JEMS

Victor Lie

Pointwise convergence of Fourier series (I). On a conjecture of Konyagin

Abstract. We provide a near-complete classification of the Lorentz spaces Λ_φ for which the sequence $\{S_n\}_{n\in\mathbb{N}}$ of partial Fourier sums is almost everywhere convergent along lacunary subsequences. Moreover, under mild assumptions on the fundamental function φ, we identify $\Lambda_\varphi := L\log\log L\log\log\log\log L$ as the *largest* Lorentz space on which the lacunary Carleson operator is bounded as a map to $L^{1,\infty}$. As a consequence,

- we disprove a conjecture stated by Konyagin in his 2006 ICM address;
- we provide a negative answer to an open question related to the Halo conjecture.

Our proof relies on a newly introduced concept of a "Cantor multi-tower embedding," a special geometric configuration of tiles that can arise within the time-frequency tile decomposition of the Carleson operator. This geometric structure plays an important role in the behavior of Fourier series near L^1, being responsible for the unboundedness of the weak-L^1 norm of a "grand maximal counting function" associated with the mass levels.

Keywords. Time-frequency analysis, Carleson's Theorem, lacunary subsequences, pointwise convergence

1. Introduction

1.1. Historical background

In this paper we address the problem of pointwise convergence of Fourier series along lacunary subsequences. Regarded in a broader context, the problem of pointwise convergence of Fourier series has a rich history, tracing back to the cornerstone set by Fourier in his study on heat propagation [14]. Since then, there has been a series of major advancements, of which we only mention those closest to our topic: in 1873 du Bois-Reymond [12] offered an example of a continuous function whose Fourier series diverges on the set of rational points. This surprising result stimulated the search for new grounds upon which one could reformulate the question of pointwise convergence for larger classes of functions by focusing only on the "almost everywhere" behavior of the series, and thus

V. Lie: Department of Mathematics, Purdue University, West Lafayette, IN 46907 USA, and Institute of Mathematics of the Romanian Academy,
P.O. Box 1-764, 014700 Bucharest, Romania; e-mail: vlie@math.purdue.edu

allowing for pathologies on "negligible" sets (of "measure 0"). The appropriate setting was developed by Lebesgue in his theory of integration [29]. Within this new framework, Luzin [34] conjectured that if a function f is square integrable then its Fourier series converges to f Lebesgue-almost everywhere. In 1922, Kolmogorov [19] constructed an example of an L^1-integrable function whose Fourier series diverges almost everywhere, suggesting that Luzin's conjecture may be false. However, after several decades of misconceptions, in 1966, the breakthrough work of L. Carleson [5] confirmed the conjecture. Shortly thereafter, Hunt [17] extended the techniques of [5], showing that Carleson's result holds for any $f \in L^p(\mathbb{T})$ as long as $1 < p < \infty$.

At this point, we should mention that though not providing a new result, the second proof of the almost everywhere convergence of the Fourier series for L^2 functions offered by C. Fefferman [13] marked a fundamental advancement in understanding the topic described here. A third proof of Luzin's conjecture was given in 2000 by Lacey and Thiele [28] using the tools they developed to address the boundedness of the bilinear Hilbert transform [26], [27].

1.2. Formulation of the main problem(s); context

We start this section by formulating (at first in a looser language) one of the main open questions in the area of Fourier series:

Main Question. *What can be said about the (almost everywhere) pointwise convergence of Fourier series between the two known cases for the Lebesgue spaces $L^p(\mathbb{T})$:*

- *$p = 1$, divergence of Fourier series (Kolmogorov),*
- *$p > 1$, convergence of Fourier series (Carleson–Hunt)?*

In order to make this Main Question precise, let us first introduce the following:

Definition 1.1. Let Y be a r.i. (quasi-)Banach space.[1] We say that Y is a $\mathcal{C}$-*space* if there exists $C_0 = C_0(Y) > 0$ such that the *Carleson operator* $C : C^\infty(\mathbb{T}) \to L^\infty(\mathbb{T})$ defined[2] by

$$Cf(x) := \sup_{N \in \mathbb{N}} \left| \int_{\mathbb{T}} e^{2\pi i N(x-y)} \cot(\pi(x-y)) f(y) \, dy \right|, \tag{1.1}$$

obeys the relation[3]

$$\|Cf\|_{1,\infty} \le C_0 \|f\|_Y \quad \forall f \in Y. \tag{1.2}$$

With this definition, the Main Question can be reformulated as follows:

---

[1] For basic definitions and concepts of the theory of rearrangement invariant Banach spaces, including Lorentz spaces, see the Appendix.

[2] Depending on the context, we identify the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with either $[-1/2, 1/2]$ or $[0, 1]$.

[3] Recall that the weak-L^1 quasinorm is given by $\|f\|_{1,\infty} := \sup_{\lambda > 0} \lambda |\{x \mid |f(x)| > \lambda\}|$.

Open Problem A. (1) *Give a satisfactory description of the Lorentz spaces $Y \subseteq L^1(\mathbb{T})$ that are also $\mathcal{C}$-spaces. Describe the maximal Lorentz $\mathcal{C}$-space Y_0, if such exists.*

(2) *Let Y be a r.i. (quasi-)Banach space. Provide necessary and sufficient conditions for Y to be a $\mathcal{C}$-space.*

The best known results relating to the above problem are:

- On the *negative* side: Konyagin [23], [24] proved that if $\phi(u) = o\left(u\sqrt{\frac{\log u}{\log\log u}}\right)$ as $u \to \infty$ then the space $\phi(L) = \Lambda_{\bar{\phi}}$ is not a $\mathcal{C}$-space, where $\bar{\phi}(t) := \int_0^t s\phi(1/s)\,ds$. Thus, there exists $f \in \phi(L)$ with

$$\overline{\lim_{m \to \infty}} \, S_m f(x) = \infty \quad \text{for all } x \in \mathbb{T}, \tag{1.3}$$

  where $S_m f$ stands for the mth partial Fourier sum associated with f.
- On the *positive* side: Antonov [1] showed that (1.2) holds for the Lorentz space $Y = L \log L \log\log\log L$; later Arias-de-Reyna [3] showed that Y can be enlarged to a rearrangement invariant quasi-Banach space, named QA, and strictly containing[4] $L \log L \log\log\log L$.

We add here that the first results along these lines were obtained on the negative side by Chen [8], Prohorenko [35] and Körner [25], and on the positive side by Sjölin [36] and Soria [38], [39].

It is worth noting that all the progress mentioned above on the positive side involved tools from extrapolation theory. Recently, using methods that rely entirely on time-frequency arguments, the author was able to reprove all the positive results by a unified approach [30].

Now recall that both the maximal Hardy–Littlewood operator and the (maximal) Hilbert transform are bounded from $L \log L$ to L^1. At a heuristic level, the Carleson operator may be thought of as a superposition of the maximal Hardy–Littlewood operator and modulated copies of the (maximal) Hilbert transform. Thus, one is naturally led to the following

Conjecture 1. *The Lorentz space $Y = L \log L$ is a $\mathcal{C}$-space.*

As a simplified model for better understanding the difficulties of Open Problem A (and of the conjecture above) one can formulate its *lacunary* version. Recall that a sequence $\{n_j\}_{j \in \mathbb{N}} \subset \mathbb{N}$ is called *lacunary* if

$$\lim_{j \to \infty} \frac{n_{j+1}}{n_j} > 1. \tag{1.4}$$

---

[4] For an interesting study of the properties of QA and relationship(s) between Antonov and Arias-de-Reyna spaces, see [7]. In the same paper, the authors prove that under suitable conditions on the function φ the space $\Lambda_\varphi = L \log L \log\log\log L$ is the largest Lorentz space contained in QA.

Now, by analogy with the previous situation, we first introduce the following

Definition 1.2. 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 $\mathcal{C}_L^{\{n_j\}_j}$-*space* if there exists $C_1 = C_1(Z, \{n_j\}_j) > 0$ such that the $\{n_j\}_{j\in\mathbb{N}}$-*lacunary Carleson operator* defined by

$$C_{\mathrm{lac}}^{\{n_j\}_j} : C^\infty(\mathbb{T}) \to L^\infty(\mathbb{T})$$

with

$$C_{\mathrm{lac}}^{\{n_j\}_j} f(x) := \sup_{j\in\mathbb{N}} \left| \int_{\mathbb{T}} e^{2\pi i n_j(x-y)} \cot(\pi(x-y)) f(y)\,dy \right|, \qquad (1.5)$$

obeys

$$\|C_{\mathrm{lac}}^{\{n_j\}_j} f\|_{1,\infty} \le C_1 \|f\|_Z \quad \forall f \in Z. \qquad (1.6)$$

(ii) We say that Z is a $\mathcal{C}_L$-*space* if it is a $\mathcal{C}_L^{\{n_j\}_j}$-space for *any* lacunary sequence $\{n_j\}_{j\in\mathbb{N}}$. Throughout the paper, if Z is a $\mathcal{C}_L$-space, we will (often) express this as[5]

$$\|C_{\mathrm{lac}} f\|_{1,\infty} \lesssim \|f\|_Z \quad \forall f \in Z, \qquad (1.7)$$

where C_{lac} stands for "the generic" lacunary Carleson operator.[6]

We can now formulate the analogue of Open Problem A:

Open Problem B. (1) *Give a satisfactory description of the Lorentz spaces Z that are also $\mathcal{C}_L$-spaces. Describe the maximal Lorentz $\mathcal{C}_L$-space Z_0, if such exists.*[7]

(2) *Let Z be a r.i. (quasi-)Banach space. Provide necessary and sufficient conditions for Z to be a $\mathcal{C}_L$-space.*

In a more general context, initial progress on this problem was made by Zygmund [42] who showed that $Z = L \log L$ is a Lorentz $\mathcal{C}_L$-space. On the negative side, Konyagin [21] proved that if $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing function with $\phi(0) = 0$ and $\phi(u) = o(u \log \log u)$ as $u \to \infty$ then $\phi(L) = \Lambda_{\bar{\phi}}$ is not a $\mathcal{C}_L$-space. This last result was reproved later in a slightly modified context by Antonov [2].[8]

In his invited talk at the 2006 International Congress of Mathematicians in Madrid, Konyagin stated the following

Conjecture 2 (Konyagin, [22]). *The Lorentz space $L \log \log L$ is a $\mathcal{C}_L$-space.*

---

[5] Given $A, B > 0$, we write $A \lesssim B$ and $B \gtrsim A$ to mean that there exists $C > 0$ such that $A \le CB$ and $B \le CA$ respectively.

[6] In (1.7), 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 $f \in Z$.

[7] One can formulate a more specific question by prescribing a lacunary sequence $\{n_j\}_{j\in\mathbb{N}}$ and asking for a satisfactory description of the Lorentz spaces Z that are also $\mathcal{C}_L^{\{n_j\}_j}$-spaces.

[8] As an immediate application of the concepts developed in this paper, one can obtain a simplified proof of the results in [21] and [2]—see Section 13.

At this point, we should say that one can phrase an analogue of the above conjecture for Walsh–Fourier series (is it true that (1.7) holds for $Z = L \log \log L$ with C_{lac} replaced by the lacunary Walsh–Carleson operator?). In this latter context, Do and Lacey [11] were the first to make progress by showing that if one takes $Z = L \log \log L \log \log \log L$ then (1.7) holds for the Walsh form of the lacunary Carleson operator. Their proof relies on a projection argument which is not transferable to the Fourier series case.[9]

In [31], we were able to prove the following

Theorem 1.3 ([31]). *Let $\mathcal{W}$ be the quasi-Banach space defined by*[10]

$$\mathcal{W} := \{f : \mathbb{T} \to C \mid f \text{ measurable}, \ \|f\|_{\mathcal{W}} < \infty\},$$

where

$$\|f\|_{\mathcal{W}} := \inf \left\{ \sum_{j=1}^{\infty} (1 + \log j) \|f_j\|_1 \log \log \frac{4\|f_j\|_{\infty}}{\|f_j\|_1} \ \middle| \ \begin{array}{c} f = \sum_{j=1}^{\infty} f_j, \\ \sum_{j=1}^{\infty} |f_j| < \infty \text{ a.e.,} \\ f_j \in L^{\infty}(\mathbb{T}) \end{array} \right\}.$$

Then

$$\|C_{\text{lac}} f\|_{1,\infty} \lesssim \|f\|_{\mathcal{W}}. \tag{1.8}$$

Thus $Z = \mathcal{W}$ is a $\mathcal{C}_L$-space. Moreover, it contains $L \log \log L \log \log \log L$.

Taking Theorem 1.3 above as a black-box, Di Plinio [9] proved that the space $L \log \log L \log \log \log \log L$ is a $\mathcal{C}_L$-space. Indeed, relying entirely[11] on standard extrapolation techniques, he showed that

$$Z' := L \log \log L \log \log \log \log L \subseteq \mathcal{W},$$

which in view of (1.8) immediately implies that Z' is a $\mathcal{C}_L$-space.

1.3. Main results

In this section we present the main results of the paper. They are based on a new concept of "Cantor multi-tower embedding" (**CME**) whose nature will be detailed in the next subsection. With these being said, we state the following

Main Theorem 1. *There exists a lacunary sequence $\{n_j\}_j$ and a sequence $\{f_k\}_{k \in \mathbb{N}}$ of (positive) functions such that:*

- *each f_k is in $L^{\infty}(\mathbb{T})$ with*

$$\|f_k\|_{L \log \log L} \approx 1; \tag{1.9}$$

---

[9] In the same paper, using previous results from extrapolation theory, the authors proved that the Walsh form of (1.7) holds for a slightly larger quasi-Banach space $Z = Q_D$. This last space turns out to be isomorphic to the space $\mathcal{W}$ introduced in [31], though we have designed $\mathcal{W}$ by different means independent of extrapolation theory.

[10] Throughout the paper, $\log k$ stands for $\log_2 k$.

[11] The difficult combinatorial and time-frequency techniques are nedeed precisely in order to show that $\mathcal{W}$ is a $\mathcal{C}_L$-space. The present paper shows that Theorem 1.3 cannot be essentially improved.

- *we have*

$$\lim_{k \to \infty} \|f_k\|_{L \log \log L \log \log \log \log L} = \infty; \tag{1.10}$$

- *there exists an absolute constant $C > 0$ such that for any $k \in \mathbb{N}$,*

$$\|C_{\text{lac}}^{\{n_j\}_j} f_k\|_{L^{1,\infty}} \geq C \|f_k\|_{L \log \log L \log \log \log \log L}. \tag{1.11}$$

The next result states that the conclusion of the above theorem remains true for *any* lacunary sequence $\{n_j\}_j$. More precisely one has

Theorem 1.4. *Given any lacunary sequence $\{n_j\}_j$, there exists a sequence $\{f_k\}_{k \in \mathbb{N}}$ of* (*positive*) *functions such that* (1.9)–(1.11) *hold.*

Main Theorem 2.

(1) *Define $\varphi_0 : [0, 1] \to \mathbb{R}_+$ by*

$$\varphi_0(s) := s \log \log \frac{17}{s} \log \log \log \log \frac{17}{s}.$$

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

  (i) *If $\underline{\lim}_{s \to 0+} \varphi(s)/\varphi_0(s) > 0$ then Λ_φ is a $\mathcal{C}_L$-space.*
  (ii) *If $\overline{\lim}_{s \to 0+} \varphi(s)/\varphi_0(s) = 0$ then Λ_φ is not a $\mathcal{C}_L$-space.*
  (iii) *If $\underline{\lim}_{s \to 0+} \varphi(s)/\varphi_0(s) = 0 < \overline{\lim}_{s \to 0+} \varphi(s)/\varphi_0(s)$ then both scenarios are possible.*

(2) *Let $\varphi : [0, 1] \to \mathbb{R}_+$ be a quasi-concave function. Consider the following statements:*

  (A) *The function φ obeys*

$$\int_0^1 \frac{-s\varphi_0''(s)}{\varphi(s)} \, ds \approx \int_0^1 \frac{\varphi_0(s)}{\varphi(s)} \frac{ds}{s \log \frac{4}{s} \log \log \frac{4}{s}} < \infty. \tag{1.12}$$

  (B) *Any r.i. Banach space X with fundamental function $\varphi_X = \varphi$ is a $\mathcal{C}_L$-space.*

  Then:

  (i) (A) *implies* (B);
  (ii) (B) *implies $\underline{\lim}_{s \to 0+} \varphi_0(s)/\varphi(s) = 0$;*
  (iii) *if there exists $\epsilon > 0$ such that $s \mapsto \varphi_0(s)/\varphi(s)$ is increasing on $(0, \epsilon)$ then* (A) *is equivalent to* (B).

Remark. In fact, one can derive Main Theorem 2 from Main Theorem 1 and Theorem 1.5 below. However, we prefer to give special attention to Main Theorems 1 and 2 since these statements include the more conceptual nature of our results.

Theorem 1.5. *Let $k \in \mathbb{N}$ and let $\{r_j\}_{1 \leq j \leq k}$ be positive real numbers. For $1 \leq j \leq k$ define $y_j = 2^{-\log j \, 2^{2^j}}$. Then, for any $x_j \in (y_{j+1}, y_j]$, one can construct measurable sets $F_j \subset \mathbb{T}$ such that:*

- *the sets $\{F_j\}_{j \leq k}$ are pairwise disjoint;*
- *$|F_j| = x_j$;*

- *there exist absolute constants $C_1 \geq C_2 > 0$ such that the function*

$$f_k := \sum_{j=1}^{k} r_j \chi_{F_j} \tag{1.13}$$

  obeys the estimate

$$C_2 \|f_k\|_{\mathcal{W}} \leq \|C_{\text{lac}}^{\{2^j\}_j} f_k\|_{1,\infty} \leq C_1 \|f_k\|_{\mathcal{W}}. \tag{1.14}$$

Consequences of Main Theorems 1 and 2. From point (1)(iii) of Main Theorem 2 (see also the corresponding proof) one can deduce that there exist (infinitely many) Lorentz $\mathcal{C}_L$-spaces Λ_φ such that

$$L \log \log L \log \log \log \log L \subsetneq \Lambda_\varphi \subsetneq \mathcal{W}.$$

While these Λ_φ spaces are non-canonical, their fundamental functions φ still share at infinitely many space locations the same structure as that of φ_0. Thus, under suitable, mild conditions on φ, Λ_{φ_0} becomes the *largest* Lorentz $\mathcal{C}_L$-space, this being simply the content of the following:

Corollary 1.6 (Maximal characterization). *Let $\varphi : [0, 1] \to \mathbb{R}_+$ be a non-decreasing concave function with $\varphi(0) = 0$. Assume that*

$$\lim_{s \to 0+} \frac{\varphi(s)}{\varphi_0(s)} \in [0, \infty] \text{ exists.} \tag{1.15}$$

Then the largest Lorentz $\mathcal{C}_L$-space Λ_φ for which φ obeys (1.15) is

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

Taking in (1.15) the function $\varphi(s) = s \log \log \frac{4}{s}$, one further deduces

Corollary 1.7 (Resolution of Konyagin's conjecture). *Conjecture 2 is false.*

At this point, we record this observation, surprising at first glance:

Observation 1.8. Define[12]

- the $\{n_j\}_j$-lacunary Lacey–Thiele *discretized Carleson periodic model* by

$$\tilde{C}^{\{n_j\}_j} f(x) := \sup_{j \in \mathbb{N}} \left| \sum_{P \in \mathbb{P}^{0,0,0,+}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(n_j) \right|; \tag{1.16}$$

- the $\{n_j\}_j$-lacunary *discretized Walsh–Carleson operator* by

$$\tilde{C}_W^{\{n_j\}_j} f(x) = \sup_{j \in \mathbb{N}} \left| \sum_{R \in \mathcal{R}} \langle f, w_{R_l} \rangle w_{R_l}(x) \chi_{\omega_{R_u}}(n_j) \right|. \tag{1.17}$$

---

[12] For more on the definitions, notation and properties of the discrete Carleson and Walsh models see Section 12.

- the $\{n_j\}_j$-lacunary *Walsh–Carleson operator* by

$$C_W^{\{n_j\}_j} f(x) := \sup_{j \in \mathbb{N}} \left| \sum_{k=0}^{n_j} \langle f, w_k \rangle w_k(x) \right|, \tag{1.18}$$

where $w_N(x)$ stands for the Nth Walsh mode regarded as a periodic function on $\mathbb{R}$;
- the $\{n_j\}_j$-lacunary *averaged Walsh–Carleson model* by

$$C_{AW}^{\{n_j\}_j} f(x) := \sup_{j \in \mathbb{N}} \left| \int_{\mathbb{T}} w_{n_j}(x) w_{n_j}(-y) \cot(\pi(x-y)) f(y) \, dy \right|. \tag{1.19}$$

(It is worth mentioning that Thiele [41] proved that, unlike the Fourier case, there is no distinction between the discretized and the standard (non-discretized) Walsh–Carleson operator, that is, $C_W^{\{n_j\}_j} f = \tilde{C}_W^{\{n_j\}_j} f$.)

Now the following are true:

- Theorem 1.4 *holds* for the operator $C_{AW}^{\{n_j\}_j}$ (and obviously for $C^{\{n_j\}_j}$);
- Theorem 1.4 *does not hold* for the operators $\tilde{C}^{\{n_j\}_j}$ and $C_W^{\{n_j\}_j}$.

All these facts will be discussed in great detail in Section 12. Notice that this is the first time when we are witnessing a sharp distinction between the behavior of the Carleson operator and that of the corresponding Lacey–Thiele discretized Carleson model. This also provides a first instance when Fefferman's type discretization—*which leaves the Carleson operator unchanged*—is a necessity and not a choice.

Observation 1.9. The next corollary answers an open question related to the so called *Halo conjecture*[13] (see [16]), regarding whether or not, given a sublinear, *translation invariant* operator T, the following are equivalent:

- T is of restricted weak type (Λ_φ, L^1);
- T is of weak type (Λ_φ, L^1).

Here Λ_φ is some generic Lorentz space.

Corollary 1.10 (Restricted weak type does not imply weak type). *The $\{2^j\}_{j \in \mathbb{N}}$-lacunary Carleson operator obeys the following:*

- $C_{\text{lac}}^{\{2^j\}_j}$ *is a sublinear, translation invariant operator.*
- (*Theorem* 1.3, [31]) $C_{\text{lac}}^{\{2^j\}_j}$ *is of restricted weak type* $(L \log \log L, L^1)$; *thus there exists an absolute constant $C > 0$ such that*

$$\|C_{\text{lac}}^{\{2^j\}_j} \chi_E\|_{1,\infty} \le C|E| \log \log \frac{4}{|E|} \tag{1.20}$$

for any measurable set $E \subseteq \mathbb{T}$.
- (*Main Theorem* 1) $C_{\text{lac}}^{\{2^j\}_j}$ *is not of weak type* $(L \log \log L, L^1)$.

---

[13] For more details on the connections between the Halo conjecture and pointwise convergence of Fourier series the interested reader is referred to [37], [38] and [16].

The next result explains why extrapolation techniques are not suitable for attempting to prove *sharp* bounds near L^1 for the (lacunary) Carleson operator (see also Observation 1.12).

Indeed, if one attempts to regard $L^{1,\infty}$ as a limiting space of the scale $\{L^{p,\infty}\}_{p>1}$, one obtains the following:

Corollary 1.11 (Limitations of extrapolation theory). *The $\{2^j\}_{j\in\mathbb{N}}$-lacunary Carleson operator obeys:*

- *There exists $c > 0$ such that for any $1 < p \leq 2$,*

$$\|C_{\mathrm{lac}}^{\{2^j\}_j} f\|_{p,\infty} \leq c \log\left(2 + \frac{1}{p-1}\right)\|f\|_p. \tag{1.21}$$

- *There exists no $C > 0$ such that*

$$\|C_{\mathrm{lac}}^{\{2^j\}_j} f\|_{1,\infty} \leq C\|f\|_{L \log \log L} \quad \forall f \in L \log \log L. \tag{1.22}$$

The fact that (1.21) holds can be easily derived from [31, proof of Theorem 1] as noticed in [10].

Now standard interpolation/extrapolation[14] results show that if (1.22) were true then (1.21) would immediately follow. However in the light of our Main Theorems 1 and 2 this implication is false, that is, (1.21) does not imply (1.22).

Observation 1.12. As a consequence of the last two corollaries we have the following conclusion: No general *equivalence*[15] can be established between weak-L^1 type bounds and either the corresponding restricted weak-type L^1 bounds or weak-L^p bounds ($p > 1$). Moreover, extrapolation theory by itself is not suitable to provide sharp answers to endpoint questions on pointwise convergence of Fourier series near L^1. To get such answers, one needs to take advantage of the special structure of the Carleson operator and hence to exploit time-frequency analysis methods.

Finally, in relation to some previous work of the author, we have:

Corollary 1.13 (Lack of uniform control for Calderón–Zygmund tile partition). *The question raised in [30] has a negative answer. More precisely, with the notation therein, there is no absolute constant $C > 0$ such that, for $\alpha \in \mathbb{N}$,*

$$\|T^{\mathbb{P}^\alpha} f\|_1 \leq C\|f\|_1 \quad \forall f \in L^1(\mathbb{T}).$$

Moreover, there exists $f \in L^1(\mathbb{T})$ such that if one partitions $\mathbb{P}^\alpha = \bigcup_{n=1}^N \mathbb{P}_n^\alpha$ with each $\mathbb{P}_n^\alpha$ having constant mass (i.e. $A(P) \approx 2^{-n}$ for any $P \in \mathbb{P}_n^\alpha$) then

$$\|T^{\mathbb{P}^\alpha} f\|_{1,\infty} \gtrsim N\|f\|_1.$$

---

[14] See [10, Section 1] for details.

[15] For large classes of operators that include the family of (maximal) operators associated with partial Fourier sums.

1.4. The fundamental idea

In this subsection we will describe, at a philosophical level, a key aspect of the present work—that of introducing the concept of a "Cantor multi-tower embedding" (**CME**)[16]:

- *What is it?* It refers to a special geometric configuration of a set of tiles that, potentially, could be part of the time-frequency decomposition of the (lacunary) Carleson operator. The essence of this geometric configuration is that it is an *extremizer* for the $L^{1,\infty}$ norm of a "grand maximal counting function" (see (6.3)–(6.3)), a new object, which turns out to play a fundamental role in the behavior of the pointwise convergence of Fourier series near L^1.

  The existence of such tile configurations is a manifestation of the "mass transfer" phenomenon from "heavy" tiles ($P \in \mathbb{P}_n$ with $n \in \mathbb{N}$ close to 1) to "light" tiles (i.e. $P \in \mathbb{P}_n$ for large $n \in \mathbb{N}$) that is capable of realizing a Cantor set structure for each of the sets $E(P)$ corresponding to a P within the given tile configuration.

  Thus, in constructing a **CME**, a key role is played by the *structure* of the sets $E(P)$ and not only by their relative *size*.

- *Context within the literature.* This particular configuration of tiles and the central role played by the corresponding grand maximal counting function are novel facts, which, to the author's knowledge, do not have a direct counterpart in the previous time-frequency literature. However, the more elementary concept of a counting function has been used in many time-frequency papers, and in the framework of the Carleson operator it was first considered in [13].

  Regarded in a broader context, the idea of studying extreme geometric configurations along with their potential key role in deciding the answer to a (harmonic analysis) problem has been successfully applied in many instances. Two such classical examples are given by

  – the (un)boundedness properties of certain (sub)linear operators, e.g. Besicovitch/ Kakeya sets related to the ball multiplier or Bochner–Riesz problems;
  – special topological/additive structure properties of sets, e.g. Cantor sets.

- *What is its purpose?* Based on the geometric location of the tiles within a **CME**, and in particular on the lacunary structure of the frequencies, we will first split the mass parameter n into dyadic blocks. Then, for each block, say B_j, we will construct a corresponding set $F_j \subset \mathbb{T}$ that realizes the alignment of the sign of all the components $\{T_P^*1(x)|_{x \in F_j}\}_{P \in \mathcal{F}_j}$ where $\mathcal{F}_j$ is the collection of all tiles P inside the above mentioned **CME** that have mass parameter $n \in B_j$. In the process we will make essential use of the fact that the exponentials $\{e^{2^j 2\pi i \cdot}\}_{j \in \mathbb{N}}$ oscillate *independently* in [0, 1], behaving similarly to a sequence of i.i.d. random variables. As a consequence, we will be able to "erase" the sign of the operators associated to various trees (of tiles), transforming

---

[16] We warn the reader that these explanations can be truly understood only by experts in the time-frequency area since this concept lies deeply at the heart of the time-frequency methods involved in analyzing the boundedness properties of the Carleson operator. A reader unfamiliar with these techniques might choose to skip this subsection and return to it only after being gradually exposed to the construction of our counterexample.

the adjoint lacunary Carleson operator restricted to this tile configuration into a positive operator. At this point, taking the input function $f = \sum_j r_j \chi_{F_j}$ with $\{r_j\}$ arbitrary positive coefficients and $\{F_j\}$ as above, one concludes that

$$\|C_{\mathrm{lac}} f\|_{1,\infty} \approx \|f\|_{\mathcal{W}}.$$

The precise definition of **CME** is somewhat intricate at the technical and notational level, and therefore we will defer it until later (see Definition 7.2 below). For the time being, sacrificing a bit in the way of rigor, we state the following

Theorem 1.14 (Existence). *The **CME** structures are compatible with the tile decomposition of the* (*lacunary*) *Carleson operator.*

More precisely, these structures can arise within the process of the time-frequency decomposition of the (lacunary) Carleson operator and are composed by tiles that, on the one hand, contain some prescribed "amount" of the graph of the measurable function N appearing in the linearization of the operator and, on the other hand, have some specific relative position to one another. That is why the existence of such structures within the time-frequency decomposition process is non-trivial.

