THE POINTWISE CONVERGENCE OF FOURIER SERIES
(II)
STRONG $L^1$ CASE FOR THE LACUNARY CARLESON OPERATOR

VICTOR LIE

Abstract. We prove that the lacunary Carleson operator is bounded from $L \log L$ to $L^1$. This result is sharp.

The proof is based on two newly introduced concepts:
(1) the time-frequency regularization of a measurable set and
(2) the set-resolution of the time-frequency plane at $0$-frequency.

These two concepts will play the central role in providing a special tile decomposition adapted to the interaction between the structure of the lacunary Carleson operator and the corresponding structure of a fixed measurable set.

Another key insight of our paper is that it provides for the first time a simultaneous treatment of families of tiles with distinct mass parameters. This should be regarded as a fundamental feature/difficulty of the problem of the pointwise convergence of Fourier Series near $L^1$, context in which, unlike the standard $L^p$, $p > 1$ case, no decay in the mass parameter is possible.

1. Introduction

In this paper we continue the author’s dedicated work in investigating one of the central and oldest themes in the area of harmonic analysis that of the almost everywhere convergence of the Fourier Series.

The main result of our paper is the following:
Main Theorem. Let \( \{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N} \) be a lacunary sequence. Define the lacunary Carleson operator associated with \( \{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N} \) as

\[
C^{(n_j)}_{\text{lac}} : C^\infty(\mathbb{T}) \mapsto L^\infty(\mathbb{T})
\]

with

\[
C^{(n_j)}_{\text{lac}} f(x) := \sup_{j \in \mathbb{N}} \left| \int_{\mathbb{T}} e^{2\pi i n_j (x-y)} \cot(\pi (x-y)) f(y) \, dy \right|.
\]

Then the following holds:

\[
\|C^{(n_j)}_{\text{lac}} f\|_1 \leq C \|f\|_{L \log L},
\]

where here \( C = C(\{n_j\}_{j}) > 0 \) is constant independent of \( f \).

Moreover, this result is sharp.

The proof of this result relies on the following key relation:

\[
\| \chi_F C^*_{\text{lac}} g \|_{L^1} \leq C_1 \|F\| \log \frac{4}{|F|} \|g\|_\infty,
\]

for any \( F \subseteq \mathbb{T} \) measurable and any function \( g \in L^\infty(\mathbb{T}) \). Here \( C^*_{\text{lac}} \) stands for the adjoint lacunary Carleson operator and \( C_1 > 0 \) is an absolute constant.

In order to show (3) and thus prove the central part of our Main Theorem, we rely on three new insights within the generic problem of the pointwise convergence of the Fourier Series near \( L^1 \):

- we introduce the concept of the time-frequency regularization (TFR) of a set - this is based on an algorithm designed to identify structures within the set \( F \) in (3) and relies on embedding the level sets of the Hardy-Littlewood maximal operator applied to \( F \) into a union of subsets having what we call “uniform” structure, that translates, in particular, into the fact that each such subset can be represented as a union of same-length disjoint intervals that roughly have similar \( F \)-density.
- relying on the first item, we further introduce the fundamental concept of the set-resolution of the time-frequency plane (at 0-frequency) - this connects the structure of the lacunary Carleson operator, more precisely its corresponding tile decomposition, with the structure

---

1Recall here that a sequence \( \{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N} \) is called lacunary iff \( \lim_{j \to \infty} \frac{n_{j+1}}{n_j} > 1 \).

2Throughout this paper, for notational simplicity, we ignore the p.v. symbol.

3For a more detailed account of this, please see the statement of the Main Theorem B at the end of this section.

4This should be understood as the adjoint of the linearized Carleson operator - see (25), where the linearized function \( N(\cdot) \) is regarded as an arbitrary measurable function independent of the input function.

5Throughout the paper, all the sets consisting of real numbers are assumed to be measurable.

6If \( I \) is a given interval one should think at its \( F \)-density (or mass) as given by \( \approx \frac{|I \cap F|}{|I|} \).
of the set $F$. Our entire proof is indissolubly tied to this type of structure-analysis/compatibility.

- in the final/main part of our proof we develop a first approach to the *simultaneous treatment* of the tile families with distinct mass-parameter that further relies on the notion of tree foliation of the time-frequency plane introduced in Section 12.2.2. This approach to the mass-analysis of our tiles is a key aspect with which one must deal in order to get sharp results on the pointwise convergence of Fourier Series near $L^1$. Indeed, the independent treatment of the tile families having distinct mass parameters is only possible when one can get decay in the mass parameter, situation that is bound to the $L^p-$case with $p > 1$.

In a nutshell, this work aims to reveal some subtle connections between the structure of the underlying set $F$ and the corresponding structure of the frequencies of the tiles appearing in the time-frequency decomposition of the (lacunary) Carleson operator with the latter further reflected in the properties of the linearizing measurable function $N$ in (4). Based on our insight and current progress, we modify the classical $L \log L$-conjecture on the pointwise convergence of the Fourier Series near $L^1$ - see Section 1.2. - into two distinct, sharper\(^7\) conjectures that address separately the $L^{1,\infty}$-bound and the $L^1$-bound of the Carleson operator - for this please see Conjecture 1 and Conjecture 2, respectively. One should regard our main result in the present paper as a support and indication for Conjecture 2.\(^8\)

Finally, we reiterate our belief that any significant progress on either Conjecture 1 or 2 must eventually involve additive combinatoric techniques together with a very careful analysis of the structural properties of the set $F$ relative to the structure of the set of frequencies of the tiles in the decomposition of the Carlson operator. This will very likely rely on and develop some of the concepts introduced in this paper and listed in the above itemization. We plan to investigate these directions in our future work.

1.1. **Historical background.** The history of this problem was initiated by J. Fourier, who in his study on the heat propagation ([12]), had the idea of representing a (suitable) function as a (possibly infinite) superposition of sinususes and cosines at distinct frequencies.

The question of when such (pointwise) representations make sense is naturally related with the regularity properties of the function that is represented and was one of the main themes of research within the mathematical analysis during the 19th century and early 20th century. After successive investigations of key leading figures in mathematics such as Dirichlet, Cauchy and Riemann it came as a complete surprise when Du Bois Reymond ([3])
showed that there are continuous periodic functions that do not admit everywhere such a representation - that is, there are continuous functions that do not agree everywhere with their Fourier series representation. After H. Lebesgue ([11]) developed his theory of integration introducing the formalism about sets of (Lebesgue) “measure” zero, N. Luzin ([22]) was the first one to formulate a “reasonable” conjecture, that is: for every square-integrable function \( f \in L^2(\mathbb{T}) \) the correspondent Fourier Series converges Lebesgue almost everywhere to \( f \). Surprisingly, his student A. Kolmogorov ([16]) was able to show in 1922 that there exists an \( L^1(\mathbb{T}) \) -function whose corresponding Fourier series diverges (almost) everywhere.

From this point on the general belief was that Luzin’s conjecture must be false. After more than forty years of misconceptions, L. Carleson ([5]) proved that in fact Luzin’s conjecture is true. Shortly afterwards, R. Hunt ([15]) extended Carleson’s result to functions \( f \in L^p(\mathbb{T}) \) with \( 1 < p < \infty \). In this context, we mention that after Carleson’s result two more proofs of Luzin’s conjecture were provided: one due to C. Fefferman ([13]) that became a landmark in the newly developing area of time-frequency analysis, and much later, another one due to M. Lacey and C. Thiele ([23]).

1.2. The Carleson operator: formulation of the main question; the two fundamental conjectures. Having in mind this evolution of the topic, one naturally reaches the following general question (formally stated in [28]):

**Main Question (Heuristic).** What is the behavior of the (almost everywhere) pointwise convergence of the Fourier Series between the two known cases for the Lebesgue-scale spaces \( L^p(\mathbb{T}) \):

- \( p = 1 \), divergence of the Fourier Series (Kolmogorov)
- \( p > 1 \), convergence of the Fourier Series (Carleson-Hunt) ?

Following [28], we now introduce

**Definition 1.** Let \( Y \) be a r.i. (quasi-)Banach space. We say that \( Y \) is a \( C \)-space iff \( C_0 = C_0(Y) > 0 \) such that the Carleson operator defined by

\[
C : C^{\infty}(\mathbb{T}) \mapsto L^\infty(\mathbb{T})
\]

with

\[
Cf(x) := \sup_{N \in \mathbb{N}} \left| \int_{\mathbb{T}} e^{2\pi i N (x-y)} \cot(\pi (x-y)) f(y) \, dy \right|,
\]

obeys the relation

\[
\|Cf\|_{1,\infty} \leq C_0 \|f\|_Y \quad \forall \ f \in Y .
\]

\[9\]Moreover, there exists a continuous function for which its Fourier Series is divergent at a given point (and in fact at any rational point).

\[10\]Recall that the weak-\( L^1 \) quasinorm is given by \( \|f\|_{1,\infty} := \sup_{\lambda > 0} \lambda \{x \ | \ |f(x)| > \lambda \} \).
With this definition, the Main Question above can be reformulated as follows:

**Main Question (Formal)** Give a satisfactory description of the Lorentz (or more general r.i. quasi-Banach) spaces \( Y \subseteq L^1(T) \) that are also \( C^- \)spaces. If such exists, describe the maximal Lorentz \( C^- \)space \( Y_0 \).

Once at this point we mention the classical conjecture on the near-\( L^1 \) behavior/convergence of Fourier Series that goes back to Carleson’s work in [5]:

**The classical \( L \log L \) Conjecture.**

The Lorentz space \( L \log L \) is a \( C^- \)space. Equivalently, we have that there exists \( c > 0 \) absolute constant such that

\[
\| Cf \|_{1, \infty} \leq c \| f \|_{L \log L} \quad \forall f \in L \log L.
\]

Returning to the formal enounce of the Main Question, we are now formulating the two fundamental conjectures that we believe are characterizing the behavior of the Carleson operator. These should be regarded as refinements of the above classical conjecture.

The first conjecture, regards per se the problem of the largest Lorentz space where one has the almost everywhere convergence of the Fourier Series:

**Conjecture 1.** [The \( L^{1, \infty} \)-behavior of the Carleson operator]

The largest Lorentz space \( Y_0 \subseteq L^1(T) \) that is also a \( C^- \)space is given by

\[
Y_0 = L \sqrt{\log L}.
\]

In particular we have that

\[
C : L \sqrt{\log L} \to L^{1, \infty}(T).
\]

**Heuristic Conjecture 1:** The elements that hint toward the formulation of this conjecture rely on the nature of the cancelations appearing among the tiles that belong to a multi-tower (see Section 4 in [28] for the definition). Due to space limitation we will not provide here more details on this topic, but instead refer our reader to a result of T. Tao and J. Wright ([33]) that can serve as a very basic/rudimentary model for understanding the behavior of the Carleson operator restricted to a tower: if \( T \) is a convolution operator arising from a Marcinkiewicz multiplier then \( T : L \sqrt{\log L} \to L^{1, \infty} \) sharply.
In contrast with the above, one may also ask about the strong bounds for the Carleson operator. More precisely, by a symmetric reasoning with the one in the formulation of the Main Question above, we can ask: “what is the largest Lorentz space \( Y_1 \subseteq L^1(\mathbb{T}) \) such that

\[
(9) \quad C : Y_1 \mapsto L^1(\mathbb{T})?
\]

Heuristic Conjecture 2: The Carleson operator encompasses both the behavior of the Hardy-Littlewood maximal operator and that of the Hilbert transform. However, both these operators map sharply \( L_{\log L} \) into \( L^1 \). As a consequence, one naturally arises to the following

**Conjecture 2.** [The \( L^1 \)-behavior of the Carleson operator]

The largest Lorentz space \( Y_1 \subseteq L^1(\mathbb{T}) \) that obeys (9) is given by

\[
(10) \quad Y_1 = L \log L.
\]

Once at this point, let us briefly comment on the up to date progress on the above conjectures:

- regarding the \( L^{1,\infty} \)-behavior of the Carleson operator:
  - on the positive side, the best known result belongs to N. Antonov ([1]) who showed that \( Y = L \log L \log \log \log L \) is a Lorentz \( C \)-space. A bit later, Arias-de-Reyna ([4]) proved that \( Y \) can be enlarged to a rearrangement invariant quasi-Banach space denoted \( QA \) that strictly contains \( L \log L \log \log \log L \). However, it is worth saying that this result is only apparently stronger, since, in [8], the authors show that under mild conditions on the fundamental function \( \phi \) the correspondent largest Lorentz space \( \Lambda_{\phi} \) contained in \( QA \) is precisely given by Antonov’s space \( \Lambda_{\phi} = L \log L \log \log \log L \).
  - Previous near-\( L^1 \) results were obtained by Sjölin ([29]) and F. Soria ([30], [31]).
  - on the negative side, Konyagin ([19], [20]) proved that if \( \phi(u) = o(u \sqrt{\log u \log \log u}) \) as \( u \to \infty \) then the space \( \phi(L) = \Lambda_{\bar{\phi}} \) is not a Lorentz \( C \)-space, where here \( \bar{\phi}(t) := \int_0^t s \phi(\frac{s}{t}) \, ds \). Previous negative results were obtained by Chen ([7]), Prokhorov ([32]) and Körner ([21]).

- regarding the \( L^1 \)-behavior of the Carleson operator:

\(^{11}\)Throughout the paper, if \( \varphi : \mathbb{T} \to [0, \infty) \) is a function with the following properties: \( \varphi(0) = 0 \) with \( \varphi \) nondecreasing and concave we define the Lorentz space \( \Lambda_{\varphi} \) as the r.i. Banach space of all the measurable functions \( f \in L^0(\mathbb{T}) \) such that \( \|f\|_{\Lambda_{\varphi}} := \int_0^1 f^*(s) \, d(\varphi(s)) < \infty \). The function \( \varphi \) becomes now the fundamental function of \( \Lambda_{\varphi} \).
– the best and only known results are due to Sjölin ([29]) for the Walsh-Fourier case and to the author ([24]) for the case of Fourier Series. These results amount to saying that (9) holds for \( Y_1 = L(\log L)^2 \).

We end the commentary on the behavior of the full Carleson operator by mentioning that all the (best) positive results mentioned above were based on extrapolation techniques. Few years ago, by using an approach that relied entirely on time-frequency methods, ([24]), the author re-proved all these best known positive results by a unified approach.

1.3. The model problem - the lacunary Carleson operator: formulation of the main question; the two fundamental conjectures - analogues. Given the deepness and difficulty of these questions it is natural to search for relevant model problems. One of the most natural and prominent such models was born in the early 20th century once that the Littlewood-Paley theory of block-dyadic summation of the Fourier Series developed. In such a context, appeared naturally the question about the pointwise convergence of the partial Fourier Sums along lacunary subsequences.

Now, by analogy with our previous section presentation, we first introduce the following definition (initially developed in [28])

**Definition 2.** Let \( Z \) be a r.i. (quasi-)Banach space.

i) Assume \( \{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N} \) is a lacunary sequence. We say that \( Z \) is a \( \mathcal{C}_{L_{n_j}} \)-space iff \( \exists C_1 = C_1(Z, \{n_j\}_j) > 0 \) such that the \( \{n_j\}_{j \in \mathbb{N}} \) - lacunary Carleson operator defined by

\[
C_{\text{lac}}^{\{n_j\}_j} : \mathcal{C}^\infty(\mathbb{T}) \to L^\infty(\mathbb{T})
\]

with

\[
C_{\text{lac}}^{\{n_j\}_j} f(x) := \sup_{j \in \mathbb{N}} \left| \int_\mathbb{T} e^{2\pi i n_j(x-y)} \cot(\pi(x-y)) f(y) dy \right|
\]

obeys the relation

\[
\|C_{\text{lac}}^{\{n_j\}_j} f\|_{1, \infty} \leq C_1 \|f\|_Z \quad \forall \; f \in Z.
\]

ii) We say that \( Z \) is a \( \mathcal{C}_L \)-space iff \( Z \) is a \( \mathcal{C}_{L_{n_j}} \)-space for any lacunary sequence \( \{n_j\}_{j \in \mathbb{N}} \).

Moreover, throughout the paper, if \( Z \) is a \( \mathcal{C}_L \)-space, we will (often) express this as \(^{12}\)

\[
\|C_{\text{lac}} f\|_{1, \infty} \lesssim \|f\|_Z \quad \forall \; f \in Z,
\]

\(^{12}\)Given \( A, B > 0 \), throughout the paper, we will use the notations \( A \lesssim B \) and \( B \gtrsim A \) to specify that there exists \( C > 0 \) such that \( A \leq C B \) and \( B \leq C A \) respectively.
where here $C_{lac}$ stands for “the generic” lacunary Carleson operator. 

We can now formulate the analogue of the Main Question raised in the previous section:

**Main Model Question (Formal).** Give a satisfactory description of the Lorentz (or more general r.i. quasi-Banach) spaces $Z$ that are also $C_L$-spaces. If such exists, describe the maximal Lorentz $C_L$-space $Z_0$.

Following now line by line the analogy with the full Carleson operator case, one can formulate the analogues of the fundamental Conjectures 1 and 2. Thus, in the context of the largest Lorentz space for which one has the almost everywhere convergence for any lacunary subsequence we have:

**Model Conjecture 1.** [The $L^{1,\infty}$-behavior of the lacunary Carleson operator]

(S. Konyagin, ICM, Madrid 2006, [18].)

The largest Lorentz space $Z_0 \subseteq L^1(\mathbb{T})$ that is also a $C_L$-space is given by

$$Z_0 = L \log \log L.$$  

**Heuristic - Model Conjecture 1**: The first important observation is that if $\{n_j\}_{j \in \mathbb{N}}$ is a lacunary sequence then the corresponding characters $\{e^{2\pi in_j}\}_j$ behave similarly to a sequence of i.i.d random variables. As a consequence one has the “trigonometric analogue” of Hincin’s inequality under the name of Zygmund’s inequality: if $\{a_j\}_{j \in \mathbb{N}} \in l^2(\mathbb{N})$, $\{n_j\}_{j \in \mathbb{N}}$ lacunary, and $F \subseteq \mathbb{T}$ measurable then

$$\| \sum_{j=1}^{\infty} a_j e^{2\pi in_j x} \|_{L^1(F)} \lesssim |F| (\log \frac{4}{|F|})^{1/2} \|\{a_j\}_j\|_{l^2(\mathbb{N})}.$$  

Fixing now $n$, $N \in \mathbb{N}$ and assuming $|F| \approx 2^{-N}$, one has from (15)

$$\| \sum_{j=1}^{2^n} a_j e^{2\pi in_j x} \|_{L^1(F)} \lesssim |F| N^{1/2} \left( \sum_{j=1}^{2^n} |a_j|^2 \right)^{1/2},$$

while a trivial $L^1$-summation argument gives

$$\| \sum_{j=1}^{2^n} a_j e^{2\pi in_j x} \|_{L^1(F)} \lesssim |F| \left( \sum_{j=1}^{2^n} |a_j| \right).$$

---

13In (13), the implicit constant is allowed to depend on the specific choice of the lacunary sequence and on the space $Z$ but not on the function $f \in Z$.

14In this very succinct heuristic we will make use of the notations and definitions introduced in Sections 2 and 5.

15In the Appendix we will provide a nice short proof of Zygmund’s inequality based on the two newly introduced concepts of the time-frequency regularization of a set and of the set resolution of the time-frequency plane at a fix frequency.
From this we deduce that if $|a_j| \approx 2^{-n}$ for $1 \leq j \leq 2^n$, then
\[
\| \sum_{j=1}^{2^n} a_j e^{2\pi i n_j x} \|_{L^1(F)} \lesssim |F| \min\{1, N^{\frac{1}{2}} 2^{-\frac{n}{2}}\}.
\]

Conclude that if we were in an ideal setting with all the tiles $P$ appearing in the time-frequency portrait of the lacunary Carleson operator $C_{lac} \equiv T$ having uniform $F$–mass, then
\[
\| \sum_{n} \sum_{P \in \mathbb{P}_n} T_P g \|_{L^1(F)} \lesssim |F| \sum_n \min\{1, N^{\frac{1}{2}} 2^{-\frac{n}{2}}\} \|g\|_\infty 
\lesssim |F| \log \log \frac{4}{|F|} \|g\|_\infty,
\]
which would justify at least the hope for the restricted weak-type form of the Model Conjecture 1, that is
\[
\|C_{lac}(h \chi_F)\|_{1,\infty} \lesssim |F| \log \log \frac{4}{|F|} \|h\|_\infty.
\]

Turning now towards the strong $L^1$ bounds for the lacunary Carleson operator, by analogy with (9), one can ask: "what is the largest Lorentz space $Z_1 \subseteq L^1(T)$ such that
\[
C_{lac} : Z_1 \rightarrow L^1(T)?
\]

Heuristic - Model Conjecture 2: Notice that the same arguments served as a motivation for the full Carleson operator remain valid in the current setting: indeed, the lacunary Carleson operator $C_{lac}$ can also be thought of as subsuming the behavior of both the Hardy-Littewood maximal operator and that of the Hilbert transform.

With these, we have

**Model Conjecture 2.** [The $L^1$-behavior of the lacunary Carleson operator]

*The largest Lorentz space $Z_1 \subseteq L^1(T)$ that obeys (17) is given by
\[
Z_1 = L \log L.
\]

1.4. **Resolution of the model problem and of the corresponding conjectures. Main results.** Once at this point, we have the following remarkable fact: we can fully answer both of the above model conjectures, and moreover, provide the answer to the Main Model Question, thus completely solving the lacunary Carleson operator case.
1.4.1. Resolution of the Main Model Question and Model Conjecture 1. From author’s previous work we first have:

**Answer Main Model Question ([25], [28]):**

Define \( \varphi_0 : [0, 1] \to \mathbb{R}_+ \) as \( \varphi_0(s) := s \log \log \frac{17}{s} \log \log \log \log \frac{17}{s} \).

Let now \( \varphi : [0, 1] \to \mathbb{R}_+ \) be a non-decreasing concave function with \( \varphi(0) = 0 \). Then we have:

i) If \( \lim_{s \to 0^+} \frac{\varphi(s)}{\varphi_0(s)} > 0 \) then the Lorentz space \( \Lambda_\varphi \) is a \( \mathcal{C}_L \)-space.

ii) If \( \lim_{s \to 0^+} \frac{\varphi(s)}{\varphi_0(s)} = 0 \) then the Lorentz space \( \Lambda_\varphi \) is not a \( \mathcal{C}_L \)-space.

iii) If \( \lim_{s \to 0^+} \frac{\varphi(s)}{\varphi_0(s)} = 0 < \lim_{s \to 0^+} \frac{\varphi(s)}{\varphi_0(s)} \) then both scenarios are possible.

Moreover, letting \( W \) be the quasi-Banach space defined by\(^{16}\):

\[
W := \{ f : T \mapsto C \mid f \text{ measurable, } \|f\|_W < \infty \},
\]

where

\[
\|f\|_W := \inf \left\{ \sum_{j=1}^{\infty} (1 + \log j) \|f_j\|_1 \log \log \frac{4\|f_j\|_\infty}{\|f_j\|_1} \mid f = \sum_{j=1}^{\infty} f_j, \sum_{j=1}^{\infty} |f_j| < \infty \text{ a.e.} \right\}.
\]

we have

\[
\|C_{\text{lac}}(f)\|_{1,\infty} \lesssim \|f\|_W.
\]

and thus \( Z = W \) is a \( \mathcal{C}_L \)-space.

From this, we immediately deduce:

**Corollary 3. [Maximal characterization] ([28]):**

Let \( \varphi : [0, 1] \to \mathbb{R}_+ \) be a non-decreasing concave function with \( \varphi(0) = 0 \). Assume that there exists

\[
\lim_{s \to 0^+} \frac{\varphi(s)}{\varphi_0(s)} \in [0, \infty].
\]

Then the largest Lorentz \( \mathcal{C}_L \)-space \( \Lambda_\varphi \) for which \( \varphi \) obeys (20) is given by

\[
Z_0 = L \log \log L \log \log \log \log L.
\]

In particular, taking in (20) the function \( \varphi(s) = s \log \log \frac{4}{s} \), one further deduces that the Model Conjecture 1 is false.

Finally, we record

**Corollary 4. [Restricted weak-type does not imply weak-type]:**

The lacunary Carleson operator obeys the following:

- \( C_{\text{lac}} \) is a sublinear, translation invariant operator;
- \( C_{\text{lac}} \) is of restricted weak type \((L \log \log L, L^1)\) and hence (16) holds validating our heuristic from before.

\(^{16}\)Throughout the paper we will use the following convention: \( \log k \) stands for \( \log_2 k \).
• $C_{lac}$ is not of weak type $(L \log \log L, L^1)$.

This last corollary disproved the so-called “generalized Halo Conjecture”, conjecturing the equivalence of weak-type and restricted–weak-type boundedness for translation-invariant sublinear operators on Lorentz (Orlicz) spaces.

We mention here that previous results on the $L^1, \infty$-behavior of the lacunary Carleson operator (or its Walsh analogue) were obtained chronologically in [34], [17], [2], [18], [10], [25] and [9] (for more details on this, please consult [28]).

1.4.2. Resolution of the Model Conjecture 2. Main results. Regarding the Model Conjecture 2, no previous results were known besides the partial answer for the stronger case of the full Carleson operator provided in [29] (for the Fourier-Walsh case) and [24] (for the Fourier case). This however, is now settled through the present paper. Indeed, we have:

Main Theorem A. [Sharp $L^1$-strong bounds for $C_{lac}$] 
Model Conjecture 2 is true.

Our Main Theorem A above is a consequence of the following result:

Main Theorem B. [Restricted type version] 
Let $\{n_j\} \in \mathbb{N}$ be any given lacunary sequence. Then the following are true:

i) There exists $C_1 > 0$ such that for any $F \subseteq \mathbb{T}$ measurable and any function $g \in L^\infty(\mathbb{T})$ the adjoint lacunary Carleson operator $C_{lac}^*$ obeys

$$\left\| \chi_F C_{lac}^*(g) \right\|_{L^1} \leq C_1 |F| \log \frac{4}{|F|} \|g\|_\infty .$$

ii) There exist $C_2 > 0$ such that for any $\eta \in (0, 1]$ there exists $F(\eta) = F \subseteq \mathbb{T}$ with $|F| = \eta$ such that

$$\left\| C_{lac}(\chi_F) \right\|_{L^1} \geq C_2 |F| \log \frac{4}{|F|} .$$

1.5. Structure of the paper. In this final (sub)section of the introduction we detail the structure of the paper:

• following the discretization in [13], Section 2 takes care of the standard decomposition of the (lacunary) Carleson operator into a superposition of operators that are well-localized in the time-frequency plane (i.e. adapted to tiles of area one).
• in Section 3 we show that Main Theorem B implies Main Theorem A and also prove the easy part of Main Theorem B, that is part ii).
• Section 4 introduces the key concept of the time-frequency regularization (TFR) of a set.
• Section 5 unravels the first chapter of the tile discretization discussing the notions of a mass and $F -$mass of a tile as well as the ordering relation among tiles. The section ends with an algorithm describing the so called tree $*$-foliation of the time-frequency plane.
• in Section 6 we introduce the new key concept of the set resolution of the time frequency plane at a fix frequency which connects the structure of the set $F$ with that of the lacunary Carleson operator. Our reasonings relies on the TFR concept introduced in Section 4. This is the second (and final) chapter of our tile discretization.
• in Section 7 we split the proof of the Main Theorem B part i) into four theorems\textsuperscript{17} - Theorems 18, 19, 20 and 21. In the same section we state and prove the Main Lemma.
• in Section 8 we present the proof of Theorem 19 that can be reduced to the proofs of the $L^2$- boundedness of the Carleson operator and of the Main Lemma.
• in Section 9 we prove Theorem 20 by closely following the spirit of our approach in \cite{25}.
• Section 10 is the most technical component of our paper. It seeks to reduce the proof of Theorem 21 to that of Proposition 37. In this section we will use the entire machinery developed by the author in the previous papers that concern the Carleson operator as well as the tools introduced in the present paper with a key emphasis on the properties of the (second) tile discretization developed in Section 6.
• in Section 11 we treat the error term appearing in Proposition 37.
• Section 12 deals with the main term in Proposition 37. This is the part where we develop a new approach in understanding the subtle interaction between the structure of the linearizing function $N$ encapsulated in the properties of set $E := \{E(P)\}_{P \in \mathbb{P}}$ and the corresponding structure of the measurable set $F$ that helps us to treat simultaneously the family of tiles with distinct mass parameter as well as those with distinct $F-$mass parameter.
• Section 13 incorporates several final remarks.
• Section 14 constitutes the Appendix of our paper and provides some light into the motivation and relevance that are hidden behind the time-frequency regularization of a set, concept that, as we will see, reveals interesting connections between additive combinatorics and time-frequency analysis areas.

