1,0)}_{\widehat{S}(a)} F_I, h_{\hat{I}} \right\rangle_{L^2(\mathbb{R})} \right| = \left| \left\langle \mathsf{P}^{(1,1)}_a F_I, h_{\hat{I}} \right\rangle \right|_{L^2(\mathbb{R})} \leq \left\| \mathsf{P}^{(1,1)}_a \right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \left(\sum_{J\subset I} \left| \widehat{S}(a)_J \right|^2 \right)^{\frac{1}{2}}. \tag{2.9}$$

Finally, we compute

$$\begin{aligned}\left|\left\langle \mathsf{P}^{(1,0)}_{\widehat{S}(a)} F_I, h_{\hat{I}}\right\rangle_{L^2(\mathbb{R})}\right| &= \left|\sum_{K\in\mathcal{D}} \widehat{S}(a)_K \left\langle F_I, h_K\right\rangle_{L^2(\mathbb{R})} \left\langle h^1_K, h_{\hat{I}}\right\rangle_{L^2(\mathbb{R})}\right| \\ &= \left|\sum_{K\subset I} \left|\widehat{S}(a)_K\right|^2 \left\langle h^1_K, h_{\hat{I}}\right\rangle_{L^2(\mathbb{R})}\right| \\ &= \frac{1}{\sqrt{\left|\hat{I}\right|}} \sum_{K\subset I} \left|\widehat{S}(a)_K\right|^2 . \end{aligned} \tag{2.10}$$

Combining (2.9) and (2.10) yields

$$\frac{1}{\sqrt{|I|}} \sum_{K\subset I} \left|\widehat{S}(a)_K\right|^2 \lesssim \left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \left(\sum_{J\subset I} \left|\widehat{S}(a)_J\right|^2\right)^{\frac{1}{2}},$$

which gives

$$\left(\frac{1}{|I|} \sum_{J\subset I} \left|\widehat{S}(a)_J\right|^2\right)^{\frac{1}{2}} \lesssim \left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}$$

and then taking the supremum over $I \in \mathcal{D}$ gives.

$$\left\|\widehat{S}(a)\right\|_{CM} \lesssim \left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} . \tag{2.11}$$

Combining (2.7) and (2.11) gives

$$\|E(a)\|_{\ell^\infty} + \left\|\widehat{S}(a)\right\|_{CM} \lesssim \left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} . \tag{2.12}$$

Then (2.6) and (2.12) prove the Corollary. ∎

When studying the composition $\mathsf{P}^{(1,0)}_b \circ \mathsf{P}^{(0,1)}_d$ identity (2.1) along with Corollary 2.2 yields the following result.

Corollary 2.3. *The operator norm of the composition* $\mathsf{P}^{(1,0)}_b \circ \mathsf{P}^{(0,1)}_d$ *satisfies*

$$\left\|\mathsf{P}^{(1,0)}_b \circ \mathsf{P}^{(0,1)}_d\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} = \left\|\mathsf{P}^{(1,1)}_{b\circ d}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \left\|\widehat{S}(b\circ d)\right\|_{CM} + \|E(b\circ d)\|_{\ell^\infty} .$$

It is clear that the above Corollary proves Theorem 1.2.

When the sequence $a = \{a_I\}_{I\in\mathcal{D}}$ is given by non-negative terms, then we have the following estimate that will be useful as well. It is proved simply by applying the Carleson Embedding Theorem.

Proposition 2.4. *Let* $a = \{a_I\}_{I\in\mathcal{D}}$ *be a sequence of non-negative numbers. Then*

$$\left\|\mathsf{P}^{(1,1)}_a\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \lesssim \left\|a^{\frac{1}{2}}\right\|^2_{CM} . \tag{2.13}$$

Proof. Let $f, g \in L^2(\mathbb{R})$. Then we have

$$\begin{aligned}\left|\left\langle \mathsf{P}_a^{(1,1)} f, g\right\rangle\right| &= \left|\sum_{I\in\mathcal{D}} a_I \left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})} \left\langle g, h_I^1\right\rangle_{L^2(\mathbb{R})}\right| \\ &\leq \sum_{I\in\mathcal{D}} a_I \left|\left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})} \left\langle g, h_I^1\right\rangle_{L^2(\mathbb{R})}\right| \\ &\leq \left(\sum_{I\in\mathcal{D}} a_I \left|\left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})}\right|^2\right)^{\frac{1}{2}} \left(\sum_{I\in\mathcal{D}} a_I \left|\left\langle g, h_I^1\right\rangle_{L^2(\mathbb{R})}\right|^2\right)^{\frac{1}{2}}.\end{aligned}$$

Now apply Lemma 2.1 and (2.3) to see that

$$\left(\sum_{I\in\mathcal{D}} a_I \left|\left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})}\right|^2\right)^{\frac{1}{2}} \lesssim \left\{\sup_{I\in\mathcal{D}} \frac{1}{|I|} \sum_{J\subset I} a_J\right\}^{\frac{1}{2}} \|f\|_{L^2(\mathbb{R})},$$

and so we have

$$\left|\left\langle \mathsf{P}_a^{(1,1)} f, g\right\rangle\right| = \left|\sum_{I\in\mathcal{D}} a_I \left\langle f, h_I^1\right\rangle_{L^2(\mathbb{R})} \left\langle g, h_I^1\right\rangle_{L^2(\mathbb{R})}\right| \lesssim \left\{\sup_{I\in\mathcal{D}} \frac{1}{|I|} \sum_{J\subset I} a_J\right\} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}$$

However,

$$\left\|a^{\frac{1}{2}}\right\|_{CM}^2 = \sup_{I\in\mathcal{D}} \frac{1}{|I|} \sum_{J\subset I} a_J,$$

and so the Proposition follows. ∎

3. Transplantation

We have

$$\mathsf{P}_b^{(\alpha,\beta)} h_I = \sum_{K\in\mathcal{D}} b_K \left\langle h_I, h_K^\beta\right\rangle_{L^2(\mathbb{R})} h_K^\alpha = \begin{cases} b_I\, h_I & \text{if} \quad (\alpha,\beta) = (0,0) \\ b_I\, h_I^1 & \text{if} \quad (\alpha,\beta) = (1,0) \\ \sum_{K\subsetneq I} b_K\, \widehat{h_K^1}(I)\; h_K & \text{if} \quad (\alpha,\beta) = (0,1) \\ \sum_{K\subsetneq I} b_K\, \widehat{h_K^1}\;(I)\, h_K^1 & \text{if} \quad (\alpha,\beta) = (1,1)\,. \end{cases}$$

From these formulas we can compute the Gram matrices of the composition of paraproducts. We will then choose an appropriate representation of Hilbert space on which to analyze a given Gram matrix. It is the simplicity of these formulas when $\beta = 0$ that accounts for our success in characterizing boundedness of products with type $(0, \beta, \gamma, 0)$.

At this point we also set forth some notation that will be used through out the remainder of this section. For the dyadic grid $\mathcal{D}$ we let $\ell^2(\mathcal{D})$ denote the standard space of square integrable sequences indexed by the dyadic intervals. For a weight function $\omega : \mathcal{D} \to \mathbb{R}_+$ we let $\ell^2(\omega)$ denote the sequences $\{a_I\}_{I\in\mathcal{D}}$ for which

$$\sum_{I\in\mathcal{D}} \omega(I)\, |a_I|^2 < \infty.$$

Recall now that we can identify the dyadic grid $\mathcal{D}$ on the real line with the standard Bergman tree of Carleson tiles on the upper plane by associating each $I \in \mathcal{D}$ with the

Carleson tile

$$T(I) \equiv I \times \left[\frac{|I|}{2}, |I|\right].$$

Also set

$$Q(I) \equiv I \times [0, |I|] = \bigcup_{J \subset I} T(J);$$

which is the Carleson square associated with $I \in \mathcal{D}$.

Let $\mathcal{H}$ denote the upper half plane, and so in particular we see that $\mathcal{H} = \bigcup_{I \in \mathcal{D}} T(I)$. We will let $L^2(\mathcal{H})$ denote the standard L^2 space on the upper half plane, and for a non-negative function σ we will let $L^2(\mathcal{H};\sigma)$ denote the functions that are square integrable with respect to $\sigma\, dA$, i.e,

$$\|f\|^2_{L^2(\mathcal{H})} \equiv \int_{\mathcal{H}} |f(z)|^2 \, dA(z) \quad \text{and} \quad \|f\|^2_{L^2(\mathcal{H};\sigma)} \equiv \int_{\mathcal{H}} |f(z)|^2 \sigma(z)\, dA(z).$$

Now consider the Hilbert subspace $L^2_c(\mathcal{H})$ which denotes the set of functions that are square integrable on $\mathcal{H}$, but are constant on tiles. Namely, $f : \mathcal{D} \to \mathbb{C}$ and can be represented as

$$f = \sum_{I \in \mathcal{D}} f_I \mathbf{1}_{T(I)}.$$

Then we have that

$$L^2_c(\mathcal{H}) \equiv \left\{ f : \mathcal{D} \to \mathbb{C} : \sum_{I \in \mathcal{D}} |f(I)|^2 |I|^2 < \infty \right\},$$

with norm $\|f\|_{L^2_c(\mathcal{H})} = \sqrt{\frac{1}{2}\sum_{I \in \mathcal{D}} |f(I)|^2 |I|^2}$.

