The Journal of Geometric Analysis
Volume 16, Number 4, 2006

Distributional Estimates for the Bilinear Hilbert Transform

By *Dmitriy Bilyk and Loukas Grafakos*

ABSTRACT. *We obtain size estimates for the distribution function of the bilinear Hilbert transform acting on a pair of characteristic functions of sets of finite measure, that yield exponential decay at infinity and blowup near zero to the power $-2/3$ (modulo some logarithmic factors). These results yield all known L^p bounds for the bilinear Hilbert transform and provide new restricted weak type endpoint estimates on $L^{p_1} \times L^{p_2}$ when either $1/p_1 + 1/p_2 = 3/2$ or one of p_1, p_2 is equal to 1. As a consequence of this work we also obtain that the square root of the bilinear Hilbert transform of two characteristic functions is exponentially integrable over any compact set.*

1. Introduction

Certain bilinear singular integral operators can be expressed as averages of bilinear Hilbert transforms in a way analogous to that which linear singular integrals can be written as averages of linear directional Hilbert transforms. The bilinear Hilbert transforms were introduced in the early 1960s to play exactly this role by A. Calderón in his study of the first commutator. Properties of these operators remained elusive until the appearance of the fundamental work of Lacey and Thiele [11, 12] in the late 1990s who established their boundedness on certain products of Lebesgue spaces. This work was based on a remarkable set of techniques called time-frequency analysis and revealed a fundamental and deep connection with almost everywhere convergence of Fourier series and in particular, the boundedness of the Carleson-Hunt operator, see Lacey and Thiele [13]; on the latter the work of Fefferman [3] was influential. The Carleson-Hunt operator is defined as

$$\mathcal{C}(f)(x) = \sup_{N>0} \left| \int_{-N}^{+N} \widehat{f}(\xi) e^{2\pi i x \xi} \, d\xi \right|$$

where $\widehat{f}(\xi) = \int_{\mathbf{R}} f(x) e^{-2\pi i x \xi} \, dx$ is the Fourier transform of the function f on the line. Carleson [2] answered a longstanding conjecture posed by Lusin by establishing the boundedness of the operator $\mathcal{C}$ on L^2. A few years later, Hunt [8] obtained its L^p boundedness for $1 < p < \infty$

Math Subject Classifications. 46B70, 42B99.
Key Words and Phrases. Multilinear operators, distributional estimates.
Acknowledgements and Notes. Both authors have been supported by the National Science Foundation under grant DMS 0400387 and by the University of Missouri Research Council.

ISSN 1050-6926

as a consequence of the following powerful distributional estimate

$$\left|\left\{\mathcal{C}(\chi_F) > \lambda\right\}\right| \leq C\,|F| \begin{cases} \frac{1}{\lambda}\left(1+\log\frac{1}{\lambda}\right) & \text{when } \lambda \leq 1 \\ e^{-c\lambda} & \text{when } \lambda > 1\,. \end{cases} \tag{1.1}$$

This estimate holds for some fixed constants c, C and for all measurable sets F of finite measure and all $\lambda > 0$. Using standard interpolation, estimate (1.1) easily implies the L^p boundedness of $\mathcal{C}$ for $1 < p < \infty$. Moreover, recent extrapolation techniques by Antonov [1] and their refinement by Sjölin and Soria [15] show that estimate (1.1) implies the boundedness of $\mathcal{C}$ on $L \log L \log\log\log L$ of every compact set; this implies the almost everywhere convergence of the partial Fourier integrals of functions locally in this class.

The main purpose of this article is to prove an estimate analogous to that in (1.1) for the bilinear Hilbert transform H_α. This operator is defined for a parameter $\alpha \in \mathbf{R}$ by

$$H_\alpha(f,g)(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{|t|\geq\varepsilon} f(x-t)g(x+\alpha t)\frac{dt}{t}, \qquad x \in \mathbf{R}$$

for functions f, g on the line. In the aforementioned work [11, 12], Lacey and Thiele proved that the operator H_α maps $L^{p_1} \times L^{p_2}$ to L^p, whenever $1 < p_1, p_2 \leq \infty$, $\frac{1}{p_1}+\frac{1}{p_2}=\frac{1}{p}$, and $\frac{2}{3} < p < \infty$, $\alpha \neq -1$. At this time it remains unknown whether the bilinear Hilbert transform is bounded for values of $p \leq \frac{2}{3}$, even near $L^1 \times L^1 \to L^{1/2}$. It is quite clear that the time-frequency techniques will not resolve this issue without new ideas (see a counterexample in [10]).

Our approach uses the model sum reduction of Lacey and Thiele [11, 12], a tree analysis based on a selection inspired by Lacey [9], and relies on an "improved energy estimate" that appeared in the proof of (1.1) by Grafakos, Tao, and Terwilleger [6]. A variant of this energy estimate had previously appeared in the related work of Muscalu, Thiele, and Tao [14].

The main result of the article is the following.

Theorem 1.1. *Let $2 \leq p_2 < \infty$ and $\alpha \in \mathbf{R}\setminus\{0,-1\}$. Then there exist constants $C = C(\alpha, p_2)$, $c = c(\alpha, p_2)$ such that for all measurable sets F_1, F_2 of finite measure we have*

$$|\{|H_\alpha(\chi_{F_1},\chi_{F_2})| > \lambda\}| \leq C\left(|F_1|\,|F_2|^{\frac{1}{p_2}}\right)^{\frac{p_2}{p_2+1}} \begin{cases} \lambda^{-\frac{p_2}{p_2+1}}\left(1+\log\frac{1}{\lambda}\right)^{\frac{2p_2}{p_2+1}} & \textit{when } \lambda < 1\,, \\ e^{-c\sqrt{\lambda}} & \textit{when } \lambda \geq 1. \end{cases} \tag{1.2}$$

Analogously, the following estimate is valid for $2 \leq p_1 < \infty$:

$$|\{|H_\alpha(\chi_{F_1},\chi_{F_2})| > \lambda\}| \leq C\left(|F_1|^{\frac{1}{p_1}}\,|F_2|\right)^{\frac{p_1}{p_1+1}} \begin{cases} \lambda^{-\frac{p_1}{p_1+1}}\left(1+\log\frac{1}{\lambda}\right)^{\frac{2p_1}{p_1+1}} & \textit{when } \lambda < 1\,, \\ e^{-c\sqrt{\lambda}} & \textit{when } \lambda \geq 1\,. \end{cases} \tag{1.3}$$

These estimates correspond to the line segments $\{(\frac{1}{p_1},\frac{1}{p_2}) : \ p_1 = 1,\ 2 \leq p_2 < \infty\}$ and $\{(\frac{1}{p_1},\frac{1}{p_2}) : \ 2 \leq p_1 < \infty,\ p_2 = 1\}$. As a corollary we obtain the following distributional estimate corresponding to the line segment $\{(\frac{1}{p_1},\frac{1}{p_2}) :\ 1 \leq p_1, p_2 \leq 2,\ \frac{1}{p_1}+\frac{1}{p_2}=\frac{3}{2}\}$.

Corollary 1.2. *For any $\alpha \in \mathbf{R}\setminus\{0,-1\}$ there exist constants $C = C(\alpha)$, $c = c(\alpha)$ such that for all measurable sets F_1, F_2 of finite measure we have*

$$|\{|H_\alpha(\chi_{F_1},\chi_{F_2})| > \lambda\}| \leq C\left(|F_1|^{\frac{1}{2}}|F_2|^{\frac{1}{2}}\min(|F_1|,|F_2|)^{\frac{1}{2}}\right)^{\frac{2}{3}} \begin{cases} \lambda^{-\frac{2}{3}}\left(1+\log\frac{1}{\lambda}\right)^{\frac{4}{3}} & \textit{for } \lambda < 1 \\ e^{-c\sqrt{\lambda}} & \textit{for } \lambda \geq 1. \end{cases} \tag{1.4}$$

Remark. In the distributional estimate (1.4), the expression $|F_1|^{\frac{1}{2}}|F_2|^{\frac{1}{2}}\min(|F_1|^{\frac{1}{2}},|F_1|^{\frac{1}{2}})$ is dominated by $|F_1|^{\frac{1}{p_1}}|F_2|^{\frac{1}{p_2}}$, where $1 \le p_j \le 2$ and $\frac{1}{p_1}+\frac{1}{p_2}=\frac{3}{2}$. Thus, this estimate (up to a logarithmic term) is similar to a restricted weak type estimate for such exponents.

Notice that the exponential decay at infinity for the distribution function of H_α is not as strong as in the case of the Carleson-Hunt operator and at the moment we do not know if it is sharp. Estimates (1.2), (1.3), and (1.4) not only capture the boundedness of H_α on products of Lebesgue spaces but also yield other crucial quantitative information such as local exponential integrability of H_α and also its boundedness on other rearrangement invariant spaces even at the endpoint cases.

We state the exponential integrability of H_α in the form of corollary.

Corollary 1.3. *Let $\alpha \in \mathbf{R}\setminus\{0,-1\}$ and $c = c(\alpha)$ be as in Corollary 1.2. Then there is a constant $C' = C'(\alpha)$ such that for any bounded measurable set K and for all measurable sets F_1, F_2 of finite measure the following holds:*

$$\int_K e^{c'|H_\alpha(\chi_{F_1},\chi_{F_2})(x)|^{\frac{1}{2}}}\,dx \le C'\left(|K| + \left(|F_1|^{\frac{1}{2}}|F_2|^{\frac{1}{2}}\min(|F_1|,|F_2|)^{\frac{1}{2}}\right)^{\frac{2}{3}}\right)$$

for any $0 < c' < c$.

2. Decomposition of the bilinear Hilbert transforms

In the sequel we will drop the dependence of H_α on α and simply denote it by H. We will use the notation $|A|$ for the Lebesgue measure of a set A and $\langle f, g\rangle$ for the complex inner product $\int f(x)\overline{g(x)}\,dx$. For a number $a > 0$ and an interval I we denote aI an interval of length $a|I|$ concentric with I and by $a \otimes I$ the interval $[ap, aq]$ if $I = [p, q]$. We will use the notation $\lesssim$ to express that a certain quantity is at most a constant multiple of another one.

Our goal will be to study the trilinear form

$$(f_1, f_2, f_3) \to \int H(f_1, f_2)(x) f_3(x)\,dx$$

for three functions f_1, f_2, f_3 which will be characteristic functions of sets of finite measure, i.e., $f_1 = \chi_{F_1}$, $f_2 = \chi_{F_2}$, and $f_3 = \chi_{E'}$.

We fix L to be the smallest integer greater than $2^{10}\max\{|\alpha|, \frac{1}{|\alpha|}, \frac{1}{|1+\alpha|}\}^3$. The dependence of the bounds on α will enter the proof through polynomial dependence on L.