\textsuperscript{17}Theorem 18 turns out it can be reduced to the classical statement that the maximal Hilbert transform acts boundedly from $L \log L$ to $L^1$. 
2. Operator discretization - First stage

Since the maximal operator under discussion is nothing else than a lacunary version of the Carleson operator, as usual in such context, we will use time-frequency methods to analyze it.

Now, the study of our operator

\[(23) \quad C_{\text{lac}}^{\{n_j\}} f(x) \approx \sup_j |S_{n_j} f(x)| \quad \text{with} \quad f \in C^1(\mathbb{T}),\]

may be canonically reduced to the analysis of

\[(24) \quad Tf(x) := \sup_{j \in \mathbb{N}} \left| \int_{\mathbb{T}} \frac{1}{x - y} e^{2\pi i n_j (x - y)} f(y) \, dy \right| ,\]

where here \(\{n_j\}_j\) is a prescribed lacunary sequence of positive integers.

Applying Fefferman’s approach, \((13)\), we perform the following steps:

- firstly, linearizing our operator, we construct \(N : \mathbb{T} \to \{n_j\}_j\) measurable function such that

\[(25) \quad Tf(x) = \int_{\mathbb{T}} \frac{1}{x - y} e^{-2\pi i N(x)y} f(y) \, dy .\]

- next, using the dilation symmetry of the kernel, we decompose

\[\frac{1}{y} = \sum_{k \geq 0} \psi_k(y) \quad \forall \ 0 < |y| < 1 ,\]

where \(\psi_k(y) := 2^k \psi(2^k y)\) (with \(k \in \mathbb{N}\)) and \(\psi\) an odd \(C^\infty\) function such that \(\text{supp} \ \psi \subseteq \{y \in \mathbb{R} \mid 2 < |y| < 8\}\).

- deduce now that in the newly created context the following holds

\[(26) \quad Tf(x) = \sum_{k \geq 0} \int_{\mathbb{T}} e^{-2\pi i N(x)y} \psi_k(x - y) f(y) \, dy .\]

- given \(k \in \mathbb{N}\), we partition the time-frequency plane in tiles\(^{19}\) of the form

\[P = [\omega, I],\]

where \(\omega, I\) are dyadic intervals\(^{20}\) such that \(|\omega| = |I|^{-1} = 2^k\). With this done, we let \(\mathbb{P}(k)\) be the collection of tiles at scale \(k\) and define the collection of all the tiles be

\[(27) \quad \mathbb{P} = \bigcup_{k \in \mathbb{N}} \mathbb{P}(k) .\]

- to each tile \(P = [\omega, I] \in \mathbb{P}\) we associate a set encapsulating the “amount” of the graph of \(N\) contained in \(P\), that is

\[E(P) := \{x \in I \mid N(x) \in \omega\} .\]
finally, for $P = [\omega, I] \in \mathbb{P}(k)$, we define the operators

$$T_P f(x) = \left\{ \int_T e^{-2\pi i N(x) y} \psi_k(x - y) f(y) \, dy \right\} \chi_{E(P)}(x),$$

and conclude that

$$T f(x) = \sum_{P \in \mathbb{P}} T_P f(x).$$

(28)

Notice that if we think at $N : T \to \{n_j\}_j$ as a predefined measurable function, then, the above decomposition does not depend on the function $f$. Using this perspective will be enough to show that the bounds on $T$ do not depend on $N$.

**Observation 5.** Since our procedure will involve the support of the adjoint operators $\{T_P\}_{P \in \mathbb{P}}$ we first isolate an elementary piece $T_P$ (and the corresponding $T^*_P$) and briefly introduce several notations that we will use in our later reasonings:

For $P = [\omega_P, I_P] \in \mathbb{P}$ we set $c(I_P)$ the center of the interval $I_P$ and define

$$I_{P^*} = [c(I_P) - \frac{17}{2}|I_P|, c(I_P) - \frac{3}{2}|I_P|] \cup [c(I_P) + \frac{3}{2}|I_P|, c(I_P) + \frac{17}{2}|I_P|].$$

We then have the following properties:

(29) \hspace{1cm} \text{supp} T_P \subseteq I_P \quad \text{and} \quad \text{supp} T^*_P \subseteq I_{P^*}.

Notice that we can express the set containing the support of $T^*_P$ as

(30) \hspace{1cm} I_{P^*} = \bigcup_{r=1}^{14} I^r_{P^*},

with each $I^r_{P^*}$ a dyadic interval of length $|I_P|$.

We now set

(31) \hspace{1cm} \tilde{I}_P = [c(I_P) - \frac{17}{2}|I_P|, c(I_P) + \frac{17}{2}|I_P|],

and notice that $I_P \cup I_{P^*} \subset \tilde{I}_P = 17I_P$.

Finally, using a standard reasoning, we can reduce our analysis to the following situation that will apply from now on throughout the entire paper:

(32) \hspace{1cm} \text{if} \ P_1, P_2 \in \mathbb{P} \text{ s.t.} |I_{P_1}| \neq |I_{P_2}| \Rightarrow |I_{P_1}| \leq 2^{-10} |I_{P_2}| \text{ or } |I_{P_2}| \leq 2^{-10} |I_{P_1}|.

As a consequence, we deduce that if $P_1, P_2 \subset \mathbb{P}$ such that $I_{P_1} \subsetneq I_{P_2}$ then

(33) \hspace{1cm} \text{supp} T^*_P \cap \text{supp} T^*_Q = \emptyset.
3. Reducing our Main Theorem A to the Main Theorem B.

Proof of Main Theorem B part ii)

3.1. Main Theorem B ⇒ Main Theorem A. Assuming for the moment that Main Theorem B holds we have:

Part ii) trivially implies that no larger space than $L \log L$ can satisfy (17) while part i) proves that $L \log L$ indeed obeys (17). To see this last fact, we proceed as follows:

For each $l \in \mathbb{Z}$ define

$$F_l := \{ x \in T \mid |f(x)| \in [2^l, 2^{l+1}) \}.$$  

Then notice that we have

$$\|f\|_{L \log L} \approx \sum_{l \in \mathbb{Z}} 2^l |F_l| \log \frac{4}{|F_l|}. \quad (34)$$

After the linearization of the lacunary Carleson operator, we deduce that for a suitable $g \in L^\infty$ with $\|g\|_\infty = 1$, one has

$$\|C_{\text{lac}}^{(n_j)}(\chi_{F_l})\|_1 \geq |F| \log \frac{4}{|F|}. \quad (35)$$

3.2. Proof of the Main Theorem B, part ii). Fix $\eta \in (0, \frac{1}{2}]$. Take now $F = [\frac{1}{2} - \eta, \frac{1}{2}]$. We will show that

$$\|C_{\text{lac}}^{(n_j)}(\chi_F)\|_1 \gtrsim |F| \log \frac{4}{|F|}. \quad (36)$$

Choosing now in (25) the linearizing function $N(x) \equiv 0$ we deduce that $|C_{\text{lac}}^{(n_j)}(\chi_F)| \geq |H(\chi_F)|$ where here $H$ stands for the Hilbert transform.

Consequently, using standard duality, we have that

$$\|C_{\text{lac}}^{(n_j)}(\chi_F)\|_1 \geq \int_T |H(\chi_F)| \geq \int_T \chi_F H^*(\chi_{[\frac{1}{2}, 1]}) = \int_T \chi_F H(\chi_{[\frac{1}{2}, 1]}).$$

In the same spirit with (26), we write

$$H(\chi_{[\frac{1}{2}, 1]})(x) = \sum_{k \geq 0} H_k(\chi_{[\frac{1}{2}, 1]})(x) = \sum_{k \geq 0} \int_T \psi_k(x - y) (\chi_{[\frac{1}{2}, 1]})(y) \, dy. \quad (37)$$

$^2$For the sake of the philosophy behind the cancelation properties of the kernel $\frac{1}{y}$, we want to present a proof that avoids the precise computation of the Hilbert transform of a characteristic function of an interval.
Choosing now $k_F \in \mathbb{N}$ such that $2^{-k_F} \approx |F|$ and setting $H_{\leq k_F} := \sum_{k \leq k_F} H_k$ and $H_{> k_F} := \sum_{k > k_F} H_k$ we further have

\[
| \int_T \chi_F H(\chi_{[\frac{1}{2},1]}) | \geq | \int_T \chi_F H_{\leq k_F}(\chi_{[\frac{1}{2},1]}) | - | \int_T \chi_F H_{> k_F}(\chi_{[\frac{1}{2},1]}) | \\
\geq \sum_{k \leq k_F} \int_T \chi_F |H_{\leq k_F}(\chi_{[\frac{1}{2},1]}) | - |F|^{\frac{1}{2}} (\int_F |H_{> k_F}(\chi_{[\frac{1}{2},1]})|^2)^{\frac{1}{2}} \\
\geq k_F |F| - c |F| \approx |F| \log \frac{4}{|F|},
\]

where in the last line $c > 0$ is a suitable absolute constant derived from the operator norm $\| \cdot \|_{L^2 \rightarrow L^2}$ of the maximal Hilbert transform.

4. The time-frequency regularization of a set (at a fixed frequency)

In this section we introduce a new concept that stays at the foundation of our approach in the present paper and has the potential of becoming a useful tool in other related problems. For more about the usefulness of this concept and for a suggestive application of it, please see the Appendix.

Let $F \subseteq \mathbb{T}$ be a measurable set. In what follows we present the time-frequency regularization (TFR) of the set $F$ (relative to the 0-frequency).

4.1. The family of 0-frequency tiles. In our construction we need to consider special structured family of tiles at 0-frequency arising from the properties of the level sets of the Hardy-Littlewood maximal operator associated with the characteristic function of $F$. The tiles involved in these structures will be selected from within a particular family of tiles called from now on 0-frequency tiles. Indeed, $R = [\omega_R, I_R] \in \mathbb{P}$ is called a 0-frequency tile iff $0 \in \omega_R$. Notice that the family of all 0-frequency tiles - denoted throughout the paper with $\mathcal{R}$ - consists from precisely those tiles which are “sitting” on the real axis. In what follows, we will keep the letter $R$ for elements within the set $\mathcal{R}$, while $\mathcal{R}$ will be reserved for subfamilies of $\mathcal{R}$.

Observe from the very definition of a 0-frequency tile that any such $R \in \mathcal{R}$ is uniquely determined by $I_R$. We will from now on identify these two objects (in particular any dyadic time interval will give rise to a corresponding 0–frequency tile). Thus given $I \subseteq \mathbb{T}$ dyadic interval, we will often refer to the corresponding 0–frequency tile as $R(I)$.

4.2. Spacial decomposition according to the level sets of the Hardy-Littlewood Maximal function $M(\chi_F)$. Define $k_F := \lceil \log \frac{1}{|F|} \rceil + 1$ and for each $k \in \mathbb{N}$ with $k \leq k_F$ we let $\mathcal{I}_k$ be the collection of maximal dyadic intervals $I$ such that

\[
\frac{|F \cap I|}{|I|} > 2^{-k},
\]
and set
\[(39) \bar{I}_k = \bigcup_{I \in \mathcal{I}_k} I.\]

Notice that we have the following natural inclusion relation: for any \(1 < k < k_F\) and \(I \in \mathcal{I}_{k-1}\) there exists \(J \in \mathcal{I}_k\) such that \(I \subseteq J\) and hence \(\bar{I}_{k-1} \subseteq \bar{I}_k\).

4.3. TFR of \(F\): the algorithm. In what follows we present the algorithm of the time-frequency \(k\)-regularization of \(F\). This new concept introduced here will prove fundamental in the tile-discretization of our lacunary Carleson operator performed in Section 6.

Fix from now on \(k \in \mathbb{N}\), with \(1 < k \leq k_F\); also fix \(I \in \mathcal{I}_k\). The central iterative body of the algorithm is given by:

- **Input:** we are given a non-empty collection\(^{22}\) of dyadic intervals \(A \subseteq \mathcal{I}_{k-1}(I)\).
- **Output:**
  - a threshold frequency \(\alpha_A\);
  - three sets\(^{23}\) of dyadic intervals: \(B, B^U\) and \(B^L\).

- **Properties of the output:**
  1) *Selection of \(\alpha_A\):* From the set \(\{|J|\}_{J \in A}\) pick the smallest possible size \(|J_0|\) such that the following saturation condition is satisfied:

\[(40) \sum_{|J| \leq |J_0|} |J| \geq \frac{1}{2} \sum_{J \in A} |J|.\]

Set \(\alpha_A := |J_0|^{-1}\).
  2) *Selection of \(B\):*

\[(41) B := \left\{ J \subset I \mid J \text{ dyadic} \quad \exists J' \in A \text{ s.t. } J' \subset J \right\}.\]

3) *Selection of \(B^U\):*

\[(42) B^U := \{ J \in A \mid |J|^{-1} > \alpha_A \}.\]

4) *Selection of \(B^L\):*

\[(43) B^L := \{ J \in A \mid |J|^{-1} < \alpha_A \}.\]

With this done, we iterate the above central body of the algorithm as follows:

\(^{22}\)This will be specified at each step of our algorithm, see below.

\(^{23}\)It is possible for some of them to be empty. However if all the three sets are empty then the algorithm will stop at that step.
• **Step 1.** Initialize $\mathcal{A} := \mathcal{I}_{k-1}(I)$ and denote the output with:
  - $\alpha := \alpha_{\mathcal{A}}$;
  - $\mathcal{C}(I) := \mathcal{B}$;
  - $\mathcal{S}^U(I) := \mathcal{B}^U$;
  - $\mathcal{S}^L(I) := \mathcal{B}^L$.

• **Step 2.** Initialize $\mathcal{A} := \mathcal{S}^U(I)$ and denote the output:
  - $\alpha^U := \alpha_{\mathcal{A}}$;
  - $\mathcal{C}^U(I) := \mathcal{B}$;
  - $\mathcal{S}^{UU}(I) := \mathcal{B}^U$;
  - $\mathcal{S}^{UL}(I) := \mathcal{B}^L$.
  Then, we initialize $\mathcal{A} := \mathcal{S}^L(I)$ and denote the output:
  - $\alpha^L := \alpha_{\mathcal{A}}$;
  - $\mathcal{C}^L(I) := \mathcal{B}$;
  - $\mathcal{S}^{LU}(I) := \mathcal{B}^U$;
  - $\mathcal{S}^{LL}(I) := \mathcal{B}^L$.

• **Step r, $r \geq 3$.** We continue inductively. From step $r - 1$, for $s_j \in \{U, L\}$ with $j \in \{1, \ldots, r - 1\}$, we have
  - $2^{r-2}$ threshold frequencies of the form $\alpha^{s_1\ldots s_{r-2}}$;
  - $2^{r-2}$ sets of the form $\mathcal{C}^{s_1\ldots s_{r-2}}(I)$;
  - $2^{r-2}$ sets of the form $\mathcal{S}^{s_1\ldots s_{r-2}U}(I)$;
  - $2^{r-2}$ sets of the form $\mathcal{S}^{s_1\ldots s_{r-2}L}(I)$.

Now, we are ready to apply the main body of our algorithm: Initialize $\mathcal{A} := \mathcal{S}^{s_1\ldots s_{r-1}}(I)$ and denote the output with:
  - $\alpha^{s_1\ldots s_{r-1}} := \alpha_{\mathcal{A}}$;
  - $\mathcal{C}^{s_1\ldots s_{r-1}}(I) := \mathcal{B}$;
  - $\mathcal{S}^{s_1\ldots s_{r-1}U}(I) := \mathcal{B}^U$;
  - $\mathcal{S}^{s_1\ldots s_{r-1}L}(I) := \mathcal{B}^L$.

We will run this algorithm until first $r$ for which all the sets $\mathcal{C}^{s_1\ldots s_{r-1}}(I)$ are empty; this last fact is guaranteed by the harmless assumption that the original set $\mathcal{I}_{k-1}(I)$ can be represented by a finite union of dyadic intervals.

With these done, we define the following sets:

• for $r = 0$, we simply define $\mathcal{R}^0_k[I] := R(I)$,
and set
\[ R_k^{F,0} := \bigcup_{I \in \mathcal{I}_k} \mathcal{R}_k^0[I]. \]

- for \( r = 1 \), we let
  \[ R_k^1[I] := \{ R(J) \}_{J \in \mathcal{C}(I)}, \]
  and set
  \[ R_k^{F,1} := \bigcup_{I \in \mathcal{I}_k} \mathcal{R}_k^1[I]. \]

- for general \( r \geq 2 \), we let
  \[ R_k^{s_1 \ldots s_{r-1}}[I] := \bigcup_{J \in \mathcal{C}^{s_1 \ldots s_{r-1}}(I)} R(J), \]
  \[ R_k^r[I] := \bigcup_{j=1}^{r-1} \bigcup_{s_j \in \{U,L\}} R_k^{s_1 \ldots s_{r-1}}[I], \]
  and finally
  \[ R_k^F[I] := \bigcup_{r \geq 0} R_k^r[I]. \]

- for \( 1 < k \leq k_F \), we define the time-frequency \( k \)-regularization of \( F \) as the collection of \( \mathcal{R} \)-tiles given by:

  \[ R_k^F := \bigcup_{I \in \mathcal{I}_k} \mathcal{R}_k^F[I]. \]

  - for \( k = 1 \), we define the time-frequency \( 1 \)-regularization of \( F \) as

    \[ R_1^F := \bigcup_{I \in \mathcal{I}_1} R(I). \]

Finally the (global) time-frequency regularization (TFR) of \( F \) is defined as the subcollection of tiles in \( \mathcal{R} \) given by

\[ R^F := \bigcup_{1 \leq k \leq k_F} R_k^F. \]

### 4.4. Key properties of TFR of a set.

In this subsection we analyze several of the important properties that characterize the above construction.

For notational simplicity we will drop from now on the super-index \( F \) from the definition of the sets \( \mathcal{R} \) above. Also, from now on we will refer to \( R_k \setminus \mathcal{R}_k^0 \) as simply \( R_k^{>0} \).

We now present several definitions:

As before, fix \( k \in \mathbb{N} \) with \( 1 \leq k \leq k_F \) and further fix \( I \in \mathcal{I}_k \).
Definition 6. [Base of \( R \)]  
Assume \( R \in \mathcal{R}_k^>0 \). We define the base of \( R \), denoted by
\[
\mathcal{R}_k \, \text{,}
\]
the unique tile \( R' \in \mathcal{R}_k \) such that
\[
\text{(51) } R' \text{ is minimal relative to the inclusion } I_R \subseteq I_{R'} .
\]

Definition 7. [Children of \( \mathcal{R}_k \)]  
For \( R \in \mathcal{R}_k \) we set
\[
\mathcal{R}_k^{\text{chi}}(R) := \{ R' \in \mathcal{R}_k \ | \ R' = R \} .
\]

Definition 8. [Base-support \( \mathcal{R}_k^{s_1 \ldots s_r} [I] \)]  
We let the base-support of \( \mathcal{R}_k^{s_1 \ldots s_r} [I] \) be
\[
\mathcal{R}_k^{s_1 \ldots s_r} [I] := \{ R \ | \ \exists R' \in \mathcal{R}_k^{s_1 \ldots s_r} [I] \text{ s.t. } R = R' \} .
\]

Definition 9. Let \( \alpha \) be a given frequency and \( R \in \mathcal{R} \). We say
\begin{itemize}
  \item \( \alpha \in R \) iff \( \alpha \in \omega_R \);
  \item \( \alpha \notin R \) iff \( \alpha \in R \) and \( \alpha \notin R \).
\end{itemize}

Notice now that the following hold:
\begin{itemize}
  \item given \( \alpha \) frequency and \( R \in \mathcal{R}_k^{s_1 \ldots s_r} [I] \) such that \( \alpha \in R \) then \( \alpha \in R' \) for any \( R' \in \mathcal{R}_k^{s_1 \ldots s_r} [I] \); this is a direct consequence of the relation
\[
\forall R, R' \in \mathcal{R}_k^{s_1 \ldots s_r} [I] \text{ we have } |I_R| = |I_{R'}| .
\]
  \item if \( \mathcal{F} \) is a given finite set of frequencies we let
\[
\mathcal{F}[R] := \{ \alpha \in \mathcal{F} \ | \ \alpha \in R \} .
\]
Then, based on the previous item, we have that
\[
\mathcal{F}[R] = \mathcal{F}[R'] \quad \forall R, R' \in \mathcal{R}_k^{s_1 \ldots s_r} [I] .
\]

Definition 10. Let \( \mathcal{F} \) be a finite set of frequencies. From the above items, we notice that it makes sense to define
\[
\mathcal{F}(\mathcal{R}_k^{s_1 \ldots s_r} [I]) := \mathcal{F}[R] ,
\]
for some \( R \in \mathcal{R}_k^{s_1 \ldots s_r} [I] \).

With this, let us list several key properties of our construction:
\begin{itemize}
  \item The final products of the algorithm above, \textit{i.e.} (47), (48) and (49), are in fact \textit{generalized} (\textit{i.e.} non-convex) \textit{trees} of 0-frequency tiles relative to the standard order relation “ \( \leq \) ”.\footnote{Throughout the paper we adopt the following convention: when setting the parameter \( r \in \mathbb{N} \) we allow the value \( r = 0 \), and, in this case, we simply set
\[
\mathcal{R}_k^{s_1 \ldots s_r} [I] := \mathcal{R}_k^0 [I] \text{ and } \mathcal{R}_k^{s_1 \ldots s_r} [I] := \mathcal{R}_k^1 [I] .
\]}
\end{itemize}
• Let $R \in \mathcal{R}_k^{s_1 \ldots s_r}[I]$ and $R' \in \mathcal{R}_k^{s'_1 \ldots s'_{r'}}[I]$ with $r < r'$. Assume there exists a frequency such that $\alpha \in R$ and $\alpha \in R'$. Then, we must have that
  - $s'_1 = s_1, \ldots, s'_r = s_r$;
  - $s'_{r+1} = L$.
• Given $r < r'$ and assuming
  \[ \mathcal{R}_k^{s'_1 \ldots s'_{r'}}[I] \cap \mathcal{R}_k^{s_1 \ldots s_r}[I] \neq \emptyset, \]
  one must have
  - $s'_1 = s_1, \ldots, s'_r = s_r$;
  - $s'_j = L \quad \forall j \in \{r+1, \ldots, r'\}$.
• Based on the previous two items we deduce that if $F$ is a given set of frequencies, then
  \[ F(\mathcal{R}_k^{s_1 \ldots s_r}[I]) \supseteq \bigcup_{s'_1 = s_1, \ldots, s'_r = s_r} \bigcup_{s'_{r+1} = L} F[R]. \]
• Similarly, if $R, R'$ as in the second item, then assuming that there exists $x \in I_R$ and $x \in I_{R'}$ we must have
  - $s_1 = s'_1, \ldots, s_r = s'_r$;
  - $s_{r+1} = U$.
• The following geometric decay holds: if $r' \geq r+2$ and $s = (s_1, \ldots, s_r)$, $s' = (s'_1, \ldots, s'_{r'})$ given such that $s'_j = s_j$ for any $j \in \{1, \ldots, r\}$, then
  \[ \sum_{R' \in \mathcal{R}_k^{s'_1 \ldots s'_{r'}}[I]} |I_{R'}| \leq \frac{1}{2} \sum_{R \in \mathcal{R}_k^{s_1 \ldots s_r}[I]} |I_R|. \]

5. Tile Discretization I: mass, $F$–mass decompositions; tree foliations

In this section we proceed with the decomposition of the family of tiles $\{ P = [\omega_P, I_P] \}_{P \in \mathcal{P}}$ according to\textsuperscript{26}

- the $F$–mass of a tile - this depends on how much of the information carried by $F$ lies within the spacial support of a tile, i.e. given $P = [\omega_P, I_P] \in \mathcal{P}$ we classify our tiles depending on the quantity $A_F(P) \approx |F|/|I_P|$. 
- the mass ($F$–smooth adapted version) of a tile - this is a refined version of the original concept of a mass of a tile introduced by C. Fefferman in [13]. It combines the structural properties of the set $F$ with those of the assigned measurable function $N(x)$, with the latter component being further reflected into the properties of the

\textsuperscript{26}The explanations provided at each of the items in the listing below, are only to provide the general picture for each of the criteria. The more precise descriptions will follow immediately afterwards.
sets \( \{E(P)\}_{P \in \mathcal{F}} \). Roughly, the mass of a tile \( P \) should be imagined as represented by \( A(P) \approx \frac{|E(P)|}{|I_P|} \).

- the tree \(+\)-foliation of the time-frequency plane - based on the standard partial order relation among tiles “\( \leq \)" and on the concepts introduced in the items above, we will partition our collection of tiles into suitable sub-collections - called descending trees - that have uniform mass and \( F \)-mass parameters.

5.1. \( F \)-mass decomposition. Before presenting our first decomposition, we take advantage of the fact that we are in the lacunary case and thus that the frequencies \( \{n_j\}_j \subseteq \mathbb{N} \) have the property that there exists a constant \( \tilde{C} > 0 \) depending only on our choice of the lacunary sequence such that\(^{27}\)

\[
\sum_{j=1}^{k} n_j < \tilde{C} n_{k+1} \quad \text{for any} \ k \in \mathbb{N} .
\]

Now relation (60) implies in particular two key facts:

1. the 0 frequency plays a special role in the time-frequency decomposition of our operator;
2. the family of all tiles \( \mathcal{P} \) can be decomposed into three subsets

\[
\mathcal{P} = \mathcal{P}(0) \cup \mathcal{P}_{\text{cluster}} \cup \mathcal{P}_{\text{sep}},
\]

such that:

- \( \mathcal{P}(0) \) does not contribute at all to the time-frequency decomposition of \( T \) (all the \( P \in \mathcal{P}(0) \) have the property that \( E(P) = \emptyset \));
- \( \mathcal{P}_{\text{cluster}} \) consists of tiles that dilated by a fixed amount (depending only on \( \tilde{C} \) in (60)) intersect the real axis;
- \( \mathcal{P}_{\text{sep}} \) is given by a union of well separated trees, with each tile \( P = [\omega, I] \in \mathcal{P}_{\text{sep}} \) having the property that there exists \( j \in \mathbb{N} \) such that \( n_j \in \omega \).