For $f \in L^2(\mathcal{H})$, let $\widetilde{f} = \frac{f}{\|f\|_{L^2(\mathcal{H})}}$ denote the normalized function. Then it is immediate that $\left\{\widetilde{\mathbf{1}}_{T(I)}\right\}_{I \in \mathcal{D}}$ is an orthonormal basis of $L^2_c(\mathcal{H})$ and easy to see that $\left\{\widetilde{\mathbf{1}}_{Q(I)}\right\}_{I \in \mathcal{D}}$ is a Riesz basis of $L^2_c(\mathcal{H})$.

For $\lambda \in \mathbb{R}$ and $a \equiv \{a_I\}_{I \in \mathcal{D}}$ the multiplication operator $\mathcal{M}^\lambda_a$ is defined on basis elements $\widetilde{\mathbf{1}}_{T(K)}$ by

$$\mathcal{M}^\lambda_a \widetilde{\mathbf{1}}_{T(K)} = a_K |K|^\lambda \widetilde{\mathbf{1}}_{T(K)}.$$

Note that $\mathcal{M}^{-1}_a$ is <u>not</u> the inverse of $\mathcal{M}_a$!. We will also let $\overline{b} \equiv \left\{\overline{b_I}\right\}_{I \in \mathcal{D}}$.

Recall that in (1.3) for an interval $I \in \mathcal{D}$ and a function $f \in L^2(\mathbb{R})$ we let

$$\mathsf{Q}_I f \equiv \sum_{J \subset I} \langle f, h_J \rangle_{L^2(\mathbb{R})} h_J$$

denote the projection of the function f onto the span of the Haar functions supported within the interval I. For sequences $a \equiv \{a_I\}_{I \in \mathcal{D}}$ in (1.4) the operator Q_I takes the following form:

$$\mathsf{Q}_I a \equiv \sum_{J \subset I} a_J h_J.$$

We now study each remaining composition type in turn. The idea we will employ in the sections below is to study the composition $\mathsf{P}^{\epsilon,\delta}_b \circ \mathsf{P}^{\gamma,\eta}_d$ by relating it to an equivalent operator $\mathsf{T}^{(\epsilon,\delta,\gamma,\eta)}_{b,d}$ via a unitary operator that exchanges $L^2(\mathbb{R})$ with $L^2(\mathcal{H})$. The choice of unitary operator will be dictated by the signatures of the paraproducts being studied. By

transplanting the composition to this new operator, we can then study the boundedness $\mathsf{T}_{b,d}^{(\epsilon,\delta,\gamma,\eta)}$ through two-weight methods of harmonic analysis. Once an answer is obtained in that setting, the unitary operator lets us pull the answer back to the setting in $L^2(\mathbb{R})$ and provide an answer purely in terms of the original composition $\mathsf{P}_b^{\epsilon,\delta} \circ \mathsf{P}_d^{\gamma,\eta}$.

3.1. **Type $(0,1,1,0)$ Compositions.** The Gram matrix $\mathfrak{G}_{\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(1,0)}} = [G_{I,J}]_{I,J\in\mathcal{D}}$ of the operator $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ relative to the Haar basis $\{h_I\}_{I\in\mathcal{D}}$ has entries

$$\begin{aligned} G_{I,J} &= \left\langle \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)} h_J, h_I \right\rangle_{L^2(\mathbb{R})} = \left\langle \mathsf{P}_d^{(1,0)} h_J, \mathsf{P}_b^{(1,0)} h_I \right\rangle_{L^2(\mathbb{R})} \\ &= \left\langle d_J\, h_J^1, b_I\, h_I^1 \right\rangle_{L^2(\mathbb{R})} \\ &= \overline{b_I} d_J \frac{|I\cap J|}{|I|\,|J|} = \begin{cases} \overline{b_I} d_J \frac{1}{|I|} & \text{if} \quad J\subset I \\ \overline{b_I} d_J \frac{1}{|J|} & \text{if} \quad I\subset J \\ 0 & \text{if} \quad I\cap J=\emptyset. \end{cases} \end{aligned}$$

Define an operator $\mathsf{T}_{b,d}^{(0,1,1,0)}$ on $L_c^2(\mathcal{H})$ by

$$\mathsf{T}_{b,d}^{(0,1,1,0)} \equiv \mathcal{M}_{\bar{b}}^0 \left(\sum_{K\in\mathcal{D}} \widetilde{\mathbf{1}}_{T(K)} \otimes \widetilde{\mathbf{1}}_{Q(K)} \right) \mathcal{M}_d^{-1}.$$

Then the Gram matrix $\mathfrak{G}_{\mathsf{T}_{b,d}^{(0,1,1,0)}} = [G_{I,J}]_{I,J\in\mathcal{D}}$ of $\mathsf{T}_{b,d}^{(0,1,1,0)}$ relative to the basis $\left\{\widetilde{\mathbf{1}}_{T(I)}\right\}_{I\in\mathcal{D}}$ has entries

$$\begin{aligned} G_{I,J} &= \left\langle \mathsf{T}_{b,d}^{(0,1,1,0)} \widetilde{\mathbf{1}}_{T(J)}, \widetilde{\mathbf{1}}_{T(I)} \right\rangle_{L^2(\mathcal{H})} \\ &= \left\langle \mathcal{M}_{\bar{b}} \left(\sum_{K\in\mathcal{D}} \widetilde{\mathbf{1}}_{T(K)} \otimes \widetilde{\mathbf{1}}_{Q(K)} \right) \mathcal{M}_d^{-1} \widetilde{\mathbf{1}}_{T(J)}, \widetilde{\mathbf{1}}_{T(I)} \right\rangle_{L^2(\mathcal{H})} \\ &= \sum_{K\in\mathcal{D}} \left\langle \left\langle \widetilde{\mathbf{1}}_{Q(K)}, \mathcal{M}_d^{-1} \widetilde{\mathbf{1}}_{T(J)} \right\rangle_{L^2(\mathcal{H})} \mathcal{M}_{\bar{b}} \widetilde{\mathbf{1}}_{T(K)}, \widetilde{\mathbf{1}}_{T(I)} \right\rangle_{L^2(\mathcal{H})} \\ &= \sum_{K\in\mathcal{D}} \overline{b_K} d_J\, |J|^{-1} \left\langle \widetilde{\mathbf{1}}_{Q(K)}, \widetilde{\mathbf{1}}_{T(J)} \right\rangle_{L^2(\mathcal{H})} \left\langle \widetilde{\mathbf{1}}_{T(K)}, \widetilde{\mathbf{1}}_{T(I)} \right\rangle_{L^2(\mathcal{H})} \\ &= \overline{b_I} d_J\, |J|^{-1} \left\langle \widetilde{\mathbf{1}}_{Q(I)}, \widetilde{\mathbf{1}}_{T(J)} \right\rangle_{L^2(\mathcal{H})} \left\langle \widetilde{\mathbf{1}}_{T(I)}, \widetilde{\mathbf{1}}_{T(I)} \right\rangle_{L^2(\mathcal{H})} \\ &= \overline{b_I} d_J \sqrt{2} \frac{|Q(I)\cap T(J)|}{|I|\,|J|^2} = \frac{1}{\sqrt{2}} \begin{cases} \overline{b_I} d_J \frac{1}{|I|} & \text{if} \quad J\subset I \\ 0 & \text{if} \quad J\not\subset I. \end{cases} \end{aligned}$$

Thus, up to an absolute constant, $\mathfrak{G}_{\mathsf{T}_{b,d}^{(0,1,1,0)}}$ matches $\mathfrak{G}_{\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(1,0)}}$ in the lower triangle where $J\subset I$.

