ANALYSIS & PDE

Volume 16 No. 10 2023

YUQIU FU, LARRY GUTH AND DOMINIQUE MALDAGUE

DECOUPLING INEQUALITIES FOR SHORT GENERALIZED DIRICHLET SEQUENCES

msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 10, pp. 2401–2464
DOI: 10.2140/apde.2023.16.2401

DECOUPLING INEQUALITIES FOR SHORT GENERALIZED DIRICHLET SEQUENCES

YUQIU FU, LARRY GUTH AND DOMINIQUE MALDAGUE

We study decoupling theory for functions on $\mathbb{R}$ with Fourier transform supported in a neighborhood of short Dirichlet sequences $\{\log n\}_{n=N+1}^{N+N^{1/2}}$, as well as sequences with similar convexity properties. We utilize the wave packet structure of functions with frequency support near an arithmetic progression.

1. Introduction

We study decoupling theory for functions $f : \mathbb{R} \to \mathbb{C}$ with Fourier support near certain convex sequences. As a model case of decoupling, consider the truncated parabola $\mathbb{P}^1 = \{(t, t^2) : |t| \le 1\}$. Let $R \ge 1$ be a large parameter and write $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ as a disjoint union of caps $\theta = \mathcal{N}_{R^{-1}}(\mathbb{P}^1) \cap (I \times \mathbb{R})$, where I is an $R^{-1/2}$-interval. The decoupling inequality of [Bourgain and Demeter 2015] says that if $2 \le p \le 6$, then for any $\varepsilon > 0$ there exists C_ε such that

$$\left\| \sum_\theta f_\theta \right\|_{L^p(\mathbb{R}^2)} \le C_\varepsilon R^\varepsilon \left(\sum_\theta \| f_\theta \|_{L^p(\mathbb{R}^2)}^2 \right)^{\frac{1}{2}}$$

whenever $f_\theta : \mathbb{R}^2 \to \mathbb{C}$ are Schwartz functions satisfying $\operatorname{supp} \widehat{f_\theta} \subset \theta$.

This paper explores analogues between decoupling for $\mathbb{P}^1$ and short Dirichlet sequences $\{\log n\}_{n=N+1}^{N+N^{1/2}}$, as well as sequences with similar convexity properties described in the following definition.

MSC2020: 42A38, 42B20.
Keywords: decoupling inequalities, wave packet decomposition, Dirichlet polynomials, convex sequences.

Definition 1.1. Let $N \geq 2$. We call $\{a_n\}_{n=1}^N$ a generalized Dirichlet sequence (with parameter N) if it satisfies the property

$$a_2 - a_1 \in \left[\frac{1}{4N}, \frac{4}{N}\right], \quad (a_{i+2} - a_{i+1}) - (a_{i+1} - a_i) \in \left[\frac{1}{4N^2}, \frac{4}{N^2}\right]. \tag{1}$$

We will call $\{a_n\}_{n=1}^{N^{1/2}}$ satisfying (1) an $N^{1/2}$- short generalized Dirichlet sequence.

For simplicity, we say short (generalized) Dirichlet sequence to mean $N^{1/2}$-short (generalized) Dirichlet sequence, unless otherwise specified. Note that the reflected short Dirichlet sequence,

$$\{-\log(N + N^{-\frac{1}{2}} - n + 1)\}_{n=1}^{N^{1/2}},$$

satisfies (1).

Now we describe our decoupling set-up. From now on $C, c > 0$ will denote absolute constants that may vary from line to line. For convenience of reading, we may regard C, c as 1. For $1 \leq L \leq cN^{1/2}$ and each $j = 1, \ldots, N^{1/2}/L$, define

$$I_j = \bigcup_{i=(j-1)L+1}^{jL} B_{L^2/N^2}(a_i),$$

where $B_{L^2/N^2}(a_i)$ means the L^2/N^2 interval centered at a_i. Let Ω be the L^2/N^2-neighborhood of $\{a_n\}_{n=1}^{N^{1/2}}$. We consider the partition

$$\Omega = \bigsqcup_j I_j. \tag{2}$$

We choose the L^2/N^2-neighborhood of $\{a_n\}_{n=1}^{N^{1/2}}$ because every I_j is essentially an L^2/N^2-neighborhood of an arithmetic progression, which we call a fat AP. To see this we calculate, for $1 \leq n \leq N^{1/2} - L$,

$$a_{n+L} - a_n - L(a_{n+1} - a_n) = \sum_{m=1}^{L} (a_{n+m} - a_{n+m-1} - (a_{n+1} - a_n)) \sim \sum_{m=1}^{L} \frac{m-1}{N^2} \sim \frac{L^2}{N^2}.$$

So indeed I_j lies in a CL^2/N^2-neighborhood of an L-term AP with common difference $a_{(j-1)L+1} - a_{(j-1)L}$ and starting point $a_{(j-1)L}$. Also, note that the common differences for distinct I_j are cL/N^2-separated.

We denote the partition $\{I_j\}_{j=1}^{N^{1/2}/L}$ by $\mathcal{I}$. The first main result of this paper is the following decoupling theorem for $\Omega = \bigsqcup_{I \in \mathcal{I}} I$.

Theorem 1.2. *Let Ω and $\mathcal{I}$ be defined as in the last paragraphs. Then for $2 \leq p \leq 6$ and every $\varepsilon > 0$*

$$\left\| \sum_{I \in \mathcal{I}} f_I \right\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \left(\sum_{I \in \mathcal{I}} \| f_I \|_{L^p(\mathbb{R})}^2 \right)^{\frac{1}{2}} \tag{3}$$

for functions f_I with supp $\widehat{f_I} \subset I$.

The range of p is sharp in the sense that (3) cannot hold for $p > 6$, which can be seen by taking $\widehat{f_I}$ to be a smooth bump with height 1 adapted to I for every I. Indeed for this choice of f_I, we have $\left|\sum_I f_I\right| \sim (L^2/N^2)N^{1/2}$ on $B_{cN^{1/2}}(0)$, and $\| f_I \|_{L^p(\mathbb{R})} \sim \| \widehat{f_I} \|_{L^{p'}(\mathbb{R})} \sim (L(L^2/N^2))^{1/p'}$, where $1/p + 1/p' = 1$.

So

$$\Big\|\sum_{I\in\mathcal{I}} f_I\Big\|_{L^p(\mathbb{R})} \gtrsim \frac{L^2}{N^2} N^{\frac{1}{2}} (N^{\frac{1}{2}})^{\frac{1}{p}}, \quad \Big(\sum_{I\in\mathcal{I}} \|f_I\|^2_{L^p(\mathbb{R})}\Big)^{\frac{1}{2}} \sim \Big(\frac{N^{\frac{1}{2}}}{L}\Big)^{\frac{1}{2}} \Big(\frac{L^3}{N^2}\Big)^{1-\frac{1}{p}}.$$

Then (3) would imply

$$\Big(\frac{N^{\frac{1}{2}}}{L}\Big)^{\frac{1}{2}-\frac{3}{p}} \lesssim_\varepsilon N^\varepsilon,$$

and hence $p \le 6$. We shall compare Theorem 1.2 with the $\ell^2 L^p$ decoupling inequality of the parabola in [Bourgain and Demeter 2015], which has the same critical exponent 6. Indeed we will see many similarities between short generalized Dirichlet sequences and $\mathbb{P}^1$ from a Fourier analytic point of view.

The notion of strict convexity of a sequence $\{a_n\}$ in $\mathbb{R}$ will parallel the role of curvature of the parabola in decoupling. Some key geometric aspects in the proof of decoupling for $\mathbb{P}^1$ are identifying caps θ as approximate $R^{-1/2} \times R^{-1}$ rectangles, which give rise to dual tubes θ^* of dimension $R^{1/2} \times R$, and noting that θ are separated in angle and so are θ^*. The $|f_\theta|$ are roughly constant on translates of θ^*.

In the $\{a_n\}_{n=1}^{N^{1/2}}$ setting, corresponding to f_θ we have f_{I_j} which are functions $f_{I_j} : \mathbb{R} \to \mathbb{C}$ satisfying $\operatorname{supp} \widehat{f_{I_j}} \subset I_j$. We may identify the (L^2/N^2)-neighborhood of I as approximately an (L^2/N^2)-neighborhood of an arithmetic progression (called a fat AP), giving rise to dual I^* defined in Definition 2.1, which are also fat APs, and note that distinct I are separated in step-size of the corresponding arithmetic progressions (and the same for I^*). The $|f_I|$ are also roughly constant on translates of I^* [Bourgain 1991; 1993].

Bourgain [1991; 1993] made use of this locally constant property to connect a conjecture of Montgomery with the Kakeya conjecture. To prove a decoupling inequality we need to identify another geometric analogy, the "ball", which is roughly the smallest set restricting to which in the physical space essentially preserves the frequency support.

For the R^{-1}-neighborhood of the parabola, the "ball" is a ball B_R of radius of R. We will define the "ball" $P(L)$ in the short generalized Dirichlet sequence setting in Section 3B. $P(L)$ will be a fat AP which sometimes degenerates to a Euclidean ball. With these notions of caps, tubes, and balls in the short generalized Dirichlet sequence setting, we are able to exploit the wave packet structure of a function with frequency support on $I \in \mathcal{I}$, and prove a bilinear Kakeya-type estimate (Proposition 3.3) and a bilinear restriction-type estimate (Proposition 3.5) that look almost identical to those in the parabola setting. The choice of $N^{1/2}$ plays an important role in making this resemblance possible, which we will discuss at the end of Section 7.

The proof of Theorem 1.2 is based on the high-low decomposition method in [Guth et al. 2022]. We do not intend to get a logarithmic decoupling constant as in that work, but we want to prove a refined decoupling inequality as in [Guth et al. 2020], which creates some technical differences.

The partition $\Omega = \bigsqcup_{I\in\mathcal{I}} I$ is maximal in the sense that if $\Omega = \bigsqcup I'$, where I' is the union of more than CL many adjacent intervals, then I' is no longer essentially a fat AP. Because of this, we will call $\Omega = \bigsqcup_{I\in\mathcal{I}} I$ the canonical partition and refer to Theorem 1.2 as decoupling for the canonical partition, or simply decoupling. In the spirit of small cap decoupling as in [Demeter et al. 2020], we may also

consider the "small cap" decoupling for short generalized Dirichlet sequences. Now we let $L_1 \in [1, L]$ be an integer, and we partition Ω into L_1 consecutive intervals J_j:

$$\Omega = \bigsqcup_{j=1}^{N^{1/2}/L_1} J_j = \bigsqcup_{j=1}^{N^{1/2}/L_1} \Big(\bigcup_{i=(j-1)L_1+1}^{jL_1} B_{L^2/N^2}(a_i) \Big). \tag{4}$$

We let $\mathcal{J}$ denote the partition $\{J_j\}_{j=1}^{N^{1/2}/L_1}$. The next decoupling result in this paper is small-cap-type decoupling inequalities.

Theorem 1.3. *Let $1 \le L_1 \le L \le N^{1/2}$, and $\{J\}_{J\in\mathcal{J}}$ be defined as in the paragraph above. Suppose $p \ge 4$. Then, for every $\varepsilon > 0$,*

$$\Big\| \sum_{J\in\mathcal{J}} f_J \Big\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \Big(\frac{N^{\frac12-\frac2p} L^{\frac2p}}{L_1^{1-\frac2p}} + \Big(\frac{N^{\frac12}}{L_1}\Big)^{\frac12-\frac1p} \Big) \Big(\sum_{J\in\mathcal{J}} \|f_J\|_{L^p(\mathbb{R})}^p \Big)^{\frac1p} \tag{5}$$

for a function $f_J : \mathbb{R} \to \mathbb{C}$ with supp $\widehat{f_J} \subset J$.

Inequality (5) is sharp up to $C_\varepsilon N^\varepsilon$ for every fixed p, L, L_1 satisfying the condition in Theorem 1.3. The first factor in front of $\big(\sum_{J\in\mathcal{J}} \|f_J\|_{L^p(\mathbb{R})}^p\big)^{1/p}$ is sharp because of the example $\widehat{f_J}$ equals a smooth bump adapted to J with height 1 for every $J \in \mathcal{J}$. The calculation is similar to the one in the paragraph below Theorem 1.2. The second factor is sharp because of the example $\widehat{f_J}$ equals a random sign times a smooth bump adapted to a ball of radius L^2/N^2 inside J with height 1 for every $J \in \mathcal{J}$, where the random signs are chosen so that $\int_{\mathbb{R}} \big|\sum_J f_J\big|^p \sim \int_{\mathbb{R}} \big(\sum_J |f_J|^2\big)^{p/2}$ by Khintchine's inequality.

The structure of the proof of Theorem 1.3 is similar to that of Theorem 3.1 in [Demeter et al. 2020], consisting of three ingredients: refined decoupling for the canonical partition, refined flat decoupling, and an incidence estimate. Refined decoupling for the canonical partition is a refined version of Theorem 1.2, which we will prove in Sections 4, 5, and 6 in order to derive Theorem 1.2. We show the other two counterparts in Section 8.

1A. ***L^p estimates for short generalized Dirichlet polynomials.*** A straight corollary of Theorem 1.3 is essentially sharp L^p estimates for short generalized Dirichlet polynomials $\sum_{n=1}^{N^{1/2}} b_n e^{ita_n}$.

Corollary 1.4. *Let $\{a_n\}_{n=1}^{N^{1/2}}$ be a short generalized Dirichlet sequence. Suppose $p \ge 4$ and $N \le T \le N^2$. We have for every $\varepsilon > 0$*

$$\Big\| \sum_{n=1}^{N^{1/2}} b_n e^{ita_n} \Big\|_{L^p(B_T)} \lesssim_\varepsilon N^\varepsilon (N^{\frac12} + T^{\frac1p} N^{\frac14-\frac{1}{2p}}) \|b_n\|_{\ell^p} \tag{6}$$

for every B_T and every $\{b_n\}_{n=1}^{N^{1/2}} \subset \mathbb{C}$.

If we let $L \in [1, N^{1/2}]$ be the integer such that $N^2/L^2 = T$, then Corollary 1.4 follows from Theorem 1.3 with that L, and $L_1 = 1$, applied to functions $f_J(t) = b_n e^{ita_n} \phi(t)$ for every J, where ϕ is a Schwartz function adapted to B_T with Fourier support inside $B_{T^{-1}}(0)$.

The inequality (6) is sharp up to $C_\varepsilon N^\varepsilon$. This is from discrete versions of the examples described below Theorem 1.3, taken with $L_1 = 1$: $b_n = 1$ for every n, and b_n equal to random signs.

We will in fact prove a more general version of Theorem 1.3 which allows us to get essentially sharp (ℓ^q, L^p) estimates for $\sum_{n=1}^{N^{1/2}} b_n e^{ita_n}$ in the range $p \geq 4$, $\frac{1}{p} + \frac{3}{q} \leq 1$. See Theorem 8.5 and Corollary 8.2.

After this work was done we learned from James Maynard a general transference method, which can in particular transfer the L^p estimate on a short generalized Dirichlet polynomial to a 2-dimensional L^p estimate on an exponential sum with frequency support near a convex curve in $\mathbb{R}^2$. This allows us to derive Corollary 1.4 directly from the small cap decoupling inequalities for the parabola in [Demeter et al. 2020]. We provide that particular argument in detail in the Appendix.

The starting point of this paper was to see whether decoupling methods could be used to make progress on Montgomery's conjecture on Dirichlet polynomials [1971; 1994]. Our investigation led us in a different direction, proving decoupling inequalities for short generalized Dirichlet sequences.

Conjecture 1.5 (Montgomery's conjecture). *For every $p \geq 2$ and every $\varepsilon > 0$ we have*

$$\Big\| \sum_{n=N+1}^{2N} b_n n^{it} \Big\|_{L^p(B_T)} \leq C_\varepsilon T^\varepsilon N^{\frac{1}{2}} (N^{\frac{p}{2}} + T)^{\frac{1}{p}} \|b_n\|_{\ell^\infty} \tag{7}$$

for every ball B_T of radius T and every $\{b_n\}_{n=N+1}^{2N} \subset \mathbb{C}$.

Conjecture 1.5 is widely open. In fact it has significant implications which are also hard conjectures. It is shown in [Montgomery 1971] that Conjecture 1.5 implies the density conjecture for the Riemann zeta function. Bourgain [1991; 1993] observed that a stronger version of Conjecture 1.5 on large value estimate of Dirichlet polynomials implies the Kakeya maximal operator conjecture in all dimensions. Conjecture 1.5 itself also implies a weaker statement that a Kakeya set has full Minkowski dimension; see [Green 2002].

Our Corollary 1.4 proves some L^p estimates for "short" Dirichlet polynomials which do not directly connect to Montgomery's conjecture. In fact we believe to make progress on Montgomery's conjecture significant new ideas are needed.

On the other hand, combining Theorem 1.2 with flat decoupling we obtain $\ell^2 L^p$ decoupling inequalities for generalized Dirichlet sequences (with N many terms instead of $N^{1/2}$), and the decoupling inequalities we get are essentially sharp for the class of generalized Dirichlet sequences. As a corollary we have essentially sharp (ℓ^2, L^p) estimates on generalized Dirichlet polynomials, but the Dirichlet polynomial $\sum_{n=N+1}^{2N} b_n e^{it \log n}$ has more structure and admits better estimates. This has to do with examples of generalized Dirichlet sequences containing a $cN^{1/2}$-term AP with common difference $CN^{-1/2}$, which $\{\log n\}_{n=N+1}^{2N}$ cannot contain by a number theory argument. We discuss these in detail in Section 7.

The paper is structured as follows. In Section 2 we will illustrate the wave packet structure of functions with frequency support in a fat AP. In Section 3 we prove a bilinear Kakeya-type estimate and a bilinear restriction-type estimate for functions with frequency support in a neighborhood of a short generalized Dirichlet sequence $\{a_n\}_{n=1}^{N^{1/2}}$. Sections 4, 5, and 6 are dedicated to proving Theorem 1.2. Section 4 introduces a refined decoupling inequality for the canonical partition (Theorem 4.4), which implies Theorem 1.2, and which we will actually prove. Section 5 sets up a high-low frequency decomposition for square functions at different scales, and in Section 6 we finish the proof of Theorem 4.4. Section 7 discusses the decoupling problem for (N-term) generalized Dirichlet sequences. In Section 8 we prove

Theorem 1.3. The Appendix is about the transference method for one-dimensional exponential sum estimates like (6).

Notation. C will denote a positive absolute constant that may vary from line to line, and it may be either small or large. $A \lesssim B$ means $A \le CB$, and $A \sim B$ means $A \lesssim B$ and $B \lesssim A$. We will also use $\mathcal{O}(A)$ to denote a quantity that is less than or equal to CA. $A \lesssim_q B$ will mean $A \le C_q B$ for some constant depending on q. Similarly $\mathcal{O}_q(A)$ denotes a quantity that is less than or equal to $C_q A$. There will be a parameter N and $A \lessapprox B$ denotes $A \lesssim_\varepsilon N^\varepsilon B$ for every $\varepsilon > 0$.

2. Locally constant property

We set up some notation and describe the locally constant property related to fat APs in this section.

Definition 2.1. We let $P_v^\delta(a)$ denote the δ-neighborhood of the arithmetic progression on $\mathbb{R}$ which contains a and has common difference v. We call $P_v^\delta(x_0) \cap B_R(x_0)$, or simply $P_v^\delta \cap B_R$, a fat AP with thickness δ, common difference v, and diameter R. We will call $P_{v^{-1}}^{R^{-1}} \cap B_{\delta^{-1}}$ a fat AP dual to $P_v^\delta \cap B_R$.

To exploit the locally constant property of a function with frequency support in a fat AP, we first construct a family of functions $\psi_k : \mathbb{R} \to \mathbb{C}$ adapted to a fat AP (in the frequency space).

Lemma 2.2. *For every $x_0 \in \mathbb{R}$, $\delta \le v/2$, $M \ge 1$, and $k \ge 1$ there exists a function $\psi_k : \mathbb{R} \to \mathbb{C}$ with the property*

$$\widehat{\psi}_k(\xi) = 1 \quad \text{on } P_v^\delta(x_0) \cap B_{Mv}(x_0), \qquad \operatorname{supp} \widehat{\psi}_k \subset P_v^{2\delta}(x_0) \cap B_{8^k Mv}(x_0), \tag{8}$$

and ψ_k decays at order k outside of the dual fat AP $P_{v^{-1}}^{(Mv)^{-1}}(0) \cap B_{\delta^{-1}}(0)$:

$$(M\delta) 1_{P_{v^{-1}}^{(Mv)^{-1}}(0) \cap B_{\delta^{-1}}(0)} \lesssim_k |\psi_k(x)| \lesssim_k M\delta \left(1 + \frac{d(x, v^{-1}\mathbb{Z})}{(Mv)^{-1}}\right)^{-k} \left(1 + \frac{d(x, B_{\delta^{-1}}(0))}{\delta^{-1}}\right)^{-k}. \tag{9}$$

We say such a ψ_k is adapted to the fat AP $P_v^\delta(x_0) \cap B_{Mv}(x_0)$ in the frequency space with order of decay k.

Proof. Since translation in frequency space corresponds to modulation in the physical space, we may assume $x_0 = 0$.

We start with the Dirichlet kernel

$$D_M(x) = \sum_{|j| \le M} e^{2\pi i j x} = \frac{\sin((2M+1)\pi x)}{\sin(\pi x)}.$$

We define $\widetilde{D}_1(x) = D_M(x)$. Then we define $\widetilde{D}_k(x)$ inductively by

$$\widetilde{D}_k(x) = d_k^{-1} \widetilde{D}_{k-1}(x) D_{8^{k-1}M/2}(x),$$

where $d_k = \|\widehat{D}_{8^{k-1}M/2}\|_{L^1(\mathbb{R})}$ is the total measure of $\widehat{D}_{8^{k-1}M/2}$. Equivalently we can define $\widetilde{D}_k$ explicitly as

$$\widetilde{D}_k = \tilde{d}\, D_M \prod_{1 \le s \le k-2} D_{8^s M/2}$$

for some suitable constant $\tilde{d} > 0$.

Since $\widetilde{D}_1 = D_M$ has the property

$$\widehat{\widetilde{D}}_1(\xi) = \sum_{|j|\le M} \delta_0(\xi - j),$$

by induction we can show that

$$\widehat{\widetilde{D}}_k(\xi) = \sum_{|j|\le M} \delta_0(\xi - j) + \sum_{M<|j|\le 8^k M/4} b_{j,k}\delta_0(\xi - j)$$

for some $0 \le b_{j,k} \le 1$. From the explicit expression of the Dirichlet kernel we see that $\widetilde{D}_1$ decays at order 1 outside of $P_1^{M^{-1}}(0)$:

$$|\widetilde{D}_0(x)| = |D_M(x)| \lesssim \frac{M}{1 + d(x,\mathbb{Z})/M^{-1}}.$$

By induction on k we obtain $\widetilde{D}_k$ decays at order k outside of $P_1^{M^{-1}}(0)$:

$$|\widetilde{D}_k(x)| \lesssim_k M\left(1 + \frac{d(x,\mathbb{Z})}{M^{-1}}\right)^{-k}. \tag{10}$$

Now let $\phi(x)$ be a Schwartz function such that $\widehat{\phi}$ is a smooth bump adapted to $B_1(0)$:

$$\widehat{\phi}(\xi) = 1 \quad \text{on } B_1(0), \quad \operatorname{supp}\widehat{\phi} \subset B_2(0).$$

Let $\phi_{\delta^{-1}}(x)$ be the function $\phi(\delta x)$. Note that $\phi_{\delta^{-1}}$ decays rapidly outside of $B_{\delta^{-1}}(0)$. Let ψ_k be given by

$$\widehat{\psi}_k := \widehat{\phi}_{\delta^{-1}} * \widehat{\widetilde{D}}_k(v^{-1}\xi)/v = \sum_{|j|\le M} \widehat{\phi}_{\delta^{-1}}(\xi - jv) + \sum_{M<|j|\le 8^k \frac{M}{4}} b_{j,k}\widehat{\phi}_{\delta^{-1}}(\xi - jv).$$

From this definition we immediately see property (8) holds. Writing ψ_k as

$$\psi_k(x) = \phi_{\delta^{-1}}(x)\widetilde{D}_k(vx)$$

we observe from (10) and the rapid decay of $\phi_{\delta^{-1}}$ outside $B_{\delta^{-1}}(0)$ that (9) holds. □

For every fat AP $P = P_{v^{-1}}^{(Mv)^{-1}}(x_0) \cap B_{\delta^{-1}}(x_0)$ with $\delta \le v$, and every $k \ge 100$, let $W_{P,k}$ be the weight function

$$W_{P,k}(x) = \left(1 + \frac{d(x, x_0 + v^{-1}\mathbb{Z})}{(Mv)^{-1}}\right)^{-k}\left(1 + \frac{d(x, B_{\delta^{-1}}(x_0))}{\delta^{-1}}\right)^{-k}.$$

We will use the notation

$$\int_{W_{P,k}} f(x)\,dx := \int_{\mathbb{R}} f(x)W_{P,k}(x)\,dx,$$

$$\fint_{W_{P,k}} f(x)\,dx := \frac{1}{\|W_{P,k}\|_{L^1(\mathbb{R})}}\int_{\mathbb{R}} f(x)W_{P,k}(x)\,dx,$$

$$\|f\|_{L^p(W_{P,k})} := \left(\fint_{W_{P,k}} |f|^p(x)\,dx\right)^{\frac{1}{p}}.$$

For measurable sets $E \subset \mathbb{R}$ we use similar notation for average integrals and L^p norms:

$$\fint_E f(x)\,dx := \frac{1}{|E|}\int_E f(x)\,dx,$$
$$\|f\|_{L^p(E)} := \left(\fint_E |f|^p(x)\,dx\right)^{\frac{1}{p}}.$$

For a fat AP P, consider another fat AP $P' \subset P$. Let $P' \subset P$ denote an indexing set of translates of P' which form an $\mathcal{O}(1)$-overlapping tiling of P. Then we have the pointwise inequality

$$1_P(x) \lesssim_k \sum_{P' \subset P} W_{P',k}(x) \lesssim_k W_{P,k}(x). \tag{11}$$

If we look at translated copies P'' of P, we have

$$\sum_{P'' \subset \mathbb{R}} W_{P'',k}(x) W_{P,k}(P'') \lesssim_k W_{P,k}(x). \tag{12}$$

Here $\sum_{P'' \subset \mathbb{R}}$ means summing over a tiling (with $\mathcal{O}(1)$ overlap) of $\mathbb{R}$ by P'', and $W_{P,k}(P'')$ is defined to be $W_{P,k}(\sup P'')$, which is comparable to $W_{P,k}(x)$ for any $x \in P''$.

Proposition 2.3 (locally constant property). *Suppose f satisfies* $\operatorname{supp} \widehat{f} \subset P_v^\delta \cap B_{Mv}$. *Then for every dual fat AP $P = P_{v^{-1}}^{(Mv)^{-1}} \cap B_{\delta^{-1}}$ and every $1 \le q < p < \infty$ we have*

$$\|f\|_{L^p(W_{P,k})} \lesssim_{p,q,k} \|f\|_{L^q(W_{P,\frac{qk}{p}})} \quad \text{if } \frac{qk}{p} \ge 100,$$
$$\|f\|_{L^\infty(P)} \lesssim_k \|f\|_{L^1(W_{P,k})}.$$

Proof. We first prove the second inequality. Fix $k \ge 100$. From (8) we have

$$f(x) = f * \psi_k(x) = \int_{\mathbb{R}} f(y)\psi_k(x-y)\,dy,$$

where ψ_k is the function in Lemma 2.2 adapted to $P_v^\delta \cap B_{Mv}$ in the frequency space with order of decay k. Therefore for $x \in P$ we have

$$\begin{aligned}|f(x)| &\le \int_{\mathbb{R}} |f(y)||\psi_k(x-y)|\,dy \\ &\le \int_{\mathbb{R}} |f(y)| \sup_{x \in P} |\psi_k(x-y)|\,dy \\ &\lesssim_k \delta M \int_{\mathbb{R}} |f(y)| W_{P,k}(y)\,dy \sim_k \fint_{W_{P,k}} |f(y)|\,dy.\end{aligned}$$

For the third inequality we used (9). Now we prove the first inequality in the proposition. We claim that from (12) (applied with k replaced by qk/p) and the assumption $q < p$ we only need to show

$$\|f\|_{L^p(P)} \lesssim_{p,q,k} \|f\|_{L^q(W_{P,k})}. \tag{13}$$

Indeed if (13) holds, then

$$\begin{aligned}\int_{W_{P,k}} |f|^p &\lesssim_k \sum_{P'\subset\mathbb{R}} \int_{P'} |f|^p W_{P,k}(P') \\ &\lesssim_{p,q,k} |P|^{1-\frac{p}{q}} \sum_{P'\subset\mathbb{R}} W_{P,k}(P') \Big(\int_{W_{P',qk/p}} |f|^q \Big)^{\frac{p}{q}} \\ &\le |P|^{1-\frac{p}{q}} \Big(\int_{\mathbb{R}} |f(x)|^q \sum_{P'\subset\mathbb{R}} W_{P,k}(P')^{\frac{q}{p}} W_{P',\frac{qk}{p}}(x)\, dx \Big)^{\frac{p}{q}} \\ &\lesssim_{p,q,k} |P|^{1-\frac{p}{q}} \Big(\int |f(x)|^q \sum_{P'\subset\mathbb{R}} W_{P,\frac{qk}{p}}(P') W_{P',\frac{qk}{p}}(x)\, dx \Big)^{\frac{p}{q}} \\ &\lesssim_{p,q,k} |P|^{1-\frac{p}{q}} \Big(\int |f|^q W_{P,\frac{qk}{p}} \Big)^{\frac{p}{q}} \quad \text{(by (12))},\end{aligned}$$

which is exactly the first inequality in the proposition. To show (13) we observe that the second inequality in the proposition together with Hölder's inequality implies that

$$\|f\|_{L^p(P)} \le \|f\|_{L^\infty(P)} \lesssim_{p,q,k} \|f\|_{L^q(W_{P,\frac{qk}{p}})},$$

which is (13). □

3. Bilinear Kakeya-type and restriction-type estimates

Kakeya and restriction-type estimates are closely related to decoupling, and we will use the bilinear version of them in the proof of Theorem 1.2, but first we need to introduce a more general decoupling set-up for the purpose of induction.

3A. ***General set-up.*** To prove Theorem 1.2 we will do a broad-narrow argument which involves rescaling of a segment of $\{a_n\}_{n=1}^{N^{1/2}}$. To properly set up our induction hypothesis we consider the following more general class of generalized Dirichlet sequences.

Definition 3.1 (generalized Dirichlet sequence). Let $\theta \in (0,1]$ and $N \ge 2$. We call $\{a_n\}_{n=1}^{N}$ a generalized Dirichlet sequence (with parameters N, θ) if it satisfies the property

$$a_2 - a_1 \in \Big[\frac{1}{4N}, \frac{4}{N}\Big], \quad (a_{i+2} - a_{i+1}) - (a_{i+1} - a_i) \in \Big[\frac{\theta}{4N^2}, \frac{4\theta}{N^2}\Big]. \tag{14}$$

We will call $\{a_n\}_{n=1}^{N^{1/2}}$ satisfying (1) an $N^{1/2}$-short generalized Dirichlet sequence (with parameters N, θ).

As before we write "short" for "$N^{1/2}$-short" for simplicity. Comparing with Definition 1.1 we see an extra parameter θ which measures the convexity of the sequence. From now on we use Definition 3.1 for the definition of generalized Dirichlet sequence.

We shall also incorporate θ in our decoupling set-up. Let $\{a_n\}_{n=1}^{N^{1/2}}$ be a short generalized Dirichlet sequence with parameter $\theta \in (0,1]$. From the spacing property (14) of $\{a_n\}_{n=1}^{N^{1/2}}$ we see that, for every $1 \le j \le N^{1/2}/L$, $\{a_n\}_{n=(j-1)L+1}^{jL}$ is essentially contained in an $L^2\theta/N^2$-neighborhood of an arithmetic progression. Indeed, if we define $v_j = a_{(j-1)L+2} - a_{(j-1)L+1}$, then $\{a_n\}_{n=(j-1)L+1}^{jL}$ is contained in

the $CL^2\theta/N^2$-neighborhood of the arithmetic progression containing a_{jL} with common difference v_j, that is,

$$\{a_n\}_{n=(j-1)L+1}^{jL} \subset P_{v_j}^{CL^2\theta/N^2}(a_{jL}) \cap B_{CL/N}(a_{jL}).$$

Now we let Ω be the $\theta L^2/N^2$-neighborhood of $\{a_n\}_{n=1}^{N^{1/2}}$. For $1 \le L \le cN^{1/2}$ and each $j = 1, \dots, N^{1/2}/L$, define

$$I_j = \bigcup_{i=(j-1)L+1}^{jL} B_{\theta L^2/N^2}(a_i).$$

We denote the collection of I_j by $\mathcal{I}$, and consider the partition

$$\Omega = \bigsqcup_{I \in \mathcal{I}} I.$$

This will be our new decoupling set-up for the canonical partition, and from now on the notation here supersedes that in the Introduction. For small-cap-type decoupling we postpone the description of the corresponding general set-up to Section 8.