We will now make our decomposition in (61) precise. For simplicity, assuming without loss of generality\(^{28}\) that

\[
n_k = \alpha^k \quad \text{with} \ \alpha \in \mathbb{R}_+, \ \alpha > 1 ,
\]

we define\(^{29}\)

\[
\mathcal{P}_{\text{cluster}} := \{ P = [\omega P, I_P] \in \mathcal{P} \mid 0 \in c(\alpha) \omega P \} ,
\]

\(^{27}\)In fact relation (60) turns out to be essentially equivalent with the lacunarity requirement; more precisely, we have that any lacunary sequence \( \{n_j\}_j \subseteq \mathbb{N} \) obeys (60) and, conversely, any (increasing) sequence of natural numbers obeying (60) can be decomposed as union of no more than \( \tilde{C} > 0 \) (depending only on \( \tilde{C} \)) lacunary (sub)sequences.

\(^{28}\)The sequence \( \{n_k\}_k \) lacunary implies \( \lim \inf_{k \to \infty} \frac{n_{k+1}}{n_k} = \alpha > 1 \).

\(^{29}\)Throughout this paper we will use the following standard notation: if \( I \) is an (open) interval having the center \( c \), then for any \( b > 0 \) we set \( bI := (c - \frac{b}{2}I, c + \frac{b}{2}I) \). The constant \( c(\alpha) \) in (63) can be taken as \( 10(1 + \lfloor \frac{1}{\alpha - 1} \rfloor) \) where here \( \lfloor x \rfloor \) stands for the eire part of \( x \in \mathbb{R} \).
\[ P_{\text{sep}} := \{ P = [\omega_P, I_P] \in \mathbb{P} \setminus \mathbb{P}_{\text{cluster}} | \{ n_j \}_j \cap \omega_P \neq \emptyset \} . \]

and

\[ P(0) := \mathbb{P} \setminus (\mathbb{P}_{\text{cluster}} \cup \mathbb{P}_{\text{sep}}) . \]

In what follows we will focus on the only truly relevant set of tiles, that is \( \mathbb{P}_{\text{sep}} \).

The tile discretization algorithm based on the time-frequency locations of the tiles in \( \mathbb{P}_{\text{sep}} \) relative to the information carried by the set \( F \) is related with the analysis in [24] and is now presented below:

- we set

\[ P^1 := \{ P = [\omega_P, I_P] \in \mathbb{P}_{\text{sep}} | \exists I \in \mathcal{I}_1 \text{ s.t. } 5I_P \subset 200I \& |I_P| \leq |I| \} . \]

- proceed by induction and assume we have constructed the set \( \mathbb{P}^{k-1} \) for a suitable \( k-1 < k_F \); let

\[ \mathbb{P}^{k-1} := \mathbb{P}_{\text{sep}} \setminus \bigcup_{l \leq k-1} \mathbb{P}^l , \]

and define

\[ \mathbb{P}^k := \{ P = [\omega_P, I_P] \in \mathbb{P}^{k-1} | \exists I \in \mathcal{I}_k \text{ s.t. } 5I_P \subset 200I \& |I_P| \leq |I| \} . \]

- notice now that we have the partition

\[ \mathbb{P}_{\text{sep}} = \bigcup_{1 \leq k \leq k_F} \mathbb{P}^k . \]

Finally, we record that from (64) and (68), one has

\[ \mathbb{P} = \mathbb{P}(0) \cup \mathbb{P}_{\text{cluster}} \cup \bigcup_{1 \leq k \leq k_F} \mathbb{P}^k . \]

We end this subsection with the following useful

Definition 11. (\( F \)-mass of a tile)

Let \( P \in \mathbb{P}^k \subset \mathbb{P}_{\text{sep}} \) where here \( k \in \{1, \ldots, k_F \} \).

We then define the \( F \)-mass of \( P \) as

\[ A_F(P) := 2^{-k} . \]

5.2. Partial ordering; Mass decomposition - an \( F \) smooth version;

Trees. Our main focus in this subsection will be on performing a mass decomposition of our family of tiles that is adapted to the structural properties of the lacunary Carleson operator and “behaves smoothly” relative to \( F \). In doing this we transform Fefferman’s mass into an adaptative concept relative to carefully chosen subfamilies of tiles. Though of different flavor, this modification shares some common features with the approach embraced by the author in defining the mass of a tile in [26].

We start by introducing the standard partial ordering among tiles (see [13]):
Definition 12. \textit{Partial ordering of the tiles}

Given any two $P = [\omega, I], P' = [\omega', I']$ in $\mathbb{P}$ we say that
\begin{equation}
P \leq P',
\end{equation}
iff 
$I \subseteq I'$ and $\omega \supseteq \omega'$.

We are now ready to introduce the main definition of this subsection:

Definition 13. \textit{Mass of a tile (F-smooth version)}

Let $P = [\omega, I] \in \mathbb{P}_{sep}$. From (68) we know that there exists a unique $k \in \{1, \ldots, k_F\}$ such that $P \in \mathbb{P}_k$.

We then define the mass of $P$ as
\begin{equation}
A(P) := \sup_{P' = [\omega', I'] \in \mathbb{P}_k, P \leq P'} \frac{|E(P')|}{|P|}.
\end{equation}

With this, we define
\begin{equation}
P_n := \{ P \in \mathbb{P}_{sep} \mid A(P) \in (2^{-n-1}, 2^{-n}] \},
\end{equation}
and deduce based on (68) that
\begin{equation}
\mathbb{P} = \mathbb{P}(0) \cup \mathbb{P}_{cluster} \cup \bigcup_{n \in \mathbb{N}} \mathbb{P}_n.
\end{equation}

Now, once we have constructed $\mathbb{P}_n$ and $\mathbb{P}_k$, it will be useful to consider for later the families
\begin{equation}
\mathbb{P}_n := \mathbb{P}_n \cap \mathbb{P}_k.
\end{equation}

With this, we notice the refinement of (74) in the form
\begin{equation}
\mathbb{P} = \mathbb{P}(0) \cup \mathbb{P}_{cluster} \cup \bigcup_{n \in \mathbb{N}} \bigcup_{k=1}^{k_F} \mathbb{P}_n.
\end{equation}

The next definition addresses the concept of a tree that should be regarded as encoding the time-frequency representation of a modulated, scaled (maximal) Hilbert transform:

Definition 14. \textit{Tree}

We say that a set of tiles $\mathcal{P} \subset \mathbb{P}$ is a tree with top $P_0$ if the following conditions hold:
1) $\forall P \in \mathcal{P} \Rightarrow P \leq P_0$;
2) if $P_1, P_2 \in \mathcal{P}$ and $P_1 \leq P \leq P_2$ then $P \in \mathcal{P}$.

We will need one more definition, as given by

Definition 15. \textit{Row}

We say that a set of tiles $\mathcal{P} \subset \mathbb{P}_{sep}$ is a row if there exists a collection of families of tiles $\{\mathcal{P}_j\}_{j \geq 1}$ such that
- each $\mathcal{P}_j$ is a tree with top $P_j = [\omega_j, I_j]$.
• for any \( j \neq k \) we have that \( I_j \cap I_k = \emptyset \);
• the set \( \mathcal{P} \) can be represented as

\[
\mathcal{P} = \bigcup_j P_j .
\]

We end this section by recording several useful facts:

**Observation 16.** 1) Remark that the simultaneous mass and \( F \)-mass partition of \( \mathbb{P}_{sep} \) conserves the convexity property of the trees on which their \( L^p \)-boundedness is heavily relying. More precisely, for \( 1 < k < k_F \), we have that

\[
|I \cap F| < 2^{-k+10} \quad \forall I \in CZ(I_{P,\min}^*)[0,1].
\]

In particular, deduce that if \( P = [\omega_P, I_P] \in \mathbb{P}_n^k \) then

\[
\frac{|I_P \cap F|}{|I_P|} < 2^{-k+10} .
\]

2) Let \( \mathcal{P} \in \mathbb{P}_n^k \) be a tree. Recalling the notations from Observation 5, we define \( I_{P,\min}^* \) to be the collection of minimum (relative to inclusion) time intervals \( \{I_P^*\} \subset \mathcal{P} \), and set \( CZ(I_{P,\min}^*)[0,1] \) the Calderon-Zygmund decomposition of the interval \([0,1]\) with respect to \( I_{P,\min}^* \).

Then, from (67) we deduce the following key property:

\[
\frac{|I \cap F|}{|I|} < 2^{-k+10} \quad \forall I \in CZ(I_{P,\min}^*)[0,1].
\]

3) Let \( \mathcal{P} \subset \mathbb{P}_n \) be any given collection of tiles. Then, following the spirit of [13] - and closer to our definition of mass - that of [26], one can use a relatively complex combinatorial procedure involving suitable tree selections partitioning \( \mathcal{P} \) to show that\(^{31}\)

\[
\|T_P f\|_2 \lesssim 2^{-\frac{n}{2}} \|f\|_2 .
\]

Since these type of estimates - often referred as \( L^2 \)-mass control - were treated extensively in [26] in a more difficult context, we will not give supplementary details here and take (81) as granted throughout the entire present paper.

\(^{30}\)We recall here that given a collection \( \mathcal{A} \) of dyadic intervals in \([0,1]\) we say that \( CZ(\mathcal{A})[0,1] \) is the Calderon-Zygmund decomposition of the interval \([0,1]\) with respect to \( \mathcal{A} \) iff \( CZ(\mathcal{A})[0,1] \) can be written as \( \mathcal{A} \cup \mathcal{B} \) a collection of dyadic intervals in \([0,1]\) such that: 1) \( \mathcal{A} \cup \mathcal{B} \) forms a partition of \([0,1]\); 2) for any \( I \in \mathcal{A} \) and \( J \in \mathcal{B} \) one has \( 2I \not\subset J \) and \( 2J \not\subset I \); 3) \( \mathcal{B} \) is a collection of maximal dyadic intervals obeying 1) and 2).

\(^{31}\)Further refinements of this decomposition were introduced in Section 5 of [26], though they are not necessary in our present situation. In particular, each family \( \mathbb{P}_n^k \) can be reduced to a \( BMO \)-forest of \( n^{th} \) generation - for the definition see Section 4 in [26].
5.3. Tree \( \ast \)-foliation of the time-frequency plane depending on the mass and \( F \)-mass parameters. In this section we intend to organize our family of tiles \( \mathcal{P}_{\text{sep}} \) into maximal trees having uniform mass and \( F \)-mass parameters according to a "descending foliation pattern" that focuses on the adjoint support of these trees.

We start now the description of our tree selection:

Since our focus is on the family of tiles \( \mathcal{P}_{\text{sep}} \) from our hypothesis and conventions we deduce that one can split the total family of separated tiles into maximal trees

\[
\mathcal{P}_{\text{sep}} = \bigcup_{l \in \mathbb{N}} \mathcal{P}(l),
\]

with each tree \( \mathcal{P}(l) \) living at a given frequency \( \omega_l \) such that the sequence \( \{\omega_l\}_l \) is a strictly increasing lacunary sequence with \( \omega_{l+1} \geq \alpha \omega_l \) where \( \alpha > 1 \) - recall the assumption (62).

Next, we decompose each tree \( \mathcal{P}(l) \) into maximal rows with uniform \( F \)-mass parameter, i.e.

\[
\mathcal{P}(l) = \bigcup_{m=1}^{k_F} \mathcal{P}(l)^{(m)},
\]

where each row

\[
\mathcal{P}(l)^{(m)} := \mathbb{P}^m \cap \mathcal{P}(l),
\]

represents a union of disjoint maximal trees of uniform \( F \)-mass \( \approx 2^{-m} \) all living at the same frequency \( \omega_l \).

With this done, we take each uniform \( F \)-mass row \( \mathcal{P}(l)^{(m)} \) and decompose it into maximal (disjoint) trees

\[
\mathcal{P}(l)^{(m)} = \bigcup_{a \geq 1} \mathcal{P}(l,a)^{(m)},
\]

with each \( \mathcal{P}(l,a)^{(m)} \) being a maximal tree of uniform \( F \)-mass \( \approx 2^{-m} \) and living at the given frequency \( \omega_l \).

Now, for each\(^{32} \) \( a \geq 1 \), we isolate \( \mathcal{P}(l,a)^{(m)} \) - assuming that this set of tiles is non-void - and decompose our tree in maximal sub-rows of uniform mass from the lowest to the largest value of the mass; to be more precise, we write:

\[
\mathcal{P}(l,a)^{(m)} = \bigcup_n \mathcal{P}(l,a)^{(n,m)}.
\]

Notice that under the assumption that \( \mathbb{P}_{\text{sep}} \) is a finite family of tiles we have that given \( m \in \mathbb{N} \) such that the LHS in (86) is nonempty, we have that the set \( \{n | \mathcal{P}(l,a)^{(n,m)} \neq \emptyset\} \) is a finite convex subset of the set of natural numbers.

\(^{32}\)Finitely many such values of \( a \in \mathbb{N} \).
Next, we take each non-empty maximal row $\mathcal{P}^n_{(l),a}$ and decompose it into maximal (disjoint) trees

$$\mathcal{P}^n_{(l),a} = \bigcup_b \mathcal{P}^n_{(l),a,b},$$

so that each $\mathcal{P}^n_{(l),a,b}$ is maximal tree with uniform $F$–mass $\approx 2^{-m}$, uniform mass $\approx 2^{-n}$ and living at the given frequency $\omega_l$.

Of key importance for us in what follows, will be the collection of tiles having a fixed frequency and a given mass parameter, that is

$$\mathcal{P}^n_{(l)} := \bigcup_{a,b,m} \mathcal{P}^n_{(l),a,b}.$$

Given now $l$ and $n$, the last stage of our decomposition regroups our constructed trees $\{\mathcal{P}^n_{(l),a,b}\}_{l,n,m,a,b}$ inside $\mathcal{P}^n_{(l)}$ following a “descending $\ast$–foliation pattern”, i.e. at the heuristic level - we iterate over maximal families of maximal trees so that inside each family, the trees within it have their corresponding adjoint support pairwise disjoint:

- let $P^n_{(l),a,b}$ be the top of the tree $\mathcal{P}^n_{(l),a,b}$. Let $\mathcal{P}^{n,\text{tot}}_{(l)}$ be the collection of all the trees $\{\mathcal{P}^n_{(l),a,b}\}_{m,a,b}$ and $T^n_l$ be the collection of their tops.

We now define the first tree-layer of our decomposition as

$$\mathcal{P}^{n,\text{[1]}}_{(l),\ast},$$

according to the following rule: a tree $\mathcal{P}^n_{(l),a,b} \in \mathcal{P}^{n,\text{[1]}}_{(l),\ast}$ iff its top $P := P^n_{(l),a,b} \in T^n_l$ has the property that there exists $x \in T$ obeying

$$x \in \tilde{I}_P;$$

- $P$ is maximal with this property, that is, if there exists $P_1 \in T^n_l$ with $x \in \tilde{I}_{P_1}$ then we must have $|I_{P_1}| \leq |I_P|$.

Let $T^{n,\text{[1]}}_{l,\ast}$ be the collection of all the tops of the trees inside $\mathcal{P}^{n,\text{[1]}}_{(l),\ast}$.

Notice that we do have the uniform intersection property

$$\sum_{P \in T^{n,\text{[1]}}_{l,\ast}} \chi_{\tilde{I}_P}(x) \leq 100.$$

Indeed, to see this, it is enough to prove that there do not exist

$$P_1, P_2, P_3 \in T^n_{l,\ast}[1] \text{ with } |I_{P_1}| > |I_{P_2}| > |I_{P_3}|,$$

such that

$$\tilde{I}_{P_1} \cap \tilde{I}_{P_2} \cap \tilde{I}_{P_3} \neq \emptyset.$$

Assuming by contradiction that (93) is violated and hence that there exists $x_0 \in \tilde{I}_{P_1} \cap \tilde{I}_{P_2} \cap \tilde{I}_{P_3}$ with $P_1, P_2, P_3$ obeying (92) we notice that based on convention (32) we must have that either $\tilde{I}_{P_1} \supset \tilde{I}_{P_2}$
or $\tilde I_{P_1} \cup \tilde I_{P_2} \supset \tilde I_{P_3}$. However none of these situations are consistent with the requirement $P_1, P_2, P_3 \in T_{l,t}^n[1]$.

Assume now that $P_{n,(m)}^{n,(m)} \in P_{n,(l)}^{n,(l)}$. Then we define

$$\tilde P_{n,(l),a,b}^{n,(m),[1]} := \left\{ P_{n,(l),a',b'}^{n,(m)} \in P_{n,(l)}^{n,(l)} \mid \tilde I_{P_{n,(l),a,b}^{n,(m)}} \cap \tilde I_{P_{n,(l),a',b'}^{n,(m)}} = \text{nontrivial interval} \right\}$$

We define now

$$\tilde P_{n,(l),a,b}^{n,[1]} := \bigcup_{P_{n,(l),a,b}^{n,(m)} \in \tilde P_{n,(l),a,b}^{n,(m),[1]}} P_{n,(l),a,b}^{n,(m),[1]}$$

and symmetrically

$$T_{l,t}^n[1] := \{ P_{n,(l),a,b}^{n,(m)} \mid P_{n,(l),a,b}^{n,(m)} \in P_{n,(l),a,b}^{n,[1]} \}$$

- Next, we update our collections $P_{n,(l),a,b}^{n,\text{tot}} := P_{n,(l),a,b}^{n,\text{tot}} \setminus \tilde P_{n,(l),a,b}^{n,[1]}$ and $T_{l,t}^n := T_{l,t}^n \setminus T_{l,t}^n[1]$.

We now have two possible situations: either $P_{n,(l),a,b}^{n,\text{tot}} = T_{l,t}^n = \emptyset$ case in which our algorithm stops or $P_{n,(l),a,b}^{n,\text{tot}} \cap T_{l,t}^n = \emptyset$ case in which we repeat our first step and construct the corresponding sets

$$P_{n,(l),a,b}^{n,[2]}, T_{l,t}^n[2], \tilde P_{n,(l),a,b}^{n,\text{tot},[2]}, \tilde P_{n,(l),a,b}^{n,[2]} \text{ and } T_{l,t}^n[2].$$

- We continue inductively this process so that - assuming that our algorithm didn’t stop earlier - at step $p$ we construct

$$P_{n,(l),a,b}^{n,[p]}, T_{l,t}^n[p], \tilde P_{n,(l),a,b}^{n,\text{tot},[p]}, \tilde P_{n,(l),a,b}^{n,[p]} \text{ and } T_{l,t}^n[p].$$

- Notice that from our construction, this algorithm ends after $p_0 \leq k_F + 1$ steps. Indeed, this is an immediate consequence of the following observation: given $p_1 \leq p_0$, $x \in \mathbb{T}$ and any $P_{n,(l),a',b'}^{n,(m_1)},[p_1] \in P_{n,(m_1),[p_1]}^{n,\text{tot}}$ (non-void) such that $x \in \tilde I_{P_{n,(l),a',b'}^{n,(m_1)},[p_1]}^{n,\text{tot}}$ then for any $1 \leq p < p_1$ we have that there exists $P_{n,(l),a',b'}^{n,(m_1),[p]}$ such that $x \in \tilde I_{P_{n,(l),a',b'}^{n,(m_1),[p]}}^{n,\text{tot}}$ and $m_1 < m$.

We end this section with a useful notation: in order to identify the layer to which a given tree $P_{n,(l),a,b}^{n,(m)} \in P_{n,(l),a,b}^{n,\text{tot}}$ belongs we write

$$P_{n,(l),a,b}^{n,\text{tot},[p]} := P_{n,(l),a,b}^{n,(m)} \ \text{iff} \ P_{n,(l),a,b}^{n,(m)} \in P_{n,(l),a,b}^{n,[p]}$$

Notice that this is a well defined notion and that for each $P_{n,(l),a,b}^{n,(m)}$ there exists a unique $p$ for which (99) holds.
6. Tile Discretization II: the set resolution of the time-frequency plane at a fix frequency

In this section we will perform a final decomposition of our tiles depending on the deep geometric and additive combinatoric properties of the set $F$ encapsulated in the time-frequency regularization of the set $F$ developed in Section 4.3. The procedure described here is a major improvement over the initial approach in [25] that was designed to treat the $L^{1,\infty}$ behavior of the lacunary Carleson operator.\(^{33}\)

- Fix $1 < k < k_F$ and $I \in \mathcal{I}_k$. Let $R \in \mathcal{R}^{0}_k$ such that $R < R(I)$.

Then, we define the $(R, 2)$—family of tiles as

$$\mathcal{P}^{2}_{k}[R] := \left\{ P = [\omega_P, I_P] \in \mathbb{P}^{>k} \mid \begin{array}{l} \hat{I}_P \cap I \neq \emptyset \\ \omega_P \cap \omega_R = \emptyset \end{array} \right\}. \quad (100)$$

**Observation 17.** Let $R \in \mathcal{R}_k$ with $I_R \subseteq I$ and $I \in \mathcal{I}_k$. Set

$$\mathcal{R}^{s_1\ldots s_r}_{k,\chi}[R] := \mathcal{R}^{\chi}[R] \cap \mathcal{R}^{s_1\ldots s_r}_k[I]. \quad (101)$$

Notice that if $R, R' \in \mathcal{R}^{s_1\ldots s_r}_{k,\chi}[R]$ then $|I_R| = |I_{R'}|$. Deduce from this that

$$\mathcal{P}^{2}_{k}[R] = \mathcal{P}^{2}_{k}[R']. \quad (102)$$

- Let $1 < k < k_F$ and $I \in \mathcal{I}_k$. Assume $R \in \mathcal{R}^{0}_k$ with $R = R(I)$.

We then define the $(R, 1)$—family of tiles as

$$\mathcal{P}^{1}_{k}[R] := \left\{ P = [\omega_P, I_P] \in \mathbb{P}^{k+1} \mid \begin{array}{l} I_R \cap \hat{I}_P \neq \emptyset \\ \omega_P \cap \omega_R \neq \emptyset \end{array} \right\}. \quad (103)$$

- for the limiting case $k = k_F$ and any $R \in \mathcal{R}^{F}_{k_F}$ we set

$$\mathcal{P}^{1}_{k_F}[R] \equiv \mathcal{P}^{2}_{k_F}[R] = \emptyset. \quad (104)$$

- we will also introduce two extra sets of tiles for the limiting case $k = 1$, corresponding to the rectangles $R \in \{R(I)\}_{I \in \mathcal{I}_1}$:

$$\mathcal{P}^{2}_{1}[R] := \left\{ P = [\omega_P, I_P] \in \mathbb{P}^{sep} \mid \begin{array}{l} \hat{I}_P \cap I_R \neq \emptyset \\ \omega_P \cap \omega_R = \emptyset \end{array} \right\}. \quad (104)$$

and

$$\mathcal{P}^{1}_{1}[R] := \left\{ P = [\omega_P, I_P] \in \mathbb{P}^{1} \cup \mathbb{P}^{2} \mid \hat{I}_P \cap I_R \neq \emptyset \& \omega_P \cap \omega_R \neq \emptyset \right\}. \quad (105)$$

- finally, we let $\mathbb{P}[F, 0]$ be the collection of tiles

$$P = [\omega_P, I_P] \in \mathbb{P}^{sep} \text{ such that } |\hat{I}_P \cap F| = 0. \quad (106)$$

\(^{33}\)The origin of the tile decomposition here traces back to the so called $(f, \lambda)$—lacunary decomposition introduced by the author in [25]. If properly modified/extended this tile-decomposition has the potential of becoming a very useful tool in approaching other time-frequency questions asking about end-points behavior of various modulation-invariant operators. We plan to make this explicit in a future work.
Define now the following

- for $1 < k < k_F$ we let
  \[ \mathbb{P}[\mathcal{R}_k] := \bigcup_{R \in \mathcal{R}_k^1} \mathcal{P}_1^1[R] \cup \bigcup_{R \in \mathcal{R}_k^{>0}} \mathcal{P}_2^2[R] . \]  

- for $k = 1$ we let
  \[ \mathbb{P}[\mathcal{R}_1] := \bigcup_{R \in \mathcal{R}_1} \mathcal{P}_1^1[R] \cup \mathcal{P}_2^2[R] . \]  

- for $k = k_F$ we simply set
  \[ \mathbb{P}[\mathcal{R}_{k_F}] := \emptyset . \]  

Notice that with these done we have that
\[ \mathbb{P} = \mathbb{P}(0) \cup \mathbb{P}_{\text{cluster}} \cup \mathbb{P}[F,0] \cup \bigcup_{1 \leq k < k_F} \mathbb{P}[\mathcal{R}_k] . \]

As in [25], we notice that (110) does not express $\mathbb{P}$ as a disjoint union (partition) of sets.

However, as we will see later, when transferred into operator language with a further spacial localization, our tile decomposition behaves as good as a partition.

We end this section with an adaptation of our construction in Section 5.3 to a generic $(\mathcal{R}, 2)$–family $\mathcal{P}^2_k[R]$.

For this, we first fix $1 \leq k < k_F$ and assume throughout what follows that $l$ is fixed and $i \in \{1, 2\}$. Setting now
\[ \mathcal{P}_{i,k}[R] := \mathcal{P}_i^1[R] \cap \mathcal{P}_k[0] , \]
we remark that
\[ \mathcal{P}_{i,k}[R] := \bigcup_l \mathcal{P}_{i,k,l}[R] . \]

With this, we define the following families of tiles:

- using (84), we let
  \[ \mathcal{P}_{i,k,l}^{(m)}[R] := \mathcal{P}_{i,k,l}^i[R] \cap \mathcal{P}_k^{(m)}[l] ; \]

- using (88), we let
  \[ \mathcal{P}_{i,k,l}^{n,(m)}[R] := \mathcal{P}_{i,k,l}^i[R] \cap \mathcal{P}_k^{n,(m)}[l] . \]

Now for the specific case $i = 2$ we will need a more refined analysis of our families of tiles. For this reason, we need to define the following

- using (87), we let
  \[ \mathcal{P}_{i,k,l,a,b}^{2, n,(m)}[R] := \mathcal{P}_{i,k,l}^2[R] \cap \mathcal{P}_{i,k,l,a,b}^{2, n,(m)}[R] ; \]
using (99) and then the analogue of (95), we let
\[ P_{k(l),a,b}^{2,n,(m),[p]}[R] := P_{k(l),a,b}^{2}[R] \cap P_{(l),a,b}^{n,(m),[p]} \]
and
\[ \mathcal{P}_{k(l),a,b}^{2,n,(m),[p]}[R] := P_{k(l),a,b}^{2}[R] \cap \mathcal{P}_{(l),a,b}^{n,(m),[p]} \]

finally, using (98), we let
\[ P_{k(l)}^{2,n,[p]}[R] := P_{k(l)}^{2}[R] \cap \mathcal{P}_{(l),a,b}^{n,[p]} \] and
\[ \mathcal{P}_{k(l)}^{2,n,[p]}[R] := P_{k(l)}^{2}[R] \cap \mathcal{P}_{(l),a,b}^{n,[p]} \]

With this we set \( P_{i,n}^{i,n,[R]} := P_{i}^{i,[R]} \cap \mathbb{P}_{n} \) and notice that
\[ P_{i,n}^{i,n,[R]} = \bigcup_{l} P_{k(l)}^{i,n,[R]} \]

In the case \( i = 2 \), we further have
\[ P_{k}^{2,n,[R]} = \bigcup_{l,p} \mathcal{P}_{k(l)}^{2,n,[p]}[R] = \bigcup_{l,m,p} \mathcal{P}_{k(l),a,b}^{2,n,(m),[p]}[R] \]

Throughout the paper, we only retain those trees \( \mathcal{P}_{k(l),a,b}^{2,n,(m),[p]}[R] \) and families \( \mathcal{P}_{k(l)}^{2,n,[p]}[R], \mathcal{P}_{k(l),a,b}^{2,n,(m),[p]}[R] \) that are non-empty.

