THE (WEAK-L^2) BOUNDEDNESS OF THE QUADRATIC CARLESON OPERATOR

Victor Lie

Abstract. We prove that the generalized Carleson operator with polynomial phase function of degree two is of weak type (2,2). For this, we introduce a new approach to the time-frequency analysis of the quadratic phase.

1 Introduction

The historical motivation for the subject of this paper is rooted in Luzin's conjecture (1913), which asserts that the Fourier series of a function $f \in L^2(\mathbb{T})$ converges pointwise to f Lebesgue-almost everywhere. In 1966, L. Carleson gave a positive answer to this conjecture in the celebrated paper [C]. His result can be essentially reformulated, via [S1], as follows:

(In the sequel, we will always omit principal value notation.)

Theorem 0. *If for $f \in C^1(\mathbb{T})$ we define the expression*

$$Cf(x) := \sup_{a>0} \left| \int_{\mathbb{T}} \frac{1}{y} e^{iay} f(x-y) dy \right|, \tag{1}$$

then C is of weak type $(2,2)$, i.e.

$$\|Cf\|_{L^{2,\infty}(\mathbb{T})} \leq A\|f\|_{L^2(\mathbb{T})}, \tag{2}$$

where here, by convention, $\mathbb{T} = [-1/2, 1/2]$ and $A > 0$ is an absolute constant.

In addition to Carleson's proof, we point out two more proofs of the above result: one due to Fefferman [F], using a very beautiful geometric combinatorial argument, and the other due to Lacey and Thiele [LT1], inspired by the subtle techniques they developed for proving the Calderón conjecture ([LT2] and [LT3]). Now, given the statement of Theorem 0, it is natural to hope that this result may be set in a broader context. Following this direction, Stein conjectured that the generalized Carleson operator defined by

$$C_d f(x) := \sup_{\deg(P)=d} \left| \int_{\mathbb{T}} \frac{1}{y} e^{iP(y)} f(x-y) dy \right| \tag{3}$$

(here $d \in \mathbb{Z}$, $d \geq 2$, P is a polynomial of degree d, and $f \in C^1(\mathbb{T})$) obeys the same bounds as C.

In [S2] he proved this conjecture, subject to the key restriction that the supremum in (3) be taken in the class of quadratic polynomials with no linear term. Further, using the TT^* method and a variant of van der Corput's lemma, Stein and Wainger [SW] extended this result for polynomials of any degree, but again without the first degree term.

Our aim in this paper is to provide a positive answer to this conjecture for the case $d = 2$. (We mention here that, using similar techniques to those in [LT1], M. Lacey published in [L2] a proof of this result, but as was revealed later by A. Ionescu, this was incorrect – for details see [L1].)

The main result of the article is given by

Theorem 1. *Let $1 \leq p < 2$; then the expression*

$$Tf(x) := C_2 f(x) = \sup_{a,b \in \mathbb{R}} \left| \int_{\mathbb{T}} e^{i\{ay+by^2\}} \frac{1}{y} f(x-y) dy \right| \tag{4}$$

satisfies

$$\|Tf\|_{L^p(\mathbb{T})} \lesssim_p \|f\|_{L^2(\mathbb{T})}.$$

Combining this result with the techniques developed by Stein in [S1], we easily deduce

COROLLARY 1. *T is of weak type $(2, 2)$.*

The proof of Theorem 1 is a combination of analytic and geometric facts; it relies on a new perspective of the time-frequency localization of the quadratic phase to which we adapt the techniques presented in [F].

One particular feature of this paper is that it presents for the first time a time-frequency proof of the boundedness of a maximal operator which is invariant under quadratic modulations.

Another novelty of this paper is that we show that one can prove the (quadratic) Carleson theorem using a single dyadic grid partition, on each axis defining the time-frequency plane. (For this, we will involve in our reasoning certain dilation factors of our tiles; see section 7 – forest decomposition algorithm.)

Finally, given the powerful geometric intuition developed in Fefferman's paper, and also the fact that many of the reasonings here rely on his work, we have chosen to present our paper maintaining the structure and some of the notation appearing in [F].

2   Preliminaries and Outline of the Proof

As our problem is of a time-frequency nature, it will be based on two steps:

(A) – a discretization procedure, in which we split our operator into "small pieces" that are well localized in both time and frequency;

(B) – a selection algorithm, which relies on finding (qualitative and quantitative) criteria depending on which we decide how to glue the above-mentioned pieces together to obtain a global estimate on our operator.

For task (A), we first need to study the symmetries of our operator. This is because these symmetries will determine the geometric properties of the time-frequency portrait of our operator, properties that will provide a significant indication of how to naturally decompose the operator "along its fibers."

We define the following classes of symmetries (since the symmetries 1 and 2 do not preserve the periodicity of the object on which they are acting, in what follows one should regard $L^2(\mathbb{T})$ as the space of functions which are L^2-integrable on any given unit interval):

1. Modulations:
$$M_a : L^2(\mathbb{T}) \longrightarrow L^2(\mathbb{T}) \ (a \in \mathbb{R}) \quad \text{by} \quad M_a f(x) := e^{iax} f(x);$$

2. Quadratic Modulations:
$$Q_b : L^2(\mathbb{T}) \longrightarrow L^2(\mathbb{T}) \ (b \in \mathbb{R}) \quad \text{by} \quad Q_b f(x) := e^{ibx^2} f(x);$$

3. Translations:
$$\tau_y : L^2(\mathbb{T}) \longrightarrow L^2(\mathbb{T}) \ (y \in \mathbb{R}) \quad \text{by} \quad \tau_y f(x) := f(x - y);$$

4. Dilations:
$$D_\lambda : L^2(\mathbb{T}) \longrightarrow L^2(\mathbb{T}) \ (\lambda \in \mathbb{N}) \quad \text{by} \quad D_\lambda f(x) := f(\lambda x).$$

The key observation is that we can recover the operators C and T from the action of these symmetries (particularly 1 and 2) on the Hilbert transform (strictly speaking, the kernel of the Hilbert transform should be $\cot \pi y$; for convenience, we work instead with $1/y$), defined by

$$H : \ L^2(\mathbb{T}) \longrightarrow L^2(\mathbb{T}) \qquad Hf(x) := \int_{\mathbb{T}} \frac{1}{y} f(x - y) dy.$$

Indeed, the periodic Carleson operator

$$Cf(x) = \sup_{a \in \mathbb{R}} \left| \int_{\mathbb{T}} \frac{1}{y} e^{iay} f(x - y) dy \right|$$

can be rewritten as

$$Cf(x) = \sup_{c \in \mathbb{R}} \left| M_c H M_c^* f(x) \right|, \tag{5}$$

while our periodic quadratic Carleson operator

$$Tf(x) = \sup_{a,b \in \mathbb{R}} \left| \int_{\mathbb{T}} e^{i\{ay + by^2\}} \frac{1}{y} f(x - y) dy \right|$$

can be rewritten as

$$Tf(x) = \sup_{b,c \in \mathbb{R}} \left| M_c Q_b H Q_b^* M_c^* f(x) \right|. \tag{6}$$

(Note that in the previous formulas the action of translations and dilations is hidden in the structure of the Hilbert transform, which is the unique – up to identity – L^2-bounded linear operator that commutes with both symmetries.)

These facts help us to conclude that C essentially obeys the relations

1. $C\tau_y = \tau_y C$
2. $CD_\lambda = D_\lambda C$
3. $CM_a = C$

while for the operator T, besides the analogous relations we have the extra condition

4. $TQ_b = T$.

(The relations 1-4 are literally true if we work in the setting of $\mathbb{R}$ rather than $\mathbb{T}$. Relations 1, 3, and 4 remain true in the torus case, while 2 serves as a useful heuristic inherited from the real case.)

We now analyze the effect of these symmetries on the time-frequency decomposition of our operator T. To help build up intuition, we will consider three cases of increasing complexity: the Hilbert transform H, the Carleson operator C, and finally the quadratic Carleson operator T.

As announced, we first look at the simplest object, namely the Hilbert transform; we begin by isolating the kernel and splitting it – taking advantage of the dilation symmetry of H – as follows:

$$\frac{1}{y} = \sum_{k \in \mathbb{N}} \psi_k(y)$$

where $\psi \in C_0^\infty$ is an odd function supported away from the origin and $\psi_k(y) = 2^k \psi(2^k y)$, $k \in \mathbb{N}$; consequently,

$$Hf(x) = \sum_{k \in \mathbb{N}} \int \psi_k(y) f(x - y) dy. \tag{7}$$

Now for each scale k we take the collection $\{I_{k,j}\}_j$ of all dyadic intervals in $[0, 1]$ of length 2^{-k}. Using the translation invariant property of H we write

$$Hf(x) = \sum_{k,j} H_{k,j} f(x) = \sum_{k,j} \left\{ \int \psi_k(y) f(x - y) dy \right\} \chi_{I_{k,j}}(x), \tag{8}$$

where χ_I is, as usual, the characteristic function of I.

Now each $H_{k,j}f$ has time support included in $I_{k,j}$ while on the frequency side it is "morally" supported near the origin, in an interval of length $|I_{k,j}|^{-1}$. Consequently, the time-frequency picture of H is as given in Figure 1.

The above story can be expressed more intuitively as follows: Observe that the translation symmetry acts on the j-direction, while the dilation symmetry acts on the k-direction. If we approximate the piece $H_{1,1}f$ by a smooth compactly supported function φ_0, then the time-frequency portrait of φ_0 is a square of area one located near the origin. Since Hf is, roughly speaking, just a sum of dilations and translations of φ_0, by basic properties of the Fourier transform we obtain Figure 1 as the time-frequency picture of Hf. From the figure we also note that the origin plays a special role in this decomposition.

We now consider the Carleson operator as described in (5). In this case we will have to deal with one more symmetry given by the modulation invariance property,
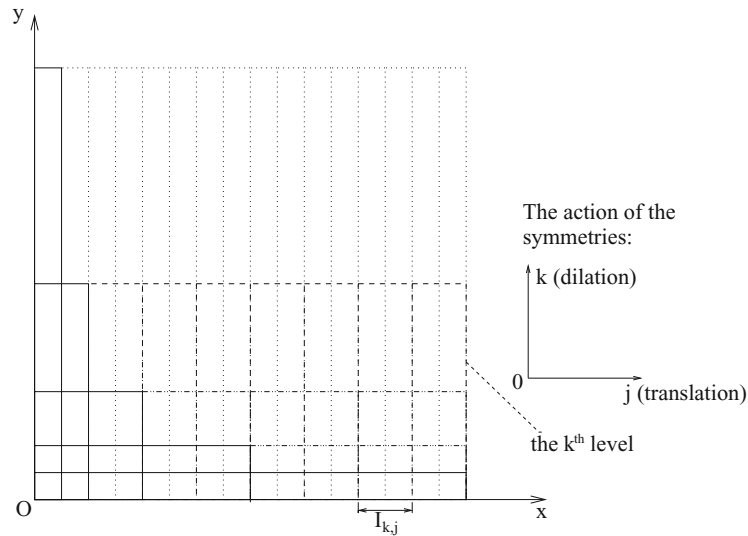
Figure 1: The time-frequency decomposition of the Hilbert transform

so we will try first to understand a simpler situation, namely how M_c acts on a smooth compactly supported function φ. As we may remark from Figure 2, in the time-frequency plane, M_c will translate the rectangle representing the localization of φ by c units in the frequency direction.

Now, as in (7), we have that

$$M_c H M_c^* f(x) = \sum_{k,j} M_c H_{k,j} M_c^* f(x) = \sum_{k,j} \left\{ \int (M_c \psi_k)(y) f(x-y) dy \right\} \chi_{I_{k,j}}(x),$$

and so combining this with the previous observation, we deduce that the time-frequency picture of $M_c H M_c^*$ will be nothing more than a frequency-translation by c units of the corresponding picture of H.

Exploiting this fact in the form of (5), we conclude that the time-frequency localization of C is as presented in Figure 3.

Note that, unlike the Hilbert transform case, there is no preferential point in the splitting of C. Also, this picture suggests that C may be written (after a linearization procedure) as

$$Cf = \sum_P C_P f$$

with each C_P a linear operator localized in a certain (Heisenberg) rectangle P (Figure 3). This is a key observation used explicitly in both [F] and [LT1].

We finalize part (A) of our program with the analysis of our operator T. As before, we begin by isolating the extra symmetry, Q_b, that adds to those appearing in the previous cases. We will approach the study of the time-frequency representation of this quadratic symmetry from two perspectives: (q1) a *restrictive* one and (q2) a *relational* one.
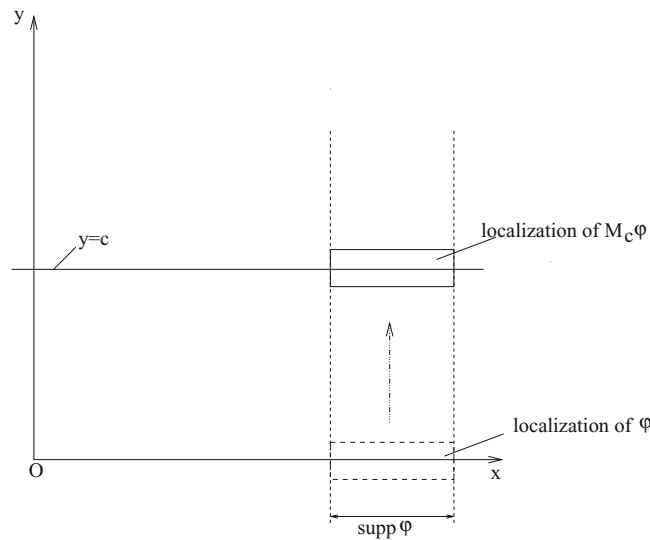
Figure 2: The time-frequency localization of $M_c\varphi$

($q1$) The *restrictive* perspective relies on the following basic approach: given an object (Schwartz function on $\mathbb{R}$) – call it h – describe (in terms of a picture) the space and frequency regions (also called the "moral" support for h and $\hat{h}$, respectively) where "most" of the information carried by the function is located. As one may notice this is an *absolute* way of quantifying the object since it relies on studying the distribution of the L^∞-norm of h (and respectively $\hat{h}$) and not on how h may relate (interact) with some other objects (functions) living in a given environment.

Reasoning in this spirit (for φ defined as above), we have that the "moral" support of $Q_b\varphi$ is given by the support of φ (here we rely on the equality $\operatorname{supp} Q_b\varphi = \operatorname{supp}\varphi$) while, with the notation of Figure 4, we have that the "moral support" of $\widehat{Q_b\varphi}$ is identified with the frequency-interval U. (Note that Q_b is a multiplication operator and hence preserves the time localization of the object on which it acts – this being the main reason for which we will split our operator T into pieces that are compactly supported in time.) At this point, we observe that we lose the (global) Heisenberg principle, this being one of the main difficulties that stood against solving this conjecture. (Throughout this paper, we use the term "Heisenberg principle" to refer to the optimal Heisenberg localization, *i.e.* the product of the sizes of the time and frequency moral supports are comparable with 1.)

One may improve this time-frequency portrait if one further decomposes φ into pieces which are better adapted to the oscillation of the quadratic factor imposed by Q_b; more exactly, writing

$$\varphi = \sum_j \varphi_j$$

with each $\varphi_j \in C_0^\infty$ and $|\operatorname{supp}\varphi_j| \approx \min((2b)^{-1/2}, |\operatorname{supp}\varphi|)$ we squeeze the previous localization to a sequence of area-one blocks concentrated near the diagonal of the
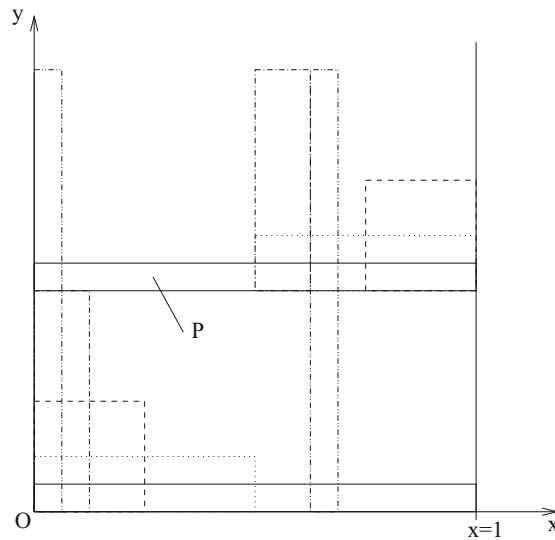
Figure 3: The time-frequency decomposition of the Carleson operator

initial "big" rectangle. Now, even though on each such block – reflecting the time-frequency portrait of a φ_j – we recover the Heisenberg principle, the parallelogram formed by their union still offers a poor (global) localization of $Q_b\varphi$. Using this viewpoint, one cannot do better.

($q2$) The *relational* (relative) perspective, as the name suggests, focuses on determining a contextual representation of our object depending on how it interacts with other objects "living" in a given environment.

More exactly, in our case the environment is formed by objects like $M_c\varphi$, $Q_b\varphi$ and the interaction is given by the scalar product in $L^2(\mathbb{T})$. (Eventually, we will increase the complexity of these objects by composing the symmetries.)

Now taking, for example, the interaction

$$\big|\langle M_c\varphi, M_{c'}\varphi\rangle\big| \tag{9}$$

(here $c, c' \in \mathbb{R}$) we see that, applying the (non-)stationary phase principle, (9) is controlled by a quantity depending on the ratio of $|V|^{-1}$ and the distance between the lines $y = c$ and $y = c'$ (where φ is a smooth function which is L^2-normalized and adapted to V, with $\operatorname{supp}\varphi = V$). By varying c', this quantity suggests that (on the frequency side) the information carried by $M_c\varphi$ should be localized "near the line" $y = c$, and that this information is roughly constant on intervals of length $|V|^{-1}$. As a consequence we may interpret the relative time-frequency localization of $M_c\varphi$ as being given by the region (rectangle) centered near the line $y = c'$ of width $|V|^{-1}$ (measured on frequency axis) and with space support in the interval V. (It is not surprising, in this case, that the relative time-frequency picture coincides with the restrictive one described above, given how the Fourier transform acts on modulation, translation, and dilation.)
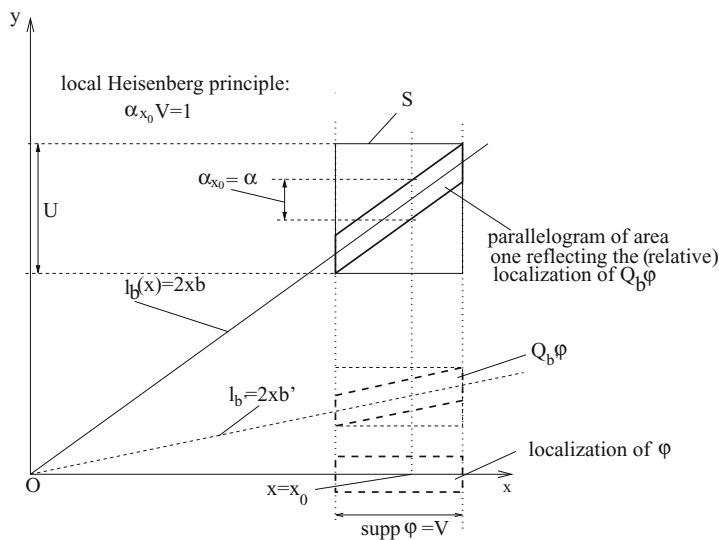
Figure 4: The (relative) time-frequency localization of $Q_b\varphi$

By analogy with the above description, we will now treat the following interaction:

$$\left|\langle Q_b\varphi, Q_{b'}\varphi\rangle\right|. \tag{10}$$

As before, applying the (non-)stationary phase principle we remark that (10) is controlled by a quantity depending on the ratio of $|V|^{-1}$ and the distance between the lines $y = 2bx$ and $y = 2b'x$ obtained by differentiating the polynomial phase. (Here the appropriate notion of distance is given by $\sup_{x\in V}|2bx - 2b'x|$ rather than $\inf_{x\in V}|2bx - 2b'x|$.) This fact invites us to think of the relative time-frequency localization of $Q_b\varphi$ as being given by the region (parallelogram) centered near the line $y = 2bx$ of width $|V|^{-1}$ (measured on the frequency axis) and with space support in the interval V. Indeed, this perspective will prove to give an accurate geometric representation of the relations among our objects. (See section 5.)