1.5. Remarks

1) The present paper sheds new light on the topic of pointwise convergence of Fourier series near L^1:

- Our Main Theorem 2 establishes near-optimal necessary and sufficient conditions for a Lorentz space Λ_φ to be a $\mathcal{C}_L$-space. It also provides a very good description of when, given a quasi-convex function φ, *any* r.i. Banach space X with $\varphi_X = \varphi$ is a $\mathcal{C}_L$-space. Moreover, it *essentially* states (see (1.14)) that the largest r.i. quasi-Banach space on which C_{lac} is $L^{1,\infty}$-bounded is the space $\mathcal{W}$ introduced in [31].
  In the literature regarding the (almost everywhere) pointwise convergence of Fourier series, almost sharp results of this type constitute a novelty.
- A second item is the *method* of approaching the difficult problem of pointwise convergence of Fourier series near L^1. Until very recently, all progress on this topic, on the positive side, was based on extrapolation theory.
  Using a different perspective—relying only on time-frequency reasonings—the author reproved [30] the best current positive results. The present paper goes beyond offering an alternative approach, as our results (Main Theorems 1 and 2) *cannot*[17] be attained by pure extrapolation methods. Thus, in the context of L^1 methods, this constitutes a first instance when the efficiency of time-frequency techniques is overtaking the canonical extrapolation approach used until now and serves for the idea advocated by the author that *in order to make substantial progress on the problem of convergence of Fourier series near L^1, one needs to leave the general extrapolation theory framework and make* essential *use of the special* structure *of the* (*lacunary*) *Carleson operator.*

---

[17] See for example Corollaries 1.10 and 1.11.

- Thirdly, the spaces $Y = L \log L \log \log \log L$ and $Y = QA$ (i.e. the best known positive results for Open Problem A), viewed for a long time as mere byproducts of extrapolation techniques, are now revealed to be direct manifestations of the "positive behavior" of the operators associated with generic **CME**. Indeed, as a consequence of the ideas introduced here[18] we have the following

Informal principle. *The oscillation and mass transference from heavy to light tiles encapsulated in a generic **CME** structure represent the real challenge in advancing on Open Problem A. If one can reduce the behavior of the adjoint Carleson operator restricted to a general **CME** to the corresponding positive operator (i.e. the absolute sum of the operators associated with the maximal, weight-uniform trees within it) then*, essentially, *the largest Lorentz $\mathcal{C}$-space is precisely Antonov's space $L \log L \log \log \log L$. If on the contrary, considering phase oscillation, one can remove (due to extra cancelation) the threat represented by the operators associated with these geometric structures, then Conjecture* 1 *can be answered affirmatively.*

In view of this heuristic, we can now summarize as follows:

In the *lacunary* situation, based on the *existence* of **CME** configurations and on the oscillatory independence of the lacunary trigonometric system $\{e^{2^j 2\pi i \cdot}\}_{j \in \mathbb{N}}$, one will be able to reduce the adjoint lacunary Carleson operator restricted to this special geometric configuration of tiles (and applied to a special input function) to the corresponding positive operator. Thus, as mentioned earlier, one concludes that $\mathcal{W}$ is essentially the largest C_L-space.

In the *general* situation (i.e. that of the *full* sequence of partial Fourier sums), while one can easily adapt the extremal tile configuration to the new context, there is no analogue of the oscillatory independence of the lacunary trigonometric system. Thus the frequency locations of the tiles play now a determinative role in the boundedness properties of the Carleson operator: if one forms the positive counterpart[19] of the adjoint of C associated with a generic **CME**—call it C_+^*—then one discovers that, essentially, the largest r.i. quasi-Banach space for which C_+ is weak-L^1 bounded is given by QA (notice the analogy with the lacunary case). Conversely, improving on Antonov's and Arias-de-Reyna's results would require precisely showing that there is some cancelation inside the operators associated with a **CME**. This is the key point where techniques from additive combinatorics might play an important role.

As a last remark, one may notice the following interesting analogy: in both the case of the Carleson operator and that of the bilinear Hilbert transform, the current technology (producing the best results to date) stops at the point where one needs to consider the sign/oscillation of some terms associated to particular structures:

– in the Carleson operator case, the current methods cannot do better than provide bounds for the *positive* adjoint Carleson operator C_+^* restricted to a generic **CME**;

---

[18] We here skip detailed explanations as this goes beyond the purpose of the current paper; however, in forthcoming work, we will clarify many of the considerations discussed in the informal principle.

[19] See the informal principle for the way in which this positive operator is defined.

– in the bilinear Hilbert transform case (see e.g. [27]), the current methods can only deal with estimating the absolute values of the elementary building blocks in the Gabor decomposition of the model operator.

This is the point where we believe that further progress (in either of the directions) requires innovative ideas—very likely connected with the additive combinatorial structure of the frequencies of the trees in the time-frequency decomposition of the corresponding operators. Though not in the same context or of similar nature, the paper [40] is a confirmation of the usefulness of additive combinatorics techniques in related time-frequency problems.

2) Passing now to the negative results (i.e. finding "the smallest" r.i. Banach space X with $L \log L \subseteq X \subset L^1$ on which we have divergence of Fourier series), it is likely that part of the ideas in the present paper will help to improve the result(s) in [23] and [24].

3) Finally, as briefly mentioned earlier, the geometry of the tile configuration introduced here is in fact the expression of the behavior of a specific "grand maximal counting function" near L^1. Though we will not detail this subject here, it is worth saying that this function controls each of the counting functions of order n ($n \in \mathbb{N}$), i.e. those functions that count the number of top maximal trees of mass 2^{-n} above each point $x \in [0, 1]$. One should also add that the BMO behavior of each of these counting functions of order n played a fundamental role in removing the exceptional sets in the discretization of the Carleson operator. This last fact generated a first direct proof (i.e. without using interpolation, see [32]) of the *strong L^2* boundedness of the (polynomial) Carleson operator (for an earlier approach on weak-L^2 bounds and strong L^p bounds with $1 < p < 2$ see also [33]).

We plan to detail many of the above considerations regarding pointwise convergence of the *full* sequence of partial Fourier sums near L^1 in a subsequent paper.

Observation 1.15. In what follows we will build a sequence of steps to construct a sharp counterexample to the conjecture of Konyagin. We proceed as follows:

- Section 2 presents a very brief overview of the nature of the counterexample.
- Section 3 reviews the discretization of the Carleson operator following Fefferman's approach [13]. It turns out to be important that this is an *exact* discretization of the Carleson operator, unlike the one provided by Lacey and Thiele [28].
- Section 4 introduces the main definitions required for our further reasonings; it is technically involved.
- In Section 5 we present the main heuristic for our approach and also test the efficiency of the definitions and concepts introduced in Section 4 on a toy model of our problem that already strengthens the best results known to date; it is aimed to prepare the reader for the very technical sections to follow (especially Sections 7–10);
- Section 6 discusses a key concept introduced in the present paper: the grand maximal counting function.
- Section 7 presents the generic construction of a Cantor multi-tower embedding (**CME**).
- Section 8 explains in detail the construction of the input functions corresponding to the sets $\{F_j\}_{j=k/2+1}^k$.

- Section 9 is meant to eliminate "the background noise" arising from the error terms; it can be skipped at the first reading.
- Section 10 contains the proof of Main Theorem 1.
- Section 11 presents the proof of Main Theorem 2 and can be read independently of all the other sections; it relies on extrapolation techniques.
- Section 12 was added at the request of the referee and does not contribute "per se" to either Main Theorems 1 or 2; it can be completely skipped at the first reading. It explains the sharp contrast between the behavior of the (lacunary) Carleson operator (using Fefferman's discretization) and the corresponding behavior of the Lacey–Thiele discretized Carleson model and the (lacunary) Walsh–Carleson operator respectively.
- Section 13 presents several final remarks.
- Finally in the Appendix we recall several standard facts about rearrangement invariant Banach spaces.

We encourage the reader to be patient with the sequence of technical definitions that will soon follow and, at their first glance through the paper, focus more on the main heuristics and "big picture" information provided in Sections 2, 5 and 6.

2. Construction of the counterexample—an overview

In this section we present the general strategy for proving Corollary 1.7.

We will show that for each $k = 2^{2^K}$ with $K \in \mathbb{N}$ large, there exists a function $f_k \in L^\infty$ with the following properties:

- f_k is given by an expression of the form

$$f_k = \sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} \chi_{F_j}, \tag{2.1}$$

where χ_{F_j} designates the characteristic function of F_j and each set F_j has some prescribed properties that will be detailed shortly;
- the $L \log \log L$ norm is under control:

$$\|f_k\|_{L \log \log L} \approx 1.$$

- the weak-L^1 norm of $C_{\text{lac}}^{\{2^j\}_j} f_k$ is large:

$$\|C_{\text{lac}}^{\{2^j\}_j} f_k\|_{L^{1,\infty}} \gtrsim \log k.$$

The construction of each F_j requires some technicalities and will be detailed later. As of now, we limit ourselves to revealing the following properties:

$$\begin{aligned} &\bullet \ F_j \subseteq [0,1] \text{ is a measurable set;} \\ &\bullet \ |F_j| \approx 2^{-\log k \, 2^{2^j}} \cdot 2^{-j} \cdot \tfrac{1}{k}; \\ &\bullet \ F_j \text{ has a finite Cantor type structure.} \end{aligned} \tag{2.2}$$

We end this section by mentioning that the construction of the sets $\{F_j\}_j$ will directly depend on the choice of the measurable function N in the linearization of the $\{2^j\}_j$-lacunary Carleson operator (see next section). Consequently, understanding/designing the structure of the set $\mathbb{P}$ of tiles appearing in the decomposition of the $\{2^j\}_j$-lacunary Carleson operator $C_{\text{lac}}^{\{2^j\}_j}$ is a precondition for assigning the precise properties to each F_j.

3. Discretization of our operator

Let us first recall the main object of our study[20]

$$C_{\text{lac}}^{\{2^j\}_j} f(x) \approx \sup_{j \in \mathbb{N}} \left| \int_{\mathbb{T}} \frac{1}{x-y} e^{2\pi i 2^j (x-y)} f(y)\, dy \right|. \tag{3.1}$$

Applying Fefferman's discretization [13], we follow the same steps as in [31]:

- We linearize our operator and write

$$Tf(x) := \int_{\mathbb{T}} \frac{1}{x-y} e^{-2\pi i N(x) y} f(y)\, dy,$$

  where $N : \mathbb{T} \to \{2^j\}_{j \in \mathbb{N}}$ is a measurable function. (Here, for technical reasons, we erase the term $N(x)x$ in the phase of the exponential, as later in the proof this will simplify the structure of the adjoint operators T^*.)
- We use the dilation symmetry of the kernel and express

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

  where $\psi_k(y) := 2^k \psi(2^k y)$ (with $k \in \mathbb{N}$) and ψ is an odd C^∞ function such that

$$\operatorname{supp} \psi \subseteq \{y \in \mathbb{R} \mid 2 < |y| < 8\}. \tag{3.2}$$

- We write[21]

$$Tf(x) = \sum_{k \in \mathbb{N}} \int_{\mathbb{T}} e^{-2\pi i N(x) y} \psi_k(x-y) f(y)\, dy.$$

- For each $k \in \mathbb{N}$, we partition the time-frequency plane into *tiles* (rectangles of area one) of the form $P = [\omega, I]$ with ω, I dyadic intervals (with respect to the canonical dyadic grids on $\mathbb{R}$ and $[0, 1]$ respectively) such that $|\omega| = |I|^{-1} = 2^k$. The set of all such tiles will be denoted by $\bar{\mathbb{P}}^k$. Further, we set $\bar{\mathbb{P}} = \bigcup_{k \in \mathbb{N}} \bar{\mathbb{P}}^k$.

---

[20] Standard reasoning reduces the study of C_{lac} to the corresponding operator having $\cot(\pi(x-y))$ replaced by $\frac{1}{x-y}$.

[21] Throughout the paper we use the convention that $0 \in \mathbb{N}$, thus $\mathbb{N} = \{0, 1, 2, \ldots\}$.

- To each $P = [\omega, I] \in \bar{\mathbb{P}}$ we assign the set

$$E(P) := \{x \in I \mid N(x) \in \omega\},$$

responsible for the mass (or "weight") of the tile, $|E(P)|/|I|$. The mass concept will later play a key role in partitioning the set $\bar{\mathbb{P}}$.
- For $P = [\omega, I] \in \bar{\mathbb{P}}^k$ with $k \in \mathbb{N}$ we define the operators

$$T_P f(x) = \left\{ \int_{\mathbb{T}} e^{-2\pi i N(x)y} \psi_k(x - y) f(y) \, dy \right\} \chi_{E(P)}(x), \qquad (3.3)$$

and conclude that

$$T f(x) = \sum_{P \in \bar{\mathbb{P}}} T_P f(x). \qquad (3.4)$$

Notice that if we think of $N : \mathbb{T} \to \{2^j\}_j$ as a predefined measurable function then the above decomposition is *independent* of the function f.

Observation 3.1. For $P = [\omega_P, I_P] \in \bar{\mathbb{P}}$ let $c(I_P)$ be the center of the interval I_P and define $I_{P^*} = [c(I_P) - \frac{17}{2}|I_P|, c(I_P) - \frac{3}{2}|I_P|] \cup [c(I_P) + \frac{3}{2}|I_P|, c(I_P) + \frac{17}{2}|I_P|]$. From (3.2) and (3.3), we deduce that

$$\operatorname{supp} T_P \subseteq I_P, \qquad (3.5)$$

while the adjoint operator of T_P denoted by T_P^* obeys[22]

$$\operatorname{supp} T_P^* \subseteq I_{P^*}. \qquad (3.6)$$

As a consequence, if $P_1, P_2 \subset \mathbb{P}$ are such that $I_{P_1} \subset I_{P_2}$ and $|I_{P_1}| \leq 2^{-10}|I_{P_2}|$, then

$$\operatorname{supp} T_{P_1}^* \cap \operatorname{supp} T_{P_2}^* = \emptyset. \qquad (3.7)$$

By standard reasoning we will be able to arrange that the following holds: if $P_1, P_2 \in \bar{\mathbb{P}}$ and $|I_{P_1}| \neq |I_{P_2}|$ then $|I_{P_1}| \leq 2^{-10}|I_{P_2}|$ or $|I_{P_2}| \leq 2^{-10}|I_{P_1}|$. Thus (3.7) is automatically guaranteed if $I_{P_1} \subsetneq I_{P_2}$.

We will make repeated use of this observation in our construction process.

Notation. Throughout the paper, if I is a (dyadic) interval of center $c(I)$, and $d > 0$ a positive constant, then dI designates the interval having the same center $c(I)$ and length $|dI| := d|I|$. Also, if $P = [\omega, I]$ and $a > 0$ then we define the tile-dilation $aP := [a\omega, I]$.

---

[22] This is a direct consequence of (3.5) and of the fact that ψ_k is compactly supported.

4. Main definitions and preparations

In this section we will introduce several of the basic concepts which will be used later in the proof. The first three definitions were introduced in [13], while Definitions 4.5 and 4.6 were first developed in [32].

Definition 4.1 (weighting the tiles). We define the *mass* of $P = [\omega, I] \in \bar{\mathbb{P}}$ as

$$A(P) := \sup_{\substack{P' = [\omega', I'] \in \bar{\mathbb{P}} \\ I \subseteq I'}} \frac{|E(P')|}{|I'|} \frac{1}{(1 + \operatorname{dist}(10\omega, 10\omega')/|\omega|)^{N_0}} \tag{4.1}$$

where N_0 is a fixed large natural number and if $A, B \subseteq \mathbb{R}$ then we write $\operatorname{dist}(A, B) = \inf_{a \in A, b \in B} |a - b|$.

We also refer to the *restricted mass* or r-mass of $P = [\omega, I] \in \bar{\mathbb{P}}$ as

$$A_0(P) := |E(P)|/|I_P|.$$

Definition 4.2 (ordering the tiles). Let $P_j = [\omega_j, I_j] \in \bar{\mathbb{P}}$ with $j \in \{1, 2\}$. We say that $P_1 \leq P_2$ if $I_1 \subseteq I_2$ and $\omega_1 \supseteq \omega_2$. We write $P_1 < P_2$ if $P_1 \leq P_2$ and $|I_1| < |I_2|$.

Notice that $\leq$ defines a partial order relation on the set $\bar{\mathbb{P}}$.

Observation 4.3. We will define various families of tiles with prescribed analytic and geometric properties (relating to the mass of a tile and to the order relation $\leq$, respectively). In order to do so, we will introduce several refinements of the set $\bar{\mathbb{P}}$, always keeping in mind that the analytic and geometric properties that we will describe are strongly influenced by the key fact

$$\operatorname{Image}(N) \subseteq \{2^j\}_{j \in \mathbb{N}}. \tag{4.2}$$

Let us first define

$$\mathbb{P}(0) := \{P \in \bar{\mathbb{P}} \mid 0 \in 100\omega_P\}, \quad \bar{\mathbb{P}}_0 := \{P \in \bar{\mathbb{P}} \setminus \mathbb{P}(0) \mid A_0(P) = 0\}, \quad \mathbb{P} := \bar{\mathbb{P}} \setminus (\mathbb{P}(0) \cup \bar{\mathbb{P}}_0).$$

For $n \in \mathbb{N}$, we further set

$$\mathbb{P}_n := \{P \in \mathbb{P} \mid A(P) \in (2^{-n-1}, 2^{-n}]\}. \tag{4.3}$$

From now on, we will say that a tile P has *weight n* if $P \in \mathbb{P}_n$.

Later on, in our construction of a **CME**, will be useful to impose the following restriction on the measurable function N:

$$\operatorname{Image}(N) \subseteq \{2^{2^{2^{2^k}} + 100m}\}_{m \in \{0, \ldots, \log_k 2^{2^k - 1}\}}, \tag{4.4}$$

where we recall that $k \in \mathbb{N}$ is a fixed large parameter.

Also we will ask that each tile $P \in \mathbb{P}$ obeys

$$A(P) \geq 2^{-2^{2k}}. \tag{4.5}$$

Consequently, we deduce that

$$\mathbb{P} = \bigcup_{n \leq 2^{2k}} \mathbb{P}_n. \tag{4.6}$$

In particular we will only work under the assumption that $\mathbb{P}$ is finite.

Definition 4.4 (tree). We say that a set $\mathcal{P} \subset \mathbb{P}$ of tiles is a *tree* with *top* $P_0 \in \mathbb{P}$ if:

(1) $P \leq P_0$ for all $P \in \mathcal{P}$;
(2) if $P_1, P_2 \in \mathcal{P}$ and $P_1 \leq P \leq P_2$ then $P \in \mathcal{P}$.

Definition 4.5 (sparse tree). We say that a tree $\mathcal{P} \subset \mathbb{P}$ is *sparse* if for any $P \in \mathcal{P}$ we have

$$\sum_{\substack{P' \in \mathcal{P} \\ I_{P'} \subseteq I_P}} |I_{P'}| \leq C|I_P|, \tag{4.7}$$

where $C > 0$ is an absolute constant.

Definition 4.6 (forest; L^∞ control over union of trees). Fix $n \in \mathbb{N}$. We say that $\mathcal{P} \subseteq \mathbb{P}_n$ is an L^∞-*forest* (of nth generation) if:

(i) $\mathcal{P}$ is a collection of *separated* trees, i.e. $\mathcal{P} = \bigcup_{j \in \mathbb{N}} \mathcal{P}_j$ with each $\mathcal{P}_j$ a tree with top $P_j = [\omega_j, I_j]$ and such that

$$\forall j' \neq j, \; P \in \mathcal{P}_j \quad 2P \nleq 10P_{j'}; \tag{4.8}$$

(ii) the *counting function*

$$\mathcal{N}_{\mathcal{P}}(x) := \sum_j \chi_{I_j}(x) \tag{4.9}$$

obeys the estimate $\|\mathcal{N}_{\mathcal{P}}\|_{L^\infty} \lesssim 2^n$.

Further, if $\mathcal{P} \subseteq \mathbb{P}_n$ only consists of sparse separated trees then we refer to $\mathcal{P}$ as a *sparse L^∞-forest*.

Observation 4.7. In this paper we focus on the decomposition of our set $\mathbb{P}$ of tiles into L^∞-forests with some prescribed properties (see below). For this reason, unlike for the preceding tile decompositions in [32] and [30] (where we have introduced the concept of a BMO-forest), we will refer to an L^∞-forest simply as a *forest*.

Definition 4.8 (generalized forest, GF). Let $r, n \in \mathbb{N}$ with $r \leq n$. We say that $\mathcal{P} \subseteq \mathbb{P}$ is a *generalized forest* of generation (r, n) if we can decompose

$$\mathcal{P} = \bigcup_{j=r}^{n} \mathcal{P}[j] \tag{4.10}$$

so that:

- each $\mathcal{P}[j]$ is an (L^∞-)forest of jth generation;
- if $\mathcal{P}[j] = \bigcup_l \mathcal{P}_l[j]$ is the decomposition of $\mathcal{P}[j]$ into maximal separated trees, then for any pair (j, j') with $r \leq j < j' \leq n$, any l and any $P \in \mathcal{P}_l[j]$ there exist l' and $P' \in \mathcal{P}_{l'}[j']$ such that $P \leq P'$.

Definition 4.9 (saturated generalized forest, SGF). Let $r, n \in \mathbb{N}$ with $r \leq n$. We say that $\mathcal{P} \subseteq \mathbb{P}$ is a *saturated generalized forest* of generation (r, n) if:

- $\mathcal{P}$ is a GF of generation (r, n);
- if there is a $P \in \mathbb{P}_j$ such that $P > P'$ for some $P' \in \mathcal{P}[j]$, then $P \in \mathcal{P}[j]$;
- if there is a $P \in \mathbb{P}_j$ such that $P < P'$ for some $P' \in \mathcal{P}[j]$, then $P \in \mathcal{P}[j]$.

Observation 4.10. All the previous definitions make perfect sense in a general context, with no particular restriction on the linearization function N. The structures introduced in the next two definitions, though, are not present (in a non-trivial form) for an arbitrary choice of N. However, in the context of this paper, we will have the liberty of *choosing N*, and thus guarantee their existence. The precise form of the definitions below is chosen in order to simplify as much as possible the general tile-configuration of the counterexample. For other, more general purposes, the requirements in these definitions can be significantly relaxed.

Definition 4.11 (uniform saturated generalized forest, USGF). Let $r, n \in \mathbb{N}$ with $r \leq n$. We say that $\mathcal{P} \subseteq \mathbb{P}$ is a *uniform saturated generalized forest* of generation (r, n) if:

- $\mathcal{P}$ is a SGF of generation (r, n);
- for each $j \in \{r, \dots, n\}$ there exists $C_j \in (0, 1]$ such that

$$|I_{P_l^j}| = C_j \quad \forall l,$$

where, with the notation above, P_l^j stands for the top of $\mathcal{P}_l[j]$.

Definition 4.12 (uniform saturated generalized top-forest, USGTF). We say that $\mathcal{P} \subseteq \mathbb{P}$ is a *uniform saturated generalized top-forest* of generation (r, n) if $\mathcal{P}$ is a USGF of generation (r, n) and each tree $\mathcal{P}_l[j]$ in the definition above consists of just a single tile, its top.

In this paper we will only work with a *special* type of USGTF's. We will describe their properties in what follows, but first we need some more notation.

If $J \subseteq [0, 1]$ is a dyadic interval and $m \in \mathbb{N}$, then we define

$$\mathcal{I}_m(J) := \{I \subseteq J \mid I \text{ dyadic}, \ |I| = |J|2^{-m}\}. \tag{4.11}$$

From now on, we will always apply the following convention: if $\{I_s\}_s$ is a collection of disjoint dyadic space intervals, then the indexing s reflects the relative position of these intervals in $[0, 1]$ from left to right, i.e., if $s_1 < s_2$ then $c(I_{s_1}) < c(I_{s_2})$.

Observation 4.13. In what follows we will use an alternative description of a tile $P = [\omega, I]$, namely $P := I \times \alpha$, where $\omega = [l(\omega), r(\omega))$ and $\alpha := l(\omega)$. This is justified since knowing I and $l(\omega)$ completely determines P (recall that the area of the rectangle determined by P is always assumed to be one). Notice that due to the definition of $C_{\text{lac}}^{\{2^j\}_j}$ we can always assume that $l(\omega) \in \{2^j\}_{j \in \mathbb{N}}$.

Observation 4.14. From now on, whenever we refer to a family $\mathcal{F} \subset \mathbb{P}$ as a USGTF (of generation (r, n)) we will specify three sets of parameters, which, following the algorithm described below, will completely determine $\mathcal{F}$:

The three sets of parameters:

- the collection of disjoint dyadic same-length space intervals
$$\mathrm{ITop}(\mathcal{F}) := \{I_j\}_j.$$

- the collection of distinct frequencies (arranged in increasing order)
$$\alpha(\mathcal{F}) := \{\alpha_u\}_{u=1}^{2^{n-1}}. \tag{4.12}$$

- the collection of disjoint dyadic (same-length) space intervals
$$\mathrm{IBtm}(\mathcal{F}) := \bigcup_j \mathcal{I}_{n-r}(I_j).$$

The algorithm (*which completely determines* $\mathcal{F}$). We *define* $\mathcal{F}$ to be the USGTF (of generation (r, n)) obeying:

- The collection of tiles of weight n in $\mathcal{F}$, denoted by $\mathcal{F}[n]$, is given by
$$\mathcal{F}[n] := \{I \times \alpha \mid I \in \mathrm{ITop}(\mathcal{F}) \,\&\, \alpha \in \alpha(\mathcal{F})\}.$$

  Also each $P \in \mathcal{F}[n]$ has
$$A(P) = A_0(P) = 2^{-n}. \tag{4.13}$$

- The collection of tiles of weight r in $\mathcal{F}$, denoted by $\mathcal{F}[r]$, is given by
$$\mathcal{F}[r] := \bigcup_j \bigcup_{l=1}^{2^{n-r}} \big\{ I_l \times \alpha \mid I_l \in \mathcal{I}_{n-r}(I_j) \,\&\, \alpha \in \{\alpha_{(l-1)2^{r-1}+1}, \ldots, \alpha_{l\,2^{r-1}}\} \big\}.$$

  As before, we require that $P \in \mathcal{F}[r]$ implies
$$A(P) = A_0(P) = 2^{-r}. \tag{4.14}$$

  Thus, if $P \in \mathcal{F}[n]$, we impose the condition that
$$\text{there exists a unique } P' \in \mathcal{F}[r] \text{ such that } P' < P, \tag{4.15}$$

  and on top of that we require
$$E(P) = E(P'). \tag{4.16}$$

Thanks to the above requirements, with a particular emphasis on (4.15) and (4.16), all the tiles in the intermediate families $\{\mathcal{F}[l]\}_{r<l<n}$ are now completely determined.

Thus, $\mathcal{F}$ is indeed completely determined by the three sets of parameters in (4.12) once we agree that we always run the above algorithm.

Notation. Let $\mathcal{A} = \{\mathcal{A}_j\}_j$, $\mathcal{B} = \{\mathcal{B}_k\}_k$ be two collections of disjoint dyadic intervals. We write $\mathcal{A} \prec \mathcal{B}$ if each $\mathcal{A}_j$ contained in some $\mathcal{B}_k$. Also we define
$$\tilde{\mathcal{A}} := \bigcup_j \mathcal{A}_j.$$

If $\mathcal{A}_j^{\mathrm{lt}}$ designates the left child of the interval $\mathcal{A}_j$ then we set

$$\mathcal{A}^{\mathrm{lt}} := \{\mathcal{A}_j^{\mathrm{lt}}\}_j.$$

The same convention applies to $\mathcal{A}^{\mathrm{rt}}$, the collection of right children of the intervals in $\mathcal{A}$.