Finally, we set\(^{34}\)
\[ P_{k(l)}^{i,n} := \bigcup_{R \in R_{k}^{i}} P_{k(l)}^{i,n,[R]} \]
and notice that all the tiles in \( P_{k(l)}^{1,n} \) have (roughly) uniform \( F \)-mass \( \approx 2^{-k} \).

7. Preparatives; subdiving the Main Theorem B, part i)
7.1. Operator Discretization - The second stage. With these facts, making use of one more observation:
\[ \forall I \text{ interval } I \subset \mathcal{I}_{k+1} \text{ and } I \cap \mathcal{I}_{k} = \emptyset \Rightarrow |I \cap F| = 0 \]
we conclude that for any \( g \in L^{\infty}(\mathbb{T}) \) with \( \|g\|_{\infty} = 1 \) one has
\( ^{34} \)For \( i = 1 \) the summation in (122) ranges over \( R \in R_{k}^{0} \) while for \( i = 2 \) the summation is over \( R \in R_{k}^{0,0} \).
we are given a tree $P$ collection of disjoint dyadic intervals satisfying the following properties:

**Main Lemma** on the following key result:

7.3. The Main Lemma. In order to prove the above theorems we will rely on the following key result:

**Main Lemma** Let $n \in \mathbb{N}$ and $k \in \{1, \ldots, k_F\}$ be fixed parameters. Assume we are given a tree $\mathcal{P}$ with top $\bar{P}$. Further, let $I_{\mathcal{P}} := \bar{I}_{\mathcal{P}}$ and let $\mathcal{I}_{\mathcal{P}}$ be a collection of disjoint dyadic intervals satisfying the following properties:

$$\| \sum_{P \in \mathcal{P}} T_P^* (g) \|_{L^1(F)} \leq \| \sum_{P \in \mathcal{P}_{\text{cluster}}} T_P^* (g) \|_{L^1(F)}$$

$$+ \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}^1}^*[R \cup \mathcal{P}^2_R] (g) \|_{L^1(F)}$$

$$+ \| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^0} \chi_{I_R} T_{\mathcal{P}^1}^*[R] (g) \|_{L^1(F)}$$

$$+ \| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^0} \chi_{I_R} T_{\mathcal{P}^2}^*[R] (g) \|_{L^1(F)} .$$

(123)

7.2. Subdividing the Main Theorem B, part i): the key statements.

The proof of our Main Theorem B i) follows immediately from the following four theorems:

**Theorem 18.** Let $F \subseteq \mathbb{T}$ be a measurable set and $g \in L^\infty(\mathbb{T})$. Then

$$\| (T_{\mathcal{P}_{\text{cluster}}}^*)^*(g) \|_{L^1(F)} \lesssim |F| \log \left( \frac{4}{|F|} \right) \|g\|_{\infty} .$$

(124)

The proof of Theorem 18 is trivial if one uses the key observation that $\mathcal{P}_{\text{cluster}}$ is just a $c(\alpha)$–dilation of a tree where here $c(\alpha)$ is the same as the one appearing in (63). Indeed, heuristically, the proof reduces to the fact that the (maximal) Hilbert transform is bounded from $L \log L$ to $L^1$. We leave the details for the reader.

**Theorem 19.** Let $F \subseteq \mathbb{T}$ be a measurable set and $g \in L^\infty(\mathbb{T})$. Then

$$\left\| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}^1}^*[R \cup \mathcal{P}^2_R] (g) \right\|_{L^1(F)} \lesssim |F| \log \left( \frac{4}{|F|} \right) \|g\|_{\infty} .$$

(125)

**Theorem 20.** Let $F \subseteq \mathbb{T}$ be a measurable set and $g \in L^\infty(\mathbb{T})$. Then

$$\left\| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^0} \chi_{I_R} T_{\mathcal{P}^1}^*[R] (g) \right\|_{L^1(F)} \lesssim |F| \log \left( \frac{4}{|F|} \right) \|g\|_{\infty} .$$

(126)

**Theorem 21.** Let $F \subseteq \mathbb{T}$ be a measurable set and $g \in L^\infty(\mathbb{T})$. Then

$$\left\| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^0} \chi_{I_R} T_{\mathcal{P}^2}^*[R] (g) \right\|_{L^1(F)} \lesssim |F| \log \left( \frac{4}{|F|} \right) \|g\|_{\infty} .$$

(127)

7.3. The Main Lemma. In order to prove the above theorems we will rely on the following key result:
• for any $P \in \mathcal{P}$ we have a uniform control on the mass parameter:

\begin{equation}
\frac{|E(P)|}{|I_P|} \leq 2^{-n}.
\end{equation}

• let

\begin{equation}
I_{P,\text{min}}^*,
\end{equation}

be the collection of minimum time intervals inside the set $\{I_{P,s}^r\}_{r \in \{1, \ldots, 14\}}$. If we denote with $CZ^*(I_P)$ the Calderon-Zygmund decomposition of the set $I_P$ relative to the set $I_{P,\text{min}}^*$ the following hold:

– for any $I \in I_P$ and any $J \in CZ^*(I_P)$ either $I \cap J = \emptyset$ or $I \subseteq J$;
– for each $I \in I_P$ there exists $k(I) \in \{1, \ldots, k_F\}$ such that for any $J \in CZ^*(I_P)$ one has

\begin{equation}
\sum_{I \in I_P} 2^{-k(I)} |I| \lesssim 2^{-k} |J|.
\end{equation}

Then, the following holds:

\begin{equation}
\sum_{I \in I_P} 2^{-k(I)} \int_I |T_P^*(g)|^2 \lesssim 2^{-k} 2^{-2n} |I_P| \|g\|_\infty^2.
\end{equation}

**Proof.** We start by noticing that without loss of generality by shifting the tree at zero frequency, we can assume $\omega_{T_{opt}}(P) = 0$.

Define now the function

\begin{equation}
h := \sum_{I \in I_P} 2^{-k(I)} \chi_I.
\end{equation}

Then, relation (132) is equivalent with showing that for any $f, g$ real valued functions with $f \in L^2(T)$ and $g \in L^\infty(T)$, one has

\begin{equation}
|\int f h^\frac{1}{2} T_P^* g| \lesssim 2^{-\frac{k}{2}} 2^{-n} |I_P|^{\frac{1}{2}} \|g\|_\infty \|f\|_2.
\end{equation}

Given a generic $u \in L^1(T)$ and $I \subseteq T$ dyadic interval we set

\begin{equation}
L_I(u) := \frac{\int_I u(s) \, ds}{|I|} \chi_I.
\end{equation}

Using now (135) we define

\begin{equation}
L_P(u) := \sum_{J \in CZ^*(I_P)} L_J(u).
\end{equation}

With these definitions we now have

\begin{equation}
V_P := \int f h^\frac{1}{2} T_P^* g = \int T^P(f h^\frac{1}{2}) g.
\end{equation}
\[
E(\mathcal{P}) := \bigcup_{P \in \mathcal{P}} E(P) .
\]

If \( \mathcal{P} \subset \mathbb{P}_n \) is a tree with root \( \bar{P} \) and \( I_P := \widetilde{I}_P \), then one has
\[
|E(\mathcal{P})| \lesssim 2^{-n} |I_\mathcal{P}| .
\]

We are now ready to pass to our proof; we will start with the analysis of the last term:

\[
|B| \lesssim \|g\|_\infty |E(\mathcal{P})|^\frac{1}{2} \|T^P(\mathcal{L}_P(f h^{\frac{1}{2}}))\|_2
\]
\[
\lesssim \|g\|_\infty |E(\mathcal{P})|^\frac{1}{2} 2^{-\frac{n}{2}} \|\mathcal{L}_P(f h^{\frac{1}{2}}))\|_2
\]
\[
\lesssim \|g\|_\infty 2^{-n} |I_\mathcal{P}|^\frac{1}{2} \left( \sum_{J \in CZ(\mathcal{P}^*)} \left( \int_J f h^{\frac{1}{2}} \frac{1}{|J|} \right)^2 |J| \right)^\frac{1}{2}
\]
\[
\lesssim_{CS} \|g\|_\infty 2^{-n} |I_\mathcal{P}|^\frac{1}{2} \left( \sum_{J \in CZ(\mathcal{P}^*)} \frac{\int_J f^2 \int_J h}{|J|} \right)^\frac{1}{2} .
\]

At this point we notice that (131) implies
\[
\int_J \frac{f h}{|J|} \lesssim 2^{-k} \quad \text{for any } J \in CZ(\mathcal{P}^*) .
\]

Inserting now (141) in the last inequality and using the mutual disjointness of the intervals in \( CZ(\mathcal{P}^*) \) we conclude
\[
B \lesssim 2^{-n} 2^{-\frac{k}{2}} |I_\mathcal{P}|^\frac{1}{2} \|g\|_\infty \|f\|_2 .
\]

We pass now to the treatment of \( A \).

Following now [24] and [25], we use the following key relation
\[
\int_J [u - \mathcal{L}_P(u)] = 0 \quad \forall J \in CZ(\mathcal{P}^*) .
\]

Using this, we deduce that
\[
A = \int (f h^{\frac{1}{2}} - \mathcal{L}_P(f h^{\frac{1}{2}})) (T^P g - \mathcal{L}_P(T^P g)) .
\]

Once in this point, we recall the following property proved in [25]: for any \( J \in CZ(\mathcal{P}^*) \) and \( x \in J \) fixed we have
\[
\left| T^P g(x) - \frac{1}{|J|} \int_J T^P g(s) ds \right| \chi_J \lesssim \|g\|_\infty \frac{|J|}{|I_{P^*} \supseteq J|} \sum_{|I_P|} |J|^2 |E(P)| .
\]
For completeness, we recall the proof of (145): assuming \( x \in J \) we have
\[
\left| T^{P^*} g(x) - \frac{1}{|J|} \int_J T^{P^*} g(s) \, ds \right| =
\]
\[
\left| \frac{1}{|J|} \int_J \left\{ \sum_{P \in \mathcal{P}} \int_T [\varphi_k(x-y) - \varphi_k(s-y)] \chi_{E(P)}(y) \, dy \right\} \, ds \right|
\]
\[
\lesssim \frac{1}{|J|} \int_J \left\{ \sum_{I_P \supseteq J} |I_P|^{-1} |J| f_E(P) \right\} \, ds .
\]

Returning now to (144) and making use of (145) we have
\[
|A| \lesssim \|g\|_\infty \sum_{J \in C(Z(P^*))} \frac{1}{|J|} \int_J \left| f h^{\frac{1}{2}} - L_P(f h^{\frac{1}{2}}) \right| \sum_{I_P \supseteq J} |I_P|^2 |E(P)|
\]
\[
\lesssim \|g\|_\infty 2^{-n} \sum_{J \in C(Z(P^*))} \left| J \right|^{\frac{1}{2}} \left( \int_J |f|^2 \right)^{\frac{1}{2}} \left( \frac{\int h}{|J|} \right)^{\frac{1}{2}} \sum_{I_P \supseteq J} |I_P|
\]
\[
\lesssim \|g\|_\infty 2^{-n} 2^{-\frac{k}{2}} \sum_{J \in C(Z(P^*))} \left| J \right|^{\frac{1}{2}} \left( \int_J |f|^2 \right)^{\frac{1}{2}} \lesssim_{CS} 2^{-n} 2^{-\frac{k}{2}} |I_P|^{\frac{1}{2}} \|g\|_\infty \|f\|_2 .
\]
This ends the proof of our Main Lemma. \( \square \)

**Corollary 23.** Let \( \mathcal{P}^{n,(m),(p)}_{(l),ab} \) as given by (98) via (94).

Then, denoting with \( I^{n,(m),(p)}_{p,ab} \) the spacial interval of the top of the tree \( \mathcal{P}^{n,(m),(p)}_{(l),ab} \), we have that
\[
\|T^{n,(m),(p)*}_{(l),ab} (g)\|_{L^2(F)} \lesssim 2^{-n} (2^{-m} |I^{n,(m),(p)}_{p,ab}|)^{\frac{1}{2}} \|g\|_\infty .
\]

**Proof.** First of all, let us notice that from the very definition of \( \mathcal{P}^{n,(m),(p)}_{(l),ab} \) we have that this family of tiles can be written as
\[
\mathcal{P}^{n,(m),(p)}_{(l),ab} = \bigcup_{\mathcal{P}^{n,(m),(p)}_{l,c} \in \mathcal{P}^{n,(m)}_{l,c}} \mathcal{P}^{n,(m)}_{l,c} ,
\]
with \( \{ \mathcal{P}^{n,(m)}_{l,c} \} \) maximal disjoint trees of uniform \( F \)-mass \( \approx 2^{-m} \) and mass \( \approx 2^{-n} \) and such that the corresponding extending tops obey the relation
\[
\sum_{\mathcal{P}^{n,(m),(p)}_{l,c} \in \mathcal{P}^{n,(m),(p)}_{(l),ab}} |\tilde{I}^{\mathcal{P}^{n,(m),(p)}_{l,c}}_{l,c}| \lesssim |I^{n,(m),(p)}_{(l),ab}| .
\]

As in the proof of the Main Lemma, we will use duality theory. Thus, taking \( f \in L^2(\mathbb{T}) \) with \( \|f\|_2 = 1 \), we deduce that (147) is equivalent with (150)
\[
\left| \int \left( \sum_{P_{l,c}^{n,(m)} \in \mathcal{P}_{l,a,b}^{n,(m)}} T_{P_{l,c}^{n,(m)}}^{(m)} (f \chi_F) \right) g \right| \leq 2^{-n} \left( 2^{-m} | \tilde{I}_{P_{l,a,b}^{n,(m) \star}}^{(m)} | \right)^{\frac{1}{2}} \| f \|_2 \| g \|_{\infty}.
\]

Let \( \mathcal{T}_{\min}^\star \) be the collection of minimum time intervals inside the set \( \{ I_{P_{l,a,b}^{n,(m) \star}}^{(m)} \} \) and let \( \tilde{I}_{\text{row}} = \bigcup_{P_{l,a,b}^{n,(m) \star}} I_{P_{l,a,b}^{n,(m) \star}}^{(m)} \). Define now

\[
CZ^{\star}(\tilde{I}_{\text{row}}),
\]

be the Calderon-Zygmund decomposition of \( \tilde{I}_{\text{row}} \) relative to the collection of intervals \( \mathcal{T}_{\min}^\star \).

We now define as in the Main Lemma (see (136)), the function

(151) \[ \mathcal{L}_{\text{row}}(u) := \sum_{J \in CZ^{\star}(\tilde{I}_{\text{row}})} L_J(u). \]

Following the same circle of ideas as before, we have the analogue of (137) given by

(153) \[
V_{\text{row}} := \int T_{P_{l,a,b}^{n,(m)}}^{(m) \star} (f \chi_F) g = \int T_{P_{l,a,b}^{n,(m)}}^{(m) \star} (f \chi_F - \mathcal{L}_{\text{row}}(f \chi_F)) g + \int T_{P_{l,a,b}^{n,(m)}}^{(m) \star} (\mathcal{L}_{\text{row}}(f \chi_F)) g = A + B.
\]

Using now (148) and (149) together with the mass single tree estimate that can be deduced from Observation 16 point 3) we deduce that

(154) \[ \| T_{P_{l,a,b}^{n,(m)}}^{(m) \star} \|_{2 \to 2} \lesssim 2^{-\frac{m}{2}}. \]

Also, notice that one has a natural correspondent to the key estimate (141) in the following

(155) \[ \| \mathcal{L}_{\text{row}}(f \chi_F) \|_{L^2(J)} \lesssim 2^{-\frac{m}{2}} \| f \|_{L^2(J)} \quad \text{for any } J \in CZ(\mathcal{P}^\star). \]

With this, one can now follow the same strategy as in the Main Lemma for the corresponding terms \( A \) and \( B \), see (146) and (140), respectively.
We now start the proof of our Main Theorem B i) by approaching our theorems stated at the beginning of this section in their increasing order of difficulty. For simplicity, wlog we will assume throughout the entire paper that \( \|g\|_\infty = 1 \).

8. Proof of Theorem 19.

This result follows easily by exploiting

- the almost orthogonality of the tiles within the set \( \{ \mathcal{P}_1^1[R] \}_{R \in \mathcal{R}_1} \), once that the integration is extended to the full torus spacial support \([0, 1]\);
- the almost orthogonality of the family of tiles within the set \( \{ \mathcal{P}_2^1[R] \}_{R \in \mathcal{R}_1} \), this time exploiting the well separated frequencies relative to the spacial support represented by the set \( \mathcal{I}_1 \);

Define now

\[
I_{1,F} := \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_1^1[R]}^*(g) \|_{L^1(F)},
\]

and

\[
I_{2,F} := \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_2^1[R]}^*(g) \|_{L^1(F)},
\]

and recall that our goal is to show that:

\[
I_{1,F} := I_{1,F} + I_{2,F} \lesssim |F| \log\left(\frac{4}{|F|}\right) \|g\|_\infty.
\]

Using now notation (121) and Cauchy-Schwarz we have:

\[
I_{1,F} = I_{1,F} + I_{2,F} \leq \sum_{n \in \mathbb{N}} \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_1^{1,n}[R]}^*(g) \|_{L^1(F)} + \sum_{n \in \mathbb{N}} \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_2^{1,n}[R]}^*(g) \|_{L^1(F)} \lesssim |F|^\frac{1}{2} \sum_{n \in \mathbb{N}} \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_1^{1,n}[R]}^*(g) \|_{L^2} + |F|^\frac{1}{2} \sum_{n \in \mathbb{N}} \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_2^{1,n}[R]}^*(g) \|_{L^2}.
\]

Next, appealing to (115), (114) and the definition of \( \mathcal{P}_1^{1,n}[R] \) for \( R \in \mathcal{R}_1 \), we apply Cauchy-Schwarz and the standard \( L^2 \)-bound for uniform mass tree of order \( \approx 2^{-n} \) following from (81), and deduce

\[
I_{1,F} := \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_1^{1,n}[R]}^*(g) \|_{L^2} \leq \| \sum_{l,a,b,m \in \{1,2\}} T_{\mathcal{P}_1^{1,n,(m)}[l,a,b]}^*(g) \|_{L^2(\mathcal{I}_2)} \lesssim \| \sum_{l,a,b,m \in \{1,2\}} |T_{\mathcal{P}_1^{1,n,(m)}[l,a,b]}^*(g)|^2 \|_{L^2(\mathcal{I}_2)} \lesssim 2^{-\frac{n}{2}} |\mathcal{I}_2|^\frac{1}{2}.
\]

Thus we deduce that

\[
I_{1,F} \lesssim |F|^\frac{1}{2} \sum_{n \in \mathbb{N}} 2^{-\frac{n}{2}} |\mathcal{I}_2|^\frac{1}{2} \lesssim |F|.
\]
In a similar fashion, using (116) - (119), for the second term we have:

\[ I_{1,F}^{2,n} := \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_1^2(R)}^*(g) \|_{L^2} \]

\[ \lesssim \sum_{p} \| \sum_{R \in \mathcal{R}_1} \chi_{I_R} T_{\mathcal{P}_1^2(R)}^* [p] (g) \|_{L^2} \]

\[ \lesssim \sum_{p} \left( \sum_{R \in \mathcal{R}_1} \left| T_{\mathcal{P}_1^2(R)}^* [p] (g) \right|^2 \right)^{1/2} \]

\[ \lesssim \sum_{p} \left( \sum_{l,m,a,b} \left| T_{\mathcal{P}_1^2(l,m,a,b)}^* [p] (g) \right|^2 \right)^{1/2} \|_{L^2(I)} \).

Applying now Main Lemma and Cauchy-Schwarz, we have

(160)

\[ I_{1,F}^2 \lesssim |F| \frac{k}{k_F} \sum_{n \in \mathbb{N}} \sum_{p} \left( \sum_{l,m,a,b} 2^{-2n} 2^{-m} \left| \tilde{T}_{\mathcal{P}_1^2(l,m,a,b)}^* [p] (g) \right| \right)^{1/2} \]

\[ \lesssim |F| \frac{k}{k_F} \sum_{n \in \mathbb{N}} 2^{-\frac{k}{2}} \left( \sum_{l,a,b,m \leq k_F} 2^{-2n} 2^{-m} \left| \tilde{T}_{\mathcal{P}_1^2(l,m,a,b)}^* [p] (g) \right| \right)^{1/2} \]

\[ \lesssim |F| \frac{k}{k_F} \left( \sum_{n \in \mathbb{N}} 2^{-\frac{k}{2}} \right)^{1/2} \frac{1}{k_F} \|g\|_\infty \lesssim |F| \log \frac{4}{|F|} \cdot \]

Combining now (159) and (160), we conclude that (158) holds.

9. Proof of Theorem 20.

Our approach here is similar in spirit with the one embraced for proving Proposition 3 in [25]. For convenience we will provide all the required details within this section.

Recall that our intention is to prove that

(161) \[ \| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_1^k} \chi_{I_R} T_{\mathcal{P}_1^2(R)}^* [p] (g) \|_{L^1(F)} \lesssim |F| \log \left( \frac{4}{|F|} \right) \|g\|_\infty , \]

for \( F \subseteq \mathbb{T} \) measurable set and \( g \in L^\infty(\mathbb{T}) \) with \( \|g\|_\infty = 1 \).

We start by making precise the following heuristic: each term of the form \( T_{\mathcal{P}_1^2(R)}^* (g) \) is essentially a sum of constant coefficients times suitable imaginary exponentials acting on the set \( F \cap I_R \).

For this, we need to introduce first some notations:

Fix \( R \in \mathcal{R}_k^0 \) and using (112) we decompose

\[ \mathcal{P}_k^1[R] = \bigcup_{l} \mathcal{P}_{k,(l)}^1[R] \]
into maximal trees at distinct frequencies $\{\omega_l\}_l$.

Recalling now (135), we define

$$T^*_c P^1_k[R] g(\cdot) := \sum_l c^{2\pi i \omega_l} L_R(T^*_c P^1_k[R,0] g(\cdot)),$$

where here we denote with $\omega_l$ the frequency corresponding to the tree $P^1_k(l)[R]$ and $P^1_k(l)[R,0]$ designates the shift of $P^1_k(l)[R]$ at the origin.

**Lemma 24.** Let $R \in \mathbb{R}^0_k$. Then the following holds

$$\|T^*_c P^1_k[R] g - T^*_c P^1_k[R,0] g\|_{L^\infty(R)} \lesssim \sum_{P \in P^1_k(R)} \frac{|I_R| \int_{E(P)} |g|}{|I_P|}. \quad (163)$$

**Proof.** Set for notational simplicity $T_l = T^*_c P^1_k(l)[R,0] g - L_R(T^*_c P^1_k(l)[R,0] g)$.

For fixed $l$ and $x \in I_R$ we have

$$|T_l(x)| = \left| T^*_c P^1_k(l)[R,0]^* (g)(x) - \frac{1}{|I_R|} \int_{I_R} T^*_c P^1_k(l)[R,0]^* (g) ds \right| =$$

$$\left| \frac{1}{|I_R|} \int_{I_R} \sum_{P \in P^1_k(l)[R,0]} \int_{|F| \leq |I_R|} \int_T [\varphi_j(x - y) - \varphi_j(s - y)] g(y) \chi_{E(P)}(y) dy ds \right| \leq \sum_{P \in P^1_k(l)[R,0]} \frac{|I_R| \int_{E(P)} |g|}{|I_P|}. \quad (164)$$

Summing now in $l$ we deduce that (163) holds.

**Lemma 25.** With the previous notations, the following holds

$$\| \sum_{1 < k \leq k_F} \sum_{R \in \mathbb{R}^0_k} \chi_{I_R} T^*_c P^1_k[R] g \|_{L^1(F)} \lesssim \| \sum_{1 < k \leq k_F} \sum_{R \in \mathbb{R}^0_k} \chi_{I_R} T^*_c P^1_k[R] g \|_{L^1(F)} + |F| \log(\frac{4}{|F|}) \|g\|_{\infty}.$$  

**Proof.** Appealing to the statement of Lemma 24 we deduce

$$\| \sum_{1 < k \leq k_F} \sum_{R \in \mathbb{R}^0_k} \chi_{I_R} T^*_c P^1_k[R] g \|_{L^1(F)} \lesssim \sum_{1 < k \leq k_F} \sum_{R \in \mathbb{R}^0_k} \chi_{I_R} T^*_c P^1_k[R] g \|_{L^1(F)} \quad (165)$$
\[ + \sum_{k \leq k_F} \sum_{R \in \mathcal{R}_k^0} \sum_{P \in \mathcal{P}_1^1[R]} \frac{|I_R|}{|I_P|} \frac{\int_{E(P)} |g|}{|I_P|} |I_R \cap F|. \]

Now denoting with \( U \) the last term above, we deduce that
\[ U \lesssim \|g\|_{\infty} \sum_{k \leq k_F} 2^{-k} \sum_{R \in \mathcal{R}_k^0} \sum_{P \in \mathcal{P}_1^1[R]} \frac{|I_R|^2}{|I_P|^2} |E(P)|. \]

Thus, for proving our lemma it is enough to show that:
\[ (166) \quad U \lesssim |F| \log(\frac{4 |F|}{|\tilde{I}_k|}) \|g\|_{\infty}. \]

Fix \( k \in \{1, \ldots, k_F\} \). Then (166) follows immediately once we show that
\[ (167) \quad U_k := \sum_{R \in \mathcal{R}_k^0} \sum_{P \in \mathcal{P}_1^1[R]} \frac{|I_R|^2}{|I_P|^2} |E(P)| \lesssim |\tilde{I}_k|. \]

Set \( \mathcal{I}(\mathcal{P}_1^1[R]) := \{I_P \mid P \in \mathcal{P}_1^1[R]\} \). Then, (167) is a consequence of
\[ (168) \quad U_k \lesssim \sum_{R \in \mathcal{R}_k^0} \sum_{I_P \in \mathcal{I}(\mathcal{P}_1^1[R])} \frac{|I_R|^2}{|I_P|^2} |I_P| \lesssim \sum_{R \in \mathcal{R}_k^0} |I_R| \lesssim |\tilde{I}_k|. \]

Based on the above two lemmas, we can justify now the following main heuristic for families of the type \( \mathcal{P}_1^1[R] \):

\[ (169) \quad \text{On the interval } I_R \text{ our operator } T^{\mathcal{P}_1^1[R]}(g) \text{ is morally constant.} \]

In order to prove the above heuristic, we start as before by decomposing
\[ (170) \quad \mathcal{P}_1^1[R] = \bigcup_l \mathcal{P}_{k,(l)}^1[R], \]
into maximal trees living at distinct (dyadic) frequencies \( \omega_l \). Once at this point, recalling the definition of \( \mathcal{P}_1^1[R] \) and (60) we deduce that for all \( l \) and \( R \) as in decomposition (170) we have
\[ (171) \quad \sum_l \omega_l \leq \tilde{C} |\omega_R|. \]

Next, taking any \( y_R \in I_R \), we notice that the Taylor expansion
\[ (172) \quad e^{2\pi i \omega_l (x-y_R)} = \sum_{k \geq 0} (2\pi i)^k \frac{\omega_l (x - y_R)^k}{k!}, \]
is absolutely and uniformly convergent for any \( x, y_R \in I_R \).
Recall now (162) and notice that

\[ T_{P_{k}^{1}[R]}^{*} = \sum_{l} e^{2\pi i \omega_{l} \cdot T_{P_{k}(l)[R,0]}^{*}}. \]

From (172) and (173) we now deduce that

\[ \| T_{C_{m}[R]}^{*} (g) \|_{L^{1}(F)} \lesssim |I_{R} \cap F| \sum_{l} e^{2\pi i \omega_{l} y_{R}} L_{I_{R}} (T_{C_{m}[R]}^{*} (g)) (y_{R}) \]

\[ + |I_{R} \cap F| \sup_{l} \int_{I_{R}} |T_{C_{m}[R]}^{*} (g)| |I_{R} \cap F| \log(\frac{4}{|I_{R}|}) \|g\|_{\infty}. \]

Using now (175) and Lemmas 24 and 25 we have

**Lemma 26.** The following holds

\[ \| \sum_{1 < k \leq k_{F}} \sum_{R \in \mathcal{R}_{k}^{0}} \chi_{I_{R}} T_{C_{m}[R]}^{*} (g) \|_{L^{1}(F)} \]

\[ \lesssim \sum_{1 < k \leq k_{F}} 2^{-k} \| \sum_{R \in \mathcal{R}_{k}^{0}} T_{C_{m}[R]}^{*} (g) \|_{L^{1}} \]

\[ + \sum_{1 < k \leq k_{F}} 2^{-k} \sum_{R \in \mathcal{R}_{k}^{0}} \sup_{l} \int_{I_{R}} |T_{C_{m}[R]}^{*} (g)| + |F| \log(\frac{4}{|F|}) \|g\|_{\infty}. \]

**Finishing the proof of Theorem 20.**

Recalling (112) and (114) we set

\[ I[n] := \| \sum_{1 < k \leq k_{F}} \sum_{R \in \mathcal{R}_{k}^{0}} \chi_{I_{R}} T_{C_{m}[R]}^{*} (g) \|_{L^{1}(F)}, \]

and notice that a similar relation with (176) will hold for this corresponding \( n \)-level.