As a consequence, this should be the "true" time-frequency "story" reflected in pictures (see Figure 4); it is of a *relative* nature since it tells us about the interaction of $Q_b\varphi$ with an exterior object and not about $Q_b\varphi$ itself. (Remark – see Figure 4 – that using this approach, we recover, on each fiber, a local Heisenberg principle. For the remainder of the paper, "time-frequency portrait" will refer to the *relative* representation described in ($q2$).) This time-frequency interpretation can be regarded as a way of drawing pictures in which besides the magnitude we also encode the oscillation of our function. (The point is that while $|\widehat{Q_b\varphi}|$ is big on the whole interval U, when tested against same-structure functions, as in the expression $\langle \widehat{Q_b\varphi}, \widehat{Q_{b'}\varphi}\rangle$, the oscillations of $\widehat{Q_b\varphi}$ come into play, canceling out most of the oscillations of $\widehat{Q_{b'}\varphi}$ up to the level given by the interaction of the corresponding parallelograms. For further study of this interaction behavior as well as for some other local properties, see section 5.)

The moral of this story is that while M_c translates the time-frequency picture up and down, the operator Q_b realizes a shearing of the same picture.

The idea presented above will be essential in the proof of Theorem 1, and might be quite productive in a series of other problems involving quadratic time-frequency analysis.

Now, coming back to our decomposition, if we let M_c interfere with Q_b we obtain the "elementary cell" of our operator modeled in $M_cQ_b\varphi$; from the previous discussion, this will be considered as being localized in a parallelogram of area one living near the line $l(x) = c + 2bx$ and with the same time localization as before (Figure 5).
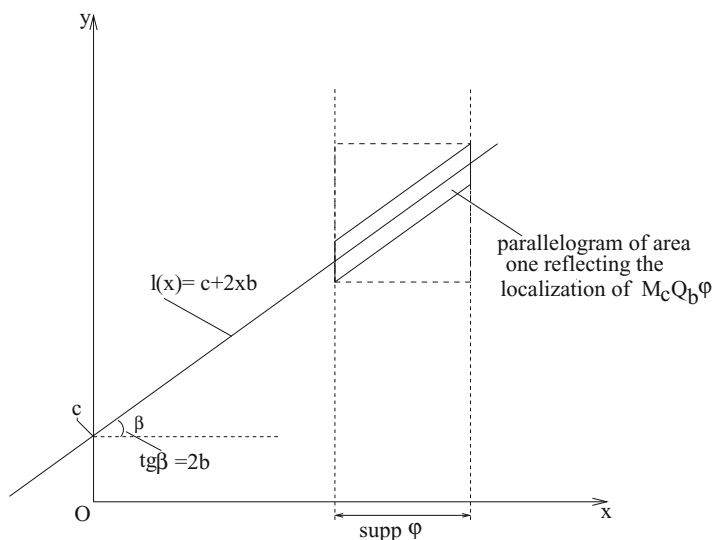
Figure 5: The (relative) time-frequency localization of $M_cQ_b\varphi$

Once we have gained this intuition, given the form (6), it is natural to split T into pieces that will be localized in the same (relative) region as our "elementary cells" $M_cQ_b\varphi$ mentioned above. Consequently, we will divide our time-frequency plane in parallelograms of area one as reflected in Figure 6.

The exact procedure will be described in section 4, and will have as a consequence

$$T = \sum_{P \in \mathcal{P}} T_P \,,$$

with each piece T_P having the time-frequency picture represented by the tile $P = [\alpha, \omega, I]$ (see section 3 for notation).

This way we have highlighted the dual nature of this problem: an analytic formulation (providing L^2 bounds for a certain object) visualized in terms of geometric interactions of some families of parallelograms (tiles). Consequently, there will be no surprise that in the second stage of our program that we now initiate – the selection algorithm – the geometric point of view in quantifying different interactions among the "small pieces" T_P will play the essential role.

Figure 6: The time-frequency decomposition of the quadratic Carleson operator

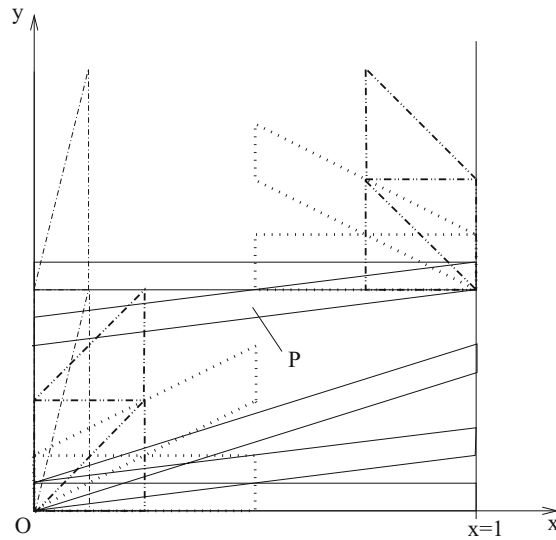Indeed, we start by defining a measurable map that assigns to each point $x \in [0,1]$ a line $l_x \in \mathcal{L}$ in $\mathbb{R}^2$; then we can regard $T_P f$ as assigning the values

$$x \overset{T_P f}{\longmapsto} 0 \quad \text{if } x \notin I \text{ or } l_x \notin P, \tag{11}$$

$$x \overset{T_P f}{\longmapsto} (\text{a quantity "oscillating along } l_x\text{"}) \quad \text{if } x \in I \text{ and } l_x \in P. \tag{12}$$

(See section 3 for notation.)

This way $T_P f$ (and similarly $T_P{}^* f$) encodes two different types of information: (11) forces us to consider the density of the "flow" $\{l_x\}_{x \in I}$ through the tile P (this concept will be made precise in section 5 – see (24) – and will be called the "density factor" of P), while (12) implies that on Fourier side, the information given by $\widehat{T_P f}$ is localized near the central line of P denoted l_P. The interplay between these two features of T_P (or $T_P{}^*$) will be discussed in detail in section 5, and it is the key fact in providing good bounds for the expression

$$\left\| \sum_{P \in \mathcal{P}} T_P{}^* f \right\|_2^2 = \sum_{P, P' \in \mathcal{P}} \langle T_P{}^* f, T_{P'}{}^* f \rangle, \tag{13}$$

where here $\mathcal{P}$ is a certain finite collection of tiles and f some fixed element in $L^2(\mathbb{T})$. In dealing with this problem, we first need to understand the quantity

$$\left| \langle T_P{}^* f, T_{P'}{}^* f \rangle \right|. \tag{14}$$

To obtain some intuition, we explain first the two possible extreme cases:

– When $P = P'$ (*i.e.* the diagonal term) the relevant point of view is given by (11); this is natural since $T_P{}^* f$ and $T_{P'}{}^* f$ oscillate in the same region of the time frequency-plane, making the information offered by (12) useless. Consequently, the norm $\|T_P\|_2$ will measure the density of P (see (23) and (24)).

– When P and P' are far apart from one another, (14) is small either due to the time localization of $T_P{}^*f$ or due to the relation (12) that comes into play by forcing $\widehat{T_P{}^*f}$ and $\widehat{T_{P'}{}^*f}$ to have different "moral supports".

Consequently, via (11) and (12) (which also determines the time-frequency localization of T_Pf and $T_P{}^*f$) we expect the following principle to be true:

The magnitude of (14) is

- big – *when P, P' have large overlaps and high density*;
- small – *when P, P' have small overlaps (are disjoint) or*   (15)
  low density.

Now this principle simultaneously offers and demands a lot of information:

I. On the one hand, it suggests that to obtain good control of (13) we may need to split the family of tiles $\mathcal{P}$ into sub-collections $\mathcal{P}_j$ with each $\mathcal{P}_j$ having uniform characteristics (all the tiles inside it must have comparable densities and any interaction between two of them must have the same degree of overlapping), estimate separately each

$$\|T^{\mathcal{P}_j*}f\|_2 := \Big\| \sum_{P \in \mathcal{P}_j} T_P{}^*f \Big\|_2$$

with bounds depending on the previously mentioned characteristics of $\mathcal{P}_j$, and then sum them up for obtaining the desired global bound.

II. On the other hand, it requires a clear formulation of the concepts:

a. the density of a tile

b. the degree of the overlapping between two tiles.

Point II will be the object of our study in section 5. While II.a will be straightforward, for II.b we will introduce two ways of measuring the corresponding concept: a *qualitative* one, by defining an "almost" order relation between tiles, "$\leq$", (Definition 3) and a *quantitative* one, the actual measurement of how much two tiles P_1, P_2 overlap, that can be recovered from the geometric factor of the pair (P_1, P_2) (Definition 1).

Now, guided by the observation made in I, our proof will be based on two propositions corresponding to the two main (geometric) possibilities appearing in the study of a family of tiles (having uniform density): Proposition 1 will treat the case where our family consists of "disjoint" (*i.e.* not comparable under "$\leq$") tiles, while Proposition 2 will deal with a family, called a "forest", that can be organized into a controlled number of clustered sets of tiles (*i.e.* trees).

With this done we will proceed (roughly) as follows:

We will decompose $\mathbb{P}$ into $\bigcup_{n=0}^{\infty} \mathcal{P}_n$ with

$$\mathcal{P}_n = \big\{ P \in \mathbb{P} \mid 2^{-n-1} < \text{the density factor of } P \leq 2^{-n} \big\}.$$

Using a combinatorial argument, we will further prove that $\mathcal{P}_n$ may be written as a disjoint union of at most n sets, $\bigcup_{j=1}^{n} \mathcal{P}_{nj}$, such that

$$\mathcal{P}_{nj} = \mathcal{A}_{nj} \cup \mathcal{B}_{nj}$$

where, for each j, $\mathcal{A}_{nj}$ is a family of at most n disjoint tiles and $\mathcal{B}_{nj}$ is a forest. Now, denoting

$$T^{\mathcal{P}_{nj}} := \sum_{P \in \mathcal{P}_{nj}} T_P \,,$$

Propositions 1 and 2 will imply that

$$\|T^{\mathcal{P}_{nj}}\|_2 \lesssim 2^{-n\eta}$$

for some absolute constant $\eta > 0$, from which we conclude that

$$\|T\|_2 = \left\| \sum_{n=0}^{\infty} T^{\mathcal{P}_n} \right\|_2 \leq \sum_{n=0}^{\infty} \sum_{j=1}^{n} \|T^{\mathcal{P}_{nj}}\|_2 \lesssim 1 \,.$$

(Throughout this paper we will denote by $\|T\|_2$ the operator norm of T acting from L^2 to L^2.)

3   Notation

Take the canonical dyadic grid in $[0,1] = \mathbb{T}$ and in $\mathbb{R}$. (For convenience, from now on we may choose to identify $\mathbb{T}$ with any unit interval, not necessarily $[-1/2, 1/2]$.) A tile $P = [\alpha, \omega, I]$ consists of dyadic (half open) intervals $\alpha, \omega \subset \mathbb{R}$ and $I \subset [0,1]$ with the property that $|\alpha| = |\omega| = |I|^{-1}$ (here $|I| = m(I)$ where $dm = dx$ stands for the Lebesgue measure on $[0,1]$). The collection of all tiles P will be denoted by $\mathbb{P}$. For simplicity of notation, P will encode two meanings (depending on the context): a triple of intervals as defined above or the parallelogram formed by these intervals in the time-frequency plane.
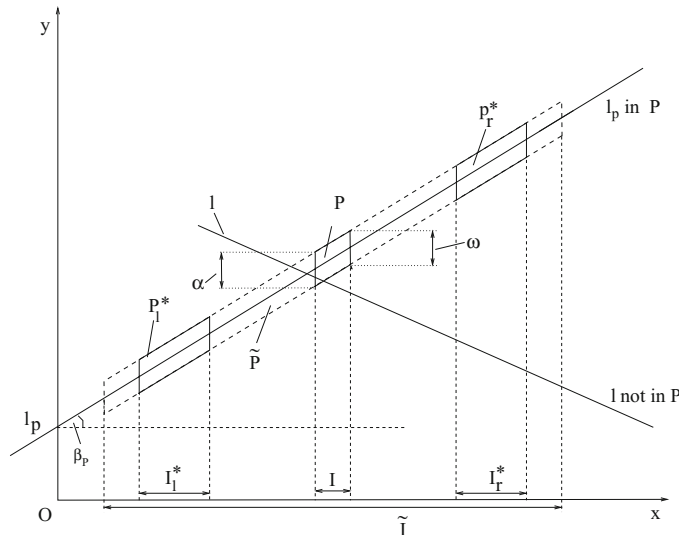

Figure 7: Notation

If I is any (dyadic) interval we denote by $c(I)$ the center of I. Let I_r be the "right brother" of I, with $c(I_r) = c(I) + |I|$ and $|I_r| = |I|$; similarly, the "left brother" of

I will be denoted I_l with $c(I_l) = c(I) - |I|$ and $|I_l| = |I|$. If $a > 0$ is some real number, by aI we mean the interval with the same center $c(I)$ and with length $|aI| = a|I|$. Now for $P = [\alpha, \omega, I] \in \mathbb{P}$, we set $aP := [a\alpha, a\omega, I]$. Also, if $\mathcal{P} \subseteq \mathbb{P}$ then by convention $a\mathcal{P} := \{aP \mid P \in \mathcal{P}\}$.

Set $\mathcal{L} := \{$all lines in the plane not parallel with the y-axis$\}$. Then, for each $P = [\alpha, \omega, I] \in \mathbb{P}$ and $l \in \mathcal{L}$, we write "$l \in P$" iff l intersects both edges of P which are parallel with the y-axis. Also, for any tile P as before, we will associate the "central line" l_P - the unique line $l \in \mathcal{L}$ that passes through the midpoints of the vertical edges (this line will give the "angle of P", denoted β_P, and given by the formula $\tan(\beta_P) = \frac{c(\omega) - c(\alpha)}{|I|}$).

Now, for $\beta \in \arctan(\mathbb{Z})$, define
$$\mathcal{P}(k, \beta) = \left\{ P = [\alpha, \omega, I] \in \mathbb{P} \,\middle|\, |I| = 2^{-k} \,\&\, \beta_P = \beta \right\}.$$
Then this collection of disjoint (similar) parallelograms (tiles) defines a partition of the band $\mathbb{T} \times \mathbb{R}$. Fixing $P = [\alpha, \omega, I] \in \mathcal{P}(k, \beta)$, denote the "upper brother" of P by $P_u = [\alpha_r, \omega_r, I] \in \mathcal{P}(k, \beta)$; similarly, the lower brother of P will be $P_l = [\alpha_l, \omega_l, I] \in \mathcal{P}(k, \beta)$.

For any dyadic interval $I \subseteq [0, 1]$ define the (non-dyadic) intervals
$$I_r^* = \left[c(I) + \tfrac{7}{2}|I|, c(I) + \tfrac{11}{2}|I| \right) \quad \& \quad I_l^* = \left[c(I) - \tfrac{11}{2}|I|, c(I) - \tfrac{7}{2}|I| \right)$$
$I^* = I_r^* \cup I_l^*$ and $\tilde{I} = 13I$.

Similarly, for $P = [\alpha, \omega, I] \in \mathbb{P}$ we adopt the following notation:

– P_r^* for the tile (parallelogram of area two) with time interval I_r^* and the same central line l_P as P
– P_l^* for the tile (parallelogram of area two) with time interval I_l^* and the same central line l_P as P.

The same procedure applies to P^* and $\tilde{P}$ (see Figure 7).

Throughout the paper, for $f \in L^2(\mathbb{T})$, we denote by
$$Mf(x) = \sup_{x \in I} \frac{1}{|I|} \int_I |f|$$
the Hardy–Littlewood maximal function associated to f.

If $\{I_j\}$ is a collection of pairwise disjoint intervals in $[0, 1]$ and $\{E_j\}$ a collection of sets such that for a fixed $\delta \in (0, 1)$
$$E_j \subset I_j \quad \& \quad \frac{|E_j|}{|I_j|} \leq \delta \quad \forall j \in \mathbb{N}, \tag{16}$$
then we denote
$$M_\delta f(x) := \begin{cases} \sup_{I \supset I_j} \frac{1}{|I|} \int_I |f|, & \text{if } x \in E_j, \\ 0, & \text{if } x \notin E_j. \end{cases} \tag{17}$$
Remark that $\forall r > 1$ we have
$$\|M_\delta f\|_r^r \lesssim \delta \|f\|_r^r. \tag{18}$$
For $A, B > 0$ we say $A \lesssim B$ ($\gtrsim$) if there exist an absolute constant $C > 0$ such that $A < CB$ ($>$); if the constant C depends on some quantity $\delta > 0$ then we may choose to stress this fact by writing $A \lesssim_\delta B$.

If $C^{-1}A < B < CB$ for C some small (positive) absolute constant then we write $A \approx B$. For $x \in \mathbb{R}$ we set $\lceil x \rceil := \frac{1}{1 + |x|}$.

The exponents η and ϵ may change throughout the paper.

4   Discretization

Our aim is to "properly" decompose the operator
$$Tf(x) = \sup_{b,c\in\mathbb{R}} \left| M_c Q_b H Q_b^* M_c^* f(x) \right| = \sup_{l\in\mathcal{L}} \left| T_l f(x) \right|$$
where
$$T_l f(x) = \int_{\mathbb{T}} \frac{1}{y} e^{i(l(x)y - by^2)} f(x-y) dy$$
with $l \in \mathcal{L}$ given by $l(x) = c + 2bx$.

Now linearizing T we can write
$$Tf(x) = T_{l_x} f(x) = \int_{\mathbb{T}} \frac{1}{y} e^{i(l_x(x)y - b(x)y^2)} f(x-y) dy$$
where by l_x we understand a line in $\mathcal{L}$ given by $l_x(z) = c(x) + 2zb(x)$ where $c(\cdot)$ and $b(\cdot)$ are certain measurable functions. (This procedure is often referred to as the Kolmogorov linearization method.)