Recalling the definition of $\mathcal{I}_m(J)$ in (4.11), for $\mathcal{A}$ as before we set

$$\mathcal{I}_m(\mathcal{A}) = \bigcup_j \mathcal{I}_m(\mathcal{A}_j).$$

Also, for J a dyadic interval, if $\mathcal{I}_m(J) = \{I_l\}_l$ then $\mathcal{I}_m^{\mathrm{lt}}(J) := \{I_l^{\mathrm{lt}}\}_l$ and $\mathcal{I}_m^{\mathrm{rt}}(J) := \{I_l^{\mathrm{rt}}\}_l$. Further, $\mathcal{I}_m^{\mathrm{lt}}(\mathcal{A}) = \bigcup_j \mathcal{I}_m^{\mathrm{lt}}(\mathcal{A}_j)$, and similarly for $\mathcal{I}_m^{\mathrm{rt}}(\mathcal{A})$.

If $\mathcal{A} = \mathrm{IBtm}(\mathcal{F})$ for $\mathcal{F}$ a USGTF, then we set $\tilde{\mathrm{I}}\mathrm{Btm}(\mathcal{F}) = \tilde{\mathcal{A}}$, $\mathrm{I}^{\mathrm{rt}}\mathrm{Btm}(\mathcal{F}) = \mathcal{A}^{\mathrm{rt}}$ and $\mathrm{I}^{\mathrm{lt}}\mathrm{Btm}(\mathcal{F}) = \mathcal{A}^{\mathrm{lt}}$. With the obvious changes, the same applies to $\mathcal{A} = \mathrm{ITop}(\mathcal{F})$.

Definition 4.15 (tower). We say that $\mathcal{P} \subseteq \mathbb{P}$ is a *tower* of generation (r, n) if there exists $m \in \mathbb{N}$, $m \geq 1$, such that $\mathcal{P} = \bigcup_{l=1}^{m} \mathcal{P}_l$ and

- each $\mathcal{P}_l$ is a USGTF of generation (r, n);
- $\mathrm{ITop}(\mathcal{P}_{l+1}) \prec \mathrm{IBtm}(\mathcal{P}_l)$ for any $l \in \{1, \ldots, m-1\}$.

The two items above imply that for all $l \neq l'$ and all $P \in \mathcal{P}_l$ and $P' \in \mathcal{P}_{l'}$ one has

$$P \not\leq P' \quad \text{and} \quad P' \not\leq P. \tag{4.17}$$

In particular[23]

$$l \neq l' \implies \alpha(\mathcal{P}_l) \cap \alpha(\mathcal{P}_{l'}) = \emptyset. \tag{4.18}$$

The number of USGTF's is called the *height* of the tower $\mathcal{P}$, while $\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{P}_1)$ stands for its *basis*; we write

$$\mathrm{Height}(\mathcal{P}) = m \quad \text{and} \quad \mathrm{Basis}(\mathcal{P}) = \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{P}_1).$$

Definition 4.16 (multi-tower). We say that $\mathcal{M} \subseteq \mathbb{P}$ is a *multi-tower* of generation (r, n) if one can decompose it as $\mathcal{M} = \bigcup_l \mathcal{M}_l$ in such a way that

- each $\mathcal{M}_l$ is a tower of generation (r, n);
- $\mathrm{Basis}(\mathcal{M}_l) \cap \mathrm{Basis}(\mathcal{M}_{l'}) = \emptyset$ for any $l \neq l'$.

Definition 4.17 (multi-tower embedding). If $\mathcal{F}^1$, $\mathcal{F}^2$ are two (multi-)towers, with $\mathcal{F}^j$ of generation (r_j, n_j), we say that $\mathcal{F}^1$ *embeds* into $\mathcal{F}_2$, and write $\mathcal{F}^1 \sqsubset \mathcal{F}^2$, if

$$n_1 \leq r_2, \quad \forall P \in \mathcal{F}^1[n_1] \ \exists P' \in \mathcal{F}^2[r_2] \quad P \leq P'.$$

In particular, if $\mathcal{F}^1$ and $\mathcal{F}^2$ are USGTF's, we must have

$$\mathrm{ITop}(\mathcal{F}_1) \prec \mathrm{IBtm}(\mathcal{F}_2) \quad \text{and} \quad \alpha(\mathcal{F}_1) \subset \alpha(\mathcal{F}_2).$$

This finishes the preparations for presenting the main components of our proof.

---

[23] If (4.18) were not true then $a \in \alpha(\mathcal{P}_l) \cap \alpha(\mathcal{P}_{l'})$ would imply that there exist $P \in \mathcal{P}_l$ and $P' \in \mathcal{P}_{l'}$ with $\alpha(P) = \alpha(P') = a$; this together with the second item in Definition 4.15 would imply $P \leq P'$ or $P' \leq P$, contradicting (4.17).

5. Heuristics and a warm-up example

In order to smooth out the transition between two technical sections of our paper and to help clarify the "big picture" in our reasonings, we start with

Main Heuristic. *Our aim is to design some special function N that will give rise to a family of embedded multi-towers, i.e., a **multi-tower chain** with respect to the embedding relation "$\sqsubset$". This chain will loosely have the form*

$$\mathcal{F} = \bigcup_{k/2 < j \le k} \mathcal{F}_j, \tag{5.1}$$

such that

- *each $\mathcal{F}_j$ is a multi-tower of generation $(2^{j-1}, 2^j)$;*
- *$\mathcal{F}_j \sqsubset \mathcal{F}_{j+1}$.*

*At the informal level, a (lacunary) **CME** will be a chain that maximizes the $L^{1,\infty}$ norm of a **grand maximal counting function**, a notion that will be our main focus in the section to follow.*
*As a consequence of this requirement, for $\mathcal{F}$ a **CME** and a generic tile $P \in \mathcal{F}$, the set $E(P)$ has a Cantor-type distribution inside I_P.*

Next, we would like to motivate the necessity of considering the **CME** concept and why we were required to develop the notions of multi-tower and chain of multi-towers. For this, we will first discuss a simpler toy model, naturally developing from our introductory discussion in Section 1.

As mentioned in the Introduction, from [21] (see also [2]) we know that if $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing with $\phi(0) = 0$ and $\phi(u) = o(u \log \log u)$ as $u \to \infty$ then $\phi(L) = \Lambda_{\bar{\phi}}$ is not a $\mathcal{C}_L$-space. This result can now be easily deduced from the following stronger claim:

Proposition 5.1. *There exists an absolute constant $C > 0$ and a sequence $\{F_k\}_{k \in \mathbb{N}}$ of measurable sets with the following properties:*

- *$F_k \subseteq \mathbb{T}$ with $|F_k| \to 0$ as $k \to \infty$;*
- *for any $k \in \mathbb{N}$,*

$$\|C_{\text{lac}}^{\{2^j\}_j} \chi_{F_k}\|_{1,\infty} \ge C |F_k| \log \log \frac{4}{|F_k|}. \tag{5.2}$$

The *idea* of the proof relies on the newly introduced concept of tower: we let $\mathcal{F}_k$ be a tower of height 1 that is a single USGTF! More precisely, using the language introduced in Observation 4.14, we define the collection $\mathcal{F}_k$ of tiles to be the USGTF of generation $(0, 2^k)$ given by the following characteristics:

- the collection of disjoint dyadic same-length space intervals

$$\text{ITop}(\mathcal{F}_k) := \{[0, 1]\};$$

- the collection of distinct frequencies[24]

$$\alpha(\mathcal{F}_k) := \{2^{2^{2^{2^k}} + 100m}\}_{m \in \{1, \ldots, 2^{2^k}\}};$$

- the collection of disjoint dyadic (same-length) space intervals

$$\mathrm{IBtm}(\mathcal{F}_k) := \bigcup_j \mathcal{I}_{2^k}([0, 1]).$$

Next, one can construct a set F_k with the following properties:

- the *size* given by

$$|F_k| \approx 2^{-2^{2^k}},$$

- the *structure* such that[25]

$$\int \mathrm{Re}(\chi_{F_k} T_P^*(1))(\cdot) \approx \int |\chi_{F_k} T_P^*(1)(\cdot)| \quad \text{for all } P \in \mathcal{F}_k.$$

Then, recalling that $\mathcal{F}_k[n]$ stands for the tiles in $\mathcal{F}_k$ of weight n, we conclude that

$$\|C_{\mathrm{lac}}^{\{2^j\}_j} \chi_{F_k}\|_{1,\infty} \gtrsim \sum_{n=1}^{2^k} \sum_{P \in \mathcal{F}_k[n]} 2^{-n} \frac{|I_P \cap F_k|}{|I_P|} |I_P| \geq 2^k |F_k|, \tag{5.3}$$

thus proving our proposition.

Observation 5.2. We stress here that for the proof of the above proposition there was no need to consider a chain of towers since we were not aiming at the extra $\log \log \log \log$ term in (5.2); equivalently, our reasoning involved a single characteristic function of a set instead of an input function f_k expressed as in (2.1), as a linear combination of $\approx k$ characteristic functions of sets. It is thus natural that once we turn to proving our Main Theorem 1, we need to consider the more involved concept of multi-tower and finally that of **CME**.

6. The grand maximal counting function

As announced, in this section we will elaborate on and motivate the introduction of the new concept of *grand maximal counting function*, which is defined as

$$\mathcal{N} := \sup_j \mathcal{N}_j, \tag{6.1}$$

where, recalling (4.3), we set

$$\mathcal{N}_j := \frac{1}{2^{j-1}} \sum_{n=2^{j-1}+1}^{2^j} \frac{1}{2^{n-1}} \sum_{P \in \mathbb{P}_n^{\max}} \chi_{I_P}, \tag{6.2}$$

with $\mathbb{P}_n^{\max}$ designating the maximal tiles[26] P in $\mathbb{P}_n$ such that $A(P) > 2^{-n-1}$.

---

[24] One can notice that here we made a minor modification of the USGTF model described in Observation 4.14 by requiring that the stack of tiles having mass 2^k have height 2^{2^k} instead of 2^{2^k-1}.

[25] The mechanism of realizing the size and structure conditions of F_k is described in full generality (i.e. in the multi-tower case) in Section 8.1, and thus it will not be detailed here.

[26] Relative to "$\leq$".

Observation 6.1. The motivation for defining the counting functions $\mathcal{N}_j$ and $\mathcal{N}$ originates in [32] where the author used a complex greedy algorithm (involving more basic counting functions) in order to remove the exceptional sets arising in the time-frequency discretization and to provide direct strong $(2, 2)$ bounds for the (standard) Carleson operator. Notice that $\mathcal{N}_j$ roughly controls the average spacial density location[27] of the maximal trees of mass parameter $n \approx 2^j$, while $\mathcal{N}$ picks the worst (largest) such density location among all the possible dyadic mass scales. The normalization factor $1/2^{j-1}$ in (6.2) preserves a uniform control of the BMO norm of $\mathcal{N}_j$, that is, $\|\mathcal{N}_j\|_{\mathrm{BMO}(\mathbb{T})} \leq 10$ for any $j \in \mathbb{N}$.

We move now to a further elaboration of the Main Heuristic whose key message is: "a **CME** is a chain that maximizes the $L^{1,\infty}$ norm of the grand maximal counting function".

Observation 6.2. Define the *k-truncated grand maximal counting function* as

$$\mathcal{N}^{[k]} := \sup_{1 \leq j \leq k} \mathcal{N}_j, \tag{6.3}$$

and notice that under the assumptions (4.4)–(4.6) we trivially have

$$\mathcal{N}^{[2k]} = \mathcal{N}. \tag{6.4}$$

Next, as a consequence of [32], for each $j \in \mathbb{N}$ with $j \leq 2k$ we have

$$\|\mathcal{N}_j\|_{\mathrm{BMO}} \lesssim 1. \tag{6.5}$$

This, together with the standard John–Nirenberg inequality, gives, as we will see momentarily,

$$\|\mathcal{N}^{[2k]}\|_{1,\infty} \leq \|\mathcal{N}^{[2k]}\|_1 \lesssim \log k. \tag{6.6}$$

The crux of our main results is that one can construct special configurations inside $\mathbb{P}$, corresponding to chains as in (5.1), such that the inequality (6.6) can be reversed. In these instances, one can thus show that

$$\|\mathcal{N}^{[2k]}\|_{1,\infty} \approx \log k. \tag{6.7}$$

We now start a more detailed analysis of the properties of the grand maximal counting function $\mathcal{N}$ by first presenting a short proof of (6.6).

For this we start by defining

$$\overline{\mathcal{N}}_n := \frac{1}{2^{n-1}} \sum_{P \in \mathbb{P}_n^{\max}} \chi_{I_P},$$

and notice that

$$\mathcal{N}_j = \frac{1}{2^{j-1}} \sum_{n=2^{j-1}+1}^{2^j} \overline{\mathcal{N}}_n. \tag{6.8}$$

---

[27] That is, the number of maximal trees sitting above a point x as x runs through the interval $[0, 1]$.

Following the reasoning in [32], we find that for every $n \in \mathbb{N}$ the function $\overline{\mathcal{N}}_n$ belongs to BMO($\mathbb{T}$) with

$$\|\overline{\mathcal{N}}_n\|_{\text{BMO}(\mathbb{T})} \leq 10.$$

Now, since $\mathcal{N}_j$ is an arithmetic mean of functions of the type $\overline{\mathcal{N}}_n$, we further deduce that

$$\|\mathcal{N}_j\|_{\text{BMO}(\mathbb{T})} \leq 10, \tag{6.9}$$

and hence applying the John–Nirenberg inequality we find that there exists a universal constant $c > 0$ such that for any $\gamma > 0$ and $j \in \mathbb{N}$,

$$|\{x \in \mathbb{T} \mid \mathcal{N}_j(x) > \gamma\}| \leq e^{-c\gamma}. \tag{6.10}$$

Next, for any $C > 0$, we have

$$\|\mathcal{N}^{[k]}\|_1 \leq C \log k + \sum_{j=1}^{k} \|\mathcal{N}_j|_{\mathcal{N}_j > C \log k}\|_1. \tag{6.11}$$

Choosing now in (6.10) $\gamma = C \log k$ with $C = 1/c$ and inserting it in (6.11) we deduce that

$$\|\mathcal{N}^{[k]}\|_1 \leq (C + 1) \log k, \tag{6.12}$$

thus proving (6.6).

We pass now to the proof of (6.7). (For the moment, the reader is invited to think of $\mathcal{F}$ in the more vague terms described in the Main Heuristic corresponding to (5.1); later on, if desired, one can consult the precise version given in Definition 7.2.)

Recall that we want to show that if for a given large $k \in \mathbb{N}$ the family $\mathbb{P}$ contains a family $\mathcal{F} = \mathcal{F}(k)$ of tiles as in (5.1) which is also a **CME**, then there exists $\bar{C} > 0$ such that

$$\|\mathcal{N}^{[k]}\|_{1,\infty} \geq \bar{C} \log k. \tag{6.13}$$

First we present the heuristic for why one would believe such a statement. This is based on the following list of loosely stated observations:

- for $\gamma \gg C \log k$ the level sets $\{\{\mathcal{N}_j > \gamma\}\}_{j=1}^{k}$ do not significantly contribute to the norm $\|\mathcal{N}^{[k]}\|_{1,\infty}$;
- similarly, for $\gamma \ll C \log k$ the level sets $\{\{\mathcal{N}_j < \gamma\}\}_{j=1}^{k}$ cannot provide an estimate of type (6.13); $\tag{6.14}$

- there exists a set $\mathbb{P}$ of tiles such that for suitable absolute positive constants C_1, C_2,

$$|\{\mathcal{N}_j > C_1 \log k\}| \geq C_2/k \quad \forall k/2 < j \leq k. \tag{6.15}$$

The first two items are just simple consequences of (6.10) and (6.11). The third item will be a direct byproduct of the construction of the **CME** $\mathcal{F} = \bigcup_{j=k/2+1}^{k} \mathcal{F}_j$ presented in the next section.

Now, if on top of property (6.15) one could arrange that the functions $\{\mathcal{N}_j\}_{j=k/2+1}^{k}$ behaved morally as if they were *independent* random variables, then we would immediately conclude that

$$|\{\mathcal{N}^{[k]} > C_1 \log k\}| \gtrsim \sum_{j=k/2+1}^{k} |\{\mathcal{N}_j > C_1 \log k\}| \overset{(6.15)}{\geq} C_2/2, \tag{6.16}$$

thus proving (6.13).

The main point of the construction in Section 6 is that the **CME** as given by Definition 7.2 provides exactly the mutual independence behavior of $\{\mathcal{N}_j\}_{j=k/2+1}^{k}$ mentioned above.

Properties (6.15) and (6.16) will be a byproduct of the construction of $\mathcal{F}$ presented in the next section (see Definition 7.2). Indeed, writing, with the usual notation, $\mathcal{F} = \bigcup_{j=k/2+1}^{k} \mathcal{F}_j$, one will be able to decompose each multi-tower $\mathcal{F}_j$ into a controlled number of towers, namely $\mathcal{F}_j = \bigcup_{l=1}^{\log k} \mathcal{F}_j^l$, and then deduce the key properties[28]

$$|\mathrm{Basis}(\mathcal{F}_j^{\log k})| \geq \bar{c}/k \quad \forall j \in \{k/2+1, \ldots, k\} \tag{6.17}$$

and

$$\mathrm{Basis}(\mathcal{F}_{j_1}^{\log k}) \cap \mathrm{Basis}(\mathcal{F}_{j_2}^{\log k}) = \emptyset \quad \forall j_1 \neq j_2 \in \{k/2+1, \ldots, k\}, \tag{6.18}$$

where $\bar{c} > 0$ is an absolute constant.

Relations (6.17) and (6.18) will then imply[29]

- for all $j \in \{k/2+1, \ldots, k\}$ and $m \in \{2^{j-1} + \log \log k, \ldots, 2^j\}$,

$$|\{\overline{\mathcal{N}}_m \geq \log k\}| \geq \bar{c}/k, \tag{6.19}$$

- for all $m_1 \in \{2^{j_1-1} + \log \log k, \ldots, 2^{j_1}\}$, $m_2 \in \{2^{j_2-1} + \log \log k, \ldots, 2^{j_2}\}$ and $j_1 \neq j_2 \in \{k/2+1, \ldots, k\}$ one has

$$\{\overline{\mathcal{N}}_{m_1} \geq \log k\} \cap \{\overline{\mathcal{N}}_{m_2} \geq \log k\} = \emptyset. \tag{6.20}$$

This ends our discussion on the motivation and main properties of the grand maximal counting function.

---

[28] In what follows we refer to $\mathrm{Basis}(\mathcal{F}_j^{\log k})$ as $\bigcup_{\mathcal{F}_j^{\log k, r}} \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{\log k, r})$ where $\mathcal{F}_j^{\log k, r}$ ranges through the decomposition of $\mathcal{F}_j^{\log k}$ into maximal USGTF's. For more details, consult Section 7.

[29] The appearance of the $\log \log k$ term is a technical artifact resulting from the construction of a **CME** in Section 7. For a further illuminating discussion on this topic see item 2) in Section 13.

7. The structure of the set $E := \{E(P)\}_{P \in \mathbb{P}}$. Definition of CME

In this section we will make a certain choice for N. This will not be done directly but through the structure that we impose on the set $\mathbb{P}$ of tiles. More precisely, as described above, we will run an algorithm for constructing a chain of multi-towers with some pre-scribed properties, this way giving rise to the concept of **CME**.

   We start with several general observations/heuristics:

- Our construction of the tile configurations will focus on the set[30]

$$\mathcal{F} \approx \bigcup_{j=2^{k/2}}^{2^k} \mathbb{P}_j.$$

   Later we will show that, as a consequence of our choice of the tile structure, the con-tribution of the tiles in $\overline{\mathbb{P}} \setminus \mathcal{F}$ to the $L^{1,\infty}$ "norm" of our operator T is small in an appropriate sense.
- Depending on the mass parameter, we will partition the set $\mathcal{F}$ into $k/2$ (dyadic) levels (preparing thus the future generations):

$$\mathcal{F} = \bigcup_{l=k/2+1}^{k} \mathcal{F}_l \quad \text{with} \quad \mathcal{F}_l \approx \bigcup_{j=2^{l-1}+1}^{2^l} \mathbb{P}_j.$$

- Our main task will be to design our set of tiles in such a way that each $\mathcal{F}_l$ is a multi-tower of generation (roughly) $(2^{l-1} + 1, 2^l)$ and of height $\log k$. This construction will be realized through an *inductive process* that will move downwards from $l = k$ to $l = k/2 + 1$.

We now present an *outline* of this process:

- At the first stage, we will design

$$\mathcal{F}_k = \bigcup_{l=1}^{\log k} \mathcal{F}_k^l$$

   so that $\mathcal{F}_k$ is a tower of generation (roughly) $(2^{k-1} + 1, 2^k)$ and of height $\log k$. For this we require that

---

[30] In reality the set $\mathcal{F}$ will have a more complicated structure; for example, $\mathcal{F}$ will also contain tiles from the families $\{\mathbb{P}_j\}_{j < 2^{k/2}}$. Indeed, during the construction we will express $\mathcal{F} = \bigcup_{j=k/2}^{k} \mathcal{F}_j$ with each tower $\mathcal{F}_j$ further decomposed as $\mathcal{F}_j = \bigcup_{l=1}^{\log k} \mathcal{F}_j^l$. While each of the families $\{\mathcal{F}_j^l\}_{l=1}^{\log k - 2}$ will only have tiles in $\bigcup_{j \geq 2^{k/2}} \mathbb{P}_j$, the remaining families $\mathcal{F}_j^{\log k - 1}$ and $\mathcal{F}_j^{\log k}$ will also contain tiles from $\bigcup_{j < 2^{k/2}} \mathbb{P}_j$. However, as described in Section 9, the tile set $\bigcup_{j < 2^{k/2}} \mathbb{P}_j$ will play a secondary role in the behavior of $\|Tf_k\|_{1,\infty}$.

- each $\mathcal{F}_k^l$ is a USGTF of generation[31] $\approx (2^{k-1} + 1, 2^k)$;
- the set of frequencies $\alpha(\mathcal{F}_k^l)$ sits entirely below and is largely separated from that corresponding to $\alpha(\mathcal{F}_k^{l+1})$;
- $\mathrm{ITop}(\mathcal{F}_k^{l+1}) \prec \mathrm{IBtm}(\mathcal{F}_k^l)$.

- In general, having constructed the (multi-)tower $\mathcal{F}_{j+1}$ of generation $(2^j + 1, 2^{j+1})$ and height $\log k$ we will divide it into maximal USGTF's $\{\mathcal{F}_{j+1,r}\}_r$ and within each $\mathcal{F}_{j+1,r}$ we will embed a specially designed multi-tower $\mathcal{F}_j[\mathcal{F}_{j+1,r}]$ of generation $(2^{j-1}+1, 2^j)$ and height $\log k$.
- We will repeat this algorithm until we exhaust the family $\mathcal{F}$ by reaching the level $j = k/2 + 1$.

Now, let us make the above description precise.

First stage: Constructing the tower $\mathcal{F}_k$. As mentioned before, we will split our family $\mathcal{F}_k$ into $\log k$ sets,

$$\mathcal{F}_k = \bigcup_{l=1}^{\log k} \mathcal{F}_k^l,$$

with each $\mathcal{F}_k^l$ being a USGTF of generation $(2^{k-1} + \log\log k, 2^k)$ (except for $\mathcal{F}_k^{\log k - 1}$ and $\mathcal{F}_k^{\log k}$, which are USGTF's of generation $(1, 2^k)$). Based on the description made in the previous section it will be enough to specify three parameters: the top, the bottom and the frequency set of each $\mathcal{F}_k^l$. We will proceed by induction.

Step 1: Defining $\mathcal{F}_k^1$. The key parameters of $\mathcal{F}_k^1$ are:

- the top

$$\mathrm{ITop}(\mathcal{F}_k^1) := \{[0,1]\},$$

- the frequency set

$$\alpha(\mathcal{F}_k^1) := \{2^{2^{2^{2^k}} + 100m}\}_{m \in \{0, 2^{k-1}-1\}},$$

- the bottom

$$\mathrm{IBtm}(\mathcal{F}_k^1) := \mathcal{I}_{2^{k-1} - \log\log k}([0,1]).$$

---

[31] In the actual construction process, for technical reasons, we will require that $\mathcal{F}_k^l$ be a USGTF of generation $(2^{k-1} + \log\log k, 2^k)$. The same observation applies to the other multi-towers at level j, i.e. the actual generation will be $(2^{j-1} + \log\log k, 2^j)$. The appearance of $\log\log k$ relies on the following loose statement: within the structure formed by the tiles at the bottom scale of each USGTF of generation $(2^{j-1} + \log\log k, 2^j)$ we can embed towers of generation $(2^{j-2} + \log\log k, 2^{j-1})$ and height precisely $\log k$, this being the height threshold that plays an important role in our proof. Another way of saying this is that the bottom structure—mass and number of tiles at the bottom—of a USGTF determines the height of a tower of a given (smaller) generation that can be embedded within it.

Step 2: Defining $\mathcal{F}_k^2$. The parameters of $\mathcal{F}_k^2$ are

$$\mathrm{ITop}(\mathcal{F}_k^2) := \mathcal{I}_{2^{k-1}-\log\log k}^{\mathrm{rt}}([0,1]),$$

$$\alpha(\mathcal{F}_k^2) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\{2^{2^k-1},\,2^{2^k}-1\}},$$

$$\mathrm{IBtm}(\mathcal{F}_k^2) := \mathcal{I}_{2^{k-1}-\log\log k}[\mathcal{I}_{2^{k-1}-\log\log k}^{\mathrm{rt}}([0,1])].$$

Step l: Defining $\mathcal{F}_k^l$ from $\mathcal{F}_k^{l-1}$ ($2 \le l \le \log k - 2$). Assume that $\mathcal{F}_k^{l-1}$ has

$$\mathrm{ITop}(\mathcal{F}_k^{l-1}) := \mathcal{I},$$

$$\alpha(\mathcal{F}_k^{l-1}) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\{(l-2)2^{2^k-1},\,(l-1)2^{2^k-1}-1\}},$$

$$\mathrm{IBtm}(\mathcal{F}_k^{l-1}) := \mathcal{I}_{2^{k-1}-\log\log k}[\mathcal{I}].$$

Then $\mathcal{F}_k^l$ is given by

$$\mathrm{ITop}(\mathcal{F}_k^l) := \mathcal{I}_{2^{k-1}-\log\log k}^{\mathrm{rt}}[\mathcal{I}],$$

$$\alpha(\mathcal{F}_k^l) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\{(l-1)2^{2^k-1},\,l\,2^{2^k-1}-1\}},$$

$$\mathrm{IBtm}(\mathcal{F}_k^l) := \mathcal{I}_{2^{k-1}-\log\log k}[\mathcal{I}_{2^{k-1}-\log\log k}^{\mathrm{rt}}[\mathcal{I}]].$$

Steps $\log k - 1$ and $\log k$: Defining $\mathcal{F}_k^{\log k-1}$ and $\mathcal{F}_k^{\log k}$. For the last two USGTF's we make some minor changes. We will require that both $\mathcal{F}_k^{\log k-1}$ and $\mathcal{F}_k^{\log k-1}$ be of generation $(1, 2^k)$, and assuming that we are given $\mathcal{F}_k^{\log k-2}$ we define $\mathcal{F}_k^{\log k-1}$ by setting

$$\mathrm{ITop}(\mathcal{F}_k^{\log k-1}) := \mathcal{I}_0^{\mathrm{rt}}[\mathrm{IBtm}(\mathcal{F}_k^{\log k-2})],$$

$$\alpha(\mathcal{F}_k^{\log k-1}) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\{(\log k-2)2^{2^k-1},\,(\log k-1)2^{2^k-1}-1\}},$$

$$\mathrm{IBtm}(\mathcal{F}_k^{\log k-1}) := \mathcal{I}_{2^k-1}[\mathrm{ITop}(\mathcal{F}_k^{\log k-1})].$$

Finally, the set $\mathcal{F}_k^{\log k}$ is given by

$$\mathrm{ITop}(\mathcal{F}_k^{\log k}) := \mathrm{ITop}(\mathcal{F}_k^{\log k-1}),$$

$$\alpha(\mathcal{F}_k^{\log k}) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\{(\log k-1)2^{2^k-1},\,\log k\,2^{2^k-1}-1\}},$$

$$\mathrm{IBtm}(\mathcal{F}_k^{\log k}) := \mathrm{IBtm}(\mathcal{F}_k^{\log k-1}).$$

This ends the process of defining the set $\mathcal{F}_k$.

Second stage: Constructing the family $\mathcal{F}_{k-1}$. We start with the following observation: the family $\mathcal{F}_{k-1}$ has $\log k$ disjoint components according to the information carried by

the graph of N within each of the previously constructed families $\{\mathcal{F}_k^l\}_{l \in \{1, \ldots, \log k\}}$. Thus we actually have

$$\mathcal{F}_{k-1} = \bigcup_{l=1}^{\log k} \mathcal{F}_{k-1}[\mathcal{F}_k^l].$$

For the particular case $l = \log k - 1$ and $l = \log k$ we have already determined $\mathcal{F}_{k-1}[\mathcal{F}_k^l]$, since the sets $\mathcal{F}_k^l$ are themselves completely determined (up to tiles of mass one) by the requirement that they be USGTF's of generation $(1, 2^k)$. (This is in contrast with the case $l < \log k - 1$ where we only require that $\mathcal{F}_k^l$ be a USGTF of generation $(2^{k-1} + \log \log k, 2^k)$.)