3B. ***Analogies between*** $\{a_n\}_{n=1}^{N^{1/2}}$ ***and*** $\mathbb{P}^1$**.** For $I = I_j \in \mathcal{I}$, we let

$$\widetilde{I}_j := P_{v_j}^{CL^2\theta/N^2}(a_{jL}) \cap B_{CL/N}(a_{jL}),$$

with C large enough so that

$$I = I_j \subset \widetilde{I}_j = \widetilde{I}.$$

Here $v_j = a_{(j-1)L+2} - a_{(j-1)L+1}$ and $v_j \sim N^{-1}$.

For each $I \in \mathcal{I}$, we denote by $P_I(x)$ the fat AP dual to $\widetilde{I}$ and centered at x, that is,

$$P_I(x) := P_{v_j^{-1}}^{CN/L}(x) \cap B_{CN^2/(L^2\theta)}(x) \tag{15}$$

if $I = I_j$, and we simply write P_I if stressing the center x is unnecessary. For $I = I_j$, we also write v_I to denote v_j. We let $P(L, y)$ denote a larger fat AP

$$P(L, y) := P_{v_1^{-1}}^{CN^{3/2}/L^2}(y) \cap B_{CN^2/(L^2\theta)}(y), \tag{16}$$

and we simply write $P(L)$ if stressing the center y is unnecessary. If $L \lesssim N^{1/4}$ we have $N^{3/2}/L^2 \ge N$ and in that case $P(L)$ is a ball $B_{CN^2/(L^2\theta)}$. Comparing (15) and (16), we see $P(L)$ has a larger thickness size $CN^{3/2}/L^2$. We will see shortly (Lemma 3.2 and the paragraph following it) that $CN^{3/2}/L^2$ is the smallest thickness that allows us to fit a P_I in any fixed $P(L)$ for every $I \in \mathcal{I}$.

The starting point of this paper is to make use of an analogy between the extension operator on $\{a_n\}_{n=1}^{N^{1/2}}$

$$\{b_n\}_{n=1}^{N^{1/2}} \mapsto \sum_{n=1}^{N^{1/2}} b_n e^{ita_n}$$

and the extension operator on the truncated parabola $\mathbb{P}^1$

$$f \mapsto \int_{[-1,1]} f(\xi)\, e^{i(x\xi + t\xi^2)}\, d\xi.$$

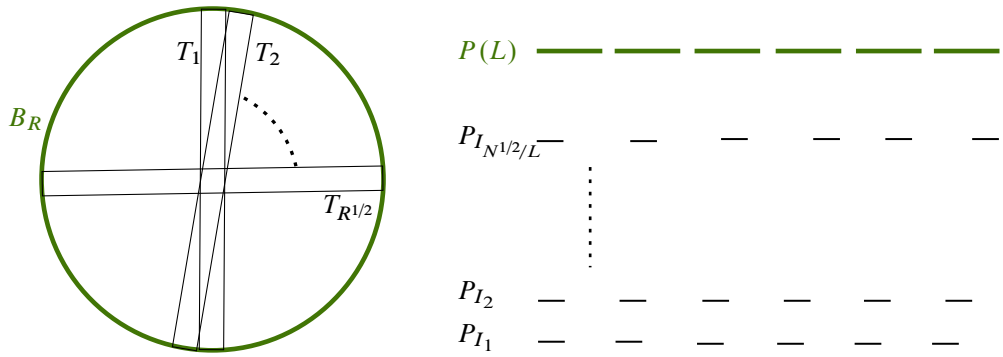

Figure 1. The ball $B_R \subset \mathbb{R}^2$ contains the union of tubes T_i having the same center, each of which is dual to θ_i, where $\bigsqcup_i \theta_i$ partitions $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$. On the right, we see analogous dual fat APs, one P_{I_i} per I_i which partition Ω into L consecutive intervals. We see that $P(L)$ contains the union of the P_{I_i} which have the same starting point.

We list the correspondence between objects in this paper and in the parabola setting. For simplicity we assume $\theta = 1$ in the following list:

(1) The parameter $L \in [1, N^{1/2}]$ is the length of the "cap" that we are looking at, and that determines a canonical neighborhood Ω with width L^2/N^2. The corresponding parameter in the parabola setting is R, which determines the length ($R^{-1/2}$) of the cap and a canonical neighborhood with width R^{-1}.

(2) The $\widetilde{I}$, P_I defined above is analogous to the cap and tube in the context of parabola decoupling. Let Θ be a partition of $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$, the R^{-1}-neighborhood of the truncated parabola $\mathbb{P}^1$ (over $[-1, 1]$), into $R^{-1/2} \times R^{-1}$ caps θ. The dual object of θ is a tube T of dimension $R^{1/2} \times R$.

(3) $P(L)$ is defined to be the smallest fat AP with the property that, for a function F with frequency support on Ω, "restricting" F in the physical space to $P(L)$ will essentially preserve its frequency support. The corresponding object for the parabola is B_R, a ball of radius R.

See Figure 1 which illustrates the analogous properties of tubes T with the ball B_R and fat APs P_I with $P(L)$. Bourgain [1991; 1993] made use of the first two analogies. The new ingredient we need is the third analogy, which gives an appropriate notion of ball in the short generalized Dirichlet sequence setting. It is very important that we define $P(L)$ to be the smallest fat AP with such a property. If we naively use B_{N^2/L^2} as the ball $P(L)$, the whole argument that follows will break down.

To make the third point precise, we prove the following lemma. We introduce one more notation. For a general fat AP $P = P_v^\delta(x_0) \cap B_{Mv}(x_0)$ and $s > 0$, sP will denote the fat AP $P_v^{s\delta}(x_0) \cap B_{sMv}(x_0)$.

Lemma 3.2. *Fix a $P(L)$. For every $I \in \mathcal{I}$ and every P_I with $P_I \cap P(L) \neq \varnothing$, P_I is contained in $2P(L)$.*

Proof. In fact for every j, the difference of differences hypothesis in (14) implies that $|v_j - v_1| \lesssim N^{-3/2}\theta$. It follows that $|v_j^{-1} - v_1^{-1}| \lesssim N^{1/2}\theta$. Therefore $P_I \cap P(L) \neq \varnothing$ implies

$$d(P_I, P(L)) \lesssim (N^{\frac{1}{2}}\theta)\frac{N^2/(L^2\theta)}{N} = \frac{N^{\frac{3}{2}}}{L^2}, \tag{17}$$

which implies $P_I \subset 2P(L)$ if C is large enough in the definition of $P(L)$. □

To put it in another way, the proof above shows that, for every $x \in \mathbb{R}$,

$$\bigcup_{I \in \mathcal{I}} P_I(x) \subset P(L, x) \tag{18}$$

if C is large enough in the definition of $P(L)$. Since the inequality (17) is sharp up to a constant, the choice of $CN^{3/2}/L^2$ as the thickness in the definition of $P(L)$ makes (18) and Lemma 3.2 barely hold.

We note that the above lemma holds if we replace $I, \mathcal{I}, P_I, P(L)$ by θ, Θ, T, B_R respectively.

3C. ***Transversality and bilinear Kakeya-type estimate.*** We say $I, J \in \mathcal{I}$ are transversal if $|v_I^{-1} - v_J^{-1}| \gtrsim N^{1/2}\theta$, or equivalently, if $d(I, J) \gtrsim N^{-1/2}$ on $\mathbb{R}$. We now prove a bilinear Kakeya-type estimate for two transversal families of P_I.

Proposition 3.3 (bilinear Kakeya-type estimate). *Suppose $g_1 = \sum_I a_I 1_{P_I}$ and $g_2 = \sum_J b_J 1_{P_J}$, where a_I, b_J are positive real numbers, $I, J \in \mathcal{I}$ and P_I are transversal to P_J. Then*

$$\fint_{P(L)} g_1 g_2 \lesssim \fint_{2P(L)} g_1 \fint_{2P(L)} g_2. \tag{19}$$

For comparison we state the bilinear Kakeya-type estimates for $R^{1/2} \times R$ tubes in $\mathbb{R}^2$.

Proposition 3.4. *Suppose $g_1 = \sum_i a_i 1_{T_i}$ and $g_2 = \sum_j b_j 1_{T_j}$, where a_i, b_j are positive real numbers, T_i, T_j are $R^{1/2} \times R$ tubes and every T_i is transversal to every T_j (in the sense that the angle between T_i, T_j is $\gtrsim 1$). Then*

$$\fint_{B_R} g_1 g_2 \lesssim \fint_{2B_R} g_1 \fint_{2B_R} g_2.$$

Proof of Proposition 3.3. For simplicity of notation we assume $C = 1$ in (15), (16). For general C the argument works the same way. Since

$$\fint_{P(L)} g_1 g_2 \le \sum_{I,J:\, P_I \cap P(L) \neq \varnothing,\, P_J \cap P(L) \neq \varnothing} a_I b_J |P(L)|^{-1} |P_I \cap P_J|$$

it suffices to show that for I, J transversal we have

$$|P_I \cap P_J| \lesssim \frac{|P_I|^2}{|P(L)|}. \tag{20}$$

We consider two cases $L \ge C_1 N^{1/4}$ and $L \le C_1 N^{1/4}$ separately, where C_1 is a sufficiently large constant that will be chosen.

Case 1: $L \ge C_1 N^{1/4}$. Without loss of generality we assume P_I, P_J both start at the origin (meaning that the first term of the underlying AP is 0). Let $P_{I,k}$ denote the k-th interval in P_I. If V_I, V_J are the common difference of P_I, P_J respectively, then from the transversality assumption we have $|V_I - V_J| \sim N^{1/2}\theta$. So for some integer

$$K \sim \frac{N/L}{N^{1/2}\theta} = \frac{N^{1/2}}{L\theta}$$

we have

$$d(P_{I,k}, P_{J,k}) \le \frac{N}{L} \quad \text{if } 1 \le k \le K$$

and

$$d(P_{I,k}, P_{J,k}) \in \left[\frac{N}{L}, N\right] \quad \text{if } K \le k \lesssim \frac{N}{N^{\frac{1}{2}}\theta} = \frac{N^{\frac{1}{2}}}{\theta}.$$

Since $L \ge C_1 N^{1/4}$ we know that if C_1 is sufficiently large then $N^{1/2}/\theta N = N^{3/2}/\theta$ is larger than $N^2/(L^2\theta)$, which is the diameter of P_I. Therefore we have

$$|P_I \cap P_J| \lesssim \frac{N^{\frac{1}{2}}}{L\theta}\frac{N}{L} = \frac{N^{\frac{3}{2}}}{L^2\theta} = \frac{|P_I|^2}{|P(L)|}.$$

Case 2: $L \le C_1 N^{1/4}$. From the first case we know that

$$|P_I \cap P_J \cap B_{CN^{3/2}/\theta}| \lesssim \frac{N^{\frac{3}{2}}}{L^2\theta}.$$

Therefore by the triangle inequality we have

$$|P_I \cap P_J| \lesssim \frac{N^{\frac{3}{2}}}{L^2\theta}\frac{N^2/(L^2\theta)}{N^{\frac{3}{2}}/\theta} = \frac{N^2}{L^4\theta} = \frac{|P_I|^2}{|P(L)|}.$$

Here we recall that $P(L)$ degenerates to the Euclidean ball $B_{N^2/(L^2\theta)}$ if $L \le N^{1/4}$. So we have shown (20) and hence (19). □

3D. ***Bilinear restriction-type estimate.*** To prove a bilinear restriction estimate, we will use the above bilinear Kakeya estimate and induction on L. First we identify where the (square of the) square function $\sum_{I\in\mathcal{I}} |f_I|^2$ is locally constant on. Note that $\operatorname{supp} \widehat{|f_I|^2} \subset I - I \subset P_{v_I}^{CL^2\theta/N^2}(0) \cap B_{CL/N}(0)$. Since $|v_I - v_1| \lesssim N^{-3/2}\theta$ for every $I \in \mathcal{I}$, we have

$$\bigcup_{I\in\mathcal{I}} (I - I) \subset P_{v_1}^{CL\theta/N^{3/2}} \cap B_{CL/N}.$$

Therefore $\sum_I |f_I|^2$ is locally constant on dual fat AP of the form $P_{v_1^{-1}}^{CN/L} \cap B_{CN^{3/2}/(L\theta)}$. Observe that if we define $L_1 = (N^{1/2}L)^{1/2}$, then

$$P_{v_1^{-1}}^{CN/L} \cap B_{CN^{3/2}/(L\theta)} = P_{v_1^{-1}}^{CN^{3/2}/L_1^2} \cap B_{CN^2/(L_1^2\theta)} = CP(L_1).$$

Now suppose I', I'' are unions of I in $\mathcal{I}$, and I', I'' are transversal in the sense that $d(I', I'') \gtrsim N^{-1/2}$ on $\mathbb{R}$. Then we have the following bilinear restriction estimate. The proof closely resembles the multilinear Kakeya implies multilinear restriction proof in [Bennett et al. 2006].

Proposition 3.5 (bilinear restriction-type estimate). *Suppose* $\operatorname{supp} \widehat{F}_1 \subset I'$ *and* $\operatorname{supp} \widehat{F}_2 \subset I''$. *Then we have*

$$\fint_{P(L)} |F_1|^2|F_2|^2 \lesssim_\varepsilon N^\varepsilon |P(L)|^{-2} \int_{\mathbb{R}} |F_1|^2 \int_{\mathbb{R}} |F_2|^2. \tag{21}$$

Before proving the proposition, we remark that under the conditions of Proposition 3.5, the seemingly stronger inequality

$$\fint_{P(L)} |F_1|^2|F_2|^2 \lesssim_\varepsilon N^\varepsilon |P(L)|^{-2} \int_{\mathbb{R}} |F_1|^2 W_{P(L),100} \int_{\mathbb{R}} |F_2|^2 W_{P(L),100} \tag{22}$$

holds. This is essentially by applying Proposition 3.5 to the functions $F_1\psi, F_2\psi$, where ψ is from Lemma 2.2 and is adapted to the fat AP dual to $P(L)$, with order of decay 100.

Proof of Proposition 3.5. We define $\mathrm{BR}(L)$ to be the smallest constant such that

$$\fint_{P(L)} |F_1|^2|F_2|^2 \le \mathrm{BR}(L)|P(L)|^{-2}\int_{\mathbb{R}}|F_1|^2\int_{\mathbb{R}}|F_2|^2$$

holds for all F_1, F_2 with $\operatorname{supp}\widehat{F}_1 \subset I'$ and $\operatorname{supp}\widehat{F}_2 \subset I''$. We let $\mathrm{BK}(L)$ be the smallest constant such that

$$\fint_{P(L)} g_1 g_2 \le \mathrm{BK}(L)|P(L)|^{-2}\int_{\mathbb{R}} g_1\int_{\mathbb{R}} g_2$$

holds for all $g_1 = \sum a_I 1_{P_I}$ and $g_2 = \sum b_J 1_{P_J}$, where a_I, b_J are positive real numbers and $I, J \in \mathcal{I}$ with $I \subset I'$, $J \subset I''$. Equivalently, we have

$$|P_I|^{-2}\fint_{P(L)}\Big(\sum_{I\subset I'} g_{1,I} * 1_{P_I(0)}\Big)\Big(\sum_{J\subset I''} g_{2,J} * 1_{P_J(0)}\Big) \le \mathrm{BK}(L)|P(L)|^{-2}\Big(\int_{\mathbb{R}}\sum_I g_{1,I}\Big)\Big(\int_{\mathbb{R}}\sum_J g_{2,J}\Big) \tag{23}$$

for all finite measures $g_{1,I}$, $g_{2,J}$ which are linear combinations of Dirac measures with nonnegative coefficients. By a density argument (linear combinations of Dirac measures are dense in the weak* topology on $C_0(\mathbb{R})^*$), (23) also holds for all finite measures $g_{1,I}$, $g_{2,J}$. In particular, (23) holds for all nonnegative L^1 functions $g_{1,I}$, $g_{2,J}$.

We have shown in Proposition 3.3 that

$$\mathrm{BK}(L) \lesssim 1.$$

Now we want to show $\mathrm{BR}(L) \lesssim_\varepsilon N^\varepsilon$. First we prove

$$\mathrm{BR}(L) \lesssim \mathrm{BR}(L_1)\,\mathrm{BK}(L). \tag{24}$$

From the definition of BR and local L^2 orthogonality (Lemma 3.6 below) we have

$$\begin{aligned}\fint_{P(L)}|F_1F_2|^2 &\lesssim \fint_{P(L)}\|F_1F_2\|^2_{L^2(P(L_1,x))}\,dx\\ &\lesssim \mathrm{BR}(L_1)\fint_{P(L)}\|F_1\|^2_{L^2(W_{P(L_1,x),200})}\|F_2\|^2_{L^2(W_{P(L_1,x),200})}\\ &\lesssim \mathrm{BR}(L_1)\fint_{P(L)}\Big(\sum_{I\subset I'}\|F_{1,I}\|^2_{L^2(W_{P(L_1,x),200})}\Big)\Big(\sum_{J\subset I''}\|F_{2,J}\|^2_{L^2(W_{P(L_1,x),200})}\Big).\end{aligned}$$

We claim that

$$\fint_{P(L)}\sum_{I,J}\|F_{1,I}\|^2_{L^2(W_{P(L_1,x),200})}\|F_{2,J}\|^2_{L^2(W_{P(L_1,x),200})} \lesssim \mathrm{BK}(L)|P(L)|^{-2}\|F_1\|^2_{L^2(\mathbb{R})}\|F_2\|^2_{L^2(\mathbb{R})}, \tag{25}$$

which together with previous arguments will imply (24). Since $\sum_{P(L_1)\subset\mathbb{R}} W_{P(L_1,x),200}(P(L_1)) \lesssim 1$, it suffices to show that

$$\fint_{P(L)}\sum_{I,J}\|F_{1,I}\|^2_{L^2(P(L_1,x))}\|F_{2,J}\|^2_{L^2(P(L_1,x))} \lesssim_k \mathrm{BK}(L)|P(L)|^{-2}\|F_1\|^2_{L^2(\mathbb{R})}\|F_2\|^2_{L^2(\mathbb{R})}.$$

We choose $\psi_{I,200}$ adapted to $P_I(0)$ in the frequency space with order of decay 200 as in Lemma 2.2. Let $\phi_I := \check{\psi}_{I,200}/|P_I|$. If we define $G_{1,I} = (\widehat{F}_{1,I}/\widehat{\phi}_I)^{\vee}$, then due to the support property of $\widehat{F}_{1,I}$ we have pointwise

$$|\widehat{G}_{1,I}| \sim |\widehat{F}_{1,I}|. \tag{26}$$

Also by definition we have $F_{1,I} = G_{1,I} * \phi_I$. We define $G_{2,J} = (\widehat{F}_{2,J}/\widehat{\phi}_J)^{\vee}$ for $F_{2,J}$ in the same way. Now for $y \in \mathbb{R}$ such that $x + y \in P(L_1, x)$, we have

$$|F_{1,I}(x+y)|^2 = |(G_{1,I} * \phi_I)(x+y)|^2 \lesssim (|G_{1,I}|^2 * |\phi_I|)(x+y) \lesssim |G_{1,I}|^2 * 1_{CP_I}/|P_I|,$$

where we used Jensen's inequality for the first inequality. Therefore we have

$$\|F_{1,I}\|^2_{L^2(P(L_1,x))} \lesssim |G_{1,I}|^2 * 1_{CP_I}/|P_I|.$$

and similarly

$$\|F_{2,J}\|^2_{L^2(P(L_1,x))} \lesssim |G_{2,J}|^2 * 1_{CP_J}/|P_I|.$$

Hence using (23) we obtain

$$\begin{aligned}
\fint_{P(L)} \sum_{I,J} \|F_{1,I}\|^2_{L^2(P(L_1,x))} \|F_{2,J}\|^2_{L^2(P(L_1,x))} &\lesssim |P_I|^{-2} \sum_{I,J} \fint_{P(L)} (|G_{1,I}|^2 * 1_{CP_I})(|G_{2,J}|^2 * 1_{CP_J}) \\
&\lesssim \mathrm{BK}(L)|P(L)|^{-2} \left(\int_{\mathbb{R}} \sum_I |G_{1,I}|^2 \right) \left(\int_{\mathbb{R}} \sum_J |G_{2,J}|^2 \right) \\
&\lesssim \mathrm{BK}(L)|P(L)|^{-2} \left(\int_{\mathbb{R}} \sum_I |F_{1,I}|^2 \right) \left(\int_{\mathbb{R}} \sum_J |F_{2,J}|^2 \right) \\
&\lesssim \mathrm{BK}(L)|P(L)|^{-2} \left(\int_{\mathbb{R}} |F_1|^2 \right) \left(\int_{\mathbb{R}} |F_2|^2 \right),
\end{aligned}$$

where the second-to-last inequality is due to (26). So we have proved (25) and therefore (24). Now we prove $\mathrm{BR}(L) \lesssim_\varepsilon N^\varepsilon$. Define $L_m = (L_{m-1} N^{1/2})^{1/2}$. Fix an $\varepsilon > 0$. We define M to be the smallest integer such that $L_M \gtrsim N^{1/2-\varepsilon}$. So $M \lesssim_\varepsilon 1$. Plugging in $\mathrm{BK}(L_m) \lesssim 1$ and applying (24) repeatedly we get

$$\mathrm{BR}(L) \le C^M \, \mathrm{BR}(L_M).$$

Since $\mathrm{BR}(L_M) \lesssim_\varepsilon N^{C\varepsilon}$ for some universal constant C (because of the locally constant property Proposition 2.3) we conclude $\mathrm{BR}(L) \lesssim_\varepsilon N^{C\varepsilon}$, which is what we want. □

The L^4 bilinear restriction inequality for the parabola in $\mathbb{R}^2$ has a more straightforward proof exploiting the fact that $\#\{(\theta_3, \theta_4) : d(\theta_3, \theta_4) \gtrsim 1,\ \mathcal{N}_{R^{-1/2}}(\theta_3 + \theta_4) \cap \mathcal{N}_{R^{-1/2}}(\theta_1 + \theta_2)\} \lesssim 1$ for every fixed θ_1, θ_2, with $d(\theta_1, \theta_2) \gtrsim 1$, where θ_i are $R^{-1/2} \times R^{-1}$ caps that cover the compact parabola [Cordoba 1977; Fefferman 1973]. However, it is not obvious whether a similar property would hold for $\mathcal{I}$ in our setting, so we took the approach in [Bennett et al. 2006] instead.

Now we give a proof of the local L^2 orthogonality used in the proof above. We denote $(LN^{1/2})^{1/2}$ by L'. So $P(L') = P(L_1) = P_{v_1^{-1}}^{CN/L} \cap B_{CN^{3/2}/(L\theta)}$.

Lemma 3.6 (local L^2 orthogonality). *For every f_I with* supp $\widehat{f_I} \subset I$ *we have*

$$\Big\| \sum_{I\in\mathcal{I}} f_I \Big\|^2_{L^2(W_{P(L'),k})} \lesssim_k \sum_{I\in\mathcal{I}} \|f_I\|^2_{L^2(W_{P(L'),k})} \tag{27}$$

Proof. Due to (12) it suffices to prove

$$\Big\| \sum_{I\in\mathcal{I}} f_I \Big\|^2_{L^2(P(L'))} \lesssim_k \sum_{I\in\mathcal{I}} \|f_I\|^2_{L^2(W_{P(L'),k})}.$$

We choose ψ_k adapted to $P(L')^* := P_{v_1}^{CL\theta/N^{3/2}}(0) \cap B_{CL/N}(0)$ in the frequency space with order of decay k as in Lemma 2.2. Here $P(L')^*$ is dual to $P(L')$. Since supp $\widehat{\psi}_k \subset 8^k P(L')^*$, and $\{I + 8^k P(L')^*\}_{I\in\mathcal{I}}$ is $\mathcal{O}_k(1)$-overlapping, we conclude

$$\begin{aligned}\Big\| \sum_{I\in\mathcal{I}} f_I \Big\|^2_{L^2(P(L'))} &\lesssim_k |P(L')| \Big\| \sum_{I\in\mathcal{I}} f_I \psi_k \Big\|^2_{L^2(\mathbb{R})} \\ &\lesssim_k |P(L')| \sum_{I\in\mathcal{I}} \|f_I \psi_k\|^2_{L^2(\mathbb{R})} \lesssim_k \sum_{I\in\mathcal{I}} \|f_I\|^2_{L^2(W_{P(L'),k})}.\end{aligned}$$

□

4. Decoupling for the canonical partition

We focus on proving Theorem 1.2 in Sections 4, 5, and 6, and in these three sections decoupling will refer to decoupling for the canonical partition.

Recall that $\{a_n\}_{n=1}^{N^{1/2}}$ satisfies

$$a_{i+1} - a_i \sim \frac{1}{N}, \quad (a_{i+2} - a_{i+1}) - (a_{i+1} - a_i) \sim \frac{\theta}{N^2}, \tag{28}$$

where, here, $\sim$ means within a factor of 4. The parameter θ is in $(0, 1]$, Ω is the $L^2\theta/N^2$-neighborhood of $\{a_n\}_{n=1}^{N^{1/2}}$, and

$$\Omega = \bigsqcup_{I\in\mathcal{I}} I,$$

where each I is an $L^2\theta/N^2$-neighborhood of L consecutive terms in $\{a_n\}_{n=1}^{N^{1/2}}$.

We restate Theorem 1.2 but for all short generalized Dirichlet sequences with $\theta \in (0, 1]$.

Theorem 4.1. *Let Ω and $\mathcal{I}$ be defined as in the last paragraphs. Then for $2 \le p \le 6$ and every $\varepsilon > 0$ we have*

$$\Big\| \sum_{I\in\mathcal{I}} f_I \Big\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1} + 1) \Big(\sum_{I\in\mathcal{I}} \|f_I\|^2_{L^p(\mathbb{R})} \Big)^{\frac{1}{2}} \tag{29}$$

for functions f_I with supp $\widehat{f_I} \subset I$.

Comparing (29) with (3) we see an extra factor $\log^C(\theta^{-1} + 1)$. This factor appears as a consequence of dyadic pigeonholing in our proof.

4A. ***Local decoupling and refined decoupling inequalities.*** We first formulate a local decoupling inequality which implies (in fact is equivalent to) the global decoupling inequality (29).

Proposition 4.2. *Let $p \geq 2$. Suppose that, for some $k \geq 100$,*

$$\Big\| \sum_{I \in \mathcal{I}} f_I \Big\|_{L^p(P(L))} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\sum_{I \in \mathcal{I}} \| f_I \|^2_{L^p(W_{P(L),k})} \Big)^{\frac{1}{2}} \tag{30}$$

holds for every f_I with supp $\widehat{f_I} \subset I$. *Then (29) is true.*

Proof. Suppose (30) holds for some $k \geq 100$. Since $\sum_{P(L) \subset \mathbb{R}} W_{P(L),k} \lesssim_k 1$ and $p \geq 2$, by Minkowski's inequality, we have

$$\begin{aligned} \Big\| \sum_I f_I \Big\|^p_{L^p(\mathbb{R})} \leq \sum_{P(L) \subset \mathbb{R}} \int_{P(L)} |f|^p &\lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \sum_{P(L)} \Big(\sum_I \| f_I \|^2_{L^p(W_{P(L),k})} \Big)^{\frac{p}{2}} \\ &\lesssim N^\varepsilon \log^C(\theta^{-1}+1) \Big(\sum_I \| f_I \|^2_{L^p(\sum_{P(L)} W_{P(L),k})} \Big)^{\frac{p}{2}} \\ &\lesssim N^\varepsilon \log^C(\theta^{-1}+1) \Big(\sum_I \| f_I \|^2_{L^p(\mathbb{R})} \Big)^{\frac{p}{2}}, \end{aligned}$$

which is (29). □

The following local decoupling inequality will imply Theorem 4.1 by Proposition 4.2.

Theorem 4.3 (local decoupling). *Suppose $2 \leq p \leq 6$. Then*

$$\Big\| \sum_{I \in \mathcal{I}} f_I \Big\|_{L^p(P(L))} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\sum_{I \in \mathcal{I}} \| f_I \|^2_{L^p(W_{P(L),100})} \Big)^{\frac{1}{2}} \tag{31}$$

for f_I with supp $\widehat{f_I} \subset I$.

Theorem 4.3 is a consequence of the following refined decoupling theorem, which we focus on proving in the next two sections. The analogous result for the parabola can be found in [Guth et al. 2020; Demeter et al. 2020]. We will show how Theorem 4.4 implies Theorem 4.3 in Section 6E.

Theorem 4.4 (refined decoupling). *Suppose $2 \leq p \leq 6$. For every $P(L)$ and every $X \subset P(L)$, we have*

$$\Big\| \sum_I f_I \Big\|_{L^p(X)} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\sup_{x \in X} \sum_I \| f_I \|^2_{L^2(W_{P_I(x),100})} \Big)^{\frac{1}{2}-\frac{1}{p}} \Big(\sum_I \| f_I \|^2_{L^2(W_{P(L),100})} \Big)^{\frac{1}{p}} \tag{32}$$

for f_I with supp $\widehat{f_I} \subset I$.

We remark that Theorem 4.4 implies that for every $X \subset P$, where P is a fat AP larger than $P(L)$ in the sense that $P(L) \subset P$ for at least one $P(L)$, and, for $2 \leq p \leq 6$,

$$\Big\| \sum_I f_I \Big\|_{L^p(X)} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\sup_{x \in X} \sum_I \| f_I \|^2_{L^2(W_{P_I(x),100})} \Big)^{\frac{1}{2}-\frac{1}{p}} \Big(\sum_I \| f_I \|^2_{L^2(W_{P,100})} \Big)^{\frac{1}{p}} \tag{33}$$

for f_I with supp $\widehat{f_I} \subset I$. Indeed, (33) follows from taking (32) to the p-th power and summing over $X \cap P(L)$ with $P(L) \subset P$.

4B. ***Induction scheme for proving Theorem 4.4.*** We fix p, L and let $\mathrm{Dec}(N,\theta) = \mathrm{Dec}_p(N,L,\theta)$ denote the smallest constant such that

$$\left\|\sum_I f_I\right\|_{L^p(X)} \le \mathrm{Dec}(N,\theta)\left(\sup_{x\in X}\sum_I \|f_I\|^2_{L^2(W_{P_I(x),100})}\right)^{\frac12-\frac1p}\left(\sum_I \|f_I\|^2_{L^2(W_{P(L),100})}\right)^{\frac1p} \tag{34}$$

holds for every sequence $\{a_n\}_{n=1}^{N^{1/2}}$ satisfying (14), every $P(L)$, every $X \subset P(L)$, and every f_I with $\operatorname{supp}\widehat{f_I} \subset I$. For a specific choice of the short generalized Dirichlet sequence $\{a_n\}_{n=1}^{N^{1/2}}$ satisfying (14) we will call the smallest constant the refined decoupling constant of $\{a_n\}_{n=1}^{N^{1/2}}$ such that (34) holds for every $X \subset P(L)$, and every f_I with $\operatorname{supp}\widehat{f_I} \subset I$. Note that $\mathrm{Dec}_p(N,L,\theta)$ is the supremum of all refined decoupling constants of sequences $\{a_n\}_{n=1}^{N^{1/2}}$ satisfying (14).

We will deduce Theorem 4.4, which now is equivalent to $\mathrm{Dec}(N,\theta) \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1)$, from the following main proposition.

Proposition 4.5. *For every $\varepsilon > 0$ and every $1 \le K \le N^{\varepsilon/2}$ satisfying $N^{1/2}/K \ge L$,*

$$\mathrm{Dec}(N,\theta) \lesssim_\varepsilon \sup_{\theta'\in[\theta/4,\theta]} \mathrm{Dec}\left(\frac{N}{K^2},\frac{\theta'}{K^2}\right) + K^D N^\varepsilon \log^D(\theta^{-1}+1). \tag{35}$$

Here D is an absolute constant.

We postpone the proof of Proposition 4.5 to Section 6. Here we show how it implies Theorem 4.4.