We begin by noting that the distribution p.v.$\frac{1}{t}$ that appears in the definition of H can be written as $c_1\delta_0 + c_2\gamma$ for some constants c_1, c_2, where δ_0 is the Dirac mass at the origin and γ is another distribution that satisfies $\widehat{\gamma} = \chi_{(0,\infty)}$. Since all the estimates that we are going to be proving in this article are trivial for δ_0, we may restrict our attention to γ. Let θ be a smooth function which is equal to 1 on $(-\infty, 2L)$ and 0 on $(3L, \infty)$. Define

$$\widehat{\psi}(\xi) = \theta(\xi) - \theta(2\xi)\ .$$

Observe that $\widehat{\psi}$ is nonzero and is supported in $[L, 3L]$. For each integer k we define

$$\psi_k(x) = 2^{-\frac{k}{2}}\psi\left(2^{-k}x\right)\ .$$

Then we have

$$\gamma = \sum_{k\in\mathbf{Z}} 2^{-\frac{k}{2}}\psi_k\ .$$

Indeed, if we look at the Fourier transform of the right-hand side of the identity above, we get a telescopic sum:

$$\Big(\sum_{k\in\mathbf{Z}} 2^{-\frac{k}{2}}\psi_k\Big)\widehat{\ }(\xi) = \lim_{N\to\infty}\sum_{-N}^{N}\big[\theta\big(2^k\xi\big)-\theta\big(2^{k+1}\xi\big)\big] = \lim_{N\to\infty}\big[\theta\big(2^{-N}\xi\big)-\theta\big(2^{N+1}\xi\big)\big] = \widehat{\gamma}\ .$$

It clearly suffices to study the trilinear form

$$\Lambda(f_1, f_2, f_3) := \sum_{k\in\mathbf{Z}} 2^{-\frac{k}{2}} \int\int f_1(x-t) f_2(x+\alpha t) f_3(x)\psi_k(t)\, dt\, dx\ . \tag{2.1}$$

We can further break the function ψ into a sum of at most $2L$ functions $\psi^{(M)}$ such that $\widehat{\psi^{(M)}}$ is supported in the interval $[M-\frac{1}{2}, M+\frac{1}{2}]$ for $L \le M \le 2L$. It would suffice to study each piece separately. For notational convenience, we will omit the dependence on M and will just write ψ.

For further decomposition we fix a Schwartz function ϕ of L^2 norm 1, with Fourier transform supported in $[-\frac{1}{2}, \frac{1}{2}]$, which also has the property that for all $\xi \in \mathbf{R}$ we have

$$\sum_{l\in\mathbf{Z}} \big|\widehat{\phi}(\xi - l/2)\big|^2 \equiv C_0$$

for some constant $C_0 > 0$.

Let $u = I_u \times \omega_u$ be a rectangle in $\mathbf{R}^2$ and set

$$\phi_u(x) = |I_u|^{-\frac{1}{2}}\phi\Big(\frac{x-c(I_u)}{|I_u|}\Big)e^{2\pi i c(\omega_u)x}\ ,$$

where $c(J)$ denotes the center of the interval J.

For each $k \in \mathbf{Z}$ we consider the set of dyadic rectangles of scale k:

$$\mathbf{S}_k = \big\{\big(2^k n, 2^k(n+1)\big)\times\big(2^{-k}m/2, 2^{-k}(m/2+1)\big)\,|\, m, n \in \mathbf{Z}\big\}\ .$$

Then $\mathbf{S} = \bigcup_k \mathbf{S}_k$ is the set of all dyadic rectangles of area 1 in $\mathbf{R}^2$.

It is an easy calculation to verify that for all $f \in L^2$

$$f = \frac{1}{C_0}\sum_{u\in\mathbf{S}_k}\langle f, \phi_u\rangle\phi_u$$

where the convergence is in L^2. Moreover, the series also converges a.e. for all $f \in L^p$, $1 < p < \infty$, see [5]. Using this decomposition of the identity in the k^{th} term of (2.1), as in [12], we obtain

$$\Lambda(f_1, f_2, f_3) := \sum_{k\in\mathbf{Z}}\ \sum_{u_1,u_2,u_3\in\mathbf{S}_k} C_{k,u_1,u_2,u_3}\Lambda_{k,u_1,u_2,u_3}(f_1, f_2, f_3)\ , \tag{2.2}$$

where

$$C_{k,u_1,u_2,u_3} = C_0^{-3}\int_{\mathbf{R}}\int_{\mathbf{R}}\phi_{u_1}(x-t)\phi_{u_2}(x+\alpha t)\phi_{u_3}(x)\psi_k(t)\, dt\, dx$$

and

$$\Lambda_{k,u_1,u_2,u_3}(f_1, f_2, f_3) = 2^{-\frac{k}{2}}\langle f_1, \phi_{u_1}\rangle\langle f_2, \phi_{u_2}\rangle\langle f_3, \phi_{u_3}\rangle\ .$$

We now take a closer look at the coefficients C_{k,u_1,u_2,u_3} in two different ways. First,

$$
\begin{aligned}
&|C_{k,u_1,u_2,u_3}| \\
&\le C_0^{-3}\int\!\!\int\left|\phi\Big(\frac{x-t-c(I_{u_1})}{|I_{u_1}|}\Big)\phi\Big(\frac{x+\alpha t-c(I_{u_2})}{|I_{u_2}|}\Big)\phi\Big(\frac{x-c(I_{u_3})}{|I_{u_3}|}\Big)\psi\big(2^{-k}t\big)\right|2^{-2k}\,dt\,dx \\
&= C_0^{-3}\int\!\!\int\Big|\phi\big(x-t-A_1\big)\phi\big(x+\alpha t-A_2\big)\phi\big(x-A_3\big)\psi(t)\Big|\,dt\,dx\ ,
\end{aligned}
$$

where $A_i=\frac{c(I_{u_i})}{|I_{u_i}|}$ for $i=1,2,3$ (these numbers are half-integers). Observe that

$$
\begin{aligned}
A_2-A_1 &= (x-t-A_1)-(x+\alpha t-A_2)+(1+\alpha)t\ ,\\
A_3-A_1 &= (x-t-A_1)-(x-A_3)+t\ ,\\
A_3-A_2 &= (x+\alpha t-A_2)-(x-A_3)-\alpha t\ .
\end{aligned}
$$

This implies that at least one of the arguments in the last displayed double integral has to have size at least $\frac{1}{4L}\mathrm{diam}\{A_i\}$. Since ϕ and ψ are Schwartz functions, it follows that, for any positive integer m, there exists a constant C_m such that

$$
|C_{k,u_1,u_2,u_3}|\le C_m\left(1+\frac{\mathrm{diam}\{A_i\}}{4L}\right)^{-m}=C_m\left(1+\frac{\max_{i,j}|c(I_{u_i})-c(I_{u_j})|}{2^k\,4L}\right)^{-m}. \qquad (2.3)
$$

Secondly, we set $F_1(x,t)=\phi_{u_1}(x-t)\phi_{u_2}(x+\alpha t)$, $F_2(x,t)=\phi_{u_3}(x)\psi_k(t)$. These are Schwartz functions of two variables. We have

$$
\begin{aligned}
\widehat{F_1}(\xi,\tau) &= \frac{1}{1+\alpha}\,\widehat{\phi_{u_1}}\left(\frac{\alpha\xi-\tau}{1+\alpha}\right)\widehat{\phi_{u_2}}\left(\frac{\xi+\tau}{1+\alpha}\right),\\
\widehat{F_2}(\xi,\tau) &= \frac{1}{1+\alpha}\,\widehat{\phi_{u_3}}(\xi)\widehat{\psi_k}(\tau)\ .
\end{aligned}
$$

Thus, applying the two-dimensional Plancherel formula, we obtain

$$
|C_{k,u_1,u_2,u_3}|\le\frac{C}{|1+\alpha|}\int\!\!\int\left|\widehat{\phi}\left(\frac{\alpha\xi-\tau}{1+\alpha}-B_1\right)\widehat{\phi}\left(\frac{\xi+\tau}{1+\alpha}-B_2\right)\widehat{\phi}(\xi-B_3)\widehat{\psi}(\tau)\right|d\xi\,d\tau\ , \qquad (2.4)
$$

where $B_i=\frac{c(\omega_{u_i})}{|\omega_{u_i}|}=2^kc(\omega_{u_i})$ (notice that this is an integer or a half-integer).

Assume that the integral above is not zero. Then we must have

$$
\frac{\alpha\xi-\tau}{1+\alpha}-B_1\in\left[-\frac12,\frac12\right],\ \frac{\xi+\tau}{1+\alpha}-B_2\in\left[-\frac12,\frac12\right],\ \xi-B_3\in\left[-\frac12,\frac12\right],\ \tau\in\left[M-\frac12,M+\frac12\right],
$$

which imply

$$
B_1\in\left[\frac{\alpha}{1+\alpha}B_3-\frac{1}{1+\alpha}M-\frac{1+|\alpha|+|1+\alpha|}{2|1+\alpha|},\ \frac{\alpha}{1+\alpha}B_3-\frac{1}{1+\alpha}M+\frac{1+|\alpha|+|1+\alpha|}{2|1+\alpha|}\right] \qquad (2.5)
$$

and

$$
B_2\in\left[\frac{1}{1+\alpha}B_3+\frac{1}{1+\alpha}M-\frac{2+|1+\alpha|}{2|1+\alpha|},\ \frac{1}{1+\alpha}B_3+\frac{1}{1+\alpha}M+\frac{2+|1+\alpha|}{2|1+\alpha|}\right]. \qquad (2.6)
$$

This means that the triple of parameters B_1, B_2, B_3 really depends only on the parameter B_3 as for each value of B_3, the quantities B_1 and B_2 can take only a finite number of values depending on α. Also, (2.5) and (2.6) show that $B_1 + B_2 = B_3$ up to an error that can only take a finite number of integer values (depending on α.)

We introduce parameters $\nu_1, \nu_2, \mu_1, \mu_2$ by setting

$$A_1 = A_3 + \nu_1, \qquad A_2 = A_3 + \nu_2, \qquad B_1 = \frac{\alpha}{\alpha+1} B_3 + \mu_1, \qquad B_2 = \frac{1}{\alpha+1} B_3 + \mu_2 .$$

We also set $\nu = \max |\nu_i|$. We aim to reduce the sum over $u_1, u_2, u_3 \in \mathbf{S}_k$ as the rapidly converging sum over $\nu_1, \nu_2, \mu_1, \mu_2$ of the sum over the tiles u_3.

