J. Eur. Math. Soc. 26, 875–917 (2024) DOI 10.4171/JEMS/1295

Larry Guth · Dominique Maldague · Hong Wang

Improved decoupling for the parabola

Received September 19, 2020; revised April 1, 2021

Abstract. We prove an (ℓ^2, L^6) decoupling inequality for the parabola with constant $(\log R)^c$. In the appendix, we present an application to the sixth-order correlation of the integer solutions to $x^2 + y^2 = m$.

Keywords. Decoupling

1. Introduction and main results

Let $f : \mathbb{R}^n \to \mathbb{C}$ be in the Schwartz class $\mathcal{S}$ with Fourier support contained in $\mathcal{N}_{R^{-1}}(\mathbb{P}^{n-1})$, the R^{-1}-neighborhood of $\mathbb{P}^{n-1} := \{(\xi, |\xi|^2) : |\xi| \le 1,\ \xi \in \mathbb{R}^{n-1}\}$. Let $\{\theta\}$ be a tiling of $\mathcal{N}_{R^{-1}}(\mathbb{P}^{n-1})$ by approximately $R^{-1/2} \times \cdots \times R^{-1/2} \times R^{-1}$ rectangular boxes θ and define $f_\theta = (\widehat{f} \chi_\theta)^\vee$.

Let $D_{n,p}(R)$ denote the smallest constant such that

$$\|f\|_{L^p(\mathbb{R}^n)} \le D_{n,p}(R) \Big(\sum_\theta \|f_\theta\|^2_{L^p(\mathbb{R}^n)} \Big)^{1/2}, \tag{1.1}$$

for any $f \in \mathcal{S}$ with supp $\widehat{f} \subset \mathcal{N}_{R^{-1}}(\mathbb{P}^{n-1})$.

A trivial estimate using Cauchy–Schwarz and the triangle inequality yields $D_{n,p}(R) \le R^{(n-1)/2}$. And we have $D_{n,p}(R) \ge 1$ by taking $f = f_\theta$. Bourgain and Demeter [3] proved that for $2 \le p \le \frac{2(n+1)}{n-1}$, $D_{n,p}(R) \le C_\epsilon R^\epsilon$ for any small $\epsilon > 0$. Such estimates are possible due to the curvature of $\mathbb{P}^{n-1}$ and are sharp up to R^ϵ-loss. The estimates have many applications in harmonic analysis, PDE and number theory. It was conjectured that $D_{n,p} \le C_p$ for $1 \le p < \frac{2(n+1)}{n-1}$.

Larry Guth: Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02142-4307, USA; lguth@math.mit.edu

Dominique Maldague: Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02142-4307, USA; dmal@mit.edu

Hong Wang: Department of Mathematics, UCLA, Los Angeles, CA 90095, USA; hongwang@math.ucla.edu

Mathematics Subject Classification (2020): 42B10

In this paper, we focus on the case $n = 2$ and write $D_p(R) = D_{n,p}(R)$. At the end point $p = 6$, Bourgain [2] proved that $D_6(R) \gtrsim (\log R)^{1/6}$. Based on the Bourgain–Demeter decoupling, Zane Li [7] proved that $D_6(R) \lesssim \exp(O(\frac{\log R \log\log\log R}{\log\log R}))$. Then, by adapting ideas from efficient congruencing, he proved [9] that $D_6(R) \lesssim \exp(O(\frac{\log R}{\log\log R}))$. This was the best previous bound for $D_6(R)$. In this paper, we prove

Theorem 1.1. *$D_6(R) \lesssim (\log R)^{c'}$ for an absolute constant c'.*

Theorem 1.1 is a corollary of our main theorem, which estimates the L^6-norm of f on a subset of $\mathbb{R}^2$.

Theorem 1.2. *There exists $c > 0$ such that the following holds. If $f \in \mathcal{S}$ has Fourier support contained in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ and $Q_R \subset \mathbb{R}^2$ is any cube of sidelength R, then*

$$\|f\|_{L^6(Q_R)}^6 \le (\log R)^c \Big(\sum_\theta \|f_\theta\|_{L^\infty(\mathbb{R}^2)}^2\Big)^2 \sum_\theta \|f_\theta\|_{L^2(\mathbb{R}^2)}^2 \quad \forall R \ge 2.$$

The proof of Theorem 1.2 is related to an incidence estimate between points and rectangles used in [4, 5]. These arguments are based on the following idea. We consider the square function $g = \sum_\theta |f_\theta|^2$, and we divide it into a high frequency part and a low frequency part. For the high frequency part, the different terms $|f_\theta|^2$ are essentially orthogonal, and this gives a powerful tool when the high frequency part of g dominates. When the low frequency part of g dominates, we try to reduce the whole problem to a similar problem at a coarser scale.

We use these tools to give a different proof of decoupling for the parabola. Compared to the previous two proofs (by Bourgain–Demeter [3] and Li [9]), our proof leans less heavily on induction on scales, and we think this is the main reason it gives a stronger estimate. In order to obtain the $(\log R)^{c'}$ bound, we also need to deal carefully with a number of technical difficulties. These include a wave packet decomposition using Gaussian partitioning of unity, carefully modifying the function at each scale and reducing to a well-spaced frequency case. We give an intuitive explanation of the argument in Section 1.