By the above computations we have

$$\left\| \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)} \right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \leq \left\| \mathsf{T}_{b,d}^{(0,1,1,0)} \right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})} + \left\| \mathsf{T}_{d,b}^{(0,1,1,0)} \right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})}$$

and we will further show below that

$$\begin{aligned}\left\|\mathsf{T}_{b,d}^{(0,1,1,0)}\right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})} &\approx C_1\\ \left\|\mathsf{T}_{d,b}^{(0,1,1,0)}\right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})} &\approx C_2\end{aligned}$$

with C_1 and C_2 the best constants in the testing inequality. However, for each of these constants we have

$$C_j \leq \left\|\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})},$$

see the argument just after (3.8), and so we obtain

$$\left\|\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \left\|\mathsf{T}_{b,d}^{(0,1,1,0)}\right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})} + \left\|\mathsf{T}_{d,b}^{(0,1,1,0)}\right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})}.$$

Now the operator norm $\left\|\mathsf{T}_{b,d}^{(0,1,1,0)}\right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})}$ equals the best constant in a certain two weight inequality for the positive operator U on $L_c^2(\mathcal{H})$, where

$$\mathsf{U} \equiv \sum_{K\in\mathcal{D}} \widetilde{\mathbf{1}}_{T(K)} \otimes \widetilde{\mathbf{1}}_{Q(K)}.$$

The inequality we wish to characterize is

$$\left\|\mathcal{M}_b^0 \mathsf{U} \mathcal{M}_d^{-1} f\right\|_{L_c^2(\mathcal{H})} = \left\|\mathsf{T}_{b,d}^{(0,1,1,0)} f\right\|_{L_c^2(\mathcal{H})} \lesssim \|f\|_{L_c^2(\mathcal{H})}, \tag{3.1}$$

which we first recast in the language of trees as in [1]. To do this, we suppose that f is constant on tiles $T(K)$ in the upper half space, and view f as the sequence $f : \mathcal{D} \longrightarrow \mathbb{C}$ given by its averages

$$f(K) \equiv \left\langle \frac{1}{|T(K)|}\mathbf{1}_{T(K)}, f\right\rangle_{L^2(\mathcal{H})}.$$

Define the adjoint tree integral $\mathcal{I}^* f$ by

$$\mathcal{I}^* f(K) \equiv \sum_{L\in\mathcal{D}:\ L\subset K} f(L), \qquad K \in \mathcal{D},$$

and define the special weight sequence $s(K) \equiv |K|$, $K \in \mathcal{D}$. Then for f constant on tiles $T(K)$ in the upper half space we have

$$\begin{aligned}\mathsf{U}f &= \sum_{K\in\mathcal{D}} \widetilde{\mathbf{1}}_{T(K)} \otimes \widetilde{\mathbf{1}}_{Q(K)} f = \sum_{K\in\mathcal{D}} \left\langle \widetilde{\mathbf{1}}_{Q(K)}, f\right\rangle_{L^2(\mathcal{H})} \widetilde{\mathbf{1}}_{T(K)}\\ &= \sum_{K\in\mathcal{D}} \left\langle \frac{1}{\sqrt{|Q(K)|}} \sum_{L\subset K} \mathbf{1}_{T(L)}, f\right\rangle_{L^2(\mathcal{H})} \frac{1}{\sqrt{|T(K)|}}\mathbf{1}_{T(K)}\\ &= \sum_{K\in\mathcal{D}} \frac{1}{\sqrt{|Q(K)|}} \sum_{L\subset K} |T(L)| \left\langle \frac{1}{|T(L)|}\mathbf{1}_{T(L)}, f\right\rangle_{L^2(\mathcal{H})} \frac{1}{\sqrt{|T(K)|}}\mathbf{1}_{T(K)}\\ &= \frac{1}{\sqrt{2}} \sum_{K\in\mathcal{D}} \left\{\sum_{L\subset K} \frac{1}{2} s(L)^2 f(L)\right\} \frac{1}{\frac{1}{2}|K|^2}\mathbf{1}_{T(K)}\\ &= \sqrt{2} \sum_{K\in\mathcal{D}} \mathcal{I}^*\left(s^2 f\right)(K) \frac{1}{s(K)^2}\mathbf{1}_{T(K)},\end{aligned}$$

which shows that

$$(\mathsf{U}f)(K) = \sqrt{2}\frac{1}{s(K)^2}\mathcal{I}^*\left(s^2 f\right)(K).$$

Since $\mathcal{M}_d^{-1}$ and $\mathcal{M}_{\overline{b}}^0$ are multiplication by $\frac{d_K}{|K|} = \frac{d_K}{s(K)}$ and $\overline{b_K}$ respectively on the tile $T(K)$, which for convenience we abbreviate as $\frac{d}{s}$ and $\overline{b}$ respectively, we see that the two weight inequality (3.1) is equivalent to

$$\left\| s\left|\overline{b}\right| \frac{1}{s^2}\mathcal{I}^*\left(s^2\frac{|d|}{s}f\right)\right\|_{\ell^2(\mathcal{D})} \lesssim \|sf\|_{\ell^2(\mathcal{D})}. \tag{3.2}$$

Now if we set

$$\begin{aligned} s|d|f &= g\omega, \\ (sf)^2 &= g^2\omega, \\ \left(\frac{|b|}{s}\right)^2 &= \sigma, \end{aligned}$$

then

$$\omega = \frac{(g\omega)^2}{g^2\omega} = \frac{(s|d|f)^2}{(sf)^2} = d^2,$$

and (3.2) is equivalent to

$$\|\mathcal{I}^*(g\omega)\|_{\ell^2(\sigma)} \leq C\|g\|_{\ell^2(\omega)}. \tag{3.3}$$

At this point we can apply the characterization of the two weight tree inequality in [1]. Now $\mathcal{D}$ is a rootless tree, and the inequality in [1] is stated for a rooted tree, but the monotone convergence theorem immediately extends the characterization in [1] to rootless trees as well. Thus the best constant C in (3.3) is comparable to the best constant C_1 in the corresponding truncated testing condition with $g = \mathbf{1}_{\{L\in D:\ L\subset I\}}$ for $I \in \mathcal{D}$:

$$\begin{aligned} \sum_{J\subset I}\left(\sum_{L\subset J}\omega(L)\right)^2\sigma(J) &= \sum_{J\subset I}\left[\mathcal{I}^*\omega(J)\right]^2\sigma(J) \leq \|\mathcal{I}^*(g\omega)\|_{\ell^2(\sigma)}^2 \\ &\leq C_1^2\|g\|_{\ell^2(\omega)} = C_1^2\mathcal{I}^*\omega(I) = C_1^2\sum_{L\subset I}\omega(L), \qquad I\in\mathcal{D}, \end{aligned}$$

i.e.

$$\sum_{J\subset I}\left(\sum_{L\subset J}|d_L|^2\right)^2\frac{|b_J|^2}{|J|^2} \leq C_1^2\sum_{L\subset K}|d_L|^2, \qquad K\in\mathcal{D}. \tag{3.4}$$

It is now convenient to relabel our weights by introducing the <u>different</u> notation,

$$\begin{aligned} w &\equiv \sum_{I\in\mathcal{D}}|b_I|^2\mathbf{1}_{T(I)} \\ \sigma &\equiv \sum_{I\in\mathcal{D}}\frac{|d_I|^2}{|I|^2}\mathbf{1}_{T(I)}, \end{aligned}$$

in which the testing condition (3.4) is equivalent to

$$\left\|\mathbf{1}_{Q(I)}\mathsf{U}\left(\sigma\mathbf{1}_{Q(I)}\right)\right\|_{L^2(\mathcal{H};w)}^2 \leq C_1^2\left\|\mathbf{1}_{Q(I)}\right\|_{L^2(\mathcal{H};\sigma)}^2, \tag{3.5}$$

since

$$\begin{aligned}
\left\|\mathbf{1}_{Q(I)}\mathsf{U}\left(\sigma\mathbf{1}_{Q(I)}\right)\right\|_{L^2(\mathcal{H};w)}^2 &= 2\left\|\sum_{J\subset I}\frac{\left\langle\sigma\mathbf{1}_{Q(I)},\widetilde{\mathbf{1}}_{Q(J)}\right\rangle_{L^2(\mathcal{H})}}{|J|}\mathbf{1}_{T(J)}\right\|_{L^2(\mathcal{H};w)}^2 \\
&= 2\sum_{J\subset I}|b_I|^2\left(\left\langle\sigma\mathbf{1}_{Q(I)},\widetilde{\mathbf{1}}_{Q(J)}\right\rangle_{L^2(\mathcal{H})}\right)^2 \\
&= 2\sum_{J\subset I}|b_I|^2\frac{1}{|J|^2}\left(\sum_{L\in\mathcal{D}}\frac{|d_L|^2}{|L|^2}\int_{\mathcal{H}}\mathbf{1}_{T(L)}\mathbf{1}_{Q(I)}\mathbf{1}_{Q(J)}\,dA\right)^2 \\
&= \frac{1}{2}\sum_{J\subset I}|b_I|^2\frac{1}{|J|^2}\left(\sum_{L\subset J}|d_L|^2\right)^2.
\end{aligned}$$

and

$$\|\mathsf{Q}_I d\|_{L^2(\mathbb{R})}^2 = \sum_{L\subset I}|d_L|^2 = \left\|1_{Q(I)}\right\|_{L^2(\mathcal{H};\sigma)}^2. \tag{3.6}$$

The testing condition (3.5) thus gives a characterization of the boundedness of the paraproduct composition $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(1,0)}$ on $L^2(\mathbb{R})$. However, we now want to rephrase this as a testing condition, but only on the operator $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(1,0)}$.