For N sufficiently large we have

$$|\Lambda(f_1, f_2, f_3)| \\ \le \sum_{\nu=0}^{\infty} C_N \big(1 + \tfrac{\nu}{4L}\big)^{-N} \sum_{\substack{(\nu_1,\nu_2): \\ \max|\nu_i|=\nu}} \sum_{\mu_1} \sum_{\mu_2} \Bigg| \sum_{k\in\mathbf{Z}} \sum_{u_3 \in \mathbf{S}_k} \varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,u_3} \Lambda_{k,u_1,u_2,u_3}(f_1, f_2, f_3) \Bigg| \tag{2.7}$$

where $u_1 = u_1(u_3)$ and $u_2 = u_2(u_3)$ are uniquely determined by u_3 in terms of $\nu_1, \nu_2, \mu_1, \mu_2$, $\varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,u_3}$ is a constant of modulus at most 1, and $\mu_1 \in \frac{1}{2}\mathbf{Z} - \frac{\alpha}{1+\alpha}\frac{1}{2}\mathbf{Z}$ and $\mu_2 \in \frac{1}{2}\mathbf{Z} - \frac{1}{1+\alpha}\frac{1}{2}\mathbf{Z}$ range in the intervals

$$\mu_1 \in \left[-\frac{1}{1+\alpha} M - \frac{1+|\alpha|+|1+\alpha|}{2|1+\alpha|}, -\frac{1}{1+\alpha} M + \frac{1+|\alpha|+|1+\alpha|}{2|1+\alpha|} \right],$$
$$\mu_2 \in \left[\frac{1}{1+\alpha} M - \frac{2+|1+\alpha|}{2|1+\alpha|}, \frac{1}{1+\alpha} M - \frac{2+|1+\alpha|}{2|1+\alpha|} \right].$$

Thus, μ_1 and μ_2 take only a finite number of values depending on α. (Note that $\varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,s_3}$ is the ratio of C_{k,u_1,u_2,u_3} by $C_N(1 + \frac{\max|\nu_i|}{4L})^{-N}$.)

It will clearly suffice to study the boundedness of the expression inside the absolute values in (2.7) and to obtain bounds independent of μ_i and polynomial in ν, since for each ν, there are of the order of ν pairs (ν_1, ν_2) with $\max|\nu_i| = \nu$.

Next, we further separate the triples in such a way that for two triples (u_1, u_2, u_3) and (u_1', u_2', u_3') from the same group the following conditions hold:

$$\text{if} \quad k \neq k', \qquad \text{then} \quad \big|k - k'\big| > L^{10} , \tag{2.8}$$
$$\text{if} \quad A_3 \neq A_3', \quad \text{then} \quad \big|A_3 - A_3'\big| > \nu L^{10} , \tag{2.9}$$
$$\text{if} \quad B_3 \neq B_3', \quad \text{then} \quad \big|B_3 - B_3'\big| > L^{10} . \tag{2.10}$$

Obviously, the number of such groups is polynomial in L and ν.

To facilitate the study of the sums above, we introduce *tri-tiles.* A tri-tile is a rectangle $s = I_s \times \omega_s$ and three subrectangles s_1, s_2, s_3 built in the following way.

Let (u_1, u_2, u_3) be a triple of rectangles participating in the sum in (2.7). Define $I_s = I_{s_i} = I_{u_3}$. Defining the frequency projections requires a little bit more work, as we cannot just use the dyadic grid. We want these projections to satisfy the following properties:

$$\mathcal{J} = \bigcup_{s\in S} \big(\omega_s \cup \omega_{s_1} \cup \omega_{s_2} \cup \omega_{s_3}\big) \text{ is a grid} . \tag{2.11}$$

$$\text{If} \quad \omega_{s_i} \subsetneqq J \text{ for some} \quad J \in \mathcal{J}, \text{then} \quad \omega_{s_j} \subsetneqq J \text{ for some } J \in \mathcal{J} \text{ for all} \quad j = 1, 2, 3 . \tag{2.12}$$

$$\omega_{s_i} \neq \omega_{s_j} \text{ for} \quad i \neq j . \tag{2.13}$$

We build these intervals by induction on the cardinality of the set U of triples of rectangles. If this set is nonempty, we pick the triple u_1, u_2, u_3, such that k, where $|\omega_{u_3}| = 2^k$, is maximal. Let $U' = U \setminus (u_1, u_2, u_3)$. By induction we find the intervals ω_s, ω_{s_i} ($i = 1, 2, 3$), corresponding to the elements of U'. If there is an element $u' \in U'$ such that $\omega_{u'_3} = \omega_{u_3}$, then we define ω_s, ω_{s_i} ($i = 1, 2, 3$) to be the same as the corresponding intervals for u'. Otherwise we define (for $i = 1, 2$, or 3) ω_{s_i} to be the convex hull of the interval $C_i \omega_{u_i}$ ($C_1 = \frac{1+\alpha}{\alpha}$, $C_2 = 1 + \alpha$, $C_3 = 1$) and all sets $\omega_{s'}$ that intersect it. (Note that because of the separation of scales what we get is only slightly smaller than the interval itself.) Next, we define ω_s as follows: Take $[a, b]$ to be the convex hull of ω_{s_i} ($i = 1, 2, 3$), then set ω_s to be the convex hull of $[a, b]$ and all intervals $\omega_{s'}$ that intersect $[a, b]$. Properties (2.11) and (2.12) are obvious in view of (2.10) and (2.8). Also $|\omega_s|$ and $|\omega_{s_i}|$ are comparable to 2^k with a factor depending on L. Property (2.13) follows from (2.5), (2.6), and the separation of scales.

We define the functions adapted to the tri-tile s with parameters $\nu_1, \nu_2, \mu_1, \mu_2$ as follows:

$$
\begin{aligned}
\varphi_{s_1}^{\nu_1,\mu_1,\alpha}(x) &= |I_s|^{-\frac{1}{2}} \phi\left(\frac{x - c(I_s)}{|I_s|} - \nu_1\right) e^{2\pi i \left(\frac{\alpha}{\alpha+1} c(\omega_{s_1}) + \theta_{s_1} |\omega_{s_1}|\right) x} = \phi_{u_1}(x) , \\
\varphi_{s_2}^{\nu_2,\mu_2,\alpha}(x) &= |I_s|^{-\frac{1}{2}} \phi\left(\frac{x - c(I_s)}{|I_s|} - \nu_2\right) e^{2\pi i \left(\frac{1}{\alpha+1} c(\omega_{s_2}) + \theta_{s_2} |\omega_{s_2}|\right) x} = \phi_{u_2}(x) , \\
\varphi_{s_3}^{\alpha}(x) &= |I_s|^{-\frac{1}{2}} \phi\left(\frac{x - c(I_s)}{|I_s|}\right) e^{2\pi i \left(c(\omega_{s_3}) + \theta_{s_3} |\omega_{s_3}|\right) x} = \phi_{u_3}(x) ,
\end{aligned}
$$

where the error terms θ_{s_i} in the modulations are chosen so that $\frac{\alpha}{\alpha+1} c(\omega_{s_1}) + \theta_{s_1} |\omega_{s_1}| = c(\omega_{u_1})$, $\frac{1}{\alpha+1} c(\omega_{s_2}) + \theta_{s_2} |\omega_{s_2}| = c(\omega_{u_2})$, and $c(\omega_{s_2}) + \theta_{s_3} |\omega_{s_3}| = c(\omega_{u_3})$. Obviously, $|\theta_{s_i}| \leq CL$.

Then the expression inside the absolute values in (2.7) becomes exactly

$$
\sum_{s_3 \in \bigcup_{k \in \mathbf{Z}} \mathbf{S}_k} |I_s|^{-\frac{1}{2}} \varepsilon_{\nu_1,\nu_2,\mu_1,\mu_2,s} \left\langle f_1, \varphi_{s_1}^{\nu_1,\mu_1,\alpha} \right\rangle \left\langle f_2, \varphi_{s_2}^{\nu_2,\mu_2,\alpha} \right\rangle \left\langle f_3, \varphi_{s_3}^{\alpha} \right\rangle .
$$

This expression needs to be controlled with bounds that grow polynomially in the parameters ν_1, ν_2, and are independent of μ_1, μ_2. We will work with sums over finite sets of tri-tiles and get bounds independent of the choice of the finite set, which is clearly sufficient by a limiting argument.

Note that if ω_{u_i} and $\omega_{u'_i}$ were not disjoint, then neither are ω_{s_i} and $\omega_{s'_i}$. Thus,

$$
\text{if} \quad \omega_{s_i} \cap \omega_{s'_i} = \emptyset, \text{ then } \quad \langle \varphi_{s_i}, \varphi_{s'_i} \rangle = 0 .
$$

For notational convenience, in the sequel we will suppress the dependence of the functions φ_{s_j} on the parameters $\nu_1, \nu_2, \mu_1, \mu_2$. Notice that

$$
|\varphi_{s_k}(x)| \leq C \left(1 + \left|\frac{x - c(I_s)}{|I_s|} - \nu_k\right|\right)^{-10} \leq C \left(1 + \left|\frac{x - c(I_s)}{|I_s|}\right|\right)^{-10} (1 + \nu)^{10} .
$$

3. Estimates for the model sums. The case $I_s \subseteq \Omega$

Let S be a finite set of tri-tiles with fixed data ν_1, ν_2, μ_1, and μ_2. Then we define the "model sum" associated with S as follows:

$$
H_S(f_1, f_2)(x) = \sum_{s \in S} |I_s|^{-\frac{1}{2}} \varepsilon_s \langle f_1, \varphi_{s_1} \rangle \langle f_2, \varphi_{s_2} \rangle \varphi_{s_3}(x) .
$$

We set

$$\Omega = \left\{ x : M(\chi_{F_1})(x) > 8 \min\Big(1, \frac{|F_1|}{|E|}\Big) \right\} \bigcup \left\{ x : M(\chi_{F_2})(x) > 8 \min\Big(1, \frac{|F_2|}{|E|}\Big) \right\},$$

where M is the Hardy-Littlewood maximal function. Since M if of weak type $(1, 1)$ with constant at most 2, it is easy to see that $|\Omega| < \frac{1}{2}|E|$. We now set $E' = E \setminus \Omega$. Obviously, then $|E'| \geq \frac{1}{2}|E|$.

The main purpose of this article is to obtain a good estimate for the expression

$$\int_{E'} H_S(\chi_{F_1}, \chi_{F_2})(x)\, dx = \langle H_S(\chi_{F_1}, \chi_{F_2}), \chi_{E'} \rangle .$$

To do so we will break the model sum into two parts: The sum over those $s \in S$ for which $I_s \subseteq \Omega$ (easier case) and the sum over tiles with $I_s \nsubseteq \Omega$.