We start our decomposition by choosing ψ to be an odd C^∞ function such that $\operatorname{supp}\psi \subseteq \{y \in \mathbb{R} \mid 2 < |y| < 8\}$ and having the property
$$\frac{1}{y} = \sum_{k\geq 0} \psi_k(y) \quad \forall 0 < |y| < 1\,,$$
where by definition $\psi_k(y) := 2^k \psi(2^k y)$ (with $k \in \mathbb{N}$). As a consequence, we deduce that
$$Tf(x) = \sum_{k\geq 0} T_k f(x) := \sum_{k\geq 0} \int_{\mathbb{T}} e^{i\{l_x(x)y - b(x)y^2\}} \psi_k(y) f(x-y) dy\,.$$

Now for each $P = [\alpha, \omega, I] \in \mathbb{P}$ let $E(P) := \{x \in I \mid l_x \in P\}$. Also, if $|I| = 2^{-k}$ ($k \geq 0$) we define the operators T_P on $L^2(\mathbb{T})$ by
$$T_P f(x) = \left\{ \int_{\mathbb{T}} e^{i\{l_x(x)y - b(x)y^2\}} \psi_k(y) f(x-y) dy \right\} \chi_{E(P)}(x)\,.$$

Clearly, as P runs through $\mathbb{P}_k := \{P = [\alpha, \omega, I] \in \mathbb{P} \mid |I| = 2^{-k}\}$, for fixed k, the $\{E(P)\}$ form a partition of $[0,1]$, and so
$$T_k f(x) = \sum_{P\in\mathbb{P}_k} T_P f(x)\,.$$

Consequently, we have
$$Tf(x) = \sum_{k\geq 0} T_k f(x) = \sum_{P\in\mathbb{P}} T_P f(x)\,.$$

This ends our decomposition.

We finish this section with some remarks.

1. Because we want better separation properties between the support of $T_P f$ and that of $T_P^* f$ (for fixed P and f), by further splitting ψ as
$$\psi(y) = \sum_{j=1}^{13} \psi^j(y)$$
(with each ψ^j an odd, smooth function with $\operatorname{supp}\psi^j \subset \left\{1 + \frac{j}{2} < |y| < 2 + \frac{j}{2}\right\}$) we may assume, relabeling for example ψ^6 with ψ, that
$$\operatorname{supp}\psi \subseteq \left\{y \in \mathbb{R} \mid 4 < |y| < 5\right\}\,.$$

Consequently, for a tile $P = [\alpha, \omega, I]$, the associated operator has the properties
$$\operatorname{supp} T_P \subseteq I \quad \& \quad \operatorname{supp} T_P^* \subseteq \big\{ x \mid 3|I| \leq \operatorname{dist}(x, I) \leq 5|I| \big\} = I^*.$$
where here T_P^* denotes, as usual, the adjoint of T_P.

2. In what follows (splitting $\mathbb{P} = \bigcup_{j=0}^{9} \bigcup_{k \geq 0} \mathbb{P}_{10k+j}$) we can suppose that our collection $\mathbb{P}$ is sparse enough; namely, if $P_j = [\alpha_j, \omega_j, I_j] \in \mathbb{P}$ with $j \in \{1, 2\}$ such that $|I_1| \neq |I_2|$ then $|I_1| \leq 2^{-10}|I_2|$ or $|I_2| \leq 2^{-10}|I_1|$.

5 Quantifying the Interactions Between Tiles

Our aim in this section is to isolate the appropriate quantities that arise in the behavior of the expression
$$\big| \langle T_{P_1}^* f, T_{P_2}^* g \rangle \big| \tag{19}$$
and further to show how they control this interaction.

We begin our study by presenting a summary of the main properties shared by the operator(s) involved in our considerations.

5.1 Properties of T_P and T_P^*.

For $P = [\alpha, \omega, I] \in \mathbb{P}$ with $|I| = 2^{-k}$, $k \in \mathbb{N}$, we have
$$\begin{aligned}
T_P f(x) &= \left\{ \int_{\mathbb{T}} e^{i(l_x(x)y - b(x)y^2)} \psi_k(y) f(x - y) dy \right\} \chi_{E(P)}(x) , \\
T_P^* f(x) &= -\left\{ \int_{\mathbb{T}} e^{i(l_{x-y}(x-y)y + b(x-y)y^2)} \psi_k(y) \big(\chi_{E(P)} f \big)(x - y) dy \right\} .
\end{aligned} \tag{20}$$

Notice that based on the previous interpretation of the symmetry Q_b (see section 2), we may conclude

- the time-frequency localization of T_P is "morally" given by the tile P;
- the time-frequency localization of T_P^* is "morally" given by the (bi)tile P^*.

$$\tag{21}$$

Also, we have the pointwise estimate
$$\big| T_P f(x) \big| \lesssim \frac{\int_{I^*} |f(y)| dy}{|I^*|} \chi_{E(P)}(x) \tag{22}$$
and the norm-estimate
$$\|T_P\|_2 \approx \left(\frac{|E(P)|}{|I|} \right)^{1/2} . \tag{23}$$

5.2 Factors associated to a tile.

Now, once we have understood what the main features of T_P and T_P^* are, we will relate them to concepts regarding the associated tile P. Indeed, taking into account relations (22) and (23), and respectively (21), for a tile $P = [\alpha, \omega, I]$ we are naturally led to the following two quantities:

(a) an *absolute* one (which may be regarded as a self-interaction) that measures how many lines from $\{l(x)\}_{x \in I}$ pass through P relative to the length of I; more exactly, we define the *density (analytic) factor of P* to be the expression
$$A_0(P) := \frac{|E(P)|}{|I|} . \tag{24}$$

Notice from (23) that $A_0(P)$ determines the L^2 norm of T_P. Consequently, we expect this quantity to play an important role in organizing and estimating the family $\{T_P\}_{P\in\mathcal{P}}$.

(b) a *relative* one (interaction of P with something exterior to it) which is of geometric type: let be $l \in \mathcal{L}$ a line and $P \in \mathbb{P}$ a tile as in Figure 8.
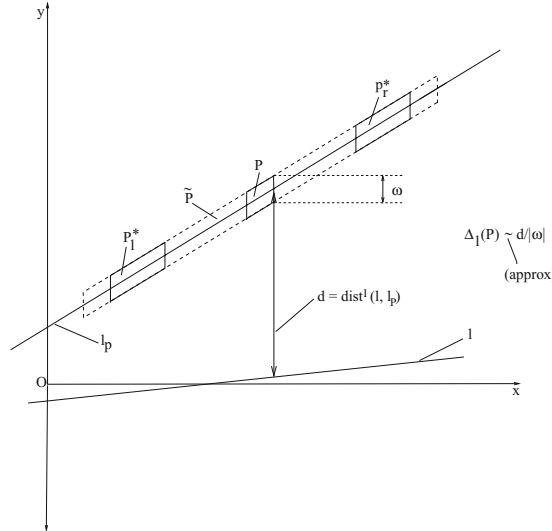
Figure 8: The geometric factor of P with respect to l

For $l_1, l_2 \in \mathcal{L}$ we introduce the following notation:

$$\mathrm{dist}_{x_0}(l_1, l_2) = \left| l_1(x_0) - l_2(x_0) \right| \quad \& \quad \mathrm{dist}^A(l_1, l_2) = \sup_{x \in A} \left\{ \mathrm{dist}_x(l_1, l_2) \right\}.$$

Then we define the *geometric factor of P with respect to l* to be the term

$$\left\lceil \Delta_l(P) \right\rceil,$$

where

$$\Delta_l(P) := \frac{\inf_{l_1 \in P}\{\mathrm{dist}^I(l, l_1)\}}{|\omega|}. \tag{25}$$

5.3 The resulting estimates. We now make the final step by observing how the above quantities relate in controlling the interaction in (19).

Given the heuristic (21) and the form of (19), we need to quantify the relative position of P_1^* with respect to P_2^*. To this end, we will need to adapt expression (25) to our context. (In the following we consider only the nontrivial case $I_{P_1}^* \cap I_{P_2}^* \neq \emptyset$.)

DEFINITION 1. *Given two tiles P_1 and P_2 (suppose that $|I_1| \geq |I_2|$), we define the geometric factor of the pair (P_1, P_2) by*

$$\left\lceil \Delta(P_1, P_2) \right\rceil,$$

where

$$\Delta(P_1, P_2)\, (= \Delta_{1,2}) := \frac{\displaystyle\inf_{\substack{l_1 \in P_1 \\ l_2 \in P_2}} \mathrm{dist}^{I_2}(l_1, l_2)}{|\omega_2|}.$$

With this notation, remark that we have

$$\lceil \Delta_{1,2} \rceil \approx \max \left\{ \lceil \Delta_{l_{P_1}}(P_2) \rceil, \lceil \Delta_{l_{P_2}}(P_1) \rceil \right\}.$$

We will also need to define the $(\epsilon_0\text{-})$*critical intersection interval* $I_{1,2}$ of the pair (P_1, P_2) as

$$I_{1,2} = [x_{1,2}^i - \gamma_{1,2}, x_{1,2}^i + \gamma_{1,2}] \cap I_2^* \cap I_1^*$$

(see Figure 9). Here $(x_{1,2}^i, y_{1,2}^i) := l_{P_1} \cap l_{P_2}$ (if l_{P_1} and l_{P_2} are parallel we set $x_{1,2}^i = \infty$), and $\gamma_{1,2}$ is chosen to obey the relation

$$\frac{\gamma_{1,2}}{\min(|I_1|, |I_2|)} = \lceil \Delta_{1,2} \rceil^{\frac{1}{2} - \epsilon_0} \tag{26}$$

for ϵ_0 some small fixed positive number.
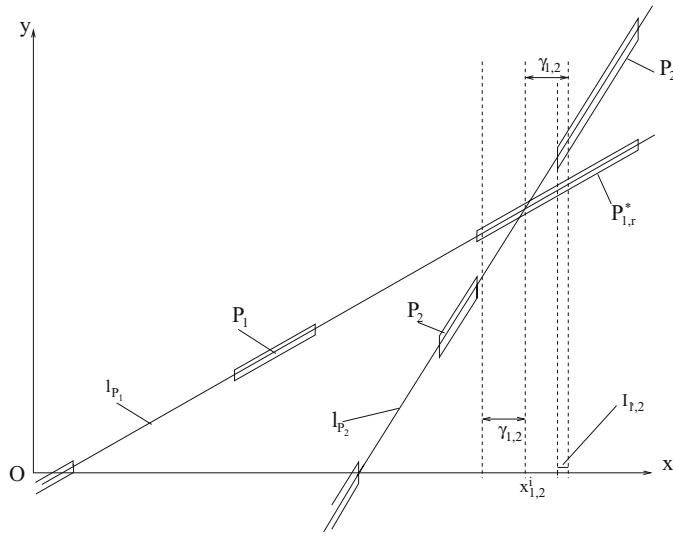
Figure 9: The interaction of two tiles

With these preparations done, we have the following result:

LEMMA 0. *Let $P_1, P_2 \in \mathbb{P}$; then we have*

$$\left| \int \tilde{\chi}_{I_{1,2}^c} T_{P_1}^* f \overline{T_{P_2}^* g} \right| \lesssim_{n,\epsilon_0} \lceil \Delta(P_1, P_2) \rceil^n \frac{\int_{E(P_1)} |f| \int_{E(P_2)} |g|}{\max(|I_1|, |I_2|)} \quad \forall n \in \mathbb{N} \tag{27}$$

$$\left| \int_{I_{1,2}} T_{P_1}^* f \overline{T_{P_2}^* g} \right| \lesssim \lceil \Delta(P_1, P_2) \rceil^{\frac{1}{2} - \epsilon_0} \frac{\int_{E(P_1)} |f| \int_{E(P_2)} |g|}{\max(|I_1|, |I_2|)} \tag{28}$$

where $\tilde{\chi}_{I_{1,2}^c}$ is a smooth variant of the corresponding cut-off.

Applying same methods for the limiting case $\epsilon_0 = 0$, we obtain

$$\|T_{P_1} T_{P_2}^*\|_2^2 \lesssim \min \left\{ \frac{|I_2|}{|I_1|}, \frac{|I_1|}{|I_2|} \right\} \lceil \Delta(P_1, P_2) \rceil A_0(P_1) A_0(P_2). \tag{29}$$

The proof of Lemma 0 relies on the (non-)stationary phase principle and is left to the reader.

6   The Preparation – Main Ingredients

As the title suggests, the role of this section is to present the important concepts and results on which the proof of our theorem relies.

We start on our way by introducing some quantitative and qualitative notions that will help us later to organize our family of tiles.

The first step is to define a quantity that inherits relevant features from both the analytic and geometric factors.

DEFINITION 2.   *For $P = [\alpha, \omega, I] \in \mathbb{P}$ we define the* mass *of P as being*

$$A(P) := \sup_{\substack{P' = [\alpha', \omega', I'] \in \mathbb{P} \\ I \subseteq I'}} \frac{|E(P')|}{|I'|} \left\lceil \Delta(2P, 2P') \right\rceil^N \tag{30}$$

where N is a fixed large natural number.

Next, we introduce a qualitative concept that characterizes the overlapping relation between tiles.

DEFINITION 3.   *Let $P_j = [\alpha_j, \omega_j, I_j] \in \mathbb{P}$ with $j \in \{1, 2\}$. We say that*
- *$P_1 \leq P_2$ iff $I_1 \subseteq I_2$ and $\exists l \in P_2$ s.t. $l \in P_1$,*
- *$P_1 \trianglelefteq P_2$ iff $I_1 \subseteq I_2$ and $\forall l \in P_2 \Rightarrow \in P_1$.*

OBSERVATION 1.   i. Remark that $\leq$ is not transitive while $\trianglelefteq$ is. However, $\leq$ is not so far from being a (partial) order relation; this may be encoded in the fact that if $P_1 \leq P_2$ then $2P_1 \trianglelefteq 2P_2$.

ii. Notice that the above definition can be meaningfully extended (in the obvious manner) to any dilated tiles, *i.e.* it makes sense to speak about $a_1 P_1 \leq a_2 P_2$ and respectively $a_1 P_1 \trianglelefteq a_2 P_2$ (here $a_1, a_2 > 0$); in addition, we say that $a_1 P_1 \lneq a_2 P_2$ iff $a_1 P_1 \leq a_2 P_2$ and $|I_1| < |I_2|$.

iii. There is a nice connection between the qualitative and quantitative concepts that measure the overlapping of the tiles P_1(or P_1^*) and P_2(or P_2^*): if $I_1 \subseteq I_2$ then $\lceil \Delta(P_1, P_2) \rceil = 1$ ($\Leftrightarrow \Delta(P_1, P_2) = 0$) $\Leftrightarrow \{aP_1 \leq P_2, \forall a > 1\}$.

OBSERVATION 2.   Notice that the notion of mass of a tile P is dependent on the environment. This definition offers many advantages, two of which we will mention here:

- the *monotonicity property (mp)*: if $P \leq P'$ (or $2P \trianglelefteq 2P'$) then $A(P) \geq A(P')$;
- the *smoothness property (sp)*: if P and P' are two tiles such that $I_P \approx I_{P'}$ (*i.e.* $2^{-a} I_P \subseteq I_{P'} \subseteq 2^a I_P$ for a some small positive integer) and $\lceil \Delta(P, P') \rceil \approx 1$ then $A(P) \approx A(P')$.

NOTATION.   To avoid the boundary problems that may arise from working with a single dyadic grid partition, we will define the concept of the top (of a tree – see the next definition) as being a set of tiles $\tilde{P} =^{def} \{P^j\}_{j \in \{1, \ldots, s\}}$ with $s \in \mathbb{N}$, $s \leq 4$ and $\{P^j = [\alpha^j, \omega^j, I^j]\}_j$ having the properties:

1. $I^j = I^k$, $\forall j, k \in \{1, \ldots s\}$;
2. $4P^j \leq 4P^k$, $\forall j, k \in \{1, \ldots s\}$.

(This technicality is introduced only for smooth handling of the tree selection argument from section 7.)

For $P \in \mathbb{P}$ we write $P \leq \tilde{P}$ iff $\exists j \in \{1, \ldots, s\}$ such that $P \leq P^j$. In what follows, it will also be convenient to work with a representative of the top $\tilde{P}$ - call it P - which is some tile from the collection $\{P^j\}_{j \in \{1,\ldots,s\}}$. (The reader may imagine a top as consisting of only one "fat" tile; indeed, in the following definitions and results, the accent will always fall on a representative of a top which may be regarded as a "specialization" of the top itself.)

Using the relation just defined, we now introduce the fundamental (geometric) set-configuration that will govern most of our reasonings.

DEFINITION 4.   *We say that a set of tiles $\mathcal{P} \subset \mathbb{P}$ is a tree (relative to "$\leq$") with top $\tilde{P}_0$ if the following conditions are satisfied:*

1. *$\forall P \in \mathcal{P} \Rightarrow \frac{3}{2}P \leq \tilde{P}_0$;*
2. *If $P \in \mathcal{P}$ and $\frac{3}{2}P_u \leq \tilde{P}_0$ then $P_u \in \mathcal{P}$ (analogously for P_l);*
3. *If $P_1, P_2 \in \mathcal{P}$ and $P_1 \leq P \leq P_2$ then $P \in \mathcal{P}$.*

OBSERVATION 3.   a. While conditions 1 and 3 have clear correspondents in [F] (the only difference appearing here is the factor $3/2$ in 1 which is used for overcoming the boundary problems that will arise later – see section 7.2), the second condition, added here, is the extra twist that offers our trees the advantage of being "centered". More exactly, the central line of the top (representative) splits the time-frequency representation of our tree in two "halves".

b. Sometimes we may exclude the (tiles forming the) top of the tree from the collection $\mathcal{P}$. Also, we say that a tree has (top) frequency line l if l is the central line of one of the tiles (representative) belonging to the top.

In this framework, we can state the results that will be used for proving our theorem; their proofs will be postponed until section 8.

PROPOSITION 1.   *There exists $\eta \in (0, 1/2)$ s.t. if $\mathcal{P}$ is any given family of incomparable tiles (i.e. no two of them can be related through "$\leq$") with the property that*

$$A(P) \leq \delta, \quad \forall P \in \mathcal{P},$$

then

$$\|T^{\mathcal{P}}\|_2 \lesssim \delta^\eta.$$

PROPOSITION 2.   *Let $\{\mathcal{P}_j\}_j$ be a family of trees with tops $\{\tilde{P}_j\}_j$ and respective representatives $\{P_j = [\alpha_j, \omega_j, I_j]\}_j$. Suppose that*

1. *$A(P) < \delta \ \forall j, P \in \mathcal{P}_j$.*
2. *$\forall k \neq j$ and $\forall P \in \mathcal{P}_j, 2P \nleq 2\tilde{P}_k$.*
3. *No point of $[0, 1]$ belongs to more than $K\delta^{-2}$ of the I_j.*

Then there is an absolute constant $\eta \in (0, 1/2)$ and a set $F \subset \mathbb{T}$ with $|F| \lesssim \delta^{50}K^{-1}$ such that $\forall f \in L^2(\mathbb{T})$ we have

$$\left\| \sum_j T^{\mathcal{P}_j} f \right\|_{L^2(F^c)} \lesssim \delta^\eta \log K \|f\|_2.$$

(*Remark: Any collection of tiles $\mathcal{P}$ that can be represented as $\cup_j \mathcal{P}_j$ with the family $\{\mathcal{P}_j\}$ respecting the conditions mentioned above will be called a "forest".*)

OBSERVATION 4.   One may notice the similarity between the above propositions and the corresponding statements in [F] (Lemma 2 and Main Lemma); this is not surprising since the "only" difference between the quadratic case and the linear case is that we have to deal with slanted rectangles. While the proof of Proposition 1 is basically the same as in [F], for the second proposition we will have to deal with the extra overlaps of our parallelograms.