Proof of Theorem 4.4 assuming Proposition 4.5. For some sufficiently large S_0 we have $\mathrm{Dec}(N,\theta) \le C_s N^s \le C_s N^s \log^D(\theta^{-1}+1)$ for $s \ge S_0$. Now suppose $\mathrm{Dec}(N,\theta) \le C_s N^s \log^D(\theta^{-1}+1)$ for some $s \le S_0$. Then from (35) we have, for every $\varepsilon > 0$ and K with $N^{1/2}/K \ge L$,

$$\begin{aligned}
\mathrm{Dec}(N,\theta) &\le C_\varepsilon\left(\sup_{\theta'\in[\theta/4,\theta]} C_s\left(\frac{N}{K^2}\right)^s \log^D(K^2(\theta')^{-1}+1) + K^D N^\varepsilon \log^D(\theta^{-1}+1)\right).\\
&\le C_\varepsilon\left(C C_s\left(\frac{N}{K^2}\right)^s \log^D(K^2\theta^{-1}+1) + K^D N^\varepsilon \log^D(\theta^{-1}+1)\right)\\
&\le C_\varepsilon\left(C C_s\left(\frac{N}{K^2}\right)^s (C\log^D(\theta^{-1}+1) + C\log^D(K^2)) + K^D N^\varepsilon \log^D(\theta^{-1}+1)\right).
\end{aligned}$$

If we choose ε to be $s/2$ and let $N^s/K^{2s} = K^D N^\varepsilon = K^D N^{s/2}$, that is, $K = N^{s/(2(2s+D))}$, then for some constant C_s' depending only on s,

$$\mathrm{Dec}(N,\theta) \le C_s' N^{s\left(1-\frac{1}{2s+D}\right)}(\log^D(\theta^{-1}+1) + \log^D N)$$

if $N^{1/2}N^{-s/(2(2s+D))} \ge L$. If $N^{1/2}N^{-s/(2(2s+D))} \le L$, then $|\mathcal{I}| \lesssim N^{s/(2(2s+D))}$ and by the triangle inequality and Cauchy–Schwarz inequality we have

$$\mathrm{Dec}(N,\theta) \lesssim N^{\frac{s}{2(2s+D)}}.$$

We can assume that D is large enough such that $\max\{2, S_0\} \le D$. Then $1/(2s+D) \sim D^{-1}$ and $K \le N^{\varepsilon/2}$, so for some absolute constant $c > 0$,

$$\mathrm{Dec}(N,\theta) \lesssim_s N^{s(1-c)} \log^D(\theta^{-1}+1).$$

Therefore we conclude

$$\mathrm{Dec}(N,\theta) \lesssim_\varepsilon N^\varepsilon \log^D(\theta^{-1}+1)$$

for every $\varepsilon > 0$, since $S_0(1-c)^m$ is arbitrarily small for large enough m. □

4C. ***Two applications.*** Before ending this section, we record two applications of Theorem 4.1. Technically these are corollaries of the $\ell^2 L^6$ decoupling inequality for the parabola in [Bourgain and Demeter 2015], by deriving the corresponding (ℓ^2, L^6) estimate on short generalized Dirichlet polynomials using the method described in the Appendix.

First we may estimate approximate solutions to the equation $a_{n_1} + a_{n_2} + a_{n_3} = a_{n_4} + a_{n_5} + a_{n_6}$ for a short generalized Dirichlet sequence $\{a_n\}_{n=1}^{N^{1/2}}$. The number of exact solutions of such equations for general convex sequences was studied in [Iosevich et al. 2006].

Corollary 4.6. *Let $\{a_n\}_{n=1}^{N^{1/2}}$ be a short generalized Dirichlet sequence with parameter $\theta \in (0,1]$. Then*

$$\#\{(a_{n_1},\dots,a_{n_6}) : 1 \le n_i \le N^{\frac{1}{2}}, |(a_{n_1}+a_{n_2}+a_{n_3}) - (a_{n_4}+a_{n_5}+a_{n_6})| \le \theta/N^2\} \lesssim_\varepsilon \log^C(\theta^{-1}+1) N^{\frac{3}{2}+\varepsilon}. \quad (36)$$

This estimate is sharp up to $C_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1)$ due to $N^{3/2}$ many diagonal solutions.

In particular if we take $a_n = \log(n+N+1)$ in the above corollary, then $\theta \sim 1$ and (36) reads

$$\#\{(n_1,\dots,n_6) : N+1 \le n_i \le N+N^{\frac{1}{2}}, |n_1n_2n_3 - n_4n_5n_6| \lesssim N\} \lesssim_\varepsilon N^{\frac{3}{2}+\varepsilon}. \quad (37)$$

We note that the triple products $n_1n_2n_3$ with $N+1 \le n_1,n_2,n_3 \le N+N^{1/2}$ lies in the interval $[N^3, N^3+CN^{5/2}]$. So (37) implies that the triple products $\{n_1n_2n_3 : N+1 \le n_1,n_2,n_3 \le N+N^{1/2}\}$ are roughly evenly distributed in $[N^3, N^3+CN^{5/2}]$ with cN separation. Indeed if we split the interval $[N^3, N^3+CN^{5/2}]$ into intervals of length cN and let E_λ denote the number of cN-intervals which contain at least λ many triple products $n_1n_2n_3$, then (37) says that

$$\lambda^2 E_\lambda \le C_\varepsilon N^{\frac{3}{2}+\varepsilon}.$$

Consequently if we choose $\lambda \ge 10C_\varepsilon N^\varepsilon$, then we have $\lambda E_\lambda \le \frac{9}{10}N^{3/2}$, and λE_λ is the number of triple products $n_1n_2n_3$ that lie in a cN-interval which contains at least λ many triple products. The total number of triple products is $N^{3/2}$ so we can conclude most of the triple products lie in cN-intervals, each of which contains few triple products.

Proof of Corollary 4.6. We let ϕ be a Schwartz function whose Fourier transform is given by a smooth bump function adapted to $B_{\theta/N^2}(0)$:

$$\widehat{\phi} = 1 \quad \text{on } B_{\theta/N^2}(0), \quad \operatorname{supp}\widehat{\phi} \subset B_{2\theta/N^2}(0), \quad 0 \le \widehat{\phi} \le 1, \quad \widehat{\phi} \text{ is even.}$$

Applying Theorem 4.1 with $p=6, L=1$ we obtain

$$\int_{\mathbb{R}} \Big|\sum_{n=1}^{N^{1/2}} e^{ia_nx}\phi(x)\Big|^6 \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1)\Big(\sum_{n=1}^{N^{1/2}} \|e^{ia_nx}\phi(x)\|_{L^6(\mathbb{R})}^2\Big)^3 \lesssim N^\varepsilon \log^C(\theta^{-1}+1) N^{\frac{3}{2}}\theta^5 N^{-10}. \quad (38)$$

We expand the left-hand side of (38) as

$$\int_{\mathbb{R}}\Big|\sum_{n=1}^{N^{1/2}} e^{ia_n x}\phi(x)\Big|^6 dx = \sum_{n_1,\dots,n_6}\int_{\mathbb{R}} e^{i(a_{n_1}+a_{n_2}+a_{n_3}-a_{n_4}-a_{n_5}-a_{n_6})x}|\phi|^6\,dx$$
$$= \sum_{n_1,\dots,n_6}\widehat{|\phi|^6}(a_{n_1}+a_{n_2}+a_{n_3}-a_{n_4}-a_{n_5}-a_{n_6}).$$

Since $\widehat{\phi}$ is even we know that ϕ is real-valued and hence $\widehat{|\phi|^6} = \widehat{\phi}*\cdots*\widehat{\phi}$ is nonnegative and $\widehat{|\phi|^6} \gtrsim \theta^5 N^{-10}$ on $B_{c\theta/N^2}(0)$ for some small absolute constant $c>0$. Therefore

$$\int_{\mathbb{R}}\Big|\sum_{n=1}^{N^{1/2}} e^{ia_n x}\phi(x)\Big|^6$$
$$\gtrsim \theta^5 N^{-10}\#\{(a_{n_1},\dots,a_{n_6}) : 1\le n_i\le N^{\frac12}, |(a_{n_1}+a_{n_2}+a_{n_3})-(a_{n_4}+a_{n_5}+a_{n_6})|\le \theta/N^2\}.$$

Combining the above estimate and (38) we obtain (36). □

Another application of Theorem 4.1 is estimating the size of the intersection of an AP with a generalized Dirichlet sequence.

Corollary 4.7. *Let $\{a_n\}_{n=1}^N$ be a generalized Dirichlet sequence with parameter $\theta\in(0,1]$ and let $a=N^{-\alpha}$ with $\alpha\in[0,2]$. Then*

$$|\{a_n\}_{n=1}^{n=N}\cap a\mathbb{Z}| \lesssim \begin{cases} N^{\alpha} & \text{if } \alpha\in\left[0,\frac12\right],\\ C_\varepsilon N^{\varepsilon}\log^C(\theta^{-1}+1)N^{\frac13+\frac{\alpha}{3}} & \text{if } \alpha\in\left[\frac12,2\right].\end{cases}$$

When $\theta=1$, Corollary 4.7 is sharp for $\alpha\in\left[0,\frac12\right]$ (see Lemma 7.3), but we do not know if it is sharp for $\alpha\in\left[\frac12,2\right]$. Corollary 4.7 has a slight connection to a conjecture of Rudin which states in an N-term AP we can find at most $\mathcal{O}(N^{1/2})$ many squares (numbers of the form n^2 for some $n\in\mathbb{Z}$). The best result so far seems to be in [Bombieri and Zannier 2002], which proves at most $\mathcal{O}(N^{3/5}\log^{\mathcal{O}(1)}N)$ many squares can be found in an N-term AP. We note that $\{n^2/N^2\}_{n=N+1}^{2N}$ is a generalized Dirichlet sequence. However we shall not expect to solve Rudin's conjecture exploiting only the convexity of the sequence $\{n^2 : n\in\mathbb{N}\}$, as shown by the example given in Lemma 7.3.

Proof of Corollary 4.7. The case $\alpha\in\left[0,\frac12\right]$ is trivial as $\{a_n\}_{n=1}^N$ is contained in a ball of radius $\lesssim 1$ and $a\mathbb{Z}$ has at most $\lesssim a^{-1}=N^{\alpha}$ many terms in such a ball. Now we suppose $\alpha\in\left[\frac12,2\right]$. It suffices to show that, for a short generalized Dirichlet sequence $\{a_n\}_{n=1}^{N^{1/2}}$, $H := |\{n : 1\le n\le N^{1/2},\ a_n\in a\mathbb{Z}\}|$ satisfies

$$H \lesssim_\varepsilon C_\varepsilon\log^C(\theta^{-1}+1)N^{\frac{\alpha}{3}-\frac16+\varepsilon}.$$

We consider the function

$$f(x) = \sum_{n:1\le n\le N^{1/2},\, a_n\in a\mathbb{Z}} e^{2\pi i t a_n}.$$

Case 1: $\alpha\in[1,2]$. We apply Theorem 4.3 with $p=6$, $L=1$ and $P(L)=P(L,0)$. Since $|f|\ge H/10$ on $\mathcal{N}_{cN^{1/2}}(a^{-1}\mathbb{Z})$ with $c\gtrsim 1$, we obtain

$$H\left(\frac{N^2\theta^{-1}}{N^{\alpha}}N^{\frac12}\right)^{\frac16} \lesssim_\varepsilon N^{\varepsilon}\log^C(\theta^{-1}+1)H^{\frac12}(N^2\theta^{-1})^{\frac16},$$

where we used that $P(L)$ is approximately an $N^2\theta^{-1}$ interval. Simplifying the above displayed math,

$$H \lesssim_\varepsilon C_\varepsilon \log^C(\theta^{-1}+1) N^{\frac{\alpha}{3}-\frac{1}{6}+\varepsilon}.$$

Case 2: $\alpha \in \left[\frac{1}{2}, 1\right]$. We apply Theorem 4.3 with $p=6$, $L = N^{1-\alpha}$ and $P(L) = P(L,0)$. Since $|f| \geq H/10$ on $\mathcal{N}_{cN^{1/2}}(a^{-1}\mathbb{Z})$ with $c \gtrsim 1$, we obtain

$$H\left(\frac{N^{2\alpha}2\theta^{-1}}{N^\alpha} N^{\frac{1}{2}}\right)^{\frac{1}{6}} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) H^{\frac{1}{2}} (N^{2\alpha}2\theta^{-1})^{\frac{1}{6}},$$

that is,

$$H \lesssim_\varepsilon C_\varepsilon \log^C(\theta^{-1}+1) N^{\frac{\alpha}{3}-\frac{1}{6}+\varepsilon}. \qquad \square$$

5. High-low frequency decomposition for the square function

The proof of Proposition 4.5 is based on the method in [Guth et al. 2022], which uses a high-low frequency decomposition for the square function. Such a decomposition is also used in [Guth et al. 2019] to study incidence estimates for tubes. We set up the preliminaries in this section and prove Proposition 4.5 in Section 6. We begin in Section 5A with an overview of the argument, at a symbolic and heuristic level, and refer readers to Section 2 of [Guth et al. 2022] for a more detailed description of the intuition behind this method.

5A. ***Overview of the argument.*** Let $2 \leq p \leq 6$. We will present a heuristic overview of the high-low proof of Theorem 4.4 (which is our goal to prove via Proposition 4.5). By a pigeonholing argument, we may assume that there is a parameter $\alpha > 0$ so that

$$\int_X \left|\sum_I f_I\right|^p \sim \alpha^p |U_\alpha|,$$

where $U_\alpha = \{x \in X : \left|\sum_I f_I(x)\right| \sim \alpha\}$. A "broad/narrow" argument (written in our context in Section 6A) roughly allows us to reduce to the case that, on most of U_α, $\left|\sum_I f_I\right|$ is bounded by a bilinear expression $\left|\sum_{I_1 \subset I'} f_{I_1} \sum_{I_2 \subset I''} f_{I_2}\right|^{1/2}$ where I', I'' are transverse, meaning $d(I', I'') \gtrsim N^{-1/2}$. The high-low frequency proof of decoupling involves upgrading the bilinear restriction theorem (Proposition 3.5) to the refined decoupling theorem (Theorem 4.4).

We split U_α into $\leq \varepsilon^{-1}$ many sets on which we know certain square functions are high- or low-frequency dominated. Consider scales $1 \leq L \leq L_{m+1} \leq L_m \leq N^{1/2}$, where $L_m/L_{m+1} \leq N^\varepsilon$. Define the (square of the) square functions $g_m = \sum_{I_m} |f_{I_m}|^2$, $g_{m+1} = \sum_{I_{m+1}} |f_{I_{m+1}}|^2$, where I_m, I_{m+1} are unions of L_m, L_{m+1} many consecutive intervals in Ω, respectively. Also write $g = \sum_I |f_I|^2$. Suppose that on most of U_α, $g_{m+1}(x) \lesssim g(x)$. Observe the pointwise inequality that, for $x \in U_\alpha$ satisfying $g_{m+1}(x) \lesssim g(x)$,

$$\begin{aligned}
\alpha \sim \left|\sum_{I_{m+1}} f_{I_{m+1}}(x)\right| &\lesssim \sum_{I_{m+1}: |f_{I_{m+1}}(x)| > N^\varepsilon \frac{g(x)}{\alpha}} |f_{I_{m+1}}(x)| + \left|\sum_{I_{m+1}: |f_{I_{m+1}}(x)| \leq N^\varepsilon \frac{g(x)}{\alpha}} f_{I_{m+1}}(x)\right| \\
&\lesssim \frac{\alpha}{N^\varepsilon g(x)} \sum_{I_{m+1}: |f_{I_{m+1}}(x)| > N^\varepsilon \frac{g(x)}{\alpha}} |f_{I_{m+1}}(x)|^2 + \left|\sum_{I_{m+1}: |f_{I_{m+1}}(x)| \leq N^\varepsilon \frac{g(x)}{\alpha}} f_{I_{m+1}}(x)\right| \\
&\lesssim \frac{\alpha}{N^\varepsilon g(x)} g_{m+1}(x) + \left|\sum_{I_{m+1}: |f_{I_{m+1}}(x)| \leq N^\varepsilon \frac{g(x)}{\alpha}} f_{I_{m+1}}(x)\right|.
\end{aligned}$$

This type of reasoning means that on most of U_α we may perform a wave packet decomposition of f at scale I_{m+1} and replace f with a version which only preserves the "small" wave packets, ensuring the property that $\|f_{I_{m+1}}\|_{L^\infty(\mathbb{R})} \lessapprox \|g\|_{L^\infty(X)}/\alpha$.

Case 1: high-dominance. Suppose that on most of U_α, $g_m(x) \lesssim |g_m * \check{\eta}_{\geq L_{m+1}/N}(x)|$, where $\eta_{\geq L_{m+1}/N}$ is a smooth bump function with support in $L_{m+1}/N \leq |\omega| \leq 2$, on most of U_α. A combination of a broad-narrow argument, Proposition 3.5, the locally constant property, and the assumption of high-frequency dominance of g_m leads to the inequality

$$\alpha^4 |U_\alpha| \lesssim \int |g_m * \check{\eta}_{\geq L_{m+1}/N}|^2.$$

Next, by Plancherel's theorem, we analyze the integral on the right-hand side. A geometric argument shows that the supports of the $\widehat{|f_{I_m}|^2}$ from $\hat{g}_m$ are sparsely overlapping on the support of $\eta_{\geq L_{m+1}/N}$. This allows us to bound the right-hand side of the previous displayed inequality by

$$C_\varepsilon N^\varepsilon \sum_{I_m} \int |f_{I_m}|^4,$$

which is bounded by $C_\varepsilon N^{10\varepsilon} \sum_{I_{m+1}} \int |f_{I_{m+1}}|^4$ using Cauchy–Schwarz. Finally, use the good L^∞ bound for each $f_{I_{m+1}}$ from the pruning of the wave packets to get

$$\sum_{I_{m+1}} \int |f_{I_{m+1}}|^4 \lessapprox \frac{\|g\|^2_{L^\infty(X)}}{\alpha^2} \sum_{I_{m+1}} \int |f_{I_{m+1}}|^2.$$

A pigeonholing argument may be used to show that without loss of generality, we may assume that $\|g\|_{L^\infty(X)} \lesssim \alpha^2$. By L^2 orthogonality, the integral on the right-hand side of the previous displayed line equals $\sum_I \int |f_I|^2$. The conclusion of the argument in this case is then

$$|U_\alpha| \lessapprox \frac{\|g\|^2_{L^\infty(X)}}{\alpha^6} \sum_I \|f_I\|^2_{L^2(\mathbb{R})} \lesssim \frac{\|g\|^{\frac{p}{2}-1}_{L^\infty(X)}}{\alpha^p} \sum_I \int |f_I|^2,$$

which is a version of the statement of Theorem 4.4.

Case 2: low-dominance. The remaining case is if $g_m(x) \leq |g_m * \check{\eta}_{<L_{m+1}/N}(x)|$ on most of U_α. A local L^2-orthogonality argument shows that $|g_m * \check{\eta}_{<L_{m+1}/N}(x)|$ is bounded by $g_{m+1} * |\check{\eta}_{<L_{m+1}/N}|(x)$, which by the locally constant heuristic, is roughly the same as $g_{m+1}(x)$. We conclude in this case that, on most of U_α, $g_m(x) \lesssim g_{m+1}(x) \lesssim g(x)$. This is the same type of assumption we made before consider the cases, except at the scale L_m instead of L_{m+1}. This allows us to reinitiate the argument beginning with the assumption that $g_m(x) \lesssim g(x)$ in place of $g_{m+1}(x) \lesssim g(x)$.

In the case that we are "low"-dominated for ε^{-1} many scales, then

$$|U_\alpha| \approx |\{x \in U_\alpha : g_1(x) \lesssim g(x)\}|,$$

where g_1 is a square function corresponding to partitions of Ω into I_1, which are N^ε many adjacent intervals. Since $|\sum_I f_I(x)| \lesssim N^\varepsilon g_1(x)$ by Cauchy–Schwarz, the statement of Theorem 4.4 becomes trivial. In the next sections, we set up the argument in full technical detail.

5B. ***Wave-packet decomposition.*** We start with a few definitions. Write $f = \sum_{I\in\mathcal{I}} f_I$, where f_I will always denote a function with frequency support in I.

Fix $2 \le p \le 6$ and $\varepsilon > 0$. For $m \in \mathbb{N}$, let $L_m = N^{1/2}N^{-\varepsilon m}$. Without loss of generality we assume $L_M = L$ for some $M \in \mathbb{N}$. So $M \lesssim_\varepsilon 1$. For every $1 \le m \le M$ we let $\mathcal{I}_m$ be the partition of Ω into $N^{1/2}/L_m$ many I_m, each of which is the union of L_m-consecutive intervals in Ω. L_m can be thought of as scales.

Note that

$$I_m \subset P_{v_m}^{CL_m^2\theta/N^2} \cap B_{CL_m/N},$$

where $v_m \sim \frac{1}{N}$. We denote the right-hand side as $\widetilde{I}_m$:

$$\widetilde{I}_m := P_{v_m}^{CL_m^2\theta/N^2} \cap B_{CL_m/N}.$$

Let $\mathcal{P}_{I_m}$ be a tiling of $\mathbb{R}$ by P_{I_m}. For each I_m, we will now construct a partition of unity $\{\phi_{I_m}\}_{P_{I_m}\in\mathcal{P}_{I_m}}$ which will be used to perform the wave packet decomposition

$$f_{I_m} = \sum_{P_{I_m}} \phi_{P_{I_m}} f_{I_m}.$$

We regard each summand $\phi_{P_{I_m}} f_{I_m}$ as a wave packet. Specifically, we let ψ_{I_m} be adapted to $\widetilde{I}_m - \widetilde{I}_m$, which is of the form $P_{v_0}^{CL_m^2\theta/N^2}(0) \cap B_{CL_m/N}(0)$, in the frequency space as in Lemma 2.2, with order of decay 200 outside of the dual fat AP P_{I_m}. For each $P_{I_m} \in \mathcal{P}_{I_m}$, define

$$\phi_{P_{I_m}} := \|\psi_{I_m}^2\|_{L^1(\mathbb{R})}^{-1} \int_{P_{I_m}} |\psi_{I_m}(x-y)|^2\, dy. \tag{39}$$

Proposition 5.1 (wave-packet decomposition). *$\{\phi_{P_{I_m}}\}_{P_{I_m}\in\mathcal{P}_{L_m}}$ forms a partition of unity, that is, $\sum \phi_{P_{I_m}} = 1$, $\phi_{P_{I_m}} \ge 0$. Each $\phi_{P_{I_m}}$ is a translated copy of the others, and*

$$\operatorname{supp}\widehat{\phi}_{P_{I_m}} \subset 8^{400}(\widetilde{I}_m - \widetilde{I}_m), \quad 1_{P_{I_m}} \lesssim \phi_{P_{I_m}} \lesssim W_{P_{I_m},200}. \tag{40}$$

Proof. By definition we see that $\phi_{P_{I_m}}$ forms a partition of unity, and each $\phi_{P_{I_m}}$ is a translated copy of the others. Also it follows from the definition that

$$1_{P_{I_m}} \lesssim |\phi_{P_{I_m}}|.$$

Note that $\phi_{P(L_m)}$ equals $\|\psi_{I_m}^2\|_{L^1(\mathbb{R})}^{-1}|\psi_{I_m}|^2 * 1_{P_{I_m}}$. Therefore ψ_{I_m} decays at order 200 outside $P_{I_m}(0)$ implies that $\phi_{P(L_m)}$ decays at order 400 outside P_{I_m}, and in particular

$$|\phi_{P_{I_m}}| \lesssim W_{P_{I_m},200}.$$

The support property $\operatorname{supp}\widehat{\phi}_{P_{I_m}} \subset 8^{400}(\widetilde{I}_m - \widetilde{I}_m)$ follows from the fact that

$$\widehat{\phi}_{P_{I_m}} = \|\psi_{I_m}^2\|_{L^1(\mathbb{R})}^{-1}\widehat{|\psi_{I_m}|^2}\,\widehat{1}_{P_{I_m}}$$

and from Lemma 2.2. □

5C. ***A pruning process and modified square functions.*** Now we define "square functions" (squared) at scales L_m, which differ from the usual square functions by a pruning process of wave packets and taking

spatial averages. The pruning process will depend on two parameters α and r, which can be thought of as the values of $|f|$ and $\sum_{I_M} |f_{I_M}|^2 = \sum_I |f_I|^2$ which dominate the L^p norm of f. We define $\lambda = \lambda(\alpha, r)$ by

$$\lambda = \widetilde{C}_\varepsilon N^\varepsilon \frac{r}{\alpha}, \tag{41}$$

where $\widetilde{C}_\varepsilon$ is a sufficiently large constant depending on ε which will be chosen later in the proof of Lemma 5.4.

We first do the pruning process (with parameters α, r), which inductively removes wave packets at each scale whose height exceeds λ. As we shall see (Lemma 5.4), those wave packets do not play a dominant role in the L^p norm of f. This process produces a family of functions f_{m,I_m}, $f_{m,I_{m-1}}$, f_m that depend on α, r, which is implicit in the notation. We will write $f_{m,I_m,\alpha,r}$, $f_{m,I_{m-1},\alpha,r}$, $f_{m,\alpha,r}$ to emphasize such dependence when necessary.

Let $\mathcal{P}_{I_M,\lambda} = \{P_{I_M} \in \mathcal{P}_{L_M} : \|\phi_{P_{I_M}} f_{I_M}\|_{L^\infty(\mathbb{R})} \le \lambda\}$, and define

$$f_{M,I_M} := \sum_{P_{I_M} \in \mathcal{P}_{I_M,\lambda}} \phi_{P_{I_m}} f_{I_M}, \quad f_M := \sum_{I_M} f_{M,I_M}.$$

We let $f_{M,I_{M-1}} = \sum_{I_M \subset I_{M-1}} f_{M,I_M}$. Now we define f_m and f_{m,I_m} inductively for $m = 1, \dots, M-1$ by

$$f_{m,I_m} := \sum_{P_{I_m} \in \mathcal{P}_{I_m,\lambda}} \phi_{P_{I_m}} f_{m+1,I_m}, \quad f_m := \sum_{I_m} f_{m,I_m}, \tag{42}$$

where $f_{m+1,I_m} = \sum_{I_{m+1} \subset I_m} f_{m+1,I_{m+1}}$ and $\mathcal{P}_{I_m,\lambda} = \{P_{I_m} \in \mathcal{P}_{I_m} : \|\phi_{P_{I_m}} f_{m+1,I_m}\|_{L^\infty(\mathbb{R})} \le \lambda\}$. For notational convenience we also define $f_{M+1} = f$ and $f_{M+1,I_M} := f_{I_M} = f_I$.

We note that

(i) $f_m = \sum_{I_m} f_{m,I_m} = \sum_{I_{m-1}} f_{m,I_{m-1}}$,

(ii) $\operatorname{supp} \widehat{f_{m,I_m}} \subset C\widetilde{I}_m$,

(iii) $\operatorname{supp} \widehat{f_{m,I_{m-1}}} \subset C\widetilde{I}_{m-1}$,

(iv) $|f_{m,I_m}| \le |f_{m+1,I_m}|$ pointwise.

Item (i) follows from the definitions, and (iv) holds because $\{\phi_{P_{I_m}}\}_{P_{I_m}}$ is a partition of unity. To see (ii) and (iii) we may induct on m and note that

$$\bigcup_{I_m \subset I_{m-1}} C\widetilde{I}_m \subset 2\widetilde{I}_{m-1}$$

when N is sufficiently large depending on ε.

To define the "square function" g_m at scale L_m we introduce ρ_{I_m}, which is an L^1-normalized nonnegative function adapted to $P_{I_m}(0)$ with decay order 100

$$|P_{I_m}|^{-1} 1_{P_{I_m}(0)}(x) \lesssim \rho_{I_m}(x) \lesssim \frac{W_{P_{I_m}(0),100}(x)}{\|W_{P_{I_m}(0),100}\|_{L^1(\mathbb{R})}}, \tag{43}$$

and $\operatorname{supp} \widehat{\rho}_{I_m} \subset C(\widetilde{I}_m - \widetilde{I}_m)$. Such a function can be constructed by taking $|\psi|^2/\|\psi^2\|_{L^1}$ for ψ adapted to $\widetilde{I}_m$ with decay order 100 as in Lemma 2.2.

Finally we define the "square function" by

$$g_m := \sum_{I_m} |f_{m+1,I_m}|^2 * \rho_{I_m}$$

for $1 \le m \le M-1$ and for $m = M$ we define

$$g_M := \sum_{I_M} |f_{I_M}|^2 * \rho_{I_M}.$$

We note here that g_m for $1 \le m \le M-1$ implicitly depends on α, r, and we will write $g_{m,\alpha,r}$ to emphasize such dependence when necessary; g_M does not depend on α, r.

5D. ***High-low decomposition.*** To set up a high-low frequency decomposition for g_m, we let $\eta_m(\xi)$ be an even smooth bump function that equals to 1 on $B_{L_{m+1}/N}(0)$ and vanishes outside $B_{2L_{m+1}/N}(0)$ for every $1 \le m \le M-1$. We also assume that η_m are rescalings of each other.

Define, for $1 \le m \le M-1$,

$$g_m^\ell := g_m * \check{\eta}_m \quad \text{and} \quad g_m^h := g_m - g_m^\ell,$$

which are low- and high-frequency parts of g_m. Both g_m^ℓ and g_m^h satisfy some proprieties. We discuss them in the following two lemmas.

Lemma 5.2 (low lemma). *For $1 \le m \le M-1$, we have the pointwise inequality*

$$|g_m^\ell| \lesssim g_{m+1}.$$

Proof. By definition

$$g_m^\ell = \Big(\sum_{I_m} |f_{m+1,I_m}|^2\Big) * \rho_{I_m} * \check{\eta}_m = \Big(\sum_{I_m} |f_{m+1,I_m}|^2\Big) * \check{\eta}_m * \rho_{I_m}.$$

Using Plancherel's theorem,

$$\begin{aligned} |f_{m+1,I_m}|^2 * \check{\eta}_m(x) &= \int |f_{m+1,I_m}(y)|^2 \check{\eta}_m(x-y)\,dy \\ &= \int (\widehat{f}_{m+1,I_m} * \widehat{\overline{f}}_{m+1,I_m})(\xi)\, e^{2\pi i x\xi} \eta_m(\xi)\,d\xi \\ &= \sum_{I_{m+1}, I'_{m+1} \subset I_m} \int (\widehat{f}_{m+1,I_{m+1}} * \widehat{\overline{f}}_{m+1,I'_{m+1}})(\xi)\, e^{2\pi i x\xi} \eta_m(\xi)\,d\xi. \end{aligned} \tag{44}$$

We note that $\widehat{f}_{m+1,I_{m+1}} * \widehat{\overline{f}}_{m+1,I'_{m+1}}$ is supported in $C\widetilde{I}_{m+1} - C\widetilde{I}'_{m+1}$ and $\widetilde{I}_{m+1}$ is of the form $P_{v_{I_{m+1}}}^{CL^2\theta/N^2} \cap B_{CL_{m+1}/N}$. Since η_m is supported on $B_{2L_{m+1}/N}(0)$ we conclude that for every fixed I_{m+1} there are only $\mathcal{O}(1)$ many I'_{m+1} such that the integral in (44) is nonzero, and for those I'_{m+1} we write $I'_{m+1} \sim I_{m+1}$. We let $\psi_{I_{m+1}}$ be adapted to $C(\widetilde{I}_{m+1} - \widetilde{I}_{m+1})$ as in Lemma 2.2 with order of decay 200. Then, using Cauchy–Schwarz in the first two inequalities, we have

$$\begin{aligned} ||f_{m+1,I_m}|^2 * \check{\eta}_m(x)| &= \sum_{I_{m+1} \subset I_m} \sum_{I'_{m+1} \sim I_{m+1}} f_{m+1,I_{m+1}}\, \overline{f}_{m+1,I'_{m+1}} * \check{\eta}_m \\ &\le \sum_{I_{m+1} \subset I_m} \sum_{I'_{m+1} \sim I_{m+1}} (|f_{m+1,I_{m+1}}|^2 * |\check{\eta}_m|)^{\frac12} (|f_{m+1,I'_{m+1}}|^2 * |\check{\eta}_m|)^{\frac12} \end{aligned}$$

$$\lesssim \sum_{I_{m+1}\subset I_m} |f_{m+1,I_{m+1}}|^2 * |\check{\eta}_m|$$
$$\lesssim \sum_{I_{m+1}\subset I_m} |f_{m+1,I_{m+1}}|^2 * |\check{\psi}_{I_{m+1}}| * |\check{\eta}_m|$$
$$\leq \sum_{I_{m+1}\subset I_m} |f_{m+2,I_{m+1}}|^2 * |\check{\psi}_{I_{m+1}}| * |\check{\eta}_m|,$$

where the last inequality is because of $|f_{m+1,I_{m+1}}| \leq |f_{m+2,I_{m+1}}|$ pointwise. Now to finish the proof, it suffices to observe that

$$|\check{\eta}_m| * |\check{\psi}_{I_{m+1}}| * \rho_{I_m} \lesssim \rho_{I_{m+1}},$$

since $|\check{\eta}_m|$ decays rapidly outside $B_{N/L_{m+1}}(0)$, $|\check{\psi}_{I_{m+1}}|$ decays at order 200 outside $P_{I_{m+1}}(0)$, ρ_{I_m} decays at order 100 outside $P_{I_m}(0)$, and $B_{L_{m+1}/N}(0) + P_{I_m}(0) \subset CP_{I_{m+1}}(0)$. □

Recall that

$$P(L_m) = P_{v_1^{-1}}^{CN^{3/2}/L_m^2} \cap B_{CN^2/(L_m^2\theta)}$$

(which degenerates to $B_{CN^2/(L_m^2\theta)}$ if $L_m \leq CN^{1/4}$) as defined in (16). Let $\phi_{P(L_M)}$ be a function such that

$$\operatorname{supp}\widehat{\phi_{P(L_M)}} \subset P_{v_1}^{CL_M^2\theta/N^2}(0) \cap B_{CL_M^2/N^{3/2}}(0) \subset \bigcap_{I\in\mathcal{I}} (\widetilde{I}-\widetilde{I}), \quad \text{where } 1_{P(L_M)} \lesssim |\phi_{P(L_M)}| \lesssim W_{P(L_M),200}.$$

To construct such a function we can take a ψ in Lemma 2.2 adapted to certain fat AP and apply a translation in the physical space to it.