We begin with the easier case. For a dyadic interval J we set

$$\omega(x) = \left(1 + \Big(\frac{|x - c(J)|}{|J|}\Big)^2\right)^5$$

and

$$S_J = \{s \in S : I_s = J\} .$$

We have the following inequalities (for $i = 1, 2, 3$):

$$\|(\langle f, \varphi_{s_i}\rangle)\|_{\ell^\infty(S_J)} \lesssim (1+\nu)^{10}|J|^{-\frac{1}{2}}\|f\|_{L^1(\omega^{-1})} \tag{3.1}$$

$$\Big\| \sum_{s\in S_J} \alpha_s \varphi_{s_i} \Big\|_{L^2(\omega)} \lesssim (1+\nu)^{10}\|(\alpha_s)\|_{\ell^2(S_J)} \tag{3.2}$$

$$\|(\langle f, \varphi_{s_i}\rangle)\|_{\ell^2(S_J)} \lesssim (1+\nu)^{10}\|f\|_{L^2(\omega^{-1})} . \tag{3.3}$$

Indeed, to prove (3.1), for any $s \in S_J$ we have

$$\begin{aligned} |\langle f, \varphi_{s_i}\rangle| &= \left| \int_{\mathbb{R}} f(x)|J|^{-\frac{1}{2}}\varphi\Big(\frac{x - c(J)}{|J|} - \nu_i\Big) e^{2\pi i x (C_i c(\omega_{s_i}) + \theta|\omega_{s_i}|)}\, dx \right| \\ &\leq C\,(1+\nu)^{10}|J|^{-\frac{1}{2}}\|f\|_{L^1(\omega^{-1})} . \end{aligned}$$

Next, we prove (3.2), which is an analog of Bessel's inequality. Although, the functions φ_{s_i} are no longer orthogonal in the weighted space $L^2(\omega)$, we will see that they are "almost" orthogonal in this space. It is straightforward to check that

$$\begin{aligned} |\langle \varphi_{s_i}, \varphi_{s_i'}\rangle_\omega| &= \Big| \big(|\varphi(y-\nu_i)|^2(1+y^2)^5\big)\widehat{\ }\Big(\big(K_i(c(\omega_{s_i}) - c(\omega_{s_i'})) + (\theta_{s_i} - \theta_{s_i'})|\omega_{s_i}|\big)|J|\Big)\Big| \\ &\leq C(1+\nu)^{10}\big(1 + K_i|c(\omega_{s_i}) - c(\omega_{s_i'})|\,|J|\big)^{-10} , \end{aligned}$$

since $|\varphi(y-\nu_i)|^2(1+y^2)^5$ (and its Fourier transform) is a Schwartz function (here $K_1 = \frac{\alpha}{\alpha+1}$,

$K_2 = \frac{1}{\alpha+1}$, $K_3 = 1$). Now we have

$$\begin{aligned}\Big\| \sum_{s\in S_J} \alpha_s \varphi_{s_i} \Big\|^2_{L^2(\omega)} &\le \sum_{s,s'\in S_J} |\alpha_s|\,|\alpha_{s'}|\,|\langle \varphi_{s_i}, \varphi_{s'_i}\rangle_\omega| \\ &\le C(1+\nu)^{10} \sum_{k,m} |\alpha_k|\,|\alpha_m|(1+|k-m|)^{-10} \\ &\le 2C(1+\nu)^{10} \sum_{k\in\mathbb{Z}} \Big(|\alpha_k|^2 \sum_{m\in\mathbb{Z}} (1+|k-m|)^{-10} \Big) \\ &\le C'(1+\nu)^{10} \|(\alpha_k)\|^2_{\ell^2} \,.\end{aligned}$$

Note that (3.3) is the dual statement of (3.2).

Let M be the Hardy-Littlewood maximal function and $M_2(f) = M(f^2)^{1/2}$. We prove the following estimate.

Lemma 3.1. *For $A > 1$ we have*

$$\|H_{S_J}(\chi_{F_1}, \chi_{F_2})\|_{L^1((AJ)^c)} \le (1+\nu)^{20} C_M A^{-M}\, |J| \inf_{x\in J} M(\chi_{F_1})(x) \inf_{x\in J} M_2(\chi_{F_2})(x) \,.$$

Proof. If we write $H_{S_J}(\chi_{F_1}, \chi_{F_2}) = (H_{S_J}(\chi_{F_1}, \chi_{F_2})\omega^{\frac12})\,\omega^{-\frac12}$ and use Hölder's inequality, we obtain:

$$\begin{aligned}\|H_{S_J}(\chi_{F_1}, \chi_{F_2})\|_{L^1((AJ)^c)} &\le \big\|\omega^{-\frac12}\big\|_{L^2((AJ)^c)} \|H_{S_J}(\chi_{F_1}, \chi_{F_2})\|_{L^2(\omega)} \\ &\le C\,A^{-M}|J|^{\frac12} \Big\| \sum_{s\in S_J} |J|^{-\frac12} \langle \chi_{F_1}, \varphi_{s_1}\rangle \langle \chi_{F_2}, \varphi_{s_2}\rangle \varphi_{s_3} \Big\|_{L^2(\omega)} \\ &\le C\,A^{-M} \|(\langle \chi_{F_1}, \varphi_{s_1}\rangle \langle \chi_{F_2}, \varphi_{s_2}\rangle)\|_{\ell^2(S_J)} \\ &\le C\,A^{-M} \|(\langle \chi_{F_1}, \varphi_{s_1}\rangle)\|_{\ell^\infty(S_J)} \|(\langle \chi_{F_2}, \varphi_{s_2}\rangle)\|_{\ell^2(S_J)} \\ &\le C\,(1+\nu)^{20} A^{-M} |J|^{-\frac12} \|\chi_{F_1}\|_{L^1(\omega^{-1})} \|\chi_{F_2}\|_{L^2(\omega^{-1})} \\ &\le C\,(1+\nu)^{20} A^{-M} |J|^{-\frac12} |J| \inf_{x\in J} M(\chi_{F_1})(x)\, |J|^{\frac12} \inf_{x\in J} M_2(\chi_{F_2})(x) \,.\end{aligned}$$

In the last estimate we have used the fact that

$$1 + \Big(\frac{|x-c(J)|}{|J|}\Big)^2 \ge 1 + \Big(\frac{|x-\theta|}{|J|} - \frac12\Big)^2$$

for all $\theta \in J$. □

The main conclusion is the following.

Lemma 3.2.

$$\left| \int_{E'} \sum_{s:I_s\subseteq\Omega} |I_s|^{-\frac12} \langle \chi_{F_1}, \phi_{s_1}\rangle \langle \chi_{F_2}, \phi_{s_2}\rangle \phi_{s_3}(x)\,dx \right| \le C_\nu (\min|F_1|, |F_2|)^{\frac12} |F_1|^{\frac12} |F_2|^{\frac12} |E|^{-\frac12} \,. \quad (3.4)$$

where $C_\nu \le C\,(1+\nu)^{20}$.

Proof. Since the roles of F_1 and F_2 are symmetric, it will suffice to prove that (3.4) holds with the expression $C\,|F_1|\,|F_2|^{\frac{1}{2}}|E|^{-\frac{1}{2}}$ on the right-hand side of the inequality.

We organize all dyadic intervals $J \subseteq \Omega$ into sets $\mathcal{F}_k$ ($k \geq 0$) in the following way:

$$\mathcal{F}_k = \left\{ J : 2^k J \subseteq \Omega,\ 2^{k+1} J \nsubseteq \Omega \right\}.$$

We note that

$$\sum_{J \in \mathcal{F}_k} |J| \leq 4|\Omega| \leq 2|E|\,.$$

Indeed, assume $J_{\max}$ is a maximal element of $\mathcal{F}_k$ with respect to inclusion. If $J \subseteq J_{\max}$ and $|J| < |J_{\max}|$, then J must have a common endpoint with $J_{\max}$ (otherwise, we would have $2^{k+1}J = 2^k(2J) \subseteq 2^k J_{\max} \subseteq \Omega$, thus $J \notin \mathcal{F}_k$). Thus, for each particular scale, $J_{\max}$ may contain at most 2 intervals belonging to $\mathcal{F}_k$. Therefore

$$\sum_{J \in \mathcal{F}_k,\, J \subseteq J_{\max}} |J| \leq \sum_{k=0}^{\infty} 2^{-k+1} |J_{\max}| \leq 4|J_{\max}|\,.$$

Since the maximal elements of $\mathcal{F}_k$ are disjoint, summing over them we obtain the required conclusion.

Also, for any $J \in \mathcal{F}_k$ we have $E' \subseteq (\Omega)^c \subseteq (2^k J)^c$. Thus, we have:

$$\begin{aligned}
&\left| \int_{E'} H_{\{I_s \subseteq \Omega\}}(\chi_{F_1}, \chi_{F_2})\, dx \right| \\
&\leq \sum_{J \subseteq \Omega} \left| \int_{E'} H_{S_J}(\chi_{F_1}, \chi_{F_2})\, dx \right| \\
&= \sum_{k=0}^{\infty} \sum_{J \in \mathcal{F}_k} \left| \int_{E'} H_{S_J}(\chi_{F_1}, \chi_{F_2})\, dx \right| \\
&\leq \sum_{k=0}^{\infty} \sum_{J \in \mathcal{F}_k} \| H_{S_J}(\chi_{F_1}, \chi_{F_2}) \|_{L^1((2^k J)^c)} \\
&\leq C_M (1+\nu)^{20} \sum_{k=0}^{\infty} \sum_{J \in \mathcal{F}_k} |J| 2^{-kM} \inf_{x \in J} M(\chi_{F_1}) \inf_{x \in J} M_2(\chi_{F_2}) \\
&\leq C_M (1+\nu)^{20} \sum_{k=0}^{\infty} 2^{-kM} C_0^{2k+2} \sum_{J \in \mathcal{F}_k} |J| \inf_{2^{k+1}J} M(\chi_{F_1}) \inf_{2^{k+1}J} M_2(\chi_{F_2}) \\
&\leq C' (1+\nu)^{20} \sum_{k=0}^{\infty} 2^{-kM} C_0^{2k+2} \sum_{J \in \mathcal{F}_k} |J| \frac{|F_1|}{|E|} \left(\frac{|F_2|}{|E|} \right)^{\frac{1}{2}} \\
&\leq C (1+\nu)^{20} |F_1|\,|F_2|^{\frac{1}{2}} |E|^{-\frac{1}{2}}\,.
\end{aligned}$$

□

4. Estimates for model sums. The case $I_s \nsubseteq \Omega$

We will now deal with the harder case $I_s \nsubseteq \Omega$. This part of the proof is based on an adaptation of the $L^2 \times L^2 \to L^{1,\infty}$ estimate in [9].

We denote by P the set of all tri-tiles $s \in S$, for which $I_s \nsubseteq \Omega$. Tri-tiles admit a partial order. We say that $s < s'$ if $I_s \subseteq I_{s'}$ and $\omega_{s'} \subseteq \omega_s$. We note that s and s' intersect as rectangles if and only if they are comparable under "$<$."

The separation of scales (2.8) allows to say that if $s < s'$, then $\omega_{s'} \subseteq \omega_{s_i}$ for some $i = 1, 2, 3$ or it is disjoint with all ω_{s_i}'s.

We say that a collection of tri-tiles T is a tree with top t if for all $s \in T$, $s < t$. Every finite collection of tri-tiles S is a union of trees. Indeed, if we denote by S^* the set of all elements in S which are maximal under "$<$," and, for each $t \in S^*$, T_t is the maximal tree in S with top t, then $S = \cup_{t \in S^*} T_t$. We refine the notion of the tree by saying that T is a *j-tree* ($j = 1, 2, 3$) if T is a tree with top T and for every $s \in T$, $\omega_{s_j} \cap \omega_t = \emptyset$.

For a tree T, $s \in T$, $s \neq t$, at most one of the intervals ω_{s_i} can intersect ω_t. Thus, if we denote $T_k = \{s \in T : \omega_{s_k} \cap \omega_t \neq \emptyset\}$, $k = 1, 2, 3$, then T_k is a j-tree for $j \neq k$ (there are also elements such that $\omega_{s_i} \cap \omega_t = \emptyset$ for all $i = 1, 2, 3$, but those may be added to any of the T_k's). Then $T = \cup_{k=1}^{3} T_k$, i.e., any tree is a union of at most three subtrees which are j-trees for at least two choices of j.