Applying first Cauchy-Schwarz and then the Main Lemma to each maximal row \( \mathcal{P}_{k,l}^{n} \) resulted from \( \mathcal{P}_{l} \cap \mathcal{P}_{k+1}^{n} \), we have

\[ \sum_{n} I[n] \]
$$\lesssim \sum_{n} \sum_{1 < k \leq k_F} 2^{-k} \left\| \sum_{R \in \mathcal{R}_k^0} T_{e}^{P^{1,n}_k[R]}(g) \right\|_{L^1}$$

$$+ \sum_{n} \sum_{1 < k \leq k_F} 2^{-k} \sum_{R \in \mathcal{R}_k^0} \sup_l \int_{I_R} |T_{e}^{P^{1,n}_k[R]}(g)| + |F| \log\left(\frac{4}{|F|}\right)$$

$$\lesssim \sum_{n} \sum_{1 < k \leq k_F} 2^{-k} \sum_{R \in \mathcal{R}_k^0} \left| I_R \right|^{1/2} \left\| \sum_{l} T_{e}^{P^{1,n}_k[R]}(g) \right\|_{L^2(I_R)}$$

$$+ \sum_{n} \sum_{1 < k \leq k_F} 2^{-k} \left| I_k \right|^{1/2} \sup_l \left( \int_{I_R} |T_{e}^{P^{1,n}_k[R]}(g)|^2 \right)^{1/2} + |F| \log\left(\frac{4}{|F|}\right)$$

$$\lesssim \sum_{n} \sum_{1 < k \leq k_F} 2^{-k} \left| I_k \right|^{1/2} \left( \sum_{l} |T_{e}^{P^{1,n}_k[R]}(g)|^2 \right)^{1/2} \left\| \sum_{l} T_{e}^{P^{1,n}_k[R]}(g) \right\|_{L^1} + |F| \log\left(\frac{4}{|F|}\right)$$

$$\lesssim \sum_{n} \sum_{1 < k \leq k_F} 2^{-k} \left| I_k \right|^{1/2} \left\| \sum_{l} T_{e}^{P^{1,n}_k[R]}(g) \right\|_{L^1} + |F| \log\left(\frac{4}{|F|}\right) \lesssim |F| \log\left(\frac{4}{|F|}\right),$$

where in the second to last line we used notation (122).

With this we conclude that (161) holds.

10. **Reduction of Theorem 21 to Proposition 37.**

Our goal here is to show that

$$\left\| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^0} \chi_{I_R} T_{e}^{P^{2}_k[R]}(g) \right\|_{L^1(F)} \lesssim |F| \log\left(\frac{4}{|F|}\right) \|g\|_{\infty}.$$  

This will be the most difficult and in the same time interesting part of our paper in which the newly introduced concept of a set-resolution of the time-frequency plane at 0-frequency will play a key role. The remainder part of the paper is entirely dedicated for proving (177).

10.1. **An intermediate step: reduction to Proposition 31.** We start by adapting (162) to our new context: fix $R \in \mathcal{R}_k^0$ and recall the meaning of $\mathcal{R}$. Applying the decomposition

$$\mathcal{P}^2_k[R] = \bigcup_l \mathcal{P}^2_{k,l}[R]$$

into maximal rows at distinct frequencies, we now define

$$T_{e}^{P^{2}_k[R]}(g)(\cdot) := \sum_{l} e^{2\pi i \omega_l} L_{I_R}\left( T_{e}^{P^{2}_{k,l}[R]}(R,0) \right)(\cdot),$$

where, as before, $\mathcal{P}^2_{k,l}[R,0]$ designates the shift of $\mathcal{P}^2_{k,l}[R]$ at the origin and $\omega_l$ stands for the frequency corresponding to the tree $\mathcal{P}^2_{k,l}[R]$.

With the obvious modifications, the lemma below is the adaptation of the corresponding Lemma 24:
Lemma 27. Let \( R \in \mathcal{R}_k^> \). Then the following holds

\[
\| T^p_k[R]^\ast (g) - T^p_{c_k}[R]^\ast (g) \|_{L^\infty(I_R)} \lesssim \sum_{P \in \mathcal{P}^2[R]} \frac{|I_R|}{|I_P|} \frac{\int_{E(P)} |g|}{|I_P|}.
\]

Proof. Adapting the definition of \( T_l \) from Lemma 24 to our context, we set

\[
T_l := T^p_{c_k}[R,0]^\ast (g) - L_{I_R}(T^p_{c_k}[R,0]^\ast (g)).
\]

With this, proceeding as in the analogue proof, for fixed \( l \) and \( x \in I_R \) we deduce

\[
|T_l(x)| \lesssim \sum_{P \in \mathcal{P}^2_k[R,0]} \frac{|I_R|}{|I_P|} \frac{\int_{E(P)} |g|}{|I_P|},
\]

which, after summing in \( l \), gives us the desired estimate. \( \square \)

The next result is the analogue of Lemma 25:

Lemma 28. With the previous notations, the following holds

\[
\| \sum_{1<k \leq k_F} \sum_{R \in \mathcal{R}_k^>^0} \chi_{I_R} T^p_k[R]^\ast (g) \|_{L^1(F)}
\]

\[
\lesssim \| \sum_{1<k \leq k_F} \sum_{R \in \mathcal{R}_k^>^0} \chi_{I_R} T^p_{c_k}[R]^\ast (g) \|_{L^1(F)} + |F| \log(\frac{4}{|F|}) \|g\|_\infty.
\]

Proof. Again, this is just an adaptation of the proof of Lemma 25 via Lemma 27. Indeed, the analogue of (165) reads now

\[
\| \sum_{1<k \leq k_F} \sum_{R \in \mathcal{R}_k^>^0} \chi_{I_R} T^p_k[R]^\ast (g) \|_{L^1(F)}
\]

\[
\lesssim \| \sum_{1<k \leq k_F} \sum_{R \in \mathcal{R}_k^>^0} \chi_{I_R} T^p_{c_k}[R]^\ast (g) \|_{L^1(F)}
\]

\[
+ \sum_{k \leq k_F} \sum_{R \in \mathcal{R}_k^>^0} \sum_{P \in \mathcal{P}^2_k[R]} \frac{|I_R|}{|I_P|} \frac{\int_{E(P)} |g|}{|I_P|} |I_R \cap F|.
\]

As before, letting

\[
U := \sum_{k \leq k_F} \sum_{R \in \mathcal{R}_k^>^0} \sum_{P \in \mathcal{P}^2_k[R]} \frac{|I_R|}{|I_P|} \frac{\int_{E(P)} |g|}{|I_P|} |I_R \cap F|,
\]
we deduce that
\[ U \lesssim \|g\|_{\infty} \sum_{k \leq k_F} 2^{-k} \sum_{R \in \mathcal{R}_k^{>0}} \sum_{P \in \mathcal{P}_k^2[R]} \frac{|I_R| |I_R|}{|I_P|^2} |E(P)|. \]

Fixing \( k \in \{1, \ldots, k_F\} \), it remains to show that
\[ U_k := \sum_{R \in \mathcal{R}_k^{>0}} \sum_{P \in \mathcal{P}_k^2[R]} \frac{|I_R| |I_R|}{|I_P|^2} |E(P)| \lesssim |\tilde{I}_k|. \]

Setting \( \mathcal{I}(\mathcal{P}_k^2[R]) := \{ I_P \mid P \in \mathcal{P}_k^2[R] \} \), (184) follows from (185)
\[ U_k \lesssim \sum_{R \in \mathcal{R}_k^{>0}} \sum_{I_P \in \mathcal{I}(\mathcal{P}_k^2[R])} \frac{|I_R| |I_R|}{|I_P|^2} |I_P| \]
\[ \lesssim \sum_{I \in \mathcal{I}_k} \sum_{R \in \mathcal{R}_k^{>0}[I]} \frac{|I_R|}{|I_P|} \sum_{I_P \in \mathcal{I}(\mathcal{P}_k^2[R])} \frac{|I_R|}{|I_P|} \lesssim \sum_{I \in \mathcal{I}_k} \sum_{R \in \mathcal{R}_k^{>0}[I]} |I_R| \lesssim |\tilde{I}_k|. \]

From the construction of the \( F \)-resolution of the time-frequency plane at \( 0 \)-frequency and from the definition of a generic \( \mathcal{P}_k^2[R] \) we have that

**Observation 29.** Let \( P \in \mathcal{P}_k^2[R] \). Then, the phase corresponding to \( P \), referred to as \( e^{2\pi i \alpha P} \cdot \) with \( \alpha P \in \omega_P \), does oscillate over the length of the interval \( I_R \) but is morally constant over the interval \( I_R \).

Applying this observation and invoking a similar argument with the one in (174)-(175) we deduce based on Lemma 28 that

**Lemma 30.** With the previous notations, we have that
\[ \| \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^{>0}} \chi_{I_R} T^{\mathcal{P}_k^2[R]^*}(g) \|_{L^1(F)} \]
\[ \lesssim \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k^{>0}} \frac{|I_R \cap F|}{|I_R|} \int_{I_R} \left| T_{c}^{\mathcal{P}_k^2[R]^*}(g) \right| + |F| \log\left(\frac{4}{|F|}\right) \|g\|_{\infty} \]

Fix now \( 1 < k \leq k_F \) and \( R \in \mathcal{R}_k^{>0} \). Recall the definitions at the end of Section 6. Now, in order to analyze the RHS of inequality (186), we linearize the main term and thus, for \( h, g \in L^\infty \) with \( \|h\|_{\infty} = \|g\|_{\infty} = 1 \), we define the following expression:
\[ \mathcal{E}(h, g) := \sum_{n \geq 1} \sum_{p \leq k_F} \mathcal{E}_n^p(h, g), \]
with

\[ E_p^p(h, g) := \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}^0} \frac{|I \cap F|}{|I|} \int_{I_R} h T^{p^{2,n,|p|}[R]^*}_c (g). \]

For any \( R \in \mathcal{R}_k[I] \) we also define

\[ T_{cc}^{p^{2,n,|p|}[R]^*}_k (g) := \sum_l e^{2\pi i \omega_l \cdot} L_f(T^{p^{2,n,|p|}[R]^*}_c (g)). \]

Depending on the context, we may choose to replace in (188) the expression \( T^{p^{2,n,|p|}[R]^*}_c (h, g) \) by \( T_{cc}^{p^{2,n,|p|}[R]^*}_k (g) \) or simply \( T^{p^{2,n,|p|}[R]^*}_c (g) \). This is fine since, the difference between any two such representations of \( E_p^p(h, g) \), even after summing in \( p \), represents an admissible error of the form \( |F| \log(\frac{1}{|F|}) \). This last claim can be verified by following similar reasoning with the ones used in the proof of Theorem 20 (more specifically Lemma 25).

With these, we remark that (188) can be then written in any of the following fundamental forms depending on our convenience:

\[ E_p^p(h, g) = \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}^0} \frac{|I \cap F|}{|I|} \int_{I_R} h T^{p^{2,n,|p|}[R]^*}_c (g) \]

\[ \approx \sum_{1 < k \leq k_F} \frac{|I \cap F|}{|I|} \sum_{R \in \mathcal{R}^0} \sum_{R \in \mathcal{R}^{h_1}(R)} \int_{I_R} h T^{p^{2,n,|p|}[R]^*}_c (g) \]

Once at this point, from Lemmas 27 - 30 and relation (190), one trivially deduces that Theorem 21 is a direct consequence of

**Proposition 31.** The following holds:

\[ |E(h, g)| \lesssim |F| \log\left(\frac{4}{|F|}\right) \|h\|_{\infty} \|g\|_{\infty}. \]

The remaining part of the paper is dedicated to proving Proposition 31.

10.2. A first reduction: small and large \( n \)-mass terms. \[ ^{36} \]

In this subsection we will split our term \( E \) in (187) into a small and large \( n \)-mass component and show that the large \( n \) component is under control.

Define now \( n_F \in \mathbb{N} \) to be the smallest natural number \( n \in \mathbb{N} \) such that

\[ 2^n \geq k_F \approx \log \left(\frac{4}{|F|}\right). \]

\[ ^{35} \text{In these relations the symbol } \approx \text{ should be understood as equality up to admissible error terms that are controlled from above by } |F| \log\left(\frac{1}{|F|}\right). \]

\[ ^{36} \text{Here we use the label } n \text{-mass for a tile } P \text{ to designate the fact that } P \in \mathbb{P}_n. \]
Thus, we have that \( n_F \approx \log \log |F| \).

Recalling now (190), we define

\[
E_< := \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_n^p,
\]

\[
E_> := \sum_{n > n_F} \sum_{p \leq k_F} \mathcal{E}_n^p,
\]

and notice from (187) that

\[
E = E_+ + E_>.
\]

As announced earlier, in what follows we will show that the second term is under control, that is

\[
\text{Proposition 32. With the previous notations, one has}
\]

\[
|E_>(h, g)| \lesssim |F| \log(\frac{4}{|F|}) \|h\|_\infty \|g\|_\infty.
\]

The above result is based on the next two key lemmas:

\[
\text{Lemma 33. Fix } I \in \mathcal{I}_k, r \in \mathbb{N}, \text{ and } s_j \in \{U, L\} \text{ and take } \mathcal{R}_k^{s_1, ..., s_r}[I]. \text{ With the definitions and notations from the previous sections the following holds:}
\]

\[
\sum_{R \in \mathcal{R}_k^{s_1, ..., s_r}[I]} \sum_{R' \in \mathcal{R}_k^{s_1, ..., s_r}[R]} \int_{I_R} \left| \overline{T}^{2,n,p}_{cc}[R][g] \right|^2 \lesssim \left( \sum_{R \in \mathcal{R}_k^{s_1, ..., s_r}[I]} \right)^{1/2} \left( \sum_{R' \in \mathcal{R}_k^{s_1, ..., s_r}[I]} \int_{I_{R'}} \left| \overline{T}^{2,n,p}_{cc}[R][g] \right|^2 \right)^{1/2}.
\]

\[
\text{Proof. We start by noticing that the LHS in (197) can be rewritten as}
\]

\[
V_{\mathcal{R}_k^{s_1, ..., s_r}[I]} := \sum_{s'_{u} = \ldots = s_{r+1} = U}^{s_{1} = \ldots = s_{r} = L} \int_{I_{R'}} \left| \overline{T}^{2,n,p}_{cc}[R][g] \right|^2.
\]

Let us define

\[
I_{\mathcal{R}_k^{s_1, ..., s_r}[I]} := \bigcup_{R \in \mathcal{R}_k^{s_1, ..., s_r}[I]} I_R.
\]

Also notice that since for any \( R, R' \in \mathcal{R}_k^{s_1, ..., s_r}[I] \) one has

\[
\overline{T}^{2,n,p}_{cc}[R][g] = \overline{T}^{2,n,p}_{cc}[R'][g],
\]

\[\text{37Here, as before, we apply the usual convention } \overline{T}^{2,n,p}_{k, \text{chi}}[\mathcal{R}_k^{s_1, ..., s_r, U}[R]] := \overline{T}^{2,n,p}_{k}[R] \text{ for some } R \in \mathcal{R}_k^{s_1, ..., s_r, U}[R].\]
and thus one further deduces that the following is a well defined object:

\[ T_{cc}^{\mathcal{P}^{2, n, [p]; R_k^{s_1 \ldots s_r'}} [I] (g) = T_{cc}^{\mathcal{P}^{2, n, [p]; R_k^{s_1 \ldots s_r'}} (g) \text{ for some } R \in R_k^{s_1 \ldots s_r'} [I]. \]

Notice that based on (59), for any \( r' > r, \) we have that

\[
(200) \quad |I_{R_k^{s_1 \ldots s_r'} [I]}| \lesssim \frac{1}{2^{r-r'}} |I_{R_k^{s_1 \ldots s_r'} [I]}|.
\]

Fix now \( s_1 \ldots s_r. \) Throughout the entire proof when talking about \( s_1 \ldots s_r' \)
we assume \( r' > r \) and \( s'_u = s_u \) if \( u \leq r, \) \( s'_r = U \) and \( s'_u = L \) if \( u > r + 1. \)

For \( w \in \mathbb{N} \) we define

\[
(201) \quad I_{R_k^{s_1 \ldots s_r'}} ^{w} [I] := \frac{1}{|I_{R_k^{s_1 \ldots s_r'}} [I]|} \int \frac{\mathcal{P}_{cc}^{2, n, [p]; R_k^{s_1 \ldots s_r'}} [I] (g) (x)}{2^{w}} \quad \| \mathcal{P}_{cc}^{2, n, [p]; R_k^{s_1 \ldots s_r'}} [I] (g) \|_{L^2 (I_{R_k^{s_1 \ldots s_r'}} [I])}.
\]

Deduce now from Chebyshev inequality that

\[
(202) \quad |I_{R_k^{s_1 \ldots s_r'}} ^{w} [I]| \lesssim 2^{-2w} |I_{R_k^{s_1 \ldots s_r'}} [I]|.
\]

We now estimate (198) as follows

\[
V_{R_k^{s_1 \ldots s_r}} [I] \approx \sum_{s'_1 = s_1, \ldots, s'_r = s_r, s'_r+1 = U} \sum_{w \in \mathbb{N}} \int_{R_k^{s_1 \ldots s_r'}} I_{R_k^{s_1 \ldots s_r'}} ^{w} [I] (g) \quad \| \mathcal{P}_{cc}^{2, n, [p]; R_k^{s_1 \ldots s_r'}} [I] (g) \|_{L^2 (I_{R_k^{s_1 \ldots s_r'}} [I])}.
\]

Using now the fact that

\[
(203) \quad \| \mathcal{P}_{cc}^{2, n, [p]; R_k^{s_1 \ldots s_r'}} [I] (g) \|_{L^2 (I_{R_k^{s_1 \ldots s_r'}} [I])} \lesssim \| \mathcal{P}_{cc}^{2, n, [p]; R_k^{s_1 \ldots s_r, U}} [I] (g) \|_{L^2 (I_{R_k^{s_1 \ldots s_r'}} [I])},
\]

together with (202) and (200) we further have

\[
V_{R_k^{s_1 \ldots s_r}} [I] \lesssim \| I_{R_k^{s_1 \ldots s_r'}} [I] \|_{L^2 (I_{R_k^{s_1 \ldots s_r'}} [I])} \| \mathcal{P}_{cc}^{2, n, [p]; R_k^{s_1 \ldots s_r, U}} [I] (g) \|_{L^2 (I_{R_k^{s_1 \ldots s_r'}} [I])}.
\]
with the notations introduced in the previous proof, our task is to prove the following lemma.

**Lemma 34.** Let \( I \in I_k \) and \( R_k^s \) as before. Then, the following holds:

\[
\sum_{s'_{1} = s_1, \ldots, s'_{r} = s_r, s'_{r+1} = L} \sum_{R \in R_k^{s_1 \ldots s_r}} \left( \sum_{R' \in R_k^{s'_{1} \ldots s'_{r}}} \left| I_R \right| \right)^{\frac{1}{2}} \left( \sum_{R' \in R_k^{s'_{1} \ldots s'_{r}}} \int_{I_R} \left| T_{cc}^{p,n,\nu \nu} [R_k^{s_1 \ldots s_r}, U] \right|^2 (g) \right)^{\frac{1}{2}} \leq 2 \left( \sum_{R \in R_k^{s_1 \ldots s_r}} \left| I_R \right| \right)^{\frac{1}{2}} \left( \sum_{R' \in R_k^{s'_{1} \ldots s'_{r}}} \int_{I_R} \left| T_{cc}^{p,n,\nu \nu} [R_k^{s_1 \ldots s_r}, U] \right|^2 (g) \right)^{\frac{1}{2}}.
\]

**Proof.** With the notations introduced in the previous proof, our task is to control the term

\[
V := \sum_{s'_{1} = s_1, \ldots, s'_{r} = s_r, s'_{r+1} = L} \sum_{R \in R_k^{s_1 \ldots s_r}} \left| I_{R_k^{s_1 \ldots s_r}} \right|^{\frac{1}{2}} \left( \sum_{R' \in R_k^{s'_{1} \ldots s'_{r}}} \int_{I_R} \left| T_{cc}^{p,n,\nu \nu} [R_k^{s_1 \ldots s_r}, U] \right|^2 (g) \right)^{\frac{1}{2}}.
\]

Throughout this proof, we fix the sequence \( s_1 \ldots s_r \) and whenever considering \( s'_1 \ldots s'_{r'} \), we assume \( r' > r \) and \( s'_u = s_u \) if \( u \leq r \) and \( s'_{r+1} = L \).

With the above notations and conventions, we have now the following important observations:

1. \( \left| I_{R_k^{s_1 \ldots s_r}} \right| \lesssim 2^{\frac{1}{2} + \frac{1}{r+1}} \left| I_{R_k^{s_1 \ldots s_r}} \right| \).
2. If \( s'_1 \ldots s'_{r'} \) and \( s''_1 \ldots s''_{r''} \), are two distinct sequences such that \( r' = r'' \) then
Using now (206), (208), (209) and Cauchy-Schwarz we treat (205) as follows:

- given \( r_0 \geq r + 1 \) we have
  \[
  \bigcup_{r' = r_0} \mathcal{P}^2 \mathcal{n}[\mathcal{R}_k^{s_1' \ldots s_r'} I]\bigcap \mathcal{P}^2 \mathcal{n}[\mathcal{R}_k^{s_1' \ldots s_r'} I] = \emptyset.
  \]

  - same, if \( r_0 \geq r + 1 \) then, we have
    \[
    \sum_{s_1' = s_1, \ldots , s_r' = s_r} |I_{\mathcal{R}_k^{s_1' \ldots s_r'} [I]}| \lesssim |I_{\mathcal{R}_k^{s_1' \ldots s_r} [I]}|.
    \]

  - finally, using the previous notations and the second item above, one deduces
    \[
    \sum_{s_1' = s_1, \ldots , s_r' = s_r} \mathcal{R}_k^{s_1' \ldots s_r'} I = \mathcal{R}_k^{s_1 \ldots s_r} I
    \]

Using now (206), (208), (209) and Cauchy-Schwarz we treat (205) as follows:

\[
\sum_{r_0 \geq r + 1} \frac{1}{2 |r_0 - r_0'|^4} |I_{\mathcal{R}_k^{s_1' \ldots s_r'} [I]}|^{\frac{1}{2}} \left( \sum_{s_1' = s_1, \ldots , s_r' = s_r} |I_{\mathcal{R}_k^{s_1' \ldots s_r'} [I]}| \right)^{\frac{1}{2}}
\]
Families of the form \( R_k \) and \( R_k^{(r)} \) can be derived from (197) and (204). We have

\[
\sum_{s'_{i+1} = L} \sum_{s'_{i} = s_{i}, r'_{i} = r_{0}} \sum_{s_{i} = 1} \sum_{s_{i} \leq L} \sum_{r'_{i} \leq r_{0}} \left| I_{\mathcal{R}_{k}^{s_{1} \ldots s_{r}}} \right| \left| I_{\mathcal{R}_{k}^{s'_{1} \ldots s'_{r}}} \right|
\]

\[
\sum_{r_{0} \geq r_{1} + 1} \frac{1}{2} \left| I_{\mathcal{R}_{k}^{s_{1} \ldots s_{r}}} \right| \left| I_{\mathcal{R}_{k}^{s'_{1} \ldots s'_{r}}} \right| \left| I_{\mathcal{R}_{k}^{s_{1} \ldots s_{r}}} \right|
\]

With these completed, we resume the proof of our Proposition 32:

First, we notice that introducing the notation for \( U \) in \( R_k^{U[I]} := R_k^{r^U[I]} \),

\[
(212)
\]

based on (197) and (204), we have

\[
(213)
\]

Next, we record the following two key properties

- for any \( 1 \leq k \leq k_F \) any \( n \geq 0 \) and any \( I \in \mathcal{I}_k \), based on (59) one has

\[
(214)
\]

- for any \( 1 \leq k \leq k_F \) any \( r \neq r' \) and any \( I \in \mathcal{I}_k \) one has

\[
(215)
\]
Applying now Cauchy-Schwarz and an almost-orthogonality argument one has

\[ |E_>(h, g)| \lesssim \sum_{n>n_F} \sum_{p \leq k_F} \left( \sum_{1<k \leq k_F} 2^{-k} \sum_{I \in I_k} \sum_r \sum_{R \in R_k^U[I]} |I_R| \right)^{\frac{1}{2}} \]

\[ \times \left( \sum_{1<k \leq k_F} 2^{-k} \sum_{I \in I_k} \sum_r \sum_{R \in R_k^U[I]} \int_{I_R} \sum_l \left| T_{cc, k(l)}^{p^{2,n,[p]}[R^{(r+1)}U[R]]^*}(g) \right|^2 \right)^{\frac{1}{2}} \]

Making use of (214), we have:

\[ (216) \sum_{I \in I_k} \sum_r \sum_{R \in R_k^U[I]} |I_R| \lesssim |\tilde{I}_k|, \]

Using now (216), Main Lemma and Corollary 23, we deduce from above that

\[ |E_>(h, g)| \lesssim \sum_{n>n_F} \sum_{p \leq k_F} (|F| \log \frac{4}{|F|})^{\frac{1}{2}} \]

\[ \times \left( \sum_{1<k \leq k_F} 2^{-k} \sum_{I \in I_k} \sum_r \sum_{R \in R_k^U[I]} \int_{I_R} \sum_l \left| T_{cc, k(l)}^{p^{2,n,[p]}[R^{(r+1)}U[R]]^*}(g) \right|^2 \right)^{\frac{1}{2}} \]

\[ \lesssim (|F| \log \frac{4}{|F|})^{\frac{1}{2}} \sum_{n>n_F} \sum_{p \leq k_F} \left( \sum_l 2^{-2n} \left( \sum_{r^{n,(m),[p]} \in \mathcal{P}_{(l),a,b}^{2,n,[p]}} 2^{-m} |I_{p^{2,n,[p]}[R]}|[R] \right) \right)^{\frac{1}{2}} \]

\[ \lesssim (|F| \log \frac{4}{|F|})^{\frac{1}{2}} \sum_{n>n_F} k_F 2^{-\frac{3}{2}} |F|^\frac{3}{2} \lesssim |F| \log \frac{4}{|F|}, \]

where in the second inequality we used the convention

\[ \mathcal{P}_{(l)}^{2,n,[p]} := \bigcup_{k \leq k_F, R \in R_k^U[I]} \mathcal{P}_{k(l)}^{2,n,[p]}[R^{(r+1)}U[R]]. \]

This proves the desired estimate (196).