7   Proof of "Pointwise Convergence"

We now present the proof of Theorem 1.

7.1   Organizing the family of tiles.   We start by breaking up $\mathbb{P}$ into $\bigcup_{n=0}^{\infty} \mathcal{P}_n$ where

$$\mathcal{P}_n = \left\{ P \in \mathbb{P} \mid 2^{-n-1} < A(P) \leq 2^{-n} \right\}.$$

Thus we have

$$T = \sum_{n=0}^{\infty} T^{\mathcal{P}_n}.$$

Here is the plan of our proof:

Step 1 (the remaining part of section 7.1): We modify each $\mathcal{P}_n$ so that the resulting set gains a certain structure: all the elements inside it have comparable mass and are clustered near some "well-arranged" maximal elements.

Step 2 (section 7.2): Taking advantage of the above-mentioned structure, we further show that each such $\mathcal{P}_n$ may be decomposed (up to a negligible family of tiles, in the sense of Proposition 1) into a certain number of forests.

Step 3 (section 7.3): Using Proposition 2, we will combine the estimates for each forest into an estimate for the operator $T^{\mathcal{P}_n}$, which allows us to obtain the desired bound for T.

As announced, we start the first part of our program by modifying (cutting) some parts of the set $\mathcal{P}_n$. For this, we first define $\{\bar{P}_k\}$, $\bar{P}_k = [\bar{\alpha}_k, \bar{\omega}_k, \bar{I}_k]$ to be the set of maximal triples with respect to "$\leq$" that obey the relation $|E(P)|/|I_P| \geq 2^{-n-1}$. Set $\mathcal{P}_n^0$ to be

$$\mathcal{P}_n^0 = \{ P \in \mathcal{P}_n \mid \exists k \in \mathbb{N} \ s.t. \ 4P \lhd \bar{P}_k \}, \tag{31}$$

and define also

$$\mathcal{C}_n = \left\{ P \in \mathcal{P}_n \mid \text{there are no chains } P \lneqq P_1 \lneqq \ldots \lneqq P_n \ \& \ \{P_j\}_{j=1}^n \subseteq \mathcal{P}_n \right\}.$$

With this notation we claim that

$$\mathcal{P}_n \setminus \mathcal{C}_n \subseteq \mathcal{P}_n^0.$$

Indeed, if $P \in \mathcal{P}_n \setminus \mathcal{C}_n$ then there exists a family of tiles $\{P_j\}_{j=1}^n \subseteq \mathcal{P}_n$ such that $P \lneqq P_1 \lneqq \cdots \lneqq P_n$. Now, since $2^{-n-1} < A(P_n) \leq 2^{-n}$, we deduce that $\exists P' = [\alpha', \omega', I']$ with $I_n \subseteq I'$ such that $|E(P')|/|I'| \geq 2^{-n-1}$ and $\Delta(P_n, P') < 2^{n/3}$. From the maximality condition, we have that $\exists k \in N$ s.t. $P' \leq \bar{P}_k$ and so $\Delta(P_n, \bar{P}_k) < 2^{n/2}$.

On the other hand, from the chain condition, we deduce that $\Delta(P, P_n) < 3/2$ and $|\omega_P| \geq 2^n |\omega_{P_n}| \geq 2^n |\omega_{\bar{P}_k}|$. Consequently, we have that

$$\Delta(P, \bar{P}_k) < 3/2 + 2^{-n/2},$$

which implies that $4P \lhd \bar{P}_k$ as we wanted.

Let $\mathcal{D}_n \subseteq \mathcal{C}_n$ be the set such that $\mathcal{P}_n \setminus \mathcal{D}_n = \mathcal{P}_n^0$; then $\mathcal{D}_n$ (or, in general, any subset of $\mathcal{C}_n$) contains no (ascending) chains of length $n + 1$ and so breaks up as a disjoint union of a most n sets $\mathcal{D}_{n1} \cup \mathcal{D}_{n2} \cup \ldots \cup \mathcal{D}_{nn}$ with no two tiles in the same $\mathcal{D}_{nj}$ comparable. Consequently, from Proposition 1, we have

$$\exists \eta \in (0, 1/2) \ s.t. \ \big\|T^{\mathcal{D}_{nj}}\big\|_2 \lesssim 2^{-n\eta}, \quad \forall j \in \{1, \ldots n\},$$

which applied to $\mathcal{D}_n$ translates into

$$\big\|T^{\mathcal{D}_n}\big\|_2 \leq \sum_{j=1}^n \big\|T^{\mathcal{D}_{nj}}\big\|_2 \lesssim \sum_{j=1}^n 2^{-n\eta} \lesssim 2^{-n\eta}. \tag{32}$$

As a consequence, we can now erase the set $\mathcal{D}_n$ without affecting our plan. The resulting structure of the collection $\mathcal{P}_n^0$ will help us later to further split our collection into forests, but for the moment we turn our attention towards the set $\{\bar{P}_k\}$, with the intention of obtaining a rough bound for the counting function N (defined below) associated to the intervals $\{\bar{I}_k\}$. For this we notice that $\{E(\bar{P}_k)\}$ are pairwise disjoint, which implies that $\sum_k |E(\bar{P}_k)| \leq 1$. Now, using the definition of $\bar{P}_k$, we deduce

$$\|N\|_1 = \sum_k |\bar{I}_k| \leq 2^{n+1}|E(\bar{P}_k)| \leq 2^{n+1} \text{ where } N(x) =^{def} \sum_k \chi_{\bar{I}_k}(x).$$

Therefore the set defined as

$$G_n = \big\{x \in \mathbb{T} \mid \text{ is contained in more than } 2^{2n}K \text{ of the } |\bar{I}_k|\big\}$$

has measure $|G_n| \lesssim (2^n K)^{-1}$. Because we want some control on the geometry of $\{\bar{P}_k\}$, we will use G_n for deleting more tiles from $\mathcal{P}_n^0$; indeed, if

$$\mathcal{P}_n^G = \big\{P = [\alpha, \omega, I] \in \mathcal{P}_n^0 \mid I \nsubseteq G_n\big\},$$

we have that

$$T^{\mathcal{P}_n^G} f(x) = T^{\mathcal{P}_n^0} f(x) \quad \forall f \in L^2(\mathbb{T}) \ \& \ \in G_n^c. \tag{33}$$

(Since we have good control of the measure of G_n, we will focus on estimating $T^{\mathcal{P}_n^0}$ only on G_n^c.)

We delete from $\{\bar{P}_k\}$ all $\bar{P}_k$ with $\bar{I}_k \subseteq G_n$. Then the resulting set $\mathcal{P}_n^G$ has the following properties:

1. $A(P) \leq 2^{-n}, \forall P \in \mathcal{P}_n^G$;
2. $\forall P \in \mathcal{P}_n^G \Rightarrow \exists k \in \mathbb{N} \ s.t. \ 4P \unlhd \bar{P}_k$;
3. No $x \in \mathbb{T}$ belongs to more than $K2^{2n}$ of the $\bar{I}_k$'s.

7.2 Decomposing into forests. Now we shall prove that $\mathcal{P}_n^G$ decomposes (up to a family of chains with length controlled by an absolute constant) as a disjoint union of at most $M = 2n \log K$ forests $\mathcal{B}_{n0} \cup \mathcal{B}_{n1} \cup \mathcal{B}_{n2} \cup \cdots \cup \mathcal{B}_{nM}$, where each $\mathcal{B}_{nk}$ satisfies the hypotheses of Proposition 2. In order to make the decomposition, we first define

$$B(P) = \#\{j \mid 4P \unlhd \bar{P}_j\} \quad \forall P \in \mathcal{P}_n^G.$$

Clearly $1 \leq B(P) \leq 2^M$. Now let's define the sets
$$\mathcal{P}_{nj} := \left\{ P \in \mathcal{P}_n^G \mid 2^j \leq B(P) < 2^{j+1} \right\} \quad \forall j \in \{0, \dots, M\}.$$

To better understand their behavior, we develop the following procedure: fix a family of tiles $\mathcal{P}_{nj}$ as defined before and

1. Select the tiles $\{P^r\}_{r \in \{1, \dots s\}} \subseteq \mathcal{P}_{nj}$ with the property that $4P^r$ are maximal elements with respect to the relation "$\leq$" inside the set $4\mathcal{P}_{nj}$. (Here we use the following convention: let be $\mathcal{D}$ a collection of tiles; P is maximal relative to "$\leq$" in $\mathcal{D}$ iff $\forall P' \in \mathcal{D}$ such that $P \leq P'$ we also have $P' \leq P$.)

2. From the maximality, we have that
$$4P^l \leq 4P^k \Rightarrow I_l = I_k, \tag{34}$$
$$\forall P \in \mathcal{P}_{nj} \; \exists P^l \text{ s.t. } 4P \leq 4P^l. \tag{35}$$

3. From the definition of $\mathcal{P}_{nj}$ we deduce
$$\text{if } P \in \mathcal{P}_{nj} \text{ s.t. } \exists k \neq l \text{ with } \begin{cases} 4P \trianglelefteq 4P^l \\ 4P \trianglelefteq 4P^k \end{cases}, \text{ then } \begin{cases} 4P^k \leq 4P^l \\ 4P^l \leq 4P^k \end{cases}.$$

4. Define
$$\mathcal{A}_{nj} := \left\{ P \in \mathcal{P}_{nj} \mid \forall P^l \Rightarrow \tfrac{3}{2}P \not\leq P^l \right\} \cup$$
$$\left\{ P \mid |\exists l \text{ s.t. } |I_P| = |I_{P^l}|, \; \tfrac{3}{2}P \leq P^l \; \& \; P \neq P^k \; \forall k \right\} = \mathcal{A}_{nj}^1 \cup \mathcal{A}_{nj}^2,$$
and set
$$\mathcal{P}_{nj} = \mathcal{A}_{nj} \cup \mathcal{B}_{nj}.$$

Now, we claim that

(a) $\mathcal{A}_{nj}$ can be split into a controlled number of sets containing no chains (with respect to the relation "$\leq$").

(b) The collection $\mathcal{B}_{nj}$ defines a forest (up to a negligible family of tiles).

We start with the proof of (a), by supposing that we can find $P_1, P_2 \in \mathcal{A}_{nj}^1$ such that $P_1 \lneq P_2$; suppose also (see (35)) that
$$4P_2 \leq 4P^l \tag{36}$$
for some l.

Now from the definition of $\mathcal{A}_{nj}^1$ we have that $\tfrac{3}{2}P_1 \not\leq P^l$, but observing that
$$|\omega_{P_1}| \geq 2^{10}|\omega_{P_2}| \tag{37}$$
we contradict relation (36). The fact that the remaining set $\mathcal{A}_{nj}^2$ contains no chains comes trivially from the maximality of the tiles $\{P^r\}_{r \in \{1, \dots, s\}}$.

For part (b), we proceed as follows. We choose $k \in \{1, \dots, s\}$ and define
$$S_k = \left\{ P \in \mathcal{B}_{nj} \mid \tfrac{3}{2}P \lneq P^k \right\}. \tag{38}$$

We now collect all $\{P^k\}_k$ for which $S_k = \emptyset$ and erase them using Proposition 1. Consequently, by relabeling the remaining maximal tiles we can always suppose that for each P^k we have $S_k \neq \emptyset$ and that $\mathcal{B}_{nj} = \bigcup_k \{S_k \cup P^k\}$. Further, we want to study the separation properties of the family $\{S_k\}_k$. For this, we first introduce the following relation: we say that
$$S_k \propto S_l$$
if and only if $\exists P_1 \in S_k$ and $\exists P_2 \in S_l$ such that $2P_1 \leq 2P_2$ or $2P_2 \leq 2P_1$.

With this done, we first claim that
$$S_k \propto S_l \Rightarrow 4P^k \leq 4P^l \Rightarrow I^k = I^l. \tag{39}$$
Indeed, suppose that $S_k \propto S_l$, and so (without loss of generality) we know that $k \neq l$ and there are $P_1 \in S_k$ and $P_2 \in S_l$ such that $2P_1 \leq 2P_2$. Then, since $\frac{3}{2}P_2 \leq P^l$ and $|\omega_1| \geq |\omega_2| \geq 2^{10}|\omega^l|$, we must have $4P_1 \trianglelefteq 4P^l$. On the other hand, since $P_1 \in S_k$, we also have $4P_1 \trianglelefteq 4P^k$, but this forces (see 3) $4P^k \leq 4P^l$.

We now construct the sets
$$\bar{S}_k := S_k \cup P^k, \quad k \geq 1,$$
and observe that with a similar reasoning as in (39) we obtain
$$\bar{S}_k \propto \bar{S}_l \Rightarrow 4P^k \leq 4P^l \Rightarrow I^k = I^l. \tag{40}$$
The point is that with respect to $\{\bar{S}_k\}_k$, $\propto$ becomes an equivalence relation. Indeed, let us check the transitivity of our relation. Suppose that $\bar{S}_k \propto \bar{S}_l \propto \bar{S}_m$. Now, since $\bar{S}_k \propto \bar{S}_l$, we deduce from (40) that $4P^k \leq 4P^l$, and since $S_k \neq \emptyset$ we also have that $\exists P_1 \in S_k$ with $\frac{3}{2}P_1 \lneq P^k$. On the other hand, from $\bar{S}_l \propto \bar{S}_m$, we have that $4P^l \leq 4P^m$. Putting these facts together, we have that $10P^k \leq 10P^m$, $I^k = I^m$, and since $|\omega_1| \geq 2^{10}|\omega^k|$ we deduce $2P_1 \trianglelefteq 2P^m$, which proves our claim.

Now let $\hat{k} := \{m \mid \bar{S}_m \propto \bar{S}_k\}$ (observe that the size of the orbit of each k ($\bar{S}_k$) is bounded from above by an absolute constant). Denote
$$\hat{S}_k := \bigcup_{m \in \hat{k}} \bar{S}_m.$$

Now, choosing a unique representative in each equivalence class, and relabeling the resulting elements in a consecutive order, we deduce that $\hat{S}_k \cap \hat{S}_l = \emptyset$ for any $k \neq l$, which implies $\{\hat{S}_k\}_k$ is a partition of $\mathcal{B}_{nj}$.

We need some final modifications to each set $\hat{S}_k$. First, we denote by $\tilde{P}^k$ the set of all the maximal tiles $\{P^l\}_l$ contained in the collection $\hat{S}_k$; now, using Proposition 1, we delete, for each k, all the elements belonging to $\tilde{P}^k$. Secondly, we define (for each k) the set
$$\hat{S}_k^{\min} = \{P \in \hat{S}_k \mid \forall P' \in \hat{S}_k \text{ s.t. } I_{P'} \cap I_P \neq \emptyset \Rightarrow I_P \subseteq I_{P'}\}.$$
By construction, the set $\{\hat{S}_k^{\min}\}_k$ contains only pairwise incomparable tiles, so, again applying Proposition 1, we can erase the set $\hat{S}_k^{\min}$ from each $\hat{S}_k$ and consider $\mathcal{B}_{nj} = \bigcup_k \hat{S}_k$. In what follows we will prove that each $\hat{S}_k$ is a tree with top $\tilde{P}^k$.

Indeed, fix a collection $\hat{S}_k$ of tiles; we will now verify conditions 1–3 in Definition 4. Take $P \in \hat{S}_k$; first observe that condition 1 holds trivially since by construction $\frac{3}{2}P \lneq \tilde{P}^k$. Suppose now that $P \in \hat{S}_k$ with $\frac{3}{2}P_u \leq \tilde{P}^k$. Then to show $P_u \in \hat{S}_k$ it is enough to prove that $P_u \in \mathcal{B}_{nj}$. For this, we need first to prove that $P_u \in \mathcal{P}_n^G$. Since $\frac{3}{2}P_u \leq P^k$ (for some P^k an element of $\tilde{P}^k$) the above statement reduces to $P_u \in \mathcal{P}_n$. But we know that $2P_u \trianglelefteq 2P^k$, and since $A(P^k) > 2^{-n-1}$, using (mp), we deduce that also $A(P_u) > 2^{-n-1}$. At this point, we recall that (following the previous procedure)
$$\exists P_0 \in \hat{S}_k^{\min} \text{ s.t. } I_{P_0} \subsetneq I_P. \tag{41}$$
Consequently, using (41), we have $2P_0 \trianglelefteq 2P_u$, and so by (mp) $A(P_u) \leq A(P_0) \leq 2^{-n}$. For the second part, we need $P_u \in \mathcal{P}_{nj}$, but this comes from the fact that
$$4P_u \trianglelefteq 4P^k \Rightarrow B(P_u) \geq 2^j$$

and
$$4P_0 \trianglelefteq 4P_u \Rightarrow B(P_u) < 2^{j+1}\,.$$

From this, we conclude that $P_u \in \hat{S}_k$, so condition 2 is true.

The convexity condition 3 is trivial since if $P_1 < P_2 < P_3$ with P_1 & $P_3 \in \hat{S}_k$ we have $\frac{3}{2}P_1 \trianglelefteq \frac{3}{2}P_2 \trianglelefteq \frac{3}{2}P_3 \leq \frac{3}{2}\tilde{P}^k$, which implies $P_2 \in \hat{S}_k$.

Consequently, we have proven $\hat{S}_k$ is a tree with top $\tilde{P}^k$. Now, from the previous considerations, we have that
$$\hat{S}_k \propto \hat{S}_l \Rightarrow k = l\,,$$

and since $\mathcal{B}_{nj} = \bigcup_k \hat{S}_k$ we deduce that $\mathcal{B}_{nj}$ becomes a forest as defined in Proposition 2.

7.3 Ending the proof.   Now, we may conclude as in [F].

We first apply Proposition 2 for each family $\mathcal{B}_{nj}$ and obtain that
$$\left\| T^{\mathcal{P}_{nj}} f \right\|_{L^2(F_{nj}^c)} \lesssim 2^{-n\eta} \log K \|f\|_2\,,$$

where F_{nj} is a small set with measure $|F_{nj}| \lesssim 2^{-n} K^{-1}$. As a result, denoting $F_n = \cup_j F_{nj}$, we have that
$$\left\| T^{\mathcal{P}_n^G} f \right\|_{L^2(F_n^c)} \leq \sum_{j=1}^{2n \log K} \left\| T^{\mathcal{P}_{nj}} f \right\|_{L^2(F_{nj}^c)} \lesssim n 2^{-n\eta} (\log K)^2 \|f\|_2 \tag{42}$$

with $|F| \lesssim n \log K / 2^n K$.