Lemma 5.3 (high lemma). *For $1 \leq m \leq M-1$ we have*

$$\int |g_m^h|^2 W_{P(L_M),100} \lesssim N^{\varepsilon} \int \sum_{I_m} |f_{m+1,I_m}|^4 W_{P(L_M),100}.$$

Proof. Because of (12), it suffices to show for every $P(L_M)$

$$\int_{P(L_M)} |g_m^h|^2 \lesssim N^{\varepsilon} \int \sum_{I_m} |f_{m+1,I_m}|^4 W_{P(L_M),100}.$$

Calculate

$$\int |g_m^h|^2 W_{P(L_M),100} \lesssim \int |g_m^h \phi_{P(L_M)}|^2 = \int \left| \sum_{I_m} \widehat{(|f_{m+1,I_m}|^2)} \hat{\rho}_{I_m}(1-\eta_m) * \widehat{\phi_{P(L_M)}} \right|^2.$$

Note that

$$\operatorname{supp}(\widehat{(|f_{m+1,I_m}|^2)} \hat{\rho}_{I_m}(1-\eta_m) * \widehat{\phi_{P(L_M)}}) \subset C(\widetilde{I}_m - \widetilde{I}_m) \setminus B_{L_{m+1}/(2N)}(0).$$

Indeed, the high-frequency cutoff $(1-\eta_m)$ removes the ball $B_{L_{m+1}/N}(0)$. The support of $\widehat{\phi_{P(L_M)}}$ is contained in a ball of radius $\leq \frac{1}{2}L_M^2/N^{3/2}$ (if the C in the definition of $P(L)$ as in (16) is large enough), so convolution with $\widehat{\phi_{P(L_M)}}$ shrinks the high-frequency cutoff by an amount smaller than $L_{m+1}/(2N)$. The structure of $\widetilde{I}_m - \widetilde{I}_m$ is unchanged by convolution by $\widehat{\phi_{P(L_M)}}$ because the thickness of $\widetilde{I}_m$ is $\sim L_m/N$

and $\frac{1}{2}L_M^2/N^{3/2} \le L_{m+1}/(2N) \le N^{-\varepsilon}L_m/N$. We claim that at every point on $\mathbb{R}$, the collection of sets $\{C(\widetilde{I}_m - \widetilde{I}_m) \setminus B_{L_{m+1}/(2N)}(0)\}_{I_m}$ has at most $\mathcal{O}(N^\varepsilon)$ overlap. Assuming this claim, by the Cauchy–Schwarz inequality we obtain

$$\int |g_m^h|^2 W_{P(L_M),100} \lesssim N^\varepsilon \int \sum_{I_m} \big|\widehat{(|f_{m+1,I_m}|^2)}\widehat{\rho}_{I_m}(1-\eta_m) * \widehat{\phi_{P(L_M)}}\big|^2.$$

So we have

$$\begin{aligned}
&\int |g_m^h|^2 W_{P(L_M),100} \\
&\quad\lesssim N^\varepsilon \sum_{I_m} \int \big||f_{m+1,I_m}|^2 * \rho_{I_m} * \widehat{(1-\eta_m)}\big|^2 |\phi_{P(L_M)}|^2 \\
&\quad\lesssim N^\varepsilon \sum_{I_m} \bigg(\int \big||f_{m+1,I_m}|^2 * \rho_{I_m}\big|^2 |\phi_{P(L_M)}|^2 + \int \big||f_{m+1,I_m}|^2 * \rho_{I_m} * |\check{\eta}_m|\big|^2 |\phi_{P(L_M)}|^2 \bigg) \\
&\quad\lesssim N^\varepsilon \sum_{I_m} \bigg(\int |f_{m+1,I_m}|^4 (|\phi_{P(L_M)}|^2 * \rho_{I_m}) + \int |f_{m+1,I_m}|^4 (|\phi_{P(L_M)}|^2 * \rho_{I_m} * |\check{\eta}_m|) \bigg),
\end{aligned}$$

where we used Cauchy–Schwarz and that ρ_{I_m} and $\check{\eta}_m$ have L^1 norms ~ 1 to justify

$$\big||f_{m+1,I_m}|^2 * \rho_{I_m}\big|^2 \lesssim |f_{m+1,I_m}|^4 * \rho_{I_m}, \quad \big||f_{m+1,I_m}|^2 * \rho_{I_m} * |\check{\eta}_m|\big|^2 \lesssim |f_{m+1,I_m}|^4 * \rho_{I_m} * |\check{\eta}_m|.$$

Noting that $|\phi_{P(L_M)}|^2 * \rho_{I_m} \lesssim W_{P(L_M),100}$ and $|\phi_{P(L_M)}|^2 * \rho_{I_m} * |\check{\eta}_m| \lesssim W_{P(L_M),100}$, we conclude

$$\int |g_m^h|^2 W_{P(L_M),100} \lesssim N^\varepsilon \sum_{I_m} \int |f_{m+1,I_m}|^4 W_{P(L_M),100}.$$

Now we prove the claim. Recall that $\widetilde{I}_m$ is a fat AP of the form $P_{v_{I_m}}^{CL^2\theta/N^2} \cap B_{CL_m/N}$, where $v_{I_m} \sim N^{-1}$. Suppose $x \in C(\widetilde{I}_m - \widetilde{I}_m) \setminus B_{L_{m+1}/(2N)}(0)$ and $x \in C(\widetilde{I}'_m - \widetilde{I}'_m) \setminus B_{L_{m+1}/N}(0)$ for distinct $\widetilde{I}_m$ and $\widetilde{I}'_m$. We denote the common difference of $\widetilde{I}_m$ and $\widetilde{I}'_m$ by v and v' respectively. Recalling that v_{I_m} are $C\theta L_m/N^2$ separated, and the maximal separation is $C(N^{1/2}/L_m)(\theta L_m/N^2) = C\theta/N^{3/2}$, we have

$$\theta L_m/N^2 \lesssim |v - v'| \lesssim \theta/N^{\frac{3}{2}}.$$

Suppose $x \in B_{CL_m^2\theta/N^2}(kv)$ and $x \in B_{CL_m^2\theta/N^2}(k'v')$ for some $k, k' \in \mathbb{N}$. Then since $x \notin B_{L_{m+1}/(2N)}(0)$, $L_{m+1} \lesssim k, k' \lesssim L_m$. By definition $L_m = N^\varepsilon L_{m+1} \le N^{1/2-\varepsilon}$, so we have

$$L_{m+1}\frac{\theta L_m}{N^2} \gtrsim N^{-\varepsilon}\frac{\theta L_m^2}{N^2}, \quad L_m\frac{\theta}{N^{\frac{3}{2}}} \le \frac{\theta}{N^{1+\varepsilon}} \le \frac{1}{N^{1+\varepsilon}}.$$

It follows that $|k - k'| \lesssim 1$ and

$$\text{either} \quad |v - v'| \lesssim N^\varepsilon \theta \frac{L_m}{N^2} \quad \text{or} \quad |v - v'| \gtrsim \frac{1}{N^{\frac{3}{2}-\varepsilon}}.$$

The second case cannot happen if N is sufficiently large (depending on ε). Since common differences v are $\mathcal{O}(\theta L_m/N^2)$-separated, we conclude that there are at most $\mathcal{O}(N^\varepsilon)$ many $\widetilde{I}'_m$ such that $x \in C(\widetilde{I}'_m - \widetilde{I}'_m) \setminus B_{L_{m+1}/(2N)}(0)$. □

5E. ***The sets $\Omega_{m,\alpha,r}$ and $U_{\alpha,r}$.*** The last part of our high-low decomposition set-up is to partition $P(L_M)$ into $\Omega_{m,\alpha,r}$, for a fixed pair (α, r). For $1 \le m \le M-1$ we define $\Omega_{m,\alpha,r}$ to be

$$\Omega_{m,\alpha,r} := \{x \in P(L_M) : g_m(x) \le 2|g_m^h(x)|, g_{m+1}(x) \le 2|g_{m+1}^\ell(x)|, \dots, g_{M-1}(x) \le 2|g_{M-1}^\ell(x)|\}.$$

Here $g_k = g_{k,\alpha,r}$. Also define $\Omega_{0,\alpha,r}$ to be

$$\Omega_{0,\alpha,r} := \{x \in P(L_M) : g_1(x) \le 2|g_1^\ell(x)|, g_2(x) \le 2|g_2^\ell(x)|, \dots, g_{M-1}(x) \le 2|g_{M-1}^\ell(x)|\}.$$

Clearly

$$P(L_M) = \bigcup_{0 \le m \le M-1} \Omega_{m,\alpha,r}$$

for every α, r. For notational convenience we let $\Omega_{M,\alpha,r} = P(L_M)$.

We define $U_{\alpha',r'}$ by

$$U_{\alpha',r'} := \{x \in P(L_M) : r'/2 < g_M(x) \le 2r',\ \alpha'/2 < |f(x)| \le 2\alpha'\}. \tag{45}$$

Recall that $g_M = \sum_{I_M} |f_{I_M}|^2 * \rho_{I_M}$ is defined without the pruning process so in particular it does not depend on the pruning parameters α, r.

We prove the following lemma, which shows that, on $U_{\alpha,r} \cap \Omega_{m,\alpha,r}$, $|f_m - f_{m,\alpha,r}|$ is very small so that $|f_m| \sim |f_{m,\alpha,r}|$. We define $f_0 = f_1$ for notational convenience. Also recall we have defined $f_{M+1} = f$ and $f_{M+1,I_M} = f_{I_M} = f_I$.

Lemma 5.4. *If the constant $\widetilde{C}_\varepsilon$ in the definition of λ is large enough depending on ε, then for every α, r, every $1 \le m \le M-1$, and any subset $\mathcal{S}$ of the partition $\mathcal{I}_m = \{I_m\}$, we have*

$$\left| \sum_{I_m \in \mathcal{S}} f_{I_m} - \sum_{I_m \in \mathcal{S}} f_{m,\alpha,r,I_m} \right| \le \frac{\alpha}{100}$$

on $U_{\alpha,r} \cap \Omega_{m,\alpha,r}$, and also on $U_{\alpha,r} \cap \Omega_{0,\alpha,r}$ if $m = 1$. In particular if $\widetilde{C}_\varepsilon$ in the definition of λ is large enough depending on ε, then for every α, r, every $0 \le m \le M-1$,

$$|f_{m,\alpha,r}| \in \left[\frac{\alpha}{4}, 4\alpha\right],$$

on $U_{\alpha,r} \cap \Omega_{m,\alpha,r}$.

Proof. Fix α, r. In the following proof g_k means $g_{k,\alpha,r}$, and f_{k,I_k}, $f_{k,I_{k-1}}$, f_k mean $f_{k,I_k,\alpha,r}$, $f_{k,I_{k-1},\alpha,r}$, $f_{k,\alpha,r}$ respectively. First suppose $1 \le m \le M-1$. By the definition of $\Omega_{m,\alpha,r}$ and Lemma 5.2 we know that on $U_{\alpha,r} \cap \Omega_{m,\alpha,r}$,

$$g_{m+1} \lesssim g_{m+2} \lesssim \cdots \lesssim g_M \lesssim r.$$

We also have by the Cauchy–Schwarz inequality $g_m \lesssim_\varepsilon N^\varepsilon g_{m+1}$. Recall that $M \lesssim_\varepsilon 1$ so we have, for $m \le k \le M$,

$$g_k \lesssim_\varepsilon N^\varepsilon r \quad \text{on } U_{\alpha,r} \cap \Omega_{m,\alpha,r}.$$

Let m' be an integer between m and M and let $I_{m'} \in \mathcal{I}_{m'}$. By the definition of $f_{m',I_{m'}}$ and $f_{m'+1,I_{m'}}$ we have for $x \in U_{\alpha,r} \cap \Omega_{m,\alpha,r}$

$$\begin{aligned}
|f_{m',I_{m'}}(x) - f_{m'+1,I_{m'}}(x)| &= \Big| \sum_{P_{I_{m'}} \notin \mathcal{P}_{I_{m'},\lambda}} \phi_{P_{I_{m'}}}(x) f_{m'+1,I_{m'}}(x) \Big| \\
&\lesssim \sum_{P_{I_{m'}} \notin \mathcal{P}_{I_{m'},\lambda}} |\phi^{\frac{1}{2}}_{P_{I_{m'}}}(x) f_{m'+1,I_{m'}}(x)| \phi^{\frac{1}{2}}_{P_{I_{m'}}}(x) \\
&\lesssim \sum_{P_{I_{m'}} \notin \mathcal{P}_{I_{m'},\lambda}} \lambda^{-1} \|\phi_{P_{I_{m'}}} f_{m'+1,I_{m'}}\|_{L^\infty(\mathbb{R})} \|\phi^{\frac{1}{2}}_{P_{I_{m'}}} f_{m'+1,I_{m'}}\|_{L^\infty(\mathbb{R})} \phi^{\frac{1}{2}}_{P_{I_{m'}}}(x) \\
&\lesssim \lambda^{-1} \sum_{P_{I_{m'}} \notin \mathcal{P}_{I_{m'},\lambda}} \|\phi^{\frac{1}{2}}_{P_{I_{m'}}} f_{m'+1,I_{m'}}\|^2_{L^\infty(\mathbb{R})} \phi^{\frac{1}{2}}_{P_{I_{m'}}}(x) \\
&\lesssim \lambda^{-1} \sum_{P_{I_{m'}} \notin \mathcal{P}_{I_{m'},\lambda}} \sum_{\widetilde{P}_{I_{m'}}} \|\phi_{P_{I_{m'}}} f^2_{m'+1,I_{m'}}\|_{L^\infty(\widetilde{P}_{I_{m'}})} \phi^{\frac{1}{2}}_{P_{I_{m'}}}(x) \\
&\lesssim \lambda^{-1} \sum_{P_{I_{m'}}} \sum_{\widetilde{P}_{I_{m'}}} \|\phi_{P_{I_{m'}}}\|_{L^\infty(\widetilde{P}_{I_{m'}})} \|f^2_{m'+1,I_{m'}}\|_{L^1(W_{\widetilde{P}_{I_{m'}}})} \phi^{\frac{1}{2}}_{P_{I_{m'}}}(x),
\end{aligned}$$

where we used $\phi_{P_{I_{m'}}} \lesssim \phi^{1/2}_{P_{I_{m'}}}$. We also used the locally constant property Proposition 2.3 for the last inequality. If we use $\phi_{I_{m'}}(\widetilde{P}_{I_{m'}})$ to denote $\phi_{I_{m'}}(\sup \widetilde{P}_{I_{m'}})$, which is comparable to $\phi_{I_{m'}}(y)$ for any $y \in \widetilde{P}_{I_{m'}}$, then we have

$$\begin{aligned}
|f_{m',I_{m'}}(x) - f_{m'+1,I_{m'}}(x)| &\lesssim \lambda^{-1} |P_{I_{m'}}|^{-1} \sum_{P_{I_{m'}}} \sum_{\widetilde{P}_{I_{m'}}} \Big(\int W_{\widetilde{P}_{I_{m'}}} \phi_{P_{I_{m'}}}(\widetilde{P}_{I_{m'}}) |f_{m'+1,I_{m'}}|^2 \Big) \phi^{\frac{1}{2}}_{P_{I_{m'}}}(x) \\
&\lesssim \lambda^{-1} |P_{I_{m'}}|^{-1} \sum_{\widetilde{P}_{I_{m'}}} \Big(\int W_{\widetilde{P}_{I_{m'}}} |f_{m'+1,I_{m'}}|^2 \Big) \phi^{\frac{1}{2}}_{\widetilde{P}_{I_{m'}}}(P_{I_{m'}}(x)) \\
&\lesssim \lambda^{-1} |P_{I_{m'}}|^{-1} \int |f_{m'+1,I_{m'}}|^2(y) \sum_{\widetilde{P}_{I_{m'}}} W_{\widetilde{P}_{I_{m'}}}(y)\, \phi^{\frac{1}{2}}_{\widetilde{P}_{I_{m'}}}(P_{I_{m'}}(x))\, dy \\
&\lesssim \lambda^{-1} |P_{I_{m'}}|^{-1} \int |f_{m'+1,I_{m'}}|^2(y)\, \phi^{\frac{1}{2}}_{P_{I_{m'}}(x)}(y)\, dy.
\end{aligned}$$

Noting that $|P_{I_{m'}}|^{-1} \phi^{1/2}_{P_{I_{m'}}(x)}(y) \lesssim \rho_{I_{m'}}(x-y)$, we get

$$|f_{m',I_{m'}}(x) - f_{m'+1,I_{m'}}(x)| \lesssim \lambda^{-1} |f_{m'+1,I_{m'}}|^2 * \rho_{I_{m'}}(x).$$

Summing the above over $I_{m'} \subset \bigcup_{I_m \in \mathcal{S}} I_m$ we conclude

$$\begin{aligned}
\Big| \sum_{I_{m'} \subset \bigcup_{I_m \in \mathcal{S}} I_m} f_{m',I_{m'}}(x) - \sum_{I_{m'} \subset \bigcup_{I_m \in \mathcal{S}} I_m} f_{m'+1,I_{m'}}(x) \Big| &\le \lambda^{-1} \sum_{I_{m'} \in \mathcal{I}_{m'}} |f_{m'+1,I_{m'}}|^2 * \rho_{I_{m'}}(x) \\
&= \lambda^{-1} g_{m'}(x) \lesssim_\varepsilon N^\varepsilon \frac{r}{\lambda}.
\end{aligned}$$

Therefore if we choose the constant $\widetilde{C}_\varepsilon$ in the definition of $\lambda = \widetilde{C}_\varepsilon N^{\varepsilon \frac{r}{\alpha}}$ to be large enough depending on ε, then we have, for $x \in U_{\alpha,r} \cap \Omega_{m,\alpha,r}$,

$$\sum_{m \le m' \le M} \left| \sum_{I_{m'} \subset \bigcup_{I_m \in \mathcal{S}} I_m} f_{m',I_{m'}}(x) - \sum_{I_{m'} \subset \bigcup_{I_m \in \mathcal{S}} I_m} f_{m'+1,I_{m'}}(x) \right| \le \frac{\alpha}{100}.$$

Since by definition $\sum_{I_{m'} \subset \bigcup_{I_m \in \mathcal{S}} I_m} f_{m',I_{m'}} = \sum_{I_{m'-1} \subset \bigcup_{I_m \in \mathcal{S}} I_m} f_{m',I_{m'-1}}$, we have by the triangle inequality that

$$\left| \sum_{I_m \in \mathcal{S}} f_{I_m} - \sum_{I_m \in \mathcal{S}} f_{m,I_m} \right| \le \frac{\alpha}{100}.$$

The case $m = 0$ follows from the above argument for $m = 1$ as by definition $f_0 = f_1$. □

From now on we will assume that $\widetilde{C}_\varepsilon$ is chosen large enough such that the conclusion of Lemma 5.4 holds.

6. Proof of Proposition 4.5

We prove Proposition 4.5 in this section, and consequently Theorem 4.4. We also give the proof of Theorem 4.3 assuming Theorem 4.4 in the last subsection. Still fix $2 \le p \le 6$, $\varepsilon > 0$, and $P(L_M) \subset \mathbb{R}$.

Suppose $1 \le K \le N^{\varepsilon/2}$ and $N^{1/2}/K \ge L$. Let $\mathcal{I}'$ be a partition of $\mathcal{N}_{N^{-1}K^{-1}}(\{a_n\}_{n=1}^{N^{1/2}})$ into K many I', which is a union of $N^{1/2}/K$ consecutive intervals in $\mathcal{N}_{N^{-1}K^{-1}}(\{a_n\}_{n=1}^{N^{1/2}})$. We call $I', I'' \in \mathcal{I}'$ nonadjacent if there exist at least two other $I''' \in \mathcal{I}'$ between I' and I'' on the real line. Alternatively, we can list $I' \in \mathcal{I}'$ as I'_j so that I'_{j+1} is on the right side of I'_j on the real line for every j. Then we define $I'_j, I'_{j'}$ to be nonadjacent if $|j - j'| \ge 3$. In displayed math we write "nonadj." as the shorthand for nonadjacent.

For f with $\operatorname{supp} \widehat{f} \subset \Omega$, we let $f_{I'}$ denote the projection of f to I' in the frequency space. So $f_{I'} = \sum_{I_M \subset I'} f_{I_M}$.

6A. ***Broad-narrow decomposition.*** The following lemma is a broad-narrow analysis on f with some complication. For parameters $\alpha, r > 0$ and m, $0 \le m \le M-1$, define

$$f_{m,\alpha,r,I'} := \sum_{I_m \subset I'} f_{m,\alpha,r,I_m},$$

where we recall that f_{m,α,r,I_m} is defined in (42).

Lemma 6.1. *For every $X \subset P(L_M)$, there exist some α, r with $\alpha \ge r^{1/2}$ and some m such that $0 \le m \le M-1$ and*

$$\int_X |f|^p \lesssim_\varepsilon \sum_{I' \in \mathcal{I}'} \int_X |f_{I'}|^p + (\log N \log(\theta^{-1}+1))^C \frac{K^C}{\alpha^{4-p}} \max_{\substack{I',I'' \\ \text{nonadj.}}} \int_{X \cap U_{\alpha,r} \cap \Omega_{m,\alpha,r}} |f_{m,\alpha,r,I'}|^2 |f_{m,\alpha,r,I''}|^2$$
$$+ \left(\sup_{x \in X} \sum_I \|f_I\|^2_{L^2(W_{P_I(x),100})} \right)^{\frac{p}{2}-1} \left(\sum_I \|f_I\|^2_{L^2(W_{P(L),100})} \right). \quad (46)$$

First we prove a technical lemma which is a pointwise broad-narrow analysis.

By taking all parameters to have dyadic values, we may assume that for each I_m, $0 \le m \le M$, and any I', either $I_m \subset I'$ or $I_m \cap I' = \varnothing$.

Lemma 6.2. *For every $\alpha, r > 0$ and $0 \le m \le M-1$,*

$$|f_{m,\alpha,r}(x)|^2 \lesssim \max_{I'} |f_{I'}(x)|^2 + K^C \max_{\substack{I',I'' \\ \text{nonadj.}}} |f_{m,\alpha,r,I'}(x)||f_{m,\alpha,r,I''}(x)|$$

for every $x \in X \cap U_{\alpha,r} \cap \Omega_{m,\alpha,r}$.

Proof. Let $x \in X \cap U_{\alpha,r} \cap \Omega_{m,\alpha,r}$. If there exist $I', I'' \in \mathcal{I}'$ nonadjacent such that $|f_{m,\alpha,r,I'}|, |f_{m,\alpha,r,I''}| \ge \frac{1}{100K}|f_{m,\alpha,r}(x)|$, then we have

$$|f_{m,\alpha,r}(x)|^2 \lesssim K^2 \max_{\substack{I',I'' \\ \text{nonadj.}}} |f_{m,\alpha,r,I'}(x)||f_{m,\alpha,r,I''}(x)|. \tag{47}$$

Now we assume there do not exist $I', I'' \in \mathcal{I}'$ nonadjacent with $|f_{m,\alpha,r,I'}|, |f_{m,\alpha,r,I''}| \ge \frac{1}{100K}|f_{m,\alpha,r}(x)|$. Note that $f_{m,\alpha,r}(x) = \sum_{I'} f_{m,\alpha,r,I'}(x)$ and the number of I' is bounded by K. So if we choose $I''' \in \mathcal{I}'$ with $|f_{m,\alpha,r,I'''}(x)| = \max_{I' \in \mathcal{I}'} |f_{m,\alpha,r,I'}(x)|$, then

$$|f_{m,\alpha,r,I'''}(x)| \ge \tfrac{1}{2}|f_{m,\alpha,r}(x)|. \tag{48}$$

By Lemma 5.4 we have $|f_{m,\alpha,r}(x)| \in [\alpha/4, 4\alpha]$, and $|f_{m,\alpha,r,I'''}(x) - f_{I'''}(x)| \le \frac{\alpha}{100}$. Therefore by the triangle inequality and (48) we obtain

$$|f_{I'''}(x)| \gtrsim \alpha \sim |f_{m,\alpha,r}(x)|.$$

This combined with (47) proves the lemma. □

Proof of Lemma 6.1. Since $P(L_M) = \bigsqcup_{\alpha,r\text{: dyadic}} U_{\alpha,r}$, we have

$$\int_X |f|^p \le \sum_{\alpha,r\text{: dyadic}} \int_{X \cap U_{\alpha,r}} |f|^p.$$

Without loss of generality we assume

$$\left(\sup_{x \in X} \sum_I \|f_I\|^2_{L^2(W_{P_I(x),100})}\right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_I \|f_I\|^2_{L^2(W_{P(L),100})}\right)^{\frac{1}{p}} = 1. \tag{49}$$

Then $X \cap U_{\alpha,r} = \varnothing$ if $\max\{\alpha, r\} \ge CN^C\theta^{-C}$ for some sufficiently large constant C. Also

$$\left(\int_{X \cap (\bigcup_{\min\{\alpha,r\} \le C^{-1}N^{-C}\theta^C} U_{\alpha,r})} |f|^p\right)^{\frac{1}{p}} \lesssim 1$$

if C is sufficiently large. So now we write

$$\int_X |f|^p \le \sum_{\alpha,r} \int_{X \cap U_{\alpha,r}} |f|^p + C, \tag{50}$$

where the number of pairs (α, r) in the summation is $\mathcal{O}(\log N \log(\theta^{-1}+1))^2$, since the number of dyadic numbers between $C^{-1}N^{-C}\theta^C$ and $CN^C\theta^{-C}$ is $\mathcal{O}(\log N + \log(\theta^{-1}+1)) = \mathcal{O}(\log N \log(\theta^{-1}+1))$.

We also observe that by Hölder's inequality and Fubini's theorem we have

$$\int_{X\cap\bigcup_{\alpha\le r^{1/2}}U_{\alpha,r}}|f|^p\lesssim\int_X\Big(\sum_I|f_I|^2*\rho_I\Big)^{\frac{p}{2}}\lesssim\Big\|\sum_I|f_I|^2*\rho_I\Big\|_{L^\infty(X)}^{\frac{p}{2}-1}\Big(\sum_I\|f_I\|^2_{L^2(W_{P(L),100})}\Big).$$

Since

$$\Big\|\sum_I|f_I|^2*\rho_I\Big\|_{L^\infty(X)}\le\sup_{x\in X}\sum_I|f_I|^2*\rho_I(x)\lesssim\sup_{x\in X}\sum_I\|f_I\|^2_{L^2(W_{P_I(x),100})},$$

we obtain

$$\int_{X\cap\bigcup_{\alpha\le r^{1/2}}U_{\alpha,r}}|f|^p\lesssim\Big(\sup_{x\in X}\sum_I\|f_I\|^2_{L^2(W_{P_I(x),100})}\Big)^{\frac{p}{2}-1}\Big(\sum_I\|f_I\|^2_{L^2(W_{P(L),100})}\Big)=1.$$

So in summary

$$\int_X|f|^p\lesssim\sum_{\alpha,r:\alpha\ge r^{1/2}}\int_{X\cap U_{\alpha,r}}|f|^p+1. \tag{51}$$

Next we further decompose $X\cap U_{\alpha,r}$ into $\bigcup_m(X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r})$:

$$\int_{X\cap U_{\alpha,r}}|f|^p\le\sum_{m=0}^{M-1}\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f|^p.$$

By Lemma 5.4 we have, for $0\le m\le M-1$,

$$\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f|^p\sim\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f_{m,\alpha,r}|^p.$$

It then follows from Lemmas 6.2 and 5.4 that

$$\begin{aligned}&\int_X|f|^p\\&\lesssim 1+\sum_{\alpha,r:\alpha\ge r^{1/2}}\sum_{m=0}^{M-1}\Big(\sum_{I'\in\mathcal{I}'}\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f_{I'}|^p+\frac{K^C}{\alpha^{4-p}}\max_{\substack{I',I''\\\text{nonadj.}}}\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f_{m,\alpha,r,I'}|^2\,|f_{m,\alpha,r,I''}|^2\Big)\\&\lesssim 1+C_\varepsilon\sum_{I'\in\mathcal{I}'}\int_X|f_{I'}|^p+\sum_{\alpha,r:\alpha\ge r^{1/2}}\sum_m\frac{K^C}{\alpha^{4-p}}\max_{\substack{I',I''\\\text{nonadj.}}}\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f_{m,\alpha,r,I'}|^2\,|f_{m,\alpha,r,I''}|^2,\end{aligned}$$

where we used $M\lesssim_\varepsilon 1$ in the last inequality. Recall that the number of pairs (α,r) in the summation is $\mathcal{O}(\log N\log(\theta^{-1}+1))^2$ (see (50)); by the pigeonhole principle we have

$$\begin{aligned}&\sum_{\alpha,r:\alpha\ge r^{1/2}}\sum_m\frac{K^C}{\alpha^{4-p}}\max_{\substack{I',I''\\\text{nonadj.}}}\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f_{m,\alpha,r,I'}|^2|f_{m,\alpha,r,I''}|^2\\&\qquad\lesssim_\varepsilon(\log N\log(\theta^{-1}+1))^2\frac{K^C}{\alpha^{4-p}}\max_{\substack{I',I''\\\text{nonadj.}}}\int_{X\cap U_{\alpha,r}\cap\Omega_{m,\alpha,r}}|f_{m,\alpha,r,I'}|^2|f_{m,\alpha,r,I''}|^2\end{aligned}$$

for some α,r with $\alpha\ge r^{1/2}$, $0\le m\le M-1$, which completes the proof. □

Now fix $X \subset P(L_M)$. We have identified a pair (α, r) from Lemma 6.1, and we fix that pair of α, r and suppress the dependence on α, r from now on in the notation. In particular write $g_m = g_{m,\alpha,r}$, $\Omega_m = \Omega_{m,\alpha,r}$, $f_{m,I'} = f_{m,\alpha,r,I'}$ and $f_{m,I_m} = f_{m,\alpha,r,I_m}$ where α, r are those chosen in Lemma 6.1.

We estimate the broad and narrow parts separately, which together with Lemma 6.1 will imply Proposition 4.5.

6B. *Narrow part.*

Proposition 6.3. *For every $I' \in \mathcal{I}'$ we have*

$$\int_X |f_{I'}|^p \lesssim \left(\sup_{\theta' \in [\theta/4,\theta]} \mathrm{Dec}\left(\frac{N}{K^2}, \frac{\theta'}{K^2}\right)^p\right)\left(\sup_{x\in X} \sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P_I(x),100})}\right)^{\frac{p}{2}-1}\left(\sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P(L),100})}\right). \quad (52)$$

Proof. In this proof, the notation cA with $c \in \mathbb{R}$, $A \subset \mathbb{R}$ will denote the set $\{ca : a \in A\}$.

We first prove (52) for $I' = \mathcal{N}_{L^2\theta/N^2}(\{a_n\}_{n=1}^{N^{1/2}/K})$. Note that $K^2 I' = \mathcal{N}_{K^2L^2\theta/N^2}(\{K^2 a_n\})_{n=1}^{N^{1/2}/K}$, and if we let $\tilde{a} = K^2 a_n$, $\tilde{N} = N/K^2$ and $\tilde{\theta} = \theta/K^2$, then

$$\tilde{a} - \tilde{a} \in \left[\frac{K^2}{4N}, \frac{4K^2}{N}\right] = \left[\frac{1}{4\tilde{N}}, \frac{4}{\tilde{N}}\right], \quad (\tilde{a}-\tilde{a}) - (\tilde{a}-\tilde{a}) \in \left[\frac{K^2\theta}{4N^2}, \frac{4K^2\theta}{N^2}\right] = \left[\frac{\tilde{\theta}}{4\tilde{N}^2}, \frac{4\tilde{\theta}}{\tilde{N}^2}\right],$$

and $K^2 I' = \mathcal{N}_{L^2\tilde{\theta}/\tilde{N}^2}(\{\tilde{a}\}_{n=1}^{\tilde{N}^{1/2}})$.