It remains now to prove a similar estimate for the first term \( E_>, \) that is

**Proposition 35.** With the previous notations, one has

\[ (217) |E_<(h, g)| \lesssim |F| \log \left( \frac{4}{|F|} \right) \|h\|_\infty \|g\|_\infty. \]

The proof of Proposition 35 will be subject of the remaining sections.
10.3. Decomposing $\mathcal{E}_< = \mathcal{E}_{\text{Main}} + \mathcal{E}_{\text{Err}}$. Reduction of Proposition 35 to Proposition 37. We start with the usual decomposition of the tile families under our analysis.

Fix $n \leq n_F$, $k \leq k_F$ and $p \leq k_F$. Recall from Section 6, that

\begin{equation}
\bar{\mathcal{P}}^{2,n,[p]}_k[R] = \bigcup_{l \geq 1} \mathcal{P}^{2,n,[p]}_{k,(l)}[R],
\end{equation}

with (see (115))

\begin{equation}
\mathcal{P}^{2,n,[p]}_{k,(l)}[R] = \bigcup_{a,b,m} \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R],
\end{equation}

such that each $\mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]$ is as described in (90); in particular, $\mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]$ is a tree with uniform $F$-mass $\approx 2^{-m}$ and mass $\approx 2^{-n}$ and living at frequency $\omega_l$.

With this said, we start a finer analysis of our families $\{\bar{\mathcal{P}}^{2,n,[p]}_{k,(l)}[R]\}_R$.

Throughout our discussion $R$, $n$ and $p$ are fixed. Recalling (91) and picking $C_1 \in \mathbb{N}$ with $C_1 \leq 100$ we deduce that given any $\bar{\mathcal{P}}^{2,n,[p]}_{k,(l)}[R]$ there exist at most $C_1$ triss $\{\bar{\mathcal{P}}^{2,n,(m),[p]}_{k,(l),a,b}[R]\}_{s=1}^{C_1}$ such that

\begin{equation}
\bar{\mathcal{P}}^{2,n,[p]}_{k,(l)}[R] = \bigcup_{s=1}^{C_1} \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R].
\end{equation}

Given this, by splitting $\bar{\mathcal{P}}^{2,n,[p]}_{k,(l)}[R]$ in finitely many components (at most $C_1$) we will assume in all that follows that in fact

\begin{equation}
\bar{\mathcal{P}}^{2,n,[p]}_{k,(l)}[R] = \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R],
\end{equation}

for some $a$, $b$, $m$ uniquely determined by $R$, $p$ and $l$.

Fix $\epsilon \in (0, \frac{1}{1000})$. With the convention made in (221), we split the family $\{\bar{\mathcal{P}}^{2,n,[p]}_{k,(l)}[R]\}_l$ into:

- the family of $R$ heavy rows defined by

\begin{equation}
\mathcal{P}^{2,n,[p]}_{k,he}[R] := \left\{ \begin{array}{l}
\mathcal{P}^{2,n,[p]}_{k,(l),a,b}[R] = \frac{||T_c \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]|_{L^2(I_R^*)}}{|I_R|^{1/2}} \\
\mathcal{P}^{2,n,[p]}_{k,(l),a,b}[R] = \frac{||T_c \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]|_{L^2(I_R^*)}}{|I_R|^{1/2}}
\end{array} \right. \geq 2^{\epsilon n} \frac{||T_c \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]|_{L^2(I_R^*)}}{(2^{-m} |I_R|^{1/2})}.
\end{equation}

- the family of $R$ regular rows defined by

\begin{equation}
\mathcal{P}^{2,n,[p]}_{k,re}[R] := \left\{ \begin{array}{l}
\mathcal{P}^{2,n,[p]}_{k,(l),a,b}[R] = \frac{||T_c \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]|_{L^2(I_R^*)}}{|I_R|^{1/2}} < 2^{\epsilon n} \frac{||T_c \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]|_{L^2(I_R^*)}}{(2^{-m} |I_R|^{1/2})}\end{array} \right. \frac{||T_c \mathcal{P}^{2,n,(m),[p]}_{k,(l),a,b}[R]|_{L^2(I_R^*)}}{(2^{-m} |I_R|^{1/2})}.
\end{equation}
Observation 36. It is worth noticing that the reference quantity in both (222) and (223) obeys via Corollary 23 the following estimate:

\[
\|T_{\mathcal{P}_{(l),a,b}^n,(m),[p]}^* (g)\|_{L^2(F)} \lesssim (2^{-m} |I_{\mathcal{P}_{(l),a,b}^n,(m),[p]}|)^{\frac{1}{2}} \lesssim 2^{-n} .
\]

In fact, at least at the heuristic level, one should think as \(\lesssim\) replaced by \(\approx\).

Next, assume we are given \(s = (s_1, \ldots, s_r)\). Our intention is to further analyze the set \(\mathcal{R}_{k,chi}^{s,R} \equiv \mathcal{R}_{k,chi}^{s_1,\ldots,s_r,R}\) (recall (101)):

- fix \(R \in \mathcal{R}_k\). We define the \((s,p,n)-height\) relative to the base \(R\) as being simply

\[
\mathcal{H}_{n,p}^{k,R} := \# \{ l | \mathcal{P}_{k,(l)}^{2^n,[p]}[R] \in \mathcal{P}_{k,he}^{2^n,[p]}[R] \} \quad \text{for some } R \in \mathcal{R}_{k,chi}^{s,R}.
\]

The height \(\mathcal{H}_{n,p}^{k,R}\) is indeed well defined since one can make a similar reasoning with the one in Observation 17 and based on (223) deduce that

\[
\forall R, R' \in \mathcal{R}_{k,chi}^{s,R} \Rightarrow \mathcal{P}_{k,he}^{2^n,[p]}[R] = \mathcal{P}_{k,he}^{2^n,[p]}[R'] .
\]

- for \(w \in \mathbb{N}\) we define

\[
\mathcal{R}_{k}^{s,[w]}(n,p,R) := \bigcup_{w \in \mathbb{N}} \mathcal{R}_{k}^{s,[w]}(n,p,R).
\]

Notice that with these definitions and notations one has

\[
\mathcal{R}_{k,chi}^{s,R} = \bigcup_{w \in \mathbb{N}} \mathcal{R}_{k}^{s,[w]}(n,p,R) .
\]

We split now the operator associated with a generic tree \(\mathcal{P}_{(l)}^n = \mathcal{P}_{(l),a,b}^{n,(m),[p]}\) into a main and error term respectively.

We start by defining the **error term** of \(T_{\mathcal{P}_{(l)}^n,[p]}^* (g)\) as

\[
T_{Err}^{\mathcal{P}_{(l)}^n,[p]}(g) := T_{\mathcal{P}_{(l),he}^n,[p]}^* (g) + T_{\mathcal{P}_{(l),whe}^n,[p]}^* (g) ,
\]

where:

- the **heavy** component

\[
T_{\mathcal{P}_{(l),he}^n,[p]}^* (g) := \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k} \sum_{R' \in \mathcal{R}_{k,chi}^{s,R}} T_{(\mathcal{P}_{(l)}^n,[p]\cap \mathcal{P}_{(l),he}^{2^n,[p]}[R])}^* (g) .
\]
• the \textit{w-heavy} component

\begin{equation}
T^{p_{(l)}_{,\text{whe}}} (g) := \sum_{1<k\leq k_F} \sum_{w \geq n_F} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{k,r}^w (n,p,R)} T((p_{(l)} \cap p_{k,r}^{2,n}@R)) (g) .
\end{equation}

We define the \textbf{main term} of $T^{p_{(l)}}$ as

\begin{equation}
T^{p_{(l)}}_{\text{Main}} (g) := T^{p_{(l)},1} (g) + T^{p_{(l)},2} (g) ,
\end{equation}

where:

• the $L^1$-component (or \textit{low-height} component) is given by:

\begin{equation}
T^{p_{(l)},1} (g) := \sum_{1<k\leq k_F} \sum_{w \geq n_F} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{k,r}^w (n,p,R)} T((p_{(l)} \cap p_{k,r}^{2,n}@R)) (g) .
\end{equation}

• the $L^2$-component (or \textit{high-height} component) is given by:

\begin{equation}
T^{p_{(l)},2} (g) := \sum_{1<k\leq k_F} \sum_{w \geq n_F} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{k,r}^w (n,p,R)} T((p_{(l)} \cap p_{k,r}^{2,n}@R)) (g) .
\end{equation}

Appealing to (190), we create the corresponding analogues of the above tree components\footnote{Here it is more convenient to formulate these analogues in terms of $T^P_c$ rather than $T^P$ for appropriate families $\mathcal{P}$, where here $T^P_c$ is the direct adaptation of (162) to our context.}, as follows:

We split the term

\begin{equation}
\mathcal{E}_< = \mathcal{E}_{\text{Err}} + \mathcal{E}_{\text{Main}} ,
\end{equation}

with

\begin{equation}
\mathcal{E}_{\text{Err}} := \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_{n,\text{Err}}^p ,
\end{equation}

and

\begin{equation}
\mathcal{E}_{\text{Main}} := \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_{n,\text{Main}}^p ,
\end{equation}

where
• the **error term** is further subdivided into

\[(239) \quad E_{n, Err}^p := E_{n, he}^p + E_{n, whe}^p ,\]

with

– the **heavy** component given by

\[(240) \quad E_{n, he}^p (h, g) := \sum_{1 < k \leq k_F} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{he}(R)} \frac{|I_R \cap F|}{|I_R|} \int_{I_R} h T_{c_{k, he}}^{n, [p]} (R)^*(g) ,\]

– the \(w\)–**heavy** component given by

\[(241) \quad E_{n, whe}^p := \sum_{1 < k \leq k_F} \sum_{w \geq n \epsilon} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{w}(n, p, R)} \frac{|I_R \cap F|}{|I_R|} \int_{I_R} h T_{c_{k, re}}^{n, [p]} (R)^*(g) ,\]

• the **main term** is further subdivided into

\[(242) \quad E_{n, Main}^p := E_{n, L^1}^p + E_{n, L^2}^p ,\]

with:

– the \(L^1\)-**component** (or **low-height** component) given by:

\[(243) \quad E_{n, L^1}^p (h, g) := \sum_{1 < k \leq k_F} \sum_{w < n \epsilon} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{k, \leq 10 n \epsilon}} \frac{|I_R \cap F|}{|I_R|} \int_{I_R} h T_{c_{k, re}}^{n, [p]} (R)^*(g) ,\]

– the \(L^2\)-**component** (or **high-height** component) given by:

\[(244) \quad E_{n, L^2}^p (h, g) := \sum_{1 < k \leq k_F} \sum_{w < n \epsilon} \sum_{R \in \mathcal{R}_k} \sum_{R \in \mathcal{R}_{k, > 10 n \epsilon}} \frac{|I_R \cap F|}{|I_R|} \int_{I_R} h T_{c_{k, re}}^{n, [p]} (R)^*(g) .\]

**Proposition 37.** With the above notations, we have

\[(245) \quad |E_{Main}(h, g)| \lesssim |F| \log \left( \frac{4}{|F|} \right) \|h\|_\infty \|g\|_\infty .\]

and

\[(246) \quad |E_{Err}(h, g)| \lesssim |F| \log \left( \frac{4}{|F|} \right) \|h\|_\infty \|g\|_\infty .\]
11. Treatment of the error term $\mathcal{E}_{Err}$

Our goal in this section is to prove (246).
Based on (237)-(241), we have

(247) \[ \mathcal{E}_{Err} = \mathcal{E}_{he} + \mathcal{E}_{whe} , \]

where

(248) \[ \mathcal{E}_{he} := \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_{p,n,he} , \]

and

(249) \[ \mathcal{E}_{whe} := \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_{p,n,whe} . \]

11.1. The term $\mathcal{E}_{he}$. We focus now our attention towards the term $\mathcal{E}_{n,he}$.

Recalling now (178) and the convention made in (221) we first remark the following key fact:

(250) \[
|T_{c,2,n,(m),[p]}^k [R]^* (g)(x)|_{x \in I_R} = \left\| \begin{array}{c}
T_{c,2,n,(m),[p]}^k [R]^* (g) \\
|I_R|^{\frac{1}{2}}
\end{array} \right\|_{L^2(I_R)}
\]

Next, maintaining the assumption in (221), that is

(251) \[ \mathcal{P}_{k,(l),a,b}^{2,n,(m),[p]} [R] = \mathcal{P}_{k,(l),a,b}^{2,n,(m),[p]} [R] , \]

we notice that by applying the same techniques as in the proof of Lemma 27 - see (180), we get that for any $x \in I_R \cap F$ one has

(252) \[ |T_{c,2,n,(m),[p]}^k (g)(x) - T_{c,2,n,(m),[p]}^k (g)(x)| \lesssim \sum_{p \in \mathcal{P}_{k,(l),a,b}^{2,n,(m),[p]} [R]} \frac{|I_R| \int_{E(P)} |g|}{|I_P|} . \]

Define $I_{\mathcal{P}_{n,(m),[p]}^{(l),(a,b)}}$ as the time interval of the top of the tree $\mathcal{P}_{n,(m),[p]}^{(l),(a,b)}$ and

let $\tilde{I}_{\mathcal{P}_{n,(m),[p]}^{(l),(a,b)}} := 2 \tilde{I}_{\mathcal{P}_{n,(m),[p]}^{(l),(a,b)}}$. Based on this, we notice that

$$\text{supp} T_{c,2,n,(m),[p]}^k (g) \subseteq \tilde{I}_{\mathcal{P}_{n,(m),[p]}^{(l),(a,b)}} .$$
Set now
\[
A_{t,a,b}^{n,m,p}(F) := \left\{ x \in \tilde{I}_{p_{(t),(l),a,b}}^{n,(m),[p]} \cap F \mid \| T_{p_{(t),(l),a,b}}^{n,(m),[p]} \ast (g)(x) \| L^2(F) \geq 2^{n-1} \frac{\| T_{p_{(t),(l),a,b}}^{n,(m),[p]} \ast (g) \| L^2(F)}{(2^{-m} \| \tilde{I}_{p_{(t),(l),a,b}}^{n,(m),[p]} \| F)^{1/2}} \right\}.
\]

Recalling definition (222) and based on (39)-(253), we now have
\[
\left| \mathcal{E}^p_{n,h,e}(h, g) \right| \lesssim \sum_{1 < k \leq k_F} \sum_{R \in R_k} \sum_{R \in R_k^{i(h)}} |I_R \cap F| \int_{I_R} \left| T_{R_k,h,e}^{n,m,[p]}(g) \right|
\]
\[
\lesssim \sum_{1 < k \leq k_F} \sum_{R \in R_k} \sum_{R \in R_k^{i(h)}} \sum_l \int_{I_R \cap F} \left| T_{R_k,l,h,e}^{n,m,[p]}(g) \right|
\]
\[
\lesssim \sum_l \sum_{1 < k \leq k_F} \sum_{R \in R_k} \sum_{R \in R_k^{i(h)}} |I_R \cap F| \sum_{p \in P_{R_k,l,h,e}^{n,m,[p]} \cap I_R} \frac{|I_R|}{|I_P|} \frac{|I_P|}{|I_F|}
\]
\[
+ \sum_l \int_{A_{t,a,b}^{n,m,p}(F)} |T_{p_{(t),(l),a,b}}^{n,(m),[p]} \ast g|
\]
\[
=: U^{n,[p]} + \sum_l \int_{A_{t,a,b}^{n,m,p}(F)} |T_{p_{(t),(l),a,b}}^{n,(m),[p]} \ast g|.
\]

Now from Chebyshev we deduce that
\[
|A_{t,a,b}^{n,m,p}(F)| \lesssim 2^{-2^{n}} 2^{-m} |\tilde{I}_{p_{(t),(l),a,b}}^{n,(m),[p]}|.
\]

Inserting now (255) in (254) and using Cauchy-Schwarz and Corollary 23, we deduce
\[
\left| \mathcal{E}^p_{n,h,e}(h, g) \right| \lesssim U^{n,[p]} + \sum_l \sum_{p \in P_{(t),a,b}^{n,(m),[p]} \in P_{(t),a,b}^{n,[p]}} |A_{t,a,b}^{n,m,p}(F)| \frac{1}{2} \left( \int_{F} \left| T_{p_{(t),(l),a,b}}^{n,(m),[p]} \ast g \right|^2 \right)^{1/2}
\]
\[
\lesssim U^{n,[p]} + 2^{-m} \sum_l \sum_{p \in P_{(t),a,b}^{n,(m),[p]} \in P_{(t),a,b}^{n,[p]}} 2^{-n} 2^{-m} |\tilde{I}_{p_{(t),(l),a,b}}^{n,(m),[p]}|.
\]

Finally, recalling the proof of Lemma 28 that relies on the estimate of the term defined in (183), we conclude
\[
\left| \mathcal{E}_{h,e}(h, g) \right| \lesssim
\]
\begin{equation}
U + \sum_{n \leq a_F} \sum_{p \leq k_F} 2^{-ne} \sum_{l} \sum_{t} 2^{-n} 2^{-m} |I_{P_{n,\{m\},[p]}}^{n,\{m\},[p]}_{(l),a,b} | \leq |F| \log 4 |F|.
\end{equation}

11.2. The term $\mathcal{E}_{\text{whc}}$. We pass to analyzing the term $\mathcal{E}_{\text{whc}}^m$.

Departing from definition (241), and making use of (228), we have (258)

\begin{equation}
|\mathcal{E}_{\text{whc}}^p(h, g)| \lesssim \sum_{1 < k \leq k_F} \sum_{w \geq n_w} \sum_{R \in \mathcal{R}^k} \sum_{R \in \mathcal{R}_k^{s,w}(n, p, R)} \int_{IR} |I_{\mathcal{E}}^{s, n, [p]}[R]^*| (g)
\end{equation}

\begin{equation}
\lesssim \sum_{1 < k \leq k_F} \sum_{w \geq n_w} \sum_{R \in \mathcal{R}^k} \sum_{R \in \mathcal{R}_k^{s,w}(n, p, R)} |IR| 2^w 2^{-n} 2^{ne} (\mathcal{H}_{n, p}^* \mathcal{H}_{n, p})^2.
\end{equation}

We stop here for introducing some useful notation.

First of all, for a fixed $I \in \mathcal{I}_k$ and $R \in \mathcal{R}_k^{s_1 \ldots s_r}[I]$, we notice that any children $R \in \mathcal{R}_k^{s_{1 \ldots r}}(R)$ must have the form $R \in \mathcal{R}_k^{s_{1 \ldots r}}(R)$ with $r' > r$ such that $s_{j'} = s_j$ if $j \leq r$, $s'_{r+1} = U$ and, if $r' > r + 1$ then $s'_{j} = L$ if $r + 1 < j \leq r'$.

Now, if $s = (s_1, \ldots, s_r)$ represents the standard notation for a given sequence, then we will write $R \in \mathcal{R}_k^{s_{1 \ldots r}}[I]$ as $R \in \mathcal{R}_k^s[I]$ and $R \in \mathcal{R}_k^{s_{1 \ldots r}}(R)$ with $R \in \mathcal{R}_k^{s_{1 \ldots r}}(R)$ as $R \in \mathcal{R}_k^{s_{1 \ldots r}}(R)$. This notation will become very useful when distinguishing among the generations of the children within $\mathcal{R}_k^{s_{1 \ldots r}}(R)$.

Now notice that given any two $R, R' \in \mathcal{R}_k^{s_{1 \ldots r}}(n, p, R)$ one has

\begin{equation}
\mathcal{P}_{k, re}^{2, n, [p]}[R'] = \mathcal{P}_{k, re}^{2, n, [p]}[R],
\end{equation}

and thus the following is well defined:

\begin{equation}
\mathcal{P}_{k, re}^{2, n, [p]}[\mathcal{R}_k^{s_{1 \ldots r}}(n, p, R)] := \mathcal{P}_{k, re}^{2, n, [p]}[R] \text{ for some } R \in \mathcal{R}_k^{s_{1 \ldots r}}(n, p, R).
\end{equation}

Also, we set

\begin{equation}
\mathcal{I}_{\mathcal{R}_k^{s_{1 \ldots r}}(n, p, R)} := \bigcup_{R \in \mathcal{R}_k^{s_{1 \ldots r}}(n, p, R)} IR,
\end{equation}

and for any $R \in \mathcal{R}_k^{s_{1 \ldots r}}(n, p, R)$ based on (225) and (226) it makes sense to define

\begin{equation}
\mathcal{H}_{\mathcal{R}_k^{s_{1 \ldots r}}(n, p, R)} := \mathcal{H}_{n, p}^R.
\end{equation}

Then, (258) can be rewritten as

(261)
\[ |\mathcal{E}_{n,h,g}^p(h, g)| \lesssim \sum_{1 < k \leq k_F} 2^{-k} \sum_{I \in \mathcal{I}_k} \sum_{s} \sum_{R \in \mathcal{R}_k^p[I]} \sum_{\nu \geq 0} \sum_{w \geq \nu} |\tilde{T}_{R_k^p,r,[w]}(n,p,R)| 2^w 2^{-n} 2^{n\epsilon} (\mathcal{H}_{R_k^p,r,[w]}(n,p,R))^{\frac{1}{2}}. \]

Now, corroborating the following elements:

- definition of the tiles inside \( P_{k, r, \ell}^d \) and their frequency position relative to \( R \);
- the lacunary structure of the frequencies of our tiles;
- relations (223) and (228);

we can apply Zygmund’s inequality (see [10] and also Section 4 in [25]), to deduce the following key John-Nirenberg type condition:

there exists \( c > 0 \) absolute constant such that

(262)
\[ |\tilde{T}_{R_k^p,r,[w]}(n,p,R)| \lesssim e^{-c2^{2w}} |I_R|. \]

We now have

**Lemma 38.** Fix \( I \in \mathcal{I}_k \) and \( R \in \mathcal{R}_k^p[I] \).

Then, with the above notations, for any \( 0 < \epsilon \leq \frac{1}{2} \) the following holds:

(263)
\[ \sum_{r \geq 0} |\tilde{T}_{R_k^p,r,[w]}(n,p,R)| (\mathcal{H}_{R_k^p,r,[w]}(n,p,R))^{\frac{1}{2}} \lesssim \epsilon e^{-c2^{2w}} |I_R| (\mathcal{H}_{R_k^p,0}(n,p,R))^{\frac{1}{2}+\epsilon}, \]

where above we denoted with

\[ \mathcal{H}_{R_k^p,0}(n,p,R) := \# \{ I | \tilde{T}_{k,(l)}^{2,n,[p]}[R] \in P_{k, r, \ell}^d \text{ for some } R \in \mathcal{R}_{k,c,0}^s(n,p,R) \}, \]

and here \( c > 0 \) is the same as in (262).

**Proof.** The above statement and its proof here can be regarded as a refinement/specialization of Lemma 33.

We start by noticing that given any \( r' \geq r \) one has

\[ P_{k, r, \ell}^{2,n,[p]}[R_{k, c, 0}^s U_{r, \ell}^d[n,p,R]] \subseteq P_{k, r, \ell}^d[n,p,R], \]

which in particular implies that

(264)
\[ \mathcal{H}_{R_k^p,r,[w]}(n,p,R) \leq \mathcal{H}_{R_k^p,r,[w]}(n,p,R) \leq \mathcal{H}_{R_k^p,0}(n,p,R). \]

Next, based on (264), we observe that in order to prove (263) it is enough to show that

(265)
\[ \sum_{\mathcal{H}_{R_k^p,r,[w]}(n,p,R) > \frac{1}{2}} |\tilde{T}_{R_k^p,r,[w]}(n,p,R)| (\mathcal{H}_{R_k^p,r,[w]}(n,p,R))^{\frac{1}{2}} \leq \epsilon e^{-c2^{2w}} |I_R| (\mathcal{H}_{R_k^p,0}(n,p,R))^{\frac{1}{2}+\epsilon}. \]
Indeed, once (265) is proved, one can iterate the same argument in order to obtain for \( k \in \mathbb{N} \) the general estimate

\[
\sum_{\frac{1}{2k+1} H_{k}(n,p,R) < H_{k}(n,p,R) \leq \frac{1}{2k} H_{k}(n,p,R)} \frac{\mathcal{I}_{\mathcal{R}_{k}^{s,u}(n,p,R)}}{2k+1} \leq 2^w e^{-c2^w |I_{R}|} 
\]

which summing up over \( k \) gives the desired estimate (263).

Thus, it remains now to prove (265).

We next observe that wlog we may suppose that

\[
r \neq r' \iff H_{k}(n,p,R) \neq H_{k}(n,p,R) \.
\]

Indeed, otherwise we can regroup the terms in the LHS of (265) and notice that letting \( \mathcal{I}_{\mathcal{R}_{k}^{s,u}(n,p,R)} \) be the union of the spacial intervals \( \{\mathcal{I}_{\mathcal{R}_{k}^{s,u}(n,p,R)}\} \) having the property that \( H_{k}(n,p,R) = \alpha \) for some fixed \( \alpha \in \mathbb{N} \), one still has the analogue of (262), that is

\[
|\mathcal{I}_{\mathcal{R}_{k}^{s,u}(n,p,R)}| \lesssim e^{-c2^w |I_{R}|}.
\]

With these done, we now proceed as follows: fix \( \epsilon \in (0, \frac{1}{2}] \) as appearing in our hypothesis and assume wlog that \( (\mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)})^\epsilon, (\mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)})^{1-\epsilon} \in \mathbb{N} \).

For \( i \in \{0, \ldots, \frac{1}{2} \} \) \( (\mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)})^\epsilon - 1 \) let us set

\[
A_i := \mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)} - \left(i (\mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)})^{1-\epsilon} + (i + 1) (\mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)})^{1-\epsilon} \right)
\]

and choose \( r_i \) to be the smallest \( r \) for which \( \mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)} \in A_i \).