For a k-tree T we set

$$\Delta(T, k) = \frac{1}{\|f_k\|_2} \left(|I_t|^{-1} \sum_{s \in T} |\langle f_k, \varphi_{s_k} \rangle|^2 \right)^{\frac{1}{2}}$$

and we define the k-energy of a finite set of tiles S by

$$\mathcal{E}_k(S) = \sup \Delta(T, k) \ , \tag{4.1}$$

where the supremum is taken over all k-trees $T \subseteq S$. Note that a singleton $\{s\}$ is a k-tree for all k, so for all $s \in S$,

$$|I_s|^{-\frac{1}{2}} |\langle f_k, \varphi_{s_k} \rangle| \leq \mathcal{E}_k(S) \, \|f_k\|_2 \ .$$

Now fix some $j = 1, 2, 3$ and let T be a k-tree for $k \neq j$. Applying the above estimate and the Cauchy-Schwarz inequality, we deduce

$$\begin{aligned} |\langle H_T(f_1, f_2), f_3 \rangle| &\leq \sum_{s \in T} \frac{|\langle f_j, \varphi_{s_j} \rangle|}{|I_s|^{\frac{1}{2}}} \prod_{k \neq j} |\langle f_k, \varphi_{s_k} \rangle| \\ &\leq \mathcal{E}_j(S) \, \|f_j\|_2 \sum_{s \in T} \prod_{k \neq j} |\langle f_k, \varphi_{s_k} \rangle| \\ &\leq \mathcal{E}_j(S) \, \|f_j\|_2 \, |I_t| \prod_{k \neq j} \Delta(T, k) \|f_k\|_2 \\ &\leq |I_t| \prod_{j=1}^{3} \mathcal{E}_j(S) \, \|f_j\|_2 \ . \end{aligned} \tag{4.2}$$

This is a crucial estimate on a single tree that will be used in conjunction with the idea that any tree can be written as a union of three trees of the above type.

Next, we state the main lemma which will allow us to obtain the estimates for the model sums (cf. [9]).

Lemma 4.1. *Let S be a finite set of tri-tiles. Then S can be written as a union of two sets $S = S_1 \cup S_2$, which have the following properties. Let S_1^* be the set of elements which are maximal in S_1 under "$<$" (i.e., S_1 is a union of trees with tops in S_1^*). We then have*

$$\sum_{t\in S_1^*} |I_t| \leq C_1(1+\nu)^{20}\mathcal{E}_k(S)^{-2}\ , \tag{4.3}$$

$$\mathcal{E}_k(S_2) \leq \frac{1}{2}\,\mathcal{E}_k(S)\ . \tag{4.4}$$

This lemma only yields weak-type estimates from $L^2 \times L^2$ into $L^{1,\infty}$. But the fact that we are now working with the set of tiles $P = \{s \in S : I_s \nsubseteq \Omega\}$ and all functions are characteristic of some sets gives us an advantage quantified by the following energy estimate which appeared in [6, 4], and is essentially contained in [14].

Lemma 4.2. *For $k = 1, 2$ and $f_k = \chi_{F_k}$, there exists a constant $C > 0$, such that the following estimate is valid:*

$$\mathcal{E}_k(P) \leq C|E|^{-\frac{1}{2}} \min\left[\left(\frac{|F_k|}{|E|}\right)^{\frac{1}{2}}, \left(\frac{|F_k|}{|E|}\right)^{-\frac{1}{2}}\right]. \tag{4.5}$$

With these two lemmata at hand we can derive an estimate of the model sum for the case $I_s \nsubseteq \Omega$ in the following way. We construct inductively the sequence of pairwise disjoint sets P_j such that

$$P = \bigcup_{j=-\infty}^{n_0} P_j$$

and the following properties are satisfied:

(1) $\mathcal{E}_k(P_j) \leq 2^{j+1}$ for $k = 1, 2, 3$.

(2) P_j is a union of trees T_{jk} such that $\sum_k |I_{\mathrm{top}(T_{jk})}| \leq C_0(1+\nu)^{20}2^{-2j}$ for all $j \leq n_0$.

(3) $\mathcal{E}_k(P \setminus (P_{n_0} \cup \cdots \cup P_j)) \leq 2^j$ for $k = 1, 2, 3$.

This sequence is constructed in the following way: We start the induction at the number $j = n_0$ such that $\mathcal{E}_k \leq 2^{n_0}$ for $k = 1, 2, 3$. We set $P_{n_0} = \emptyset$. Then properties (1), (2), and (3) are clearly satisfied. Assuming that we have already constructed the set P_n, we construct P_{n-1} as follows. Let $S = P \setminus (P_{n_0} \cup \cdots \cup P_n)$. First, if $\mathcal{E}_1(S) > 2^{n-1}$, then apply Lemma 4.1 to S with $k = 1$, thus obtaining the sets $S_1^{(1)}$ with a control of the sum of the tops and $S_2^{(1)}$ with small 1-energy, otherwise just skip this step (i.e., $S_1^{(1)} = \emptyset$). Then, in the same fashion, if $\mathcal{E}_2(S_2^{(1)}) > 2^{n-1}$, we apply this lemma to $S_2^{(1)}$ obtaining the set $S_1^{(2)}$ and $S_2^{(2)}$ (otherwise again skipping this step, $S_1^{(2)} = \emptyset$). And, finally, we apply Lemma 4.1 for the third time with $k = 3$ to the set $S_2^{(2)}$ to obtain $S_1^{(3)}$ and $S_2^{(3)}$ (we also skip this step, if $\mathcal{E}_1(S_1^{(2)}) \leq 2^{n-1}$). We set $P_{n-1} = S_1^{(1)} \cup S_1^{(2)} \cup S_1^{(3)}$. Observe that if all three steps were skipped, then $P_{n-1} = \emptyset$. We have to verify that properties (1)–(3) indeed hold.

First, for $k = 1, 2, 3$:

$$\mathcal{E}_k(P \setminus (P_{n_0} \cup \cdots \cup P_{n-1})) \leq \frac{1}{2}\mathcal{E}_k(P \setminus (P_{n_0} \cup \cdots \cup P_n)) \leq 2^{n-1}$$

by Lemma 4.1 (and the fact that we just skipped the corresponding step if this was already so for some k), thus verifying (3). Then,

$$\mathcal{E}_k(P_{n-1}) \leq \mathcal{E}_k(P \setminus (P_{n_0} \cup \cdots \cup P_n)) \leq 2^n = 2^{(n-1)+1}\ ,$$

which proves (1). And finally, using Lemma 4.1, we have (with the convention that $S_1^{(0)} = S$):

$$\sum_k |I_{\mathrm{top}(T_{jk})}| \le C(1+\nu)^{20} \sum_{k=1}^{3} \mathcal{E}_k\big(S_1^{(k-1)}\big)^{-2} \le 3C(1+\nu)^{20} 2^{-2(n-1)} ,$$

since the sum actually ranges over those values of k for which $\mathcal{E}_k(S_1^{(k-1)}) > 2^{n-1}$, otherwise the corresponding part of P_{n-1} is empty.

Taking into account the above families P_j, we obtain the following:

$$\begin{aligned}
&|\langle H_P(\chi_{F_1}, \chi_{F_2}), \chi_{E'}\rangle| \\
&\quad\le \sum_{j=-\infty}^{\infty} \sum_k |\langle H_{T_{jk}}(\chi_{F_1}, \chi_{F_2}), \chi_{E'}\rangle| \\
&\quad\le C \sum_{j=-\infty}^{\infty} \left(\sum_k |I_{\mathrm{top}_{T_{jk}}}|\right) \mathcal{E}_1(\chi_{F_1}, S_j)\mathcal{E}_2(\chi_{F_2}, S_j)\mathcal{E}_3(\chi_{E'}, S_j)|F_1|^{\frac12}|F_2|^{\frac12}|E|^{\frac12} \\
&\quad\le C' \sum_{j=-\infty}^{\infty} 2^{-2j} \min\left(|F_1|^{-\frac12}, \frac{|F_1|^{\frac12}}{|E|}, 2^j\right) \min\left(|F_2|^{-\frac12}, \frac{|F_2|^{\frac12}}{|E|}, 2^j\right) 2^j\, |F_1|^{\frac12}|F_2|^{\frac12}|E|^{\frac12} \\
&\quad= C' \sum_{j=-\infty}^{\infty} 2^{-j} \min\left(|F_1|^{-\frac12}, \frac{|F_1|^{\frac12}}{|E|}, 2^j\right) \min\left(|F_2|^{-\frac12}, \frac{|F_2|^{\frac12}}{|E|}, 2^j\right) |F_1|^{\frac12}|F_2|^{\frac12}|E|^{\frac12}, \quad (4.6)
\end{aligned}$$

where we used the estimate on a single tree (4.2) and the improved energy estimate (4.5).

We control (4.6) in different cases:

(A) Suppose $|E| \ge |F_2| \ge |F_1|$. Then (4.6) is bounded by

$$\begin{aligned}
&\left(\sum_{j=-\infty}^{\log \frac{|F_1|^{\frac12}}{|E|}} 2^j + \sum_{j=\log \frac{|F_1|^{\frac12}}{|E|}}^{\log \frac{|F_2|^{\frac12}}{|E|}} |F_1|^{\frac12}|E|^{-1} + \sum_{j=\log \frac{|F_2|^{\frac12}}{|E|}}^{\infty} 2^{-j}|F_1|^{\frac12}|F_1|^{\frac12}|E|^{-2}\right)|F_1|^{\frac12}|F_2|^{\frac12}|E|^{\frac12} \\
&\qquad \lesssim |F_1|\,|F_2|^{\frac12}|E|^{-\frac12}\left(1 + \log\frac{|F_2|}{|F_1|}\right).
\end{aligned}$$

So, by symmetry, in the case $|E| \ge |F_1|, |F_2|$ the expression (4.6) can be controlled by

$$\min\big(|F_1|^{\frac12}, |F_2|^{\frac12}\big)|F_1|^{\frac12}\,|F_2|^{\frac12}|E|^{-\frac12}\left(1 + \left|\log\frac{|F_2|}{|F_1|}\right|\right). \quad (4.7)$$

We may also note that in this case $|\log \frac{|F_2|}{|F_1|}| \le \log \frac{|E|^2}{|F_1|\,|F_2|}$.