Therefore, combining (32), (33) and (42), we deduce
$$\left\| T^{\mathcal{P}_n} f \right\|_{L^2(E_n^c)} \lesssim n 2^{-n\eta} (\log K)^2 \|f\|_2\,,$$

where $E_n = F_n \cup G_n$ still has measure $\lesssim n \log K / 2^n K$.

Summing now over n, we obtain
$$\|Tf\|_{L^2(E^c)} \lesssim (\log K)^2 \|f\|_2$$

with $E = \cup_n E_n$ and $|E| \lesssim \log K / K$.

In conclusion, given $\gamma > 0$, we have that for all $K > 100$
$$\left| \{ |Tf(x)| > \gamma \} \right| \leq \frac{\|Tf\|_{L^2(E^c)}^2}{\gamma^2} + |E| \lesssim (\log K)^4 \frac{\|f\|_2^2}{\gamma^2} + \frac{\log K}{K}\,.$$

Now, if we pick K to minimize the right-hand side, we arrive at the relation
$$\left| \{ |Tf(x)| > \gamma \} \right| \lesssim_\epsilon \left(\frac{\|f\|_2}{\gamma} \right)^{2-\epsilon} \quad \forall \epsilon \in (0, 2)\,,$$

which further implies
$$\|Tf\|_p \lesssim_p \|f\|_2 \quad \forall p < 2\,,$$

ending the proof of our theorem.

8   Some Technicalities – The Proofs of Propositions 1 and 2

We now present the proofs of the statements made in section 6.

Proof of Proposition 1.   The basic idea of our proof relies on combining the TT^* and maximal methods. Indeed, once we have expressed the norm of our operator as a sum of interactions among "small pieces" T_P, we split it in two terms:

– for the first one (close to the diagonal) we use some maximal methods since all our pieces T_P "oscillate" in the same region of the time-frequency plane;

– for the second one (far from the diagonal) we take advantage of the orthogonality of our terms, which is reflected in the smallness of the resulting geometric factors.

$$\int_{\mathbb{T}} \left|(T^{\mathcal{P}})^* f(x)\right|^2 dx \lesssim \left| \sum_{\substack{P' \in \mathcal{P} \\ P'=[\alpha',\omega',I']}} \int_{\mathbb{T}} f(x) \left\{ \sum_{\substack{P=[\alpha,\omega,I] \in \mathcal{P} \\ |I| \leq |I'|}} \overline{T_{P'}T_P^* f}(x) \right\} dx \right|$$

$$\lesssim \sum_{P' \in \mathcal{P}} \int_{E(P')} |f| \left\{ \sum_{P \in a(P')} \left\lceil \Delta(P,P') \right\rceil^{1/2} \frac{\int_{E(P)} |f|}{|I'|} \right\}$$

$$+ \sum_{P' \in \mathcal{P}} \int_{E(P')} |f| \left\{ \sum_{P \in b(P')} \left\lceil \Delta(P,P') \right\rceil^{1/2} \frac{\int_{E(P)} |f|}{|I'|} \right\} \overset{\text{def}}{=} A + B$$

where for the second inequality we used the estimate (cf. Lemma 0)

$$\left| T_{P'} T_P^* f(x) \right| \lesssim \left\lceil \Delta(P,P') \right\rceil^{1/2} \frac{\int_{E(P)} |f|}{|I'|} \chi_{E(P')}(x)$$

together with the following notation:

$$a(P') = \left\{ P = [\alpha,\omega,I] \in \mathcal{P}, \ |I| \leq |I'| \ \& \ I^* \cap I'^* \neq \emptyset \mid \Delta(P,P') \leq \delta^{-2\epsilon} \right\},$$

$$b(P') = \left\{ P = [\alpha,\omega,I] \in \mathcal{P}, \ |I| \leq |I'| \ \& \ I^* \cap I'^* \neq \emptyset \mid \Delta(P,P') \geq \delta^{-2\epsilon} \right\}.$$

(Here $\epsilon \in (0,1)$ is some fixed constant.)

Now the second term is easy to estimate:

$$B \lesssim \sum_{P' \in \mathcal{P}} \int_{E(P')} |f(x)| \left\{ \frac{\delta^\epsilon}{|I'|} \sum_{P \in b(P')} \int_{E(P)} |f| \right\} dx$$

$$\leq \delta^\epsilon \sum_{P' \in \mathcal{P}} \int_{E(P')} |f(x)t| \left\{ \frac{1}{|I'|} \int_{\tilde{I}'} |f| \right\} dx \leq \delta^\epsilon \int_{\mathbb{T}} |f(x)| M f(x) dx \lesssim \delta^\epsilon \int_{\mathbb{T}} |f|^2.$$

For the first term we use the following Carleson measure-type estimate:

$$\sum_{P \in a(P')} |E(P)| \lesssim \delta^{1-100\epsilon} |I'|, \tag{43}$$

which is a consequence of the smoothness property (sp) of the mass $A(P)$. Indeed, define

$$\mathcal{J}(P') = \left\{ I \mid \exists P = [\alpha,\omega,I] \in a(P') \right\}.$$

Let $\mathcal{J}_{\min}(P')$ be the set of minimal (with respect to inclusion) intervals inside $\mathcal{J}(P')$, and define

$$\check{\mathcal{J}}(P') := \left\{ I \subset 30I' \mid \text{Exactly one of the left or right halves} \right.$$
$$\left. \text{of } I \text{ contains an element of } \mathcal{J}_{\min}(P') \right\} \cup \mathcal{J}_{\min}(P').$$

Finally, set

$$\check{a}(P') = \left\{ P = [\alpha,\omega,I] \in \mathbb{P} \mid I \in \check{\mathcal{J}}(P') \ \& \ \Delta(P,P') \leq \delta^{-2\epsilon} \right\}.$$

Then using the property (sp) (with an appropriate choice of N in the definition of $A(P)$), and the fact that any two tiles inside $\mathcal{P}$ are not comparable we have

$$\sum_{P \in a(P')} |E(P)| \leq \sum_{P \in \check{a}(P')} |E(P)| \lesssim \delta^{1-100\epsilon} \sum_{I \in \check{\mathcal{J}}(P')} |I| \lesssim \delta^{1-100\epsilon} |I'|,$$

which gives us the desired estimate (43).

Now set $E_{P'} := \cup_{P \in a(P')} E(P)$; using Hölder's inequality for some fixed $1 < r < 2$, we deduce

$$A \lesssim \sum_{P' \in \mathcal{P}} \frac{\int_{E(P')} |f|}{|I'|} \int_{E_{P'}} |f| \lesssim \sum_{P' \in \mathcal{P}} \int_{E(P')} |f| \left(\frac{|E_{P'}|}{|I'|} \right)^{1-\frac{1}{r}} \left(\frac{\int_{I'} |f|^r}{|I'|} \right)^{\frac{1}{r}}$$

$$\lesssim \delta^{1-1/r-100\epsilon} \sum_{P' \in \mathcal{P}} \int_{E(P')} |f(y)| f_r^*(y) dy$$

$$\lesssim \delta^{1-1/r-100\epsilon} \int_{\mathbb{T}} [f_r^*(y)]^2 dy \lesssim \delta^{1-1/r-100\epsilon} \|f\|_2^2 .$$

(Here $f_r^*(x) = \sup_{x \in I} (\int_I |f(y)|^r dy / |I|)^{1/r}$ designates the Hardy–Littlewood maximal function of order r.)

The conclusion of our proposition now follows, by properly choosing $\epsilon > 0$.   □

The remainder of the section will be dedicated to proving Proposition 2. The natural approach is to obtain control on

– the single tree estimate (Lemma 1),
– the interaction between (separated) trees (Lemmas 2 and 3).

We now start the study of the other extremal geometric configuration, namely the tree (as opposed to the structure of the family of tiles appearing in Proposition 1). For Lemma 1, due to the structure of our family, the geometric factors will play no role, the entire effort being concentrated on properly using the (uniform) density condition and the mean zero property.

LEMMA 1.   Let $\delta > 0$ be fixed and let $\mathcal{P} \subseteq \mathbb{P}$ be a tree with top $\tilde{P}_0$, representative $P_0 = [\alpha_0, \omega_0, I_0]$, and frequency line l_0 and such that

$$A(P) < \delta \quad \forall P \in \mathcal{P} .$$

Then

$$\|T^{\mathcal{P}}\|_2 \leq \delta^{1/2} .$$

Proof. The essence of the proof below (our case is a "quadratic perturbation" that realizes a shearing of the linear tree case presented in [F]) relies on the following outlook:

"For $\mathcal{P}$ a tree, the associated operator $T^{\mathcal{P}}$ behaves

like the (maximal) Hilbert transform."   (44)

[REMARK.   The easiest way to understand this heuristic is to take a particular instance of $\mathcal{P}$ ($T^{\mathcal{P}}$): suppose that the top P_0 stays on the real axis, that $l_x \equiv 0$ for any $x \in I_0$ and that all the minimal tiles in the collection $\mathcal{P}$ (we may assume $\mathcal{P}$ finite) are in the same scale. Then, from the convexity condition 3 in Definition 4, we remark that $\exists k_0, k_1 \in \mathbb{Z}$, such that, for $x \in I_0$,

$$T^{\mathcal{P}} f(x) = \sum_{k_0 \leq k \leq k_1} \int \psi_k(y) f(x-y) dy , \tag{45}$$

i.e. $T^{\mathcal{P}}$ is a truncation of (7).]

To make this precise, we will further show that (45) is always true locally, on $\operatorname{supp} T^{\mathcal{P}}$. Indeed, let

$$k_0(x) = \inf\left\{k \in \mathbb{N} \mid \exists P \in \mathcal{P} \text{ s.t. } |I_P| = 2^{-k} \quad \chi_{E(P)}(x) \neq 0\right\},$$

$$k_1(x) = \sup\left\{k \in \mathbb{N} \mid \exists P \in \mathcal{P} \text{ s.t. } |I_P| = 2^{-k} \quad \chi_{E(P)}(x) \neq 0\right\}.$$

Using the convexity condition we then deduce

$$T^{\mathcal{P}}f(x) = \sum_{k_0(x) \leq k \leq k_1(x)} \int e^{i(l_x(x)y - b(x)y^2)} \psi_k(y) f(x-y) dy. \tag{46}$$

Since we also want to obtain some decay, we need to take advantage of the "mass" of our tree. For this, the key fact is to observe that heuristically our operator behaves as follows:

$$x \xmapsto{T^{\mathcal{P}}f} 0 \text{ if } x \in I^j \setminus E^j,$$

$$x \xmapsto{T^{\mathcal{P}}f} \operatorname{constant}(j) \text{ if } x \in E^j, \tag{47}$$

where here the sets $\{I^j\}_j$ and $\{E^j\}_j$ obey the conditions:

  – $\{I^j\}$ is a partition of I_0;
  – $E^j \subseteq I^j$ and $|E^j|/|I^j| \lesssim \delta$.

Indeed, one can define $\{I^j\}_j$ to be the maximal dyadic intervals contained in I_0 that satisfy

$$\frac{|E(l_0, I)|}{|I|} > 100\delta$$

where $E(l_0, I) := \{x \in I \mid \operatorname{dist}^I(l_x, l_0) < 2|I|^{-1}\}$. Now setting $\bar{E}^j = E(l_0, \bar{I}^j)$ and $E^j = \bar{E}^j \cap I^j$ and making use of (sp) one concludes that if $P = [\alpha, \omega, I] \in \mathcal{P}$ with $I \cap I^j \neq \emptyset$ then $\bar{I}^j \subseteq I$ & $E(P) \cap I^j \subseteq E^j$ (here $\bar{I}^j$ is the dyadic interval containing I^j and having twice the length).

Now, combining the views offered by (44) and (47) we proceed as follows: To come closer to (45), our first step is to move our tree near the real axis: set $\mathcal{T}^{\mathcal{P}} = Q_{b_0}{}^* M_{c_0}{}^* T^{\mathcal{P}} M_{c_0} Q_{b_0}$ and $g(x) = M_{c_0}^* Q_{b_0}^* f(x)$ (here $l_0(z) = c_0 + 2b_0 z$ is the central line of P_0). Then, for $x \in I_0$ fixed, we have

$$\left|T^{\mathcal{P}}f(x)\right| = \left|\mathcal{T}^{\mathcal{P}}g(x)\right|$$

$$\leq \sum_{k_0(x) \leq k \leq k_1(x)} \left\{\int_{\mathbb{T}} \left|e^{i\{(l_x(x) - l_0(x))y - (b(x) - b_0)y^2\}} - 1\right| |\psi_k(y)| |g(x-y)| dy\right\}$$

$$+ \left|\sum_{k_0(x) \leq k \leq k_1(x)} \left\{\int_{\mathbb{T}} \psi_k(y) g(x-y) dy\right\}\right| = A(x) + B(x).$$

Now for the first term, using (47) and the small oscillation of the exponential, we deduce that

$$A(x) \lesssim M_\delta f(x).$$

For the second term, as claimed initially, we remark that $B(x)$ is the local version of (45); to "achieve" (44) we need to compare B with some averages of the Hilbert transform. Here, the main ingredient is

$$\left|\sum_{k \leq K} \psi_k(y) - 2^{K-1} \int_{-2^{-K}}^{2^{-K}} R(y+z) dz\right| \lesssim \frac{2^{-K}}{|y|^2 + 2^{-2K}}, \tag{48}$$

where $R(y) = \sum_{k \in 10\mathbb{N}} \psi_k(y)$, $K \in \mathbb{N}$ and $y \in \mathbb{T}$.

Now, for $x \in I_j$ fixed, we conclude that

$$B(x) \lesssim \sup_{I \supset I_j} \left(\frac{1}{|I|} \int_I |(R * g)(y)| dy \right) + \sup_{I \supset I_j} \left(\frac{1}{|I|} \int_I |g(y)| dy \right)$$

$$= M_\delta(R * g)(x) + M_\delta g(x) .$$

Finally, combining our estimates for A and B and using the fact that the operator $g \to R * g$ is bounded on $L^2(\mathbb{T})$ (this comes from $\hat{R} \in L^\infty(\mathbb{T})$, which is an easy consequence of the fact that the function ψ is compactly supported away from the origin and has mean zero), we conclude that

$$\|T^{\mathcal{P}} f\|_2 \lesssim \left\| M_\delta(R * (M_{c_0}^* Q_{b_0}^* f)) \right\|_2 + \|M_\delta f\|_2 \lesssim \delta^{1/2} \|f\|_2 .   \square$$

At this point we have learned how to estimate basic families of tiles, having a simple geometric structure, for which we have uniform control on the density factor. The next step (Lemmas 2 and 3) will be to understand the interaction between two such basic families in the case in which we have no information about their density factors, but we know that they are located in different regions of the time-frequency plane. (Here we will use the fact that the geometric factor of pair (P_1, P_2) is small whenever P_1 and P_2 are not in the same family of tiles.)

Before presenting the lemmas, we will need several definitions.

DEFINITION 5.   *Fix a number* $\delta \in (0, 1)$. *Let be* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *two trees with (tops* $\tilde{P}_1$ *and* $\tilde{P}_2$) *representatives* $P_1 = [\alpha_1, \omega_1, I_1]$ *and* $P_2 = [\alpha_2, \omega_2, I_2]$ *respectively; we say that* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *are* $(\delta$-)*separated if either* $I_1 \cap I_2 = \emptyset$ *or else*

(i) $P = [\alpha, \omega, I] \in \mathcal{P}_1 \ \& \ I \subseteq I_2 \Rightarrow \lceil \Delta(P, P_2) \rceil < \delta$;
(ii) $P' = [\alpha', \omega', I'] \in \mathcal{P}_2 \ \& \ I' \subseteq I_1 \Rightarrow \lceil \Delta(P', P_1) \rceil < \delta$.

NOTATION.   Whenever we have two trees $\mathcal{P}_1$ and $\mathcal{P}_2$ as in Definition 5 we will denote by $x_{P_1, P_2}^i := x_{1,2}^i$ the abscissa of the intersection point of l_{P_1} with l_{P_2}. With this done, set $\frac{w}{\min(|I_1|, |I_2|)} := \frac{(\delta^{-1} \lceil \Delta(P_1, P_2) \rceil)^{1/2}}{100}$ and define

- I_s – the *separation interval (relative to the intersection)* of $\mathcal{P}_1$ and $\mathcal{P}_2$ by
$$I_s = [x_{1,2}^i - w, x_{1,2}^i + w] \cap \tilde{I}_1 \cap \tilde{I}_2 .$$

- I_c – the $(\epsilon$-)*critical intersection interval* (between $\mathcal{P}_1$ and $\mathcal{P}_2$) by
$$I_c = 3\delta^{1/2 - \epsilon} I_s$$

  where ϵ is some small fixed positive real number.

OBSERVATION 5.   (a) The two notions introduced above can be regarded as indicators of how much the quadratic symmetry is involved in the interaction of the two separated trees. Indeed, the procedure of estimating terms like (here $P_j \in \mathcal{P}_j$ with $j \in \{1, 2\}$)

$$\langle T_{P_1}{}^* f, T_{P_2}{}^* g \rangle   \tag{49}$$

will roughly obey the following scenario:

- if $\tilde{I}_{P_1} \cap I_s = \emptyset$ (or $\tilde{I}_{P_2} \cap I_s = \emptyset$) then (49) can be treated as in Fefferman's case, neglecting the quadratic modulation;
- else, guided by the results obtained in section 5 (see Lemma 0), we will split the integral in (49) in two: the first (integrated over the complement of I_c) will be treated as in the previous case, while the second term (integrated over a

set included in I_c) will be placed into a collection of objects representing the critical contribution of the quadratic symmetry. (We use the fact that, for properly chosen ϵ and ϵ_0, the (ϵ_0-)critical intersection interval $I_{1,2}$ of the pair (P_1, P_2) is always included in I_c.)

(b) Further, we will make use of two essential properties of our above-defined intervals:

1. $\forall P \in \mathcal{P}_1 \cup \mathcal{P}_2$ such that $x^i_{1,2} \in 5\tilde{I}_P$ we have $|I_P| > |I_s|$.
2. $\forall P \in \mathcal{P}_1 \cup \mathcal{P}_2$ we have (for ϵ properly chosen) $|\tilde{I}_P \cap I_c| < \delta^{1/4}|I_P|$.