We define $\widetilde{P}(L)$, $\widetilde{P}_{K^2 I}$ by (16), (15) respectively with N, L, θ, v_j replaced by $\tilde{N}, L, \tilde{\theta}, K^2 v_j$. Then for any x_0 we have $\widetilde{P}_{K^2 I}(K^{-2}x_0) = K^{-2} P_I(x_0)$, and $\widetilde{P}(L, K^{-2}x_0) \subset K^{-2} P(L, x_0)$. Now by the change of variable formula,

$$\int_X |f_{I'}(x)|^p \, dx = K^2 \int_{K^{-2}X} |f_{I'}(K^2 x)|^p \, dx.$$

We have $\operatorname{supp} \widehat{f_{I'}(K^2\cdot)} \subset K^2 I' = \mathcal{N}_{L^2\tilde{\theta}/\tilde{N}^2}(\{\tilde{a}\}_{n=1}^{\tilde{N}^{1/2}})$. Let $\tilde{f}(x)$ denote the function $f_{I'}(K^2 x)$. Therefore by the definition of the refined decoupling constant for $\mathcal{N}_{L^2\tilde{\theta}/\tilde{N}^2}(\{\tilde{a}\}_{n=1}^{\tilde{N}^{1/2}})$, and (33) (as $\widetilde{P}(L, K^{-2}x_0) \subset K^{-2}P(L, x_0)$), we have

$$\int_{K^{-2}X} |\tilde{f}(x)|^p \, dx \le \mathrm{Dec}(\tilde{N}, \tilde{\theta})^p \left(\sup_{x\in X} \sum_{I\subset I'} \|\tilde{f}\|^2_{L^2(W_{\widetilde{P}_{K^2 I}(K^{-2}x),100})}\right)^{\frac{p}{2}-1}\left(\sum_{I\subset I'} \|\tilde{f}\|^2_{L^2(W_{K^{-2}P(L),100})}\right).$$

By the change of variable formula,

$$\|\tilde{f}\|_{L^2(W_{\widetilde{P}_{K^2 I}(K^{-2}x),100})} \lesssim \|f_I\|_{L^2(W_{P_I(x),100})},$$
$$\|\tilde{f}\|^2_{L^2(W_{K^{-2}P(L),100})} \lesssim K^{-2}\|f_I\|^2_{L^2(W_{P(L),100})}.$$

So we conclude

$$\int_X |f_{I'}|^p \lesssim \mathrm{Dec}\left(\frac{N}{K^2}, \frac{\theta}{K^2}\right)^p \left(\sup_{x\in X} \sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P_I(x),100})}\right)^{\frac{p}{2}-1}\left(\sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P(L),100})}\right).$$

Now we consider a general $I' \in \mathcal{I}'$. Suppose a_l is the first term in $I' \cap \{a_n\}_{n=1}^{N^{1/2}}$, and let $v_l = a_{l+1} - a_l$. Because of (14) we have $v_l \in [v_1, 2v_1]$. So we may choose $K_l \in [K/\sqrt{2}, K]$ such that

$$K_l^2 v_l \in \left[\frac{1}{4\widetilde{N}}, \frac{4}{\widetilde{N}}\right].$$

Then

$$K_l^2((a_{n+1}-a_n)-(a_n-a_{n-1})) \in \left[\frac{\theta K_l^2}{4N^2}, \frac{4\theta K_l^2}{N^2}\right] = \left[\frac{\tilde{\theta}'}{4\widetilde{N}^2}, \frac{4\tilde{\theta}_l}{\widetilde{N}^2}\right]$$

for some $\tilde{\theta}_l \in [\tilde{\theta}/4, \tilde{\theta}]$. Let $\theta_l = K^2\tilde{\theta}_l$, which lies in $[\theta/4, 4\theta]$. So by a change of variable argument again we have

$$\int_X |f_{I'}|^p \lesssim \mathrm{Dec}\left(\frac{N}{K^2}, \frac{\theta_l}{K^2}\right)^p \left(\sup_{x\in X} \sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P_I(x),100})}\right)^{\frac{p}{2}-1} \left(\sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P(L),100})}\right).$$

Therefore we have shown (52) for every $I' \in \mathcal{I}'$. □

The proof of Proposition 6.3 actually shows that (52) holds for every f with frequency support in Ω (not only alternately spaced f) and every $X \subset P(L)$.

6C. *Broad part.*

Proposition 6.4. *For $1 \le m \le M-1$ and $I', I'' \in \mathcal{I}'$ nonadjacent we have*

$$\int_{X\cap U_{\alpha,r}\cap\Omega_m} |f_{m,I'}|^2 |f_{m,I''}|^2 \lesssim_\varepsilon N^{C\varepsilon} K^C \left(\frac{r}{\alpha}\right)^2 \left(\sum_{I\in\mathcal{I}} \|f_I\|^2_{L^2(W_{P(L),100})}\right). \tag{53}$$

Proof. Fix a $P(L'_m)$ such that $P(L'_m) \cap X \cap U_{\alpha,r} \cap \Omega_m \neq \varnothing$. Recall that $L'_m = (L_m N^{1/2})^{1/2}$ as defined in Section 3. Suppose the distance between I' and I'' is $1/K'$. Since I', I'' are nonadjacent, we have $1/K \le 1/K' \lesssim 1$. Let $\tilde{f}(x)$ denote the function $f_{m,I'}((K')^2 x)$, and $\tilde{f}(x)$ denote the function $f_{m,I''}((K')^2 x)$. Then supp $\tilde{\tilde{f}}_1 \subset (K')^2 I'$, supp $\tilde{\tilde{f}}_2 \subset (K')^2 I''$, and $d((K')^2 I', (K')^2 I'') \gtrsim 1$. By (22) and a change of variable argument similar to that in the proof of Proposition 6.3, we have

$$\int_{(K')^{-2}P(L'_m)} |\tilde{f}|^2 |\tilde{f}|^2 \lesssim_\varepsilon N^\varepsilon (K')^C |P(L'_m)|^{-1} \int |\tilde{f}|^2 W_{(K')^{-2}P(L'_m),200} \int |\tilde{f}|^2 W_{(K')^{-2}P(L'_m),200}.$$

By the local L^2 orthogonality Lemma 3.6, we further obtain

$$\begin{aligned}&\int_{(K')^{-2}P(L'_m)} |\tilde{f}|^2 |\tilde{f}|^2 \\ &\lesssim_\varepsilon N^\varepsilon (K')^C |P(L'_m)|^{-1} \int \sum_{I_m\subset I'} |(\tilde{f})_{(K')^2 I_m}|^2 W_{(K')^{-2}P(L'_m),200} \int \sum_{I_m\subset I''} |(\tilde{f})_{(K')^2 I_m}|^2 W_{(K')^{-2}P(L'_m),200}.\end{aligned}$$

Here the notation cA with $c \in \mathbb{R}$, $A \subset \mathbb{R}$ denotes the set $\{ca : a \in A\}$. Applying the change of variable $x \mapsto (K')^{-2}x$ to both sides of the above inequality, and using $K' \le K$, we get

$$\int_{P(L'_m)} |f_{m,I'}|^2 |f_{m,I''}|^2 \lesssim_\varepsilon N^\varepsilon K^C |P(L'_m)|^{-1} \int \sum_{I_m\subset I'} |f_{m,I_m}|^2 W_{P(L'_m),200} \int \sum_{I_m\subset I''} |f_{m,I_m}|^2 W_{P(L'_m),200}.$$

By Hölder's inequality,

$$\int_{P(L'_m)} |f_{m,I'}|^2 |f_{m,I''}|^2 \lesssim_\varepsilon N^\varepsilon K^C \int \left(\sum_{I_m} |f_{m,I_m}|^2 \right)^2 W_{P(L'_m),200},$$

and due to $|f_{m,I_m}| \le |f_{m+1,I_m}|$ we further have

$$\int_{P(L'_m)} |f_{m,I'}|^2 |f_{m,I''}|^2 \lesssim_\varepsilon N^\varepsilon K^C \int \left(\sum_{I_m} |f_{m+1,I_m}|^2 \right)^2 W_{P(L'_m),200}.$$

Now applying Proposition 2.3 we obtain

$$\begin{aligned}\int_{P(L'_m)} |f_{m,I'}|^2 |f_{m,I''}|^2 &\lesssim_\varepsilon N^\varepsilon K^C |P(L'_m)|^{-1} \left(\int \left(\sum_{I_m} |f_{m+1,I_m}|^2 \right) W_{P(L'_m),100} \right)^2 \\ &\lesssim N^\varepsilon K^C \int_{P(L'_m)} g_m^2.\end{aligned}$$

Note that from the definition of Ω_m and the definition of $g_m := \sum_{I_m} |f_{m+1,I_m}|^2 * \rho_{I_m}$ we have $x \in \Omega_m$ implies $|g_m(x)| \sim \sup_{y \in P(L'_m(x))} |g_m(y)| \lesssim |g_m^h(x)|$. Therefore we have (by Proposition 2.3)

$$\int_{P(L'_m)} g_m^2 \lesssim |P(L'_m)| |g_m^h(x)|^2 \lesssim \int |g_m^h|^2 W_{P(L'_m),100},$$

where $x \in P(L'_m) \cap \Omega_m$. Summing over disjoint $P(L'_m)$ that intersect $X \cap U_{\alpha,r} \cap \Omega_m$ we obtain

$$\int_{X \cap U_{\alpha,r} \cap \Omega_m} |f_{m,I'}|^2 |f_{m,I''}|^2 \lesssim_\varepsilon N^\varepsilon K^C \int |g_m^h|^2 W_{P(L_M),100} \lesssim N^{2\varepsilon} K^C \int \sum_{I_m} |f_{m+1,I_m}|^4 W_{P(L_M),100},$$

where the last inequality is due to Lemma 5.3. By Hölder's inequality and the definition of $f_{m+1,I_{m+1}}$ we have

$$\begin{aligned}\int \sum_{I_m} |f_{m+1,I_m}|^4 W_{P(L_M),100} &\lesssim N^{C\varepsilon} \int \sum_{I_{m+1}} |f_{m+1,I_{m+1}}|^4 W_{P(L_M),100} \\ &\lesssim N^{C\varepsilon} \left(\frac{r}{\alpha} \right)^2 \int \sum_{I_{m+1}} |f_{m+1,I_{m+1}}|^2 W_{P(L_M),100}.\end{aligned}$$

By the pointwise inequality $|f_{m+1,I_{m+1}}| \le |f_{m+2,I_{m+1}}|$ and local L^2 orthogonality (Lemma 3.6),

$$\begin{aligned}\int \sum_{I_{m+1}} |f_{m+1,I_{m+1}}|^2 W_{P(L_M),100} &\lesssim \int \sum_{I_{m+2}} |f_{m+2,I_{m+1}}|^2 W_{P(L_M),100} \\ &\lesssim \int \sum_{I_{m+2}} |f_{m+2,I_{m+2}}|^2 W_{P(L_M),100}.\end{aligned}$$

Continuing this process we obtain

$$\int \sum_{I_{m+1}} |f_{m+1,I_{m+1}}|^2 W_{P(L_M),100} \lesssim_\varepsilon \int \sum_{I_M} |f_{M,I_M}|^2 W_{P(L_M),100}. \tag{54}$$

Recalling that $|f_{M,I_M}| \le |f_{I_M}| = |f_I|$ we conclude

$$\int_{X\cap U_{\alpha,r}\cap\Omega_m} |f_{m,I'}|^2 |f_{m,I''}|^2 \lesssim_\varepsilon N^{C\varepsilon} K^C \left(\frac{r}{\alpha}\right)^2 \int \sum_I |f_I|^2 W_{P(L_M),100}. \qquad \square$$

Proposition 6.5. *For $I', I'' \in \mathcal{I}'$ nonadjacent we have*

$$\int_{X\cap U_{\alpha,r}\cap\Omega_0} |f_{0,I'}|^{\frac{p}{2}} |f_{0,I''}|^{\frac{p}{2}} \lesssim_\varepsilon N^\varepsilon \Bigg(\sup_{x\in X} \sum_I \|f_I\|^2_{L^2(W_{P_I(x),100})}\Bigg)^{\frac{p}{2}-1} \Bigg(\sum_I \|f_I\|^2_{L^2(W_{P(L),100})}\Bigg).$$

Proof. By the Cauchy–Schwarz inequality we have

$$\begin{aligned}\int_{X\cap U_{\alpha,r}\cap\Omega_0} |f_{0,I'}|^{\frac{p}{2}} |f_{0,I''}|^{\frac{p}{2}} &\lesssim N^\varepsilon \int_{X\cap U_{\alpha,r}\cap\Omega_0} \Bigg(\sum_{I_1} |f_{1,I_1}|^2\Bigg)^{\frac{p}{2}} \\ &\lesssim N^\varepsilon \sup_{x\in X\cap\Omega_0} \Bigg(\sum_{I_1} |f_{1,I_1}|^2\Bigg)^{\frac{p}{2}-1} \int \sum_{I_1} |f_{1,I_1}|^2 W_{P(L_M),100}.\end{aligned}$$

We have shown in the proof of Proposition 6.4 (inequality (54)) that

$$\int \sum_{I_1} |f_{1,I_1}|^2 W_{P(L_M),100} \lesssim_\varepsilon \int \sum_I |f_I|^2 W_{P(L),100}.$$

So it suffices to show

$$\sup_{x\in X\cap\Omega_0} \Bigg(\sum_{I_1} |f_{1,I_1}|^2\Bigg) \lesssim_\varepsilon \sup_{x\in X} \sum_I \|f_I\|^2_{L^2(W_{P_I(x),100})}. \tag{55}$$

From the locally constant property (Proposition 2.3) we have

$$\sum_{I_1} |f_{1,I_1}|^2(x) \lesssim \sum_{I_1} |f_{1,I_1}|^2 * \rho_{I_1}(x) \lesssim \sum_{I_1} |f_{2,I_1}|^2 * \rho_{I_1}(x) = g_1(x)$$

(recall that ρ_{I_1} is an L^1 normalized nonnegative function adapted to $P_{I_1}(0)$ satisfying (43)), and by Lemma 5.2 we have, for $x \in X \cap \Omega_0$, $g_1(x) \lesssim_\varepsilon g_M(x)$. So we conclude

$$\sup_{x\in X\cap\Omega_0} \sum_{I_1} |f_{1,I_1}|^2(x) \lesssim_\varepsilon \sup_{x\in X\cap\Omega_0} g_M(x) \lesssim \sup_{x\in X} \sum_I \|f_I\|^2_{L^2(W_{P_I(x),100})}. \qquad \square$$

6D. ***Proof of Proposition 4.5.*** Let $X \subset P(L)$. We choose α, r as in Lemma 6.1. Note that

$$r \le 2 \Bigg\| \sum_I |f_I|^2 \Bigg\|_{L^\infty(X)}$$

since otherwise $X \cap U_{\alpha,r} = \varnothing$. So

$$r \le 2 \Bigg\| \sum_I |f_I|^2 \Bigg\|_{L^\infty(X)} \lesssim \sup_{x\in X} \sum_{I\subset I'} \|f_I\|^2_{L^2(W_{P_I(x),100})}.$$

Also $\alpha \geq r^{1/2}$ implies that $r^{3-p/2}/\alpha^{6-p} \leq 1$ as $p \leq 6$. Therefore combining Propositions 6.3, 6.4, and 6.5 and Lemma 6.1 we obtain

$$\int_X |f|^p \lesssim_\varepsilon \left(\sup_{\theta' \in [\theta/4, 4\theta]} \mathrm{Dec}\left(\frac{N}{K^2}, \frac{\theta'}{K^2}\right)^p + \log^C(\theta^{-1}+1) N^{C\varepsilon} K^C \right) \times \left(\sup_{x \in X} \sum_I \|f_I\|^2_{\mathbb{L}^2(W_{P_I(x),100})} \right)^{\frac{p}{2}-1} \left(\sum_I \|f_I\|^2_{L^2(W_{P(L),100})} \right). \quad (56)$$

6E. ***Proof of Theorem 4.3.*** Finally, in this section we show how Theorem 4.4 implies Theorem 4.3. Let $f = \sum_I f_I$. Taking $X = P(L)$ in (32) we see that

$$\|f\|_{L^p(P(L))} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \left(\sup_{x \in P(L)} \sum_I \|f_I\|^2_{\mathbb{L}^2(W_{P_I(x),100})} \right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_I \|f_I\|^2_{L^2(W_{P(L),100})} \right)^{\frac{1}{p}}.$$

To prove Theorem 4.1 we will do dyadic pigeonholing on the L^2-norm of wave packets of f, using Proposition 5.1. More precisely we write

$$f = \sum_I f_I = \sum_I \sum_{P_I} \phi_{P_I} f_I = \sum_{\lambda \text{: dyadic}} \sum_{I, P_I : \|\phi_{P_I} f_I\|_{L^2(W_{P_I,100})} \in [\lambda/2, \lambda)} \phi_{P_I} f_I.$$

Without loss of generality we assume $\left(\sum_I \|f_I\|^2_{L^p(W_{P(L),100})}\right)^{1/2} = 1$. Then

$$\left\| \sum_{I, P_I : \|\phi_{P_I} f_I\|_{L^2(W_{P_I,100})} \notin [N^{-C}\theta^C, N^C \theta^{-C}]} \phi_{P_I} f_I \right\|_{L^p(P(L))} \lesssim 1$$

for sufficiently large C. Therefore there exists a λ such that

$$\|f\|_{L^p(P(L))} \lesssim C_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \left\| \sum_{I, P_I : \|\phi_{P_I} f_I\|_{L^2(W_{P_I,100})} \in [\lambda/2, \lambda)} \phi_{P_I} f_I \right\|_{L^p(P(L))} + 1.$$

By a further dyadic pigeonholing argument on I, we may assume, for every I, either

$$\#\{P_I : \|\phi_{P_I} f_I\|_{L^2(W_{P(L),100})} \in [\lambda/2, \lambda)\} = 0$$

or

$$\#\{P_I : \|\phi_{P_I} f_I\|_{L^2(W_{P(L),100})} \in [\lambda/2, \lambda)\} \in [A/2, A) \quad \text{for some constant } A.$$

We denote by $\#I$ the number of I such that $\#\{P_I : \|\phi_{P_I} f_I\|_{L^2(W_{P(L),100})} \in [\lambda/2, \lambda)\} \in [A/2, A)$. For simplicity of notation we will also drop writing the condition $\|\phi_{P_I} f_I\|_{L^2(W_{P(L),100})} \in [\lambda/2, \lambda)$ in the summation. Now apply Theorem 4.4 to get

$$\left\| \sum_{I, P_I} \phi_{P_I} f_I \right\|_{L^p(P(L))} \lesssim_\varepsilon \log^C(\theta^{-1}+1) N^\varepsilon \left(\sup_{x \in P(L)} \sum_I \left\| \sum_{P_I} \phi_{P_I} f_I \right\|^2_{\mathbb{L}^2(W_{P_I(x),100})} \right)^{\frac{1}{2}-\frac{1}{p}} \times \left(\sum_I \left\| \sum_{P_I} \phi_{P_I} f_I \right\|^2_{L^2(W_{P(L),100})} \right)^{\frac{1}{p}}. \quad (57)$$

To estimate the first factor on the right-hand side of (57) we note that, for every $x \in P(L)$,

$$\sum_I \Big\| \sum_{P_I} \phi_{P_I} f_I \Big\|^2_{L^2(W_{P_I(x),100})} \lesssim \sum_I \sum_{P_I} \|\phi_{P_I} f_I\|^2_{L^2(W_{P_I(x),100})} \lesssim (\#I)\lambda^2 |P_I|^{-1}$$

because of $\big(\sum_{P_I} \phi_{P_I}(y)\big)^2 \lesssim \sup_{P_I} \phi^2_{P_I}(y) \le \sum_{P_I} \phi^2_{P_I}(y)$ and (12). Therefore

$$\sup_{x\in P(L)} \sum_I \Big\| \sum_{P_I} \phi_{P_I} f_I \Big\|^2_{L^2(W_{P_I(x),100})} \lesssim (\#I)\lambda^2 |P_I|^{-1}.$$

To estimate the second factor on the right-hand side of (57) we calculate

$$\sum_I \Big\| \sum_{P_I} \phi_{P_I} f_I \Big\|^2_{L^2(W_{P(L),100})} \lesssim \sum_I \sum_{P_I} \|\phi_{P_I} f_I\|^2_{L^2(W_{P(L),100})} \lesssim (\#I)\lambda^2 A.$$

To summarize, (57) implies that

$$\Big\| \sum_{I,P_I} \phi_{P_I} f_I \Big\|_{L^p(P(L))} \lesssim_\varepsilon \log^C(\theta^{-1}+1) N^\varepsilon |P_I|^{\frac{1}{p}-\frac{1}{2}} (\#I)^{\frac{1}{2}} A^{\frac{1}{p}} \lambda.$$

Now by Hölder's inequality we have

$$\begin{aligned}
\Big(\sum_I \|f_I\|^2_{L^p(W_{P(L),100})} \Big)^{\frac{1}{2}} &\ge \Big(\sum_I \Big(\sum_{P_I} \|\phi_{P_I}^{\frac{1}{2}} f_I\|^p_{L^p(W_{P(L),100})} \Big)^{\frac{2}{p}} \Big)^{\frac{1}{2}} \\
&\gtrsim \Big(\sum_I \Big(\sum_{P_I} \|\phi_{P_I} f_I\|^p_{L^2(W_{P(L),100})} |P_I|^{1-\frac{p}{2}} \Big)^{\frac{2}{p}} \Big)^{\frac{1}{2}} \\
&\gtrsim |P_I|^{\frac{1}{p}-\frac{1}{2}} (\#I)^{\frac{1}{2}} A^{\frac{1}{p}} \lambda.
\end{aligned}$$

Hence we have (31).

7. A decoupling inequality for generalized Dirichlet sequences

In this section we focus only on generalized Dirichlet sequences with parameter $\theta = 1$. That is, we say $\{a_n\}_{n=1}^N$ is a generalized Dirichlet sequence if it satisfies (14) with $\theta = 1$. We will present a decoupling inequality for generalized Dirichlet sequences, by combining Theorem 4.1 and the flat decoupling (Proposition 7.2 below). Then we show that for certain choices of the generalized Dirichlet sequences $\{a_n\}_{n=1}^N$ the decoupling inequality that we obtain in this way is sharp (up to $C_\varepsilon N^\varepsilon$).

More precisely, for $1 \le L \le N^{1/2}$, we let Ω' denote the L^2/N^2-neighborhood of $\{a_n\}_{n=1}^N$, and let $\{J\}_{J\in\mathcal{J}}$ be a partition of Ω' into $\Omega' \cap B_{N^{-1/2}}$, where $B_{N^{-1/2}}$ runs over a tiling of $\mathbb{R}$ by balls of radius $N^{-1/2}$. So there are about $N^{1/2}$ many J and each J contains $\mathcal{O}(N^{1/2})$ many consecutive intervals in Ω'. For each J we let $\mathcal{I}_J$ be the partition of J into I, which is a union of L many consecutive intervals in Ω'.

We have the following decoupling inequality for the partition $\Omega' = \bigsqcup_{J\in\mathcal{J}} \bigsqcup_{I\in\mathcal{I}_J} I$.

Theorem 7.1. *For $2 \le p \le 6$, we have*

$$\|f\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^{\frac{1}{4}-\frac{1}{2p}+\varepsilon}\Big(\sum_{J\in\mathcal{J}}\sum_{I\in\mathcal{I}_J}\|f_I\|^2_{L^p(\mathbb{R})}\Big)^{\frac{1}{2}} \tag{58}$$

for every $f:\mathbb{R}\to\mathbb{C}$ with $\operatorname{supp}\widehat{f}\subset\Omega'$. There exists a choice of $\{a_n\}_{n=1}^N$ (satisfying (14) with $\theta=1$) such that the above estimate is sharp up to an N^ε factor.

7A. ***Proof of (58).*** From Theorem 4.1 we have, for every $J\in\mathcal{J}$ and $2\le p\le 6$,

$$\|f_J\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon\Big(\sum_{I\in\mathcal{I}_J}\|f_I\|^2_{L^p(\mathbb{R})}\Big)^{\frac{1}{2}}. \tag{59}$$

Next we decouple f_J into f_I using the flat decoupling:

Proposition 7.2. *Let $\mathcal{U}$ denote the partition*

$$[0,M) = \bigsqcup_{m=0}^{M-1}[m,m+1).$$

Then for $p\ge 2$ we have

$$\|f\|_{L^p(\mathbb{R})} \lesssim_p M^{\frac{1}{2}-\frac{1}{p}}\Big(\sum_{U\in\mathcal{U}}\|f_U\|^2_{L^p(\mathbb{R})}\Big)^{\frac{1}{2}}$$

for every $f:\mathbb{R}\to\mathbb{C}$ with $\operatorname{supp}\widehat{f}\subset[0,M)$.

Flat decoupling inequality is well known (see for example Proposition 2.4 in [Demeter et al. 2020]) but we include a proof here for the sake of completeness.

Proof. Fix $p\ge 2$. It suffices to prove that

$$\|f\|_{L^p(B_1)} \lesssim M^{\frac{1}{2}-\frac{1}{p}}\Big(\sum_{U\in\mathcal{U}}\|f_U\|^2_{L^p(W_{B_1,100})}\Big)^{\frac{1}{2}}$$

for f with $\operatorname{supp}\widehat{f}\subset[0,M)$. We calculate

$$\begin{aligned}
\|f\|^p_{L^p(B_1)} &\le \|f\|^{p-2}_{L^\infty(B_1)}\|f\|^2_{L^2(B_1)}\\
&\lesssim \Big(\sum_U\|f_U\|_{L^\infty(B_1)}\Big)^{p-2}\Big(\sum_U\|f_U\|^2_{L^2(W_{B_1,100})}\Big)\\
&\lesssim \Big(\sum_U\|f_U\|_{L^p(W_{B_1,100})}\Big)^{p-2}\Big(\sum_U\|f_U\|^2_{L^p(W_{B_1,100})}\Big)\\
&\lesssim M^{\frac{p-2}{2}}\Big(\sum_U\|f_U\|^2_{L^p(W_{B_1,100})}\Big)^{\frac{p-2}{2}}\Big(\sum_U\|f_U\|^2_{L^p(W_{B_1,100})}\Big)\\
&\lesssim M^{\frac{p-2}{2}}\Big(\sum_U\|f_U\|^2_{L^p(W_{B_1,100})}\Big)^{\frac{p}{2}}.
\end{aligned}$$

Here we used the locally constant property similar to Proposition 2.3 and local L^2 orthogonality similar to Lemma 3.6. □

Now we prove the decoupling inequality in Theorem 7.1.

Proof of (58) in Theorem 7.1. Combining (59) with Proposition 7.2 we obtain

$$\|f\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \Big(\sum_{J\in\mathcal{J}} \|f_J\|^2_{L^p(\mathbb{R})}\Big)^{\frac12} \lesssim N^{\frac14-\frac{1}{2p}+\varepsilon}\Big(\sum_{J\in\mathcal{J}}\sum_{I\in\mathcal{I}_J} \|f_I\|^2_{L^p(\mathbb{R})}\Big)^{\frac12}$$

for f with supp $\widehat{f}\subset\Omega'$. □

7B. ***An example and sharpness of (58).*** To prove the sharpness part, we construct a sequence $\{a_n\}_{n=1}^N$ satisfying (14) (with $\theta=1$) and for which (58) is sharp. We will use the function

$$g(x) = \frac{4x + (N^{\frac12} - \sqrt{N-4x})^2}{4N}$$

to define the sequence. For $n = 0,\dots,\frac{N}{8}$, let

$$a_n = g(n).$$

Distinguish the subsequence a_{n_k} where $n_k = kN^{1/2} - k^2$.

Lemma 7.3. *There is an absolute constant $N_0 > 0$ such that for every $N \ge N_0$, the sequence $\{a_n\}_{n=1}^{N/8}$ constructed above satisfies property (14) (with $\theta = 1$). Furthermore, there is an absolute constant $c > 0$ so that*

$$\Big\{\frac{j}{N^{\frac12}} : j = 1,\dots,\lfloor cN^{\frac12}\rfloor\Big\}$$

is a subsequence of $\{a_n\}_{n=1}^{N/8}$.

Proof. First we verify the presence of the subsequence: Let n_k and a_{n_k} be as above. Calculate directly that

$$\begin{aligned} a_{n_k} = g(n_k) &= \frac{4n_k + (N^{\frac12} - \sqrt{N-4n_k})^2}{4N} \\ &= \frac{4(kN^{\frac12}-k^2) + (N^{\frac12} - \sqrt{N-4(kN^{\frac12}-k^2)})^2}{4N} \\ &= \frac{4(kN^{\frac12}-k^2) + (N^{\frac12} - (N^{\frac12}-2k))^2}{4N} \\ &= \frac{4kN^{\frac12} - 4k^2 + 4k^2}{4N} = \frac{k}{N^{\frac12}}. \end{aligned}$$

This calculation holds as long as $k \le N^{1/2}/2$. Also note that $n_k = kN^{1/2} - k^2$ is increasing as a function of k as long as $k \le N^{1/2}/2$, so the n_k define a subsequence $a_{n_0},\dots,a_{n_K}$ where $K = \lfloor N^{1/2}/2\rfloor$.

To verify property (14), it suffices to check that for N large enough

$$a_1 - a_0 \in \Big[\frac{1}{2N}, \frac{2}{N}\Big] \tag{60}$$

and that

$$(a_{n+1} - a_n) - (a_n - a_{n-1}) \in \Big[\frac{1}{4N^2}, \frac{4}{N^2}\Big] \tag{61}$$

whenever $1 \le n \le \frac{N}{8} - 1$, since (60) together with (61) will imply $a_2 - a_1 \in \big[\frac{1}{4N}, \frac{4}{N}\big]$ for N large enough.

First we check (60). Note that $a_0 = 0$ and

$$a_1 = g(1) = \frac{4 + (N^{\frac{1}{2}} - \sqrt{N-4})^2}{4N}.$$

Then

$$a_1 - a_0 = \frac{1}{4N}\left(4 + \frac{16}{(N^{\frac{1}{2}} + \sqrt{N-4})^2}\right) \in \left[\frac{1}{2N}, \frac{2}{N}\right]$$

if N is large enough.

Next we check (61). First calculate

$$\begin{aligned} g(x+1) - g(x) &= \frac{4 + (N^{\frac{1}{2}} - \sqrt{N-4x-4})^2 - (N^{\frac{1}{2}} - \sqrt{N-4x})^2}{4N} \\ &= \frac{4 + 2N^{\frac{1}{2}}(\sqrt{N-4x} - \sqrt{N-4x-4}) - 4}{4N} \\ &= \frac{\sqrt{N-4x} - \sqrt{N-4x-4}}{2N^{\frac{1}{2}}} = \frac{2}{N^{\frac{1}{2}}(\sqrt{N-4x} + \sqrt{N-4x-4})}. \end{aligned}$$

Use this formula to calculate the difference

$$\begin{aligned} &(a_{n+1} - a_n) - (a_n - a_{n-1}) \\ &\quad = \frac{2}{N^{\frac{1}{2}}}\left(\frac{1}{\sqrt{N-4n} + \sqrt{N-4n-4}} - \frac{1}{\sqrt{N-4n+4} + \sqrt{N-4n}}\right) \\ &\quad = \frac{2}{N^{\frac{1}{2}}} \frac{\sqrt{N-4n+4} - \sqrt{N-4n-4}}{(\sqrt{N-4n} + \sqrt{N-4n-4})(\sqrt{N-4n+4} + \sqrt{N-4n})} \\ &\quad = \frac{16}{N^{\frac{1}{2}}(\sqrt{N-4n} + \sqrt{N-4n-4})(\sqrt{N-4n+4} + \sqrt{N-4n})(\sqrt{N-4n+4} + \sqrt{N-4n-4})}. \end{aligned}$$

As long as $n \le \frac{N}{8}$, and N is sufficiently large, this lies in $\left[\frac{1}{4N}, \frac{4}{N}\right]$ and we are done. □

Now we can finish the sharpness part of Theorem 7.1.