(B) Suppose that $|F_1| \le |E| \le |F_2|$ and $|E|^2 \ge |F_1|\,|F_2|$. In this case, we can bound (4.6) by

$$\begin{aligned}
&\left(\sum_{j=-\infty}^{\log \frac{|F_1|^{\frac12}}{|E|}} 2^j + \sum_{j=\log \frac{|F_1|^{\frac12}}{|E|}}^{\log |F_2|^{-\frac12}} |F_1|^{\frac12}|E|^{-1} + \sum_{j=\log |F_2|^{-\frac12}}^{\infty} 2^{-j}|F_1|^{\frac12}|F_1|^{-\frac12}|E|^{-1}\right)|F_1|^{\frac12}|F_2|^{\frac12}|E|^{\frac12} \\
&\qquad \lesssim |F_1|\,|F_2|^{\frac12}|E|^{-\frac12}\left(1 + \log\frac{|E|^2}{|F_1|\,|F_2|}\right).
\end{aligned}$$

Thus, by symmetry, in the case when $|E|$ is between $|F_1|$ and $|F_2|$ and $|E|^2 \geq |F_1|\,|F_2|$ we obtain that (4.6) is bounded by

$$\min\left(|F_1|^{\frac{1}{2}}, |F_2|^{\frac{1}{2}}\right)|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}|E|^{-\frac{1}{2}}\left(1+\log\frac{|E|^2}{|F_1|\,|F_2|}\right). \tag{4.8}$$

The other cases work in a similar way:

(C) If $|E|$ is between $|F_1|$ and $|F_2|$, but $|E|^2 \leq |F_1|\,|F_2|$, the bound is

$$\min\left(|F_1|^{\frac{1}{2}}, |F_2|^{\frac{1}{2}}\right)|E|^{\frac{1}{2}}\left(1+\log\frac{|F_1|\,|F_2|}{|E|^2}\right). \tag{4.9}$$

(D) For $|E| \leq |F_1|,\ |F_2|$, we obtain the bound

$$\min\left(|F_1|^{\frac{1}{2}}, |F_2|^{\frac{1}{2}}\right)|E|^{\frac{1}{2}}\left(1+\left|\log\frac{|F_1|}{|F_2|}\right|\right). \tag{4.10}$$

Combining the four Cases (A), (B), (C), and (D) we obtain the following inequality for the case when the tiles s satisfy $I_s \nsubseteq \Omega$:

$$\begin{aligned}&\left|\int_{E'} H_{\{s:\,I_s\nsubseteq\Omega\}}(\chi_{F_1},\chi_{F_2})\,dx\right|\\ &\quad\leq C_1\min\left(|F_1|^{\frac{1}{2}}, |F_2|^{\frac{1}{2}}\right)\min\left(\frac{|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}}{|E|^{\frac{1}{2}}},\ |E|^{\frac{1}{2}}\right)\left(1+\left|\left|\log\frac{|F_1|}{|E|}\right|-\left|\log\frac{|F_2|}{|E|}\right|\right|\right).\end{aligned} \tag{4.11}$$

As a consequence of the results so far we deduce the following.

Proposition 4.3. *There exists a constant C_1 such that, for any sets E, F_1, F_2 with the property that $|E|^2 \geq |F_1|\,|F_2|$ there exists a set $E' \subseteq E$ with $|E'| \geq \frac{1}{2}|E|$ such that for any set of tri-tiles S we have the following estimate:*

$$\left|\int_{E'} H_S(\chi_{F_1},\chi_{F_2})(x)\,dx\right| \leq C_1\min(|F_1|,|F_2|)^{\frac{1}{2}}|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}|E|^{-\frac{1}{2}}\left(1+\log\frac{|E|^2}{|F_1|\,|F_2|}\right). \tag{4.12}$$

This estimate is also valid for the bilinear Hilbert transform H.

Proof. The result for H_S follows from the estimates (3.4) and (4.9). Note that the construction of E' did not depend on the choice of the set of tri-tiles, so E' is the same for any S, and by an averaging argument this estimate is also valid for H. □

It is clear that, since both adjoints of H_S, are "essentially" the same operators, the same estimate (with different constants) also holds for them.

5. $L^{r_1} \times L^{r_2} \to L^r$ boundedness of the model sums

In this section we will show that estimates (3.4) and (4.9) imply boundedness of the model sum operator H_S from $L^{r_1} \times L^{r_2}$ to L^r for $\frac{1}{r_1}+\frac{1}{r_2}=\frac{1}{r}$, $r_1, r_2 > 1$, $r > \frac{2}{3}$. We include this section for the sake of completeness (as we will use this result in the sequel), but we point out that the reader may wish to skip it and cite the results of Lacey and Thiele [11, 12].

Take some p_1, p_2 such that $\frac{1}{p_1} + \frac{1}{p_2} = \frac{3}{2}$ and $p_1, p_2 > 1$. We will show that H_S is of restricted weak type (r_1, r_2, r) where $\frac{1}{r_1} = \frac{1}{p_1} - \varepsilon$, $\frac{1}{r_2} = \frac{1}{p_2} - \varepsilon$ and $\frac{1}{r} = \frac{3}{2} - 2\varepsilon$. By interpolation it follows that H_S is bounded from $L^{r_1} \times L^{r_2}$ to L^r when $|r_1^{-1} - r_2^{-1}| < 1/2$ and $2/3 < r < 2$. The boundedness of H_S in the remaining range of exponents follows by duality. (Note that the same conclusion may be obtained using the interpolation theorem of Grafakos and Tao [7] as the operator H_S has bounded kernel whenever S is a finite set.)

We recall that a bilinear operator T is of restricted weak type (r_1, r_2, r) if and only if the following is valid: For any sets E, F_1, F_2 of finite measure there exists a set $E' \subset E$ with $|E'| \geq \frac{1}{2}|E|$, such that

$$\left| \int_{E'} T(\chi_{F_1}, \chi_{F_2})(x)\, dx \right| \lesssim \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}} . \tag{5.1}$$

Take arbitrary sets E, F_1, F_2 of finite positive measure. It follows from (3.4) and (4.9) that

$$\left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})\, dx \right| \lesssim \frac{|F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{2}}} \left(1 + \left| \log \frac{|F_1|}{|E|} \right|\right) \left(1 + \left| \log \frac{|F_2|}{|E|} \right|\right) . \tag{5.2}$$

We will use the fact that $1 + \log x \lesssim x^\varepsilon$ for $x \geq 1$. In the case when $|E| \geq \max(|F_1|, |F_2|)$ we can estimate the right-hand side of (5.2) by the expression

$$\frac{|F_1|^{\frac{1}{p_1}} |F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{2}}} \left(1 + \log \frac{|E|}{|F_1|}\right) \left(1 + \log \frac{|E|}{|F_2|}\right) \lesssim \frac{|F_1|^{\frac{1}{p_1}-\varepsilon} |F_2|^{\frac{1}{p_2}-\varepsilon}}{|E|^{\frac{1}{2}-2\varepsilon}} = \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}} .$$

Now consider the case $|F_1| \leq |E| \leq |F_2|$ (as the case $|F_2| \leq |E| \leq |F_1|$ is symmetric). Fix some $\varepsilon_1 > 2\varepsilon$. Put $\alpha = \frac{1}{p_1} - \varepsilon + \varepsilon_1$ (ε and ε_1 have to be chosen small enough, so that $\alpha \leq 1$) and $\beta = \frac{1}{p_2} - \varepsilon_1 + \varepsilon$ (thus $\beta \leq 1$ also). We have $\alpha + \beta = \frac{3}{2}$. Thus, similarly to (5.2), we obtain:

$$\begin{aligned}
\left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})\, dx \right| &\lesssim \frac{|F_1|^{\alpha} |F_2|^{\beta}}{|E|^{\frac{1}{2}}} \left(1 + \log \frac{|E|}{|F_1|}\right) \left(1 + \log \frac{|F_2|}{|E|}\right) \\
&\lesssim \frac{|F_1|^{\alpha} |F_2|^{\beta}}{|E|^{\frac{1}{2}}} \left(\frac{|E|}{|F_1|}\right)^{\varepsilon_1} \left(\frac{|F_2|}{|E|}\right)^{\varepsilon_1 - 2\varepsilon} \\
&= \frac{|F_1|^{\frac{1}{p_1}-\varepsilon} |F_2|^{\frac{1}{p_2}-\varepsilon}}{|E|^{\frac{1}{2}-2\varepsilon}} = \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}} .
\end{aligned}$$

The remaining case is $|E| \leq \min(|F_1|, |F_2|)$. We observe that in this case the set Ω is empty, since $M(\chi_{F_i}) \leq 1$. We therefore only need to use (4.6) which for $|E|$ small yields:

$$\begin{aligned}
\left| \int_{E'} H_S(\chi_{F_1}, \chi_{F_2})\, dx \right| &\lesssim \min(|F_1|, |F_2|)^{\frac{1}{2}} |E|^{\frac{1}{2}} \left(1 + \log \frac{|F_1|}{|E|}\right) \left(1 + \log \frac{|F_2|}{|E|}\right) \\
&\lesssim |F_1|^{\frac{1}{p_1}-\frac{1}{2}} |F_2|^{\frac{1}{p_2}-\frac{1}{2}} |E|^{\frac{1}{2}} \left(\frac{|F_1|}{|E|}\right)^{\frac{1}{2}-\varepsilon} \left(\frac{|F_2|}{|E|}\right)^{\frac{1}{2}-\varepsilon} \\
&= \frac{|F_1|^{\frac{1}{p_1}-\varepsilon} |F_2|^{\frac{1}{p_2}-\varepsilon}}{|E|^{\frac{1}{2}-2\varepsilon}} = \frac{|F_1|^{\frac{1}{r_1}} |F_2|^{\frac{1}{r_2}}}{|E|^{\frac{1}{r}-1}} .
\end{aligned}$$

Thus, for any measurable sets E and F_1, F_2, H_S satisfies (5.1) and this implies that it is of restricted weak type (r_1, r_2, r). The strong type estimates for the same range of exponents can now be obtained by varying r_1 and r_2 and using the result on interpolation between adjoint operators (cf. [7]).

6. Distributional estimates corresponding to the case $p_1 = 1, 2 \leq p_2 < \infty$

Fix $2 \leq p_2 < \infty$. There are some minor differences in the treatment of the cases $p_2 = 2$ and $p_2 > 2$. In the case $p_2 = 2$ for the moment we shall assume that $|F_1| \leq |F_2|$.

Case: $p_2 = 2, |E|^{\frac{3}{2}} \geq |F_1|\,|F_2|^{\frac{1}{2}}, |F_1| \leq |F_2| \leq |E|$.

Since $|E|^{\frac{3}{2}} \geq |F_1|\,|F_2|^{\frac{1}{2}}$ and $|F_2| \geq |F_1|$, we have $|E|^2 \geq |F_1|\,|F_2|$. Using estimate (4.12) we obtain

$$\left|\int_{E'} H(\chi_{F_1}, \chi_{F_2})\,dx\right| \leq C_1 \frac{|F_1|\,|F_2|^{\frac{1}{2}}}{|E|^{\frac{1}{2}}}\left(1 + \log\frac{|E|^{\frac{3}{2}}}{|F_1|\,|F_2|^{\frac{1}{2}}}\right). \tag{6.1}$$

We note that this estimate is also valid if $|E| \geq \max|F_i|$, even when $|F_1| \geq |F_2|$. We will use this estimate in the inductive procedures below.

Case: $p_2 > 2, |E|^{1+\frac{1}{p_2}} \geq |F_1|\,|F_2|^{\frac{1}{p_2}}, |E| \geq |F_2|$.