LEMMA 2.   *Let be $\{\mathcal{P}_j\}_{j \in \{1,2\}}$ two separated trees with tops $P_j = [\alpha_j, \omega_j, I_0]$, $j \in \{1,2\}$. Then, for any $f, g \in L^2(\mathbb{T})$ and $n \in \mathbb{N}$, we have that*

$$\left|\langle T^{\mathcal{P}_1 *}f, T^{\mathcal{P}_2 *}g\rangle\right| \lesssim_n \delta^n \|f\|_{L^2(\tilde{I}_0)}\|g\|_{L^2(\tilde{I}_0)} + \left\|\chi_{I_c} T^{\mathcal{P}_1 *}f\right\|_2 \left\|\chi_{I_c} T^{\mathcal{P}_2 *}g\right\|_2. \qquad (50)$$

(REMARK: *The first term in the right-hand side of (50) expresses the result of the interaction (as if we were in the linear phase case - see [F]) far from the intersection point of the trees while the second one reflects the correction needed for handling the quadratic case in the critical region I_c.)*

Proof. We start the proof of our lemma by making a partial (Whitney) dyadic decomposition of the real axis with respect to the point $x^i_{1,2}$; more exactly (we may assume that $|I_s| = 2^{-r}$ for some $r \in \mathbb{N}$), let

- $A_0 = \left(-\infty, x^i_{1,2} - \frac{|I_0|}{2}\right) \cup \left(x^i_{1,2} + \frac{|I_0|}{2}, \infty\right)$;
- $A_1 = \left[x^i_{1,2} - \frac{|I_0|}{2}, x^i_{1,2} - \frac{|I_0|}{4}\right) \cup \left(x^i_{1,2} + \frac{|I_0|}{4}, x^i_{1,2} + \frac{|I_0|}{2}\right]$;
- $A_2 = \left[x^i_{1,2} - \frac{|I_0|}{4}, x^i_{1,2} - \frac{|I_0|}{8}\right) \cup \left(x^i_{1,2} + \frac{|I_0|}{8}, x^i_{1,2} + \frac{|I_0|}{4}\right]$;
- $A_k = \left[x^i_{1,2} - |I_s|, x^i_{1,2} + |I_s|\right]$.

For $j \in \{1,2\}$ define the following sets:

$S_{j,k} = \{P \in \mathcal{P}_j \mid x^i_{1,2} \in 5\tilde{I}_P\} \quad \mathcal{P}'_j := \mathcal{P}_j - S_{j,k}$

$S_{j,0} = \left\{P \in \mathcal{P}'_j \mid I_P{}^* \cap A_0 \neq \emptyset \ \& \ |\tilde{I}_P| \leq \frac{|I_0|}{4}\right\}$

$S_{j,l} = \left\{P \in \mathcal{P}'_j \mid I_P{}^* \cap A_l \neq \emptyset, \ P \notin S_{j,l-1} \ \& \ |\tilde{I}_P| \leq \frac{|A_l|}{3}\right\} \quad \forall l \in \{1, \ldots, k-1\}.$

With this notation it is clear that $\{S_{j,l}\}_{l=1}^k$ form a partition of $\mathcal{P}_j$.

Now setting $T^*_{j,l} = \sum_{P \in S_{j,l}} T_P{}^*$ we obtain

$$\langle T^{\mathcal{P}_1 *}, T^{\mathcal{P}_2 *}\rangle = \sum_{n,l=0}^{k} \langle T_{1,l}{}^*, T_{2,n}{}^*\rangle. \qquad (51)$$

Now let $A_{-1} = \emptyset$, $A_{k+1} = \emptyset$ and $\tilde{A}_l \stackrel{def}{=} \bigcup_{\substack{P \in S_{j,l} \\ j \in \{1,2\}}} \tilde{I}_P$; then $\operatorname{supp} T_{j,l}{}^* \subset \tilde{A}_l$ with $\tilde{A}_l \subseteq A_{l-1} \cup A_l \cup A_{l+1}$, $\forall l \in \{0, \ldots, k-2\}$ & $\tilde{A}_{k-1} \cap \frac{1}{3}A_k = \emptyset$. Consequently, to estimate (51), we need to study the following expressions:

$$U := \sum_{l=0}^{k-1} \langle T_{1,l}{}^*f, T_{2,l}{}^*g\rangle \qquad Y := \sum_{l=0}^{k-1} \langle T_{1,k}{}^*f, T_{2,l}{}^*g\rangle$$

$$Z := \sum_{l=0}^{k-1} \langle T_{1,l}{}^*f, T_{2,k}{}^*g\rangle \qquad V := \langle T_{1,k}{}^*f, T^*_{2,k}g\rangle. \qquad (52)$$

(The terms of the form $\langle T_{1,l}{}^*f, T_{2,l-1}{}^*g\rangle$ or $\langle T_{1,l}{}^*f, T_{2,l+1}{}^*g\rangle$ have a similar treatment.)

We concentrate now on the first term U.

For the beginning we will introduce several useful tools. For $j \in \{1,2\}$, let $l_{P_j}(x) = l_j(x) = c_j + 2xb_j$ and $d_{j,l} = \min\{|I| \,|\, P = [\alpha, \omega, I] \in \mathcal{P}_j \ \& \ P \in S_{j,l-1} \cup S_{j,l} \cup S_{j,l+1}\}$ (here $S_{j,-1}, S_{j,k+1} := \emptyset$); also, define a real-valued function $\varphi \in C_0^\infty(\mathbb{R})$ with the following properties:

- $\operatorname{supp}\varphi \subset \left\{\frac{1}{4} \le |x| \le \frac{1}{2}\right\}$;
- φ is even ;
- $\left|\hat\varphi(\xi) - 1\right| \lesssim_n |\xi|^n, \quad \forall |\xi| \le 1$ and n big enough ;
- $\left|\hat\varphi(\xi)\right| \lesssim_n |\xi|^{-n}, \quad \forall |\xi| \ge 1$.

$$(53)$$

Now, for $j \in \{1,2\}$ and $l \in \{1,\ldots,k-1\}$ set
$$\varphi_{j,l}(x) = (\delta^{1/3}d_{j,l})^{-1}\varphi\big((\delta^{1/3}d_{j,l})^{-1}x\big),$$
and define the operators
$$\tilde\varphi_{j,l} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}) \quad \text{by } \tilde\varphi_{j,l}f = \varphi_{j,l} * f$$
and
$$\Phi_{j,l} : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}) \quad \text{by } \Phi_{j,l} = M_{c_j}Q_{b_j}\tilde\varphi_{j,l}Q_{b_j}{}^*M_{c_j}{}^*.$$
Remark that $\tilde\varphi_{j,l}$ and $\Phi_{j,l}$ are self-adjoint for all $j \in \{1,2\}$ and $l \in \{1,\ldots,k-1\}$.

Our first aim is to prove the following:

CLAIM. *For $j \in \{1,2\}$, $l \in \{1,\ldots,k-1\}$ and $n \in \mathbb{N}$, decomposing $T_{j,l}{}^*$ as*
$$T_{j,l}{}^*f = \Phi_{j,l}T_{j,l}{}^*f + \Omega_{j,l}f \tag{54}$$
we have
$$\|\Omega_{j,l}\|_2 \lesssim_n \delta^n \tag{55}$$
and
$$\left|\langle \Phi_{1,l}T_{1,l}{}^*f, \Phi_{2,l}T_{2,l}{}^*g\rangle\right| \lesssim_n \delta^n \|f\|_2\|g\|_2. \tag{56}$$

Proof of the claim. Suppose that $d_{2,l} \le d_{1,l}$; now, since $\Phi_{j,l}$ are self-adjoint, for showing (56) it is enough to prove

$$\left|\langle \Phi_{2,l}\Phi_{1,l}\overbrace{T_{1,l}{}^*f}^{v}, u\chi_{\tilde A_l}\rangle\right| \lesssim_n \delta^n \|v\|_2\|u\|_2. \tag{57}$$

Now fix $x \in \tilde A_l$; then
$$\left|\Phi_{2,l}\Phi_{1,l}v(x)\right| = \left|\tilde\varphi_{2,l}Q_{b_2}{}^*M_{c_2}{}^*M_{c_1}Q_{b_1}\tilde\varphi_{1,l}Q_{b_1}{}^*M_{c_1}{}^*v(x)\right|$$
$$\le \int |v(s)| \left|\underbrace{\int \varphi_{2,l}(x-y)\varphi_{1,l}(y-s)e^{i[(b_1-b_2)y^2 + (c_2-c_1)y]}dy}_{\mathcal{K}_l(x,s)}\right| ds.$$

Now making the change of variable $y = x - t\delta^{1/3}d_{2,l}$ we deduce
$$\left|\mathcal{K}_l(x,s)\right|$$
$$= \left|\int_{\mathbb{R}} e^{i\{\delta^{1/3}d_{2,l}t(l_2(x)-l_1(x)) - (b_2-b_1)t^2(\delta^{1/3}d_{2,l})^2\}}\varphi(t)\varphi_{1,l}(x-s-\delta^{1/3}d_{2,l}t)dt\right|.$$

Consequently, we need to estimate an expression of the form
$$I_{\alpha,\beta}(\phi) = \left|\int_{\mathbb{R}} e^{i(\alpha t + \beta t^2)}\phi(t)dt\right|$$

where $\phi \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp}\phi \subseteq \{1/4 \le |t| \le 1/2\}$, $\alpha = \delta^{1/3}d_{2,l}(l_2(x) - l_1(x))$ and $\beta = -(\delta^{1/3}d_{2,l})^2(b_2 - b_1)$.

Now, since the trees are separated, and since $x \in \tilde{A}_l$, we have that
$$\inf_{x \in \tilde{A}_l} \left|l_2(x) - l_1(x)\right| \ge \delta^{-1}d_{2,l}^{-1} \implies |\alpha| \ge \delta^{-2/3}\,.$$

Since for $r(t) = t + \frac{\beta}{\alpha}t^2 \Rightarrow |r'(t)| \ge 1 - \frac{|2\beta|}{|\alpha|}t \ge 1 - \frac{\delta^{1/3}d_{2,l}}{2|A_l|} > 0$ we can apply the (non-)stationary phase method and deduce that
$$\left|I_{\alpha,\beta}(\phi)\right| \lesssim_n \alpha^{-n}, \quad \forall n \in \mathbb{N}\,.$$

As a consequence, we have that
$$\left|\mathcal{K}_l(x,s)\right| \lesssim \delta^n \underbrace{(\delta^{1/3}d_{1,l})^{-1}\chi_{\{|t| \le 2\delta^{1/3}d_{1,l}\}}(x-s)}_{u_{1,l}(x-s)} \quad \forall x \in \tilde{A}_l$$
$$\implies \left|\Phi_{2,l}\Phi_{1,l}v(x)\right| \lesssim \delta^n(u_{1,l} * |v|)(x) \lesssim \delta^n Mv(x)\,,$$
so (57) holds.

We now discuss the expression
$$\Omega_{j,l}f = T_{j,l}{}^*f - \Phi_{j,l}T_{j,l}{}^*f\,.$$

Keeping in mind the fact that Q_{b_j}, M_{c_j} are unitary we have the following chain of equalities:
$$\|\Omega_{j,l}f\|_2 = \|T_{j,l}{}^* - \Phi_{j,l}T_{j,l}{}^*f\|_2$$
$$= \|Q_{b_j}{}^*M_{c_j}{}^*T_{j,l}{}^*f - \tilde{\varphi}_{j,l}Q_{b_j}{}^*M_{c_j}{}^*T_{j,l}{}^*f\|_2$$
$$\overset{f := M_{c_j}Q_{b_j}h}{=} \|Q_{b_j}{}^*M_{c_j}{}^*T_{j,l}{}^*M_{c_j}Q_{b_j}h - \tilde{\varphi}_{j,l}Q_{b_j}{}^*M_{c_j}{}^*T_{j,l}{}^*M_{c_j}Q_{b_j}h\|_2\,.$$

Denote $\mathcal{T}_{j,l}^* = Q_{b_j}{}^*M_{c_j}{}^*T_{j,l}{}^*M_{c_j}Q_{b_j}$; then $\mathcal{T}_{j,l} = Q_{b_j}{}^*M_{c_j}{}^*T_{j,l}M_{c_j}Q_{b_j}$ and since $\|\Omega_{j,l}\|_2 = \|\Omega_{j,l}^*\|_2$ we have that
$$\|\Omega_{j,l}\|_2 = \|\mathcal{T}_{j,l} - \mathcal{T}_{j,l}\tilde{\varphi}_{j,l}\|_2\,.$$

Now fixing a tile $P = [\alpha, \omega, I] \in S_{j,l}$ (with $|I| = 2^{-k}$ for some $k \in \mathbb{N}$) we set
$$\mathcal{T}_{j,l}^P = Q_{b_j}{}^*M_{c_j}{}^*T_P M_{c_j}Q_{b_j}\,.$$

Then, for $x \in E(P) \subseteq I$ we have
$$\left|\mathcal{T}_{j,l}^P h(x) - \mathcal{T}_{j,l}^P\tilde{\varphi}_{j,l}h(x)\right| = \left|\int_{\mathbb{T}} e^{i\{y(l_x(x)-l_j(x)) - (b(x)-b_j)y^2\}}\varphi_k(y)[h - \varphi_{j,l}*h](x-y)dy\right|$$
$$\le \int_{\mathbb{T}} \left|(h\chi_{\tilde{A}_l})(x-y)\right|\left|r_x^P(y) - (r_x^P*\varphi_{j,l})(y)\right|dy$$

where
$$r_x^P(y) = e^{i\{y(l_x(x)-l_j(x)) - (b(x)-b_j)y^2\}}\varphi_k(y)\,. \tag{58}$$

Our next step is to provide an L^∞ bound on the expression
$$r_x^P(y) - (r_x^P*\varphi_{j,l})(y) = \int_r \hat{r_x^P}(\xi)\left(1 - \hat{\varphi_{j,l}}(\xi)\right)e^{i\xi y}dy\,.$$

For this we write
$$\left|\widehat{r_x^P}(\xi)\right| = \left|\int_r e^{i\{\frac{s}{2^k}[(l_x(x)-l_j(x))-\xi] - \frac{s^2}{2^{2k}}(b(x)-b_j)\}}\varphi(s)ds\right|\,,$$

and observe that $|l_x(x) - l_j(x)| \lesssim |\omega_P| \lesssim 2^k$; from this, since $|\xi| \gtrsim 2^k$, we can apply the method of (non-)stationary phase to obtain
$$\left|\widehat{r_x^P}(\xi)\right| \lesssim \left(1 + \frac{|\xi|}{2^k}\right)^{-n}, \quad \forall n \in \mathbb{N} \,\&\, x \in E(P)\,. \tag{59}$$

Using now (59) together with (53), we deduce that

$$\int_r \left|\widehat{r_x^P}(\xi)(1 - \widehat{\varphi_{j,l}}(\xi))\right| d\xi \lesssim \int_{|\xi| \le (\delta^{1/3} d_j)^{-1}} \left(1 + \frac{|\xi|}{2^k}\right)^{-n-2} (\delta^{1/3} d_{j,l})^n |\xi|^n d\xi$$

$$+ \int_{|\xi| > (\delta^{1/3} d_j)^{-1}} \left(1 + \frac{|\xi|}{2^k}\right)^{-n-2} d\xi \lesssim 2^k (2^k \delta^{1/3} d_{j,l})^n .$$

As a consequence we have that for any $P \in \mathcal{P}_j$ and $x \in E(P) \subseteq I_P$

$$\left|r_x^P(y) - (r_x^P * \varphi_{j,l})(y)\right| \lesssim (\delta^{1/3} 2^k d_{j,l})^n 2^k \chi_{[-2^{-k}, 2^{-k}]}(y) .$$

Then denoting $\mathcal{R}_{j,l}(y) = \sum_{2^k \le (d_{j,l})^{-1}} (\delta^{1/3} 2^k d_{j,l})^n 2^k \chi_{[-2^{-k}, 2^{-k}]}(y)$ we obtain $\|\mathcal{R}_{j,l}\|_1 \lesssim (\delta)^{n/3}$, and so

$$\|\mathcal{T}_{j,l} h - \mathcal{T}_{j,l} \tilde\varphi_{j,l} h\|_2^2 \lesssim \int_{\mathbb{T}} \left(\int_{\mathbb{T}} |h \chi_{\tilde{A}_l}(x-y)| \mathcal{R}_{j,l}(y) dy\right)^2 dx \le^{C-S}$$

$$\|\mathcal{R}_{j,l}\|_1^2 \int_{\mathbb{T}} |h\chi_{\tilde{A}_l}|^2 \lesssim \delta^{2n} \int_{\tilde{A}_l} |h|^2 .$$

Consequently, we have shown that

$$\|\Omega_{j,l}^* h\|_2 \lesssim \delta^n \left(\int_{\tilde{A}_l} |h|^2\right)^{1/2},$$

ending the proof of our claim.

Now, reformulating the previous statements, we have $\forall \in \{0, \ldots, k-1\}$

$$\begin{cases} \left\|\Omega_{j,l}^*(f)\right\|_2, \quad \left\|\Omega_{j,l}(f)\right\|_2 \lesssim_n \delta^n \left(\int_{\tilde{A}_l} |f|^2\right)^{1/2} \\ \left|\langle \Phi_{1,l} T_{1,l}^* f, \Phi_{2,l} T_{2,l}^* g\rangle\right| \lesssim_n \delta^n \left(\int_{\tilde{A}_l} |f|^2\right)^{1/2} \left(\int_{\tilde{A}_l} |g|^2\right)^{1/2} \end{cases} .$$

(Note that $\operatorname{supp} T_{j,l}$, $\operatorname{supp} T_{j,l}^*$, $\operatorname{supp} \Omega_{j,l}^*$, $\operatorname{supp} \Omega_{j,l} \subseteq \tilde{A}_l$.)

Now since

$$\langle T_{1,l}^* f, T_{2,l}^* g\rangle = \langle \Phi_{1,l} T_{1,l}^* f, \Phi_{2,l} T_{2,l}^* g\rangle + \langle \Omega_{1,l} f, \Phi_{2,l} T_{2,l}^* g\rangle$$

$$+ \langle \Phi_{1,l} T_{1,l}^* f, \Omega_{2,l} g\rangle + \langle \Omega_{1,l} f, \Omega_{2,l} g\rangle ,$$

we have that

$$|U| = \left|\sum_{l=0}^{k-1} \langle T_{1,l}^* f, T_{2,l}^* g\rangle\right| \le \sum_{l=0}^{k-1} \left|\langle \Phi_{1,l} T_{1,l}^* f, \Phi_{2,l} T_{2,l}^* g\rangle\right|$$

$$+ \sum_{l=0}^{k-1} \left|\langle \Omega_{1,l} f, \Phi_{2,l} T_{2,l}^* g\rangle\right| + \sum_{l=0}^{k-1} \left|\langle \Phi_{1,l} T_{1,l}^* f, \Omega_{2,l} g\rangle\right| + \sum_{l=0}^{k-1} \left|\langle \Omega_{1,l} f, \Omega_{2,l} g\rangle\right|$$

$$\lesssim \delta^n \sum_{l=0}^{k-1} \left(\int_{\tilde{A}_l} |f|^2\right)^{1/2} \left(\int_{\tilde{A}_l} |g|^2\right)^{1/2} + \sum_{l=0}^{k-1} \left\|\Omega_{1,l}(f)\right\|_2 \left(\int_{\tilde{A}_l} |g|^2\right)^{1/2}$$

$$+ \sum_{l=0}^{k-1} \left\|\Omega_{2,l}(g)\right\|_2 \left(\int_{\tilde{A}_l} |f|^2\right)^{1/2} + \sum_{l=0}^{k-1} \left\|\Omega_{1,l}(f)\right\|_2 \left\|\Omega_{2,l}(g)\right\|_2 \lesssim^{C-S} \delta^n \|f\|_2 \|g\|_2 .$$

The terms Y and Z can be treated similarly; we leave these details for the reader.