Proof of the sharpness part of Theorem 7.1. For $N \ge N_0$, we take $\{a_n\}_{n=1}^{N/8}$ to be the sequence constructed in Lemma 7.3, extended arbitrarily to $\{a_n\}_{n=1}^{N}$ so that (14) is satisfied with $\theta = 1$. We take $f = \sum_I f_I$ to be the function

$$\phi_{N^2/L^2}(x) \sum_{n=1}^{\lfloor cN^{1/2} \rfloor} e^{ixa_n},$$

where c is the constant in Lemma 7.3, and $\phi_{N^2/L^2}(x)$ is an L^∞-normalized Schwartz function whose Fourier transform is a smooth bump adapted to $B_{L^2/N^2}(0)$. Then we have

$$\|f\|_{L^p(\mathbb{R})} \gtrsim N^{\frac{1}{2}}\left(\frac{N^{\frac{3}{2}}}{L^2}\right)^{\frac{1}{p}}$$

since $|f(x)| \sim N^{1/2}$ on $P^C_{N^{1/2}}(0) \cap B_{CN^2/L^2}(0)$. Since $|f_I| = \phi_{N^2/L^2}$, we have

$$\left(\sum_{J\in\mathcal{J}}\sum_{I\in\mathcal{I}_J}\|f_I\|^2_{L^p(\mathbb{R})}\right)^{\frac{1}{2}} \sim N^{\frac{1}{4}}\left(\frac{N^2}{L^2}\right)^{\frac{1}{p}}.$$

Therefore (58) is sharp up to N^ε. □

7C. ***Some discussions.*** If we take $L = 1$ and $p = 4$ in Theorem 7.1, we get

$$\left\|\sum_{n=1}^{N} b_n e^{ia_n x}\right\|_{L^4(B_{N^2})} \lesssim_\varepsilon N^{\frac{1}{2}+\frac{1}{8}+\varepsilon}\|b_n\|_{\ell^2}. \tag{62}$$

On the other hand, for the Dirichlet polynomial we have, by unique factorization in $\mathbb{Z}$ and local L^2 orthogonality, that

$$\left\|\sum_{n=N+1}^{2N} b_n e^{ix\log n}\right\|_{L^4(B_{N^2})} = \left\|\sum_{m=N+1}^{2N}\sum_{n=N+1}^{2N} b_m b_n e^{ix\log(nm)}\right\|^{\frac{1}{2}}_{L^2(B_{N^2})} \lesssim_\varepsilon N^{\frac{1}{2}+\varepsilon}\|b_n\|_{\ell^2}. \tag{63}$$

Comparing (62) with (63) we see that while we can construct a generalized Dirichlet sequence that contains an AP with about $N^{1/2}$ many terms and common difference $N^{-1/2}$ so that (62) is sharp for that sequence, the Dirichlet sequence $\{\log n\}_{n=N+1}^{2N}$ does not contain such an (N^{-2}-approximate) AP and therefore allows a better estimate (63).

However we notice that the example $D_0(x) = \sum_{j=1}^{cN^{1/2}} e^{ixj/N^{1/2}}$ does not exclude the possibility that Montgomery's conjecture may hold for generalized Dirichlet polynomials. By Montgomery's conjecture for generalized Dirichlet polynomials we mean, for every $\varepsilon > 0$,

$$\left\|\sum_{n=1}^{N} b_n e^{ixa_n}\right\|_{L^p(B_T)} \lesssim_\varepsilon T^\varepsilon N^{\frac{1}{2}}(N^{\frac{p}{2}}+T)^{\frac{1}{p}}\|b_n\|_{\ell^\infty} \tag{64}$$

for every generalized Dirichlet sequence $\{a_n\}_{n=1}^N$ with $\theta = 1$. Indeed we know $|D_0(x)| \gtrsim N^{1/2}$ on $P^C_{N^{1/2}}(0)$, so

$$\|D_0\|_{L^p(B_T)} \gtrsim T^{\frac{1}{p}} N^{\frac{1}{2}-\frac{1}{2p}}.$$

On the right-hand side of (7) we have $C_\varepsilon T^\varepsilon N^{1/2}(N^{p/2}+T)^{1/p} \geq N^{1/2}T^{1/p}$. So there is no contradiction to (64). Note that if we apply Hölder's inequality $\|b_n\|_{\ell^2} \leq N^{1/2}\|b_n\|_{\ell^\infty}$ to (63) then we obtain

$$\left\|\sum_{n=N+1}^{2N} b_n e^{ix\log n}\right\|_{L^4(B_{N^2})} \lesssim_\varepsilon N^{1+\varepsilon}\|b_n\|_{\ell^\infty},$$

which is exactly (7) with $p = 4, T = N^2$. However although we know (62) is sharp (up to $C_\varepsilon N^\varepsilon$) for our example $D_0(x)$, the Hölder step $\|b_n\|_{\ell^2} \leq N^{1/2}\|b_n\|_{\ell^\infty}$ is not sharp because $D_0(x)$ has only $N^{1/2}$ many nonzero coefficients.

On the other hand we may construct a periodic generalized Dirichlet polynomial

$$f = \sum_{n=1}^{N} e^{it\frac{(N+n)}{N^2}},$$

which contradicts (64) for $p > 4$, $T > N^{2+\varepsilon_0}$ with any $\varepsilon > 0$. We notice that $|f| \gtrsim N$ on $\mathcal{N}_C(N^2\mathbb{Z})$. So

$$\|f\|_{L^p(B_T)} \gtrsim N\left(\frac{T}{N^2}\right)^{\frac{1}{p}} = N^{1-\frac{2}{p}}T^{\frac{1}{p}}.$$

Under the condition $p > 4$ we have

$$N^{1-\frac{2}{p}}T^{\frac{1}{p}} \gtrsim_{\varepsilon_0} N^{\varepsilon_1}N^{\frac{1}{2}}T^{\frac{1}{p}}$$

for some $\varepsilon_1 > 0$ depending on p. Under the condition $T > N^{2+\varepsilon_0}$ we have

$$N^{1-\frac{2}{p}}T^{\frac{1}{p}} > N^{\varepsilon_2}N$$

for some $\varepsilon_2 > 0$ depending on p. Therefore when $p > 4$ and $T > N^{2+\varepsilon_0}$ with any $\varepsilon_0 > 0$, (64) fails for the generalized Dirichlet polynomial f.

At the end of this section we discuss briefly what makes $N^{1/2}$ special. Suppose we consider the sequence $\{a_n\}_{n=1}^{N^\alpha}$ for some $\alpha \in \left(\frac{1}{2}, 1\right]$, and $\{a_n\}_{n=1}^{N}$ is a generalized Dirichlet sequence with $\theta = 1$. For simplicity we will omit constants C in the following discussion. Still we look at (L^2/N^2)-neighborhood of $\{a_n\}_{n=1}^{N^\alpha}$ with $L \geq 1$. For $L \geq N^{1/2}$, the (L^2/N^2)-neighborhood is essentially the same as the $(1/N)$-neighborhood (as long as $L \leq N$), which is an interval of length about 1. So the induction scheme in this paper fails to work for $L \geq N^{1/2}$.

Another difficulty is about the "bush" structure of $\bigcup_I (I - I)$ in the frequency space. To illustrate this, we let $L \leq N^{1/2}$, and define I, P_I as before, that is, I is the (L^2/N^2)-neighborhood of an L-segment $\{a_n\}_{n=(j-1)L+1}^{jL}$ of the sequence $\{a_n\}_{n=1}^{N^\alpha}$, and P_I denotes a fat AP of the form $P_{v_I^{-1}}^{CN/L} \cap B_{CN^2/L^2}$, where $v_I = a_{(j-1)L+2} - a_{(j-1)L+1}$ (see (15)). So now there are N^α/L many I, $v_I \sim 1/N$ are L/N^2 separated, and the maximal separation of v_I is $1/N^{2-\alpha}$. For $\alpha > \frac{1}{2}$, we no longer have an essentially linear decaying pattern of the bush $\bigcup_I (I - I)$ if $L \geq N^{1-\alpha}$, which is exploited in the proof of Lemma 5.3. To be precise, we consider the function $\sum_I 1_{I-I}(t)$, which counts the number of overlap of the sets $I - I$ at t. If $\alpha \leq \frac{1}{2}$ then we can verify that

$$\left|\sum_I 1_{I-I}(t)\right| \lesssim \frac{N/L}{|t|} \quad \text{when } \frac{1}{N} \lesssim |x| \lesssim \frac{L}{N}. \tag{65}$$

See Figure 2 for a rough graph of the function $\sum_I 1_{I-I}(t)$. However if $\alpha > \frac{1}{2}$ then we no longer have (65). This is because $\frac{1}{2}$ is the largest value for α such that for every $L \leq N^{1/2}$, the k-th intervals in all $I - I$ are within about N^{-1} distance from each other for every $1 \leq k \leq L$. For comparison, we note that for $R^{-1/2} \times R^{-1}$ caps θ that tile the R^{-1}-neighborhood of the truncated parabola, the bush $\{\theta - \theta\}$ has a

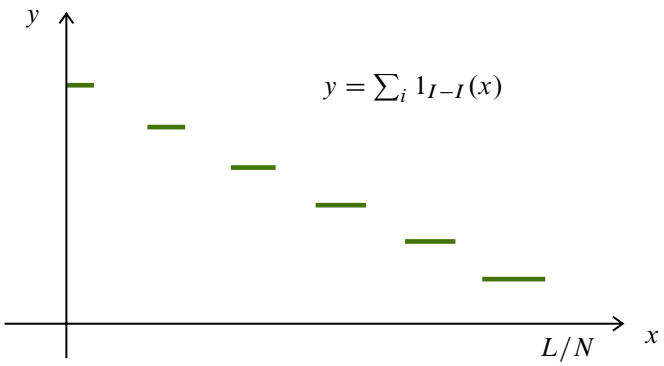

Figure 2. The overlap number of the $I - I$ has a linear decay pattern provided $L/N^{2-\alpha} \lesssim N^{-1}$. This condition is guaranteed as long as $\alpha \le \frac{1}{2}$. Controlling the overlap number of the $I - I$ outside of a certain neighborhood of the origin is a central step in Lemma 5.3.

similar linear decay pattern:

$$\left|\sum_{\theta} 1_{\theta-\theta}(x)\right| \lesssim \frac{R^{-\frac{1}{2}}}{|x|} \quad \text{when } R^{-1} \lesssim |x| \lesssim R^{-\frac{1}{2}}.$$

On the physical side, how P_I interact also becomes more complicated when $\alpha > 2$. One important property we used in the $\alpha = \frac{1}{2}$ case is that the maximal separation of v_I^{-1} (which is about $N^{1/2}$) is less than the thickness of P_I (which is about N/L) for every $1 \le L \le N^{1/2}$. However for $\alpha > \frac{1}{2}$, the maximal separation is about $N^{1-\alpha}$ which is greater than the thickness N/L for $L \ge N^{1-\alpha}$. In particular this makes the pattern of the intersection $P_I \cap P_J$ more complicated and the notion of transversal less clear.

8. Small-cap-type decoupling

In this section we prove Theorem 1.3, which is about small-cap-type decoupling inequalities in the spirit of [Demeter et al. 2020].

First we restate Theorem 1.3 but with the more general definition of generalized Dirichlet sequence. Let $\{a_n\}_{n=1}^{N^{1/2}}$ be a short generalized Dirichlet sequence with parameter $\theta \in (0, 1]$ as defined in Definition 3.1. Let L, L_1 be two integers such that $1 \le L_1 \le L \le N^{1/2}$. Denote by Ω the $\theta L^2/N^2$-neighborhood of $\{a_n\}_{n=1}^{N^{1/2}}$. We let $\{J\}_{J\in\mathcal{J}} = \{J_k\}_{k=0}^{\lfloor N^{1/2}/L_1 \rfloor}$ be the partition of Ω into unions of L_1 many consecutive intervals, that is,

$$J_k = \bigcup_{i=1}^{L_1} B_{\theta L^2/N^2}(a_{kL_1+i}).$$

Let $\{I\}_{I\in\mathcal{I}}$ be the partition of Ω into unions of L many consecutive intervals, which we called the canonical partition.

A more general version of Theorem 1.3 is the following, which we prove in the rest of this section.

Theorem 8.1. *Let $\{J\}_{J\in\mathcal{J}}$ be defined as in the above paragraph. Suppose $p\geq 4$, $\frac{1}{q}+\frac{3}{p}\leq 1$. If either of the two conditions*

(a) $L_1 = 1$,

(b) $p = q$,

is satisfied, then, for every $\varepsilon > 0$,

$$\Big\|\sum_{J\in\mathcal{J}} f_J\Big\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1)\Big(\frac{N^{\frac{1}{2}-\frac{1}{2q}-\frac{3}{2p}}L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{q}}}+\Big(\frac{N^{\frac{1}{2}}}{L_1}\Big)^{\frac{1}{2}-\frac{1}{q}}\Big)\Big(\sum_{J\in\mathcal{J}}\|f_J\|^q_{L^p(\mathbb{R})}\Big)^{\frac{1}{q}} \tag{66}$$

for all functions $f_J:\mathbb{R}\to\mathbb{C}$ with $\operatorname{supp}\widehat{f_J}\subset J$.

As a corollary we have a more general version of Corollary 1.4.

Corollary 8.2. *Let $\{a_n\}_{n=1}^{N^{1/2}}$ be a short generalized Dirichlet sequence with parameter $\theta\in(0,1]$. Suppose $p\geq 4$, $\frac{1}{q}+\frac{3}{p}\leq 1$, and $N\theta^{-1}\leq T\leq N^2\theta^{-1}$. We have, for every $\varepsilon>0$,*

$$\Big\|\sum_{n=1}^{N^{1/2}} b_n e^{ita_n}\Big\|_{L^p(B_T)} \lesssim_\varepsilon N^\varepsilon\log^C(\theta^{-1}+1)(N^{\frac{1}{2}(1+\frac{1}{p}-\frac{1}{q})}\theta^{-\frac{1}{p}}+T^{\frac{1}{p}}N^{\frac{1}{4}-\frac{1}{2q}})\|b_n\|_{\ell^q} \tag{67}$$

for every B_T and every $\{b_n\}_{n=1}^{N^{1/2}}\subset\mathbb{C}$,

To prove results of the form (66), we may use the small cap decoupling method for $\mathbb{P}^1$ developed in [Demeter et al. 2020], which is based on refined decoupling for the canonical partition, refined flat decoupling and an incidence estimate for tubes with spacing conditions. We have three analogous results in the short generalized Dirichlet sequence setting. Theorem 4.4 is the analogy of the refined canonical cap decoupling for $\mathbb{P}^1$. Now we state and prove the other two.

8A. ***An incidence estimate for fat APs.*** We start with the incidence estimate. First we introduce some notation. Suppose P, P' are fat APs such that $P = P_I(y)$ and $P' = P_{I'}(y')$ for some $I, I'\in\mathcal{I}$. We say P, P' are parallel if $I = I'$. For a collection $\mathcal{P} = \{P\}$ of fat APs, we say $x\in\mathbb{R}$ is an r-rich point if r many P contain it.

Proposition 8.3. *Let $1\leq L_1\leq L\leq N^{1/2}$ and let $\{J\}_{J\in\mathcal{J}}$, $\{I\}_{I\in\mathcal{I}}$ be defined as in the beginning of Section 8. Suppose we have a collection of fat AP $\mathcal{P} = \{P\}$ inside a fixed $P(L)$, where each $P = P_I$ for some $I\in\mathcal{I}$. Assume for every $J\in\mathcal{J}$ and every $P_J\subset P(L)$, P_J contains either M or 0 parallel $P\in\mathcal{P}$. Denote by Q_r the set of r-rich points of $\mathcal{P}$. Suppose $Q_r\neq\varnothing$. Then one of the two cases below happens:*

(1) *There exists a dyadic $s\in[1,\min\{L,N^{1/2}/L\}]$ and $M_s\in\mathbb{N}$ such that*

$$|Q_r| \lessapprox \frac{M_s}{sr^2}(\#P)|P|, \tag{68}$$

$$r \lessapprox \frac{M_s N^{1/2}}{s^2 L}, \tag{69}$$

$$M_s \lesssim sM\max\Big\{1, s\frac{L_1}{L}\Big\}. \tag{70}$$

(2) *We have*

$$|Q_r| \le |P(L)|, \tag{71}$$

$$r \lesssim (\#P)\frac{|P|}{|P(L)|}. \tag{72}$$

Here $\#P$ denotes the cardinality of $\mathcal{P}$.

Proof. For each dyadic $1 \le s \le \min\{L, N^{1/2}/L\}$, we let η_s denote a smooth bump with height 1 adapted to the annulus $|\xi| \sim (L/s)v$ in the frequency space, and let η_0 denote a smooth bump with height 1 adapted to $P_{v_1}^{C\theta L^2/N^2}(0) \cap B_{CL^2/N^{3/2}}(0)$ (which degenerates to $B_{C\theta L^2/N^2}(0)$ when $L \le N^{1/4}$) such that

$$\eta_0 + \sum_{\substack{1\le s\le \min\{L,N^{1/2}/L\},\\ s\text{: dyadic}}} \eta_s = 1 \quad \text{on } \bigcup_I (I-I).$$

For each $P \in \mathcal{P}$ we let $v_P(x)$ be a positive smooth function (with height 1) adapted to P in the physical space with frequency support in $C(I-I)$, where $P = P_I$. If we define $g = \sum_P v_P$, then we can write

$$g = g * \check{\eta}_0 + \sum_{1\le s\le \min\{L,N^{1/2}/L\}} g * \check{\eta}_s.$$

Fix $s \in [1, \min\{L, N^{1/2}/L\}]$. There exists a collection of fat APs $\mathcal{I}_s$ consisting of $I_s = P_{v_{I_s}}^{C\theta sL^2/N^2}(0) \cap B_{CL/N}(0)$ with the properties that $v_{I_s} \sim N^{-1}$ and v_{I_s} are $\sim s\theta L/N^2$ separated such that for every $I \in \mathcal{I}$, $I - I$ is contained in one and only one $I_s \in \mathcal{I}_s$. In fact we may let $v_{I_s} = v_I$ for any I with $(I-I) \subset I_s$. The cardinality of $\mathcal{I}_s$ is $N^{1/2}/(sL)$. For $I_s \in \mathcal{I}_s$ we let $\mathcal{P}_{I_s}$ be the tiling of $\mathbb{R}$ by fat APs of the form $P_{v_{I_s}^{-1}}^{\theta CsN/L} \cap B_{CN^2/(L^2\theta)}$. For every $P = P_I \in \mathcal{P}$ there exists a unique $I_s \in \mathcal{I}_s$ and $P_s \in \mathcal{P}_{I_s}$ such that $I \overset{-}{-} I \subset I_s$ and $P \subset P_s$. For every $1 \le M \lesssim s^2$, we define $\mathcal{P}_{s,M}$ be the subcollection of $\mathcal{P}$ consisting of P such that P_s contains $\sim M$ many $P' \in \mathcal{P}$. For $1 \le s \le \min\{L, N^{1/2}/L\}$ let

$$g_{s,M} = \sum_{P\in\mathcal{P}_{s,M}} v_P * \check{\eta}_s.$$

By the pigeonhole principle, for every $x \in Q_r$ there either exist an s and M_s such that $g(x) \lessapprox |g_{s,M_s}(x)|$ or $g(x) \lessapprox |g_0(x)|$. Again by the pigeonhole principle either we can find s, M_s such that, for x in a subset E of Q_r with measure $\gtrsim |Q_r|$,

$$g(x) \lessapprox |g_{s,M_s}(x)|$$

or, for x in a subset E of Q_r with measure $\gtrsim |Q_r|$,

$$g(x) \lessapprox |g_0(x)|.$$

We consider these two cases separately.

Case 1: Suppose $g(x) \lessapprox |g_{s,M_s}(x)|$ for x in a subset E of Q_r with measure $\gtrsim |Q_r|$. We write

$$g_{s,M_s} = \sum_{I_s}\sum_{P_{I_s}} \sum_{P\subset P_{I_s},\, P\in\mathcal{P}_{s,M_s}} v_P * \check{\eta}_s =: \sum_{I_s}\sum_{P_{I_s}} g_{P_{I_s}}.$$

Here the sum over P_{I_s} is over $P_{I_s} \in \mathcal{P}_{I_s}$ such that $g_{P_{I_s}}$ is nonzero.

We note that $\sum_{P_{I_s}} g_{P_{I_s}}$ with I_s varying are almost orthogonal (meaning that the Fourier support of them has $\mathcal{O}(1)$-overlap). This is because $\operatorname{supp} \widehat{g}_{P_{I_s}} \subset \left(\bigcup_{I \subset I_s} (I - I)\right) \cap \{\xi : |\xi| \sim \frac{Lv}{s}\}$, and for every distinct $I_s, I_s' \in \mathcal{I}_s$, and every $I, I' \in \mathcal{I}$ with $I \subset I_s$, $I' \subset I_s'$, the distance $d_{I,I'}$ between the $\frac{L}{s}$-th terms in I and I' satisfies

$$\frac{\theta L^2}{N^2} = \frac{s\theta L}{N^2}\frac{L}{s} \lesssim d_{I,I'} \lesssim \frac{N^{\frac{1}{2}}\theta}{N^2}\frac{L}{s} \lesssim \frac{1}{N}.$$

Therefore $\operatorname{supp} \widehat{\sum_{P_{I_s}} g_{P_{I_s}}}$ are $\mathcal{O}(1)$-overlapping.

Hence

$$|Q_r| r^2 \lessapprox \int_E g^2 \lessapprox \int_{\mathbb{R}} |g_{s,M_s}|^2 \lesssim \sum_{I_s} \int_{\mathbb{R}} \left|\sum_{P_{I_s}} g_{P_{I_s}}\right|^2.$$

We note that for $P \subset P_{I_s}$,

$$|v_P * \check{\eta}_s| \lesssim \frac{1}{s} W_{P_{I_s},100},$$

so

$$\int_{\mathbb{R}} \left|\sum_{P_{I_s}} g_{P_{I_s}}\right|^2 \lesssim \int_{\mathbb{R}} \left(\sum_{P_{I_s}} \sum_{P \subset P_{I_s},\, P \in \mathcal{P}_{s,M_s}} \frac{1}{s} W_{P_{I_s},100}\right)^2 \lesssim \sum_{P_{I_s}} \frac{M_s^2}{s^2} |P_{I_s}|.$$

Hence

$$|Q_r| r^2 \lessapprox \sum_{I_s} \sum_{P_{I_s}} |P_{I_s}| \left(\frac{M_s}{s}\right)^2.$$

Since $|P_{I_s}|/s \sim |P|$ and $\sum_{I_s} \sum_{P_{I_s}} M_s \le (\#P)$, we obtain

$$r^2 |Q_r| \lessapprox (\#P)|P| \frac{M_s}{s},$$

which is (68).

Now we show (69). We choose $x \in E$. Then

$$r \lesssim g(x) \lessapprox |g_{s,M_s}(x)| \le \sum_{I_s} \sum_{P_{I_s}} |g_{P_s}(x)| \lesssim |\mathcal{I}_s| \frac{M_s}{s} \lesssim \frac{N^{\frac{1}{2}}}{sL} \frac{M_s}{s}.$$

Finally we prove (70). When $s \le L/L_1$, every P_{I_s} is contained in a single P_J and therefore can contain $\lesssim M$ parallel $P \in \mathcal{P}$. For every P_{I_s}, there are $\lesssim s$ many $I \in \mathcal{I}$ such that there could exist P_I such that $P_I \subset P_{I_s}$, so we conclude P_{I_s} contain $\lesssim sM$ many $P \in \mathcal{P}$. When $s \ge L/L_1$, every P_{I_s} is contained in at most sL_1/L many P_J and therefore can contain $\lesssim sMsL_1/L$ many $P \in \mathcal{P}$. Hence we obtain (70).

Case 2: Suppose $g(x) \lessapprox |g_0(x)|$ for x in a subset of Q_r with measure $\gtrsim |Q_r|$. Inequality (71) is trivial since $Q_r \subset P(L)$. To show (72) we choose $x \in E$. Then

$$r \lesssim g(x) \lessapprox |g_0(x)| \lesssim (\#P) \frac{|P|}{|P(L)|},$$

where the last inequality is because

$$|g_0(x)| = |g * \check{\eta}_0(x)| \le \|g\|_{L^1} \|\check{\eta}_0\|_{L^\infty} \lesssim (\#P)|P| \frac{1}{|P(L)|} = (\#P)\frac{|P|}{|P(L)|}. \qquad \square$$

8B. ***Refined flat decoupling for fat APs.*** Next we have the following refined flat decoupling inequality for fat APs.

Proposition 8.4. *Suppose $2 \le q \le p$, and let $\{J\}_{J\in\mathcal{J}}$, $\{I\}_{I\in\mathcal{I}}$ be defined as in the beginning of Section 8. Fix $I \in \mathcal{I}$. Write $f_I = \sum_{P_I\in\mathcal{P}_I} f_{I,P_I}$ for the wave packet decomposition of f_I. Suppose that $\widetilde{\mathcal{P}}_I \subset \mathcal{P}_I$ is a collection of P_I for which f_{I,P_I} are nonzero, $\|f_{I,P_I}\|_{L^\infty(\mathbb{R})}$ are roughly constant, and for every $J \subset I$, and every P_J (in a tiling of $\mathbb{R}$), P_J contains either $\sim M$ or 0 wave packets f_{I,P_I} (in the sense that $P_I \subset P_J$). Then*

$$\Big\| \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I} \Big\|_{L^p(\mathbb{R})} \lesssim M^{\frac{1}{p}-\frac{1}{2}} \Big(\frac{L}{L_1}\Big)^{1-\frac{1}{p}-\frac{1}{q}} \Big(\sum_{J\subset I} \|f_J\|^q_{L^p(\mathbb{R})}\Big)^{\frac{1}{q}}. \tag{73}$$

Proof. Fix a P_J that contains $\sim M$ many wave packets f_{I,P_I}. We first show

$$\Big\| \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I} \Big\|_{L^p(P_J)} \lesssim M^{\frac{1}{p}-\frac{1}{2}} \Big(\frac{L}{L_1}\Big)^{1-\frac{1}{p}-\frac{1}{q}} \Big(\sum_{J\subset I} \|f_J\|^q_{L^p(W_{P_J,100})}\Big)^{\frac{1}{q}}. \tag{74}$$

Assume $\|f_{I,P_I}\|_{L^\infty(\mathbb{R})} \sim H$ for every nonzero f_{I,P_I}, $P_I \in \widetilde{\mathcal{P}}_I$. By assumption we have

$$\Big\| \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I} \Big\|_{L^p(P_J)} \lesssim H(M|P_I|)^{\frac{1}{p}}.$$

On the other hand by local L^2 orthogonality we have

$$H(M|P_I|)^{\frac{1}{2}} \lesssim \Big\| \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I} \Big\|_{L^2(P_J)} \le \|f_I\|_{L^2(P_J)} \lesssim \Big(\sum_{J\subset I} \|f_J\|^2_{L^2(W_{P_J,100})}\Big)^{\frac{1}{2}}$$

(where we used that $\big|\sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I}\big| \le |f_I|$), and by Hölder's inequality the right-hand side is bounded by

$$\Big(\frac{L}{L_1}\Big)^{\frac{1}{2}-\frac{1}{q}} |P_J|^{\frac{1}{2}-\frac{1}{p}} \Big(\sum_{J\subset I} \|f_J\|^q_{L^p(W_{P_J,100})}\Big)^{\frac{1}{q}}.$$

Noting that $|P_I|/|P_J| = L_1/L$, we conclude

$$\begin{aligned}\Big\| \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I} \Big\|_{L^2(P_J)} &\lesssim H(M|P_I|)^{\frac{1}{2}} (M|P_I|)^{\frac{1}{p}-\frac{1}{2}} \\ &\lesssim M^{\frac{1}{p}-\frac{1}{2}} \Big(\frac{L}{L_1}\Big)^{1-\frac{1}{p}-\frac{1}{q}} \Big(\sum_{J\subset I} \|f_J\|^q_{L^p(W_{P_J,100})}\Big)^{\frac{1}{q}}.\end{aligned}$$

So (74) holds. Since $q \le p$, (73) follows from (74) by raising (74) to the p-th power, summing over P_J in a tiling of $\mathbb{R}$, and applying Minkowski's inequality (see Proposition 4.2). □

8C. ***Proof of Theorem 8.1.*** Now we are ready to prove Theorem 8.1. We first show a bilinear version of Theorem 8.1 and then conclude Theorem 8.1 by a broad-narrow argument. Still let $\{J\}_{J\in\mathcal{J}}$ be defined as in the beginning of Section 8. We say two subcollections of $\mathcal{J}$, $\mathcal{J}_1$ and $\mathcal{J}_2$, are transversal if $d(J_1, J_2) \gtrsim N^{-1/2}$ for every $J_1 \in \mathcal{J}_1$, $J_2 \in \mathcal{J}_2$.

Theorem 8.5. *Suppose* $4 \le q \le p \le 6$, $\frac{1}{q} + \frac{3}{p} \le 1$. *If either of the two conditions*

(a) $L_1^{1/2-1/q} \le L^{1-3/p-1/q}$,

(b) $p = q$,

is satisfied, then, for every $\varepsilon > 0$,

$$\Big\| \prod_{i\in\{1,2\}} \Big| \sum_{J\in\mathcal{J}_i} f_J \Big|^{\frac{1}{2}} \Big\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac{2}{p}}}{L_1^{1-\frac1p-\frac1q}} + \Big(\frac{N^{\frac12}}{L_1}\Big)^{\frac12-\frac1q} \Big) \prod_{i\in\{1,2\}} \Big(\sum_{J\in\mathcal{J}_i} \|f_J\|_{L^p(\mathbb{R})}^q \Big)^{\frac{1}{2q}} \tag{75}$$

for all transversal subcollections $\mathcal{J}_1, \mathcal{J}_2$ *of* $\mathcal{J}$, *and all functions* $f_J : \mathbb{R} \to \mathbb{C}$ *with* $\operatorname{supp} \widehat{f_J} \subset J$.

Proof. By a local-to-global argument similar to Proposition 4.2, to show (75), it suffices to show, for a sufficiently large k and for every ball $B_{N^2/(\theta L^2)}$,

$$\Big\| \prod_{i\in\{1,2\}} \Big| \sum_{J\in\mathcal{J}_i} f_J \Big|^{\frac{1}{2}} \Big\|_{L^p(B_{N^2/(\theta L^2)})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac{2}{p}}}{L_1^{1-\frac1p-\frac1q}} + \Big(\frac{N^{\frac12}}{L_1}\Big)^{\frac12-\frac1q} \Big) \prod_{i\in\{1,2\}} \Big(\sum_{J\in\mathcal{J}_i} \|f_J\|_{L^p(W_{B_{N^2/(\theta L^2)},k})}^q \Big)^{\frac{1}{2q}}. \tag{76}$$

We will assume that f_J has been replaced by $f_J \psi_{B_{N^2/(\theta L^2)}}$, where $\psi_{B_{N^2/(\theta L^2)}}$ is a Schwartz function satisfying $|\psi_{B_{N^2/(\theta L^2)}}| \sim 1$ on $B_{N^2/(\theta L^2)}$, $\psi_{B_{N^2/(\theta L^2)}}$ decays rapidly away from $B_{N^2/(\theta L^2)}$, and $\operatorname{supp} \widehat{\psi}_{B_{N^2/(\theta L^2)}} \subset (-\theta L^2/N^2, \theta L^2/N^2)$. Then $f_J \psi_{B_{N^2/(\theta L^2)}}$ has Fourier support which is contained in a $(\theta L^2/N^2)$-neighborhood of J. The arguments which follow apply equally well to the $\theta L^2/N^2$ neighborhoods of J (which are contained in $2J$) as they do to J. Note also that $\|f_J \psi_{B_{N^2/(\theta L^2)}}\|_{L^p(\mathbb{R})} \lesssim_k \|f_J\|_{L^p(W_{B_{N^2/(\theta L^2)},k})}$, so abusing notation by letting f_J mean $f_J \psi_{B_{N^2/(\theta L^2)}}$ from here on in the proof, the inequality

$$\Big\| \prod_{i\in\{1,2\}} \Big| \sum_{J\in\mathcal{J}_i} f_J \Big|^{\frac{1}{2}} \Big\|_{L^p(B_{N^2/(\theta L^2)})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \Big(\frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac{2}{p}}}{L_1^{1-\frac1p-\frac1q}} + \Big(\frac{N^{\frac12}}{L_1}\Big)^{\frac12-\frac1q} \Big) \prod_{i\in\{1,2\}} \Big(\sum_{J\in\mathcal{J}_i} \|f_J\|_{L^p(\mathbb{R})}^q \Big)^{\frac{1}{2q}} \tag{77}$$

implies (76). Now we fix a $B_{N^2/\theta L^2}$ and prove (77). Write $F_1 = \sum_{J\in\mathcal{J}_1} f_J$ and $F_2 = \sum_{J\in\mathcal{J}_2} f_J$. For $i \in \{1,2\}$ we write $F_i = \sum_{P\in\mathcal{P}_i} F_{i,P}$ for the wave packet decomposition with respect to $\{I\}_{I\in\mathcal{I}}$. So

$$F_i = \sum_{I\in\mathcal{I}} F_{i,I} = \sum_{I\in\mathcal{I}} \sum_{P_I} F_{i,I,P_I} =: \sum_{P\in\mathcal{P}_i} F_{i,P}.$$

Write $\mathcal{I}_1 = \{I \in \mathcal{I} : I \subset \cup_{J\in\mathcal{J}_1} J\}$ and $\mathcal{I}_2 = \{I \in \mathcal{I} : I \subset \cup_{J\in\mathcal{J}_2} J\}$. Let $F = F_1 + F_2$. By a dyadic pigeonholing argument and rescaling which we detail in Proposition 8.6 directly following this proof, we may assume that, for every nonzero $F_{i,P}$, $\|F_{i,P}\|_{L^\infty} \sim 1$. We assume $\mathcal{P}_i$ contains only nonzero $F_{i,P}$.