Let $V: L^2(\mathbb{R})\to L_c^2(\mathcal{H})$ be the unitary operator defined on basis elements by

$$Vh_I = \widetilde{\mathbf{1}}_{T(I)}. \tag{3.7}$$

Using these unitary operators we can write

$$\begin{aligned}
\left\|\mathsf{Q}_I\mathsf{P}_b^{(0,1)}\mathsf{P}_d^{(1,0)}\mathsf{Q}_I\overline{d}\right\|_{L^2(\mathbb{R})} &= \left\|\mathsf{Q}_I V^*\left(T_{b,d}^{(0,1,1,0)}\right)V\mathsf{Q}_I\overline{d}\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathsf{Q}_I V^*\left(T_{b,d}^{(0,1,1,0)}\right)\left(\sum_{J\subset I}\overline{d_J}\,\widetilde{\mathbf{1}}_{T(J)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathsf{Q}_I V^*\mathcal{M}_{\overline{b}}^0\mathsf{U}\mathcal{M}_d^{-1}\left(\sum_{J\subset I}\overline{d_J}\,\widetilde{\mathbf{1}}_{T(J)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathsf{Q}_I V^*\mathcal{M}_{\overline{b}}^0\mathsf{U}\left(\sigma\mathbf{1}_{Q(I)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|V^*\mathbf{1}_{Q(I)}\mathcal{M}_{\overline{b}}^0\mathsf{U}\left(\sigma\mathbf{1}_{Q(I)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathbf{1}_{Q(I)}\mathcal{M}_{\overline{b}}^0\mathsf{U}\left(\sigma\mathbf{1}_{Q(I)}\right)\right\|_{L_c^2(\mathcal{H})} \\
&= \left\|\mathbf{1}_{Q(I)}\mathsf{U}\left(\sigma\mathbf{1}_{Q(I)}\right)\right\|_{L_c^2(\mathcal{H};w)}.
\end{aligned}$$

Finally, using (3.6) one sees that (3.5) is equivalent to a simple testing condition on the composition $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$:

$$\left\| \mathsf{Q}_I \mathsf{P}_b^{(0,1)} \mathsf{P}_d^{(1,0)} \mathsf{Q}_I \overline{d} \right\|_{L^2(\mathbb{R})} \lesssim \left\| \mathsf{Q}_I d \right\|_{L^2(\mathbb{R})}^2 . \tag{3.8}$$

Furthermore, it is immediate to see that (3.8) is implied by the boundedness of $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ on $L^2(\mathbb{R})$. Interchanging the roles of b and d we have the following Theorem that characterizes the boundedness of $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$, which is just a restatement of Theorem 1.3.

Theorem 3.1. *The composition $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ is bounded on $L^2(\mathbb{R})$ if and only if both*

$$\begin{aligned} \left\| \mathsf{Q}_I \mathsf{P}_b^{(0,1)} \mathsf{P}_d^{(1,0)} \left(\mathsf{Q}_I \overline{d} \right) \right\|_{L^2(\mathbb{R})}^2 &\leq C_1^2 \left\| \mathsf{Q}_I d \right\|_{L^2(\mathbb{R})}^2 ; \\ \left\| \mathsf{Q}_I \mathsf{P}_d^{(0,1)} \mathsf{P}_b^{(1,0)} \left(\mathsf{Q}_I \overline{b} \right) \right\|_{L^2(\mathbb{R})}^2 &\leq C_2^2 \left\| \mathsf{Q}_I b \right\|_{L^2(\mathbb{R})}^2 \end{aligned}$$

for all $I \in \mathcal{D}$; i.e. for all $I \in \mathcal{D}$ the following inequalities are true

$$\begin{aligned} \sum_{J \subset I} \left| b_J \right|^2 \frac{1}{|J|^2} \left(\sum_{L \subset J} \left| d_L \right|^2 \right)^2 &\leq C_1^2 \sum_{L \subset I} \left| d_L \right|^2 ; \\ \sum_{J \subset I} \left| d_J \right|^2 \frac{1}{|J|^2} \left(\sum_{L \subset J} \left| b_L \right|^2 \right)^2 &\leq C_2^2 \sum_{L \subset I} \left| b_L \right|^2 . \end{aligned}$$

Moreover, the norm of $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ on $L^2(\mathbb{R})$ satisfies

$$\left\| \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)} \right\|_{L^2(\mathbb{R}) \to L^2(\mathbb{R})} \approx C_1 + C_2$$

where C_1 and C_2 are the best constants in appearing above.

3.2. **Type $(0,1,0,0)$ Compositions.** The Gram matrix $\mathfrak{G}_{\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)}} = [G_{I,J}]_{I,J \in \mathcal{D}}$ of the operator $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)}$ relative to the Haar basis $\{h_I\}_{I \in \mathcal{D}}$ has entries given by

$$\begin{aligned} G_{I,J} &= \left\langle \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)} h_J, h_I \right\rangle_{L^2(\mathbb{R})} = \left\langle \mathsf{P}_d^{(0,0)} h_J, \mathsf{P}_b^{(1,0)} h_I \right\rangle_{L^2(\mathbb{R})} \\ &= \left\langle d_J h_J, b_I h_I^1 \right\rangle_{L^2(\mathbb{R})} \\ &= \overline{b_I} d_J \widehat{h_I^1}(J) = \begin{cases} \overline{b_I} d_J \frac{-1}{\sqrt{|J|}} & \text{if} \quad I \subset J_- \\ \overline{b_I} d_J \frac{1}{\sqrt{|J|}} & \text{if} \quad I \subset J_+ \\ 0 & \text{if} \quad J \subset I \text{ or } I \cap J = \emptyset. \end{cases} \end{aligned}$$

Now consider the operator $\mathsf{T}_{b,d}^{(0,1,0,0)}$ defined by

$$\mathsf{T}_{b,d}^{(0,1,0,0)} \equiv \mathcal{M}_{\overline{b}}^{-1} \left(\sum_{K \in \mathcal{D}} \widetilde{\mathbf{1}}_{Q_\pm(K)} \otimes \widetilde{\mathbf{1}}_{T(K)} \right) \mathcal{M}_d^{\frac{1}{2}},$$

where

$$\mathbf{1}_{Q_\pm(K)} \equiv - \sum_{L \subset K_-} \mathbf{1}_{T(L)} + \sum_{L \subset K_+} \mathbf{1}_{T(L)}. \tag{3.9}$$

A straightforward computation shows that

$$\begin{aligned}\left\|\mathbf{1}_{Q_{\pm}(K)}\right\|_{L^2(\mathcal{H})} &= \frac{|K|}{2};\\ \mathcal{M}_a^{\lambda}\mathbf{1}_{Q_{\pm}(K)} &= -\sum_{L\subset K_-} a_L\,|L|^{\lambda}\,\mathbf{1}_{T(L)} + \sum_{L\subset K_+} a_L\,|L|^{\lambda}\,\mathbf{1}_{T(L)}.\end{aligned}$$

The Gram matrix $\mathfrak{G}_{\mathsf{T}_{b,d}^{(0,1,0,0)}} = [G_{I,J}]_{I,J\in\mathcal{D}}$ of $\mathsf{T}_{b,d}^{(0,1,0,0)}$ relative to the basis $\left\{\widetilde{\mathbf{1}}_{T(I)}\right\}_{I\in\mathcal{D}}$ then has entries given by