Let $\alpha = \frac{1}{2} - \frac{1}{p_2} > 0$, $\beta = 1 - \frac{1}{p_2} > 0$. Since $|E| \geq |F_2|$ we must have $|E|^2 \geq |F_1|\,|F_2|$. Using (4.12) we obtain

$$\begin{aligned}\left|\int_{E'} H(\chi_{F_1}, \chi_{F_2})(x)\,dx\right| &\leq C_1 \min(|F_1|, |F_2|)^{\frac{1}{2}}|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}|E|^{-\frac{1}{2}}\left(1 + \log\frac{|E|^2}{|F_1|\,|F_2|}\right)\\ &\leq C_1 \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\left(\frac{|F_2|}{|E|}\right)^{\alpha}\left(1 + \log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}} + \log\frac{|E|^{\beta}}{|F_2|^{\beta}}\right)\\ &\lesssim \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\left(1 + \log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right),\end{aligned} \tag{6.2}$$

since the function $f(x) = x^{\alpha}(1 + \log\frac{1}{x^{\beta}})$ is bounded on $[0, 1]$ when $\alpha > 0$ (here $x = \frac{|F_2|}{|E|}$).

Case: $p_2 \geq 2, |E|^{1+\frac{1}{p_2}} \geq |F_1|\,|F_2|^{\frac{1}{p_2}}, |E| \leq |F_2|$ (which implies $|E| \geq |F_1|$).

In this case, we will obtain an estimate via an iterative procedure which consists of two parts. Let us denote by H^{*2} the adjoint of H with respect to the second variable. At first, we set $F_2^0 = F_2$. We will continue this part of the iteration until the first integer n such that $|F_2^n| \leq |E|$. At the j^{th} step, according to the estimates above, we choose a subset S^j of F_2^j with $|S^j| \geq \frac{1}{2}|F_2^j|$, such that:

$$\left|\int_{S^j} H^{*2}(\chi_{F_1}, \chi_E)(x)\,dx\right| \lesssim \frac{|F_1|\,|E|^{\frac{1}{p_2}}}{|F_2^j|^{\frac{1}{p_2}}}\left(1 + \log\frac{\left|F_2^j\right|^{1+\frac{1}{p_2}}}{|F_1|\,|E|^{\frac{1}{p_2}}}\right) \leq |F_1|\left(1 + \log\frac{|F_2|^{1+\frac{1}{p_2}}}{|F_1|\,|E|^{\frac{1}{p_2}}}\right).$$

Then we set $F_2^{j+1} = F_2^j \setminus S^j$. Obviously, for the number of steps n we have $n \lesssim 1 + \log\frac{|F_2|}{|E|}$.

Thus, we have

$$\begin{aligned}\left|\int_E H(\chi_{F_1},\chi_{F_2})\,dx\right| &\lesssim |F_1|\left(1+\log\frac{|F_2|^{1+\frac{1}{p_2}}}{|F_1|\,|E|^{\frac{1}{p_2}}}\right)\left(1+\log\frac{|F_2|}{|E|}\right)+\left|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\right|\\ &\lesssim \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\left(1+\log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right)^2+\left|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\right|.\end{aligned}$$

In the last line we have used the following simple inequality (with $a=\frac{|E|}{|F_1|}$, $b=\frac{|F_2|}{|E|}$):

For $a\ge 1$, $b\ge 1$, such that $ab^{-\frac{1}{p_2}}\ge 1$ we have

$$\left(1+\log\left(ab^{1+\frac{1}{p_2}}\right)\right)\left(1+\log b\right)\lesssim\left(1+\log\frac{a}{b^{\frac{1}{p_2}}}\right)^2 b^{\frac{1}{p_2}}\,. \tag{6.3}$$

To prove (6.3) we note that if $b^{\frac{1}{p_2}}\le\sqrt{a}$, then $\log\frac{a}{b^{\frac{1}{p_2}}}\ge\log\sqrt{a}=\frac{1}{2}\log a$ and we have

$$\left(1+\log\left(ab^{1+\frac{1}{p_2}}\right)\right)\left(1+\log b\right)\lesssim\left(1+\log a\right)^2\lesssim\left(1+\log\frac{a}{b^{\frac{1}{p_2}}}\right)^2 b^{\frac{1}{p_2}}\,,$$

while when $\sqrt{a}\le b^{\frac{1}{p_2}}\le a$, then

$$\left(1+\log\left(ab^{1+\frac{1}{p_2}}\right)\right)\left(1+\log b\right)\lesssim\left(1+\log b\right)^2\lesssim b^{\frac{1}{p_2}}\,.$$

It remains to estimate the term

$$\left|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\right|.$$

In the second part of the iteration process we proceed in a similar manner, only now we will be splitting either F_2 or E, depending on which one is larger in size. We set $E^n=E$. At the j^{th} step, if $|E^j|\ge|F_2^j|$, we choose $S^j\subset E^j$ such that $|S^j|\ge\frac{1}{2}|E^j|$ and

$$\begin{aligned}\left|\int_{S^j} H(\chi_{F_1},\chi_{F_2^j})\,dx\right| &\lesssim \frac{|F_1|\left|F_2^j\right|^{\frac{1}{p_2}}}{\left|E^j\right|^{\frac{1}{p_2}}}\left(1+\log\frac{\left|E^j\right|^{1+\frac{1}{p_2}}}{|F_1|\left|F_2^j\right|^{\frac{1}{p_2}}}\right)\\ &\le |F_1|\frac{\left|F_2^j\right|^{\frac{1}{p_2}}}{\left|E^j\right|^{\frac{1}{p_2}}}\left(1+\log\frac{\left|E^j\right|^{\frac{1}{p_2}}}{\left|F_2^j\right|^{\frac{1}{p_2}}}+\log\frac{|E|}{|F_1|}\right)\\ &\lesssim |F_1|\left(1+\log\frac{|E|}{|F_1|}\right),\end{aligned}$$

where we have once again made use of the fact that $f(x)=x\cdot\log\frac{1}{x}$ is bounded on $[0,1]$ $\Big(x=\frac{|F_2^j|^{\frac{1}{p}}}{|E^j|^{\frac{1}{p}}}\le 1\Big)$.

In the other case, when $|F_2^j|\ge|E^j|$, we choose $S^j\subset F_2^j$ with $|S^j|\ge\frac{1}{2}|F_2^j|$ such that

$$\left|\int_{S^j} H^{*2}\big(\chi_{F_1},\chi_{E^j}\big)\,dx\right|\lesssim\frac{|F_1|\left|E^j\right|^{\frac{1}{p_2}}}{\left|F_2^j\right|^{\frac{1}{p_2}}}\left(1+\log\frac{\left|F_2^j\right|^{1+\frac{1}{p_2}}}{|F_1|\left|E^j\right|^{\frac{1}{p_2}}}\right).$$

An identical calculation and the fact that $|F_2^j| \le |E|$ show that this can also be dominated by $|F_1|\big(1+\log\frac{|E|}{|F_1|}\big)$.

In the first case we set $F_2^{j+1} = F_2^j$, $E^{j+1} = E^j \setminus S^j$. In the second case we set $F_2^{j+1} = F_2^j \setminus S^j$, $E^{j+1} = E^j$. We proceed until the first integer m such that both $|E^m|, |F_2^m| \le |F_1|$. Obviously, the number of steps in the second part $m \lesssim \big(1+\log\frac{|E|}{|F_1|}\big)$. We now have

$$\begin{aligned}
\left|\int_E H(\chi_{F_1},\chi_{F_2^n})\,dx\right| &= \left|\int_{E^{n+1}\cup S^n} H(\chi_{F_1},\chi_{F_2^n})\,dx\right| \\
&\le \left|\int_{S^n} H(\chi_{F_1},\chi_{F_2^n})\,dx\right| + \left|\int_{E^{n+1}} H(\chi_{F_1},\chi_{F_2^{n+1}})\,dx\right| \\
&\lesssim |F_1|\left(1+\log\frac{|E|}{|F_1|}\right) + \left|\int_{E^{n+1}} H(\chi_{F_1},\chi_{F_2^{n+1}})\,dx\right| \\
&\lesssim \dots \\
&\lesssim m\,|F_1|\left(1+\log\frac{|E|}{|F_1|}\right) + \left|\int_{E^m} H(\chi_{F_1},\chi_{F_2^m})\,dx\right| \\
&\lesssim |F_1|\left(1+\log\frac{|E|}{|F_1|}\right)^2 + \left|E^m\right|^{\frac12}|F_1|^{\frac14}\left|F_2^m\right|^{\frac14} \\
&\lesssim |F_1|\frac{|F_2|^{\frac{1}{p_2}}}{|E|^{\frac{1}{p_2}}}\left(1+\log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right)^2,
\end{aligned}$$

where we made use of the boundedness of H on $L^4 \times L^4 \to L^2$ and the following inequality: For any $a \ge 1$, $b \ge 1$, such that $ba^{-\frac{1}{p_2}} \ge 1$ we have

$$(1+\log b)^2 \lesssim \left(1+\log\left(ba^{-\frac{1}{p_2}}\right)\right)^2 a^{\frac{1}{p_2}} ,$$

with $a = \frac{|F_2|}{|E|}$, $b = \frac{|E|}{|F_1|}$. The proof of this inequality is similar to that in (6.3) and is omitted.

Case: $p_2 \ge 2$, $|E|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$. (We are still assuming that $|F_1| \le |F_2|$ when $p_2 = 2$.)

Here we will need the following lemma.

Lemma 6.1. *Let $2 \le p_2 < \infty$. For all measurable sets E, F_1, F_2 of finite measure satisfying $|E|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$ (and also $|F_1| \le |F_2|$ when $p_2 = 2$) we have*

$$\left|\int_E H(\chi_{F_1},\chi_{F_2})(x)\,dx\right| \lesssim |E|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\right)^2 .$$

Proof. Let us denote $F_i^0 = F_i$ for $i = 1, 2$. We shall now employ an inductive procedure similar to the one described above. At the j^{th} step among the sets F_1^j and F_2^j we choose the one which has greater size and denote it by $F_{\max}^j$ and the other one by $F_{\min}^j$. By $H^{*\max}$ we shall denote the expression $H^{*1}(\chi_E, \chi_{F_2^j})$ in the case when $F_{\max}^j = F_1^j$ and $H^{*2}(\chi_{F_1^j}, \chi_E)$ in the other case. By (6.2) or (6.1) applied to the respective adjoint of H with the roles of E and $F_{\max}^j$ interchanged, we can choose $S^j \subset F^j_{\max}$ such that $|S^j| \geq \frac{1}{2}|F^j_{\max}|$ and

$$\left| \int_{S^j} H^{*\max}(x)\, dx \right| \lesssim \frac{|E| \left|F^j_{\min}\right|^{\frac{1}{p_2}}}{\left|F^j_{\max}\right|^{\frac{1}{p_2}}} \left(1 + \log \frac{\left|F^j_{\max}\right|^{1+\frac{1}{p_2}}}{|E| \left|F^j_{\min}\right|^{\frac{1}{p_2}}} \right) . \tag{6.4}$$

We define $F_i^{j+1} = F_i^j \setminus S_i^j$ for all $i = 1, 2$, where we set $S_i^j = S^j$ if $F^j_{\max} = F^j_i$ and $S^j_i = \emptyset$ otherwise. Let us examine the right-hand side of the inequality above. If $|E| \leq |F^j_{\min}|$, it is easy to check that

$$\frac{\left|F^j_{\max}\right|^{1+\frac{1}{p_2}}}{\left|F^j_{\min}\right|^{\frac{1}{p_2}} |E|} \leq \left(\frac{\left|F^j_{\min}\right| \left|F^j_{\max}\right|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}} \right)^{p_2+1} \leq \left(\frac{|F_{\min}|\, |F_{\max}|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}} \right)^{p_2+1} \leq \left(\frac{|F_1|\, |F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}} \right)^{p_2+1} .$$

Thus, in this case we can estimate the right-hand side by $(p_2+1)|E|\big(1 + \log \frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\big)$.