By a further dyadic pigeonholing argument we may assume that for every P_J (in a tiling of $\mathbb{R}$), P_J either contains M_i or 0 many wave packets F_{i,I,P_I}, where $J \subset I$, for $i \in \{1,2\}$. Lastly, by one more dyadic pigeonholing argument we may assume that, for each $i \in \{1,2\}$, $\|F_I\|_{L^p(\mathbb{R})}$ are comparable for nonzero F_I with $I \in \mathcal{I}_i$. For dyadic $1 \le r_1, r_2 \le N^{1/2}/L$ we let Q_{r_1,r_2} denote the collection of $P(L')$ (in the tiling of $P(L)$) that intersect $\sim r_1$ many $P \in \mathcal{P}_1$, and $\sim r_2$ many $P \in \mathcal{P}_2$. Recall that $L' = (N^{1/2}L)^{1/2}$. From the refined decoupling inequality (Theorem 4.4) we have

$$\|(F_1F_2)^{\frac12}\|_{L^6(Q_{r_1,r_2})} \le \|F_1\|^{\frac12}_{L^6(Q_{r_1,r_2})}\|F_2\|^{\frac12}_{L^6(Q_{r_1,r_2})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) r_1^{\frac16} r_2^{\frac16} \prod_{i\in\{1,2\}} \left(\sum_{I\in\mathcal{I}_i}\int |F_I|^2\right)^{\frac{1}{12}}.$$

On the other hand from bilinear restriction (Proposition 3.5) we have for every $P(L') \subset Q_{r_1,r_2}$

$$\|(F_1F_2)^{\frac12}\|_{L^4(P(L'))} \lesssim_\varepsilon N^\varepsilon r_1^{\frac14} r_2^{\frac14} |P(L')|^{\frac14}$$

and thus

$$\|(F_1F_2)^{\frac12}\|_{L^4(Q_{r_1,r_2})} \lesssim_\varepsilon N^\varepsilon r_1^{\frac14} r_2^{\frac14} |Q_{r_1,r_2}|^{\frac14}.$$

Therefore by the interpolation inequality we obtain

$$\|(F_1F_2)^{\frac12}\|_{L^p(Q_{r_1,r_2})} \lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) r_1^{\frac1p} r_2^{\frac1p} |Q_{r_1,r_2}|^{\frac32-\frac12} \prod_{i\in\{1,2\}} \left(\sum_{I\in\mathcal{I}_i}\|F_I\|_{L^2}^2\right)^{\frac14-\frac1p}. \tag{78}$$

We assumed each nonzero wave packet $F_{i,P}$ satisfies $\|F_{i,P}\|_{L^\infty} \sim 1$, so

$$\sum_{I\in\mathcal{I}_i}\|F_I\|_{L^2}^2 \sim (\#P_i)|P| \sim \sum_{I\in\mathcal{I}_i}\|F_I\|_{L^p}^p,$$

where $\#P_i$ denotes the total number of nonzero wave packets in F_i, that is, $|\mathcal{P}_i|$. Hence we may rewrite (78) as

$$\begin{aligned}&\|(F_1F_2)^{\frac12}\|_{L^p(Q_{r_1,r_2})}\\ &\quad\lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1)|Q_{r_1,r_2}|^{\frac32-\frac12} \prod_{i\in\{1,2\}} \left(r_i^{\frac2p}\left(\sum_{I\in\mathcal{I}_i}\|F_I\|_{L^p}^q\right)^{\frac1q}((\#P_i)|P|)^{\frac12-\frac3p}(\#I_i)^{\frac1p-\frac1q}\right)^{\frac12},\end{aligned}$$

where $\#I_i$ denotes the total number of $I \in \mathcal{I}_i$ such that F_I is nonzero. By Proposition 8.4 we have (note that in (73) the left-hand side involves pigeonholed wave packets while the right-hand side includes all wave packets)

$$\sum_{I\in\mathcal{I}_i}\|F_I\|_{L^p}^q \lesssim M_i^{\frac qp-\frac q2}\left(\frac{L}{L_1}\right)^{q-\frac qp-1}\left(\sum_{J\in\mathcal{J}_i}\|f_J\|_{L^p}^q\right). \tag{79}$$

Therefore we conclude

$$\begin{aligned}&\|(F_1F_2)^{\frac12}\|_{L^p(Q_{r_1,r_2})}\\ &\quad\lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1)|Q_{r_1,r_2}|^{\frac32-\frac12}\\ &\qquad\times \prod_{i\in\{1,2\}}\left(r_i^{\frac2p}((\#P_i)|P|)^{\frac12-\frac3p}(\#I_i)^{\frac1p-\frac1q}M_i^{\frac1p-\frac12}\left(\frac{L}{L_1}\right)^{1-\frac1p-\frac1q}\left(\sum_{J\in\mathcal{J}_i}\|f_J\|_{L^p}^q\right)^{\frac1q}\right)^{\frac12}.\end{aligned} \tag{80}$$

So (75) follows if we may show for $i \in \{1,2\}$,

$$|Q_{r_1,r_2}|^{\frac{3}{p}-\frac{1}{2}} r_i^{\frac{2}{p}} ((\#P_i)|P|)^{\frac{1}{2}-\frac{3}{p}} (\#I_i)^{\frac{1}{p}-\frac{1}{q}} M_i^{\frac{1}{p}-\frac{1}{2}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} \lessapprox \frac{N^{\frac{1}{2}-\frac{1}{2q}-\frac{3}{2p}} L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{q}}} + \left(\frac{N^{\frac{1}{2}}}{L_1}\right)^{\frac{1}{2}-\frac{1}{q}}. \quad (81)$$

We show (81) using Proposition 8.3. Fix $i \in \{1,2\}$. We split the proof into two cases depending on which case happens in Proposition 8.3 when applied to $\{P\}_{P\in\mathcal{P}_i}$ with $r = r_i$.

Case 1: (1) in Proposition 8.3 happens. Let s, M_s be the s, M_s given in case (1) of Proposition 8.3. By (68) we have

$$\text{LHS of (81)} \lessapprox r_i^{1-\frac{4}{p}} s^{\frac{1}{2}-\frac{3}{p}} M_s^{\frac{3}{p}-\frac{1}{2}} (\#I_i)^{\frac{1}{p}-\frac{1}{q}} M_i^{\frac{1}{p}-\frac{1}{2}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}}.$$

Case 1.1: $s \le L/L_1$. Then (70) reads $M_s \lesssim sM_i$. Note that we have

$$(\#I) \gtrsim r_i$$

since we have assumed $\|F_{i,P}\|_{L^\infty} \sim 1$. Therefore by (69) and (70) we have

$$\begin{aligned}
\text{LHS of (81)} &\lessapprox \left(\frac{M_s N^{\frac{1}{2}}}{s^2 L}\right)^{1-\frac{3}{p}-\frac{1}{q}} s^{\frac{1}{2}-\frac{3}{p}} M_s^{\frac{3}{p}-\frac{1}{2}} M_i^{\frac{1}{p}-\frac{1}{2}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} \\
&= M_s^{\frac{1}{2}-\frac{1}{q}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} s^{-\frac{3}{2}+\frac{3}{p}+\frac{2}{q}} M_i^{\frac{1}{p}-\frac{1}{2}} \\
&\lesssim (sM_i)^{\frac{1}{2}-\frac{1}{q}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} s^{-\frac{3}{2}+\frac{3}{p}+\frac{2}{q}} M_i^{\frac{1}{p}-\frac{1}{2}} \\
&= M_i^{\frac{1}{p}-\frac{1}{q}} s^{-1+\frac{3}{p}+\frac{1}{q}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}}.
\end{aligned}$$

Since $p \ge q$, $\frac{1}{q} + \frac{3}{p} \le 1$, and $s, M_i \ge 1$, we conclude

$$\text{LHS of (81)} \lessapprox \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} = \frac{N^{\frac{1}{2}-\frac{1}{2q}-\frac{3}{2p}} L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{q}}}.$$

Case 1.2: $s \ge L/L_1$. This is the case where we see the two conditions in Theorem 8.5. Now (70) reads $M_s \lesssim s^2 M_i L_1/L$. By $(\#I) \gtrsim r_i$ and (69) we have

$$\begin{aligned}
\text{LHS of (81)} &\lessapprox r_i^{1-\frac{3}{p}-\frac{1}{p}} s^{\frac{1}{2}-\frac{3}{p}} M_s^{\frac{3}{p}-\frac{1}{2}} M_i^{\frac{1}{p}-\frac{1}{2}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} \\
&\lesssim \left(\frac{M_s}{s^2}\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} s^{\frac{1}{2}-\frac{3}{p}} M_s^{\frac{3}{p}-\frac{1}{2}} M_i^{\frac{1}{p}-\frac{1}{2}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} \\
&= M_s^{\frac{1}{2}-\frac{1}{q}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} s^{-\frac{3}{2}+\frac{3}{p}+\frac{2}{q}} M_i^{\frac{1}{p}-\frac{1}{2}}.
\end{aligned}$$

Plugging in (70) we obtain

$$\text{LHS of (81)} \lessapprox \left(s^2 M_i \frac{L_1}{L}\right)^{\frac{1}{2}-\frac{1}{q}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} s^{-\frac{3}{2}+\frac{3}{p}+\frac{2}{q}} M_i^{\frac{1}{p}-\frac{1}{2}}$$
$$= M_i^{\frac{1}{p}-\frac{1}{q}} s^{-\frac{1}{2}+\frac{3}{p}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}}.$$

Since $M_i \geq 1$ and $q \leq p$, we conclude

$$\text{LHS of (81)} \lessapprox s^{-\frac{1}{2}+\frac{3}{p}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}}.$$

If we use $s \leq L$, then

$$s^{-\frac{1}{2}+\frac{3}{p}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}} \leq L^{-\frac{1}{2}+\frac{3}{p}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}}.$$

We may then verify that

$$L^{-\frac{1}{2}+\frac{3}{p}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}} \leq \frac{N^{\frac{1}{2}-\frac{1}{2q}-\frac{3}{2p}} L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{q}}}$$

if and only if

$$L_1^{\frac{1}{2}-\frac{1}{q}} \leq L^{1-\frac{3}{p}-\frac{1}{q}}.$$

On the other hand if we use $s \leq N^{1/2}/L$, then

$$\text{LHS of (81)} \lessapprox \left(\frac{N^{\frac{1}{2}}}{L}\right)^{-\frac{1}{2}+\frac{3}{p}} \left(\frac{N^{\frac{1}{2}}}{L}\right)^{1-\frac{3}{p}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}} = \left(\frac{N^{\frac{1}{2}}}{L}\right)^{\frac{1}{2}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}}.$$

The last line equals

$$\left(\frac{N^{\frac{1}{2}}}{L}\right)^{\frac{1}{2}-\frac{1}{p}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}}$$

if $p = q$. In conclusion we have shown (81) holds in this case if either condition (a) or (b) is satisfied.

Case 2: (2) in Proposition 8.3 happens. By (71), (72) we have

$$\text{LHS of (81)} \lessapprox |P(L)|^{\frac{3}{p}-\frac{1}{2}} \left(\frac{(\#P_i)|P|}{|P(L)|}\right)^{\frac{2}{p}} (\#I_i)^{\frac{1}{p}-\frac{1}{q}} M_i^{\frac{1}{p}-\frac{1}{2}} \left(\frac{L}{L_1}\right)^{1-\frac{1}{p}-\frac{1}{q}} ((\#P_i)|P|)^{\frac{1}{2}-\frac{3}{p}}. \tag{82}$$

Note that we have

$$(\#P_i) \lesssim (\#I_i) M_i \frac{|P(L)|}{|P_J|} \sim (\#I_i) M_i \frac{|P(L)|}{|P|} \frac{L_1}{L}$$

since the right-hand side is the maximal number of P one can fit into a $P(L)$ under the assumption that each P_J can contain $\lesssim M_i$ many $P \in \mathcal{P}_i$. Substituting the above for M_i in (82) and simplifying the

algebra we obtain

$$\text{LHS of (81)} \lessapprox (\#I_i)^{\frac{1}{2}-\frac{1}{q}} \left(\frac{L}{L_1}\right)^{\frac{1}{2}-\frac{1}{q}}.$$

Since $\#I_i \le N^{1/2}/L$ and $q \ge 2$, we conclude

$$\text{LHS of (81)} \lessapprox \left(\frac{N^{\frac{1}{2}}}{L_1}\right)^{\frac{1}{2}-\frac{1}{q}}.$$

Hence (81) holds in this case.

In conclusion we have shown (81) and therefore (77) and (75). □

The following proposition shows that it was justified in the proof of Theorem 8.5 to treat functions $\widetilde{F}_i = \sum_{P\in\mathcal{P}_i} F_{i,P}$ whose wave packets with respect to $\{I\}_{I\in\mathcal{I}}$ satisfied certain extra assumptions. Here, each wave packet $F_{i,P}$ equals $\phi_{P_I} f_I$ for some $I \in \mathcal{I}$ and some P_I, as in the definition of wave packet decomposition from Section 5B, except we assume the extra condition that ϕ_{P_I} decays at a rate of $10^3\varepsilon^{-2}$ away from P_I.

Write $\mathcal{I}_i = \left\{I \in \mathcal{I} : I \subset \bigcup_{J\in\mathcal{J}_i} J\right\}$. For each $I \in \mathcal{I}_i$, write

$$f_I = \sum_{P_I\in\mathcal{P}_I} f_{I,P_I},$$

where $f_{I,P_I} = \phi_{P_I} f_I$ and $\mathcal{P}_I$ denotes the collection of translates of P_I which tile $\mathbb{R}$, from the definition of wave packet decomposition. Fix collections $\mathcal{P}_J$ of translates of P_J which tile $\mathbb{R}$ and with the property that $P_J \cap P_I$ is either P_I or $\varnothing$ whenever $J \subset I$. Note that the set $\mathcal{P}_J$ does not vary for $J \subset I$.

Proposition 8.6 (pigeonholing of the wave packets). *Assume the hypotheses of Theorem 8.5. There exist subsets $\widetilde{\mathcal{I}}_i \subset \mathcal{I}_i$ and $\widetilde{\mathcal{P}}_I \subset \mathcal{P}_I$ as well as integers M_i, H_i with the following properties*:

$$\left\| \prod_{i\in\{1,2\}} |F_i|^{\frac{1}{2}} \right\|_{L^p(B_{N^2/(\theta L^2)})} \lesssim \log(\theta^{-1}+1)(\log N)^2 \left\| \prod_{i\in\{1,2\}} |\widetilde{F}_i|^{\frac{1}{2}} \right\|_{L^p(B_{N^2/(\theta L^2)})} + N^{-50}(\text{RHS of (77)}),$$

where $\widetilde{F}_i = \sum_{I\in\widetilde{\mathcal{I}}_i} \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I}$,

$$\#\{P_I \in \widetilde{\mathcal{P}}_I : P_I \subset P_J\} \sim M_i \text{ or } = 0 \quad \text{for all} \quad P_J \in \mathcal{P}_J,\ J \subset I \in \widetilde{\mathcal{I}}_i, \tag{83}$$

$$\#\widetilde{\mathcal{P}}_I \sim \#\widetilde{\mathcal{P}}_{I'} \quad \text{for all } I, I' \in \widetilde{\mathcal{I}}_i, \tag{84}$$

$$\|f_{I,P_I}\|_{L^\infty(\mathbb{R})} \sim H_i \quad \text{for all } I \in \widetilde{I}_i \text{ and } P_I \in \widetilde{\mathcal{P}}_I. \tag{85}$$

It follows that, for $\widetilde{F}_I = \sum_{P_I\in\widetilde{\mathcal{P}}_I} f_{I,P_I}$ *with* $I \in \widetilde{\mathcal{I}}_i$, $\|\widetilde{F}_I\|^p_{L^p(\mathbb{R})}$ *is within a factor of* $C_\varepsilon N^\varepsilon$ *of* $H_i^p \#\{P_I \in \widetilde{\mathcal{P}}_i\}|P_I| + N^{-500}\max_{J\in\mathcal{J}_i}\|f_J\|^p_{L^p(\mathbb{R})}$.

The collection $\mathcal{P}_i$ from the proof of Theorem 8.5 is the union of the $\widetilde{\mathcal{P}}_I$, where $I \in \widetilde{\mathcal{I}}_i$.

Proof. First we will show that $\||F_1F_2|^{1/2}\|_{L^p(B_{N^2/(\theta L^2)})} \lesssim \||F_1\widetilde{F}_2|^{1/2}\|_{L^p(B_{N^2/(\theta L^2)})}$ plus the remainder term. The argument showing $\||F_1\widetilde{F}_2|^{1/2}\|_{L^p(B_{N^2/(\theta L^2)})} \lesssim \||\widetilde{F}_1\widetilde{F}_2|^{1/2}\|_{L^p(B_{N^2/(\theta L^2)})}$ plus the remainder

term is analogous, so we omit it. Split F_2 into

$$F_2 = \sum_{I\in\mathcal{I}_2}\sum_{P_I\in\mathcal{P}_I^c} f_{I,P_I} + \sum_{I\in\mathcal{I}_2}\sum_{P_I\in\mathcal{P}_I^f} f_{I,P_I}, \tag{86}$$

where the close set is

$$\mathcal{P}_I^c := \{P_I \in \mathcal{P}_I : P_I \cap N^{10}B_{N^2/(\theta L^2)} \neq \varnothing\}$$

and the far set is

$$\mathcal{P}_I^f := \{P_I \in \mathcal{P}_I : P_I \cap N^{10}B_{N^2/(\theta L^2)} = \varnothing\}.$$

Using Hölder's inequality, Cauchy–Schwarz, and Minkowski's inequality with $q \le p$, we have

$$\begin{aligned}
&\left\| \left| F_1 \sum_{I\in\mathcal{I}_2}\sum_{P_I\in\mathcal{P}_I^f} f_{I,P_I}\right|^{\frac12}\right\|_{L^p(B_{N^2/(\theta L^2)})}\\
&\lesssim \left(\frac{N}{L_1}\right)^{1-\frac1q}\left\| \sum_{J_1\in\mathcal{J}_1} |f_{J_1}|^q \sum_{I\in\mathcal{I}_2}\sum_{J_2\subset I}\left|\sum_{P_I\in\mathcal{P}_I^f}\phi_{P_I} f_{J_2}\right|^q\right|^{\frac{1}{2q}}\right\|_{L^p(B_{N^2/(\theta L^2)})}\\
&\le \left(\frac{N}{L_1}\right)^{1-\frac1q}\left\| \sum_{J_1\in\mathcal{J}_1} |f_{J_1}|^q\right|^{\frac1q}\Big\|^{\frac12}_{L^p(B_{N^2/(\theta L^2)})}\left\|\sum_{I\in\mathcal{I}_2}\sum_{J_2\subset I}\left|\sum_{P_I\in\mathcal{P}_I^f}\phi_{P_I} f_{J_2}\right|^q\right|^{\frac1q}\Big\|^{\frac12}_{L^p(B_{N^2/(\theta L^2)})}\\
&\le \left(\frac{N}{L_1}\right)^{1-\frac1q}\left(\sum_{J_1\in\mathcal{J}_1}\|f_{J_1}\|^q_{L^p(B_{N^2/(\theta L^2)})}\right)^{\frac{1}{2q}}\left(\sum_{I\in\mathcal{I}_2}\sum_{J_2\subset I}\left\|\sum_{P_I\in\mathcal{P}_I^f}\phi_{P_I} f_{J_2}\right\|^q_{L^p(B_{N^2/(\theta L^2)})}\right)^{\frac{1}{2q}}\\
&\le \left(\frac{N}{L_1}\right)^{1-\frac1q}\left(\sum_{J_1\in\mathcal{J}_1}\|f_{J_1}\|^q_{L^p(B_{N^2/(\theta L^2)})}\right)^{\frac{1}{2q}}\max_{I\in\mathcal{I}_2}\left\|\sum_{P_I\in\mathcal{P}_I^f}\phi_{P_I}\right\|^{\frac12}_{L^\infty(B_{N^2/(\theta L^2)})}\left(\sum_{J_2\in\mathcal{J}_2}\|f_{J_2}\|^q_{L^p(B_{N^2/(\theta L^2)})}\right)^{\frac{1}{2q}}\\
&\le \frac{1}{N^{100}}\prod_{i\in\{1,2\}}\left(\sum_{J_i\in\mathcal{J}_i}\|f_{J_i}\|^q_{L^p(B_{N^2/(\theta L^2)})}\right)^{\frac{1}{2q}}.
\end{aligned}$$

This takes care of the *far* portion of F_2.

For each $I \in \mathcal{I}_2$, the close set has cardinality $\#\mathcal{P}_I^c \le N^{11}$. Let

$$H_2 = \max_{I\in\mathcal{I}_2}\max_{P_I\in\mathcal{P}_I^c}\|f_{I,P_I}\|_{L^\infty(\mathbb{R})}. \tag{87}$$

By Proposition 2.3 and Hölder's inequality,

$$H_2 \le \max_{I\in\mathcal{I}_2}\|f_I\|_{L^\infty(\mathbb{R})} \lesssim N\left(\sum_{J_2\in\mathcal{J}_2}\|f_{J_2}\|^q_{L^p(\mathbb{R})}\right)^{\frac1q}. \tag{88}$$

Split the *close* part of F_2 into

$$\sum_{I\in\mathcal{I}_2}\sum_{P_I\in\mathcal{P}_I^c} f_{I,P_I} = \sum_{I\in\mathcal{I}_2}\sum_{\theta N^{-10^3}\le\lambda\le1}\sum_{P_I\in\mathcal{P}_{I,\lambda}^c} f_{I,P_I} + \sum_{I\in\mathcal{I}_2}\sum_{P_I\in\mathcal{P}_{I,s}^c} f_{I,P_I}, \tag{89}$$

where λ is a dyadic number in the range $[\theta N^{-10^3}, 1]$,

$$\mathcal{P}^c_{I,\lambda} := \Big\{ P_I \in \mathcal{P}^c_I : \|f_{I,P_I}\|_{L^\infty(\mathbb{R})} \in \Big(\frac{\lambda H_2}{2}, \lambda H_2\Big]\Big\},$$

and

$$\mathcal{P}^c_{I,s} := \Big\{ P_I \in \mathcal{P}^c_I : \|f_{I,P_I}\|_{L^\infty(\mathbb{R})} \le \frac{\theta}{2} N^{-10^3} H_2\Big\}.$$

Handle the *small* term from (89) by

$$\begin{aligned}
&\Big\| \Big| F_1 \sum_{I\in\mathcal{I}_2} \sum_{P_I\in\mathcal{P}^c_{I,s}} f_{I,P_I} \Big|^{\frac12} \Big\|_{L^p(B_{N^2/(\theta L^2)})} \\
&\quad \le \Big(\frac{N^2}{L_1 L}\Big)^{\frac12-\frac{1}{2q}} \Big\| \Big| \sum_{J_1\in\mathcal{J}_1} |f_{J_1}|^q \sum_{I\in\mathcal{I}_2} \Big| \sum_{P_I\in\mathcal{P}^c_{I,s}} f_{I,P_I} \Big|^q \Big|^{\frac{1}{2q}} \Big\|_{L^p(B_{N^2/(\theta L^2)})} \\
&\quad \le \Big(\frac{N^2}{L_1 L}\Big)^{\frac12-\frac{1}{2q}} \Big(\sum_{J_1\in\mathcal{J}_1} \|f_{J_1}\|^q_{L^p(B_{N^2/(\theta L^2)})} \Big)^{\frac{1}{2q}} \\
&\qquad\qquad \times \big(\#\mathcal{I}_2^{\frac1q} \max_{I\in\mathcal{I}_2} \#\mathcal{P}^c_I \max_{P_I\in\mathcal{P}^c_{I,s}} \|f_{I,P_I}\|_{L^\infty(B_{N^2/(\theta L^2)})} |B_{N^2/(\theta L^2)}|^{\frac2p}\big)^{\frac12} \\
&\quad \le \Big(\sum_{J_1\in\mathcal{J}_1} \|f_{J_1}\|^q_{L^p(B_{N^2/(\theta L^2)})} \Big)^{\frac{1}{2q}} \Big(\Big(\frac{N}{L}\Big)^{\frac1q} N^{11} \theta N^{-10^3} H_2 |B_{N^2/(\theta L^2)}|^{\frac2p} \Big)^{\frac12} \\
&\quad \le N^{-150} \Big(\sum_{J_1\in\mathcal{J}_1} \|f_{J_1}\|^q_{L^p(B_{N^2/(\theta L^2)})} \Big)^{\frac{1}{2q}} H_2^{\frac12} \lesssim N^{-100} \prod_{i\in\{1,2\}} \Big(\sum_{J\in\mathcal{J}_i} \|f_J\|^q_{L^p(\mathbb{R})} \Big)^{\frac{1}{2q}}.
\end{aligned}$$

Decompose the remaining term from (89) using the fact that for $J \subset I \in \mathcal{I}_2$, $P_J \in \mathcal{P}_J$, the number $\#\{P_I \in \mathcal{P}^c_{I,\lambda} : P_I \subset P_J\}$ is in $\{0, \dots, L/L_1\}$ (and does not depend on the specific $J \subset I$), which allows us to write

$$\sum_{\theta N^{-10^3}\le\lambda\le1} \sum_{I\in\mathcal{I}_2} \sum_{P_I\in\mathcal{P}^c_{I,\lambda}} f_{I,P_I} = \sum_{\theta N^{-10^3}\le\lambda\le1} \sum_{1\le 2^k\le L/L_1} \sum_{I\in\mathcal{I}_2} \sum_{P_I\in\mathcal{P}^c_{I,\lambda,k}} f_{I,P_I}, \tag{90}$$

where, for $J \subset I$,

$$\begin{aligned}
\mathcal{P}^k_J &= \{P_J \in \mathcal{P}_J : \#\{P_I \in \mathcal{P}^c_{I,\lambda} : P_I \subset P_J\} \sim 2^k\}, \\
\mathcal{P}^c_{I,\lambda,k} &= \bigcup_{P_J\in\mathcal{P}^k_J} \{P_I \in \mathcal{P}^c_{I,\lambda} : P_I \subset P_J\}.
\end{aligned}$$

Finally, note that the number of $P_J \in \mathcal{P}_J$ which intersect $N^{10} B_{N^2/(\theta L^2)}$ is bounded by $N^{10} L_1 \le N^{11}$. Further decompose the right-hand side from (90) as

$$\sum_{\theta N^{-10^3}\le\lambda\le1} \sum_{1\le 2^k\le L/L_1} \sum_{1\le 2^j\le N^{11}} \sum_{I\in\mathcal{I}^{k,j}_{2,\lambda}} \sum_{P_I\in\mathcal{P}^c_{I,\lambda,k}} f_{I,P_I}, \tag{91}$$

where, for $J \subset I$, $\mathcal{I}^{k,j}_{2,\lambda} = \{I \in \mathcal{I}_2 : \#\mathcal{P}^k_J \sim 2^j\}$.

Because 2^k, 2^j, and λ are dyadic numbers, by the pigeonhole principle, there is a choice of (k, j, λ) so that

$$\Big\| \Big| F_1 \sum_{\theta N^{-10^3} \le \lambda \le 1} \sum_{1 \le 2^k \le L/L_1} \sum_{1 \le 2^j \le N^{11}} \sum_{I \in \mathcal{I}_{2,k,j}} \sum_{P_I \in \mathcal{P}^c_{I,\lambda,k}} f_{I,P_I} \Big|^{\frac{1}{2}} \Big\|_{L^p(B_{N^2/(\theta L^2)})} \lesssim \log(\theta^{-1}+1)(\log N)^2 \Big\| \Big| F_1 \sum_{I \in \mathcal{I}^{k,j}_{2,\lambda}} \sum_{P_I \in \mathcal{P}^c_{I,\lambda,k}} f_{I,P_I} \Big|^{\frac{1}{2}} \Big\|_{L^p(B_{N^2/(\theta L^2)})}.$$

Let $\widetilde{\mathcal{I}}_2 = \mathcal{I}^{k,j}_{2,\lambda}$ and for each $I \in \widetilde{\mathcal{I}}_2$, let $\widetilde{\mathcal{P}}_I = \mathcal{P}^c_{I,\lambda,k}$. It follows from Proposition 2.3 and properties of weight functions ϕ_{P_I} and $W_{P_I} = W_{P_I,600}$ that, for each $I \in \widetilde{\mathcal{I}}_2$,

$$\begin{aligned}
\|\widetilde{F}_I\|^p_{L^p(\mathbb{R})} &= \sum_{P'_I \in \mathcal{P}_I} \Big\| \sum_{P_I \in \widetilde{\mathcal{P}}_I} f_{I,P_I} \Big\|^p_{L^p(P'_I)} \\
&\le \sum_{P'_I \in \mathcal{P}_I} \int_{P'_I} \Big| \sum_{P_I \in \widetilde{\mathcal{P}}_I} \phi_{P_I} \Big|^p \|f_I\|^p_{L^\infty(P'_I)} \le \sum_{P'_I \in \mathcal{P}_I} |P'_I| \max_{P_I \in \widetilde{\mathcal{P}}_I} \phi_{P_I}(P'_I)^p \|f_I\|^p_{L^\infty(P'_I)} \\
&\le \sum_{P'_I \in \mathcal{P}_I} |P'_I| \sum_{P_I \in \widetilde{\mathcal{P}}_I} \phi_{P_I}(P'_I)^p \|f_I\|^p_{L^\infty(P'_I)} \lesssim \sum_{P'_I \in \mathcal{P}_I} \sum_{P_I \in \widetilde{\mathcal{P}}_I} |P'_I| \|\phi_{P_I} f_I\|^p_{L^\infty(P'_I)} \\
&\lesssim \sum_{P_I \in \widetilde{\mathcal{P}}_I} \sum_{P'_I \in \mathcal{P}_I} \int |\phi_{P_I} f_I|^p W_{P'_I} \sim \sum_{P_I \in \widetilde{\mathcal{P}}_I} \int |\phi_{P_I} f_I|^p.
\end{aligned}$$

The assumption that ϕ_{P_I} decays at order $10^3\varepsilon^{-2}$ allows us to write, for each $I \in \widetilde{\mathcal{I}}_2$ and $P_I \in \widetilde{\mathcal{P}}_I$,

$$\begin{aligned}
\Big| \int |\phi_{P_I} f_I|^p - \int_{N^\varepsilon P_I} |\phi_{P_I} f_I|^p \Big| &\le C_\varepsilon N^{-1000} \|f_I\|^p_{L^p(\mathbb{R})} \\
&\le C_\varepsilon N^{-500} \max_{J \subset I} \|f_J\|^p_{L^p(\mathbb{R})}
\end{aligned}$$

and

$$\int_{N^\varepsilon P_I} |\phi_{P_I} f_I|^p \le C_\varepsilon N^\varepsilon |P_I| B_2^p \lesssim C_\varepsilon N^\varepsilon \int |f_{I,P_I}|^p,$$

which proves the final property about $\|f_{I,P_I}\|_{L^p(\mathbb{R})}$ from the proposition. □

Proof of Theorem 8.1 using Theorem 8.5. The proof resembles Section 5.1 in [Demeter et al. 2020]. First we fix (p, q) with $4 \le p \le 6$, and either $\frac{1}{q} + \frac{3}{p} = 1$ or $p = q$. Note that under such assumption we always have $p \ge q$ and $q \ge 2$. Recall that Ω is the $(\theta L^2/N^2)$-neighborhood of $\{a_n\}_{n=1}^{N^{1/2}}$, which is a union of $N^{1/2}$ many intervals of length $C\theta L^2/N^2$. We let τ denote the union of l many consecutive intervals in Ω, and write $\ell(\tau) = l$, so in this notation $\ell(I) = L$ and $\ell(J) = L_1$. Let $F = \sum_{J \in \mathcal{J}} f_J$, and denote by F_τ the Fourier projection of F to τ, that is, $\widehat{(1_\tau F)}$. Fix $K > 1$. We have the inequality

$$|F(x)| \le \sum_{\ell(\tau) = \frac{N^{1/2}}{K}} |F_\tau(x)| \le C \max_{\ell(\tau) = \frac{N^{1/2}}{K}} |F_\tau(x)| + K^C \max_{\substack{\ell(\tau_1) = \ell(\tau_2) = \frac{N^{1/2}}{K} \\ d(\tau_1,\tau_2) \gtrsim \frac{1}{KN^{1/2}}}} |F_{\tau_1} F_{\tau_2}|^{\frac{1}{2}}.$$

Iterating this (for the first term) we obtain

$$\|F\|_{L^p(\mathbb{R})}^p \lesssim C^m \sum_{\ell(\tau)=L} \|F_\tau\|_{L^p(\mathbb{R})}^p + C^m K^C \sum_{\substack{l=\frac{N^{1/2}}{K^a} \text{ for } a\in\mathbb{Z} \\ KL\le l\le N^{1/2}}} \sum_{\tau:\ell(\tau)=l} \sum_{\substack{\tau_1,\tau_2\subset\tau \\ \ell(\tau_1)=\ell(\tau_2)=K^{-1}l \\ d(\tau_1,\tau_2)\gtrsim K^{-1}l}} \|(F_{\tau_1}F_{\tau_2})^{\frac12}\|_{L^p(\mathbb{R})}^p. \tag{92}$$

Here m satisfies $N^{1/2}/K^m = L$.