Thus it only remains to discuss the construction of the families $\{\mathcal{F}_{k-1}[\mathcal{F}_k^l]\}_{l=1}^{\log k-2}$. In our algorithm we will demand that each $\mathcal{F}_{k-1}[\mathcal{F}_k^l]$ be a multi-tower of generation $(2^{k-2} + \log \log k, 2^{k-1})$ and height $\log k$ embedded into $\mathcal{F}_k^l$.

In what follows, we will only detail the construction of $\mathcal{F}_{k-1}[\mathcal{F}_k^1]$ since the remaining multi-towers are constructed in the same way by adapting the reasonings for $\mathcal{F}_k^1$ to the case of $\mathcal{F}_k^l$.

Recall now the properties of $\mathcal{F}_k^1$:

$$\mathrm{ITop}(\mathcal{F}_k^1) := \{[0, 1]\},$$

$$\alpha(\mathcal{F}_k^1) := \{2^{2^{2^{2^k}} + 100m}\}_{m \in \{0, 2^{2^{k-1}} - 1\}},$$

$$\mathrm{IBtm}(\mathcal{F}_k^1) := \mathcal{I}_{2^{k-1} - \log \log k}([0, 1]).$$

Write now

$$\mathrm{IBtm}(\mathcal{F}_k^1) = \mathrm{I}^{\mathrm{lt}}\mathrm{Btm}(\mathcal{F}_k^1) \cup \mathrm{I}^{\mathrm{rt}}\mathrm{Btm}(\mathcal{F}_k^1),$$

and further express

$$\mathrm{I}^{\mathrm{lt}}\mathrm{Btm}(\mathcal{F}_k^1) = \{J_s\}_{s=1}^{2^{2^{k-1} - \log \log k}}.$$

(Recall here the index convention: $s_1 < s_2$ implies $c(J_{s_1}) < c(J_{s_2})$.)

Fix such an interval J_s and consider the set

$$\mathcal{I}_{\log \log k}(J_s) = \{I_r^s\}_{r=1}^{\log k}.$$

We then define the family $\mathcal{F}_{k-1}[\mathcal{F}_k^1][J_s]$ consisting of $\log k$ towers,

$$\mathcal{F}_{k-1}[\mathcal{F}_k^1][J_s] = \bigcup_{r=1}^{\log k} \mathcal{F}_{k-1}[\mathcal{F}_k^1][I_r^s].$$

Each tower can be decomposed as

$$\mathcal{F}_{k-1}[\mathcal{F}_k^1][I_r^s] = \bigcup_{l_1=1}^{\log k} \mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s],$$

with each $\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]$ a USGTF.

To see this, we first describe the maximal USGTF within each of the towers ($l_1 = 1$): $\mathcal{F}_{k-1}^1[\mathcal{F}_k^1][I_r^s]$ is a USGTF of generation $(2^{k-2} + \log \log k, 2^{k-1})$ with

$$\text{ITop}(\mathcal{F}_{k-1}^1[\mathcal{F}_k^1][I_r^s]) := \{I_r^s\},$$

$$\alpha(\mathcal{F}_{k-1}^1[\mathcal{F}_k^1][I_r^s])$$
$$:= \{2^{2^{2^{2^k}}+100m}\}_{m\in\{(s-1)(\log k\, 2^{2^{k-1}-1})+(r-1)(2^{2^{k-1}-1}),(s-1)(\log k\, 2^{2^{k-1}-1})+r2^{2^{k-1}-1}-1\}},$$

$$\text{IBtm}(\mathcal{F}_{k-1}^1[\mathcal{F}_k^1][I_r^s]) := \mathcal{I}_{2^{k-2}-\log\log k}(I_r^s).$$

Now, once we have established the base of each tower, the rest of the procedure should follow the lines of the tower construction described at Stage 1. For clarity we will specify the following:

Assume we have constructed $\mathcal{F}_{k-1}^{l_1-1}[\mathcal{F}_k^1][I_r^s]$. Then for $l_1 \le \log k - 2$, $\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]$ is a USGTF of generation $(2^{k-2} + \log\log k, 2^{k-1})$ with

$$\text{ITop}(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]) := \text{I}^{\text{rt}}\text{Btm}(\mathcal{F}_{k-1}^{l_1-1}[\mathcal{F}_k^1][I_r^s]),$$

$$\alpha(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\mathcal{A}_{k,r,s,l_1}},$$

$$\text{IBtm}(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]) := \mathcal{I}_{2^{k-2}-\log\log k}(\text{ITop}(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]))$$

where

$$\mathcal{A}_{k,r,s,l_1} := \bigcup_{m=(s-1)(\log k\, 2^{2^{k-1}-1})+(r-1+l_1-1)(2^{2^{k-1}-1})}^{(s-1)(\log k\, 2^{2^{k-1}-1})+(r+l_1-1)2^{2^{k-1}-1}-1} \{m\} \quad \text{if } r+l_1-1 \le \log k, \quad (7.1)$$

$$\mathcal{A}_{k,r,s,l_1} := \bigcup_{m=(s-1)(\log k\, 2^{2^{k-1}-1})+(r-1+l_1-1-\log k)(2^{2^{k-1}-1})}^{(s-1)(\log k\, 2^{2^{k-1}-1})+(r+l_1-1-\log k)2^{2^{k-1}-1}-1} \{m\} \quad \text{otherwise}. \quad (7.2)$$

The construction of $\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]$ with $l_1 \in \{\log k - 1, \log k\}$ follows a similar pattern with the following changes (see also the corresponding changes at the First Stage):

$\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s]$ is a USGTF of generation $(1, 2^{k-1})$ with

$$\text{ITop}(\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s]) := \text{I}^{\text{rt}}\text{Btm}(\mathcal{F}_{k-1}^{\log k-2}[\mathcal{F}_k^1][I_r^s]),$$

$$\alpha(\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s]) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\mathcal{A}_{k,r,s,\log k-1}},$$

$$\text{IBtm}(\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s]) := \mathcal{I}_{2^{k-1}-1}(\text{ITop}(\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s])),$$

where we preserve the definitions (7.1) and (7.2).

$\mathcal{F}_{k-1}^{\log k}[\mathcal{F}_k^1][I_r^s]$ is a USGTF of generation $(1, 2^{k-1})$ with

$$\text{ITop}(\mathcal{F}_{k-1}^{\log k}[\mathcal{F}_k^1][I_r^s]) := \text{ITop}(\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s]),$$

$$\alpha(\mathcal{F}_{k-1}^{\log k}[\mathcal{F}_k^1][I_r^s]) := \{2^{2^{2^{2^k}}+100m}\}_{m\in\mathcal{A}_{k,r,s,\log k}},$$

$$\text{IBtm}(\mathcal{F}_{k-1}^{\log k}[\mathcal{F}_k^1][I_r^s]) := \text{IBtm}(\mathcal{F}_{k-1}^{\log k-1}[\mathcal{F}_k^1][I_r^s]).$$

Observation 7.1. Notice the following key property of our construction:

$$\alpha(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s]) \cap \alpha(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_{r'}^s]) = \emptyset \quad \forall r \neq r' \text{ and } \forall l_1.$$

Moreover, for each r, s, the sets $\{\alpha(\mathcal{F}_{k-1}^{l_1}[\mathcal{F}_k^1][I_r^s])\}_{l_1=1}^{\log k}$ form a partition of the frequency set $\{2^{2^{2^{2^k}}+100m}\}_{m\in\{(s-1)(\log k\, 2^{2^{k-1}-1}),s\log k\, 2^{2^{k-1}-1}-1\}}$.

This ends the process of defining $\mathcal{F}_k$ and $\mathcal{F}_{k-1}$. We now repeat this algorithm and further construct by induction $\mathcal{F}_{k-2}, \ldots, \mathcal{F}_{k/2+1}$.

Third stage: Constructing a generic tower $\mathcal{F}_j$, $k/2 + 1 \leq j \leq k - 2$. Assume we have constructed the multi-tower $\mathcal{F}_{j+1}$ of height $\log k$. We first write as before its layer decomposition

$$\mathcal{F}_{j+1} = \bigcup_{l=1}^{\log k} \mathcal{F}_{j+1}^l.$$

For the sake of clarity, we mention here the process of obtaining $\{\mathcal{F}_{j+1}^l\}_l$. Thus, $\mathcal{F}_{j+1}^1$ consists of the union of maximal USGTF's of generation $(2^j + \log \log k, 2^{j+1})$,

$$\mathcal{F}_{j+1}^1 = \bigcup_m \mathcal{F}_{j+1}^{1,m},$$

such that

- $\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_{j+1}^{1,m})$ is maximal with respect to inclusion among all the sets $\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{A})$ with $\mathcal{A}$ a maximal USGTF inside $\mathcal{F}_{j+1}$;
- the sets $\{\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_{j+1}^{1,m})\}_m$ are pairwise disjoint.

Erase now $\mathcal{F}_{j+1}^1$ from $\mathcal{F}_{j+1}$ and repeat the above algorithm to obtain $\mathcal{F}_{j+1}^2$. Continue this process inductively. (Notice that once we reach $l = \log k - 1$, the generation of maximal USGTF's in the decomposition of $\mathcal{F}_{j+1}^l$ changes to $(1, 2^{j+1})$.) From our construction this process will end in precisely $\log k$ steps.

With this done, fix a family $\mathcal{F}_{j+1}^l$ (here we assume $l \leq \log k - 2$, otherwise trivial considerations), and, with the previous notation, write

$$\mathcal{F}_{j+1}^l = \bigcup_m \mathcal{F}_{j+1}^{l,m}.$$

Now taking $\mathrm{I}^{\mathrm{lt}}\mathrm{Btm}(\mathcal{F}_{j+1}^{l,m}) = \{J_s\}_s$, and then

$$\mathcal{I}_{\log\log k}(J_s) = \{I_r^s\}_{r=1}^{\log k},$$

we can initiate the same algorithm as in the case of the construction of $\mathcal{F}_{k-1}$ above. Adapting the description made at the second stage we have: given J_s, we design precisely $\log k$ towers, each tower being

- of generation $(2^{j-1} + \log \log k, 2^j)$ and height $\log k$;
- embedded in $\mathcal{F}_{j+1}^{l,m}$;
- with basis equal to the corresponding I_r^s in the partition of J_s.

To make things clear, for each J_s, we will construct the family $\mathcal{F}_j[\mathcal{F}_{j+1}^{l,m}][J_s]$ consisting of $\log k$ towers,

$$\mathcal{F}_j[\mathcal{F}_{j+1}^{l,m}][J_s] = \bigcup_{r=1}^{\log k} \mathcal{F}_j[\mathcal{F}_{j+1}^{l,m}][I_r^s].$$

Again as in the case of $\mathcal{F}_{k-1}$, we have

$$\mathcal{F}_j[\mathcal{F}_{j+1}^{l,m}][I_r^s] = \bigcup_{l_1=1}^{\log k} \mathcal{F}_j^{l_1}[\mathcal{F}_{j+1}^{l,m}][I_r^s],$$

with $\mathcal{F}_j^{l_1}[\mathcal{F}_{j+1}^{l,m}][I_r^s]$ a USGTF of generation $(2^{j-1} + \log\log k, 2^j)$ (excepting the cases $l_1 = \log k - 1$ and $l_1 = \log k$).

For expository reasons we will no longer give further details of this construction, since one follows the same steps (with obvious changes) as in the second stage of our construction.

In this way, by repeating the above algorithms, our construction process ends by specifying the multi-towers

$$\{\mathcal{F}_j\}_{j \in \{k/2+1,\ldots,k\}}.$$

Lastly, just to specify the set of tiles $P \in \mathbb{P}_n$ with $n \leq 2^{k/2}$ (though this information is not in any way *essential* for our later reasonings), we slightly modify the structure of the last constructed family $\mathcal{F}_{k/2+1}$ by requiring that all the maximal USGTF's that sit inside are of generation $(1, 2^{k/2+1})$.

Definition 7.2. We say that $\mathcal{F} \subset \mathbb{P}$ is a (lacunary) *Cantor Multi-tower Embedding* (**CME**) if

$$\mathcal{F} = \bigcup_{l=k/2+1}^{k} \mathcal{F}_j, \tag{7.3}$$

with $\mathcal{F}_j$ constructed as above.

Notice that as a byproduct of the above construction process we find that Theorem 1.14 holds.

Observation 7.3. One could relax many of the requirements in Definition 7.2 above. For example, the lacunary structure of the frequencies of each $\mathcal{F}_j$ as well as the dyadic splitting of the mass parameter (when forming the generations) are just an expression of the particular operator considered in this paper, i.e. the lacunary Carleson operator. Also, the precise number of multi-towers (in this case $k/2$) is irrelevant in general and should be adapted to the nature of the problem under discussion. Thus, the **CME** structure can easily be adapted to the study of pointwise convergence of the full sequence of partial Fourier sums near L^1.

However, the key property that should be present in every variation on the theme generated by Definition 7.2 is that a **CME** structure is required to have a tile configuration that maximizes the $L^{1,\infty}$ norm of a grand maximal counting function similar to (6.3).

8. Construction of the set(s) F_j

In this section we will focus on defining the sets $\{F_j\}_{j=k/2+1}^k$ appearing in the definition of f. Each set F_j will be constructed independently. Its structure will be completely determined solely by the *normal* component of the family $\mathcal{F}_j$, as given below:

Definition 8.1. Let $j \in \{k/2+1, k\}$ and $\mathcal{F}_j$ be constructed as before. Partition

$$\mathcal{F}_j = \bigcup_r \mathcal{F}_{j,r}$$

into maximal USGTF's. Set now

$$\mathcal{F}_{j,r}^{\mathrm{nm}} := \{P \in \mathcal{F}_{j,r} \mid I_{P^*} \cap (\mathbb{R} \setminus \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_{j,r})) = \emptyset\}. \tag{8.1}$$

We then define the *normal component* of $\mathcal{F}_j$ as

$$\mathcal{F}_j^{\mathrm{nm}} := \bigcup_r \mathcal{F}_{j,r}^{\mathrm{nm}}, \tag{8.2}$$

with the *boundary component* of $\mathcal{F}_j$ given by

$$\mathcal{F}_j^{\mathrm{bd}} := \mathcal{F}_j \setminus \mathcal{F}_j^{\mathrm{nm}}. \tag{8.3}$$

8.1. Construction of F_k

As mentioned previously, we will only work with the collection $\mathcal{F}_k$.

Recall now the following key property:

$$\mathrm{I}\mathrm{Top}(\mathcal{F}_k^l) = \mathrm{I}^{\mathrm{rt}}\mathrm{Btm}(\mathcal{F}_k^{l-1}) \quad \forall 2 \le l \le \log k - 1. \tag{8.4}$$

As a consequence,

$$|\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_k^l)| = \tfrac{1}{2}|\tilde{\mathrm{I}}\mathrm{Btm}(\mathcal{F}_k^{l-1})| \quad \forall 2 \le l \le \log k - 1. \tag{8.5}$$

We next specify the *size* of F_k and its *approximate location*. We impose on F_k the following conditions:

- the total measure of F_k is $|F_k| \approx 2^{-\log k \, 2^{2^k}} \cdot 2^{-k} \cdot \frac{1}{k}$;
- $F_k \subset \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_k^{\log k - 1}) = \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_k^{\log k})$;
- for any $P \in \mathcal{F}_k^{\log k - 1} \cup \mathcal{F}_k^{\log k}$ with $I_P \in \mathrm{IBtm}(\mathcal{F}_k^{\log k - 1}) \cup \mathrm{IBtm}(\mathcal{F}_k^{\log k})$ one has

$$\frac{|I_P \cap F_k|}{|I_P|} \approx 2^{\log k}|F_k|. \tag{8.6}$$

It remains now to specify the *concrete location* of F_k inside each of the intervals I_P mentioned in (8.6).

First we notice that from our construction, $\text{IBtm}(\mathcal{F}_k^{\log k-1}) = \text{IBtm}(\mathcal{F}_k^{\log k})$, and $\text{IBtm}(\mathcal{F}_k^{\log k})$ consists of pairwise disjoint intervals. Given this, it will be enough to specify the structure of F_k inside a given $I \in \text{IBtm}(\mathcal{F}_k^{\log k})$.

Now, fixing $I \in \text{IBtm}(\mathcal{F}_k^{\log k})$, our set F_k inside I will be determined by

- the frequencies of tiles $P \in \mathcal{F}_k^{\text{nm}}$;
- the sign on I of the functions $e^{-2\pi i l(\omega_P)} \cdot T_P^*(\chi_{[0,1]})(\cdot)$.

Notice that given the choice of our frequencies for the tiles inside $\mathbb{P}$, i.e.

$$\{2^{2^{2^{2^k}}+100m}\}_{m \in \{0,\ldots,\log k\, 2^{2^k-1}-1\}},$$

we know that for any $P \in \mathcal{F}_k$, T_P^* is highly oscillatory on I_P^*, or equivalently the oscillation period of $e^{2\pi i l(\omega_P)\cdot}$ is much smaller than $|I_P|$.

Now the process of constructing $I \cap F_k$ will follow a fractal pattern: we will start with the tiles at the lowest frequency and as we move up the frequency scale we will trim more and more from the possible location(s) of $I \cap F_k$ inside I. (A detailed construction is only done for the $j = k$ case; with some small adaptations, everything can be repeated for a general F_j).

Here are three key observations derived from Observation 3.1 (see properties (3.6), (3.7)) and Observation 4.14:

- the frequencies $\{e^{2\pi i(2^{2^{2^k}}+100m)\cdot}\}_{m \in \{0,\ldots,\log k\, 2^{2^k-1}-1\}}$ are oscillating *independently*;
- for each $P \in \mathcal{F}_k^{\text{nm}}$, $e^{-2\pi i l(\omega_P)\cdot} T_P^*(\chi_{[0,1]})(\cdot)$ keeps the same sign[32] on I, and moreover due to the smoothing effect of the convolution is morally constant on I;   (8.7)
- if $P, P' \in \mathcal{F}_k^{\text{nm}}$ ($P \neq P'$) are such that $l(\omega_P) = l(\omega_{P'})$ then

$$\text{supp}\, T_P^* \cap \text{supp}\, T_{P'}^* = \emptyset.$$

We start by focussing on the first level of our (multi-)tower, $\mathcal{F}_k^1$. In this USGTF we have 2^{2^k-1} frequencies.

Making use of observations (8.7) above, we see that the following function is well defined:

$$\mathcal{S}[I, \mathcal{F}_k^1] : \alpha(\mathcal{F}_k^1) \to \{-1, 1\},$$

---

[32] This is because in definition (3.3) the function ψ_k is smooth, odd with $x\psi_k(x) \geq 0$ for any $x \in \mathbb{R}$ and $\chi_{[0,1]} \geq 0$.

where $\mathcal{S}[I, \mathcal{F}_k^1](a) = 1$ if

$$T_P^* \equiv 0 \text{ on } I \quad \forall P \in \mathcal{F}_k^1 \cap \mathcal{F}_k^{\mathrm{nm}} \text{ with } l(\omega_P) = a, \tag{8.8}$$

and[33]

$$\mathcal{S}[I, \mathcal{F}_k^1](a) := \mathrm{sgn}\left(\int_I e^{-2\pi i a x} T_P^*(\chi_{[0,1]})(x)\, dx \right) \tag{8.9}$$

if there exists (see the observation below) a tile $P \in \mathcal{F}_k^1 \cap \mathcal{F}_k^{\mathrm{nm}}$ such that

$$l(\omega_P) = a \quad \text{and} \quad T_P^*(\chi_{[0,1]}) \text{ is not } 0 \text{ a.e. on } I. \tag{8.10}$$

Observation 8.2. Under the assumptions made in Observation 3.1, if a tile P obeying (8.10) exists, then it is unique; however, this last fact is not actually essential in defining $\mathcal{S}[I, \mathcal{F}_k^1]$ since the sign of the expression defined in the RHS of (8.9) remains the same for all the tiles $P \in \mathcal{F}_k^1 \cap \mathcal{F}_k^{\mathrm{nm}}$ obeying (8.10).

Next, for $a \in \alpha(\mathcal{F}_k^1)$, we let

$$\mathcal{U}_{\mathcal{F}_k^1}[I](a) := \{ x \in I \mid \mathrm{sgn}(\mathrm{Re}(e^{2\pi i a x})) = \mathcal{S}[I, \mathcal{F}_k^1](a) \}. \tag{8.11}$$

Note that $\mathcal{U}_{\mathcal{F}_k^1}[I](a)$ is a union of disjoint, equidistant, same-size dyadic intervals. Moreover,

$$a, b \in \alpha(\mathcal{F}_k^1) \text{ with } a < b \implies |\mathcal{U}_{\mathcal{F}_k^1}[I](a) \cap \mathcal{U}_{\mathcal{F}_k^1}[I](b)| = \tfrac{1}{2} |\mathcal{U}_{\mathcal{F}_k^1}[I](a)|. \tag{8.12}$$

We now simply impose the following requirement on F_k:

$$I \cap F_k \subseteq \bigcap_{a \in \alpha(\mathcal{F}_k^1)} \mathcal{U}_{\mathcal{F}_k^1}[I](a). \tag{8.13}$$

The ends the process of restricting the location of F_k relative to the first level $\mathcal{F}_k^1$.

Naturally, the same idea is further extended inside each of the higher level USGTF's $\{\mathcal{F}_k^l\}_{l \leq \log k}$. At the end of this inductive process, putting together all the levels, one concludes

$$I \cap F_k \subseteq \bigcap_{l=1}^{\log k} \bigcap_{a \in \alpha(\mathcal{F}_k^l)} \mathcal{U}_{\mathcal{F}_k^l}[I](a), \tag{8.14}$$

where $\mathcal{U}_{\mathcal{F}_k^l}[I](a)$ for $a \in \alpha(\mathcal{F}_k^l)$ designates the obvious notational extension from the case $l = 1$.

Let us write

$$\bigcup_m U_m(I)$$

for the decomposition of $\bigcap_{l=1}^{\log k} \bigcap_{a \in \alpha(\mathcal{F}_k^l)} \mathcal{U}_{\mathcal{F}_k^l}[I](a)$ into maximal (disjoint) dyadic intervals.

---

[33] Notice that from (3.3) and (8.10) the expression to which the sign function is applied in (8.9) is always *real*.

Notice that from (8.12) and (8.14) one has

$$\left|\bigcup_m U_m(I)\right|/|I| \approx 2^{-\log k \,(2^{2^k-1})}. \tag{8.15}$$

Also, if $a_0 = a_0(I, k) \in \alpha(\mathcal{F}_k)$ stands for the minimal $l(\omega_P)$ with $P \in \mathcal{F}_k$ and $I_P \supseteq I$ then let

$$\mathcal{U}_{F_k}[I](a_0) = \bigcup_l R_l \tag{8.16}$$

be the decomposition of $\mathcal{U}_{F_k}[I](a_0)$ into maximal disjoint dyadic intervals (as one can easily notice, for $j = k$ we actually have $a_0 = 2^{2^{2^k}}$). Notice that the intervals $\{R_l\}_l$ have the same length and are equidistant inside I. Moreover

$$|\mathcal{U}_{F_k}[I](a_0)| = |I|/2. \tag{8.17}$$

Then, we can finally determine F_k (up to small perturbations within each $U_m(I)$) by requiring that for any $I \in \mathrm{IBtm}(\mathcal{F}_k^{\log k})$ the following hold:

- the set F_k obeys (8.14), or equivalently

$$I \cap F_k \subset \bigcup_m U_m(I); \tag{8.18}$$

- the set $I \cap F_k$ is equidistributed inside $\{U_m(I)\}_m$, i.e.

$$|U_m(I) \cap F_k| = |U_{m'}(I) \cap F_k| \quad \forall m, m'.$$

- for each $R_l \subset I$ one has

$$|R_l \cap F_k|/|R_l| \approx 2^{\log k}|F_k|; \tag{8.19}$$

as a consequence, from (8.16), (8.17) and (8.19), one immediately has

$$|I \cap F_k|/|I| \approx 2^{-k} \cdot 2^{-\log k 2^{2^k}} \approx 2^{\log k}|F_k|; \tag{8.20}$$

- inside each $U_m(I)$ the set $U_m(I) \cap F_k$ consists of a single dyadic interval called from now on $U_m(I, F_k)$; its position is irrelevant for our purposes but for clarity pick this dyadic interval so that its left end-point is the center of $U_m(I)$.

This ends the construction of F_k.

8.2. Construction of an arbitrary F_j

In the general situation, we start by decomposing the multi-tower $\mathcal{F}_j$ into maximal towers:

- first we apply the layer decomposition $\mathcal{F}_j = \bigcup_{l=1}^{\log k} \mathcal{F}_j^l$;
- then we decompose each $\mathcal{F}_j^l$ into maximal USGTF's, $\mathcal{F}_j^l = \bigcup_m \mathcal{F}_j^{l,m}$;
- finally we form the maximal chains (of length $\log k$) which are precisely the maximal towers we were looking for.

For each such maximal tower we repeat the main steps from the $j = k$ case. Since this process is pretty straightforward, we will not give further details. Putting together the set specifications relative to each maximal tower, at the end of the day, we will have constructed the set F_j.

8.3. Consequences of the construction

In this section we analyze how the presence of a **CME** structure within $\mathbb{P}$ and the adapted construction of $\{F_j\}_j$ reflects on the properties of our operator T (or T^*).

In the next section, we will prove

Proposition 8.3 (heuristic). *The main component of our operator T relative to the $\|\cdot\|_{1,\infty}$ norm is given by*

$$T_M(f_k) := \sum_{k/2 < j \le k} \sum_{P \in \mathcal{F}_j^{\text{nm}}} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}). \tag{8.21}$$

In what follows we want to present a glimpse of the methods that we will employ in order to prove our main result. As expected, the fundamental role in our reasonings will be played by

- the properties of our **CME** $\mathcal{F} = \bigcup_{j=k/2+1}^{k} \mathcal{F}_j$;
- the properties of the sets $\{F_j\}_{j=k/2+1}^{k}$.

As a consequence of the above items based on the dual formulation

$$\int T_M(f_k) = \int f_k T_M^*(1), \tag{8.22}$$

we make the following fundamental observation:

the real part of the function $f_k T_M^*(1)$ is a positive *function whose integral is bounded from below by* $\|f_k\|_{L \log \log L \log \log \log L}$.

This is the main reason for proving[34]

---

[34] It is worth noticing that in general one can show a much stronger estimate for the *full* operator C_{lac} since we know that in particular $\|Hf\|_1 \approx \|f\|_{L \log L}$ where H is the Hilbert transform. However, *relative to the L^1 norm*, our operator T_M becomes an error term due to the specific *choice* of our **CME**. Indeed, one can trivially modify the proof of Lemma 8.4 and find that in fact $\|T_M(f_k)\|_1 \approx \|f_k\|_{L \log \log L \log \log \log L} \approx \log k$. In this context, the main contribution is given by the operator representing the tiles at frequency 0, that is, the operator T_O (see the notation/definitions from the next section). Notice that T_O behaves as a variant of the maximal Hilbert transform.

Lemma 8.4 (L^1 blowup). *With the previous notation,*

$$\|T_M(f_k)\|_1 \gtrsim \log k. \tag{8.23}$$

Proof. Based on the construction of our sets F_j and of our family of tiles, we have

Claim. *We have*

$$\int \operatorname{Re}(\chi_{F_j} T_P^*(1))(\cdot) \geq \frac{1}{500} \int \chi_{F_j}(\cdot) \left| \int \psi_P(\cdot - y) \chi_{E(P)}(y)\, dy \right|$$

$$= \frac{1}{500} \int |\chi_{F_j} T_P^*(1)(\cdot)| \tag{8.24}$$

for all[35] $P \in \mathcal{F}_j^{\mathrm{nm}} \setminus (\mathcal{F}_j^{\log k - 1} \cup \mathcal{F}_j^{\log k})$.

Proof of claim. We will appeal repeatedly to the fundamental relations (8.8)–(8.15). Assume without loss of generality that

$$P \in \mathcal{F}_j^{\mathrm{nm}} \cap \mathcal{F}_j^l \quad \text{for some } l \in \{1, \ldots, \log k - 2\}. \tag{8.25}$$

Now, from the construction of F_j we know that $F_j \subseteq \tilde{\mathrm{IBtm}}(\mathcal{F}_j^{\log k})$. As a consequence,

$$\int \operatorname{Re}(\chi_{F_j} T_P^*(1))(\cdot) = \sum_{I \in \tilde{\mathrm{IBtm}}(\mathcal{F}_j^{\log k})} \int_I \operatorname{Re}(\chi_{F_j} T_P^*(1))(\cdot). \tag{8.26}$$

Next, from our assumption (8.25), we have $\tilde{\mathrm{IBtm}}(\mathcal{F}_j^{\log k}) \cap I_P^* \neq \emptyset$, and if $I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k})$ with $I \cap I_P^* \neq \emptyset$ then $I \subseteq I_P^*$ and $|I| < 2^{-1000}|I_P|$ as long as k is large enough.