Let now

\[
F_0 = F_{\mathcal{R}_{k}^{s,u}(n,p,R)} := \{l \mid \mathcal{P}_{k,l}^{2,n,[p]}[R] \in \mathcal{P}_{k,l}^{2,n,[p]}[R] \text{ for some } R \in \mathcal{R}_{k,l}^{s,U}(n,p,R)\},
\]

and similarly, for \( p > 0 \), set

\[
F_r = F_{\mathcal{R}_{k}^{s,r,u}(n,p,R)} := \{l \mid \mathcal{P}_{k,l}^{2,n,[p]}[R] \in \mathcal{P}_{k,l}^{2,n,[p]}[R] \text{ for some } R \in \mathcal{R}_{k,l}^{s,r,u}(n,p,R)\}.
\]

Deduce that

\[
F_{r+1} \subsetneq F_r \text{ and that } \#F_r = \mathcal{H}_{\mathcal{R}_{k}^{s,r,u}(n,p,R)}.
\]

Observe now that (265) follows if for any \( i \in \{0, \ldots, \frac{1}{2} (\mathcal{H}_{\mathcal{R}_{k}^{s,u}(n,p,R)})^\epsilon - 1 \} \) we show that
We now rephrase (228) in the form
\[ \frac{1}{2^{-n} 2^{n \epsilon} (\#F_r)^{\frac{1}{2}}} \| \sum_{l \in F_r} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g \|_{L^\infty} \approx 2^w. \]

**Observation 39.** It is worth noticing that for any \( R \in R_{k, \chi, \ell}^{s, \ell, r, L, [w]} (n, p, R) \) with \( w >> 1 \) we have that the term \( \sum_{l \in F_r} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g \) is morally constant on \( I_R \), in the sense that
\[ \sup_{x \in I_R} | \sum_{l \in F_r} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g (x) | \approx \inf_{x \in I_R} \left| \sum_{l \in F_r} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g (x) \right| \approx 2^w 2^{-n} 2^{n \epsilon} (\#F_r)^{\frac{1}{2}}. \]
Indeed, this follows from a similar reasoning with the one used to prove heuristic 169 (see in particular (174)), together with the estimate (270).

As a consequence of the above observation, applying Zygmund’s inequality, we notice that for any \( r_i \leq r < r_{i+1} \) and for each \( R \in R_{k, \chi, \ell}^{s, \ell, r, L, [w]} (n, p, R) \) (remark from (259) that \( I_R \subseteq T_{R_h^*, r, [w]} (n, p, R) \)) we have
\[ \ln \frac{|I_R|}{|I_R|} \leq -c \max \left\{ \frac{\| \sum_{l \in F_{r_{i+1}}} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g \|_{L^\infty (I_R)}}{2^{-n} 2^{n \epsilon} (\#F_{r_{i+1}})^{\frac{1}{2}}} \right\}^2 \frac{\| \sum_{l \in F_r \setminus F_{r_{i+1}}} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g \|_{L^\infty (I_R)}}{2^{-n} 2^{n \epsilon} (\#(F_r \setminus F_{r_{i+1}}))^{\frac{1}{2}}} \]

Set now
\[ Y^i_r (R) := \frac{\| \sum_{l \in F_{r_{i+1}}} T_{c, k, (l)}^{2, n, [p]} [R_h^*, r, [w]]_n (n, p, R) g \|_{L^\infty (I_R)}}{2^{-n} 2^{n \epsilon} (\#F_{r_{i+1}})^{\frac{1}{2}}}, \]
and using now (270) and triangle inequality we rewrite (272) as (273)
\[
\ln \left| \frac{I_R}{|I_R|} \right| \lesssim -c \max \left\{ Y^i_R(R), 2^w \left( \frac{\#F_r}{\#(F_r \setminus F_{r+1})} \right)^{\frac{1}{2}} - Y^i_R(R) \left( \frac{\#F_{r+1}}{\#(F_r \setminus F_{r+1})} \right)^{\frac{1}{2}} \right\}^2
\]

Define now the following quantities:

\[
\mathcal{R}^u_{k, \chi, b}(n, p, R) := \{ R \in \mathcal{R}^u_{k, \chi, b} \mid Y^i_R(R) \approx 2^w \}, \\
\mathcal{R}^h_{k, \chi, to}(n, p, R) := \mathcal{R}^u_{k, \chi, b}(n, p, R) \setminus \mathcal{R}^u_{k, \chi, b}(n, p, R)
\]

and

\[
\tilde{I}_{\mathcal{R}^u_{k, \chi, b}}(n, p, R) := \bigcup_{R \in \mathcal{R}^u_{k, \chi, b}(n, p, R)} I_R
\]

The key two observations are the following:

\[(274) \quad \sum_{r=r_i}^{r_i+1-1} |\tilde{I}_{\mathcal{R}^u_{k, \chi, b}}(n, p, R)| \lesssim e^{-c2^w} |I_R|,
\]

and

\[(275) \quad |\tilde{I}_{\mathcal{R}^u_{k, \chi, b}}(n, p, R)| \lesssim e^{-c2^w} \frac{\#F_r}{\#(F_r \setminus F_{r+1})} |I_R|.
\]

With these, setting \( L := \mathcal{H}_{\mathcal{R}^u_k(n, p, R)} \) we now have that

\[(276) \quad \sum_{r=r_i}^{r_i+1-1} |\tilde{I}_{\mathcal{R}^u_{k, \chi, b}}(n, p, R)| = \sum_{r=r_i}^{r_i+1-1} |\tilde{I}_{\mathcal{R}^u_{k, \chi, b}}(n, p, R)| + \sum_{r=r_i}^{r_i+1-1} |\tilde{I}_{\mathcal{R}^u_{k, \chi, to}}(n, p, R)| \\
\lesssim e^{-c2^w} |I_R| \left( 1 + \sum_{r=r_i}^{r_i+1-1} e^{-c \frac{\#F_r}{\#(F_r \setminus F_{r+1})}} \right) \\
\lesssim e^{-c2^w} |I_R| \left( 1 + \sum_{l=1}^{L^{1-\epsilon}} \left( \frac{l}{L} \right)^{\frac{1}{2}} \right) \\
\lesssim e^{-c2^w} |I_R|.
\]
Returning now to estimate (261) and using (263) we have
\[
|\mathcal{E}_{n,\text{whe}}(h, g)| \lesssim \sum_{1 < k \leq k_F} 2^{-k} \sum_{I \in I_k} \sum_{s \in \mathbb{R}_k^{R_k^s(n,p,R)}} \sum_{R \in \mathbb{R}_k^{R_k^s(n,p,R)}} e^{-c2^{2w}} 2^w 2^{ne} 2^{-n} \mathcal{H}_{\mathbb{R}_k^s(n,p,R)} |I_R|
\]
\[
\lesssim 2^{-ne} \sum_{l} F_{l,n,(m),[p]} \sum_{l \in \mathbb{P}_{l,(i),a,b}} 2^{-n} 2^{-m} |\tilde{I}_{F_{l,n,(m),[p]}}| .
\]
Thus
\[
\text{(277)}
\]
\[
|\mathcal{E}_{\text{whe}}(h, g)| = \left| \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_{n,\text{whe}}(h, g) \right|
\]
\[
\lesssim \sum_{n \leq n_F} \sum_{p \leq k_F} 2^{-ne} \sum_{l} F_{l,n,(m),[p]} \sum_{l \in \mathbb{P}_{l,(i),a,b}} 2^{-n} 2^{-m} |\tilde{I}_{F_{l,n,(m),[p]}}| \lesssim |F| \log \frac{4}{|F|} .
\]

12. TREATMENT OF THE MAIN TERM $\mathcal{E}_{\text{Main}}$

12.1. Treatment of the $L^2$-term $\mathcal{E}_{L^2} := \sum_{n \leq n_F} \sum_{p \leq k_F} \mathcal{E}_{n,L^2}^p$.

We start by analyzing the term $\mathcal{E}_{n,L^2}^p$.

Making now use of (228) and (244) we deduce that
\[
\text{(278)}
\]
\[
|\mathcal{E}_{n,L^2}^p(h, g)| \lesssim \sum_{1 < k \leq k_F} \sum_{w \leq n_e} \sum_{s \in \mathbb{R}_k^{R_k^s(n,p,R)}} \sum_{R \in \mathbb{R}_k^{R_k^s(n,p,R)}} \frac{|I_R \cap F|}{|I_R|} \int_{I_R} \left| T_{c,k,(l,a,b)}^{2,n,[p]} |g| \right|^{*}.
\]
\[
\lesssim \sum_{1 < k \leq k_F} \sum_{w \leq n_e} \sum_{s \in \mathbb{R}_k^{R_k^s(n,p,R)}} \sum_{R \in \mathbb{R}_k^{R_k^s(n,p,R)}} 2^{2ne} \left( \mathcal{H}_{n,p,R} \right) \frac{1}{2} 2^{-n} |I_R \cap F|
\]
\[
\lesssim \sum_{1 < k \leq k_F} \sum_{w \leq n_e} \sum_{s \in \mathbb{R}_k^{R_k^s(n,p,R)}} \sum_{R \in \mathbb{R}_k^{R_k^s(n,p,R)}} 2^{-3ne} \mathcal{H}_{n,p,R} 2^{-n} |I_R \cap F|
\]
\[
\lesssim F_{\text{ubini}} 2^{-3ne} \sum_{l} \sum_{P_{l,(i),a,b} \in \mathbb{P}_{l,(i),a,b}} 2^{-n} |\tilde{I}_{P_{l,(i),a,b}}| \cap F| .
\]
If we now set
\[ E_{n,L}^2 := \sum_{p \leq k_F} E_{n,L}^p, \]
from above, we deduce that
\[ |E_{n,L}^2(h,g)| \lesssim 2^{-3nt} \sum_{p \leq k_F} \sum_{l} 2^{-n} |I_{p,l(a,b)} \cap F| \lesssim 2^{-3nt} |F| \log \frac{4}{|F|}, \]
which implies
\[ |E_{L}^2(h,g)| = \left| \sum_{n \leq n_F} |E_{n,L}^2(h,g)| \right| \lesssim |F| \log \frac{4}{|F|}. \]

12.2. Treatment of the $L^1$-term $E_{L}^1 := \sum_{n \leq n_F} \sum_{p \leq k_F} E_{n,L}^p$.

This entire section is dedicated to the analysis of the term $E_{L}^1$. Along the way, we will develop a methodology that will treat simultaneously families of tiles with distinct mass parameters.

Indeed, unlike the previous cases, we will not be able - for the generic case - to obtain direct decay in the mass parameter $n$ for the term $E_{n,L}^p$; instead we will need to play a subtle game between the mass and $F$-mass scales of the maximal trees involved in our decomposition. Our approach will be based on two new ideas:

- the tree foliation of the time-frequency plane introduced in Section 12.2.2;
- the $(E)$-mass versus $F$-mass dichotomy defined in Section 12.2.3.

12.2.1. Identifying the main component of $E_{L}^1$. We start by recalling the main term analyzed in this section:

\[ E_{L}^1(h,g) = \sum_{n \leq n_F} \sum_{p \leq k_F} E_{n,L}^p(h,g) \]
\[ = \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 \leq k \leq p} \sum_{w \leq n \varepsilon_{R_k(n,p,R)}} \sum_{\substack{R \in R_k \cap I_{R} \cap F \subset 210 n\varepsilon_{R_k(n,p,R)} \subset \mathbb{R}}} |I_{R} \cap F| \int_{I_{R}} h_{T_{c,\varepsilon_{R_k(n,p,R)}}}^{p^{n,p} R} (g) \]
\[ \lesssim \sum_{n \leq n_F} \sum_{p \leq k_F} 2^{-k} \sum_{1 \leq k \leq p} \sum_{w \leq n \varepsilon_{R_k} \cap \mathbb{R}} \sum_{\substack{R \in R_k \cap I_{R} \cap F \subset 210 n\varepsilon_{R_k(n,p,R)} \subset \mathbb{R}}} \int_{I_{R}} h_{T_{c,\varepsilon_{R_k(n,p,R)}}}^{p^{n,p} R} (g). \]

We introduce the following
Definition 40. Let $1 \leq p \leq k_F$. We say that $\mathcal{P}^{n,(m)}_{(l),a,b} = \mathcal{P}^n_{(l),a,b}$ is an $(L^1, F)$-saturated tree iff

(282)

\[
\sum_{1 < k \leq k_F} \sum_{w \leq n} \sum_{R \in \mathbb{R}_k} \sum_{R \in \mathbb{R}_k^{n,w}} 2^{-k} |I_R \mathcal{P}^{n,(m)}_{k,l}(a,b)(R)| \geq \frac{2^{-m}}{2^{2n\epsilon}} |I_{\mathcal{P}^{n,(m)}_{(l),a,b}}|,
\]

where, as expected, here $\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R) := \mathcal{P}^{2,n}_{k,l}(a,b)(R) \cap \mathcal{P}^{n,(m)}_{(l),a,b}$ and $I_{\mathcal{P}^{2,n,(m)}_{(l),a,b}} := I_R \cap I_{\mathcal{P}^{n,(m)}_{(l),a,b}}$.

We let $\mathcal{F}_{\text{sat}}$ be the collection of all $\mathcal{P}^{n,(m)}_{(l),a,b}$ with $n \leq n_F, p \leq k_F, a, b \geq 1$ which are $(L^1, F)$-saturated trees.

With these done, we decompose

(283)

\[
\mathcal{E}_{L^1}(h, g) = \mathcal{E}^{\text{sat}}_{L^1}(h, g) + \mathcal{E}^{\text{nsat}}_{L^1}(h, g),
\]

where

(284)

\[
\mathcal{E}^{\text{sat}}_{L^1}(h, g) := \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 < k \leq k_F} \sum_{w \leq n} \sum_{R \in \mathbb{R}_k} \sum_{R \in \mathbb{R}_k^{n,w}} \frac{|I_R \cap F|}{|I_R|}
\]

\[
\times \sum_{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R) \in \mathcal{F}_{\text{sat}}} \int_{I_R} h T_{c}^{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R)}(g)
\]

with the obvious correspondence for the term $\mathcal{E}^{\text{nsat}}_{L^1}(h, g)$.

Claim 41. The term $\mathcal{E}^{\text{nsat}}_{L^1}(h, g)$ in (283) is an error term, that is

(285)

\[
|\mathcal{E}^{\text{nsat}}_{L^1}(h, g)| \lesssim |F| \log \frac{4}{|F|} \|f\|_{\infty} \|g\|_{\infty}.
\]

Proof. Using now (223), (224) and (39) we have

\[
|\mathcal{E}^{\text{nsat}}_{L^1}(h, g)| = \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 < k \leq k_F} \sum_{w \leq n} \sum_{R \in \mathbb{R}_k} \sum_{R \in \mathbb{R}_k^{n,w}} \frac{|I_R \cap F|}{|I_R|}
\]

\[
\times \sum_{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R) \in \mathcal{F}_{\text{nsat}}} \int_{I_R} h T_{c}^{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R)}(g)
\]

\[
\lesssim \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 < k \leq k_F} \sum_{w \leq n} \sum_{R \in \mathbb{R}_k} \sum_{R \in \mathbb{R}_k^{n,w}} \frac{|I_R \cap F|}{|I_R|}
\]

\[
\times \sum_{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R) \in \mathcal{F}_{\text{nsat}}} \int_{I_R} h T_{c}^{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R)}(g)
\]

\[
\lesssim \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 < k \leq k_F} \sum_{w \leq n} \sum_{R \in \mathbb{R}_k} \sum_{R \in \mathbb{R}_k^{n,w}} \frac{|I_R \cap F|}{|I_R|}
\]

\[
\times \sum_{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R) \in \mathcal{F}_{\text{nsat}}} \int_{I_R} h T_{c}^{\mathcal{P}^{2,n,(m)}_{k,l}(a,b)(R)}(g)
\]
\[
\times \sum_{p^{2,n,\lceil m \rceil},[p] \in \mathcal{F}_{\text{sat}}^{2,n,\lceil m \rceil},[p]} |I_R| \frac{\| T^{p^{2,n,\lceil m \rceil},[p]}_{k,\langle l \rangle,\alpha,\beta} (g) \|_{L^2(I_R)}}{|I_R|^\frac{1}{2}} \leq \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 < k \leq k_F} \sum_{\omega < n_F} \sum_{R \subset \mathbb{R}_k^{\lceil m \rceil},[w]} \sum_{R \subset \mathbb{R}_k^{\lceil m \rceil},[w]} |I_R \cap F| \leq \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{\omega < n_F} |F| \log \frac{4}{|F|}.
\]

We are now left with the treatment of the saturated term \( \mathcal{E}^{\text{sat}}_{L^1} \). In this situation we can no longer use almost orthogonality techniques among various collections of trees. Thus, our very first step evolves as follows: within each 0 frequency tile \( R \) we use triangle inequality - relative to \( \| \cdot \|_1 \) - among distinct frequency trees and then for each frequency \( l \) we apply Cauchy-Schwarz and Corollary 23. That is

\[(286) \quad |\mathcal{E}^{\text{sat}}_{L^1}(h,g)| \leq \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{1 < k \leq k_F} \sum_{\omega < n_F} \int_{I_R \cap F} |I_R \cap F| \leq \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{\omega < n_F} |F| \log \frac{4}{|F|}.
\]

In what follows, for notational simplicity we will ignore the error term \( \mathcal{E}^{\text{sat}}_{L^1}(h,g) \) and thus wlog consider we have the following estimate

\[(287) \quad |\mathcal{E}_{L^1}(h,g)| \approx \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{\omega < n_F} 2^{-n} 2^{-m} |I_{p^{2,n,\lceil m \rceil},[p]}_{k,\langle l \rangle,\alpha,\beta}| .
\]
12.2.2. Tree foliation of the time-frequency plane depending on the mass and \( F \)-mass parameters. In this section, in parallel with the process described in Section 5.3, we develop an algorithm that organizes the family of tiles involved in \( E_{L,1} \) into maximal trees having uniform mass and \( F \)-mass parameters according to a “descending foliation pattern” that focuses directly on the spatial support of these trees. This contrasts with the analysis performed in Section 5.3 that was centered on the spatial support of the adjoint operators.

Recall first the decomposition made in Section 5.3, specifically relations (83) - (88).

Define now \( \mathcal{T} \) as the set of all tops \( P_{(l),a,b}^{n,(m),\{s\}} \) of the trees \( P_{(l),a,b}^{n,(m),\{s\}} \in \mathcal{F}_{\text{sat}} \).

Next, we let \( \mathcal{T}(l) \) be the collection of all tops inside \( \mathcal{T} \) which live at the frequency \( l \). Decompose this set into maximal layers \( \{\mathcal{T}(l,s)\}_{s \geq 1} \) as follows:

- for \( s = 1 \) the set \( \mathcal{T}(l,1) \) defines the collection of maximal, mutually disjoint top-trees at frequency \( l \);
- then, we remove the collection \( \mathcal{T}(l,1) \) from \( \mathcal{T}(l) \) and repeat the above selection procedure completing thus the second level (layer) \( \mathcal{T}(l,2) \).
- this process will end up in finitely many steps; in fact one can say even more: there are at most \( n_F + k_F \) such layers.

**Notation 42.** From now on, we redenote \( P \in \mathcal{T}(l,s) \) as

\[
P_{(l),a}^{n,(m),\{s\}},
\]

iff \( P \) is the top of a maximal tree \( \mathbb{P}_{(l),a}^{n,(m),\{s\}} \) with the following characteristics:

- lives at the frequency \( \omega_l \);
- has uniform mass \( \approx 2^{-n} \);
- has uniform \( F \)-mass \( \approx 2^{-m} \).

Further on, we will apply the following convention: for \( P_{(l),a}^{n,(m),\{s\}} \in \mathcal{T}(l,s) \) we define

\[
\mathcal{T}(l,s+1)[P_{(l),a}^{n,(m),\{s\}}] := \{ P_{(l),a'}^{n,(m'),\{s+1\}} \in \mathcal{T}(l,s+1) | P_{(l),a'}^{n,(m'),\{s+1\}} < P_{(l),a}^{n,(m),\{s\}} \}.
\]

We now have the following important

**Observation 43.** Recalling Definitions 11 and 13, and with the above notations, the following holds

\[
\text{if } P_{(l),a'}^{n,(m'),\{s+1\}} \in \mathcal{T}(l,s+1)[P_{(l),a}^{n,(m),\{s\}}] \text{ then}^{40}
\]

- either \( m' = m - 1 \);
- or \( n' = n - 1 \).

---

40 The itemization below is not mutually exclusive.
Also, remark that always \( m - 1 \leq m' \leq m \) while \( n' \) is not always/necessarily related to \( n \); indeed, we have that if \( m' = m \) then \( n' = n - 1 \), while if \( m' = m - 1 \) the value of \( n' \) can be anything relative to \( n \).

12.2.3. \((E-) mass versus F-mass dichotomy.\) In what follows we introduce a partition of the set \( \mathcal{T}(l) \) into two components

\[
\mathcal{T}(l) = \mathcal{T}^E(l) \cup \mathcal{T}^F(l),
\]

with

- \( \mathcal{T}^E(l) \) the \((E-) mass\) component;
- \( \mathcal{T}^F(l) \) the \(F-mass\) component.

This will be done according to the following selection algorithm:

- let \( P^{n,(m),\{s\}}_{(l),a} \in \mathcal{T}(l, s) \) for some \( s \in \mathbb{N} \); now
  - if
    \[
    \sum_{P^{n,(m-1),(s+1)}_{(l),a'} \leq \mathcal{T}(l, s+1)[P^{n,(m),\{s\}}_{(l),a}]} 2^{-(m-1)} |\tilde{I}_{P^{n,(m-1),(s+1)}_{(l),a'}}| \leq \frac{7}{8} 2^{-m} |\tilde{I}_{P^{n,(m),\{s\}}_{(l),a}}|
    \]
    then we declare \( P^{n,(m),\{s\}}_{(l),a} \in \mathcal{T}^E(l) \).
  - otherwise we select \( P^{n,(m),\{s\}}_{(l),a} \in \mathcal{T}^F(l) \).

Now, we let

\[
\mathcal{T}^E := \bigcup_l \mathcal{T}^E(l) \quad \text{and} \quad \mathcal{T}^F := \bigcup_l \mathcal{T}^F(l).
\]

Let us set now

\[
\mathcal{E}^E_{L^1} := \sum_l \sum_{n \leq n_F} \sum_{s \leq k_F} \sum_{P^{n,(m),\{s\}}_{(l),a} \in \mathcal{T}^E(l)} 2^{-n} 2^{-m} |\tilde{I}_{P^{n,(m),\{s\}}_{(l),a}}|,
\]

\[
\mathcal{E}^F_{L^1} := \sum_l \sum_{n \leq n_F} \sum_{s \leq k_F} \sum_{P^{n,(m),\{s\}}_{(l),a} \in \mathcal{T}^F(l)} 2^{-n} 2^{-m} |\tilde{I}_{P^{n,(m),\{s\}}_{(l),a}}|,
\]

and deduce from (293), (294) and (287) that

\[
|\mathcal{E}_{L^1}(h, g)| \lesssim \mathcal{E}^E_{L^1} + \mathcal{E}^F_{L^1}.
\]

\footnote{In the remaining part of the section we prefer to stress the fact that the standard mass introduced in Definition 13 quantifies the behavior of the measurable function \( N(\cdot) \) in (25) and is thus directly related with the properties of the set \( E = \{E(P)\}_{P \in \mathcal{P}_m^m} \).}
12.2.4. Treatment of the $E$–mass component. We pass now to estimating the term $\mathcal{E}_{L_{1}}^{E}$.

We start by setting

\[
TT_{E,n}(l) := \{ P_{(l),a}^{n',(m),\{s\}} \in TT_{E} \mid n' = n \}.
\]

In a similar spirit with the row selection presented in the second half of Section 12.2.1, we perform a layer type decomposition of the set $TT_{E,n}(l)$ as follows:

\[
\begin{align*}
\bullet & \text{ we define } TT_{E,n}(l)(1) \text{ to be the collection of maximal disjoint tops within the set } TT_{E,n}(l); \\
\bullet & \text{ next, we erase from } TT_{E,n}(l) \text{ the set } TT_{E,n}(l)(1) \text{ and repeat the previous procedure of selecting the collection of maximal disjoint tops into a new set } TT_{E,n}(l)(2); \\
\bullet & \text{ iterate this procedure until the first time when the new initialized set } TT_{E,n}(l) \text{ is void and then stop. This way we construct the successive layers } \{TT_{E,n}(l)(r)\}_{r \geq 1}.
\end{align*}
\]

Now first, we observe that based on the Carleson type condition (292), we deduce the following key relation:

for any $l$, $n \leq n_F$ and $P \in TT_{E,n}(l)$ we have

\[
\sum_{P' \in TT_{E,n}(l)} A_{F}(P') |I_{P'}| \leq A_{F}(P) |I_{P}|.
\]

Combining now (293) and (298) we have

\[
\mathcal{E}_{L_{1}}^{E} \approx \sum_{l} \sum_{n \leq n_F} \sum_{s \leq k_F} \sum_{r} \sum_{P_{(l),a}^{n,(m),\{s\}} \in TT_{E,n}(l)(r)} 2^{-n} 2^{-m} |I_{P_{(l),a}^{n,(m),\{s\}}}|,
\]

which further, by using Definition 40, becomes

\[
\begin{align*}
\sum_{n \leq n_F} & \sum_{p \leq k_F} \sum_{1 \leq k \leq k_F} \sum_{I \in I_{k}} \bigg( \sum_{n \leq n_F} \sum_{p \leq k_F} \sum_{R \in R_{k}^{n,p}} \sum_{(n,p,R) \in TT_{E,n}(l)(1)} 2^{n_{\epsilon}} 2^{-n} 2^{-k} |I_{R_{(n,p,R)}^{n,p,(m),\{s\}}}| \bigg) \\
\lesssim & \sum_{1 \leq k \leq k_F} \sum_{I \in I_{k}} \sum_{l \geq 0} \sum_{R \in R_{k}^{l+R_{(n,p,R)}^{n,p,(m),\{s\}}(R)}} 2^{2n_{\epsilon}} 2^{-n} 2^{-k} |I_{R}|,
\end{align*}
\]
Consequently, from (303) and (304) we deduce that
\begin{align}
\sum_{\mathcal{P}^{n,(m),\{s\}} \in \mathcal{T}\mathcal{T}^F(l,s+1)[\mathcal{P}^{n,(m),\{s\}}]} 2^{-(m-1)} |\tilde{I}_{P^{n,(m),\{s\}}}^{n,(m),\{s\}}| > \frac{7}{8} 2^{-m} |\tilde{I}_{P^{n,(m),\{s\}}}^{n,(m),\{s\}}| ,
\end{align}

which thus further implies the key relation
\begin{align}
\sum_{\mathcal{P}^{n,(m),\{s\}} \in \mathcal{T}\mathcal{T}^F(l,s+1)[\mathcal{P}^{n,(m),\{s\}}]} 2^{-m} |\tilde{I}_{P^{n,(m),\{s\}}}^{n,(m),\{s\}}| \leq \frac{1}{8} 2^{-m} |\tilde{I}_{P^{n,(m),\{s\}}}^{n,(m),\{s\}}| .
\end{align}

At this point, rephrasing Observation 43 in our context, we notice that
\begin{align}
\text{if } \mathcal{P}^{n,(m),\{s\}} \in \mathcal{T}\mathcal{T}^F(l,s+1)[\mathcal{P}^{n,(m),\{s\}}] \text{ then } n' = n - 1 .
\end{align}

Consequently, from (303) and (304) we deduce that
\begin{align}
\sum_{\mathcal{P}^{n,(m),\{s\}} \in \mathcal{T}\mathcal{T}^F(l,s+1)[\mathcal{P}^{n,(m),\{s\}}]} A(\mathcal{P}^{n,(m),\{s\}},\mathcal{P}^{n,(m),\{s\}}) \mathcal{A}_{F}(\mathcal{P}^{n,(m),\{s\}},\mathcal{P}^{n,(m),\{s\}},\mathcal{P}^{n,(m),\{s\}}) |\tilde{I}_{P^{n,(m),\{s\}}}^{n,(m),\{s\}}| \leq \frac{1}{2} A(\mathcal{P}^{n,(m),\{s\}},\mathcal{P}^{n,(m),\{s\}},\mathcal{P}^{n,(m),\{s\}}) |\tilde{I}_{P^{n,(m),\{s\}}}^{n,(m),\{s\}}| .
\end{align}

In symmetry with (296) we let
\begin{align}
\mathcal{T}\mathcal{T}^F,m(l) := \{ \mathcal{P}^{n,(m'),\{s\}} \in \mathcal{T}\mathcal{T}^F \mid m' = m \} .
\end{align}

Also, applying an analogous procedure with the one presented in (297), we decompose each set \( \mathcal{T}\mathcal{T}^F,m(l) \) into successive (finitely many) layers of maximal rows
\begin{align}
\{ \mathcal{T}\mathcal{T}^F,m(l)(r) \}_{r \geq 1} .
\end{align}
Now (305) implies the fundamental analogue of (308): for any $l$, $m \leq k_F$ and $P \in \mathcal{T} \mathcal{T}^{F,[m]}(l)$ we have

\begin{equation}
\sum_{P' \in P} A(P') A_F(P') |\tilde{I}_{P'}| \leq A(P) A_F(P') |\tilde{I}_{P}|,
\end{equation}

(308)

With all these done, applying (308), we have

\begin{equation}
\mathcal{E}_l^F \lesssim \sum_{l \leq n_F} \sum_{m \leq k_F} \sum_{P \in \mathcal{T} \mathcal{T}^{F,[m]}(l)} A(P^n, \{s\}) A_F(P^n, \{s\}) |\tilde{I}_{P^n, \{s\}}|,
\end{equation}

(309)

\begin{equation}
\lesssim \sum_{l \leq n_F} \sum_{m \leq k_F} \sum_{P \in \mathcal{T} \mathcal{T}^{F,[m]}(l)} 2^{-m} A(P) |I_P|.
\end{equation}

(310)

where in the last line we used the special definition of the mass introduced in our current paper - i.e. Definition 13, and the following key observation:

\begin{equation}
\text{if } P \in \mathcal{T} \mathcal{T}^{F,[m]}(l)(1) \text{ then } A(P) = \frac{|E(P)|}{|I_{P}|}.
\end{equation}

Indeed, if $P \in \mathcal{T} \mathcal{T}^{F,[m]}(l)(1)$ with $P \in \mathcal{P}_n$ then from Definition 13 we know that there must exist $P' \in P^n$ with $P \leq P'$ such that $A(P) = \frac{|E(P')|}{|I_{P'}|} \in (2^{-n-1}, 2^{-n}]$. However, from the construction of our trees described in the first half of Section 5.3, $P$ must be the top of a maximal tree of a given $F$-mass and mass parameters. Thus since both $P, P' \in \mathcal{P}_n$ and $P \leq P'$ one must have $P = P'$.