$$\begin{aligned}G_{I,J} &= \left\langle \mathsf{T}_{b,d}^{(0,1,0,0)}\widetilde{\mathbf{1}}_{T(J)}, \widetilde{\mathbf{1}}_{T(I)}\right\rangle_{L^2(\mathcal{H})}\\ &= \left\langle \mathcal{M}_{\overline{b}}^{-1}\left(\sum_{K\in\mathcal{D}}\widetilde{\mathbf{1}}_{Q_{\pm}(K)}\otimes\widetilde{\mathbf{1}}_{T(K)}\right)\mathcal{M}_d^{\frac{1}{2}}\widetilde{\mathbf{1}}_{T(J)}, \widetilde{\mathbf{1}}_{T(I)}\right\rangle_{L^2(\mathcal{H})}\\ &= \sum_{K\in\mathcal{D}}\left\langle\left\langle\widetilde{\mathbf{1}}_{T(K)}, \mathcal{M}_d^{\frac{1}{2}}\widetilde{\mathbf{1}}_{T(J)}\right\rangle_{L^2(\mathcal{H})}\mathcal{M}_{\overline{b}}^{-1}\widetilde{\mathbf{1}}_{Q_{\pm}(K)}, \widetilde{\mathbf{1}}_{T(I)}\right\rangle_{L^2(\mathcal{H})}\\ &= \sum_{K\in\mathcal{D}} d_J\,|J|^{\frac{1}{2}}\left\langle\widetilde{\mathbf{1}}_{T(K)}, \widetilde{\mathbf{1}}_{T(J)}\right\rangle_{L^2(\mathcal{H})}\left\langle\mathcal{M}_{\overline{b}}^{-1}\widetilde{\mathbf{1}}_{Q_{\pm}(K)}, \widetilde{\mathbf{1}}_{T(I)}\right\rangle_{L^2(\mathcal{H})}\\ &= d_J\,|J|^{\frac{1}{2}}\left\langle\mathcal{M}_{\overline{b}}^{-1}\widetilde{\mathbf{1}}_{Q_{\pm}(J)}, \widetilde{\mathbf{1}}_{T(I)}\right\rangle_{L^2(\mathcal{H})}\\ &= 2\sqrt{2}d_J\,|J|^{\frac{1}{2}}\,|J|^{-1}\,|I|^{-1}\left\langle -\sum_{L\subset J_-}\overline{b_L}\,|L|^{-1}\,\mathbf{1}_{T(L)} + \sum_{L\subset J_+}\overline{b_L}\,|L|^{-1}\,\mathbf{1}_{T(L)}, \mathbf{1}_{T(I)}\right\rangle_{L^2(\mathcal{H})}\\ &= \sqrt{2}\begin{cases} -\overline{b_I}d_J\,|J|^{-\frac{1}{2}} & \text{if} \quad I\subset J_-\\ \overline{b_I}d_J\,|J|^{-\frac{1}{2}} & \text{if} \quad I\subset J_+\\ 0 & \text{if} \quad J\subset I \text{ or } I\cap J=\emptyset.\end{cases}\end{aligned}$$

Thus, up to an absolute constant, we see that $\mathfrak{G}_{\mathsf{T}_{b,d}^{(0,1,0,0)}} = \mathfrak{G}_{\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}}$, and we obtain the following conclusion

$$\left\|\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \approx \left\|\mathsf{T}_{b,d}^{(0,1,0,0)}\right\|_{L^2(\mathcal{H})\to L^2(\mathcal{H})}.$$

Now the operator norm $\left\|\mathsf{T}_{b,d}^{(0,1,0,0)}\right\|_{L_c^2(\mathcal{H})\to L_c^2(\mathcal{H})}$ equals the best constant in a certain two weight inequality for the operator U on $L^2(\mathcal{H})$ defined by

$$\mathsf{U} \equiv \sum_{K\in\mathcal{D}}\widetilde{\mathbf{1}}_{Q_{\pm}(K)}\otimes\widetilde{\mathbf{1}}_{T(K)}.$$

This operator is not positive, but its singular character is well-behaved, and the best constant in a certain two weight inequality associated to U is in turn comparable to the best constants in the associated testing conditions. These testing conditions thus give a characterization of the boundedness of the paraproduct composition $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}$ on $L^2(\mathbb{R})$. Here are the details which provide this reduction.

By the computations above, the inequality we wish to characterize is:

$$\left\|\mathcal{M}_{\bar{b}}^{-1}\mathsf{U}\mathcal{M}_{d}^{\frac{1}{2}}f\right\|_{L^2_c(\mathcal{H})}=\left\|\mathsf{T}_{b,d}^{(0,1,0,0)}f\right\|_{L^2_c(\mathcal{H})}\lesssim\|f\|_{L^2_c(\mathcal{H})}\,. \tag{3.10}$$

Now if $f=\sum_{I\in\mathcal{D}}f_I\mathbf{1}_{T(I)}$ then $\mathcal{M}_d^{\frac{1}{2}}f=\sum_{I\in\mathcal{D}}d_I\sqrt{|I|}f_I\mathbf{1}_{T(I)}$, and if we define $g=\mathcal{M}_d^{\frac{1}{2}}f$ then inequality (3.10) is equivalent to:

$$\|\mathsf{U}g\|_{L^2_c(\mathcal{H};w)}\lesssim\|g\|_{L^2_c(\mathcal{H};\nu)}\,, \tag{3.11}$$

where the weights w and ν are given by

$$\begin{aligned} w &\equiv \sum_{I\in\mathcal{D}}|b_I|^2\,|I|^{-2}\,\mathbf{1}_{T(I)}\\ \nu &\equiv \sum_{I\in\mathcal{D}}|d_I|^{-2}\,|I|^{-1}\,\mathbf{1}_{T(I)}. \end{aligned}$$

This follows because a straightforward computation shows that for $k=\sum_{I\in\mathcal{D}}k_I\mathbf{1}_{T(I)}$

$$\left\|\mathcal{M}_{\bar{b}}^{\frac{1}{2}}k\right\|^2_{L^2_c(\mathcal{H})}=\frac{1}{2}\sum_{I\in\mathcal{D}}|b_I|^2\,|k_I|^2\,|I|^4=\|k\|^2_{L^2_c(\mathcal{H};w)}$$

and that for $f=\left(\mathcal{M}_d^{\frac{1}{2}}\right)^{-1}g=\sum_{I\in\mathcal{D}}g_Id_I^{-1}\,|I|^{-\frac{1}{2}}\,\mathbf{1}_{T(I)}$ one has

$$\|f\|^2_{L^2_c(\mathcal{H})}=\frac{1}{2}\sum_{I\in\mathcal{D}}|g_I|^2\,|d_I|^{-2}\,|I|=\|g\|^2_{L^2_c(\mathcal{H};\nu)}\,.$$

Now, let

$$\sigma\equiv\sum_{I\in\mathcal{D}}|d_I|^2\,|I|\,\mathbf{1}_{T(I)}$$

and substitute $g=h\sigma$ into (3.11) to see that (3.10) is in terms of weighted L^2 norms equivalent to

$$\|\mathsf{U}\,(\sigma h)\|_{L^2_c(\mathcal{H};w)}\lesssim\|h\|_{L^2_c(\mathcal{H};\sigma)}\,, \tag{3.12}$$

By Theorem 3.3 we have that the best constant in (3.12) is equivalent to best constants in the testing conditions given by

$$\left\|\mathsf{U}\left(\sigma\mathbf{1}_{T(I)}\right)\right\|_{L^2_c(\mathcal{H};w)}\leq C_1\left\|\mathbf{1}_{T(I)}\right\|_{L^2_c(\mathcal{H};\sigma)} \tag{3.13}$$

$$\left\|\mathbf{1}_{Q(I)}\mathsf{U}^*\left(w\mathbf{1}_{Q(I)}\right)\right\|_{L^2_c(\mathcal{H};\sigma)}\leq C_2\left\|\mathbf{1}_{Q(I)}\right\|_{L^2_c(\mathcal{H};w)}\,. \tag{3.14}$$

We now phrase these conditions in terms of paraproduct type testing conditions. For our special choice of measures σ and w given above, a simple computation shows that right-hand side of (3.13) is

$$\left\|\mathbf{1}_{T(I)}\right\|_{L^2_c(\mathcal{H};\sigma)}=\frac{|d_I|\,|I|^{\frac{3}{2}}}{\sqrt{2}}.$$

While the left-hand side of (3.13) yields

$$\begin{aligned}\left\|\mathsf{U}\left(\sigma\mathbf{1}_{T(I)}\right)\right\|_{L^2_c(\mathcal{H};w)} &= \sqrt{2}\,|I|\,|d_I|^2\left\|\widetilde{\mathbf{1}}_{Q_\pm(I)}\right\|_{L^2_c(\mathcal{H};w)}\\ &= |I|\,|d_I|^2\left(\sum_{L\subset I_-}|b_L|^2+\sum_{L\subset I_+}|b_L|^2\right)^{\frac{1}{2}}\\ &= |I|\,|d_I|^2\left(\sum_{L\subsetneq I}|b_L|^2\right)^{\frac{1}{2}}.\end{aligned}$$

Thus, for our special choice of measures, (3.13) is equivalent to

$$\sup_{I\in\mathcal{D}}|d_I|\left(\frac{1}{|I|}\sum_{L\subsetneq I}|b_L|^2\right)^{\frac{1}{2}}\lesssim 1.$$

And, since $\left(\frac{1}{|I|}\sum_{L\subsetneq I}|b_L|^2\right)^{\frac{1}{2}}=\left\|\mathsf{P}_b^{(0,1)}h_I\right\|_{L^2(\mathbb{R})}$ we conclude that (3.13) is implied by

$$|d_I|\left\|\mathsf{P}_b^{(0,1)}h_I\right\|_{L^2(\mathbb{R})}\lesssim 1\quad\forall I\in\mathcal{D}.$$