In the case when $|E| \geq |F^j_{\min}|$, we have

$$\frac{\left|F_{\max}^{1+\frac{1}{p_2}}\right|^2}{\left|F^j_{\min}\right|^{\frac{1}{p}} |E|} \leq \frac{\left|F_{\max}\right|^{1+\frac{1}{p_2}}}{\left|F_{\min}\right|^{1+\frac{1}{p_2}}} .$$

So, in this case, the right-hand side of the inequality can be estimated by

$$|E| \left(\frac{\left|F^j_{\min}\right|}{\left|F^j_{\max}\right|} \right)^{\frac{1}{p_2}} \left(1 + \log \frac{\left|F^j_{\max}\right|}{\left|F^j_{\min}\right|} \right) \lesssim |E| ,$$

since the function $f(x) = x^{\frac{1}{p_2}}(1 + \log \frac{1}{x})$ is bounded for $x \in [0, 1]$.

Thus, in each case we get

$$\left| \int_{S^j} H^{*\max}(x)\, dx \right| \leq C'' \, |E| \left(1 + \log \frac{|F_1|\, |F_2|^{\frac{1}{p}}}{|E|^{1+\frac{1}{p}}} \right) .$$

We proceed by induction and we stop at the first integer n such that

$$|E|^{1+\frac{1}{p}} \geq \left|F_1^n\right| \left|F_2^n\right|^{\frac{1}{p_2}} .$$

(Such an integer always exists since the quantity $|F_1^n|\,|F_2^n|^{\frac{1}{p_2}}$ gets smaller by at least a factor of $\frac{1}{2^{\frac{1}{p_2}}}$ when j is replaced by $j+1$.) Obviously, the number of steps $n \lesssim 1 + \log \frac{|F_1|\,|F_2|^{\frac{1}{p}}}{|E|^{1+\frac{1}{p_2}}}$.

We can now estimate

$$\begin{aligned}
\left|\int_E H(\chi_{F_1},\chi_{F_2})\,dx\right| &= \left|\int_E H(\dots,\chi_{S^0}+\chi_{F^1_{\max}},\dots)\,dx\right| \\
&\le \left|\int_{S^0} H^{*\max}(x)\,dx\right| + \left|\int_E H(\chi_{F_1^1},\chi_{F_2^1})\,dx\right| \\
&\lesssim |E|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\right) + \left|\int_E H(\dots,\chi_{S^1}+\chi_{F^2_{\max}},\dots)\,dx\right| \\
&\le \dots \\
&\lesssim n\,|E|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\right) + \left|\int_E H(\chi_{F_1^n},\chi_{F_2^n})\,dx\right| \\
&\lesssim |E|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^2}\right)^2 + |E|^{1-\frac{p_2+1}{\theta p_2}}\left|F_1^n\right|^{\frac{1}{\theta}}\left|F_2^n\right|^{\frac{1}{\theta p_2}} \\
&\le C_2\,|E|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\right)^2,
\end{aligned}$$

where in the second line from the bottom we have used the Hölder inequality and the fact that H is of strong type $(\theta,\theta p_2,\theta\frac{p_2}{p_2+1})$ for some large θ. □

In the case $p_2 > 2$ we obtain the following estimate: For any sets F_1, F_2, and E of finite measure we can find $E' \subset E$ with $|E'| \ge \frac{1}{2}|E|$ such that

$$\left|\int_{E'} H(\chi_{F_1},\chi_{F_2})\,dx\right| \lesssim |E|\min\left[1,\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{|E|^{1+\frac{1}{p_2}}}\right]\left[1+\left|\log\frac{|E|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right|\right]^2. \tag{6.5}$$

We now remove the assumption that $|F_1| \le |F_2|$ when $p_2 = 2$. For $p_2 = 2$, we can consider the (symmetric) case when $|F_1| \ge |F_2|$, proceed as above with the roles of F_1 and F_2 interchanged and putting together the two estimates we obtain: For any sets F_1, F_2, and E of finite measure we can find a set $E' \subset E$ with $|E'| \ge \frac{1}{2}|E|$ such that

$$\left|\int_{E'} H(\chi_{F_1},\chi_{F_2})\,dx\right| \lesssim |E|\min\left[1,\frac{\min(|F_i|)^{\frac{1}{2}}|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}}{|E|^{\frac{3}{2}}}\right]\left[1+\left|\log\frac{|E|^{\frac{3}{2}}}{\min(|F_i|)^{\frac{1}{2}}|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}}\right|\right]^2. \tag{6.6}$$

7. Distributional estimates for the bilinear Hilbert transform

We can now prove Theorem 1.1.

Proof. For a given $\lambda > 0$, we set

$$\begin{aligned}
E_\lambda^+ &= \{H(\chi_{F_1},\chi_{F_2}) > \lambda\}, \\
E_\lambda^- &= \{H(\chi_{F_1},\chi_{F_2}) < -\lambda\}.
\end{aligned}$$

Suppose that $|E_\lambda^+|^{1+\frac{1}{p_2}} > |F_1|\,|F_2|^{\frac{1}{p_2}}$. Then by (6.5) there is a subset S_λ^+ of E_λ^+ of at least half its measure so that

$$\frac{\lambda}{2}\left|E_\lambda^+\right| \le \left|\int_{S_\lambda^+} H(\chi_{F_1},\chi_{F_2})\,dx\right| \le C_3\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{\left|E_\lambda^+\right|^{\frac{1}{p_2}}}\left(1+\log\frac{\left|E_\lambda^+\right|^{1+\frac{1}{p_2}}}{|F_1|\,|F_2|^{\frac{1}{p_2}}}\right)^2.$$

This implies that

$$\left|E_\lambda^+\right| \le C_4\left(|F_1|\,|F_2|^{\frac{1}{p_2}}\right)^{\frac{p_2}{p_2+1}} \lambda^{-\frac{p_2}{p_2+1}} \left(1+\log\frac{1}{\lambda}\right)^{\frac{2p_2}{p_2+1}} . \tag{7.1}$$

But then this implies that there is a $\lambda_0 > 0$ such that for $\lambda > \lambda_0$ we have $|E_\lambda^+|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$. Thus, for $\lambda > \lambda_0$, $|E_\lambda^+|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$ holds and estimate (6.5) gives

$$\lambda\left|E_\lambda^+\right| \le C_5\left|E_\lambda^+\right|\left(1+\log\frac{|F_1|\,|F_2|^{\frac{1}{p_2}}}{\left|E_\lambda^+\right|^{1+\frac{1}{p_2}}}\right)^2$$

from which one easily deduces that

$$\left|E_\lambda^+\right| \le \frac{1}{2}C\, e^{-c\sqrt{\lambda}}\left(|F_1|\,|F_2|^{\frac{1}{p_2}}\right)^{\frac{p_2}{p_2+1}} . \tag{7.2}$$

Suppose now that $\lambda \le \lambda_0$. As shown, if $|E_\lambda^+|^{1+\frac{1}{p_2}} > |F_1|\,|F_2|^{\frac{1}{p_2}}$, then (7.1) is valid. If $|E_\lambda^+|^{1+\frac{1}{p_2}} \le |F_1|\,|F_2|^{\frac{1}{p_2}}$ then (7.2) holds which is even stronger.

An identical argument yields the same result for $|E_\lambda^-|$ with the same λ_0.

For $p_2 = 2$ we run the same argument for estimate (6.6) and in the end dominate the expression $\min(|F_1|^{\frac{1}{2}}, |F_2|^{\frac{1}{2}})|F_1|^{\frac{1}{2}}\,|F_2|^{\frac{1}{2}}$ by $|F_1|\,|F_2|^{\frac{1}{2}}$.

Replacing the constants C, c by different ones we may take $\lambda_0 = 1$ and thus estimate (1.2) is now proved. Estimate (1.3) is proved likewise. Finally, Corollaries 1.2 and 1.3 are easy consequences of these estimates. □

References

[1] Antonov, N. Y. Convergence of Fourier series, Proceedings of the XXth Workshop on Function Theory (Moscow 1995), *East J. Approx.* **2**, 187–196, (1996).

[2] Carleson, L. On convergence and growth of partials sums of Fourier series, *Acta Math.* **116**, 135–157, (1966).

[3] Fefferman, C. Pointwise convergence of Fourier series, *Ann. of Math.* **98**, 551–571, (1973).

[4] Grafakos, L. *Classical and Modern Fourier Analysis,* Prentice-Hall, Pearson Education, Upper Saddle River, NJ, (2003).

[5] Grafakos, L. and Lennard, C. Characterization of $L^p(R^n)$ using Gabor frames, *J. Fourier Anal. Appl.* **7**(2), 101–126, (2001).

[6] Grafakos, L., Tao, T., and Terwilleger, E. L^p bounds for a maximal dyadic sum operator, *Math. Zeit.* **246**, 321–337, (2004).

[7] Grafakos, L. and Tao, T. Multilinear interpolation between adjoint operators, *J. Funct. Anal.* **199**, 379–385, (2003).

[8] Hunt, R. A. *On the Convergence of Fourier Series, Orthogonal Expansions and their Continuous Analogues,* (Proc. Conf. Edwardsville, IL, 1967), Haimo, D. T., Ed., Southern Illinois Univ. Press, Carbondale IL, 235–255.

[9] Lacey, M. T. On the bilinear Hilbert transform, *Doc. Math.* Extra Vol. II, 647–656, (1998).

[10] Lacey, M. T. The bilinear maximal function maps into L^p, $p > 2/3$, *Ann. of Math. (2)* **151**(1), 35–57, (2000).

[11] Lacey, M. T. and Thiele, C. M. L^p bounds for the bilinear Hilbert transform, $2 < p < \infty$, *Ann. of Math.* **146**, 693–724, (1997).

[12] Lacey, M. T. and Thiele, C. M. On Calderón's conjecture, *Ann. of Math.* **149**, 475–496, (1999).

[13] Lacey, M. T. and Thiele, C. M. A proof of boundedness of the Carleson operator, *Math. Res. Lett.* **7**, 361–370, (2000).

[14] Muscalu, C., Tao, T., and Thiele, C. Multi-linear operators given by singular multipliers, *J. Amer. Math. Soc.* **15**, 469–496, (2002).

[15] Sjölin, P. and Soria, F. Remarks on a theorem by N. Yu. Antonov, *Studia Math.* **158**, 79–97, (2003).

Received May 2, 2006

Department of Mathematics, University of Missouri, Columbia, MO 65211
e-mail: bilykd@math.missouri.edu

Department of Mathematics, University of Missouri, Columbia, MO 65211
e-mail: loukas@math.missouri.edu