Now, it remains to estimate the term

$$V = \int_{\mathbb{T}} T_{1,k}^* f \, \overline{T_{2,k}^* g} \,.$$

Setting $B_c = \mathbb{T} \setminus \frac{1}{3} I_c$ we have that

$$V = \left\langle \chi_{1/3 I_c} T_{1,k}{}^* f, \chi_{1/3 I_c} T_{2,k}{}^* g \right\rangle + \left\langle \chi_{B_c} T_{1,k}{}^* f, \chi_{B_c} T_{2,k}{}^* g \right\rangle = A + B \,.$$

Clearly, only the second term requires some work; for this, we need first to introduce some adapted tools: for $j \in \{1, 2\}$ and φ as above, we define

$$\varphi_j(x) = \varphi(x) = \left(\delta^{1/2} |I_s|\right)^{-1} \varphi\left((\delta^{1/2} |I_s|)^{-1} x\right) \quad \& \quad \tilde{\varphi}_j(f) = \tilde{\varphi}(f) = \varphi * f$$

$$\Phi_j : L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R}) \quad \text{with } \Phi_j = M_{c_j} Q_{b_j} \tilde{\varphi}_j Q_{b_j}{}^* M_{c_j}{}^* \,,$$

and, finally, $\Omega_j f := \chi_{B_c} T_{j,k}{}^* f - \Phi_j \chi_{B_c} T_{j,k}{}^* f$.

Then for $j \in \{1, 2\}$, we have

$$\chi_{B_c} T_{j,k}^* f = \Phi_j \chi_{B_c} T_{j,k}^* f + \chi_{\frac{2}{3} I_c} \Omega_j f + \chi_{c\left(\frac{2}{3} I_c\right)} \Omega_j f \,.$$

Using the facts:

i. for $x \in B_c$ we have $|l_1(x) - l_2(x)| \gtrsim \delta^{-\epsilon} |\operatorname{supp} \varphi|^{-1}$,

ii. $\chi_{c\left(\frac{2}{3} I_c\right)} \Omega_j f = \chi_{c\left(\frac{2}{3} I_c\right)} T_{j,k}{}^* f - \chi_{c\left(\frac{2}{3} I_c\right)} \Phi_j T_{j,k}{}^* f$,

we can repeat the previous arguments and obtain

$$\|\Phi_j * g\|_2 \lesssim \|g\|_2$$

$$\left\| \chi_{\frac{2}{3} I_c} \Phi_j \chi_{B_c} T_{j,k}{}^* f \right\|_2 \lesssim \left\| \chi_{I_c} T_{j,k}{}^* f \right\|_2 = \left\| \chi_{I_c} T^{\mathcal{P}_j}{}^* f \right\|_2$$

$$\left\| \chi_{c\left(\frac{2}{3} I_c\right)} \Omega_j f \right\|_2 \lesssim \delta^n \|f\|_2$$

$$\left| \left\langle \Phi_1 \chi_{B_c} T_{1,k}{}^* f, \Phi_2 \chi_{B_c} T_{2,k}{}^* g \right\rangle \right| \lesssim_n \delta^n \left(\int_{\tilde{I}_0} |f|^2 \right)^{1/2} \left(\int_{\tilde{I}_0} |g|^2 \right)^{1/2} \,.$$

Putting these relations together we conclude

$$|B| \lesssim_n \delta^n \left(\int_{\tilde{I}_0} |f|^2 \right)^{1/2} \left(\int_{\tilde{I}_0} |g|^2 \right)^{1/2} + \left\| \chi_{I_c} T^{\mathcal{P}_1}{}^* f \right\|_2 \left\| \chi_{I_c} T^{\mathcal{P}_2}{}^* g \right\|_2 \,.$$

Since we trivially have

$$|A| \leq \left\| \chi_{I_c} T^{\mathcal{P}_1}{}^* f \right\|_2 \left\| \chi_{I_c} T^{\mathcal{P}_2}{}^* g \right\|_2 \,,$$

our proof is now complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

DEFINITION 6. *A tree $\mathcal{P}$ with top-representative $P_0 = [\alpha_0, \omega_0, I_0]$ is called* normal *if*

$$P = [\alpha, \omega, I] \in \mathcal{P} \Longrightarrow |I| \leq \frac{\delta^{100}}{K} |I_0| \quad \& \quad \operatorname{dist}(I, \partial I_0) > 20 \frac{\delta^{100}}{K} |I_0| \,.$$

(Here $K > 10$ is some fixed constant and ∂I_0 designates the boundary of I_0.)

OBSERVATION 6. *Notice that if $\mathcal{P}$ is a normal tree then $\operatorname{supp} T^{\mathcal{P}*} f \subseteq \{ x \in I_0 \mid \operatorname{dist}(x, \partial I_0) > 10 \frac{\delta^{100}}{K} |I_0| \}$.*

DEFINITION 7. *A* row *is a collection $\mathcal{P} = \cup_{j \in \mathbb{N}} \mathcal{P}^j$ of normal trees $\mathcal{P}^j$ with top-representatives $P_0^j = [\alpha_0^j, \omega_0^j, I_0^j]$ such that the $\{I_0^j\}$ are pairwise disjoint.*

LEMMA 3.   Let $\mathcal{P}$ be a row as above, let $\mathcal{P}'$ be a tree with top-representative $[\alpha_0', \omega_0', I_0']$ and suppose that $\forall j \in \mathbb{N}$, $I_0^j \subseteq I_0'$ and $\mathcal{P}^j, \mathcal{P}'$ are separated trees; for each j, denote by I_c^j the critical intersection interval between $\mathcal{P}^j$ and $\mathcal{P}'$.

Then for any $f, g \in L^2(\mathbb{T})$ and $n \in \mathbb{N}$ we have that

$$\left| \left\langle T^{\mathcal{P}'^*} f, T^{\mathcal{P}^*} g \right\rangle \right| \lesssim_n \delta^n \|f\|_2 \|g\|_2 + \left\| \sum_j \chi_{I_c^j} T^{\mathcal{P}'^*} f \right\|_2 \left\| \sum_j \chi_{I_c^j} T^{\mathcal{P}^j{}^*} g \right\|_2 .$$

Proof. First, observe that it is enough to show that, for a fixed j, we have

$$\left| \left\langle T^{\mathcal{P}'^*} f, T^{\mathcal{P}^j{}^*} g \right\rangle \right| \lesssim_n \delta^n \big(\|M(Mf)\|_{L_2(I_0^j)} + \big\| M(M(T^{\mathcal{P}'^*} f)) \big\|_{L_2(I_0^j)} \big) \|g\|_{L_2(I_0^j)}$$
$$+ \big\| \chi_{I_c^j} T^{\mathcal{P}'^*} f \big\|_2 \big\| \chi_{I_c^j} T^{\mathcal{P}^j{}^*} g \big\|_2 .$$

For simplicity, in what follows we will drop the index j. Repeating now the procedures from the previous lemma, we define the following objects: $\{A_l\}_{\{l \in \{0,\dots,k\}\}}$ – the dyadic decomposition with respect to the (abscissa of the) intersection point – x^i, $\mathcal{P} = \cup_{l=0}^k S_l$ – the partition of the tree $\mathcal{P}$ in the well-localized (with respect to the separation interval I_s) sets of tiles, and $\{T_l^*\}_{\{l \in \{0,\dots,k\}\}}$ the corresponding decomposition of $T^{\mathcal{P}^*}$ (so we have $T^{\mathcal{P}^*} = \sum_{l=0}^k T_l^*$). Also, for

$$d_l = \min \big\{ |I| \mid P = [\alpha, \omega, I] \in \mathcal{P} \cup \mathcal{P}' \ \& \ I \subseteq A_{l-1} \cup A_l \cup A_{l+1}, \ I \subseteq I_0 \big\}$$

define $\varphi_l, \tilde{\varphi}_l, \Phi_l$ and Ω_l as before. Finally, set

$$\Phi_l' = M_{c'} Q_{b'} \tilde{\varphi}_l Q_{b'}{}^* M_{c'}{}^* \ \& \ \Omega_l' f = T^{\mathcal{P}'^*} f - \Phi_l' T^{\mathcal{P}'^*} f .$$

Then, for $l \in \{0, \dots, k-1\}$, we have

$$\left\langle T^{\mathcal{P}'^*} f, T_l^* g \right\rangle = \left\langle \Phi_l' T^{\mathcal{P}'^*} f, \Phi_l T_l^* g \right\rangle + \left\langle \Phi_l' T^{\mathcal{P}'^*} f, \Omega_l g \right\rangle + \left\langle \Omega_l' f, T_l^* g \right\rangle . \qquad (60)$$

Now using the following relations (see the previous lemma):

$$\big| \Phi_l' \Phi_l h(x) \big| \lesssim_n \delta^n \big(u_l * |h| \big)(x)$$
$$\big| \Omega_l^* h(x) \big| \lesssim \big(\mathcal{R}_l * |h| \big)(x)$$

with $\|\mathcal{R}_l\|_1 \lesssim (\delta)^{n/3}$, where

$$\mathcal{R}_l(y) = \sum_{2^k \leq (d_l)^{-1}} (\delta^{1/3} 2^k d_l)^n 2^k \chi_{[-2^{-k}, 2^{-k}]}(y)$$

and

$$u_l(x) = (d^{1/3} d_l)^{-1} \chi_{\{|t| \leq 2\delta^{1/3} d_l\}}(x) ,$$

we deduce (recall that $\mathcal{P}$ is a normal tree) for the first two terms

$$\left| \left\langle \Phi_l' T^{\mathcal{P}'^*} f, \Phi_l T_l^* g \right\rangle \right| \lesssim_n \delta^n \left\langle M\big(T^{\mathcal{P}'^*} f\big), |T_l^* g| \right\rangle$$
$$\lesssim \delta^n \big\| M\big(T^{\mathcal{P}'^*} f\big) \big\|_{L^2(\tilde{A}_l \cap I_0)} \|g\|_{L^2(\tilde{A}_l \cap I_0)} \qquad (61)$$

and respectively,

$$\left| \left\langle \Phi_l' T^{\mathcal{P}'^*} f, \Omega_l g \right\rangle \right| \lesssim \left\langle \mathcal{R}_l * \big\{ \chi_{\tilde{A}_l \cap I_0} M(T^{\mathcal{P}'^*} f) \big\}, |g| \right\rangle$$
$$\lesssim \delta^n \big\| M(T^{\mathcal{P}'^*} f) \big\|_{L^2(\tilde{A}_l \cap I_0)} \|Mg\|_{L^2(\tilde{A}_l \cap I_0)} . \qquad (62)$$

We now treat the last term of the right-hand side of (60). For this, set first $\mathcal{P}_l' := \{ P = [\alpha, \omega, I] \in \mathcal{P}' \mid |I| \geq d_l \}$; then, for $x \in \operatorname{supp} T_l^*$, we have

$$\Omega_l' f(x) = T^{\mathcal{P}'^*} f(x) - \Phi_l' T^{\mathcal{P}'^*} f(x) = T^{\mathcal{P}_l'^*} f(x) - \Phi_l' T^{\mathcal{P}_l'^*} f(x)$$

and consequently we deduce
$$x \in \operatorname{supp} T_l^* \implies \left|\Omega_l'^* h(x)\right| \lesssim \left(\mathcal{R}_l * |h|\right)(x),$$
so
$$\left|\langle \Omega_l' f, T_l^* g\rangle\right| \lesssim \langle |f|, \mathcal{R}_l * |T_l^* g|\rangle \lesssim \delta^n \|Mf\|_{L^2(\tilde{A}_l \cap I_0)} \|g\|_{L^2(\tilde{A}_l \cap I_0)}. \tag{63}$$

Now, adding the relations (61)–(63), we obtain
$$\left|\langle T^{\mathcal{P}'^*} f, T_l^* g\rangle\right| \lesssim \delta^n \left(\|Mf\|_{L^2(\tilde{A}_l \cap I_0)} + \left\|M(T^{\mathcal{P}'^*} f)\right\|_{L^2(\tilde{A}_l \cap I_0)}\right) \|Mg\|_{L^2(\tilde{A}_l \cap I_0)}$$
and consequently from Cauchy–Schwarz we deduce
$$\left|\langle T^{\mathcal{P}'^*} f, T^{\mathcal{P}^*} g\rangle\right|$$
$$\lesssim_n \delta^n \left(\|Mf\|_{L^2(I_0)} + \left\|M(T^{\mathcal{P}'^*} f)\right\|_{L^2(I_0)}\right) \|Mg\|_{L^2(I_0)} + \left|\langle T^{\mathcal{P}'^*} f, T_k^* g\rangle\right| \tag{64}$$
where $T_k^* = \sum_{P \in S_k} T_P^*$ with $S_k = \{P \in \mathcal{P} \mid x^i \in 5\tilde{I}_P\}$.

Now, for the last term of the right-hand side of (8), we argue as follows:

Case 1: $\quad |I_s| \gtrsim \frac{\delta^{100}}{K} |I_0|$

In this situation we have no tile $P = [\alpha, \omega, I] \in \mathcal{P}$ such that $100I \cap I_s \neq \emptyset$, and consequently $T_k^* = 0$, so we have nothing to prove.

Case 2: $\quad |I_s| \lesssim \frac{\delta^{100}}{K} |I_0|$

Let $\mathcal{P}_i' = \{P \in \mathcal{P}' \mid x^i \in 5\tilde{I}_P\}$, $\dot{\mathcal{P}} = \mathcal{P}' \setminus (\mathcal{P}_n' \cup \mathcal{P}_i')$ where $\mathcal{P}_n' = \{P = [\alpha, \omega, I] \in \{\mathcal{P}' \setminus \mathcal{P}_i'\} \mid |I| \leq \frac{\delta^{100}}{K} |I_0|\}$; define $T_i'^* = \sum_{P \in \mathcal{P}_i'} T_P^*$; then obviously
$$T^{\mathcal{P}'^*} f = T^{\mathcal{P}_n'^*} f + T^{\dot{\mathcal{P}}^*} f + T_i'^* f.$$

For the first term, from Lemma 2, we deduce that
$$\left|\langle T^{\mathcal{P}_n'^*} f, T_k^* g\rangle\right| \lesssim_n \delta^n \|f\|_{L^2(I_0)} \|g\|_{L^2(I_0)}.$$

Now using that $P = [\alpha, \omega, I] \in \dot{\mathcal{P}} \Rightarrow |I| \gtrsim |I_s|$ and defining
$$\varphi(x) = \varphi(x) = \left(\delta^{1/2} |I_s|\right)^{-1} \varphi\left((\delta^{1/2}|I_s|)^{-1} x\right) \ \& \ \tilde{\varphi}(f) = \tilde{\varphi}(f) = \varphi * f$$
$$\Phi = M_c Q_b \tilde{\varphi} Q_b^* M_c^* \ \& \ \Phi' = M_{c'} Q_{b'} \tilde{\varphi} Q_{b'}^* M_{c'}^*$$
we can follow the general ideas presented above and show that
$$\left|\langle T^{\dot{\mathcal{P}}^*} f, T_k^* g\rangle\right| \lesssim_n \delta^n \left(\|Mf\|_{L^2(I_0)} + \left\|M\{\mathcal{M}(T^{\mathcal{P}'^*} f)\}\right\|_{L^2(I_0)}\right) \|Mg\|_{L^2(I_0)}$$
and
$$\left|\langle T_i'^* f, T_k^* g\rangle\right| \lesssim_n \delta^n \left(\|Mf\|_{L^2(I_0)} + \left\|M\{\mathcal{M}(T^{\mathcal{P}'^*} f)\}\right\|_{L^2(I_0)}\right) \|Mg\|_{L^2(I_0)}$$
$$+ \left\|\chi_{I_c} T^{\mathcal{P}'^*} f\right\|_2 \left\|\chi_{I_c} T^{\mathcal{P}^*} g\right\|_2,$$
where
$$\mathcal{M}(T^{\mathcal{P}'^*} f) \overset{\text{def}}{=} \sup_{m \in N} \Big| \sum_{\substack{P = [\alpha, \omega, I] \in \mathcal{P}' \\ |I| \geq 2^{-m}}} T_P^* f \Big|.$$

So to summarize, we proved that
$$\left|\langle T^{\mathcal{P}'^*} f, T^{\mathcal{P}^*} g\rangle\right| \lesssim_n \delta^n \left(\|Mf\|_{L^2(I_0)} + \left\|M\{\mathcal{M}(T^{\mathcal{P}'^*} f)\}\right\|_{L^2(I_0)}\right) \|Mg\|_{L^2(I_0)}$$
$$+ \left\|\chi_{I_c} T^{\mathcal{P}'^*} f\right\|_2 \left\|\chi_{I_c} T^{\mathcal{P}^*} g\right\|_2.$$

Now the conclusion follows if we add the observation that
$$\mathcal{M}(T^{\mathcal{P}'^*} f) \leq Mf + M(T^{\mathcal{P}'^*} f). \tag{65}$$

Indeed, we first see that (65) can be rewritten as

$$\mathcal{M}(\mathcal{T}^{\mathcal{P}'^*}f) \leq Mf + M(\mathcal{T}^{\mathcal{P}'^*}f) \tag{66}$$

where, as usual, $\mathcal{T}^{\mathcal{P}'^*} := Q_{b'_0}^* M_{c'_0}^* \mathcal{T}^{\mathcal{P}'^*} M_{c'_0}^* Q_{b'_0}^*$ and $l'(x) = c'_0 + 2b'_0 x$ is the central line associated with the top of $\mathcal{P}'$.

Now fix $m \in \mathbb{N}$, $x \in [0,1]$ and define a function $\phi \in C_0^\infty(\mathbb{R})$ with the following properties:

$$\begin{aligned} &\bullet *\quad \operatorname{supp}\phi \subset \{|x| \leq 2\} \\ &\bullet *\quad \int_{\mathbb{R}} \phi = 1 \ \& \ \phi \geq 0 \,. \end{aligned} \tag{67}$$

Let J be the dyadic interval having the properties $x \in J$ and $|J| = 2^{-m}$; set $\phi_J(x) := |J|^{-1}\phi(|J|^{-1}(x - c(J)))$; we want to estimate the expression

$$(A) = \left| \sum_{\substack{P \in \mathcal{P}' \\ |I_P| \geq |J|}} \mathcal{T}_P^* f(x) - \int_{\mathbb{R}} \phi_J(s) \mathcal{T}^{\mathcal{P}'^*} f(s) ds \right|$$

$$\leq \int_{\mathbb{R}} \phi_J(s) \left\{ \sum_{\substack{P \in \mathcal{P}' \\ |I_P| \geq |J|}} \left| \mathcal{T}_P^* f(x) - \mathcal{T}_P^* f(s) \right| \right\} ds + \left| \int_{\mathbb{R}} \phi_J(s) \left\{ \sum_{\substack{P \in \mathcal{P}' \\ |I_P| < |J|}} \mathcal{T}_P^* f(s) \right\} ds \right|$$

$$:= B + C \,.$$

We start by treating the first term; observe first, that with the notation (58) we have (up to conjugation) that

$$\mathcal{T}_P^* f(x) = \int_{Pi} r_y^P(x - y) f(y) \chi_{E(P)}(y) dy \,. \tag{68}$$

Relying on this, we further have

$$\left| \mathcal{T}_P^* f(x) - \mathcal{T}_P^* f(s) \right| \leq \int_{\mathbb{T}} \left| r_y^P(x - y) - r_y^P(s - y) \right| |f(y)| \chi_{E(P)}(y) dy \,.$$

Using now relation (59) (for $|I_P| = 2^{-k}$) we deduce

$$\left| r_y^P(x - y) - r_y^P(s - y) \right| = \left| \int_{\mathbb{R}} \hat{r}_y^P(\eta) \left(e^{i\eta(x-y)} - e^{i\eta(s-y)} \right) d\eta \right|$$

$$\lesssim \int_{\mathbb{R}} \left(1 + \frac{\eta}{2^k}\right)^{-n} \eta |x - s| d\eta \lesssim 2^{2k} |J| \,, \tag{69}$$

where we used the fact that $x, s \in 5J$ and $y \in E(P)$.