Furthermore, it is clear that this last condition is implied by the boundedness of the operator $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}$ from $L^2(\mathbb{R})\to L^2(\mathbb{R})$. For if g is any function, then we have

$$\begin{aligned}\|g\|_{L^2(\mathbb{R})}\left\|\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} &\geq \left|\left\langle\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}h_I,g\right\rangle_{L^2(\mathbb{R})}\right|\\ &= |d_I|\left|\left\langle\mathsf{P}_b^{(0,1)}h_I,g\right\rangle_{L^2(\mathbb{R})}\right|.\end{aligned}$$

Choosing $g=\mathsf{P}_b^{(0,1)}h_I$ yields,

$$|d_I|\left(\frac{1}{|I|}\sum_{L\subsetneq I}|b_L|^2\right)^{\frac{1}{2}}=|d_I|\left\|\mathsf{P}_b^{(0,1)}h_I\right\|_{L^2(\mathbb{R})}\leq\left\|\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}.$$

Turning to (3.14) one easily computes that the right-hand side is given by

$$\left\|\mathbf{1}_{Q(I)}\right\|^2_{L^2_c(\mathcal{H};w)}=\frac{1}{2}\sum_{J\subset I}|b_J|^2=\frac{1}{2}\left\|\mathsf{Q}_Ib\right\|^2_{L^2(\mathbb{R})}.$$

We now provide an alternate, equivalent, way to study the backward testing condition. Let $V:L^2(\mathbb{R})\to L^2_c(\mathcal{H})$ be the unitary operator defined in (3.7)

Next, observe that $w\mathbf{1}_{Q(I)}=\sum_{J\subset I}|b_J|^2|J|^{-2}\mathbf{1}_{T(J)}$. Then, using these unitary operators we have that

$$\begin{aligned}
\left\|\mathsf{Q}_I\mathsf{P}_d^{(0,0)}\mathsf{P}_b^{(1,0)}\mathsf{Q}_I\bar{b}\right\|_{L^2(\mathbb{R})} &= \left\|\mathsf{Q}_I V^*\left(T_{b,d}^{(0,1,0,0)}\right)^* V\mathsf{Q}_I\bar{b}\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathsf{Q}_I V^*\left(T_{b,d}^{(0,1,0,0)}\right)^*\left(\sum_{J\subset I}\overline{b_J}\,\widetilde{\mathbf{1}}_{T(J)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathsf{Q}_I V^*\mathcal{M}_{\overline{d}}^{\frac{1}{2}}\mathsf{U}^*\mathcal{M}_b^{-1}\left(\sum_{J\subset I}\overline{b_J}\,\widetilde{\mathbf{1}}_{T(J)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathsf{Q}_I V^*\mathcal{M}_{\overline{d}}^{\frac{1}{2}}\mathsf{U}^*\left(w\mathbf{1}_{Q(I)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|V^*\mathbf{1}_{Q(I)}\mathcal{M}_{\overline{d}}^{\frac{1}{2}}\mathsf{U}^*\left(w\mathbf{1}_{Q(I)}\right)\right\|_{L^2(\mathbb{R})} \\
&= \left\|\mathbf{1}_{Q(I)}\mathcal{M}_{\overline{d}}^{\frac{1}{2}}\mathsf{U}^*\left(w\mathbf{1}_{Q(I)}\right)\right\|_{L^2_c(\mathcal{H})} \\
&= \left\|\mathbf{1}_{Q(I)}\mathsf{U}^*\left(w\mathbf{1}_{Q(I)}\right)\right\|_{L^2_c(\mathcal{H};\sigma)}.
\end{aligned}$$

These computation show that backward testing is equivalent to the following

$$\left\|\mathsf{Q}_I\mathsf{P}_d^{(0,0)}\mathsf{P}_b^{(1,0)}\mathsf{Q}_I\bar{b}\right\|_{L^2(\mathbb{R})} \lesssim \left\|\mathsf{Q}_I\bar{b}\right\|_{L^2(\mathbb{R})}.$$

This condition is again clearly implied by the boundedness of the operator $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}$ on $L^2(\mathbb{R})$. Finally, we note that

$$\left\|\mathbf{1}_{Q(I)}\mathsf{U}^*(w\mathbf{1}_{Q(I)})\right\|^2_{L^2_c(\mathcal{H};\sigma)} = \sum_{J\subset I}\frac{|d_J|^2}{|J|}\left(\sum_{K\subset J_+}|b_K|^2 - \sum_{K\subset J_-}|b_K|^2\right)^2.$$

To see this, note that

$$\left\|\mathbf{1}_{Q(I)}\mathsf{U}^*(w\mathbf{1}_{Q(I)})\right\|^2_{L^2_c(\mathcal{H};\sigma)} = 4\sum_{J\subset I}\frac{|d_J|^2}{|J|}\left|\left\langle w\mathbf{1}_{Q(I)},\mathbf{1}_{Q_\pm(J)}\right\rangle_{L^2(\mathcal{H})}\right|^2.$$

But, observe that we have $\mathbf{1}_{Q_\pm(J)} = -\mathbf{1}_{Q(J_-)} + \mathbf{1}_{Q(J_+)}$, and that if $L\subset K$

$$\mathbf{1}_{Q(L)}\mathbf{1}_{Q(K)} = \mathbf{1}_{Q(L)}.$$

Using these observations, we find that

$$\begin{aligned}
4\sum_{J\subset I}\frac{|d_J|^2}{|J|}\left|\left\langle w\mathbf{1}_{Q(I)},\mathbf{1}_{Q_\pm(J)}\right\rangle_{L^2(\mathcal{H})}\right|^2 &= 4\sum_{J\subset I}\frac{|d_J|^2}{|J|}\left|\left\langle w,\mathbf{1}_{Q(J_+)}-\mathbf{1}_{Q(J_-)}\right\rangle_{L^2(\mathcal{H})}\right|^2 \\
&= \sum_{J\subset I}\frac{|d_J|^2}{|J|}\left(\sum_{K\subset J_+}|b_K|^2 - \sum_{K\subset J_-}|b_K|^2\right)^2.
\end{aligned}$$

Therefore, we have the following theorem providing the boundedness in terms of testing conditions on the paraproduct $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,0)}$. This is just a restatement of Theorem 1.4.

Theorem 3.2. *The composition* $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)}$ *is bounded on* $L^2(\mathbb{R})$ *if and only if both*

$$\begin{aligned} |d_I| \left\| \mathsf{P}_b^{(0,1)} h_I \right\|_{L^2(\mathbb{R})} &\leq C_1; \\ \left\| \mathsf{Q}_I \mathsf{P}_d^{(0,0)} \mathsf{P}_b^{(1,0)} \mathsf{Q}_I \bar{b} \right\|_{L^2(\mathbb{R})} &\leq C_2 \left\| \mathsf{Q}_I b \right\|_{L^2(\mathbb{R})} \end{aligned}$$

for all $I \in \mathcal{D}$*.; i.e. for all* $I \in \mathcal{D}$ *the following inequalities are true*

$$\begin{aligned} |d_I| \left(\frac{1}{|I|} \sum_{L \subsetneq I} |b_L|^2 \right)^{\frac{1}{2}} &\leq C_1; \\ \left(\sum_{J \subset I} \frac{|d_J|^2}{|J|} \left(\sum_{K \subset J_+} |b_K|^2 - \sum_{K \subset J_-} |b_K|^2 \right)^2 \right)^{\frac{1}{2}} &\leq C_2 \left(\sum_{L \subset I} |b_L|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Moreover, the norm of $\mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(1,0)}$ *on* $L^2(\mathbb{R})$ *satisfies*

$$\left\| \mathsf{P}_b^{(0,1)} \circ \mathsf{P}_d^{(0,0)} \right\|_{L^2(\mathbb{R}) \to L^2(\mathbb{R})} \approx C_1 + C_2$$

where C_1 *and* C_2 *are the best constants in appearing above.*

3.2.1. *A Discrete T1 Theorem with Different Bases.* We will prove the following Theorem by adapting the proof strategy from Nazarov, Treil and Volberg in [3]. Recall that for $K \in \mathcal{D}$ that we have defined

$$\mathbf{1}_{Q_\pm(K)} \equiv - \sum_{L \subset K_-} \mathbf{1}_{T(L)} + \sum_{L \subset K_+} \mathbf{1}_{T(L)}.$$

Theorem 3.3. *Let*

$$\mathsf{U} \equiv \sum_{K \in \mathcal{D}} \widetilde{\mathbf{1}}_{Q_\pm(K)} \otimes \widetilde{\mathbf{1}}_{T(K)}$$

and suppose that μ *and* ν *are positive measures on* $\mathcal{H}$ *that are constant on tiles, i.e.,*

$$\begin{aligned} \mu &\equiv \sum_{I \in \mathcal{D}} \mu_I \mathbf{1}_{T(I)} \\ \nu &\equiv \sum_{I \in \mathcal{D}} \nu_I \mathbf{1}_{T(I)}. \end{aligned}$$