Thus, it is enough to restrict our attention to a fixed $I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k})$ with $I \subseteq I_P^*$. Define now

$$\mathcal{J}_P := \left\{ J \text{ dyadic} \mid J \subset I_P^*,\ |J| = \tfrac{1}{4} l(\omega_P)^{-1} \right\}. \tag{8.27}$$

Notice that from our construction of the **CME** $\mathcal{F}$ we know that $J \in \mathcal{J}_P$ implies $|J| \ll |I|$.

With the notation from Section 6 we define[36]

$$\mathcal{U}_P[I] := \bigcup_{\substack{J \in \mathcal{J}_P \\ J \subset I}} \{x \in J \mid \operatorname{sgn}(\operatorname{Re}(e^{2\pi i l(\omega_P)x})) = \mathcal{S}[J, \mathcal{F}_j^l](l(\omega_P))\}. \tag{8.28}$$

Observe that there exists a subcollection of intervals inside I and belonging to $\mathcal{J}_P$, denoted by $\mathcal{J}_P^+[I]$, such that

$$\mathcal{U}_P[I] = \bigcup_{J \in \mathcal{J}_P^+[I]} J. \tag{8.29}$$

---

[35] The condition $P \in \mathcal{F}_j^{\mathrm{nm}} \setminus (\mathcal{F}_j^{\log k - 1} \cup \mathcal{F}_j^{\log k})$ could be relaxed here at the expense of some extra technicalities.

[36] Notice that definition (8.9) can be extended from $I \in \mathrm{IBtm}(\mathcal{F}_j^l)$ to any (dyadic) interval $J \in \mathcal{J}_P$.

Moreover, as a consequence of our F_j construction,

$$I \cap F_j = \mathcal{U}_P[I] \cap F_j. \tag{8.30}$$

Now due to the choice of the distribution of our frequencies (recall (4.4)), any two distinct frequencies $l(\omega_1) > l(\omega_2)$ in our **CME** must obey $l(\omega_1) \geq 2^{10} l(\omega_2)$. Fix now $J \in \mathcal{J}_P^+[I]$. Then, from our construction of the set F_j we deduce the following uniformity-of-distribution condition:

$$|J_1 \cap F_j| = |J_2 \cap F_j| = 2^{-5}|J \cap F_j| \quad \forall J_1, J_2 \in \mathcal{I}_5(J). \tag{8.31}$$

This implies that for any $J \in \mathcal{J}_P^+[I]$,

$$\int_J \mathrm{Re}(\chi_{F_j} T_P^*(1))(\cdot) \geq \frac{1}{4} \int \chi_{F_j \cap \frac{1}{16}J}(\cdot) \left| \int \psi_P(\cdot - y)\chi_{E(P)}(y)\, dy \right|. \tag{8.32}$$

Thus, as a consequence of (8.32), (8.30), (8.8), (8.9) and (8.26),

$$\int \mathrm{Re}(\chi_{F_j} T_P^*(1))(\cdot) = \sum_{\substack{I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k}) \\ I \subseteq I_P^*}} \sum_{J \in \mathcal{J}_P^+[I]} \int_J |\mathrm{Re}(\chi_{F_j} T_P^*(1))(\cdot)|$$

$$\geq \sum_{\substack{I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k}) \\ I \subseteq I_P^*}} \sum_{J \in \mathcal{J}_P^+[I]} \frac{1}{4} \int \chi_{F_j \cap \frac{1}{16}J}(\cdot) \left| \int \psi_P(\cdot - y)\chi_{E(P)}(y)\, dy \right|$$

$$\geq \frac{1}{500} \int \chi_{F_j}(\cdot) \left| \int \psi_P(\cdot - y)\chi_{E(P)}(y)\, dy \right|. \tag{8.33}$$

This concludes the proof of our claim.

Now, from (8.24), we deduce that

$$\int \mathrm{Re}\left(\sum_{k/2 < j \leq k} \sum_{P \in \mathcal{F}_j^{\mathrm{nm}}} 2^{\log k \, 2^{2^j}} \chi_{F_j} T_P^*(1) \right) \approx \int \sum_{k/2 < j \leq k} \sum_{P \in \mathcal{F}_j^{\mathrm{nm}}} 2^{\log k \, 2^{2^j}} \chi_{F_j} |T_P^*(1)|. \tag{8.34}$$

Thus,

$$\left| \int T_M(f_k) \right| \gtrsim \sum_{k/2 < j \leq k} \sum_{P \in \mathcal{F}_j^{\mathrm{nm}}} 2^{\log k \, 2^{2^j}} |F_j \cap I_P^*| \frac{|E(P)|}{|I_P|}$$

$$= \sum_{k/2 < j \leq k} \sum_{1 \leq l \leq \log k} \sum_{r = 2^{j-1} + \log \log k}^{2^j} 2^{\log k \, 2^{2^j}} \sum_{P \in \mathcal{F}_j^{\mathrm{nm},l}[r]} \frac{|F_j \cap I_P^*|}{|I_P|} \frac{|E(P)|}{|I_P|} |I_P|$$

$$\gtrsim \sum_{k/2 < j \leq k} \sum_{1 \leq l \leq \log k} 2^{\log k \, 2^{2^j}} 2^j \frac{|F_j \cap \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^l)|}{|\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^l)|} |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^l)|$$

$$\gtrsim \log k \sum_{k/2 < j \leq k} 2^{\log k \, 2^{2^j}} 2^j |F_j| \approx \|f_k\|_{L \log \log L \log \log \log \log L},$$

where for the third relation we have used the fact that $|F_j \cap I_P^*|/|I_P| \approx |F_j \cap \tilde{I}\text{Top}(\mathcal{F}_j^l)|/$
$|\tilde{I}\text{Top}(\mathcal{F}_j^l)|$ for all $P \in \mathcal{F}_j^{\text{nm},l}[r]$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Our goal in this paper will be to show that the above lemma remains true if one replaces
in (8.23) the L^1 norm with an $L^{1,\infty}$ estimate.

9. Removing the small terms

As already announced, in this section we show that the contribution of the $(T - T_M)(f_k)$
component of the operator is small. More precisely, we will prove that

$$\|(T - T_M)(f_k)\|_{1,\infty} \lesssim \|f_k\|_{L \log \log L}. \tag{9.1}$$

For this, let us first introduce some notation. Set

$$T_O(f_k) := \sum_{P \in \mathbb{P}(0)} T_P(f_k), \qquad T_0(f_k) := \sum_{P \in \mathbb{P}_0} T_P(f_k),$$

$$T_{R,<}(f_k) := \sum_{k/2 < j \le k} \sum_{n < 2^{j-1} + \log \log k} \sum_{P \in \mathbb{P}_n \setminus \mathcal{F}_j} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}),$$

$$T_{R,>}(f_k) := \sum_{k/2 < j \le k} \sum_{n > 2^j} \sum_{P \in \mathbb{P}_n} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}),$$

$$T_R^{\text{bd}}(f_k) := \sum_{k/2 < j \le k} \sum_{P \in \mathcal{F}_j^{\text{bd}}} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}).$$

Also recall the formula

$$T_M(f_k) := \sum_{k/2 < j \le k} \sum_{P \in \mathcal{F}_j^{\text{nm}}} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}).$$

With this, we obviously have

$$T(f_k) = T_O(f_k) + T_0(f_k) + T_{R,<}(f_k) + T_R^{\text{bd}}(f_k) + T_M(f_k) + T_{R,>}(f_k). \tag{9.2}$$

We will show that

$$\|T_O(f_k)\|_{1,\infty} \lesssim \|f_k\|_{L^1}, \tag{9.3}$$

$$\|T_0(f_k)\|_{1,\infty} \lesssim \|f_k\|_{L^1}, \tag{9.4}$$

$$\|T_{R,>}(f_k)\|_1 \lesssim \|f_k\|_{L \log \log L}, \tag{9.5}$$

$$\|T_{R,<}(f_k)\|_1 \lesssim \|f_k\|_{L \log \log L}, \tag{9.6}$$

$$\|T_R^{\text{bd}}(f_k)\|_1 \lesssim \|f_k\|_{L \log \log L}. \tag{9.7}$$

The first relation is just a consequence of the simple geometric observation that the
tiles in $\mathbb{P}(0)$ form a tree—though in this case the tree concept must be slightly modified, by

replacing the first item in Definition 4.4 with the more relaxed requirement $100 P \leq P_0$. As a consequence, T_0 behaves essentially like the maximal Hilbert transform which we know to be bounded from L^1 to $L^{1,\infty}$.[37] Thus (9.3) holds.

Next, (9.4) is a direct consequence of [30, Theorem 1.1(b)].

For (9.6), we start by decomposing

$$T_{R,<}(f_k) = T_{R,<}^{\mathrm{nm}}(f_k) + T_{R,<}^{\mathrm{bd}}(f_k),$$

where

$$T_{R,<}^{\mathrm{nm}}(f_k) := \sum_{k/2 < j \leq k} \sum_{l < j} \sum_{P \in \mathcal{F}_l^{\mathrm{nm}} \setminus \mathcal{F}_j} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}).$$

From the construction of $\{F_j\}_j$ and $\mathcal{F}$ we have

$$T_{R,<}^{\mathrm{nm}}(f_k) = 0. \tag{9.8}$$

For the boundary term, we proceed as follows: First, fixing j, for any $n < 2^{j-1} + \log \log k$ we set

$$\mathcal{F}_{<j}^{\mathrm{bd}}[n] := \{P \in \mathbb{P}_n \mid P \notin \mathcal{F}_j \text{ and } I_{P*} \cap F_j \neq \emptyset\}, \quad \mathcal{J}_{<j}^{\mathrm{bd}}[n] := \{I_P \mid P \in \mathcal{F}_{<j}^{\mathrm{bd}}[n]\}.$$

Further, we define

$$\mathcal{F}_{<j}^{\mathrm{bd}} := \bigcup_{n < 2^{j-1} + \log \log k} \mathcal{F}_{<j}^{\mathrm{bd}}[n].$$

The key observation here is the following Carleson packing property[38]: for any $n < 2^{j-1} + \log \log k$ and $J \in \mathcal{J}_{<j}^{\mathrm{bd}}[n]$,

$$\sum_{\substack{I \in \bigcup_{r \leq n} \mathcal{J}_{<j}^{\mathrm{bd}}[r] \\ I \subset 1000 J}} |I| \lesssim |J|. \tag{9.9}$$

With this in mind, we have

$$\|T_{R,<}^{\mathrm{bd}}(f_k)\|_1 \leq \left\| \sum_{k/2 < j \leq k} \sum_{P \in \mathcal{F}_{<j}^{\mathrm{bd}}} T_P(2^{\log k \, 2^{2^j}} \chi_{F_j}) \right\|_1$$

$$\lesssim \sum_{k/2 < j \leq k} \sum_{P \in \mathcal{F}_{<j}^{\mathrm{bd}}} 2^{\log k \, 2^{2^j}} \frac{|I_{P*} \cap F_j|}{|I_P|} \frac{|E(P)|}{|I_P|} |I_P| \lesssim \sum_{k/2 < j \leq k} 2^{\log k \, 2^{2^j}} 2^{\log k} |F_j| \lesssim k \|f_k\|_1.$$

Thus (9.6) holds.

The proof of (9.7) is similar in spirit to that of (9.6); we omit the details.

---

[37] A stronger estimate that implies the $L^1 \to L^{1,\infty}$ bound for a tree is the content of [30, Lemma 6.2].

[38] The Carleson packing property can be viewed as a BMO-type condition arising naturally from the concept of a Carleson measure; for more details, see [6].

We pass now to (9.5). To prove this statement we need to essentially re-do one of the main components of the proof in [31]. Setting

$$T_{R,>}[j] := \sum_{n>2^j} \sum_{P \in \mathbb{P}_n} T_P(2^{\log k\, 2^{2^j}} \chi_{F_j}), \tag{9.10}$$

it is enough to show that

$$\|T_{R,>}[j]\|_1 \lesssim 2^{\log k\, 2^{2^j}} |F_j| 2^j, \tag{9.11}$$

since

$$T_{R,>}(f_k) = \sum_{k/2 < j \le k} T_{R,>}[j].$$

Proceeding as in [31], and using the properties of the set F_j (corresponding to the analogue of (8.18)), we follow the steps below:

- First, for some $g \in L^\infty(\mathbb{T})$ with $\|g\|_\infty = 1$, we write

$$\|T_{R,>}[j,2]\|_1 = 2^{\log k\, 2^{2^j}} \int \chi_{F_j} T^*_{R,>}[j,2](g)$$

  where

$$T^*_{R,>}[j,2](g) := \sum_{n \ge 2^j} \sum_{P \in \mathbb{P}_n} T^*_P(g).$$

- Next, for $f \in L^1(\mathbb{T})$ and $I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k})$, define

$$\mathcal{L}_I(f) := \frac{\int_I f}{|I|} \chi_I, \quad \mathcal{L}(f) := \sum_{I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k})} \mathcal{L}_I(f).$$

- Define the operator

$$T^*_{R,>}[\mathrm{av}, j, 2](g)(\cdot) := \sum_{n \ge 2^j} \sum_{P \in \mathbb{P}_n} e^{2\pi i l(\omega_P)\cdot} \mathcal{L}(T^*_{P0}(g))(\cdot), \tag{9.12}$$

  where P^0 stands for the shift of P to the real axis (zero frequency) and if $|I_P| = 2^{-k}$ then $T^*_{P0}(g)(x) = -\int \psi_k(x - y) \chi_{E(P)}(y) g(y)\, dy$ is the corresponding shift of the operator $T^*_P(g)$.
- We now have the key relation (see [31])

$$\|T^*_{R,>}[j,2](g)\|_{L^1(F_j)} \lesssim \|T^*_{R,>}[\mathrm{av}, j, 2](g)\|_{L^1(F_j)} + |F_j|. \tag{9.13}$$

- Next, we decompose $\mathbb{P}_n = \bigcup \mathcal{P}$ into maximal trees, and then, for each $I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k})$ and $n \geq 2^j$, we further deduce

$$\left| \int_{I \cap F_j} \sum_{\mathcal{P} \in \mathbb{P}_n} e^{2\pi i l(\omega_{\mathcal{P}}) \cdot} \mathcal{L}_I(T^{\mathcal{P}^{0*}}(g))(\cdot) \right|$$
$$\lesssim \|\mathcal{S}_{\mathbb{P}_n}^*(g)\|_{L^2(I)} \left\{ \sum_{\mathcal{P} \in \mathbb{P}_n} \frac{|\langle \chi_{F_j \cap I}, e^{2\pi i l(\omega_{\mathcal{P}}) \cdot} \rangle|^2}{|I|} \right\}^{1/2}, \quad (9.14)$$

where $\mathcal{S}_{\mathbb{P}_n}^*$ is the square function associated with $\mathbb{P}_n$, given by

$$\mathcal{S}_{\mathbb{P}_n}^* := \left\{ \sum_{\mathcal{P} \in \mathbb{P}_n} |T^{\mathcal{P}^*}|^2 \right\}^{1/2},$$

and we use the convention $T^{\mathcal{P}} := \sum_{P \in \mathcal{P}} T_P$ and $T^{\mathcal{P}^0}$ stands for the shift of $T^{\mathcal{P}}$ at the 0 frequency.
- Recall now Zygmund's inequality:

$$\left\| \sum_j a_j e^{i m_j x} \right\|_{\exp(L^2(\mathbb{T}))} \lesssim \left\{ \sum_j |a_j|^2 \right\}^{1/2} \quad (9.15)$$

for any lacunary sequence $\{m_j\}_j \subset \mathbb{N}$. Applying this (see also [11], [31] for a similar treatment) one has

$$\left\{ \sum_{\mathcal{P} \in \mathbb{P}_n} \frac{|\langle \chi_{F_j \cap I}, e^{2\pi i l(\omega_{\mathcal{P}}) \cdot} \rangle|^2}{|I|} \right\}^{1/2} \lesssim \frac{|I \cap F_j|}{|I|} \left(1 + \log \frac{|I|}{|I \cap F_j|} \right)^{1/2} |I|^{1/2}. \quad (9.16)$$

- Due to the *global* control of the L^2 norm of $\mathcal{S}_{\mathbb{P}_n}^*$ in terms of the mass parameter n we have

$$\left\| \sum_{I \in \mathrm{IBtm}(\mathcal{F}_j^{\log k})} \chi_I \mathcal{S}_{\mathbb{P}_n}^* \right\|_2 \lesssim 2^{-n/2}. \quad (9.17)$$

- Now, putting together the j analogue of (8.18) with (9.13), (9.14), (9.16) and (9.17), one concludes

$$\|T_{R,>}^*[j,2](g)\|_{L^1(F_j)} \lesssim |F_j| + |F_j| \sum_{n \geq 2^j} 2^{-n/2} k (\log k \, 2^{2^j})^{1/2} \lesssim k^2 |F_j|. \quad (9.18)$$

Thus, we have actually improved on (9.5), proving that in fact

$$\|T_{R,>}(f_k)\|_1 \lesssim \|f_k\|_{L(\log \log \log L)^2}. \quad (9.19)$$

10. Proof of Main Theorem 1

In this section we will prove the key result of our paper, encoded in the following

Theorem 10.1 ($L^{1,\infty}$ blowup). *With the previous notation,[39]*

$$\|T_M(f_k)\|_{1,\infty} \gtrsim \log k. \tag{10.1}$$

Proof. To prove (10.1) we need to show that there exists $G \subset [0,1]$ such that for every *major set* $G' \subset G$ (i.e. $|G'| \approx |G|$), one has

$$\left| \int_{G'} T_M(f_k) \right| \gtrsim \log k. \tag{10.2}$$

In order to match Theorem 10.1 with relation (1.11) we need to make a "wise" choice of G. Recall relations (9.3) and (9.4). In particular, choosing $G_0 = [0,1]$ we can find a major set $G \subseteq G_0$ such that $|G| \geq 1 - 100^{-1000}$ and

$$\int_G (|T_O(f_k)| + |T_0(f_k)|) \lesssim \|f_k\|_1. \tag{10.3}$$

Main Proposition. *Taking G as above, for any $G' \subseteq G$ with $|G'| \geq 1 - 10^{-1000}$ we have*

$$\left| \int f_k T_M^*(\chi_{G'}) \right| \gtrsim \log k. \tag{10.4}$$

If we assume that this holds, then, putting together (9.2)–(9.7), (10.3) and (10.4), we deduce Main Theorem 1, and hence Corollary 1.7.

Now we start the proof of the Main Proposition. First, from the previous discussion,

$$\left| \int f_k T_M^*(\chi_{G'}) \right| \geq \int f_k \, \mathrm{Re}(T_M^*(\chi_{G'})) \gtrsim \sum_{k/2 < j \leq k} 2^{\log k \, 2^{2^j}} \sum_{P \in \mathcal{F}_j} \int \chi_{F_j} |T_P^*(\chi_{G'})|.$$

(Strictly speaking in the definition of T_M we only have to deal with the normal components of $\mathcal{F}$, namely $\{\mathcal{F}_j^{\mathrm{nm}}\}_j$; however, since in the previous section, we proved that the term involving the boundary component is an error term, in the following estimates we will allow terms arising from considering the full family $\mathcal{F}$.)

With this observation we begin the analysis of the structure of the multi-tower $\mathcal{F}_j$. As usual, we first apply the layer decomposition

$$\mathcal{F}_j = \bigcup_{l=1}^{\log k} \mathcal{F}_j^l.$$

Next, we decompose each $\mathcal{F}_j^l$ into maximal USGTF's:

$$\mathcal{F}_j^l = \bigcup_m \mathcal{F}_j^{l,m}.$$

---

[39] Notice that as a consequence of Theorem 1, we actually have $\|T_M(f_k)\|_{1,\infty} \approx \log k$.

We then introduce the following notation:

$$E(\mathcal{F}_j^{l,m}) := \bigcup_{P \in \mathcal{F}_j^{l,m}} E(P), \quad G'(\mathcal{F}_j^{l,m}) := E(\mathcal{F}_j^{l,m}) \cap G'.$$

Also we set

$$T^{\mathcal{F}_j} := \sum_{l=1}^{\log k} T^{\mathcal{F}_j^l}, \quad T^{\mathcal{F}_j^l} := \sum_m T^{\mathcal{F}_j^{l,m}}, \quad \text{where} \quad T^{\mathcal{F}_j^{l,m}} := \sum_{P \in \mathcal{F}_j^{l,m}} T_P.$$

We first claim that if

$$|G'(\mathcal{F}_j^{l,m})| \geq \tfrac{1}{2}|E(\mathcal{F}_j^{l,m})|, \tag{10.5}$$

then there exists an absolute constant[40] $c > 10^{-3}$ such that

$$\left| \int \chi_{F_j} (T^{\mathcal{F}_j^{l,m}})^*(\chi_{G'}) \right| \geq c 2^j |F_j \cap \tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|. \tag{10.6}$$

This is a direct consequence of the special properties of the family $\mathcal{F}$ resulting from our construction.

Indeed, first notice that $\mathcal{F}_j^{l,m}[2^{j-1} + \log \log k]$ is precisely the bottom of the family $\mathcal{F}_j^{l,m}$ (i.e. the time intervals of these tiles form precisely the set $\mathrm{IBtm}(\mathcal{F}_j^{l,m})$).

Set now

$$A_0^{G'}(P) := |E(P) \cap G'|/|I_P|,$$

and let the *heavy* component of $\mathcal{F}_j^{l,m}[2^{j-1} + \log \log k]$ be

$$\mathcal{F}_j^{l,m}[2^{j-1} + \log \log k](H) := \left\{ P \in \mathcal{F}_j^{l,m}[2^{j-1} + \log \log k] \mid A_0^{G'}(P) \geq \tfrac{1}{4} A_0(P) \right\}.$$

Now, from (10.5), we must have

$$\#\mathcal{F}_j^{l,m}[2^{j-1} + \log \log k](H) \geq \tfrac{1}{4}\#\mathcal{F}_j^{l,m}[2^{j-1} + \log \log k]. \tag{10.7}$$

Observe now that from our construction, more precisely from the USGTF properties (4.13)–(4.16), we have: given $I \in \mathrm{ITop}(\mathcal{F}_j^{l,m})$, $a \in \alpha(\mathcal{F}_j^{l,m})$ and $s \in [2^{j-1} + \log \log k, 2^j]$ there exists a *unique* $P = [\omega_P, I_P] \in \mathcal{F}_j^{l,m}$ such that

- $I_P \subseteq I$;
- $l(\omega_P) = a$; \hfill (10.8)
- $A_0(P) \in [2^{-s}, 2^{-s+1})$.

Moreover, $E(P) = E(P')$ for any $P, P' \in \mathcal{F}_j^{l,m}$ with $I_P, I_{P'} \subseteq I$ and $l(\omega_P) = l(\omega_{P'}) = a$.

---

[40] The value of the constant c here is not relevant. The condition $c > 10^{-3}$, while not at all fundamental for the reasonings to follow, can be easily verified and is used only to give explicit quantitative bounds of the sort of (10.37). Alternatively, one can choose to work with unspecified constants c but then the statement of the Main Proposition above must be adjusted correspondingly.

Thus, if for each $I \in \mathrm{ITop}(\mathcal{F}_j^{l,m})$ and $a \in \alpha(\mathcal{F}_j^{l,m})$ we specify the information $E(\bar{P}) \cap G'$ carried by the unique tile $\bar{P} = [\omega_{\bar{P}}, I_{\bar{P}}] \in \mathcal{F}_j^{l,m}[2^{j-1} + \log\log k]$ with $l(\omega_{\bar{P}}) = a$ and $I_{\bar{P}} \subseteq I$, we will then have determined the entire structure of the family $\mathcal{F}_j^{l,m}$.

Define now

$$\mathcal{F}_j^{l,m}(H) := \{P \in \mathcal{F}_j^{l,m} \mid \exists P' \in \mathcal{F}_j^{l,m}[2^{j-1} + \log\log k](H) \text{ with } P' \leq P\},$$

$$\mathcal{F}_j^{l,m}[s](H) := \mathcal{F}_j^{l,m}(H) \cap \mathcal{F}_j^{l,m}[s].$$

Then from our previous observations (10.8) and the geometry of our tiles in $\mathcal{F}_j^{l,m}$ we find that

$$\tilde{\mathcal{N}}_j^{l,m}(H) := \frac{1}{2^{j-1}} \sum_{s=2^{j-1}+\log\log k}^{2^j} \sum_{P \in \mathcal{F}_j^{l,m}[s](H)} \frac{1}{2^{s-1}} \chi_{I_P}$$

obeys

$$\|\tilde{\mathcal{N}}_j^{l,m}(H)\|_1 \geq 10^{-2} |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|. \tag{10.9}$$

Since

$$\left| \int \chi_{F_j} (T^{\mathcal{F}_j^{l,m}})^* (\chi_{G'}) \right| \gtrsim \sum_{P \in \mathcal{F}_j^{l,m}(H)} \int \chi_{F_j} |T_P^* (\chi_{G'})|$$

and

$$\frac{|I_P^* \cap F_j|}{|I_P^*|} \approx \frac{|\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m}) \cap F_j|}{|\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|} \approx 2^l |F_j| \quad \forall P \in \mathcal{F}_j^{l,m} \cap \mathcal{F}_j^{\mathrm{nm}}, \tag{10.10}$$

we deduce that (10.9) implies (10.6).