We have now a last fundamental observation to record here:

Assume that we are given a (dyadic) interval $I \subset [0,1]$ and $\mathcal{P} = \mathcal{P}(I) \in \mathcal{P}$ any family of tiles obeying

- no two tiles in $\mathcal{P}$ are comparable under “$\leq$”;
- for any $P \in \mathcal{P}$ one has that $I_P \subseteq I$.

Then, we have the following fundamental Carleson-packing type condition:

\begin{equation}
\sum_{P \in \mathcal{P}(I)} |E(P)| \leq |I|.
\end{equation}

(311)
Implementing now (311) in (309) we conclude that
\[
\sum_l \sum_{P \in \mathcal{T}^{F, [m]}(l)} |E(P)| \lesssim |\tilde{I}_m|
\]
and hence
\[
\mathcal{E}_{L^1}^F \lesssim \sum_{m \leq k_F} \sum_l \sum_{P \in \mathcal{T}^{F, [m]}(l)} 2^{-m} |E(P)|
\]
\[
\lesssim \sum_{m \leq k_F} 2^{-m} |\tilde{I}_m| \lesssim k_F |F|.
\]

13. Final remarks

1) The techniques developed in this paper can easily be extended to the Walsh case. Indeed, recall first quickly several simple facts:

Fix \(n \in \mathbb{N}\) and let

\[
n = \sum_{i=0}^{\infty} \epsilon_i 2^i \quad \text{with} \quad \epsilon_i \in \{0, 1\},
\]

be it’s dyadic decomposition. We define the Walsh system \(\{w_n\}_{n \in \mathbb{N}}\) as:

\[
\begin{align*}
\bullet & \text{ if } x \in \mathbb{R} \setminus [0, 1) \text{ then } w_n(x) = 0 \text{ for any } n \in \mathbb{N}; \\
\bullet & \text{ if } x \in [0, 1) \text{ and } n = 0 \text{ then we set } w_n(x) = 1; \\
\bullet & \text{ if } x \in [0, 1) \text{ and } n \geq 1 \text{ obeys (314) then we let}^{42} \\
& w_n(x) := \prod_{i=0}^{\infty} (\text{sgn}(\sin 2^{i+1} \pi x))^\epsilon_i.
\end{align*}
\]

Define the \(n^{th}\) partial Walsh-sum as

\[
W_n f(x) := \sum_{k=0}^{n} < f, w_k > w_k(x).
\]

Now as a direct consequence of the methods developed in this paper, we have:

**Corollary 44.** Let \(\{n_j\}_j\) be a lacunary sequence. Define the (lacunary) Walsh-Carleson operator as

\[
C_{W}^{\{n_j\}} f(x) := \sup_{j \in \mathbb{N}} |W_{n_j} f(x)|.
\]

Then we have that

\[
C_{W}^{\{n_j\}} : L \log L \to L^1.
\]

Moreover, within the class of all lacunary sequences this result is sharp.

\[\text{---}^{42}\text{Notice that in (314) only finitely many } \epsilon_i \text{'s are nonzero.}\]
The fact that (318) holds can be proved by following the same steps as in the Fourier case. The second claim, that is that the \((L \log L, L^1)\) bound is sharp within the class of all lacunary sequences can be easily proved as follows:

Choose \(\{n_j\}_j\) such that \(n_j = 2^j - 1\). Then, from the properties of the Walsh system, we have

\[
W_{n_j}(f)(x) = \langle f(\cdot), \prod_{i=0}^{j-1} (r_0(x) r_0(\cdot) + r_{2^j}(x) r_{2^j}(\cdot)) \rangle,
\]

where here \(r_{2^j}(x) := \text{sgn}(\sin 2^{i+1} \pi x)\) stands for the \(i^{th}\) Rademacher function.

Consequently, we deduce

\[
C_{W^{\{n_j\}}_j} f(x) \approx M_f(x),
\]

and choosing now \(f(x)\) within the sequence \(\{f_n\}_{n \geq 1}\) given by \(f_n(x) := \min\{2^n, \frac{1}{x}\}\) on \([0,1]\) one immediately notices that

\[
\|C_{W^{\{n_j\}}_j} f_n\|_1 \approx n^2 \approx \|f_n\|_{L \log L}.
\]

2) Our main theorem is a first result in the time-frequency literature involving \((L \log L, L^1)\) bounds for a frequency modulated maximal operator of Carleson type that incorporates both the behavior of the Hardy-Littlewood maximal operator and that of the Hilbert transform. This result is possible due to the following key two features:

\[
\begin{align*}
\text{•} & \quad \text{the new concept of the set-resolution of the time-frequency plane (at a fix frequency) which is further based on the concept of the time-frequency regularization of a set;} \\
\text{•} & \quad \text{the } E-\text{mass versus } F-\text{mass dichotomy (see Section 12.2.3) that brings a selection of the trees in two families according to (291), (292) and (303) respectively.}
\end{align*}
\]

The first item is responsible with matching the structural properties of the set \(F\) and of the lacunary Carleson operator. The second item realizes an antithesis/complementarity in terms of the impact of the sets \(E\) and \(F\) on the tiles: if we isolate all the trees at a given frequency then the tops of the trees obey one of the two Carleson packing conditions corresponding to

\[
\begin{align*}
\text{•} & \quad (298) \text{ for the } E-\text{mass component;} \\
\text{•} & \quad (308) \text{ for the } F-\text{mass component.}
\end{align*}
\]

This dichotomy is finally the one that allows the control over the summation in the terms (293) and (294) having as the final output the desired \(L \log L\) bound.

3) Let us briefly comment on Conjecture 2. Recall that in this case we ask wether or not the full Carleson operator \(C\) maps \(L \log L\) into \(L^1\). Obviously in this situation the much increased complexity is given by the
fact that the set of frequencies range over arbitrary sets of integers. Thus isolating structures within this set of frequencies relative to the structural properties of the set $F$ becomes a very challenging task. However, there are genuine hopes about approaching this problem that reside on elaborating on itemization (320): the first item there can be transformed into a modulation invariant concept that we like to call it “the set-resolution of the time-frequency plane”. This seems the right tool to be used since in this general instance no frequency plays a central role. The second item described in (320) is applied to each single fixed frequency and thus can be transferred without modifications to our general setting. One of the main difficulties, relies now on the fact that there are no simple analogues of (328) below (see the Appendix). This is the critical point where we expect that techniques from additive combinatorics will play a major role. We hope to address this topic into a future work.

14. Appendix

Interaction between additive combinatorics and time-frequency analysis via TFR

This section should be regarded as providing some motivation and further heuristics for the key concept introduced Section 4. Our aim is to explain how the newly introduced concept of the time-frequency regularization of a set exploits the additive combinatoric properties of the set relative to lacunary Fourier series.

Indeed, we have the following

**Proposition 45.** Assume we are given the following:

- a natural numbers $N \in \mathbb{N}$;
- a measurable set $F \subseteq \mathbb{T}$ with $|F| \approx 2^{-N}$;
- a lacunary sequence $\{n_j\}_{j \in \mathbb{N}} \subseteq \mathbb{N}$;
- a sequence $\{a_j\}_{j \in \mathbb{N}} \subseteq \mathbb{C}$ such that $|a_j| \approx 1$ for any $j \in \mathbb{N}$.

Then given any $M \in \mathbb{N}$ one has

\[
\int_F \left| \sum_{j=1}^{M} a_j e^{2\pi i n_j x} \right| dx \lesssim |F| \min\{N^{\frac{1}{2}}, M^{\frac{1}{2}}\} M^{\frac{1}{2}}.
\]

Moreover inequality (321) is sharp in terms of parameters $M, N$.

Assuming for the moment the above proposition we have the following immediate consequence:

**Corollary 46.** [Zygmund’s inequality]

Given a lacunary sequence $\{n_j\}_{j \in \mathbb{N}} \subseteq \mathbb{N}$ and a sequence $\{a_j\}_{j \in \mathbb{N}} \in l^2(\mathbb{N})$ one has

\[
\left\| \sum_{j=1}^{\infty} a_j e^{2\pi i n_j \cdot} \right\|_{\text{exp}(L^2)} \lesssim \left\| \{a_j\}_{j} \right\|_{l^2(\mathbb{N})}.
\]
Proof. Inspecting the proof of our Proposition 45 we deduce that the following stronger relation holds:

\[(323) \int_F \left| \sum_{j=1}^{\infty} a_j e^{2\pi i n_j x} \right| \, dx \lesssim |F| (\log |F|)^{1/2} \left( \sum_{j=1}^{\infty} |a_j|^2 \right)^{1/2}.\]

Exploiting now the duality between the Orlicz (Lorentz) spaces \(L(\log L)^{1/2}\) and \(\exp(L^2)\) we conclude that (322) holds. \(\square\)

**Proof of Proposition 45**

We first notice that the case \(M \leq N\) is trivial since in this situation (321) reduces simply to \(L^1 - L^\infty\) Hölder inequality. The fact that the estimate is sharp can be easily verified in the instance \(F\) interval, say \(F = [0, 2^{-N}]\) and \(n_j = 2^j\) for \(j \in \{1, \ldots, M\}\) with \(M \leq N\).

Thus, we remain with the following statement to prove: for any \(M > N\) one has

\[(324) \int_F \left| \sum_{j=1}^{M} a_j e^{2\pi i n_j x} \right| \, dx \lesssim |F| \left( \frac{M}{N} \right)^{1/2}.\]

Apply now the TFR algorithm for the set \(F\) and recall the properties listed in Section 4.4.

Set now the set of frequencies

\[(325) \mathcal{F} := \{n_j\}_{j=1}^{M},\]

where here \(\{n_j\}_{j=1}^{M}\) is as given in our hypothesis.

With these done, we have

\[(326) \quad S_F := \int_F \left| \sum_{j=1}^{M} a_j e^{2\pi i n_j x} \right| \, dx \leq \sum_{k=1}^{k_F} \sum_{l \in \mathbb{Z}_k} \sum_{n} \sum_{\substack{a_j \in \{L, U\} \quad R \in \mathcal{R}_k^{1:n} \quad [I] \quad R \in \mathcal{R}_k^{n:l}(R)}} \int_{F'} \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right|.\]

**Observation 47.** [Relevance of lacunarity] Here we want to record the two key relations where we encode the information that \(\mathcal{F}\) consists of lacunary frequencies:

- using similar reasonings with those from the proofs of Lemmas 24 and 25, one has for each \(R \in \mathcal{R}_k[I]\)

\[(327) \int_F \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right| \lesssim 2^{-k} \int_{I_R} \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right|.\]
• with the previous notations, one has

$$\int_{I_R} \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right|^2 \approx |I_R| \sum_{j \in F[R]} |a_j|^2. \tag{328}$$

Using now (326) and (327) one deduces

$$S_F \lesssim k_F \sum_{k=1}^{k_F} \sum_{I \in I_k} \sum_{n} \sum_{\sum_{s_j \in \{L, U\}}} \int_{I_R} \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right|. \tag{329}$$

The following result is a subcase of the more involved Lemma 33 whose proof was provided in Section 10:

**Claim 48.** Fix $I \in I_k$, $n \in \mathbb{N}$, $s_j \in \{U, L\}$ and take $\mathcal{R}_k^1 \cdots s_n[I]$. With the definitions and notations from the previous section the following holds:

$$\sum_{\mathcal{R} \in \mathcal{R}_k^1 \cdots s_n[I]} \sum_{\mathcal{R} \in \mathcal{R}_k^1 \cdots s_n[I]} \int_{I_R} \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right| \lesssim \left( \sum_{\mathcal{R} \in \mathcal{R}_k^1 \cdots s_n[I]} |I_R| \right) \left( \sum_{j \in F(\mathcal{R}_k^1 \cdots s_n, U[I])} |a_j|^2 \right)^{\frac{1}{2}}. \tag{330}$$

Then we immediately deduce that

$$\sum_{k=1}^{k_F} \sum_{I \in I_k} \sum_{n} \sum_{\sum_{s_j \in \{L, U\}}} \int_{I_R} \left| \sum_{j \in F[R]} a_j e^{2\pi i n_j x} \right| \lesssim \sum_{k=1}^{k_F} 2^{-k} \sum_{I \in I_k} \sum_{n} \sum_{\sum_{s_j \in \{L, U\}}} \left( \sum_{\mathcal{R} \in \mathcal{R}_k^1 \cdots s_n[I]} |I_R| \right) \left( \sum_{j \in F(\mathcal{R}_k^1 \cdots s_n, U[I])} |a_j|^2 \right)^{\frac{1}{2}}. \tag{331}$$

The next result is a particular instance of Lemma 34 whose proof was also provided in Section 10:

**Claim 49.** Let $I \in I_k$ and $\mathcal{R}_k^1 \cdots s_n[I]$ as before. Then, the following holds:

$$\sum_{s_1' = s_1 \cdots s_n' = s_n} \sum_{\mathcal{R} \in \mathcal{R}_k^{s_1' \cdots s_n'[I]}} \sum_{\mathcal{R} \in \mathcal{R}_k^{s_1' \cdots s_n'[I]}} |I_R| \left( \sum_{j \in F(\mathcal{R}_k^{s_1' \cdots s_n'[U[I]})} |a_j|^2 \right)^{\frac{1}{2}} \lesssim \left( \sum_{\mathcal{R} \in \mathcal{R}_k^{s_1' \cdots s_n'[I]}} |I_R| \right) \left( \sum_{j \in F(\mathcal{R}_k^{s_1' \cdots s_n'[U[I]})} |a_j|^2 \right)^{\frac{1}{2}}. \tag{332}$$
Implementing (332) in (331) and putting this together with (329), we deduce that

\[
S_F \lesssim \sum_{k=1}^{k_F} 2^{-k} \sum_{I \in \mathcal{I}_k} (\sum_{R \in \mathcal{R}_k^{U_1}[I]} |I_R|) \left( \sum_{j \in F(\mathcal{R}_k^{U_1}[I])} |a_j|^2 \right)^{\frac{1}{2}},
\]

where here we recall the notation (212).

We will use now the following two observations

- for any $1 \leq k \leq k_F$ any \textsuperscript{43} $n \geq -1$ and any $I \in \mathcal{I}_k$, based on (59) one has

\[
\sum_{R \in \mathcal{R}_k^{U_1}[I]} |I_R| \leq \frac{1}{4} \sum_{R \in \mathcal{R}_k^{U_1}[I]} |I_R|.
\]

- for any $1 \leq k \leq k_F$ any $n \neq n'$ and any $I \in \mathcal{I}_k$ one has

\[
F(\mathcal{R}_k^{nU_1}[I]) \cap F(\mathcal{R}_k^{n'U_1}[I]) = \emptyset.
\]

From the above itemization we deduce that it is enough to estimate $S^0_F$ for $n = -1$, where

\[
S^0_F := \sum_{k=1}^{k_F} 2^{-k} \sum_{I \in \mathcal{I}_k} (\sum_{R \in \mathcal{R}_k^{U_1}[I]} |I_R|) \left( \sum_{j \in F(\mathcal{R}_k^{U_1}[I])} |a_j|^2 \right)^{\frac{1}{2}},
\]

since from (334) we will get a geometric decay in $n$.

Now, using Cauchy-Schwarz, we deduce that

\[
S^0_F \lesssim \sum_{k=1}^{k_F} \sum_{I \in \mathcal{I}_k} 2^{-k} |I| \left( \sum_{j \in F(\mathcal{R}_k^{U_1}[I])} |a_j|^2 \right)^{\frac{1}{2}}
\]

\[
\lesssim (\sum_{k=1}^{k_F} \sum_{I \in \mathcal{I}_k} 2^{-k} |I|)^{\frac{1}{2}} \left( \sum_{k=1}^{k_F} \sum_{I \in \mathcal{I}_k} |I \cap F| \sum_{j \in F(\mathcal{R}_k^{U_1}[I])} |a_j|^2 \right)^{\frac{1}{2}}
\]

\[
\lesssim (\sum_{k=1}^{k_F} \sum_{I \in \mathcal{I}_k} |I \cap F|)^{\frac{1}{2}} \left( \int_{F} \sum_{k=1}^{k_F} \sum_{I \in \mathcal{I}_k} \sum_{j \in F(\mathcal{R}_k^{U_1}[I])} |a_j|^2 \right)^{\frac{1}{2}}
\]

\[
\lesssim \frac{1}{2} |F| M^{\frac{1}{2}}.
\]

This ends the proof of inequality (321).

It remains now to show that (324) is sharp.

We start with a general comment and continue with few remarks: in understanding the best possible upper bound for our inequality (321) the fundamental role is played by the interaction between the structure of the set $F$ and the corresponding structure of the frequencies $\{n_j\}_j$. Indeed, let us notice the following:

\textsuperscript{43}Recall the convention made in (211).
if $F$ is an interval, then for any lacunary sequence $\{n_j\}_j$ relation (321) can be upgraded to:

$$\int_F |\sum_{j=1}^M a_j e^{2\pi i n_j x}| dx \lesssim |F| (N + M^{\frac{1}{2}}).$$

- same holds if $F$ consists of a union of equidistant same-size intervals - this is what we call the arithmetic progression case.
- in symmetry with this, if we fix $\{n_j = 2^j\}_j$ then, depending on the structure of $F$, the best upper bound in (321) is
  - $|F| (N + M^{\frac{1}{2}})$ if $F$ an interval or an arithmetic progression with low entropy (the number of equidistant same-size intervals that form $F$ is much smaller than $\min\{M,N\}$)
  - $|F| M^{\frac{1}{2}}$ - which is the best possible bound - if $F$ is an arithmetic progression with high entropy (larger than $2^N$)
  - $|F| \min\{N^{\frac{1}{2}}, M^{\frac{1}{2}}\} M^{\frac{1}{2}}$ if $F$ has the structure of a suitable Cantor set as described a bit later below - see (339) (this corresponds to the generalized arithmetic progression case).

From the above remarks it becomes transparent that in order to find an example for which the RHS of (324) is sharp one needs to search for a set $F$ with a Cantor set structure, that is for which $F$ has the structure of a generalized arithmetic progression. Recalling that we only discuss the case $M > N$, wlog we can assume that $M = 2^s N$ for some large $s \in \mathbb{N}$.

We stop here briefly to introduce some more notations: assuming that $J \subseteq [0,1]$ dyadic interval we set $J_s(J) := \{I \subseteq J | I \text{ dyadic}, |I| = |J| 2^{-s}\}$. Also, if $A = \bigcup A_j$ is a collection of dyadic intervals then we set $A^t := \bigcup A^t_j$ where $A^t_j$ is the right child of $A$.

With this, the construction of $F$ is as follows:

(339)

- Stage 1 - we let $F_0 := J_s([0,1])$;
- Stage 2 - we define $F_1 := \bigcup_{I \in F^t_0} J_s(I)$;
- we continue inductively to construct $F_2, F_3$ and so on until we reach Stage $N$ - define $F_N := \bigcup_{I \in F^t_{N-1}} J_s(I)$;
- finally, we simply set $F := F_N$.

Choosing now $\{n_j = 2^j\}_j$ one can simply check now that relations (336) - (337) become approximate equalities\(^{44}\).

This ends our proof.

\(^{44}\)Up to harmless constants the bound from the above is of the same nature with the bound from below.
References

[1] N. Yu. Antonov. Convergence of Fourier series. In Proceedings of the XX Workshop on Function Theory (Moscow, 1995), volume 2, pages 187–196, 1996.
[2] N. Yu. Antonov. Almost everywhere divergent subsequences of Fourier sums of functions in $\phi(L) \cap H^p_1$. Mat. Zametki, 85 (4): 502–515, 2009.
[3] Du-Bois-Reymond, P. Untersuchungen über die Convergenz und Divergenz der Fouri-erschen Darstellungsformen. Abhand. Akad. München, (12): 1–103, 1876.
[4] J. Arias-de Reyna. Pointwise convergence of Fourier series. J. London Math. Soc. (2), 65(1):139–153, 2002.
[5] Lennart Carleson. On convergence and growth of partial sums of Fourier series. Acta Math., 116:135–157, 1966.
[6] Lennart Carleson. Interpolations by bounded analytic functions and the corona problem. Ann. of Math., 76 (3): 547559, 1962.
[7] Yung-Ming Chen. An almost everywhere divergent Fourier series of the class $L(\log^+ \log^+ L)^{1-\epsilon}$. J. London Math. Soc., 44:643-654, 1969.
[8] M. J. Carro, M. Mastylo, and L. Rodriguez-Piazza. Almost everywhere convergent Fourier series. J. Fourier Anal. Appl. (18), 2:266-286, 2012.
[9] Francesco Di Plinio. Lacunary Fourier and Walsh-Fourier series near $L^1$. Collect. Math. (65), 2: 219–232, 2014.
[10] Yen Do and Michael Lacey. On the Convergence of Lacunary Walsh-Fourier Series. Bull. London Math. Soc. (44), 2:241-254, 2012.
[11] Henri Lebesgue. Sur les intégrales singulières. Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. (3), 1: 25–117, 1909.
[12] Joseph Fourier. Théorie analytique de la chaleur. Reprint of the 1822 original, Editions Jacques Gabay, Paris, 1988.
[13] Charles Fefferman. Pointwise convergence of Fourier series. Ann. of Math. (2), 98:551–571, 1973.
[14] Richard P. Gosselin. On the divergence of Fourier series. Proc. Amer. Math. Soc., 9:278–282, 1958.
[15] Richard A. Hunt. On the convergence of Fourier series. In Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), pages 235–255. Southern Illinois Univ. Press, Carbondale, Ill., 1968.
[16] A. N. Kolmogoroff. Une serie de Fourier-Lebesgue divergente presque partout. Number 4, pages 324–328. 1923.
[17] S. V. Konyagin. Divergence everywhere of subsequences of partial sums of trigonometric Fourier series. Proc. Steklov Inst. Math., (Function Theory, suppl. 2):S167–S175, 2005.
[18] Sergey V. Konyagin. Almost everywhere convergence and divergence of Fourier series. In International Congress of Mathematicians. Vol. II, pages 1393–1403. Eur. Math. Soc., Zürich, 2006.
[19] Sergey V. Konyagin. On divergence of trigonomometric Fourier series everywhere. C. R. Acad. Sci. Paris, Ser. I Math., (329): 693–697, 1999.
[20] Sergey V. Konyagin. On the divergence everywhere of trigonometric Fourier series. Sb. Math., (191): 97–120, 2000.
[21] T.W. Körner. Everywhere divergent Fourier series. Colloq. Math., (45): 103–118, 1981.
[22] N. N. Luzin. The integral and the trigonometric series Dissertation, 1915, in Collected Works of N. N. Luzin, Moscow, 1953. (Russian)
[23] Michael Lacey and Christoph Thiele. A proof of boundedness of the Carleson operator. Math. Res. Lett., 7(4):361–370, 2000.
[24] Victor Lie. On the boundedness of the Carleson operator near $L^1$. Rev. Mat. Iberoam., 29(4):1239–1262, 2013.
[25] Victor Lie. On the pointwise convergence of the sequence of partial Fourier sums along lacunary subsequences. *J. Funct. Anal.*, 263 (11): 3333–3740, 2012.

[26] Victor Lie. The Polynomial Carleson operator. Submitted, https://arxiv.org/abs/1105.4504, 2011, (revised 2017).

[27] Victor Lie. The (weak-$L^2$) boundedness of the quadratic Carleson operator. *Geom. Funct. Anal.*, 2 (19): 457–497, 2009.

[28] Victor Lie. The pointwise convergence of the Fourier Series (I) - On a conjecture of Konyagin. *J. Eur. Math. Soc.*, 19 (6): 1655–1728, 2017.

[29] Per Sjölin. An inequality of Paley and convergence a.e. of Walsh-Fourier series. *Ark. Mat.*, 7:551–570 (1969), 1969.

[30] Fernando Soria. Note on differentiation of integrals and the Halo conjecture. *Studia Math.*, 81(1):29–36, 1985.

[31] Fernando Soria. On an extrapolation theorem of Carleson-Sjölin with applications to a.e. convergence of Fourier series. *Studia Math.*, 94(3):235–244, 1989.

[32] V. I. Prohorenko. Divergent Fourier series. (Russian) *Mat. Sb. (N.S.),* 75(117):185-198, 1968.

[33] Terence Tao and James Wright Endpoint multiplier theorems of Marcinkiewicz type. *Rev. Mat. Iberoam.*, 17(3):521–528, 2001.

[34] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002. With a foreword by Robert A. Fefferman.

DEPARTMENT OF MATHEMATICS, PURDUE, IN 46907 USA

E-mail address: vlie@math.purdue.edu

INSTITUTE OF MATHEMATICS OF THE ROMANIAN ACADEMY, BUCHAREST, RO 70700, P.O. BOX 1-764, ROMANIA.