Then

$$\mathsf{U}(\mu \cdot) : L^2_c(\mathcal{H}; \mu) \to L^2_c(\mathcal{H}; \nu)$$

if and only if both

$$\begin{aligned} \left\| \mathsf{U}\left(\mu \mathbf{1}_{T(I)}\right) \right\|_{L^2_c(\mathcal{H};\nu)} &\leq C_1 \left\| \mathbf{1}_{T(I)} \right\|_{L^2_c(\mathcal{H};\mu)} = \sqrt{\mu(T(I))}, \\ \left\| \mathbf{1}_{Q(I)} \mathsf{U}^* \left(\nu \mathbf{1}_{Q(I)}\right) \right\|_{L^2_c(\mathcal{H};\mu)} &\leq C_2 \left\| \mathbf{1}_{Q(I)} \right\|_{L^2_c(\mathcal{H};\nu)} = \sqrt{\nu(Q(I))}, \end{aligned}$$

hold for all $I \in \mathcal{D}$*. Moreover, we have that*

$$\left\| \mathsf{U} \right\|_{L^2_c(\mathcal{H};\mu) \to L^2_c(\mathcal{H};\nu)} \approx C_1 + C_2$$

where C_1 *and* C_2 *are the best constants appearing above.*

Proof. Note that

$$\mathsf{U}_\mu(f) = \mathsf{U}\left(f\mu\right) = \sum_{K\in\mathcal{D}} \left\langle f\mu, \widetilde{\mathbf{1}}_{T(K)}\right\rangle_{L^2(\mathcal{H})} \widetilde{\mathbf{1}}_{Q_\pm(K)}.$$

For notational simplicity, in this proof only, we let $\nu(J) \equiv \nu(Q(J))$ (i.e., we implicitly identify J with $Q(J)$). Now the weight adapted orthonormal bases are given by

$$\{h_I^\mu\}_{I\in\mathcal{D}} \text{ and } \{H_J^\nu\}_{J\in\mathcal{D}},$$

with

$$h_I^\mu \equiv \frac{\widetilde{\mathbf{1}}_{T(I)}}{\sqrt{\mu_I}} \text{ and } H_J^\nu \equiv \widetilde{\nu}\left(J\right)\left(-\frac{\mathbf{1}_{Q(J_+)}}{\nu\left(J_+\right)} + \frac{\mathbf{1}_{Q(J_-)}}{\nu\left(J_-\right)}\right),$$

where

$$\widetilde{\nu}\left(J\right) \equiv \sqrt{\frac{\nu\left(J_+\right)\nu\left(J_-\right)}{\nu\left(I_+\right)+\nu\left(I_-\right)}}.$$

Let $\widehat{f}_\mu$ denote the "Haar coefficient" of f with respect to the basis h_I^μ, i.e.,

$$\widehat{f}_\mu(I) \equiv \langle f, h_I^\mu\rangle_{L^2(\mathcal{H};\mu)},$$

and similarly for $\widehat{g}_\nu(J)$. We can now expand the function f and g with respect to these weighted orthonormal bases and write $f = \sum_{I\in\mathcal{D}} \widehat{f}_\mu\left(I\right) h_I^\mu$ and $g = \sum_{J\in\mathcal{D}} \widehat{g}_\nu\left(J\right) H_\nu^J$. Doing so, we then see that

$$\begin{aligned}
\langle \mathsf{U}_\mu f, g\rangle_{L^2_c(\mathcal{H};\nu)} &= \sum_{I,J\in\mathcal{D}} \widehat{f}_\mu\left(I\right)\widehat{g}_\nu\left(J\right)\left\langle \mathsf{U}_\mu h_I^\mu, H_\nu^J\right\rangle_{L^2_c(\mathcal{H};\nu)} \\
&= \sum_{I,J\in\mathcal{D}} \widehat{f}_\mu\left(I\right)\widehat{g}_\nu\left(J\right)\sqrt{\mu_I}\left\langle \widetilde{\mathbf{1}}_{Q_\pm(I)}, H_\nu^J\right\rangle_{L^2_c(\mathcal{H};\nu)}
\end{aligned}$$

since $\mathsf{U}_\mu h_I^\mu = \sqrt{\mu_I}\,\widetilde{\mathbf{1}}_{Q_\pm(I)}$. By a further, straightforward, computation we have

$$\begin{aligned}
\left\langle \widetilde{\mathbf{1}}_{Q_\pm(I)}, H_\nu^J\right\rangle_{L^2_c(\mathcal{H};\nu)} &= \frac{1}{|I|}\int \left(-\mathbf{1}_{Q(I_+)} + \mathbf{1}_{Q(I_-)}\right)\widetilde{\nu}\left(J\right)\left(-\frac{\mathbf{1}_{Q(J_+)}}{\nu\left(J_+\right)} + \frac{\mathbf{1}_{Q(J_-)}}{\nu\left(J_-\right)}\right)\nu dA \\
&= \begin{cases} 0 & \text{if} \quad J\subset I, \\ \frac{\pm 1}{|I|}\frac{\widetilde{\nu}(J)}{\nu(J_\pm)}\left(-\nu\left(I_+\right)+\nu\left(I_-\right)\right) & \text{if} \quad I\subset J_\pm, \\ \frac{1}{|I|}\widetilde{\nu}\left(I\right) & \text{if} \quad I = J. \end{cases} \qquad (3.15)
\end{aligned}$$

Altogether we have

$$\begin{aligned}
\langle \mathsf{U}_\mu f, g\rangle_{L^2_c(\mathcal{H};\nu)} &= \sum_{I,J\in\mathcal{D}} \widehat{f}_\mu\left(I\right)\widehat{g}_\nu\left(J\right)\sqrt{\mu_I}\left\langle \widetilde{\mathbf{1}}_{Q_\pm(I)}, H_\nu^J\right\rangle_{L^2_c(\mathcal{H};\nu)} \\
&= \left(\sum_{I=J} + \sum_{J\subset I} + \sum_{I\subset J}\right)\widehat{f}_\mu\left(I\right)\widehat{g}_\nu\left(J\right)\sqrt{\mu_I}\left\langle \widetilde{\mathbf{1}}_{Q_\pm(I)}, H_\nu^J\right\rangle_{L^2_c(\mathcal{H};\nu)} \\
&\equiv \mathbf{A}+\mathbf{B}+\mathbf{C}.
\end{aligned}$$

We then need to show that

$$\left|\langle \mathsf{U}_\mu f, g\rangle_{L^2_c(\mathcal{H};\nu)}\right| \lesssim \left(C_1+C_2\right)\|f\|_{L^2_c(\mathcal{H};\mu)}\|g\|_{L^2_c(\mathcal{H};\nu)}$$

and to accomplish this we will show the desired estimates on each of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$.

Now for the first term, by the third line in (3.15) we have that

$$\begin{aligned} |\mathbf{A}| &= \left|\sum_{I\in\mathcal{D}} \widehat{f}_\mu(I)\,\widehat{g}_\nu(I)\,\frac{\sqrt{\mu_I}}{|I|}\widetilde{\nu}(I)\right| \\ &\leq \|f\|_{L^2_c(\mathcal{H};\mu)}\,\|g\|_{L^2_c(\mathcal{H};\nu)}\left(\sup_{I\in\mathcal{D}}\frac{\sqrt{\mu_I}}{|I|}\widetilde{\nu}(I)\right), \end{aligned}$$

with the last line following by Cauchy-Schwarz and Parseval's Identity. However, the forward testing condition gives

$$\begin{aligned} \frac{C_1^2}{2}\mu_I\,|I|^2 &= C_1^2\left\|\mathbf{1}_{T(I)}\right\|^2_{L^2(\mathcal{H};\mu)} \\ &\geq \left\|\mathsf{U}_\mu\left(\mathbf{1}_{T(I)}\right)\right\|^2_{L^2(\mathcal{H};\nu)} = 8\left\|\mu_I\mathbf{1}_{Q_\pm(I)}\right\|^2_{L^2_c(\mathcal{H};\nu)} \\ &= 8\mu_I^2\left(\nu(I_+)+\nu(I_-)\right), \end{aligned}$$