Moreover, combining (10.6) with (10.10) and (2.2), we have

$$2^{\log k \, 2^{2^j}} \left| \int \chi_{F_j} (T^{\mathcal{F}_j^{l,m}})^* (\chi_{G'}) \right| \geq c \frac{1}{k} 2^l |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|. \tag{10.11}$$

Define now

$$\mathcal{H}_j^l := \bigcup_{\substack{m \\ \mathcal{F}_j^{l,m} \text{ satisfies } (10.5)}} \mathcal{F}_j^{l,m}, \tag{10.12}$$

$$\mathcal{H}_j := \bigcup_l \mathcal{H}_j^l, \tag{10.13}$$

$$\mathcal{H} := \bigcup_j \mathcal{H}_j. \tag{10.14}$$

Then, using (10.11), one deduces that

$$\int f_k \, \mathrm{Re}(T_M^*(\chi_{G'})) \gtrsim \frac{1}{k} \sum_{\mathcal{F}_j^{l,m} \in \mathcal{H}} 2^l |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})| + \sum_{\mathcal{F}_j^{l,m} \notin \mathcal{H}} 2^{\log k \, 2^{2^j}} \int \chi_{F_j} |(T^{\mathcal{F}_j^{l,m}})^*(\chi_{G'})|. \tag{10.15}$$

Take now a very large constant $C > 0$ ($C > 2^{100}$ is enough). To reach a contradiction assume that

$$\sum_{\mathcal{F}_j^{l,m} \in \mathcal{H}} 2^l |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})| < \frac{k \log k}{C}, \tag{10.16}$$

since otherwise we are done. Set

$$\mathcal{V}_j^l := \sum_{\mathcal{F}_j^{l,m} \in \mathcal{F}_j^l} 2^l |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|, \tag{10.17}$$

$$\mathcal{V}_j^l(H) := \sum_{\mathcal{F}_j^{l,m} \in \mathcal{H}_j^l} 2^l |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|, \tag{10.18}$$

$$\mathcal{V}_j^l(L) := \sum_{\mathcal{F}_j^{l,m} \in \mathcal{F}_j^l \setminus \mathcal{H}_j^l} 2^l |\tilde{\mathrm{I}}\mathrm{Top}(\mathcal{F}_j^{l,m})|. \tag{10.19}$$

Notice that

$$\mathcal{V}_j^l = \mathcal{V}_j^l(H) + \mathcal{V}_j^l(L).$$

On the one hand, from the construction of $\mathcal{F}$ we know that

$$\mathcal{V}_j^l \geq 2^{-50} \quad \forall j, l. \tag{10.20}$$

On the other hand, reformulating (10.16) we have

$$\sum_{j=k/2+1}^{k} \sum_{l=1}^{\log k} \mathcal{V}_j^l(H) < \frac{k \log k}{C}. \tag{10.21}$$

Let now

$$\mathcal{D} := \left\{ j \in \{k/2 + 1, \dots, k\} \mid \exists l, \ \mathcal{V}_j^l(L) \geq \mathcal{V}_j^l(H) \right\}. \tag{10.22}$$

Then, as a consequence of (10.20) and (10.16), we have

$$\#\mathcal{D} \geq k/100. \tag{10.23}$$

For $j \in \mathcal{D}$, let l_j be the smallest value of l that appears in (10.22). Set

$$E(j, l_j) := \bigcup_{\mathcal{F}_j^{l_j,m} \in \mathcal{F}_j^{l_j} \setminus \mathcal{H}_j^{l_j}} E(\mathcal{F}_j^{l_j,m}),$$

$$G'(j, l_j) := \bigcup_{\mathcal{F}_j^{l_j,m} \in \mathcal{F}_j^{l_j} \setminus \mathcal{H}_j^{l_j}} G'(\mathcal{F}_j^{l_j,m}),$$

$$\mathcal{A}_j := E(j, l_j) \setminus G'(j, l_j).$$

Also define

$$\mathcal{N}_j(x) := \frac{1}{2^{j-1} - \log \log k} \sum_{n=2^{j-1}+\log \log k+1}^{2^j} \frac{1}{2^{n-1}} \sum_{P \in \mathcal{F}_j[n]} \chi_{I_P}. \tag{10.24}$$

Then, for any $j \in \mathcal{D}$, we have

- $|\mathcal{A}_j| \geq \frac{1}{2}|E(j, l_j)| \geq \frac{1}{1000} \sum_{\mathcal{F}_j^{l_j,m} \in \mathcal{F}_j^{l_j}} |\tilde{\text{ITop}}(\mathcal{F}_j^{l_j,m})|;$
- $\mathcal{A}_j \subseteq \{x \mid \mathcal{N}_j(x) \geq l_j\};$ (10.25)
- $|\mathcal{A}_j| \geq \frac{1}{1000}|\{x \mid \mathcal{N}_j(x) \geq l_j\}|.$

Let us now turn our attention to the level sets of the function(s) $\mathcal{N}_j$.
For $l \in \{1, \ldots, \log k\}$ we decompose the set

$$C_j^l := \{x \mid \mathcal{N}_j(x) \geq l\} \tag{10.26}$$

into maximal disjoint (dyadic) intervals,

$$C_j^l = \bigcup_r C_j^l(r). \tag{10.27}$$

Now, from our construction of the **CME** $\mathcal{F}$, we have the following key properties:

- if $j_1 \geq j_2$ then for any pairs (l_1, r_1), (l_2, r_2),

$$\text{either} \quad C_{j_2}^{l_2}(r_2) \subset C_{j_1}^{l_1}(r_1) \quad \text{or} \quad C_{j_2}^{l_2}(r_2) \cap C_{j_1}^{l_1}(r_1) = \emptyset;$$

- if we set $C_{j_1}^{l_1}(r_1)[j_2, l_2] := \bigcup_{C_{j_2}^{l_2}(r_2) \subset C_{j_1}^{l_1}(r_1)} C_{j_2}^{l_2}(r_2)$ then we have the   (10.28)
John–Nirenberg type condition

$$|C_{j_1}^{l_1}(r_1)[j_2, l_2]| < 2^{-l_2+10}|C_{j_1}^{l_1}(r_1)|.$$

Indeed, properties (10.28) rely on the key observation that given any $C_j^l(r)$,

$$\exists \mathcal{F}_j^{l,m} \, \exists! I \in \text{ITop}(\mathcal{F}_j^{l,m}), \quad I = C_j^l(r). \tag{10.29}$$

Now the first item in (10.28) is a direct consequence of (10.29) and of the construction of $\mathcal{F}$ which requires

if $j_1 \geq j_2$ then for every $I_s \in \text{ITop}(\mathcal{F}_{j_s}^{l_{j_s}, m_s})$ with $s \in \{1, 2\}$ we have

either $I_1 \cap I_2 = \emptyset$ or $I_2 \subset I_1$.   (10.30)

The second item in (10.28) follows from (10.29), (10.30) and

if $l_1 \leq l_2$ then for all $k/2 < j \leq k$ and $I_s \in \text{ITop}(\mathcal{F}_j^{l_s, m_s})$ with $s \in \{1, 2\}$,

either $I_1 \cap I_2 = \emptyset$ or $I_2 \subset I_1$ with $|I_2| \leq 2^{10}2^{l_1-l_2}|I_1|$.   (10.31)

From (10.25) and (10.28), one deduces the following behavior of the sets $\{\mathcal{A}_j\}_{j \in \mathcal{D}}$:

$$2^{100}2^{-l_j} > |\mathcal{A}_j| > 2^{-100}2^{-l_j}, \quad |\mathcal{A}_{j_1} \cap \mathcal{A}_{j_2}| \leq 2^{100}|\mathcal{A}_{j_1}||\mathcal{A}_{j_2}|. \tag{10.32}$$

To see this, we first notice that with notations (10.26) and (10.27) and from (10.25) the first item in (10.32) trivially follows from

$$\mathcal{A}_j \subseteq C_j^{l_j} \quad \text{and} \quad |\mathcal{A}_j| \geq \tfrac{1}{1000}|C_j^{l_j}|, \tag{10.33}$$

$$2^{-55} \leq 2^{l_j}|C_j^{l_j}| \leq 2^{10}, \tag{10.34}$$

where for the last relation we have used (10.20) and the last item in (10.28) with $j_2 = j$, $j_1 = k, l_2 = l_j$ and $l_1 = 1$.

For the second item of (10.32) we use (10.33) and first notice that

$$|\mathcal{A}_{j_1} \cap \mathcal{A}_{j_2}| \leq |C_{j_1}^{l_{j_1}} \cap C_{j_2}^{l_{j_2}}|. \tag{10.35}$$

Now, assuming without loss of generality that $j_1 \geq j_2$ and making use of the last item in (10.28) we have

$$|C_{j_1}^{l_{j_1}} \cap C_{j_2}^{l_{j_2}}| \leq 2^{-l_{j_2}+10}|C_{j_1}^{l_1}| \leq 2^{70}|C_{j_1}^{l_{j_1}}||C_{j_2}^{l_{j_2}}| \leq 2^{100}|\mathcal{A}_{j_1}||\mathcal{A}_{j_2}|, \tag{10.36}$$

where we have again used (10.33) and (10.34).

At this point, notice that

$$|G \setminus G'| \geq \left|\bigcup_{j \in \mathcal{D}} \mathcal{A}_j\right| - 100^{-1000}. \tag{10.37}$$

Using now (10.37), (10.32), (10.23), (10.20) and the inclusion-exclusion principle we conclude that

$$|G \setminus G'| \geq 2^{-500}, \tag{10.38}$$

which contradicts the requirement that $|G'| \geq 1 - 10^{-1000}$, thus proving our proposition.

This concludes the proof of Theorem 10.1 and of Main Theorem 1. $\qquad\square$

11. Proof of Main Theorem 2

In this section we will give the proof of Main Theorem 2. The main ingredients that we will use are Theorem 1.3 and Main Theorem 1. To these, we will need to add: for (1)(i) the simple observation that

$$L \log \log L \log \log \log \log L \subset \mathcal{W}, \tag{11.1}$$

noticed in [9], and for (1)(iii) reasonings in the spirit of [7, proof of Theorem 2.3].

11.1. Proof of (1)(i)

In this case we just remark that

$$\Lambda_\varphi \subseteq \Lambda_{\varphi_0} = L \log \log L \log \log \log \log L,$$

and from (11.1) and Theorem 1.3 we deduce that Λ_φ is a $\mathcal{C}_L$-space.

11.2. Proof of (1)(ii)

Assume for contradiction that Λ_φ is a $\mathcal{C}_L$-space. Then there exists $C > 0$ such that

$$\|Tf\|_{L^{1,\infty}} \leq C\|f\|_{\Lambda_\varphi} \quad \forall f \in \Lambda_\varphi. \tag{11.2}$$

Take now $f = f_k$ as designed in the proof of Main Theorem 1. Further, let

$$f_k^* = \sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} \chi_{F_j^*} \tag{11.3}$$

be the decreasing rearrangement of f_k.

Recall the definition of the Λ_φ-norm:

$$\|f_k\|_{\Lambda_\varphi} = \int_0^1 f^*(t) \, d\varphi(t).$$

Using this and (11.3) we find that

$$\|f_k\|_{\Lambda_\varphi} = \sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} \int_{F_j^*} d\varphi \lesssim \sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} \varphi(|F_j|).$$

Next, we notice that $\varphi_0(|F_j|) \approx 2^{-\log k \, 2^{2^j}} (\log k)/k$. Thus,

$$\|f_k\|_{\Lambda_\varphi} \lesssim \frac{\log k}{k} \sum_{j=k/2+1}^{k} \frac{\varphi(|F_j|)}{\varphi_0(|F_j|)}. \tag{11.4}$$

At this point, from Main Theorem 1, we know that there exists $C' > 0$ such that

$$\|Tf_k\|_{L^{1,\infty}} \geq C' \log k. \tag{11.5}$$

Combining (11.4) and (11.5) with assumption (11.2), we deduce that

$$C_0 \leq \frac{1}{k} \sum_{j=k/2+1}^{k} \frac{\varphi(|F_j|)}{\varphi_0(|F_j|)}, \tag{11.6}$$

where is $C_0 > 0$ is an absolute constant.

Letting now $k \to \infty$, we notice that the hypothesis $\overline{\lim}_{s \to 0+} \varphi(s)/\varphi_0(s) = 0$ contradicts (11.5), thus proving that our assumption (11.2) cannot be true.

11.3. Proof of (1)(iii)

Assume now that

$$\underline{\lim_{s \to 0+}} \frac{\varphi(s)}{\varphi_0(s)} = 0 < \overline{\lim_{s \to 0+}} \frac{\varphi(s)}{\varphi_0(s)}. \tag{11.7}$$

We will first show that there exists φ obeying (11.7) such that

$$\Lambda_{\varphi_0} = L \log \log L \log \log \log \log L \subsetneq \Lambda_\varphi \subset \mathcal{W}, \tag{11.8}$$

and hence Λ_φ is a Lorentz $\mathcal{C}_L$-space strictly larger than Λ_{φ_0}.

Given the analogy between (11.8) and [7, Theorem 2.3], we will only outline the main steps of the proof (just a simple adaptation of the corresponding steps in [7]):

- Define $\mu : [0, 1] \to \mathbb{R}_+$ by $\mu(0) = 0$ and $\mu(t) := t \log \log(4/t)$ with $t \in (0, 1]$. Notice that

$$\|f\|_{\mathcal{W}} = \inf \left\{ \sum_{j=1}^{\infty} (1 + \log j) \|f_j\|_\infty \mu \left(\frac{\|f_j\|_1}{\|f_j\|_\infty} \right) \,\middle|\, \begin{matrix} f = \sum_{j=1}^{\infty} f_j, \\ \sum_{j=1}^{\infty} |f_j| < \infty \text{ a.e.}, \\ f_j \in L^\infty(\mathbb{T}) \end{matrix} \right\}.$$

- Preserving the notation in [7], for $s = \{s_n\}_{n \in \mathbb{N}}$ with each s_n in $[0, 1]$, define $\Lambda^{(s)}$ as the space of all measurable functions $f : \mathbb{T} \to \mathbb{C}$ such that there exists a sequence $\{f_n\}_{n \in \mathbb{N}}$ with $f_n \in L^\infty(\mathbb{T})$ satisfying $f = \sum_{n \in \mathbb{N}} f_n$ (with convergence in $L^1(\mathbb{T})$) and

$$\sum_{n=1}^{\infty} \max\{\|f_n\|_1, s_n \|f_n\|_\infty\} \frac{\mu(s_n)}{s_n} (1 + \log n) < \infty.$$

We endow $\Lambda^{(s)}$ with the norm

$$\|f\|_{\Lambda^{(s)}} := \inf \left\{ \sum_{n=1}^{\infty} \max\{\|f_n\|_1, s_n \|f_n\|_\infty\} \frac{\mu(s_n)}{s_n} (1 + \log n) \,\middle|\, f = \sum_{n \in \mathbb{N}} f_n \right\}. \tag{11.9}$$

- Using the simple observation $\mu(\alpha) \leq \max\{1, \alpha/\beta\} \mu(\beta)$ for any $\alpha, \beta \in (0, 1)$ one shows that for any sequence $s = \{s_n\}_{n \in \mathbb{N}} \in (0, 1]^{\mathbb{N}}$, $\Lambda^{(s)}$ is a r.i. Banach space such that

$$\Lambda^{(s)} \hookrightarrow \mathcal{W}, \tag{11.10}$$

with the inclusion norm ≤ 1.

- For $s \in (0, 1]^{\mathbb{N}}$ as before, let $\varphi^{(s)}$ be the quasi-concave function on $[0, 1]$ defined by $\varphi^{(s)}(0) = 0$ and

$$\varphi^{(s)}(t) = \inf_{n \in \mathbb{N}} \max\{t, s_n\} \frac{\mu(s_n)}{s_n} (1 + \log n), \quad \forall t \in (0, 1].$$

Then, for $\tilde{\varphi}^{(s)}$ the least concave majorant of $\varphi^{(s)}$, we have

$$\Lambda_{\tilde{\varphi}^{(s)}} \hookrightarrow \mathcal{W}, \tag{11.11}$$

with inclusion norm smaller than or equal to 1. To prove (11.11) one uses (11.10), (14.7) (the last observation in the Appendix) and the fact that the fundamental function of $\Lambda^{(s)}$ is precisely $\varphi^{(s)}$. This last fact is pretty straightforward and we leave its proof to the reader.

- If $s = \{s_n\}_{n\in\mathbb{N}} \in (0, 1]^{\mathbb{N}}$ is given by $s_n = 2^{-2^{2^n}}$, then

$$\Lambda_{\tilde{\varphi}^{(s)}} = L \log\log L \log\log\log\log L.$$

We can now end the proof of (11.8). Let $s = \{s_n\}_{n\in\mathbb{N}} \in (0, 1]^{\mathbb{N}}$ with $s_n = 2^{-2^{2^{2^n}}}$ and define

$$\varphi(t) = \min\{\varphi_0(t), \tilde{\varphi}^{(s)}\}. \tag{11.12}$$

Then clearly

$$\Lambda_{\varphi_0} \subseteq \Lambda_{\varphi} \subseteq \Lambda_{\tilde{\varphi}^{(s)}} + \Lambda_{\varphi_0} \subset \mathcal{W}.$$

In order to prove that

$$\Lambda_{\varphi_0} \subsetneq \Lambda_{\varphi}, \tag{11.13}$$

we just notice that

$$\varphi_0(s_n) \approx 2^{-2^{2^{2^n}}} 2^{2^n} n,$$

while

$$\varphi^{(s)}(s_n) \leq \mu(s_n)(1 + \log n) \approx 2^{-2^{2^{2^n}}} 2^{2^n} \log n.$$

We conclude that

$$\lim_{n\to\infty} \frac{\varphi(s_n)}{\varphi_0(s_n)} = 0,$$

thus showing (11.13) and ending the proof of (11.8).

We pass now to proving that there exists φ obeying (11.7) such that

$$\Lambda_{\varphi} \text{ is not a } \mathcal{C}_L\text{-space.} \tag{11.14}$$

For this, appealing again to Main Theorem 1, we notice that it will be enough to show that for a proper choice of φ,

$$\varlimsup_{k\to\infty} \frac{1}{\log k} \sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} \varphi(|F_j|) = 0. \tag{11.15}$$

For each $n \geq 100$, take $k_n = 2^{2^n}$ in the counterexample provided by Main Theorem 1 and define l_n to be the line passing through the points $A_{k_n/2+1} = (|F_{k_n/2+1}|, \varphi_0(|F_{k_n/2+1}|))$ and $A_{k_n} = (|F_{k_n}|, \varphi_0(|F_{k_n}|))$. Define now φ as follows:

$$\varphi(t) := \begin{cases} l_n(t) & \text{if } t \in [|F_{k_n/2+1}|, |F_{k_n}|], \ n \in \mathbb{N}, \ n \geq 100, \\ \varphi_0(t) & \text{otherwise.} \end{cases} \tag{11.16}$$

One can now easily check that φ satisfies (11.7). Moreover, for $k = k_n = 2^{2^n}$,

$$\sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} \varphi(|F_j|) = \sum_{j=k/2+1}^{k} 2^{\log k \, 2^{2^j}} l_n(|F_j|) \approx 2^{\log k \, 2^{2^j}} l_n(|F_j|)|_{j=k/2+1} \approx \frac{\log k}{k},$$

and thus (11.15) holds.

11.4. Proof of (2)(i)

Our goal here is to prove that if X is a r.i. Banach space with fundamental function $\varphi_X = \varphi$ obeying (1.12) then X is a $\mathcal{C}_L$-space.

Now, by (14.7), it will be enough to show that (1.12) implies $M_{\varphi_*} \subset \Lambda_{\varphi_0}$. But this last relation follows easily from the following:

- $f \in M_{\varphi_*}$ implies that there exists $C > 0$ such that

$$f^{**}(t) < C/\varphi(t) \quad \forall t \in (0, 1]. \tag{11.17}$$

- $\|f\|_{\Lambda_{\varphi_0}} = \int_0^1 f^*(t)\, d\varphi_0(t) \approx \int_0^1 f^{**}(t)[-t\varphi_0''(t)]\, dt.$

11.5. Proof of (2)(ii) and (iii)

With the notation of point (1) of Main Theorem 2, assume that

$$\|Tf_k\|_{1,\infty} \gtrsim \|f_k\|_V := \inf_{\sigma \in S_k} \sum_{j=1}^{k} r_j |F_j| 2^j \log(\sigma(j) + 1), \tag{11.18}$$

where S_k designates the group of all permutations of $\{1, \ldots, k\}$ (the proof of this statement will be outlined when proving Theorem 1.5 below).

Next, imposing the condition that $\{r_j\}_{1 \le j \le k}$ is an increasing sequence of positive numbers and using the definition of the Marcinkiewicz space M_{φ_*}, we find that

$$\|f_k\|_{M_{\varphi_*}} \approx \sup_{1 \le n \le k} \frac{\varphi(|F_n|)}{|F_n|} \sum_{j=n}^{k} r_j |F_j|. \tag{11.19}$$

Thus if we assume that (B) holds, taking $X = M_{\varphi_*}$, we deduce that there exists an absolute constant $C' > 0$ such that

$$\inf_{\sigma \in S_k} \sum_{j=1}^{k} r_j |F_j| 2^j \log(\sigma(j) + 1) \le C' \sup_{1 \le n \le k} \frac{\varphi(|F_n|)}{|F_n|} \sum_{j=n}^{k} r_j |F_j| \tag{11.20}$$

for any $k \in \mathbb{N}$ large enough.

If we choose $r_j := \frac{1}{|F_j| 2^j j}$, relation (11.20) becomes

$$\sup_{n \le k} \frac{\varphi(|F_n|)}{2^n |F_n| n} \ge C''(\log k)^2, \tag{11.21}$$

which further implies

$$\varlimsup_{s \to 0+} \frac{\varphi(s)}{\varphi_0(s)} = \infty, \tag{11.22}$$

thus proving (2)(ii).

Notice that for the choice of $r_j := \frac{1}{|F_j| 2^j j (\log j)^2}$ one can improve (11.22) to

$$\varlimsup_{s \to 0+} \frac{\varphi(s)}{\varphi_0(s) \log\log\log \frac{1}{s} \log\log\log\log \frac{1}{s}} = \infty. \tag{11.23}$$

Passing now to the proof of (2)(iii), assume without loss of generality that

$$s \mapsto \varphi_0(s)/\varphi(s) \text{ is increasing on } (0, 1), \tag{11.24}$$

i.e. $\epsilon = 1$ (otherwise trivial modifications are necessary).

In this setting it is enough to show that (B) implies (A).

Making the choice $r_j = 1/\varphi(|F_j|)$ and using (11.24) we obtain

$$\|f_k\|_{M_{\varphi_*}} \approx \sup_{1 \le n \le k} \frac{\varphi(|F_n|)}{|F_n|} \sum_{j=n}^{k} \frac{|F_j|}{\varphi(|F_j|)} \lesssim 1, \tag{11.25}$$

and

$$\|Tf_k\|_{1,\infty} \gtrsim \sum_{j=1}^{k/2} \frac{\varphi_0(|F_j|)}{\varphi(|F_j|)}. \tag{11.26}$$

Take now $|F_j|$ such that

$$\frac{|F_j|}{\varphi(|F_j|)} = \int_{y_{j+1}}^{y_j} \frac{s}{\varphi(s)} \left(-\frac{\varphi_0''(s)}{\varphi_0'(s)} \right) ds. \tag{11.27}$$

Then, from (11.25)–(11.27) and the requirement that M_{φ_*} is a $\mathcal{C}_L$-space we deduce that

$$\lim_{k \to \infty} \int_{y_{k/2}}^{1} \frac{-s\varphi_0''(s)}{\varphi(s)} ds \lesssim 1, \tag{11.28}$$

finishing the proof of (2)(iii).

11.6. Some remarks on the proof of Theorem 1.5

Notice that our claim follows if we are able to show (11.18). For this, though, all that we need is to notice that we can follow the reasonings in the proof of Main Theorem 1 and construct $F_j \subseteq \mathbb{T}$ measurable such that:

- the *size* of F_j obeys $|F_j| \in [y_{j+1}, y_j]$.
- the *location* of $\{F_j\}_j$ comes from the **CME** structure following the steps of Sections 7 and 8.

In this new setting, defining f_k as in (1.13) and proceeding as in Section 8, one obtains

$$\|Tf_k\|_{1,\infty} \gtrsim \inf_{\substack{\sum_{j=1}^{k} y_j \le 1/2 \\ y_j \ge 0}} \sum_{j=1}^{k} r_j |F_j| \log \log \frac{1}{|F_j|} \log \frac{1}{y_j} \approx \|f_k\|_V. \tag{11.29}$$

Thus in particular (11.18) holds.

It is worth mentioning that actually

$$C'\|f_k\|_{\mathcal{W}} \le \|f_k\|_V \le C''\|f_k\|_{\mathcal{W}} \tag{11.30}$$

for some absolute $C'' \ge C' > 0$.

12. A discussion regarding the (Lacey–Thiele) discretized Carleson model and the (discretized) Walsh model

This section explains in detail the comments in Observation 1.8. We will thus present a detailed antithesis between the a.e. pointwise convergence properties of the (lacunary) Carleson operator $C_{\mathrm{lac}}^{\{n_j\}_j}$ and those corresponding to the (lacunary) Walsh–Carleson operator $C_W^{\{n_j\}_j}$ and (more briefly) to the (lacunary) Lacey–Thiele discretized Carleson model.

In fact, in most of our analysis we will insist on analyzing the Walsh–Carleson operator, as the corresponding analysis of the Lacey–Thiele discretized Carleson model will then become immediately transparent.

We first list (a detailed explanation will follow) the key aspects that will make a difference in the behavior of $C_W^{\{n_j\}_j}$:

$$\begin{array}{ll} \text{(I) the } \textit{algebraic} \text{ properties of the Walsh wave-packets;} \\ \text{(II) the } \textit{discrete/dyadic} \text{ character of the Walsh–Carleson operator.} \end{array} \tag{12.1}$$

In order to better explain the above items, let us recall[41] some of the definitions/properties of the Walsh system/Walsh–Carleson operator.

We will only discuss the periodic setting, since this is the one of interest for us. Thus we define the Walsh phase plane $\mathbf{W} = [0, 1) \times [0, \infty)$. We fix the canonical dyadic grid (having origin at 0 and scales in powers of two) on both $[0, 1)$ and $[0, \infty)$. As before, a dyadic interval will be of the form $[2^{-j}n, 2^{-j}(n + 1))$ with $j \in \mathbb{N}$ (or $j \in \mathbb{Z}$ for positive real axis) and $n \in \mathbb{N}$. Keeping the notation from the Introduction, we refer to P as a tile if $P = I_P \times \omega_P \subset \mathbf{W}$ with I_P, ω_P dyadic intervals such that $|I_P| |\omega_P| = 1$, and we let $\mathbb{P}$ be the collection of all such tiles. Unlike the Fourier setting, here we will also need to work with *bitiles*, dyadic rectangles $R = I_R \times \omega_R \subset \mathbf{W}$ of area two (that is, I_R, ω_R are dyadic intervals such that $|I_R| |\omega_R| = 2$). We denote the collection of all bitiles by $\mathcal{R}$.

Next for $R \in \mathcal{R}$ with $I_R = [x_0, x_1]$ and $\omega_R = [\xi_0, \xi_1)$ we define

$$\begin{array}{ll} R_u := [x_0, x_1] \times [(\xi_0 + \xi_1)/2, \xi_1) \in \mathbb{P}, & \text{the } \textit{upper son} \text{ of } R; \\ R_l := [x_0, x_1] \times [\xi_0, (\xi_0 + \xi_1)/2) \in \mathbb{P}, & \text{the } \textit{lower son} \text{ of } R; \\ l^R := [x_0, (x_0 + x_1)/2) \times [\xi_0, \xi_1) \in \mathbb{P}, & \text{the } \textit{left son} \text{ of } R; \\ r^R := [(x_0 + x_1)/2, x_1) \times [\xi_0, \xi_1) \in \mathbb{P}, & \text{the } \textit{right son} \text{ of } R. \end{array} \tag{12.2}$$

Let us now recall the definition of the Walsh system.

Fix $n \in \mathbb{N}$ and let

$$n = \sum_{i=0}^{\infty} \epsilon_i 2^i \quad \text{with } \epsilon_i \in \{0, 1\} \tag{12.3}$$

be its dyadic decomposition. Let $\{r_j\}_{j \in \mathbb{N}}$ be the *Rademacher system*, that is, $r_j : [0, 1) \to \mathbb{R}$ with $r_0(x) = 1$ and $r_j(x) = \mathrm{sgn}(\sin 2^j \pi x)$ for $j \geq 1$ and $x \in [0, 1)$. Then we define the *Walsh system* $\{w_n\}_{n \in \mathbb{N}}$ as:

---

[41] For this we will closely follow [41].

- if $x \in \mathbb{R} \setminus [0, 1)$ then $w_n(x) = 0$ for any $n \in \mathbb{N}$;
- if $x \in [0, 1)$ and $n = 0$ then $w_n(x) = r_n(x) = 1$;
- if $x \in [0, 1)$ and $n \geq 1$ obeys (12.3) then[42]

$$w_n(x) := \prod_{i=0}^{\infty} (r_{i+1}(x))^{\epsilon_i} = \prod_{i=0}^{\infty} (\text{sgn}(\sin 2^{i+1}\pi x))^{\epsilon_i}. \tag{12.4}$$

Next, given $P \in \mathbb{P}$ with $P = [2^{-j}l, 2^{-j}(l+1)) \times [2^j n, 2^j (n+1))$ and $j, l, n \in \mathbb{N}$ we define the associated *Walsh wave-packet* as

$$w_P(x) = w_{n,l,j}(x) = 2^{j/2} w_n(2^j x - l). \tag{12.5}$$

Given $R \in \mathcal{R}$ and using now definitions (12.2), (12.4) and (12.5) we have the following key *algebraic* properties of the Walsh wave-packets:

$$w_{R_u} = \frac{1}{\sqrt{2}}(w_{l^R} - w_{r^R}), \tag{12.6}$$

$$w_{R_l} = \frac{1}{\sqrt{2}}(w_{l^R} + w_{r^R}). \tag{12.7}$$

Now, as a consequence of (12.6) and (12.7), the following key identity holds:[43]

$$W_n f(x) := \sum_{k=0}^{n} \langle f, w_k \rangle w_k(x) = \sum_{R \in \mathcal{R}} \langle f, w_{R_l} \rangle w_{R_l}(x) \chi_{\omega_{R_u}}(n) \tag{12.8}$$

for all $f \in L^1(\mathbb{T})$ and $n \in \mathbb{N}$.

Observation 12.1. We deduce from (12.8) that the (lacunary) Walsh–Carleson operator obeys

$$C_W^{\{n_j\}_j} f(x) = \sup_{j \in \mathbb{N}} \left| \sum_{R \in \mathcal{R}} \langle f, w_{R_l} \rangle w_{R_l}(x) \chi_{\omega_{R_u}}(n_j) \right|. \tag{12.9}$$

The meaning of item (II) is precisely that the Walsh–Carleson operator is of *discrete* nature, namely, it can be written as a superposition of projection operators associated with a *single* dyadic grid.

We end this discussion of the basic properties of the Walsh system by mentioning another relevant *algebraic* feature of it: as a consequence of (12.4),

$$\sum_{n=0}^{2^L-1} w_n(x) = \prod_{i=0}^{L-1} (r_0(x) + r_{i+1}(x)) \tag{12.10}$$

for all $L \in \mathbb{N}$. This implies that for any $f \in L^1(\mathbb{T})$,

$$W_{2^L-1}(f)(x) = \left\langle f(\cdot), \prod_{i=0}^{L-1} (r_0(x)r_0(\cdot) + r_{i+1}(x)r_{i+1}(\cdot)) \right\rangle. \tag{12.11}$$

---

[42] Notice that in (12.3) only finitely many ϵ_i's are non-zero.
[43] For a proof see [41].