From the previous relations, we conclude

$$B \lesssim \int_{\mathbb{R}} \phi_J(s) \left\{ \sum_{\substack{P \in \mathcal{P}' \\ J \subseteq 3\tilde{I}_P}} \left(|I_P|^{-1}|J|\right) \int_{\mathbb{R}} |f(y)| \frac{\chi_{I_P}(y)}{|I_P|} dy \right\} ds \lesssim Mf(x) \,.$$

For the second term C we use the fact that ψ has the mean zero property. Indeed, we have

$$C = \left| \sum_{\substack{P \in \mathcal{P}' \\ |I_P| < |J|}} \int_{\mathbb{R}} f(y) \chi_{E(P)}(y) \left(\int_{\mathbb{R}} \phi_J(s) r_y^P(s - y) ds \right) dy \right|$$

$$\lesssim \sum_{\substack{P \in \mathcal{P}' \\ \tilde{I}_P \subset 10J}} \int_{\mathbb{R}} |f(y)| \chi_{E(P)}(y) \left(\int_{\mathbb{R}} |\hat{\phi}_J(\eta)\hat{r}_y^P(\eta)| d\eta \right) dy \,.$$

Now, for a fixed y, we argue as follows:

$$\sum_{\substack{P \in \mathcal{P}' \\ \tilde{I}_P \subset 10J}} \chi_{E(P)}(y) \left(\int_{\mathbb{R}} |\hat{\phi}_J(\eta) \hat{r}_y^P(\eta)| d\eta \right)$$

$$\lesssim \chi_{10J}(y) \int_{\mathbb{R}} |\hat{\phi}_J(\eta)| \left(1 + \sum_{k>m} 2^{-k} |\eta| \right) d\eta \lesssim |J| \chi_{10J}(y) \, .$$

Consequently,

$$C \lesssim M f(x)$$

and replacing the bounds for B and C in (A), we conclude

$$\left| \sum_{\substack{P \in \mathcal{P}' \\ |I_P| \geq |J|}} \mathcal{T}_P^* f(x) \right| \lesssim \left| \int_{\mathbb{R}} \phi_J(s) \mathcal{T}^{\mathcal{P}'^*} f(s) ds \right| + M f(x)$$

which is what we needed for (65) to hold. $\square$

Finally, we combine the previous results to prove that we can control the L^2 norm of the operator associated to a forest.

Proof of Proposition 2. Define $F = \bigcup_j \left\{ x \in I_j \mid \text{dist}(x, \partial I_j)) \leq 100 \frac{\delta^{100}}{K^2} |I_j| \right\} =^{def} \bigcup_j F_j \, .$

We will estimate our operator only on the complement of this set. This is safe since we can control the measure of the excised set as follows:

$$|F| \leq \sum_j |F_j| \lesssim \sum_j |I_j| \frac{\delta^{100}}{K^2} \lesssim \frac{\delta^{50}}{K} \, ,$$

where the last inequality is derived from hypothesis 3.

Now, on F^c, we intend to use the previous estimates obtained in Lemma 3, but before this, we are forced to create enough space to apply the separation results. (Here it is essential that our trees are "centered"; see Observation 3.a.) In the following procedure, we will assume that there is no tree $\mathcal{P}_j$ having two tiles with same time interval; if this is not the case, then we must have (for some j) the situation $P \in \mathcal{P}_j$ and $P_u \in \mathcal{P}_j$ (or $P_l \in \mathcal{P}_j$), in which case we take the union of these two tiles and consider it as a single tile, renamed P. Consequently, we start by removing a few tiles from each tree $\mathcal{P}_j$.

Let $\mathcal{P} = \cup_j \mathcal{P}_j$; for $M = \log(K^{100} \delta^{-100})$ denote

$\mathcal{P}^+ := \{P \in \mathcal{P} \mid \text{there is no chain } P < P_1 < \cdots < P_M \text{ with all } P_j \in \mathcal{P}\}$

and

$\mathcal{P}^- := \{P \in \mathcal{P} \mid \text{there is no chain } P_1 < P_2 < \cdots < P_M < P \text{ with all } P_j \in \mathcal{P}\} \, .$

Now, it is easy to see that each such set can be split into at most M subsets with no two comparable tiles inside the same subset. Consequently, using Proposition 1, we deduce that (as mentioned in section 3, η may change from line to line)

$$\left\| T^{\mathcal{P}^+} \right\|_2, \left\| T^{\mathcal{P}^-} \right\|_2 \lesssim M \delta^\eta \lesssim \delta^\eta \log K \, .$$

We remove all the above mentioned sets from our collection $\mathcal{P}$ and decompose this new set as follows:

$$\mathcal{P} = \cup_j \mathcal{P}_j^0 \quad \text{where} \quad \mathcal{P}_j^0 = \mathcal{P}_j \cap \mathcal{P} \, .$$

Now this modified collection $\mathcal{P}$ behaves much better than the initial one; indeed, we have

1. $\forall P = [\alpha, \omega, I] \in \mathcal{P}_j^0$, $|I| \leq \delta^{100}/K^{100}|I_j|$;
2. $\forall j \neq k$, the trees $\mathcal{P}_j^0$ and $\mathcal{P}_k^0$ are δ'-separated where $\delta' = \delta^{100}/K^{100}$.

Moreover, if we split each $\mathcal{P}_j^0 = \mathcal{P}_j^N \cup \mathcal{P}_j^C$, with

$$\mathcal{P}_j^C =^{def} \left\{ P = [\alpha, \omega, I] \in \mathcal{P}_j^0 \mid I \subseteq F_j \right\},$$

we conclude that $\{\mathcal{P}_j^N\}_j$ represents a collection of normal, δ'-separated trees, while for the remaining parts of the trees we have the relation

$$\operatorname{supp} T^{\mathcal{P}_j^C} \subset F_j.$$

Consequently, on F^c we have that

$$T^{\mathcal{P}} f = \sum_j T^{\mathcal{P}_j^N} f,$$

and so our conclusion reduces to

$$\left\| \sum_j T^{\mathcal{P}_j^N} f \right\|_2 \lesssim \delta^{\eta} \log K \|f\|_2. \tag{70}$$

Now we are ready to apply the results from Lemma 3. We start by dividing $\bigcup_j \mathcal{P}_j^N$ into a union of at most $K\delta^{-2}$ rows, $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_{K\delta^{-2}}$. This is done by using an easy maximal argument: choose from $\{I_j\}_j$ a collection of maximal (disjoint) dyadic intervals - call it r_1; after that, erase the set r_1 from the previous collection and repeat the same procedure with the remaining one obtaining a new set r_2; due to condition 3 in our hypothesis, we know that this procedure will end in at most $K\delta^{-2}$ steps; now take $\mathcal{R}_j$ to be the set containing all trees that have their top inside the set r_j. Now, denoting by $T^{\mathcal{R}_j}$ the operator associated with $\mathcal{R}_j$, we claim that

(C1) $\|T^{\mathcal{R}_j}\|_{\{2\to 2\}} \lesssim \delta^{1/2}$.

(C2) $\|T^{\mathcal{R}_k *} T^{\mathcal{R}_j}\|_{\{2\to 2\}} = 0$ if $k \neq j$.

(C3) $\|T^{\mathcal{R}_k} T^{\mathcal{R}_j *}\|_{\{2\to 2\}} \lesssim \left(\frac{\delta}{K}\right)^{10}$ if $k \neq j$.

If we accept this for the moment, then applying the Cotlar–Stein lemma we deduce that

$$\left\| \sum_j T^{\mathcal{P}_j^N} f \right\|_2 = \left\| \sum_j T^{\mathcal{R}_j} f \right\|_2 \lesssim \delta^{1/2} \|f\|_2.$$

This last relation trivially implies (70), ending our proof.

We now pass to the analysis of our claims; for (C1), we just remark that since $\mathcal{R}_j$ is a row all the trees that belong to it are spatially disjoint, which together with Lemma 1 implies our statement. (C2) is trivial since for $k \neq j$, the operators $T^{\mathcal{R}_k}$ and $T^{\mathcal{R}_j}$ live in disjoint parts of the unit interval. The only interesting claim is (C3). Fix k_0, j_0 and suppose that $j_0 < k_0$. To avoid working with double indices, we will use the following notation: let $\{\mathcal{A}_j\}_j$ be the tree-decomposition of $\mathcal{R}_{j_0}$ with top time intervals $\{A_j\}_j$ and $\{\mathcal{B}_k\}_k$ the trees corresponding to $\mathcal{R}_{k_0}$ and with top time intervals $\{B_k\}_k$. Since $j_0 < k_0$ (from the way we constructed our rows) we have that $A_j \cap B_k \neq \emptyset$ implies $B_k \subseteq A_j$. Given this fact, we may assume that there exists $\{n_l\}_l \subset \mathbb{N}$ a strictly increasing sequence of natural numbers ($n_0 = 1$) such that

$$A_j \supseteq \bigcup_{n_{j-1} \leq k < n_j} B_k.$$

Now, from the fact that our trees are normal, we have

$$\left\langle T^{\mathcal{R}_{j_0}*}f, T^{\mathcal{R}_{k_0}*}g \right\rangle = \sum_{j \geq 1} \left\langle T^{\mathcal{A}_j*}f, \sum_{n_{j-1} \leq k < n_j} T^{\mathcal{B}_k*}g \right\rangle,$$

where here f, g are two arbitrary functions in $L^2(\mathbb{T})$.

Now define $I_c^{j,k}$ to be the critical intersection interval associated with the trees $\mathcal{A}_j$ and $\mathcal{B}_k$, and let $I_c^j = \bigcup_{n_{j-1} \leq k < n_j} I_c^{j,k}$. Then, applying Lemma 3 for a fixed j, we deduce

$$\left| \left\langle T^{\mathcal{A}_j*}f, \sum_{n_{j-1} \leq k < n_j} T^{\mathcal{B}_k*}g \right\rangle \right| \lesssim_n (\delta')^n \|f\|_{L^2(A_j)} \|g\|_{L^2(A_j)}$$

$$+ \left\| \chi_{I_c^j} T^{\mathcal{A}_j*}f \right\|_2 \left\| \sum_{n_{j-1} \leq k < n_j} \chi_{I_c^{j,k}} T^{\mathcal{B}_k*}g \right\|_2, \quad (71)$$

where we recall that the trees $\mathcal{A}_j$ and $\mathcal{B}_k$ are δ'-separated with $\delta' = \delta^{100}/K^{100}$.

Now, using the relations (see Observation 5.b)

$$|I_c^j \cap \tilde{I}_P| \leq \frac{\delta^{25}}{K^{25}} |I_P|, \quad \forall P \in \mathcal{A}_j, \quad (72)$$

$$|I_c^{j,k} \cap \tilde{I}_P| \leq \frac{\delta^{25}}{K^{25}} |I_P|, \quad \forall P \in \mathcal{B}_k, \quad (73)$$

together with Lemma 4 below and the fact that all trees involved are normal, we have that

$$\left\| \chi_{I_c^j} T^{\mathcal{A}_j*}f \right\|_2 \lesssim \left(\tfrac{\delta}{K} \right)^{10} \|f\|_{L^2(A_j)}$$

and

$$\left\| \sum_{n_{j-1} \leq k < n_j} \chi_{I_c^{j,k}} T^{\mathcal{B}_k*}g \right\|_2 \lesssim \left(\tfrac{\delta}{K} \right)^{10} \|g\|_{L^2(A_j)}.$$

Now replacing the last two relations in (71), we conclude

$$\left| \left\langle T^{\mathcal{A}_j*}f, \sum_{n_{j-1} \leq k < n_j} T^{\mathcal{B}_k*}g \right\rangle \right| \lesssim \left(\tfrac{\delta}{K} \right)^{10} \|f\|_{L^2(A_j)} \|g\|_{L^2(A_j)},$$

which together with an easy orthogonality argument gives us relation (C3), completing our proof. □

Heuristically, the next result is a "dual statement" of Lemma 1.

LEMMA 4. *Let $\mathcal{P}$ be a tree with top-representative $P_0 = [\alpha_0, \omega_0, I_0]$; suppose also that we have a set $A \subseteq \tilde{I}_0$ with the property that*

$$\exists \delta \in (0,1) \quad s.t. \ \forall P = [\alpha, \omega, I] \in \mathcal{P} \text{ we have } |I^* \cap A| \leq \delta|I|. \quad (74)$$

Then $\forall f \in L^2(\mathbb{T})$ we have

$$\left\| \chi_A T^{\mathcal{P}*}f \right\|_2 \lesssim \delta^{1/2} \|f\|_2. \quad (75)$$

Proof. We start the proof with the following observation: if $l_0(x) = c_0 + 2b_0 x$ is the central line of P_0 then (75) is equivalent to

$$\left\| \chi_A Q_{b_0}^* M_{c_0}^* T^{\mathcal{P}*} M_{c_0} Q_{b_0} f \right\|_2 \lesssim \delta^{1/2} \|f\|_2.$$

Consequently, we may assume that the top frequency line l_{P_0} coincides with the real axis (indeed, for the general case, taking as usual $\mathcal{T}^{\mathcal{P}*} = Q_{b_0}^* M_{c_0}^* T^{\mathcal{P}*} M_{c_0} Q_{b_0}$, one may repeat the procedure appearing below, by using relations (68) and (69) in (78)).

Another observation is that from the structure of the intervals $\{I_P^*\}$ we know that even though they are not necessarily dyadic, each I_P^* can be written as a union of at most 4 dyadic intervals with the same length - call them $\{I_{P,j}\}_{j=1}^4$. With this done, set

$$S = \left\{ I_{P,j} \mid P \in \mathcal{P} \ \& \ j \in \{1, \ldots, 4\} \right\}.$$

Suppose now that $\mathcal{P}$ is a finite collection of tiles. Also, define $\mathcal{J}$ the collection of maximal dyadic intervals I with the property

$$\forall J \in S \quad \text{if } J \cap I \neq \emptyset \text{ then } I \subseteq J. \tag{$*$}$$

Set then $\tilde{\mathcal{J}}$ to be any dyadic partition of $[0,1]$ that contains $\mathcal{J}$. Now, by inspecting (75), we remark that we may consider $A \subset \operatorname{supp} T^{\mathcal{P}^*}$. Then from the maximality of $\mathcal{J}$ and (74), we deduce

$$\forall J \in \tilde{\mathcal{J}} \quad |A \cap J| \lesssim \delta |J|. \tag{76}$$

On the other hand, we also have

$$\chi_A T^{\mathcal{P}^*} f(x) = \sum_{J \in \tilde{\mathcal{J}}} \chi_{J \cap A} \left\{ \sum_{P \in \mathcal{P}} T_P^* f(x) \right\}.$$

Now our proof relies on the relation ($x \in J$ fixed, and $J \in \tilde{\mathcal{J}}$)

$$\left| T^{\mathcal{P}^*} f(x) - \frac{1}{|J|} \int_J T^{\mathcal{P}^*} f(s) ds \right| \lesssim \frac{1}{|J|} \int_J Mf(s) ds. \tag{77}$$

If we accept this for the moment, then, denoting

$$M_{\tilde{\mathcal{J}}} f(x) = \sum_{J \in \tilde{\mathcal{J}}} \chi_J(x) \sup_{J \subseteq I} \frac{1}{|I|} \int_I |f|(s) ds,$$

we have that

$$\left| T^{\mathcal{P}^*} f(x) \right| \lesssim M_{\tilde{\mathcal{J}}} \left(T^{\mathcal{P}^*} f \right)(x) + M_{\tilde{\mathcal{J}}} (Mf)(x).$$

Now based on (76), we see that the relation (16) is satisfied for $E_J = A \cap J$, and so we conclude

$$\left| \chi_A T^{\mathcal{P}^*} f(x) \right| \lesssim M_\delta \left(T^{\mathcal{P}^*} f \right)(x) + M_\delta (Mf)(x),$$

which combined with (17) implies (75).

We return now to (77). For fixed $J \in \mathcal{J}$ and $x \in J$ we have

$$\left| T^{\mathcal{P}^*} f(x) - \frac{1}{|J|} \int_J T^{\mathcal{P}^*} f(s) ds \right|$$

$$= \left| \frac{1}{|J|} \int_J \left\{ \sum_{\substack{P \in \mathcal{P} \\ |I_P| \geq |J|}} \int_{\mathbb{T}} \left[\varphi_k(x-y) - \varphi_k(s-y) \right] f(y) \chi_{E(P)}(y) dy \right\} ds \right| \tag{78}$$

$$\lesssim \frac{1}{|J|} \int_J \left\{ \sum_{2^{-k} \geq |J|} 2^k |J| Mf(s) \right\} ds \lesssim M_{\tilde{\mathcal{J}}} (Mf)(x),$$

and the proof of our claim is now complete. $\square$

9   Remarks

1. Using interpolation methods, one can show (see also [F, §8]) that the previous results can be extended to handle the L^p case $(1 < p < \infty)$; more exactly one can show that for any $f \in L^p(\mathbb{T})$ one has $|Tf(x)| < \infty$ a.e. $x \in \mathbb{T}$.

2. The general polynomial phase case requires further generalization of the tiles to curved regions in the time-frequency plane. We hope to address this subject in the future.

References

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

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

[L1]   M. Lacey, Carleson's theorem: proof, complements, variations, Publ. Mat. 48:2 (2004), 251–307.

[L2]   M. Lacey, Carleson's theorem with quadratic phase functions, Studia Math. 153 (2002), 249–267.

[LT1]   M. Lacey, C. Thiele, A proof of boundedness of the Carleson operator, Math. Res. Lett. 7:4 (2000), 361–370.

[LT2]   M. Lacey, C. Thiele, On Calderon's conjecture, Ann. of Math. 149 (1999), 475–496.

[LT3]   M. Lacey, C. Thiele, L^p bounds for the bilinear Hilbert transform, $p > 2$, Ann. of Math. 146 (1997), 693–724.

[S1]   E.M. Stein, On limits of sequences of operators, Ann. of Math. 74 (1961), 140–170.

[S2]   E.M. Stein, Oscillatory integrals related to radon-like transforms, Proceed. of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993) (1995), 535–551.

[SW]   E.M. Stein, S. Wainger, Oscillatory integrals related to Carleson's theorem, Math. Res. Lett. 8 (2001), 789–800.

Victor Lie, Department of Mathematics, UCLA, Los Angeles CA 90095-1555, USA
                                                              vilie@math.ucla.edu

and

Institute of Mathematics of the Romanian Academy, Bucharest, RO 70700, P.O. Box 1-764 Romania