By Proposition 7.2 and Hölder's inequality we have

$$\sum_{\ell(\tau)=L} \|F_\tau\|_{L^p(\mathbb{R})}^p \lesssim \sum_{\ell(\tau)=L} \left(\frac{L}{L_1}\right)^{\frac{p}{2}-1} \left(\sum_{J\subset\tau} \|F_J\|_{L^p(\mathbb{R})}^2\right)^{\frac{p}{2}} \le \sum_{\ell(\tau)=L} \left(\frac{L}{L_1}\right)^{p-1-\frac{p}{q}} \left(\sum_{J\subset\tau} \|F_J\|_{L^p(\mathbb{R})}^q\right)^{\frac{p}{q}}.$$

Since $\frac1q + \frac3p \le 1$ and $L \le N^{1/2}$, we have

$$\frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac2p}}{L_1^{1-\frac1p-\frac1q}} \ge \frac{L^{1-\frac1q-\frac1p}}{L_1^{1-\frac1p-\frac1q}}.$$

Therefore,

$$\sum_{\ell(\tau)=L} \|F_\tau\|_{L^p(\mathbb{R})}^p \lesssim \left(\frac{L}{L_1}\right)^{p-1-\frac{p}{q}} \left(\sum_{J\in\mathcal{J}} \|F_J\|_{L^p(\mathbb{R})}^q\right)^{\frac{p}{q}} \le \left(\frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac2p}}{L_1^{1-\frac1p-\frac1q}}\right)^p \left(\sum_{J\in\mathcal{J}} \|F_J\|_{L^p(\mathbb{R})}^q\right)^{\frac{p}{q}}. \tag{93}$$

Now we estimate the second term on the right-hand side of (92). Let $s = N^{1/2}/l$. Then using the change of variable $x \mapsto s^2 x$ as in the proof of Proposition 6.3, and by Theorem 8.5 we have

$$\|(F_{\tau_1}F_{\tau_2})^{\frac12}\|_{L^p(\mathbb{R})} \lesssim_\varepsilon N^\varepsilon \log^C(\tilde\theta^{-1}+1) \left(\frac{\tilde N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} \tilde L^{\frac2p}}{\tilde L_1^{1-\frac1p-\frac1q}} + \left(\frac{\tilde N^{\frac12}}{\tilde L_1}\right)^{\frac12-\frac1q}\right) \left(\sum_{J\subset\tau} \|f_J\|_{L^p(\mathbb{R})}^q\right)^{\frac1q},$$

where $\tilde N = N/s^2,\ \tilde\theta = \theta/s^2,\ \tilde L_1 = L_1,\ \tilde L = L$. Plugging in the expressions for $\tilde N, \tilde\theta, \tilde L_1, \tilde L$ we obtain

$$\begin{aligned}&\|(F_{\tau_1}F_{\tau_2})^{\frac12}\|_{L^p(\mathbb{R})} \\ &\lesssim_\varepsilon N^\varepsilon \log^C(\theta^{-1}+1) \left(s^{-1+\frac1q+\frac3p} \frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac2p}}{L_1^{1-\frac1p-\frac1q}} + s^{-\frac12+\frac1q} \left(\frac{N^{\frac12}}{L_1}\right)^{\frac12-\frac1q}\right) \left(\sum_{J\subset\tau} \|f_J\|_{L^p(\mathbb{R})}^q\right)^{\frac1q}.\end{aligned} \tag{94}$$

We let $K = N^{\varepsilon'}$ for some $\varepsilon' > 0$ which will be chosen depending on ε. Then from (93) and (94) we conclude

$$\begin{aligned}\|F\|_{L^p(\mathbb{R})} &\lesssim_{\varepsilon,\varepsilon'} N^{\varepsilon+C\varepsilon'} \log^C(\theta^{-1}+1) \Bigg(\Bigg(\sum_{\substack{s=K^a \text{ for } a\in\mathbb{Z} \\ 1\le s\le \frac{N^{1/2}}{KL}}} s^{-1+\frac1q+\frac3p} \Bigg) \frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac2p}}{L_1^{1-\frac1p-\frac1q}} \\ &\qquad + \Bigg(\sum_{\substack{s=K^a \text{ for } a\in\mathbb{Z} \\ 1\le s\le \frac{N^{1/2}}{KL}}} s^{-\frac12+\frac1q} \Bigg) \left(\frac{N^{\frac12}}{L_1}\right)^{\frac12-\frac1q} \Bigg) \left(\sum_{J\in\mathcal{J}} \|f_J\|_{L^p(\mathbb{R})}^q\right)^{\frac1q} \\ &\lesssim_{\varepsilon,\varepsilon'} N^{\varepsilon+C\varepsilon'} \log^C(\theta^{-1}+1) \left(\frac{N^{\frac12-\frac{1}{2q}-\frac{3}{2p}} L^{\frac2p}}{L_1^{1-\frac1p-\frac1q}} + \left(\frac{N^{\frac12}}{L_1}\right)^{\frac12-\frac1q} \right) \left(\sum_{J\in\mathcal{J}} \|f_J\|_{L^p(\mathbb{R})}^q\right)^{\frac1q}.\end{aligned}$$

Therefore we have shown Theorem 8.1 under condition (a) and the extra condition $\frac{1}{q}+\frac{3}{p}=1$, $p\le 6$, or under condition (b) with the extra condition $p\le 6$.

First assume (a) and we want to remove the condition $\frac{1}{q}+\frac{3}{p}=1$, $p\le 6$. First we note that it suffices to show (66) for every (p,q) with $p\ge 4$, $\frac{1}{q}+\frac{3}{p}=1$. This is because for a general (p,q) with $p\ge 4$, $\frac{1}{q}+\frac{3}{p}\le 1$ we may consider (66) with (p,q) replaced by (p,q_0), where $\frac{1}{q_0}+\frac{3}{p}=1$. Then (66) with (p,q) follows from Hölder's inequality applied in the index J to the right-hand side of (66) with (p,q_0), since $|\mathcal{J}|\lesssim N^{1/2}/L_1$. Second we note that it suffices to show (66) for every (p,q) with $4\le p\le 6$, $\frac{1}{q}+\frac{3}{p}=1$. This is because when $p\ge 6$, we always have

$$\frac{N^{\frac{1}{2}-\frac{1}{2q}-\frac{3}{2p}}L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{q}}}\ge\left(\frac{N^{\frac{1}{2}}}{L_1}\right)^{\frac{1}{2}-\frac{1}{q}}$$

and (66) reduces to

$$\Big\|\sum_{J\in\mathcal{J}}f_J\Big\|_{L^p(\mathbb{R})}\lesssim_\varepsilon N^\varepsilon\log^C(\theta^{-1}+1)\frac{N^{\frac{1}{2}-\frac{1}{2q}-\frac{3}{2p}}L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{q}}}\Big(\sum_{J\in\mathcal{J}}\|f_J\|^q_{L^p(\mathbb{R})}\Big)^{\frac{1}{q}}.$$

So (66) with $q>6$, $\frac{1}{q}+\frac{3}{p}=1$ follows from interpolating (66) with $(p,q)=(6,2)$, and with $(p,q)=(\infty,1)$. (For the interpolation of decoupling inequalities, see Exercise 9.21 of [Demeter 2020].) When $p=\infty, q=1$, (66) becomes the triangle inequality which holds trivially. Hence we have shown Theorem 8.1 under condition (a).

Now assume (b) and we want to remove the condition $p\le 6$. As in the previous paragraph, when $p\ge 6$ we always have

$$\frac{N^{\frac{1}{2}-\frac{1}{2p}-\frac{3}{2p}}L^{\frac{2}{p}}}{L_1^{1-\frac{1}{p}-\frac{1}{p}}}\ge\left(\frac{N^{\frac{1}{2}}}{L_1}\right)^{\frac{1}{2}-\frac{1}{p}},$$

and therefore (66) with $q>6$, $p=q$ follows from interpolating (see Exercise 9.21 of [Demeter 2020]) (66) with $(p,q)=(6,6)$, and with $(p,q)=(\infty,\infty)$. So Theorem 8.1 holds under condition (b) as well. □

Appendix

Corollary 1.4 can be derived from small-cap decoupling inequalities for the parabola in [Demeter et al. 2020]. This is through a transference method which we learned from James Maynard. We record a detailed proof here. The same argument would also imply Corollary 8.2 if the corresponding $\ell^q L^p$ small cap decoupling inequalities for the parabola are known.

We first recall the small-cap decoupling inequalities in [Demeter et al. 2020].

Theorem A.1 [Demeter et al. 2020]. *Suppose $\alpha\in[\frac{1}{2},1]$, and let $\Gamma=\{\gamma\}$ be the partition of $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ into R^α many $R^{-\alpha}\times R^{-1}$ rectangles γ. Assume $p=2+\frac{2}{\alpha}$. Then for every $\varepsilon>0$ we have*

$$\Big\|\sum_{\gamma\in\Gamma}f_\gamma\Big\|_{L^p(\mathbb{R}^2)}\lesssim_\varepsilon R^{\alpha(\frac{1}{2}-\frac{1}{p})+\varepsilon}\Big(\sum_\gamma\|f_\gamma\|^p_{L^p(\mathbb{R}^2)}\Big)^{\frac{1}{p}}\tag{95}$$

for every $f_\gamma:\mathbb{R}^2\to\mathbb{C}$ with supp $\widehat{f_\gamma}\subset\gamma$.

Theorem A.1 continues to hold, by essentially the same proof, with $\mathbb{P}^1$ replaced by a C^2 curve of the form $\{(x, g(x)) : x \in [0,1]\}$, with $g'(0) = 0$, $g''(x) \sim 1$ for $x \in [0,1]$. See for example Section 7 of [Bourgain and Demeter 2015] (whose argument we think actually requires a bit more regularity of the curve than C^2), or the appendix of [Guth et al. 2022]. Additionally we may interpolate (see Exercise 9.21 of [Demeter 2020]) between (95) and the elementary inequalities

$$\left\| \sum_{\gamma \in \Gamma} f_\gamma \right\|_{L^2(\mathbb{R}^2)} \lesssim \left(\sum_\gamma \| f_\gamma \|^2_{L^2(\mathbb{R}^2)} \right)^{\frac{1}{2}}$$
$$\left\| \sum_{\gamma \in \Gamma} f_\gamma \right\|_{L^\infty(\mathbb{R}^2)} \lesssim R^\alpha (\sup_\gamma \| f_\gamma \|_{L^\infty(\mathbb{R}^2)})$$

to obtain the following version of Theorem A.1.

Theorem A.2 [Demeter et al. 2020]. *Suppose G is a C^2 convex curve of the form $\{(x, g(x)) : x \in [0,1]\}$, where $g'(0) = 0$, $g''(x) \sim 1$ for $x \in [0,1]$. Suppose $\alpha \in \left[\frac{1}{2}, 1\right]$, and let $\Gamma = \{\gamma\}$ be the partition of $\mathcal{N}_{R^{-1}}(G)$ into R^α many $R^{-\alpha} \times R^{-1}$ rectangles γ. Assume $p \geq 2$. Then for every $\varepsilon > 0$ we have*

$$\left\| \sum_{\gamma \in \Gamma} f_\gamma \right\|_{L^p(\mathbb{R}^2)} \lesssim_\varepsilon R^\varepsilon (R^{\alpha(\frac{1}{2}-\frac{1}{p})} + R^{\alpha(1-\frac{1}{p})-(1+\alpha)\frac{1}{p}}) \left(\sum_\gamma \| f_\gamma \|^p_{L^p(\mathbb{R}^2)} \right)^{\frac{1}{p}} \tag{96}$$

for every $f_\gamma : \mathbb{R}^2 \to \mathbb{C}$ with supp $\widehat{f_\gamma} \subset \gamma$.

For the rest of this section we work under the assumption of Corollary 1.4. In particular $\theta = 1$. For simplicity we assume $a_1 = 0$, and $v := a_2 - a_1 = N^{-1}$. Let $1 \leq L \leq N^{1/2}$. It suffices to show (67) for $4 \leq p \leq 6$ and we assume that (since the $p > 6$ case follows from interpolating between $p = 6$ and $p = \infty$).

By (14) we may write $a_n = (n-1)/N + e_n$, where $e_n = a_n - (n-1)/N \sim (n-1)^2/N^2$. For every $t \in \mathbb{R}$ we may write it as $t_1 + t_2$, where $t_1 \in 2\pi N\mathbb{Z}$ and $t_2 \in [0, 2\pi N)$. Without loss of generality we assume $2\pi N$ divides T, so $(2\pi)^{-1} N^{-1} T \in \mathbb{Z}$. Now we may write

$$\begin{aligned} \int_0^T \left| \sum_{n=1}^{N^{1/2}} b_n e^{ita_n} \right|^p dt &= \sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, T-2\pi N]} \int_0^{2\pi N} \left| \sum_{n=1}^{N^{1/2}} b_n e^{i(t_1+t_2)(\frac{n-1}{N}+e_n)} \right|^p dt_2 \\ &= \sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, T-2\pi N]} \int_0^{2\pi N} \left| \sum_{n=1}^{N^{1/2}} b_n e^{i(t_1 e_n + t_2 \frac{n-1}{N} + t_2 e_n)} \right|^p dt_2. \end{aligned}$$

We write $e(n) = e_n$ and let $e : [1, N^{1/2}] \to \mathbb{R}$ be the piecewise linear function such that, for every $n \in \mathbb{Z} \cap [1, N^{1/2} - 1]$, $e(x)$ is linear on $[n, n+1]$ and $e(n) = e_n$. Since $e_{n+1} - e_n \sim n/N^2$, we have $|e'(x)| \lesssim 1/N^{3/2}$ for $x \in [1, N^{1/2}] \setminus \mathbb{Z}$.

By Abel's summation formula we have

$$\left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N} + t_2 e_n\right)}\right| \le \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right| + \int_1^{N^{1/2}} \left|\sum_{n=1}^{\lfloor u \rfloor} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right| |t_2 e'(u)|\, du$$
$$\lesssim \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right| + \frac{1}{N^{\frac{1}{2}}} \int_1^{N^{1/2}} \left|\sum_{n=1}^{\lfloor u \rfloor} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right| du. \quad (97)$$

The last inequality uses $t_2 \lesssim N$.

We first estimate

$$A := \sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, T-2\pi N]} \int_0^{2\pi N} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right|^p dt_2.$$

Since $e_n \lesssim \frac{1}{N}$ for every $1 \le n \le N^{1/2}$, $\sum_{n=1}^{N^{1/2}} b_n e^{i(t_1 e_n + t_2 (n-1)/N)}$ is locally constant on intervals of length N in t_1, that is, for every $y \in \mathbb{R}$,

$$\sup_{t_1 \in [y, y+2\pi N]} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right| \lesssim \left(\int_{\mathbb{R}} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right|^p W_{[y, y+2\pi N], 100}(t_1)\, dt_1\right)^{\frac{1}{p}}.$$

We note that the above is also a special case of Proposition 2.3, applied to a fat AP that is just a single interval. Since $\sum_{y \in 2\pi N\mathbb{Z} \cap [0, T-2\pi N]} W_{[y, y+2\pi N], 100}(t_1) \lesssim W_{[0,T], 100}(t_1)$, we have

$$A \lesssim \frac{1}{N} \int_{\mathbb{R}} \int_0^{2\pi N} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right|^p dt_2 W_{[0,T], 100}(t_1)\, dt_1. \quad (98)$$

We consider two cases, $T \ge N^{3/2}$ and $T \le N^{3/2}$.

Case 1: $T \ge N^{3/2}$. We observe that $\sum_{n=1}^{N^{1/2}} b_n e^{i(t_1 e_n + t_2 (n-1)/N)}$ is $2\pi N$-periodic in t_2, so we have

$$A \lesssim \frac{1}{N} \frac{N^{\frac{3}{2}}}{T} \int_{\mathbb{R}} \int_0^{TN^{-1/2}} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right|^p dt_2 W_{[0,T], 100}(t_1)\, dt_1$$

By a change of variable $t_1 \mapsto N^1 t_1$, $t_2 \mapsto N^{1/2} t_2$, we obtain

$$A \lesssim N^{\frac{1}{2}} \frac{N^{\frac{3}{2}}}{T} \int_{\mathbb{R}} \int_{\mathbb{R}} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n N + t_2 \frac{n-1}{N^{1/2}}\right)}\right|^p W_{B_{TN^{-1}}(0), 100}(t_1, t_2)\, dt_2\, dt_1.$$

Now we let $g(x)$ be a C^2 strictly convex function defined on $[0, 1]$ such that $|g((n-1)/N^{1/2}) - e_n N| \le N^{-1}/4$ for $n = 1, \dots, N^{1/2}$. (See Lemma A.3 below.) Since $N^{-1} \le T^{-1} N$, we have for every n, the ball of radius $T^{-1}N/4$ centered at $((n-1)/N^{1/2}, e_n N)$ fits in exactly one of the γ in the partition of the $T^{-1}N$ neighborhood of $G = \{(x, g(x)) : x \in [0, 1]\}$ by $N^{-1/2} \times T^{-1} N$ rectangles. Under our assumption that $T \in [N^{3/2}, N^2]$ we have $\log(N^{-1/2})/\log(T^{-1}N) \in \left[\frac{1}{2}, 1\right]$. Therefore we may apply Theorem A.2

with $R = TN^{-1}$, $R^\alpha = N^{1/2}$ to the curve G, which yields for every $T \in [N^{3/2}, N^2]$,

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n N + t_2 \frac{n-1}{N^{1/2}}\right)}\right|^p W_{B_{TN^{-1}}(0),100}(t_1,t_2)\,dt_2\,dt_1 \lesssim_\varepsilon N^\varepsilon (T^{\frac{1}{p}} N^{\frac{1}{2}-\frac{2}{p}} + T^{\frac{2}{p}} N^{\frac{1}{4}-\frac{5}{2p}})^p \|b_n\|_{\ell^p}^p. \quad (99)$$

Hence

$$A \lesssim_\varepsilon N^\varepsilon (N^{\frac{1}{2}} + T^{\frac{1}{p}} N^{\frac{1}{4}-\frac{1}{2p}})^p \|b_n\|_{\ell^p}^p.$$

Case 2: $T \le N^{3/2}$. From (98) and a change of variable we have

$$A \lesssim N^{\frac{1}{2}} \int_{\mathbb{R}}\int_0^{2\pi N^{1/2}} \left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n N + t_2 \frac{n-1}{N^{1/2}}\right)}\right|^p dt_2 W_{[0,TN^{-1}],100}(t_1)\,dt_1.$$

Since $T \le N^{3/2}$, we may bound the right-hand side trivially by

$$N^{\frac{1}{2}} \int_{\mathbb{R}}\int_{\mathbb{R}}\left|\sum_{n=1}^{N^{1/2}} b_n e^{i\left(t_1 e_n N + t_2 \frac{n-1}{N^{1/2}}\right)}\right|^p W_{B_{N^{1/2}}(0),100}(t_1,t_2)\,dt_2\,dt_1,$$

so by (99) with $T = N^{3/2}$ we have

$$A \lesssim_\varepsilon N^\varepsilon N^{\frac{1}{2}} (N^{\frac{3}{2p}} N^{\frac{1}{2}-\frac{2}{p}} + N^{\frac{3}{2}\frac{2}{p}} N^{\frac{1}{4}-\frac{5}{2p}})^p \|b_n\|_{\ell^p}^p.$$

Since $p \ge 4$ we may verify

$$N^{\frac{3}{2p}} N^{\frac{1}{2}-\frac{2}{p}} \ge N^{\frac{3}{2}\frac{2}{p}} N^{\frac{1}{4}-\frac{5}{2p}}.$$

Hence

$$A \lesssim_\varepsilon N^\varepsilon (N^{\frac{1}{2}})^p \|b_n\|_{\ell^p}^p.$$

In conclusion we have shown

$$A \lesssim_\varepsilon N^\varepsilon (N^{\frac{1}{2}} + T^{\frac{1}{p}} N^{\frac{1}{4}-\frac{1}{2p}})^p \|b_n\|_{\ell^p}^p. \quad (100)$$

Next we estimate the second term in (97). We define

$$B := \sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, N^2/L^2 - 2\pi N]} \int_0^{2\pi N} \left| \frac{1}{N^{\frac{1}{2}}} \int_1^{N^{1/2}} \left|\sum_{n=1}^{\lfloor u \rfloor} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right| du \right|^p dt_2.$$

By Minkowski's inequality we have

$$B^{\frac{1}{p}} \le \frac{1}{N^{\frac{1}{2}}} \int_1^{N^{1/2}} \left(\sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, T - 2\pi N]} \int_0^{2\pi N} \left|\sum_{n=1}^{\lfloor u \rfloor} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)}\right|^p dt_2 \right)^{\frac{1}{p}} du.$$

Then applying (100) to the expression in the brackets we obtain

$$\begin{aligned} B^{\frac{1}{p}} &\lesssim_\varepsilon N^\varepsilon \frac{1}{N^{\frac{1}{2}}} \int_0^{N^{1/2}} (N^{\frac{1}{2}} + T^{\frac{1}{p}} N^{\frac{1}{4}-\frac{1}{2p}}) \|b_n\|_{\ell^p}\, du \\ &= N^\varepsilon (N^{\frac{1}{2}} + T^{\frac{1}{p}} N^{\frac{1}{4}-\frac{1}{2p}}) \|b_n\|_{\ell^p}. \end{aligned}$$

Combining the estimates for A and B we conclude

$$\left\| \sum_{n=1}^{N^{1/2}} b_n e^{ita_n} \right\|_{L^p(B_T)} \lesssim_\varepsilon N^\varepsilon (N^{\frac{1}{2}} + T^{\frac{1}{p}} N^{\frac{1}{4}-\frac{1}{2p}}) \|b_n\|_{\ell^p}.$$

We used the following lemma in the proof above.

Lemma A.3. *Suppose $\{a_n\}_{n=1}^{N^{1/2}}$ is a short generalized Dirichlet sequence with $\theta = 1$, $a_2 - a_1 = N^{-1}$, $a_1 = 0$. Let $e_n = a_n - (n-1)/N$. Then, for every $c > 0$, there exists a C^2 curve $g : [0,1] \to \mathbb{R}$ with $g''(x) \sim 1$ for $x \in [0,1]$ such that $|g((n-1)/N^{1/2}) - e_n N| \le cN^{-1}$ for every $n = 1, \dots, N^{1/2}$.*

Proof. We first define $g_0 : [0,1] \to \mathbb{R}$ to be a C^1 piecewise quadratic polynomial with $g_0'(0) = 0$ such that g_0 restricted to $[n/N^{1/2}, (n+1)/N^{1/2}]$ is a quadratic polynomial for every $n = 0, \dots, N^{1/2} - 1$, and

$$g_0\left(\frac{n-1}{N^{\frac{1}{2}}}\right) = e_n N.$$

Since

$$\frac{N(e_{n+1} - 2e_n + e_{n-1})}{N^{-1}} \sim 1,$$

we have $g_0'' \sim 1$ on $[0,1] \setminus N^{-1/2}\mathbb{Z}$, and consequently $\|g_0\|_{L^\infty([0,1])} \lesssim 1$ because $g_0'(0) = 0$. Now we let $g = g_0 * \phi$ be the $c'N^{-1}$ mollification of g_0. Here ϕ is an L^1-normalized smooth bump adapted to $B_{c'N^{-1}}(0)$ and $c' > 0$ is sufficiently small depending on c. Then we have, for every $x \in [0,1]$,

$$g''(x) = \int_{\mathbb{R}} g_0''(y)\phi(x-y)\,dy \sim 1,$$

and

$$\left| g\left(\frac{n-1}{N^{\frac{1}{2}}}\right) - e_n N \right| \le \int_{\mathbb{R}} \left| g_0(y) - g_0\left(\frac{n-1}{N^{\frac{1}{2}}}\right) \right| \phi\left(\frac{n-1}{N^{\frac{1}{2}}} - y\right) dy \le c'N^{-1} \sup_{y\in[0,1]} |g_0'| \le cN^{-1}$$

if $c' = c/(\|g_0'\|_{L^\infty([0,1])} + 1)$. □

We can use the same approach to transfer an L^p estimate for a longer generalized Dirichlet polynomial to an L^p estimate on an exponential sum with frequency support near a C^2 convex curve.

Suppose $\{a_n\}_{n=1}^N$ is a generalized Dirichlet sequence with $\theta = 1$, $a_2 - a_1 = 1/N$, $a_1 = 0$, and let $\alpha \in \left(\frac{1}{2}, 1\right]$. As before we write $e_n = (n-1)/N \sim ((n-1)^2)/N^2$. The same calculation as above shows that

$$\int_{[0,T]} \left| \sum_{n=1}^{N^\alpha} b_n e^{ita_n} \right|^p dt \lesssim \sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, T-2\pi N]} \int_0^{2\pi N} \left| \sum_{n=1}^{N^\alpha} b_n e^{i(t_1 e_n + t_2 \frac{n-1}{N} + t_2 e_n)} \right|^p dt_2.$$

One difficulty that appears is that we cannot treat $e^{it_2 e_n}$ as an error term as before. This is because when we apply the partial summation formula we get

$$\left| \sum_{n=1}^{N^\alpha} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N} + t_2 e_n\right)} \right| \lesssim \left| \sum_{n=1}^{N^\alpha} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)} \right| + \frac{1}{N^{1-\alpha}} \int_1^{N^\alpha} \left| \sum_{n=1}^{\lfloor u \rfloor} b_n e^{i\left(t_1 e_n + t_2 \frac{n-1}{N}\right)} \right| du.$$

However now $N^{1-\alpha} > N^{\alpha}$ and we cannot estimate the second term on the right-hand side as before using the estimate for the first term and Minkowski's inequality. We could still find a C^2 convex curve such that $((n-1)/N + e_n, e_n)$ lies in an N^{-1}-neighborhood of it, but the extra e_n doesn't allow us to use the $2\pi N$-periodicity in the t_2-variable.

Another difficulty we find is the integrand is locally constant on intervals of length $N^{2-2\alpha}$ in the t_1-variable, and since $N < N^{2-2\alpha}$, that prevents us from transferring the discrete summation into $\sum_{t_1 \in 2\pi N\mathbb{Z} \cap [0, T-2\pi N]}$ into $\int_{[0,T]}$. We may though transfer the discrete sum into an integral over a fat AP $\int_{P_{2\pi N}^{N^{2-2\alpha}} \cap B_{[0,T]}}$, and that might suggest some new decoupling problems in $\mathbb{R}^2$ that might be helpful for estimating longer generalized Dirichlet polynomials.

Finally we remark that for the Dirichlet sequence $\{\log n\}_{n=N+1}^{2N}$, we may implement this transference method to higher-order approximations of $\log n$. For examples we can write

$$\left| \sum_{n=N+1}^{N+N^{\alpha}} b_n e^{it \log n} \right| = \left| \sum_{n=1}^{N^{\alpha}} b_{n+N} e^{it \log\left(1+\frac{n}{N}\right)} \right| = \left| \sum_{n=1}^{N^{\alpha}} b_{n+N} e^{it\left(\frac{n}{N} - \frac{n^2}{2N^2} + e_n'\right)} \right|,$$

where

$$e_n' := \log\left(1 + \frac{n}{N}\right) - \frac{n}{N} + \frac{n^2}{2N^2} \sim \frac{n^3}{N^3}.$$

If we write $t = t_1 + t_2 + t_3$ with $t_1 \in 2\pi N^2\mathbb{Z}$, $t_2 \in 2\pi N\mathbb{Z}$, $t_3 \in [0, 2\pi N)$, then we could transfer L^p estimates on $\sum_{n=N+1}^{N+N^{\alpha}} b_n e^{it \log n}$ to 3-dimensional L^p estimates on exponential sums with frequency supported on a nondegenerate curve in $\mathbb{R}^3$. More generally one can exploit more terms in the Taylor expansion and get higher-dimensional estimates. We do not know how much this would help with estimates on Dirichlet polynomials using decoupling techniques.

Acknowledgements

We would like to thank James Maynard for thoughtful discussions related to this paper. In particular we learned from him the transference method described in the Appendix. Guth is supported by a Simons Investigator grant.

References

[Bennett et al. 2006] J. Bennett, A. Carbery, and T. Tao, "On the multilinear restriction and Kakeya conjectures", *Acta Math.* **196**:2 (2006), 261–302. MR Zbl

[Bombieri and Zannier 2002] E. Bombieri and U. Zannier, "A note on squares in arithmetic progressions, II", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **13**:2 (2002), 69–75. MR Zbl

[Bourgain 1991] J. Bourgain, "Remarks on Montgomery's conjectures on Dirichlet sums", pp. 153–165 in *Geometric aspects of functional analysis*, edited by V. D. Milman and J. Lindenstrauss, Lecture Notes in Math. **1469**, Springer, 1991. MR Zbl

[Bourgain 1993] J. Bourgain, "On the distribution of Dirichlet sums", *J. Anal. Math.* **60** (1993), 21–32. MR Zbl

[Bourgain and Demeter 2015] J. Bourgain and C. Demeter, "The proof of the l^2 decoupling conjecture", *Ann. of Math.* (2) **182**:1 (2015), 351–389. MR Zbl

[Cordoba 1977] A. Cordoba, "The Kakeya maximal function and the spherical summation multipliers", *Amer. J. Math.* **99**:1 (1977), 1–22. MR Zbl

[Demeter 2020] C. Demeter, *Fourier restriction, decoupling, and applications*, Cambridge Stud. Adv. Math. **184**, Cambridge Univ. Press, 2020. MR Zbl

[Demeter et al. 2020] C. Demeter, L. Guth, and H. Wang, "Small cap decouplings", *Geom. Funct. Anal.* **30**:4 (2020), 989–1062. MR Zbl

[Fefferman 1973] C. Fefferman, "A note on spherical summation multipliers", *Israel J. Math.* **15** (1973), 44–52. MR Zbl

[Green 2002] B. J. Green, "Restriction and Kakeya phenomena", lecture notes, 2002, available at http://people.maths.ox.ac.uk/greenbj/papers/rkp.pdf.

[Guth et al. 2019] L. Guth, N. Solomon, and H. Wang, "Incidence estimates for well spaced tubes", *Geom. Funct. Anal.* **29**:6 (2019), 1844–1863. MR Zbl

[Guth et al. 2020] L. Guth, A. Iosevich, Y. Ou, and H. Wang, "On Falconer's distance set problem in the plane", *Invent. Math.* **219**:3 (2020), 779–830. MR Zbl

[Guth et al. 2022] L. Guth, D. Maldague, and H. Wang, "Improved decoupling for the parabola", *J. Eur. Math. Soc.* (online publication December 2022).

[Iosevich et al. 2006] A. Iosevich, S. Konyagin, M. Rudnev, and V. Ten, "Combinatorial complexity of convex sequences", *Discrete Comput. Geom.* **35**:1 (2006), 143–158. MR Zbl

[Montgomery 1971] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math. **227**, Springer, 1971. MR Zbl

[Montgomery 1994] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Reg. Conf. Ser. Math. **84**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

Received 20 Jul 2021. Revised 25 Apr 2022. Accepted 28 May 2022.

YUQIU FU: yuqiufu@mit.edu
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, United States

LARRY GUTH: lguth@math.mit.edu
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, United States

DOMINIQUE MALDAGUE: dmal@mit.edu
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, United States

Analysis & PDE

msp.org/apde

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Analysis & PDE

Volume 16 No. 10 2023