One can reshape (12.11) for other subsequences of partial Fourier–Walsh sums or even differences of partial sums. For example, choosing $L \gg M$ with $L, M \in \mathbb{N}$ one has

$$
\begin{aligned}
& W_{2^L-1}(f)(x) - W_{2^L-2^M-1}(f)(x) \\
&\quad = \Big\langle f(\cdot), \Big(\prod_{i=M}^{L-1} r_{i+1}(x) r_{i+1}(\cdot) \Big) \Big(\prod_{i=0}^{M-1} (r_0(x) r_0(\cdot) + r_{i+1}(x) r_{i+1}(\cdot)) \Big) \Big\rangle.
\end{aligned}
\tag{12.12}
$$

Let us now switch our attention to the Carleson operator C (or its lacunary version $C_{\text{lac}}^{\{n_j\}_j}$). To make the analogy with the Walsh case more transparent, we will use a different, cleaner decomposition than the one in [28].[44]

Following the Walsh case development, we first define the real Fourier phase plane as $F_r = \mathbb{R} \times \mathbb{R}$ and the periodic Fourier phase plane[45] as $F_p = [-1, 2) \times \mathbb{R}$. Unlike the Walsh case, in order to connect C with its *discretized* model, we will need to use a continuum of grids. Let λ, μ be parameters ranging in $[0, 1)$. On two real axes we will use "dyadic" grids defined by the following structure:

- in space: $\mathcal{I}_{-m}^{-\lambda, y}$, the set of intervals $2^{-m-\lambda}[n + y, n + y + 1)$ with $m, n \in \mathbb{Z}$, $y \in [0, 1]$.
- in frequency: $\mathcal{J}_m^{\lambda, \mu}$, the set of intervals $2^{m+\lambda}[n + \mu, n + \mu + 1)$ with $m, n \in \mathbb{Z}$.

(12.13)

Define now

– the (y, λ, μ)-collection of tiles (real case)

$$
\mathbb{P}^{y,\lambda,\mu} := \bigcup_{m \in \mathbb{Z}} \mathbb{P}_m^{y,\lambda,\mu} := \bigcup_{m \in \mathbb{Z}} \mathcal{I}_{-m}^{-\lambda, y} \times \mathcal{J}_m^{\lambda, \mu};
\tag{12.14}
$$

– the $(y, \lambda, \mu, +)$-collection of tiles (periodic case)

$$
\mathbb{P}^{y,\lambda,\mu,+} := \bigcup_{m \in \mathbb{N}} \mathbb{P}_m^{y,\lambda,\mu,+} := \bigcup_{m \in \mathbb{N}} \mathcal{I}_{-m}^{-\lambda, y, +} \times \mathcal{J}_m^{\lambda, \mu},
\tag{12.15}
$$

where $\mathcal{I}_{-m}^{-\lambda, y, +} = \{ I \in \mathcal{I}_{-m}^{-\lambda, y} \mid I \subset [-1, 2) \}$.

We now define the Fourier wave-packets adapted to $\mathbb{P}^{y,\lambda,\mu}$ (with the obvious changes for the periodic case).

Let $\phi \in \mathcal{S}(\mathbb{R})$ with supp $\hat{\phi} \subseteq [-0.1, 0.1]$, $\hat{\phi} \geq 0$ and $\hat{\phi} \equiv 1$ on $[-0.07, 0.07]$.

Next, we introduce the classes of symmetries entering in the structure of the Carleson operator:

- translations: $T_z f(x) = f(x - z)$ with $x, z \in \mathbb{R}$;
- modulations: $M_\xi f(x) = e^{2\pi i x \xi} f(x)$ with $\xi \in \mathbb{R}$;
- dilations: $D_\lambda^p f(x) = \lambda^{-1/p} f(\lambda^{-1} x)$ with $\lambda > 0$ and $p \in (0, \infty]$.

---

[44] This simple, elegant decomposition of the real line Carleson operator appears in the junior paper [18] written under the supervision of E. Stein and guidance of the author.

[45] For technical reasons, in order to remove the boundary terms we enlarge the canonical interval $[0, 1)$ to $[-1, 2)$.

Let now P be a generic tile, that is, belonging to $\bar{\mathbb{P}} = \bigcup_{y,\lambda,\mu \in [0,1)} \mathbb{P}^{y,\lambda,\mu}$. If $P = I_P \times \omega_P$ (recall that P has area one) then $c(\omega_P)$ is the center of I_P, $\omega_{P_l} = (-\infty, c(\omega_P)] \cap \omega_P$ and $\omega_{P_u} = \omega_P \setminus \omega_{P_l}$, and as before define

$$P_u = I_P \times \omega_{P_u}, \quad \text{the \textit{upper son} of } P;$$
$$P_l = I_P \times \omega_{P_l}, \quad \text{the \textit{lower son} of } P.$$

We are now ready to define the wave-packet associated with P_l by

$$\phi_{P_l}(x) := M_{c(\omega_{P_l})} T_{c(I_P)} D_{|I_P|}^2 \phi(x), \tag{12.16}$$

or equivalently $\phi_{P_l}(x) = e^{2\pi i c(\omega_{P_l})x} |I_P|^{-1/2} \phi\left(\frac{x-c(I_P)}{|I_P|}\right)$.
  We now define

- the (y, λ, μ, ξ)-*discretized Carleson real model* as

$$\tilde{C}_\xi^{(y,\lambda,\mu)} f(x) := \sum_{P \in \mathbb{P}^{y,\lambda,\mu}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(\xi) \tag{12.17}$$

  for $x, \xi \in \mathbb{R}$ and $f \in L^1(\mathbb{R})$;
- the (y, λ, μ, ξ)-*discretized Carleson periodic model* as

$$\tilde{C}_\xi^{(y,\lambda,\mu)} f(x) := \sum_{P \in \mathbb{P}^{y,\lambda,\mu,+}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(\xi) \tag{12.18}$$

  for $x \in [0, 1)$, $\xi \in \mathbb{R}$ and $f \in L^1(\mathbb{R})$ with supp $f \subseteq [0, 1)$.

We have the following key result:

Proposition 12.2. *In what follows, $c \in \mathbb{R}$ is an absolute constant that is allowed to change from line to line. The following are true:*

- *For any $\mu \in [0, 1/4]$ we have for the real case*

$$\chi_{(-\infty,0]}(\xi) = c \int_0^1 \sum_{m \in \mathbb{Z}} 2^{m+\lambda} \sum_{P \in 2^{-m-\lambda}[0,1) \times \mathcal{J}_m^{\lambda,\mu}} |\hat{\phi}_{P_l}(\xi)|^2 \chi_{\omega_{P_u}}(0) \, d\lambda, \tag{12.19}$$

  and for the periodic case

$$\tilde{\chi}_{(-\infty,0]}(\xi) = c \int_0^1 \sum_{m \in \mathbb{N}} 2^{m+\lambda} \sum_{P \in 2^{-m-\lambda}[0,1) \times \mathcal{J}_m^{\lambda,\mu}} |\hat{\phi}_{P_l}(\xi)|^2 \chi_{\omega_{P_u}}(0) \, d\lambda, \tag{12.20}$$

  where $\tilde{\chi}_{(-\infty,0]}$ stands for a smooth version of $\chi_{(-\infty,0]}$, that is, $\tilde{\chi}_{(-\infty,0]} \in C^\infty(\mathbb{R})$ with supp $\tilde{\chi}_{(-\infty,0]} \subseteq (-\infty, 0]$ and $\tilde{\chi}_{(-\infty,0]}(\xi) = 1$ for $\xi \leq -1$.

- *If $N \in \mathbb{R}$ is fixed then for the real case*

$$\chi_{(-\infty, N]}(\xi) = c \int_0^1 \int_0^1 \sum_{m \in \mathbb{Z}} 2^{m+\lambda} \sum_{P \in 2^{-m-\lambda}[0,1) \times \mathcal{J}_m^{\lambda,\mu}} |\hat{\phi}_{P_l}(\xi)|^2 \chi_{\omega_{P_u}}(N) \, d\lambda \, d\mu,$$

(12.21)

while for the periodic case

$$\tilde{\chi}_{(-\infty, N]}(\xi) = c \int_0^1 \int_0^1 \sum_{m \in \mathbb{N}} 2^{m+\lambda} \sum_{P \in 2^{-m-\lambda}[0,1) \times \mathcal{J}_m^{\lambda,\mu}} |\hat{\phi}_{P_l}(\xi)|^2 \chi_{\omega_{P_u}}(N) \, d\lambda \, d\mu,$$

(12.22)

where as before $\tilde{\chi}_{(-\infty, N]}$ stands for a smooth version of $\chi_{(-\infty, N]}$.
- *For $f \in \mathcal{S}(\mathbb{R})$, define the real axis Carleson operator*

$$\mathcal{C} f(x) := \sup_{N \in \mathbb{Z}} |\mathcal{C}_N f(x)| = \sup_{N \in \mathbb{Z}} \left| \text{p.v.} \int_{\mathbb{R}} e^{2\pi i N(x-y)} \frac{1}{x-y} f(y) \, dy \right|.$$

(12.23)

Further, for $N \in \mathbb{Z}$, set

$$\tilde{\mathcal{C}}_N f(x) := \int_{-\infty}^{N} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi,$$

(12.24)

and notice that there exist absolute constants $c_1, c_2 \in \mathbb{R}$ such that

$$\mathcal{C}_N f = c_1 \tilde{\mathcal{C}}_N f + c_2 f.$$

(12.25)

As a consequence, one can reduce studying $\mathcal{C}$ to the study of

$$\tilde{\mathcal{C}} f(x) := \sup_{N \in \mathbb{Z}} |\tilde{\mathcal{C}}_N f(x)| := \sup_{N \in \mathbb{Z}} \left| \int_{-\infty}^{N} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi \right|.$$

(12.26)

Now, for $f, g \in \mathcal{S}(\mathbb{R})$, one has

$$\langle \tilde{\mathcal{C}}_N f, g \rangle = c \int_0^1 \int_0^1 \int_0^1 \sum_{P \in \mathbb{P}^{y,\lambda,\mu}} \langle f, \phi_{P_l} \rangle \langle \phi_{P_l}, g \rangle \chi_{\omega_{P_u}}(N) \, d\lambda \, d\mu \, dy.$$

(12.27)

Thus, linearizing the supremum, we deduce

$$\tilde{\mathcal{C}} f(x) = c \int_0^1 \int_0^1 \int_0^1 \sum_{P \in \mathbb{P}^{y,\lambda,\mu}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(N(x)) \, d\lambda \, d\mu \, dy,$$

(12.28)

or equivalently

$$\tilde{\mathcal{C}} f(x) = c \int_0^1 \int_0^1 \int_0^1 \tilde{\mathcal{C}}_{N(x)}^{(y,\lambda,\mu)} f(x) \, d\lambda \, d\mu \, dy.$$

(12.29)

- *In what follows, we consider $f \in C_0^\infty(\mathbb{R})$ with supp $f \subseteq [0, 1)$ and $x \in [0, 1)$. Let us recall the definition of the periodic Carleson operator:*

$$Cf(x) := \sup_{N \in \mathbb{Z}} |C_N f(x)| = \sup_{N \in \mathbb{Z}} \left| \text{p.v.} \int_{\mathbb{T}} e^{2\pi i N(x-y)} \cot(\pi(x-y)) f(y) \, dy \right|. \quad (12.30)$$

Following the real case reasonings, for $N \in \mathbb{Z}$, one defines

$$\tilde{C}_N f(x) := \int_0^1 \int_0^1 \int_0^1 \sum_{P \in \mathbb{P}^{y,\lambda,\mu,+}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(N) \, d\lambda \, d\mu \, dy. \quad (12.31)$$

Then relation (12.25) is replaced by

$$C_N f = c\tilde{C}_N f + \mathcal{A}_N f \quad (12.32)$$

with

$$\sup_N |\mathcal{A}_N f| \leq Mf, \quad (12.33)$$

where M stands for the Hardy–Littlewood maximal operator. Thus, setting

$$\tilde{C} f(x) = \int_0^1 \int_0^1 \int_0^1 \sum_{P \in \mathbb{P}^{y,\lambda,\mu,+}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(N(x)) \, d\lambda \, d\mu \, dy, \quad (12.34)$$

or equivalently

$$\tilde{C} f(x) = \int_0^1 \int_0^1 \int_0^1 \tilde{C}_{N(x)}^{(y,\lambda,\mu)} f(x) \, d\lambda \, d\mu \, dy, \quad (12.35)$$

from (12.32) and (12.33) one finds that there exists $c \in \mathbb{R}$ such that

$$|Cf - c\tilde{C}f| \leq Mf. \quad (12.36)$$

Observation 12.3. From (12.35) and (12.36) we notice that[46]

$$\left| C^{\{n_j\}_j} f(x) - c \sup_j \left| \int_0^1 \int_0^1 \int_0^1 \sum_{P \in \mathbb{P}^{y,\lambda,\mu,+}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(n_j) \, d\lambda \, d\mu \, dy \right| \right|$$
$$\leq Mf(x) \quad (12.37)$$

whenever $f \in C_0^\infty(\mathbb{R})$ with supp $f \subseteq [0, 1)$ and $x \in [0, 1)$.

Comparing now (12.37) with (12.9), we deduce that *unlike* the Walsh–Carleson operator, the Carleson operator is obtained as a *continuum average* of discrete models of type (12.18). This will play a key role in explaining the differences between the a.e. pointwise behavior of the two operators.

This ends the prerequisites about the basic definitions, concepts and properties regarding $C_W^{\{n_j\}_j}$ and $C^{\{n_j\}_j}$.

---

[46] In what follows, for notational simplicity, we drop the subscript lac from $C_{\text{lac}}^{\{n_j\}_j}$.

We pass now to a detailed motivation of the statements made in Observation 1.8.

From the proof of our Main Theorem 1, it should by now be obvious that there is no fundamental distinction between the a.e. pointwise behavior of $C^{\{n_j\}_j}$ and that of $C_{AW}^{\{n_j\}_j}$. Indeed, we immediately have the following

Corollary 12.4 ("averaged" Walsh–Carleson model[47]). *Recall definition* (1.5) *of the* (*lacunary Fourier*) *Carleson operator. Given* $N \in \mathbb{N}$ *one can replace the Fourier mode* $e^{2\pi i N x}$ *with the corresponding Walsh mode* $w_N(x)$ *and so for a given sequence* $\{n_j\}_{j\in\mathbb{N}} \subset \mathbb{N}$ *one can define an averaged Walsh–Carleson model by*

$$C_{AW}^{\{n_j\}_j} f(x) := \sup_{j\in\mathbb{N}} \left| \int_{\mathbb{T}} w_{n_j}(x) w_{n_j}(-y) \cot(\pi(x - y)) f(y)\, dy \right|. \qquad (12.38)$$

(*Here* $\{w_{n_j}\}_j$ *are regarded as periodic functions on* $\mathbb{R}$.) *Then the conclusions of Main Theorem* 1, *Main Theorem* 2, *Corollary* 1.6 *and Corollary* 1.7 *remain true for* $C_{AW}^{\{n_j\}_j}$.

Indeed, one notices that Corollary 12.4 requires only trivial modifications. In fact, the entire **CME** structure of $\mathcal{F}$ and the construction of the sets $\{F_j\}_j$ are left untouched. The only required modifications are expressed in the actual form of the (lacunary) averaged Walsh–Carleson operator $C_{AW}^{\{n_j\}_j}$ and part of the intermediate estimates provided in the last two sections.

To complete our antithesis and thus fully address Observation 1.8, it remains to discuss why

- Theorem 1.4 *holds* for $C^{\{n_j\}_j}$;
- Theorem 1.4 *does not hold* for $C_W^{\{n_j\}_j}$ (and for the corresponding Lacey–Thiele discretized Carleson model). $\qquad (12.39)$

We start our analysis with an easy but important observation on the behavior of $C^{\{n_j\}_j}$ versus its discretized model(s):

Theorem 1.4 states that for any lacunary $\{n_j\}_j \subset \mathbb{N}$ we can find a sequence $\{f_k\}_k \subset L^1(\mathbb{T})$ such that (1.9) and (1.10) hold and for some absolute constant $C > 0$,

$$\|C^{\{n_j\}_j}(f_k)\|_{L^{1,\infty}} \geq C \|f_k\|_{L \log\log L \log\log\log\log L}. \qquad (12.40)$$

(Notice that Theorem 1.4 is proved as a consequence of the proof of Main Theorem 1 and of the first item in Section 13.)

Now in view of (12.37), we know that (12.40) can be rewritten as

$$\left\| \sup_j \left| \int_0^1 \int_0^1 \int_0^1 \tilde{C}_{n_j}^{(y,\lambda,\mu)}(f_k)\, d\lambda\, d\mu\, dy \right| \right\|_{1,\infty} \geq C' \|f_k\|_{L \log\log L \log\log\log\log L}. \qquad (12.41)$$

We now present

---

[47] In the initial version of this paper, the corollary addressing the Walsh–Carleson model (labeled then as Corollary 3) was stated incorrectly. We thank Michael Lacey and the referee for pointing this out. The subtle difficulties arising from the discretization of the Walsh–Carleson operator are now discussed in great detail in the present section with the corresponding implications being summarized in Observation 1.8.

Observation 12.5. While (12.41) holds, it is possible to find triples (y, λ, μ) for which the corresponding *un-averaged* relation does *not* hold. In fact, there exist triples (y, λ, μ) and $c > 0$ such that for any $f \in L^1(\mathbb{T})$ one has

$$\left\| \sup_j |\tilde{C}_{n_j}^{(y,\lambda,\mu)}(f)| \right\|_{1,\infty} \leq c \|f\|_{L^1(\mathbb{T})}. \tag{12.42}$$

Indeed, pick $\{n_j = 2^j\}_{j \in \mathbb{N}}$ and take a closer look at the discrete operators $\{\tilde{C}_{2^j}^{(y,0,0)}\}_{j \in \mathbb{N}}$. Given definitions (12.13) and (12.18) we deduce that

$$\tilde{C}_{2^j}^{(y,0,0)} f(x) = \sum_{P \in \mathbb{P}_{j+1}^{y,0,0,+}} \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(2^j). \tag{12.43}$$

This is a *single scale* operator, and we immediately deduce that for any $f \in L^1(\mathbb{T})$,

$$\sup_j |\tilde{C}_{2^j}^{(y,0,0)}(f)| \lesssim Mf(x), \tag{12.44}$$

and hence

$$\left\| \sup_j |\tilde{C}_{2^j}^{(y,0,0)}(f)| \right\|_{1,\infty} \lesssim \|f\|_{L^1(\mathbb{T})}, \tag{12.45}$$

explaining Observation 12.5.

Exactly the same argument can be carried out for the Walsh operator, giving

$$\left\| \sup_j |C_W^{\{2^j\}_j}(f)| \right\|_{1,\infty} \lesssim \|Mf\|_{1,\infty} \lesssim \|f\|_{L^1(\mathbb{T})}. \tag{12.46}$$

Observation 12.6. Thus, we have identified a *first* reason of why Theorem 1.4 does not hold for $C_W^{\{n_j\}_j}$: relying on (12.1)(I), we find that the discrete nature of the Walsh operator makes possible a "boundary effect" where a specific choice of $\{n_j\}_j$ interacts with the choice of the dyadic grid in the decomposition of the operator, allowing a single scale behavior at the frequencies defined by the sequence. In this way, the key geometric configurations of tiles defined by a **CME** *cannot* be present in the time-frequency decomposition of $C_W^{\{n_j\}_j}$ since this structure requires a superposition of *multiple scales*[48] at *each* frequency defined by—possibly a subsequence of—our lacunary sequence.

One can now ask the following natural

Question: *Is it true that the only way in which the Walsh analogue of* (12.40) *can fail is when the time-frequency decomposition of* $C_W^{\{n_j\}_j}$ *lacks **CME** structures?*

Answer: *No!*

---

[48] The number of such scales must tend to infinity as the parameter k in (12.41) goes to infinity.

Indeed, one can easily check that shifting the previous dyadic sequence, that is, defining $n_j := 2^j - 1$ ($j \in \mathbb{N}$), one has:

- the operator $C_W^{\{2^j-1\}_j}$ maps L^1 into $L^{1,\infty}$;
- the time-frequency decomposition of $C_W^{\{2^j-1\}_j}$ does admit **CME** structures.   (12.47)

To understand the general framework that allows the above answer (as well as the underlying mechanism behind the example provided by (12.47)) we need to elaborate on a subtle element that plays a key role in the proof of Main Theorem 1 and Theorem 1.4.

Preserving the usual notation and in particular identifying $C^{\{n_j\}_j}$ with the operator T used in the proof of Main Theorem 1, suppose we have constructed our **CME** $\mathcal{F} = \bigcup_{j=k/2+1}^{k} \mathcal{F}_j$ with $\mathcal{F}_j = \bigcup_{l=1}^{\log k} \mathcal{F}_j^l$. When next constructing the corresponding sets $\{F_j\}_{j=k/2+1}^{k}$, there are two important properties that we want each set F_j to fulfill:

- for each $P \in \mathcal{F}_j$ the function $\mathrm{Re}(\chi_{F_j} T_P^*(\chi_{[0,1]}))$ is positive;
- for each $P \in \mathcal{F}_j^l$ one has $|I_P^* \cap F_j|/|I_P| \approx |I_P \cap F_j|/|I_P| \approx 2^l |F_j|$.   (12.48)

Decompose now $\mathcal{F}_j^l = \bigcup_m \mathcal{F}_j^{l,m}$ into maximal USGTF's. Next, fix an $\mathcal{F}_j^{l,m}$ and choose $a \in \alpha(\mathcal{F}_j^{l,m})$; denote by $\mathcal{F}_j^{l,m}(a)$ all the tiles $P \in \mathcal{F}_j^{l,m}$ that live at frequency a. Recall that $\mathcal{F}_j^{l,m}[n]$ stands for those tiles $P \in \mathcal{F}_j^{l,m}$ having $A(P) = A_0(P) = 2^{-n}$ with $n \in \{2^{j-1} + \log\log k, 2^j\}$. Set now $\mathcal{F}_j^{l,m}[n](a) := \mathcal{F}_j^{l,m}[n] \cap \mathcal{F}_j^{l,m}(a)$. From our **CME** construction, $\mathcal{F}_j^{l,m}[n](a)$ has precisely one element; moreover, denoting by $P_j^{l,m}(a)$ the unique element in $\mathcal{F}_j^{l,m}(a)$ with $I_{P_j^{l,m}(a)} \in \mathrm{IBtm}(\mathcal{F}_j^l)$ one has

$$E(P) = E(P_j^{l,m}(a)) \quad \forall P \in \mathcal{F}_j^{l,m}(a). \tag{12.49}$$

As a consequence of (12.48), if $P \in \mathcal{F}_j^{l,m}[n](a)$ then

$$\mathrm{Re}\left(\int \chi_{F_j}(x) T_P^*(\chi_{[0,1]})(x)\, dx \right) \approx \int \chi_{F_j}(x) |T_P^*(\chi_{[0,1]})(x)|\, dx$$
$$\approx |I_P^* \cap F_j| \frac{|E(P)|}{|I_P|} \overset{(12.49)}{\approx} 2^l |F_j| \, |E(P_j^{l,m}(a))|. \tag{12.50}$$

It is now the moment to include the above relation into the following

Observation 12.7. The proof of our Theorem 1.4 involving the operator $T \approx C^{\{n_j\}_j}$ relies crucially on the fact that for each $P \in \mathcal{F}_j^{l,m}(a)$ the quantity $\mathrm{Re}(\int \chi_{F_j} T_P^*(\chi_{[0,1]}))$ has *the same sign and approximate size* given by $2^l |F_j| \, |E(P_j^{l,m}(a))|$, a quantity *independent* (at least for our construction of USGTF's) of the scale and of the mass of the tile P.

Hence one immediately deduces that

$$\left| \sum_{P \in \mathcal{F}_j^{l,m}(a)} \int \chi_{F_j} T_P^*(\chi_{[0,1]}) \right| \approx 2^j 2^l |F_j| \, |E(P_j^{l,m}(a))|. \tag{12.51}$$

The analogue of (12.51) for the Walsh case would read

$$
\Big| \sum_{\substack{R \in \mathcal{R} \\ R_u \in \mathcal{F}_j^{l,m}(a)}} \langle \chi_{F_j}, w_{R_l} \rangle \langle \chi_{[0,1]} w_a w_{R_l} \chi_{E(P_j^{l,m}(a))} \rangle \chi_{\omega_{R_u}}(a) \Big|
$$
$$
\approx 2^j 2^l |F_j| |E(P_j^{l,m}(a))|. \quad (12.52)
$$

However, the above relation *does not hold* for arbitrary values of $a \in \mathbb{N}$.

Indeed, taking $\{n_r := 2^r - 1\}_{r \in \mathbb{N}}$ and $a = a_0 = 2^r - 1$ (for some large $r \in \mathbb{N}$), we appeal to (12.6)–(12.8) and (12.12) to deduce that

$$
\Big| \sum_{\substack{R \in \mathcal{R} \\ R_u \in \mathcal{F}_j^{l,m}(a_0)}} \langle \chi_{F_j}, w_{R_l} \rangle \langle \chi_{[0,1]} w_{a_0} w_{R_l}, \chi_{E(P_j^{l,m}(a_0))} \rangle \chi_{\omega_{R_u}}(a_0) \Big|
$$
$$
= |\langle W_{a_0}(\chi_{F_j}) - W_{a_0 - 2^{2^{j-1} + \log \log k}}(\chi_{F_j}), \chi_{[0,1]} w_{a_0} \chi_{E(P_j^{l,m}(a_0))} \rangle|
$$
$$
\lesssim \sup_{P \in \mathcal{F}_j^{l,m}} \frac{\int_{I_P} \chi_{F_j}}{|I_P|} |E(P_j^{l,m}(a_0))| \lesssim 2^l |F_j| |E(P_j^{l,m}(a_0))|. \quad (12.53)
$$

This contradicts (12.51) for large enough $j \in \mathbb{N}$.

By simple modifications of (12.12), one can see that (12.53) continues to hold for many other choices of (frequencies within) lacunary sequences $\{n_r\}_r$ that have some suitable dyadic structure. Some examples are given by $\{n_r = 2^r + m\}_{r \in \mathbb{N}}$ (with $m \in \mathbb{N}$ fixed) or more generally any lacunary sequence $\{n_r\}_{r \in \mathbb{N}}$ having the property that there exists $p \in \mathbb{N}$ such that for any $r \in \mathbb{N}$ the dyadic expansion of n_r has at most p non-zero terms.[49] All these examples have the property that

$$
C_W^{\{n_r\}_r} \text{ maps } L^1 \text{ into } L^{1,\infty}, \quad (12.54)
$$

which has been known for some time (see e.g. [20]). Indeed, based on an interesting observation of Konyagin [20], one can find many other (classes of) sequences $\{n_j\}_j$ for which $C_W^{\{n_j\}_j}$ maps L^1 to $L^{1,\infty}$ boundedly. Exactly because of this peculiar behavior, in his ICM address, Konyagin posed the following

Open Problem (OP I) ([22]). *Find a necessary and sufficient condition on a sequence* $\{n_j\}_j \subseteq \mathbb{N}$ *for which the associated Walsh–Carleson operator*

$$
C_W^{\{n_j\}_j} \text{ maps } L^1 \text{ to } L^{1,\infty}. \quad (12.55)
$$

In view of the second item in (12.39) and of the discussion following it, it is natural to ask if (12.55) can only hold for rather "exceptional" (lacunary) sequences as a manifestation of a "boundary effect" due to the special *algebraic* and *dyadic/discrete* structures of the Walsh system. In this context, we raise the following

---

[49] Of course, in this latter case the constant appearing in the second to last inequality of (12.53) depends on p.

Open Problem (OP II). *Decide if the Walsh analogue of Main Theorem 1 holds, that is, if it is true that in* (1.11)*,* $C_{\text{lac}}^{\{n_j\}_j}$ *can be replaced by* $C_W^{\{n_j\}_j}$.

However, due to the dichotomy (12.52)–(12.53), we now better understand what type of obstacles we encounter when dealing with Open Problems I and II. For example, a reasonable strategy in approaching OP II relies on the following

Observation 12.8. In order for Main Theorem 1 to hold in the Walsh case (or equivalently, for our Open Problem II to be affirmatively decided) one needs to search for lacunary sequences $\{n_r\}_r$ for which

- there is a **CME** structure compatible with the time-frequency discretization of $C_W^{\{n_r\}_r}$;
- once we identify a **CME**, relation (12.52) must be satisfied.