Then, using

$$\widetilde{\nu}(I)^2 \equiv \frac{\nu(I_+)\,\nu(I_-)}{\nu(I_+)+\nu(I_-)} \leq \min\left\{\nu(I_+),\nu(I_-)\right\} \leq \nu(I_+)+\nu(I_-),$$

we get

$$\sup_{I\in\mathcal{D}}\frac{\sqrt{\mu_I}}{|I|}\widetilde{\nu}(I) \lesssim C_1,$$

and thus, have

$$|\mathbf{A}| \lesssim C_1\,\|f\|_{L^2_c(\mathcal{H};\mu)}\,\|g\|_{L^2_c(\mathcal{H};\nu)}.$$

The second term is trivial since $\mathbf{B}=0$ by the first line in (3.15). Finally, by the second line in (3.15) we have

$$\begin{aligned} |\mathbf{C}| &= \left|\sum_{I\in\mathcal{D}}\sum_{J\supset I}\widehat{f}_\mu(I)\,\widehat{g}_\nu(J)\,\frac{\sqrt{\mu_I}}{|I|}\left(\frac{\pm\widetilde{\nu}(J)}{\nu(J_\pm)}\right)\left(-\nu(I_+)+\nu(I_-)\right)\right| \\ &= \left|\sum_{I\in\mathcal{D}}\widehat{f}_\mu(I)\,\frac{\sqrt{\mu_I}}{|I|}\left(-\nu(I_+)+\nu(I_-)\right)\left(\sum_{J\supset I}\widehat{g}_\nu(J)\,\frac{\pm\widetilde{\nu}(J)}{\nu(J_\pm)}\right)\right| \\ &= \left|\sum_{I\in\mathcal{D}}\widehat{f}_\mu(I)\,\frac{\sqrt{\mu_I}}{|I|}\left(-\nu(I_+)+\nu(I_-)\right)\left\langle g,\frac{\mathbf{1}_{Q(I)}}{\nu(I)}\right\rangle_{L^2(\mathcal{H};\nu)}\right| \\ &\leq \left(\sum_{I\in\mathcal{D}}\left|\widehat{f}_\mu(I)\right|^2\right)^{\frac12}\left(\sum_{I\in\mathcal{D}}\left|\left\langle g,\frac{\mathbf{1}_{Q(I)}}{\nu(I)}\right\rangle_{L^2(\mathcal{H};\nu)}\right|^2\frac{\mu_I}{|I|^2}\left(-\nu(I_+)+\nu(I_-)\right)^2\right)^{\frac12} \\ &= \|f\|_{L^2_c(\mathcal{H})}\left(\sum_{I\in\mathcal{D}}\left|\left\langle g,\frac{\mathbf{1}_{Q(I)}}{\nu(I)}\right\rangle_{L^2(\mathcal{H};\nu)}\right|^2\frac{\mu_I}{|I|^2}\left(-\nu(I_+)+\nu(I_-)\right)^2\right)^{\frac12}. \end{aligned}$$

Expanding $\mathsf{U}_\nu^*\left(\mathbf{1}_{Q(I)}\right)$ with respect to the basis $\left\{\widetilde{\mathbf{1}}_{T(J)}\right\}_{J\in\mathcal{D}}$ we note that the backward testing condition gives

$$\begin{aligned} C_2^2\nu\left(I\right) &= C_2^2\left\|\mathbf{1}_{Q(I)}\right\|_{L_c^2(\mathcal{H};\nu)}^2 \\ &\geq \left\|\mathbf{1}_{Q(I)}\mathsf{U}_\nu^*\left(\mathbf{1}_{Q(I)}\right)\right\|_{L_c^2(\mathcal{H};\mu)}^2 \\ &= \sum_{J\subset I}\mu_J\left|\left\langle\nu\mathbf{1}_{Q(I)},\widetilde{\mathbf{1}}_{Q_\pm(J)}\right\rangle_{L^2(\mathcal{H})}\right|^2 \\ &= \sum_{J\subset I}\frac{\mu_J}{\left|J\right|^2}\left(-\nu\left(J_+\right)+\nu\left(J_-\right)\right)^2, \end{aligned}$$

and then the Carleson Embedding Theorem shows that

$$\sum_{I\in\mathcal{D}}\left|\left\langle g,\frac{\mathbf{1}_{Q(I)}}{\nu\left(I\right)}\right\rangle_{L_c^2(\mathcal{H};\nu)}\right|^2\frac{\mu_I}{\left|I\right|^4}\left(-\nu\left(I_+\right)+\nu\left(I_-\right)\right)^2\lesssim C_2^2\left\|g\right\|_{L_c^2(\mathcal{H};\nu)}^2.$$

Therefore, we have

$$\left|\mathbf{C}\right|\lesssim C_2\left\|f\right\|_{L_c^2(\mathcal{H})}\left\|g\right\|_{L_c^2(\mathcal{H})}$$

Combining the above we get

$$\left|\left\langle\mathsf{U}_\mu f,g\right\rangle_{L_c^2(\mathcal{H};\nu)}\right|\leq\left|\mathbf{A}\right|+\left|\mathbf{C}\right|\lesssim\left(C_1+C_2\right)\left\|f\right\|_{L^2(\mathcal{H};\mu)}\left\|g\right\|_{L^2(\mathcal{H};\nu)}.$$

∎

Remark 3.1. *The paper [3] more generally studies operators that are "well localized" with respect to the Haar basis. It is clear that the method of proof in [3] can be extended to operators that are sufficiently localized with respect to a pair of bases. We do not explore this extension at this time, but will return to it at some point in the future.*

4. Conclusion

Unfortunately the methods we have used in this paper do not appear to work to handle type $(0,1,0,1)$ compositions. However, we strongly believe that the following conjecture is true:

Conjecture 4.1. $\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,1)}$ *is bounded on* $L^2\left(\mathbb{R}\right)$ *if and only if for each* $I\in\mathcal{D}$ *there exists* $L^2\left(\mathbb{R}\right)$ *functions* F_I *and* B_I *of norm* 1 *such that*

$$\begin{aligned} \left\|\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,1)}F_I\right\|_{L^2(\mathbb{R})} &\leq C_1 \\ \left\|\mathsf{P}_d^{(1,0)}\circ\mathsf{P}_b^{(1,0)}B_I\right\|_{L^2(\mathbb{R})} &\leq C_2. \end{aligned}$$

Moreover, we will have

$$\left\|\mathsf{P}_b^{(0,1)}\circ\mathsf{P}_d^{(0,1)}\right\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})}\approx C_1+C_2.$$

The choice of the families $\{F_I\}_{I\in\mathcal{D}}$ and $\{B_I\}_{I\in\mathcal{D}}$ will clearly play an important role.

References

[1] Nicola Arcozzi, Richard Rochberg, and Eric Sawyer, *Carleson measures for analytic Besov spaces*, Rev. Mat. Iberoamericana **18** (2002), no. 2, 443–510. ↑2, 14, 15

[2] F. Nazarov, *A counterexample to Sarason's conjecture* (1997), 1–17 pp., preprint, available at http://www.math.msu.edu/~fedja/Preprints/Sarason.ps. ↑1

[3] F. Nazarov, S. Treil, and A. Volberg, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** (2008), no. 3, 583–597. ↑2, 22, 25

[4] S. Petermichl, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic*, Amer. J. Math. **129** (2007), no. 5, 1355–1375. ↑5

[5] Stefanie Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris Sér. I Math. **330** (2000), no. 6, 455–460 (English, with English and French summaries). ↑5

[6] Sandra Pott, Eric T. Sawyer, Maria C. Reguera, and Brett D. Wick, *The Linear Bound for the Hilbert Transform*, preprint. ↑5

[7] Sandra Pott and Martin P. Smith, *Paraproducts and Hankel operators of Schatten class via p-John-Nirenberg theorem*, J. Funct. Anal. **217** (2004), no. 1, 38–78. ↑7

[8] Donald Sarason, *Products of Toeplitz operators*, Linear and complex analysis. Problem book 3. Part I (V. P. Havin and N. K. Nikolski, eds.), Lecture Notes in Mathematics, vol. 1573, Springer-Verlag, Berlin, 1994, pp. 318-319. ↑1

[9] Eric T. Sawyer, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. **308** (1988), no. 2, 533–545. ↑

[10] Sergei Treil, Alexander Volberg, and Dechao Zheng, *Hilbert transform, Toeplitz operators and Hankel operators, and invariant A_∞ weights*, Rev. Mat. Iberoamericana **13** (1997), no. 2, 319–360. ↑1

S. Pott, Centre for Mathematical Sciences, University of Lund, Lund, Sweden
E-mail address: sandra@maths.lth.se

M. C. Reguera, Department of Mathematics, Universitat Autònoma de Barcelona, Barcelona, Spain
E-mail address: mreguera@mat.uab.cat

E. T. Sawyer, Department of Mathematics, McMaster University, Hamilton, Canada
E-mail address: sawyer@mcmaster.ca

Brett D. Wick, School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA USA 30332-0160
E-mail address: wick@math.gatech.edu
URL: www.math.gatech.edu/~wick