Turning to the Fourier case, we continue with the following

Observation 12.9. Notice the following antithesis:

- It is possible to use a *single* (time-frequency) dyadic grid to decompose the Carleson operator. Indeed, (proceeding as in Section 3) one can write, for a generic operator $C_{\text{lac}}^{\{n_r\}_r}$, the following equality:

$$C_{\text{lac}}^{\{n_r\}_r} = \sum_{P \in \mathbb{P}} C_P. \qquad (12.56)$$

  The key fact is that in this case C_P are *not* projection operators of rank one, but convolution operators of the form imposed by (3.3). Due to this specific form, once we fix a certain frequency $N(x) = a$, all the adjoint operators C_P^* with $a \in \omega_P$ will oscillate at the *same* frequency.
- In contrast, one can use a decomposition of the Carleson operator involving *infinitely many* time-frequency "dyadic" grids. In the latter case, in the sense described in (12.34)–(12.37) one has

$$Cf(x) \approx \int_0^1 \int_0^1 \int_0^1 \tilde{C}_{N(x)}^{(y,\lambda,\mu)} f(x) \, d\lambda \, d\mu \, dy \qquad (12.57)$$

  with

$$\tilde{C}_\xi^{(y,\lambda,\mu)} = \sum_{P \in \mathbb{P}^{y,\lambda,\mu,+}} C_{P,\xi}, \qquad (12.58)$$

  where the $C_{P,\xi}$ are rank one projection operators defined by

$$C_{P,\xi} f = \langle f, \phi_{P_l} \rangle \phi_{P_l}(x) \chi_{\omega_{P_u}}(\xi). \qquad (12.59)$$

  Notice that unlike the previous case, if we now fix the frequency $N(x) = a$, the adjoint operators $\{C_{P,a}^*\}_{P \in \mathbb{P}^{y,\lambda,\mu,+}}$ will oscillate at pairwise *distinct* frequencies.

Now the entire proof of our Main Theorem 1, including the construction of a **CME**, is realised using a *single* dyadic time-frequency grid corresponding to the family of tiles $\mathbb{P} \approx \mathbb{P}^{0,0,+}$ and involving decomposition (12.56) as described in the first item above. Taking now such a **CME** $\mathcal{F}$ and using the standard notation as before, assume that $\max\{|I_P| \mid P \in \mathcal{F}_j^{l,m}(a)\} = 2^{-m_0}$ and $\min\{|I_P| \mid P \in \mathcal{F}_j^{l,m}(a)\} = 2^{-m_1}$. Then, following the same approach and in the same spirit as in (12.34)–(12.37) we obtain

$$
\sum_{P \in \mathcal{F}_j^{l,m}(a)} C_P(\chi_{F_j})(x) \approx \int_0^1 \int_0^1 \int_0^1 \sum_{\tilde{m}=m_0}^{m_1} \sum_{P \in \mathbb{P}_{\tilde{m}}^{y,\lambda,\mu,+}} \langle \chi_{F_j}, \phi_{P_l} \rangle
$$
$$
\times \phi_{P_l}(x) \chi_{E(P_j^{l,m}(a))}(x) e^{-2\pi i a x} \chi_{\omega_{P_u}}(a)\, d\lambda\, d\mu\, dy. \quad (12.60)
$$

Notice that while the RHS of (12.52) corresponds in the Fourier setting to

$$
\sum_{\substack{P \in \mathbb{P}^{0,0,0,+} \\ P \in \mathcal{F}_j^{l,m}(a)}} \langle \chi_{F_j}, \phi_{P_l} \rangle \langle \phi_{P_l}(\cdot) e^{-2\pi i a \cdot}, \chi_{E(P_j^{l,m}(a))}(\cdot) \chi_{[0,1]}(\cdot) \rangle \chi_{\omega_{P_u}}(a), \quad (12.61)
$$

after the *averaging process* we get, using (12.60) and (12.51),

$$
\sum_{P \in \mathcal{F}_j^{l,m}(a)} \int C_P(\chi_{F_j}) \chi_{[0,1]} \approx \int_0^1 \int_0^1 \int_0^1 \sum_{\tilde{m}=m_0}^{m_1} \sum_{P \in \mathbb{P}_{\tilde{m}}^{y,\lambda,\mu,+}} \langle \chi_{F_j}, \phi_{P_l} \rangle
$$
$$
\times \langle \phi_{P_l}(\cdot) e^{-2\pi i a \cdot}, \chi_{E(P_j^{l,m}(a))}(\cdot) \chi_{[0,1]}(\cdot) \rangle \chi_{\omega_{P_u}}(a)\, d\lambda\, d\mu\, dy
$$
$$
\approx 2^j 2^l |F_j| \, |E(P_j^{l,m}(a))|. \quad (12.62)
$$

Thus we can now conclude with the following

Observation 12.10. We now understand the *second* reason why Theorem 1.4 does not hold for $C_W^{\{n_j\}_j}$: while necessary, the existence[50] of **CME** structures is not sufficient for the Walsh analogue of (12.40). This is due to the *combined* effect of two facts:

- The **CME** structure has *different frequency implications* in the Fourier ("continuous") case and in the Walsh ("discrete") case. Indeed, as hinted at in Observation 12.9, in the first case, all the operators $\{C_P^*\}_{P \in \mathcal{F}_j^{l,m}(a)}$ oscillate at the *same* frequency a. In the second case, the operators $\{W_R^*\}_{R_u \in \mathcal{F}_j^{l,m}(a)}$ defined by $W_R f(x) := \langle f, w_{R_l} \rangle w_{R_l}(x) \chi_{\omega_{R_u}}|_{N(x)=a}$ will oscillate at pairwise *distinct* frequencies. The latter property is not due to the particular (algebraic) structure of the Walsh wave-packets, but is due to the *geometry* of the tile discretization: if $\tilde{\mathcal{R}}$ is any family of bitiles with the property that there exists $a \in \mathbb{R}$ such that $R \in \tilde{\mathcal{R}}$ implies $a \in \omega_{R_u}$, then any two tiles within the family $\{R_l\}_{R \in \tilde{\mathcal{R}}}$ are pairwise disjoint.

---

[50] Relative to the time-frequency decomposition of $C_W^{\{n_j\}_j}$.

- Based on the algebraic properties of the Walsh wave packets (such as (12.7)–(12.12)) one can find counterexamples to the Walsh analogue of (12.51). However, as described in Observation 12.7, relation (12.51) is of key importance in the proof of Theorem 1.4.

Finally, we notice that taking averages over discrete models makes the pointwise behavior of the lacunary Carleson operator $C^{\{n_j\}_j}$ independent of the structure of the lacunary sequence $\{n_j\}_j$.

13. Final remarks

1) We start with a discussion of Theorem 1.4. We will briefly present the relevant modifications in the proof of Main Theorem 1 that are required in order to deal with a general lacunary sequence. The key observation is that the mechanism involved in constructing the **CME** $\mathcal{F}$ is *independent* of the specific choice of the lacunary sequence, and the only thing that depends on $\{n_j\}_j$, and hence needs special care, is the construction of the sets $\{F_j\}_{j=k/2+1}^{k}$.[51] Indeed, it is natural to expect that the structure of the frequencies plays a role in the corresponding structure of the sets. In the case of a perfectly dyadic sequence, once we constructed the multi-tower $\mathcal{F} = \bigcup_{j=k/2+1}^{k} \mathcal{F}_j$ and decomposed each $\mathcal{F}_j = \bigcup_{l=1}^{\log k} \mathcal{F}_j^l$ and then each tower $\mathcal{F}_j^l$ into maximal USGTF's $\bigcup_m \mathcal{F}_j^{l,m}$, taking for simplicity $j = k$, we could arrange for (8.12) to hold for each $I \in \mathrm{IBtm}(\mathcal{F}_k^{\log k})$ and thus in turn we were able to construct F_k to obey (8.18). This further implied that

$$\begin{aligned} &\textit{for any } I \in \mathrm{IBtm}(\mathcal{F}_k^{\log k}) \textit{ the set } I \cap F_k \\ &\textit{has a fractal structure (Cantor set) of the same size.} \end{aligned} \tag{13.1}$$

For an arbitrary lacunary sequence $\{n_j\}_j$ we no longer aim at preserving the *exact* form of (13.1). Instead, we seek for a good *approximation* of (13.1), meaning that we want the sets $I \cap F_k$ to have fractal structure with "almost" the same (relative) size as I ranges through $\mathrm{IBtm}(\mathcal{F}_k^{\log k})$. The precise meaning of the above heuristic is explained in what follows:

Let $\{n_j\}_j$ be our favorite choice for a lacunary sequence. Fix as before $k = 2^{2^K}$ for some large $K \in \mathbb{N}$. We want to construct the sets $\{F_j\}_{j=k/2+1}^{k}$ such that the function

$$f_k := \frac{1}{k} \sum_{j=k/2+1}^{k} \frac{1}{|F_j| \log \log \frac{1}{|F_j|}} \chi_{F_j}$$

obeys (1.11).

Now since $C_{\mathrm{lac}}^{\{n_j\}_j}$ is a maximal operator and we are interested in a lower bound for its $L^{1,\infty}$ norm, we can always restrict our attention to any subsequence of the initial $\{n_j\}_j$. Taking $\bar{K} := 2^{2^{2^k}}$ and possibly passing to a subsequence, we can assume that $n_1 > 1$ and

$$n_{j+1}/n_j > 2^{\bar{K}} \quad \forall j \in \mathbb{N}. \tag{13.2}$$

---

[51] This should be integrated in the line of thought supporting Observation 7.3.

We now adapt (4.4) to our new context by requiring

$$\mathsf{Image}(N) \subseteq \{n_{\bar{K}+m}\}_{m\in\{0,\log k\, 2^{2k}-1\}}.\tag{13.3}$$

With this done, we follow line by line the construction of our **CME** $\mathcal{F}:=\bigcup_{l=k/2+1}^{k}\mathcal{F}_l$ in Section 7 with the only trivial change regarding the frequency locations of our tiles in each USGTF: whenever we see a frequency $\alpha(P) = 2^{2^{2^{2^k}}} + 100m$, we replace it with the corresponding analogue $\alpha(P) = n_{\bar{K}+m}$.

Having constructed our **CME** we now adapt the construction of the sets $\{F_j\}_{j=k/2+1}^{k}$ in Section 8.1 to the new setting. For simplicity and space constraints we only focus on the case $j = k$. As in Section 8.1, we fix $I \in \mathsf{IBtm}(\mathcal{F}_k^{\log k})$ and define $\mathcal{S}[I, \mathcal{F}_k^1]$ and $\mathcal{U}_{\mathcal{F}_k^1}[I](a)$ exactly as in (8.8)–(8.11).

Note that (8.12) ceases to remain true; however, based on the fact that for any $P \in \mathcal{F}$ (in particular $A_0(P) \leq 1/2$) one has $|I_P| > 2^{-2^{2k}}$ and hence $1/n_j \ll |I|$ for any $j \geq \bar{K}$, we deduce that

- for all $a \in \alpha(\mathcal{F}_k)$,

$$\left(\frac{1}{2} - \frac{1}{2^{\bar{K}/2}}\right)|I| \leq |\mathcal{U}_{\mathcal{F}_k^1}[I](a)| \leq \left(\frac{1}{2} + \frac{1}{2^{\bar{K}/2}}\right)|I|;\tag{13.4}$$

- for all $a, b \in \alpha(\mathcal{F}_k^1)$ with $a < b$,

$$\left(\frac{1}{2} - \frac{1}{2^{\bar{K}/2}}\right)|\mathcal{U}_{\mathcal{F}_k^1}[I](a)| \leq |\mathcal{U}_{\mathcal{F}_k^1}[I](a) \cap \mathcal{U}_{\mathcal{F}_k^1}[I](b)| \leq \left(\frac{1}{2} + \frac{1}{2^{\bar{K}/2}}\right)|\mathcal{U}_{\mathcal{F}_k^1}[I](a)|.\tag{13.5}$$

Notice now that (13.5) becomes a very good approximation of (8.12).

Following Section 8.1, we require that the set F_k obeys in a first instance (8.13) and then the more general (8.14). In fact, we can now slightly simplify the initial dyadic scenario and directly define F_k as the set obeying:

- $F_k = \bigcup_{I\in\mathsf{IBtm}(\mathcal{F}_k^{\log k})} I \cap F_k$;
- $I \cap F_k := \bigcap_{l=1}^{\log k} \bigcap_{a\in\alpha(\mathcal{F}_k^l)} \mathcal{U}_{\mathcal{F}_k^l}[I](a)$.

Making essential use of (13.4) and (13.5) we then get the analogue of (8.20):

$$\left(\frac{1}{2} - \frac{1}{2^{\bar{K}/2}}\right)^{\#\alpha(\mathcal{F}_k)} \leq \frac{|I \cap F_k|}{|I|} \leq \left(\frac{1}{2} + \frac{1}{2^{\bar{K}/2}}\right)^{\#\alpha(\mathcal{F}_k)},\tag{13.6}$$

which from the choice of $\bar{K}$ further implies that

$$\frac{1}{e}\,\frac{1}{2^{\log k\, 2^{2k}-1}} \leq \frac{|I \cap F_k|}{|I|} \leq e\,\frac{1}{2^{\log k\, 2^{2k}-1}}.\tag{13.7}$$

In particular, since $|\tilde{\mathsf{I}}\mathsf{Btm}(\mathcal{F}_k^{\log k})| \approx 2^{-\log k} = 1/k$, we get

$$|F_k| \approx \frac{1}{k} 2^{-\log k\, 2^{2k}-1}.\tag{13.8}$$

Notice now that from the above construction one gets another key property that our set F_k must satisfy:

$$|I_P \cap F_k|/|I_P| \approx 2^l |F_k| \quad \forall l \in \{1, \ldots, \log k\} \text{ and } \forall P \in \mathcal{F}_k^l. \tag{13.9}$$

The above process can be repeated with the obvious changes for a general $j \in \{k/2+1, k\}$ and in this situation (13.7)–(13.9) become respectively

$$\frac{1}{e} \frac{1}{2^{\log k \, 2^{2j-1}}} \leq \frac{|I \cap F_j|}{|I|} \leq e \frac{1}{2^{\log k \, 2^{2j-1}}}, \tag{13.10}$$

$$|F_j| \approx \frac{1}{k} 2^{-\log k \, 2^{2j-1}}, \tag{13.11}$$

$$|I_P \cap F_j|/|I_P| \approx 2^l |F_j| \quad \forall l \in \{1, \ldots, \log k\} \text{ and } \forall P \in \mathcal{F}_j^l. \tag{13.12}$$

Thus at the end of this process one is able to construct the desired sequence of sets $\{F_j\}_{j=k/2+1}^k$.

Finally, based again on the properties (13.2), (13.4) and (13.5) one can easily check that the fundamental Claim in Section 8.3, more precisely (8.24), holds.

With this done, the reasonings in Sections 9 and 10 can be repeated, concluding the proof of Theorem 1.4.

2) Recalling the description of the **CME** $\mathcal{F} = \bigcup_{j=k/2+1}^k \mathcal{F}_j$ (see Section 7), this remark seeks to explain *why* we chose the height of each tower $\mathcal{F}_j$ to be of the order $\log k$. As one can notice, writing $\mathcal{F}_j = \bigcup_{l=1}^m \mathcal{F}_j^l$, a first impulse would be to aim for a height m as large as possible (e.g. $m = (\log k)^2$), hoping that this would increase the lower bound of the $L^{1,\infty}$ norm of $C_{\text{lac}} f_k$. However, in the view of the discussion in Section 6, we immediately notice that this is not the case. Indeed, based on (6.10), (6.11) and (6.14) and on the fact that

$$\{\mathcal{N}_j \geq m\} \subset \text{Basis}(\mathcal{F}_j^m) \subset \{\mathcal{N}_j > m/2\}, \tag{13.13}$$

we notice that there exists an absolute constant $c > 0$ such that if $m > C \log k$ with $C > 0$ an absolute constant large enough then

$$\sum_{j=k/2+1}^k |\text{Basis}(\mathcal{F}_j^m)| \leq k e^{-cm} = o(k), \tag{13.14}$$

and thus $\bigcup_{j=k/2+1}^k \text{Basis}(\mathcal{F}_j^m)$ becomes an "exceptional" (removable) set that consequently has a negligible impact on the size of $\|C_{\text{lac}} f_k\|_{1,\infty}$. From this, one further notices the strong connection between the tile structure maximizing $\|C_{\text{lac}} f_k\|_{1,\infty}$ and the one corresponding to the maximal size of $\|\mathcal{N}^{[k]}\|_{1,\infty}$.

3) Corollaries 1.6, 1.7, 1.10 and 1.11 are straightforward applications of the first part of Main Theorem 2.

4) For Corollary 1.13, preserving the notation in Section 7, one simply takes $\mathbb{P}^\alpha := \mathcal{F}_k^1$ and $f = \chi_{F_k}$. We leave further details to the interested reader.

5) Our entire paper relies on a completely new idea revealing the deep relationship between the behavior of the grand maximal counting function and the pointwise convergence of Fourier series near L^1. This is the first result in the literature regarding this topic that provides a counterexample using multiscale analysis of wave-packets. All the previous counterexamples focused on identifying and working directly with input functions and on involved computations with local pointwise estimates of (sub)sequences of partial Fourier sums applied to such input functions. In our case, we change the point of view, by first designing a geometric construction of tiles that encode both the nature of the lacunary sequence (reflected in their frequencies) and the extremizer property relative to the $L^{1,\infty}$ norm of the grand maximal counting function. The input functions are now naturally obtained as a direct byproduct of this construction, relying on the alignment of the oscillations requirement.

Thus, the entire paper is not about a technical $\log \log \log \log L$ factor addition but is about a *conceptual* advancement that identifies a *structural* mechanism in approaching the problem of pointwise convergence of Fourier series near L^1.

14. Appendix

In this section we review some of the basic facts concerning rearrangement invariant (quasi-)Banach spaces. We follow closely the description in [7] which further relies on [4].

Denote by $L^0(\mathbb{T})$ the topological linear space of all periodic Lebesgue-measurable functions equipped with the topology of convergence in measure. Given $f \in L^0(\mathbb{T})$, we define its *distribution function* as

$$m_f(\lambda) := m(\{x \in \mathbb{T} \mid |f(x)| > \lambda\}), \tag{14.1}$$

where m stands for the Lebesgue measure on $\mathbb{T}$.

The *decreasing rearrangement* of f is defined as

$$f^*(t) := \inf\{\lambda \geq 0 \mid m_f(\lambda) \leq t\}, \quad t \geq 0. \tag{14.2}$$

All the quasi-Banach spaces X mentioned in our paper are considered as subspaces of $L^0(\mathbb{T})$.

We say that X is a (*quasi-*)*Banach lattice* if the following properties are satisfied:

– there exists $h \in X$ with $h > 0$ a.e.;
– if $|f| \leq |g|$ a.e. with $g \in X$ and $f \in L^0(\mathbb{T})$ then $\|f\|_X \leq \|g\|_X$.

A (quasi-)Banach lattice $(X, \|\cdot\|_X)$ is called a *rearrangement invariant* (*quasi-*)*Banach space* if given $f \in X$ and $g \in L^0(\mathbb{T})$ with $m_f = m_g$ one has $g \in X$ and $\|f\|_X = \|g\|_X$.

If X is a r.i. (quasi-)Banach space and χ_A stands for the characteristic functions of a measurable set $A \subseteq \mathbb{T}$, then the function

$$\varphi_X(t) = \|\chi_A\|_X \quad \text{with } m(A) = t \in [0, 1] \tag{14.3}$$

is called the *fundamental function* of X.

In what follows we introduce two classes of r.i. Banach spaces: the Marcinkiewicz and the Lorentz spaces.

We say that $\varphi : \mathbb{T} \to \mathbb{R}_+$ is a *quasi-concave function* if:

- $\varphi(0) = 0$ and $\varphi(t) > 0$ for all $t \in (0, 1)$;
- the functions $\varphi(t)$ and $\varphi_*(t) = t/\varphi(t)$ are non-decreasing on $\mathbb{T}$.

It is worth noticing that, for our purposes here, we can always replace a quasi-concave function φ with its least concave majorant $\tilde{\varphi}$, since $\tilde{\varphi}(t) \leq 2\varphi(t) \leq 2\tilde{\varphi}(t)$ for all $t \in \mathbb{T}$.

The *Marcinkiewicz space M_φ* is the r.i. Banach space of all $f \in L^0(\mathbb{T})$ such that

$$\|f\|_{M_\varphi} := \sup_{t \in (0,1]} \frac{1}{\varphi(t)} \int_0^t f^*(s)\, ds < \infty. \qquad (14.4)$$

Defining a Lorentz space always requires first defining a function $\varphi : \mathbb{T} \to [0, \infty)$ with the following properties:

- $\varphi(0) = 0$;
- φ is non-decreasing; $\qquad\qquad\qquad\qquad\qquad\qquad\qquad (14.5)$
- φ is concave.

With φ as above, we define the *Lorentz space Λ_φ* as the r.i. Banach space of all the functions $f \in L^0(\mathbb{T})$ such that

$$\|f\|_{\Lambda_\varphi} := \int_0^1 f^*(s)\, d\varphi(s) < \infty. \qquad (14.6)$$

Finally, we close with the following observation: if X is a r.i. Banach space with fundamental function φ, then

- φ is quasi-concave;
- the following continuous inclusion holds:

$$\Lambda_{\tilde{\varphi}} \hookrightarrow X \hookrightarrow M_{\varphi_*}. \qquad (14.7)$$

References

[1] Antonov, N. Yu.: Convergence of Fourier series. East J. Approx. **2**, 187–196 (1996) Zbl 0919.42004 MR 1407066

[2] Antonov, N. Yu.: Almost everywhere divergent subsequences of Fourier sums of functions in $\phi(L) \cap H_1^\omega$. Mat. Zametki **85**, 502–515 (2009) (in Russian) Zbl 1217.42008 MR 2549413

[3] Arias-de-Reyna, J.: Pointwise convergence of Fourier series. J. London Math. Soc. (2) **65**, 139–153 (2002) Zbl 1017.42002 MR 1875141

[4] Bennett, C., Sharpley, R.: Interpolation of Operators. Academic Press, Boston (1988) Zbl 0647.46057 MR 0928802

[5] Carleson, L.: On convergence and growth of partial sumas of Fourier series. Acta Math. **116**, 135–157 (1966) Zbl 0144.06402 MR 0199631

[6] Carleson, L.: Interpolations by bounded analytic functions and the corona problem. Ann. of Math. **76**, 547–559 (1962) Zbl 0112.29702 MR 0141789

[7] Carro, M. J., Mastyło, M., Rodríguez-Piazza, L.: Almost everywhere convergent Fourier series. J. Fourier Anal. Appl. **18**, 266–286 (2012) Zbl 1246.46025 MR 2898729

[8] Chen, Y. M.: An almost everywhere divergent Fourier series of the class $L(\log^+ \log^+ L)^{1-\epsilon}$. J. London Math. Soc. **44**, 643–654 (1969) Zbl 0169.39703 MR 0240539

[9] Di Plinio, F.: Lacunary Fourier and Walsh–Fourier series near L^1. Collect. Math. **65**, 219–232 (2014) Zbl 1311.42008 MR 3189278

[10] Di Plinio, F.: Weak-L^p bounds for the Carleson and Walsh–Carleson operators. C. R. Math. Acad. Sci. Paris **352**, 327–331 (2014) Zbl 1290.42045 MR 3186922

[11] Do, Y., Lacey, M.: On the convergence of lacunary Walsh–Fourier series. Bull. London Math. Soc. **44**, 241–254 (2012) Zbl 1241.42005 MR 2914604

[12] Du-Bois-Reymond, P.: Untersuchungen über die Convergenz und Divergenz der Fourierschen Darstellungsformen. Abhand. Akad. München **12**, 1–102 (1876) Zbl 09.0298.01

[13] Fefferman, C.: Pointwise convergence of Fourier series. Ann. of Math. (2) **98**, 551–571 (1973) Zbl 0268.42009 MR 0340926

[14] Fourier, J.: Théorie analytique de la chaleur. Reprint of the 1822 original, éditions Jacques Gabay, Paris (1988) Zbl 05828987 MR 2856180

[15] Gosselin, R. P.: On the divergence of Fourier series. Proc. Amer. Math. Soc. **9**, 278–282 (1958) Zbl 0090.28201 MR 0096068

[16] Hagelstein, P. A.: Problems in interpolation theory related to the almost everywhere convergence of Fourier series. In: Topics in Harmonic Analysis and Ergodic Theory, Contemp. Math. 444, Amer. Math. Soc., 175–183 (2007) Zbl 1131.42018 MR 2423628

[17] Hunt, R. A.: On the convergence of Fourier series. In: Orthogonal Expansions and their Continuous Analogues (Edwardsville, IL, 1967), Southern Illinois Univ. Press, Carbondale, IL, 235–255 (1968) Zbl 0159.35701 MR 0238019

[18] Jensen, D.: On the boundedness of the Carleson operator. Junior Thesis, Princeton Univ. (2013)

[19] Kolmogoroff, A. N.: Une série de Fourier–Lebesgue divergente presque partout. Fund. Math. **4**, 324–328 (1923) JFM 49.0205.02

[20] Konyagin, S. V. : On a subsequence of Fourier–Walsh partial sums. Mat. Zametki **54**, no. 4, 69–75, 158 (1993) (in Russian) Zbl 0804.42014 MR 1256607

[21] Konyagin, S. V.: Divergence everywhere of subsequences of partial sums of trigonometric Fourier series. Proc. Steklov Inst. Math. (Function Theory, suppl. 2), S167–S175 (2005) Zbl 1312.42006 MR 2200228

[22] Konyagin, S. V.: Almost everywhere convergence and divergence of Fourier series. In: Proc. Int. Congress of Mathematicians (Madrid, 2006), Vol. II, Eur. Math. Soc., Zürich, 1393–1403 (2006) Zbl 1105.42003 MR 2275651

[23] Konyagin, S. V.: On divergence of trigonometric Fourier series everywhere. C. R. Acad. Sci. Paris Sér. I Math. **329**, 693–697 (1999) Zbl 0942.42002 MR 1724074

[24] Konyagin, S. V.: On the divergence everywhere of trigonometric Fourier series. Sb. Math. **191**, 97–120 (2000) Zbl 0967.42004 MR 1753494

[25] Körner, T. W.: Everywhere divergent Fourier series. Colloq. Math. **45**, 103–118 (1981) Zbl 0491.42011 MR 0652607

[26] Lacey, M., Thiele, C.: L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. Ann. of Math. **146**, 693–724 (1997) Zbl 0914.46034 MR 1491450

[27] Lacey, M., Thiele, C.: On Calderón's conjecture. Ann. of Math. **149**, 475–496 (1999) Zbl 0934.42012 MR 1689336

[28] Lacey, M., Thiele, C.: A proof of boundedness of the Carleson operator. Math. Res. Lett. **7**, 361–370 (2000) Zbl 0966.42009 MR 1783613

[29] Lebesgue, H.: Sur les intégrales singulières. Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. **3**, 25–117 (1909) JFM 41.0327.02 MR 1508308

[30] Lie, V.: On the boundedness of the Carleson operator near L^1. Rev. Mat. Iberoamer. **29**, 1239–1262 (2013) Zbl 1297.42011 MR 3148602

[31] Lie, V.: On the pointwise convergence of the sequence of partial Fourier sums along lacunary subsequences. J. Funct. Anal. **263**, 3333–3740 (2012) Zbl 1259.42005 MR 2984070

[32] Lie, V.: The polynomial Carleson operator. arXiv:1105.4504v1 (2011)

[33] Lie, V.: The (weak-L^2) boundedness of the quadratic Carleson operator. Geom. Funct. Anal. **19**, 457–497 (2009) Zbl 1178.42007 MR 2545246

[34] Luzin, N. N.: Integral and trigonometric series. Dissertation, 1915; also in: Collected Works of N. N. Luzin, Moscow (1953) (in Russian)

[35] Prohorenko, V. I. : Divergent Fourier series. Mat. Sb. (N.S.) **75** (117), 185–198 (1968) (in Russian) Zbl 0179.09801 MR 0223815

[36] Sjölin, P.: An inequality of Paley and convergence a.e. of Walsh–Fourier series. Ark. Mat. **7**, 551–570 (1969) Zbl 0169.08203 MR 0241885

[37] Sjölin, P., Soria, F.: Remarks on a theorem by N. Yu. Antonov. Studia Math. **158**, 79–97 (2003) Zbl 1053.42005 MR 2014553

[38] Soria, F.: Note on differentiation of integrals and the Halo conjecture. Studia Math. **81**, 29–36 (1985) Zbl 0526.28003 MR 0818168

[39] Soria, F.: On an extrapolation theorem of Carleson–Sjölin with applications to a.e. convergence of Fourier series. Studia Math. **94**, 235–244 (1989) Zbl 0684.42001 MR 1019791

[40] Tao, T.: Cancellation for the multilinear Hilbert transform. Collect. Math. **67**, 191–206 (2016) Zbl 06577173 MR 3484017

[41] Thiele, C.: The quartile operator and pointwise convergence of Walsh series. Trans. Amer. Math. Soc. **352**, 5745–5766 (2000) Zbl 0976.42017 MR 1695038

[42] Zygmund, A.: Trigonometric Series. Vols. I, II, 3rd ed., Cambridge Math. Library, Cambridge Univ. Press, Cambridge (2002) Zbl 1084.42003 MR 1963498