The bound $(\log R)^{c'}$ is useful compared to R^ϵ in some diophantine equation problems. Let $\Lambda_m = \{(x, y) \in \mathbb{Z}^2 : x^2 + y^2 = m\}$. In [1], Bombieri and Bourgain studied the number of solutions of the system $\lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6$ with $\lambda_j \in \Lambda_m$. In [8], Li and Bourgain applied decoupling to this problem. They were able to prove a very strong bound for the number of solutions provided that Λ_m is very large. Using our stronger estimate for $D_6(R)$, we can extend their bound to a wider range of Λ_m. We present this application in the appendix.

Another corollary of Theorem 1.1 concerns the discrete Fourier restriction on $\{(n, n^2) : n \in \mathbb{Z}\}$.

Corollary 1.3. *Let $K_p(N)$ denote the smallest constant such that for any $\{a_n\}_{|n|\le N}$,*

$$\Big\| \sum_{|n|\le N} a_n e^{2\pi i(nx+n^2t)} \Big\|_{L^p(\mathbb{T}^2)} \lesssim K_p(N) \Big(\sum_{|n|\le N} |a_n|^2\Big)^{1/2}.$$

Then $K_6(N) \lesssim (\log N)^{c'}$.

Bourgain showed in [2, (2.51), Proposition 2.36] that

$$c(\log N)^{1/6} \le K_6(N) \le \exp\Big(c\frac{\log N}{\log\log N}\Big).$$

He also asked whether $K_p(N)$ is bounded independent of N for each $p < 6$.

2. Intuitive explanation of the argument

In this section, we outline the main ideas of the proof. For simplicity, we suppress some minor technical details, but at the end we will discuss the most important technical issues that come up.

Let $\mathrm{Dec}(R)$ be the optimal constant in the decoupling inequality

$$\|f\|_{L^6(Q_R)} \le \mathrm{Dec}(R)\Big(\sum_\theta \|f_\theta\|^2_{L^6(\omega_R)}\Big)^{1/2}.$$

Here the θ denote $\sim R^{-1/2}\times R^{-1}$ approximate rectangles which partition the R^{-1}-neighborhood of $\mathbb{P}^1$ and $\widehat{f_\theta} = \widehat{f}\chi_\theta$ where f is a Schwartz function. The original arguments of Bourgain and Demeter to prove that $\mathrm{Dec}(R) \le C_\varepsilon R^\varepsilon$ involve analysis of f_τ where τ is a rectangle in a neighborhood of $\mathbb{P}^1$ and τ is at various scales between 1 and the final scale R. This is also true for the proof of Theorem 1.2, which involves analysis of f at $\sim \log R$ many scales. We use the notations $A \lesssim B$ and $A \lessapprox B$ to mean $A \le CB$ and $A \le (\log R)^C B$, respectively, for some absolute constant C.

By a standard pigeonholing argument (see §5), the decoupling inequality above follows from the estimate

$$\alpha^6|\{x\in Q_R : |f|\sim\alpha\}| \lessapprox \Big(\sum_\theta \|f_\theta\|^2_{L^\infty(Q_R)}\Big)^2 \sum_\theta \|f_\theta\|^2_{L^2(Q_R)} \tag{2.1}$$

where we may assume that for each θ,

$$\|f_\theta\|_{L^\infty(Q_R)} \sim 1 \quad \text{or} \quad f_\theta = 0 \tag{2.2}$$

and that $\|f_\theta\|_{L^p(Q_R)}$ are comparable for all non-zero f_θ and all $2 \le p \le 6$. Note that inequality (2.1) is also (roughly) the statement of Theorem 1.2. In this section, we are suppressing the weight functions localized to Q_R which are present in the L^2 norms on the right-hand side above.

Recall the reverse square function estimate for L^4, which says that

$$\alpha^4|\{x\in Q_R : |f|\sim\alpha\}| \lesssim \int_{Q_R}\Big(\sum_\theta |f_\theta|^2\Big)^2. \tag{2.3}$$

The heart of our argument involves analyzing special cases where we can upgrade (2.3) to something that implies L^6 decoupling. We will describe the simplest special case of the argument in the following subsection.

2.1. *Special case: High frequency dominance*

Consider the square function that appears on the right-hand side of (2.3), $\sum_\theta |f_\theta|^2$. Each summand $|f_\theta|^2 = f_\theta \overline{f_\theta}$ has Fourier support in $\theta - \theta$, which looks like a copy of the $R^{-1/2} \times R^{-1}$-rectangle θ that is dilated by a factor of 2 and translated to the origin. Let η be a smooth approximation of the characteristic function of the ball of radius $R^{-1/2}/\log R$ centered at the origin. Define the low frequency part as

$$\Big(\sum_\theta |f_\theta|^2\Big)_\ell := \sum_\theta |f_\theta|^2 * \check{\eta}$$

and the high frequency part by

$$\Big(\sum_\theta |f_\theta|^2\Big)_h := \sum_\theta |f_\theta|^2 - \Big(\sum_\theta |f_\theta|^2\Big)_\ell.$$

In this special case, we assume that

$$\int_{Q_R} \Big(\sum_\theta |f_\theta|^2\Big)^2 \lessapprox \int_{Q_R} \Big|\Big(\sum_\theta |f_\theta|^2\Big)_h\Big|^2.$$

The Fourier transform of $(\sum_\theta |f_\theta|^2)_h$ is supported on $\bigcup_\theta (\theta - \theta)$ intersected with the complement of the ball centered at the origin of radius $R^{-1/2}/\log R$. The $\theta - \theta$ are $\sim R^{-1/2} \times R^{-1}$-rectangles centered at the origin and oriented at $\sim R^{-1/2}$-separated angles. They all intersect in the R^{-1}-ball centered at the origin, but overlap less and less as we move away from the origin. The overlap of the $\theta - \theta$ outside of $B(R^{-1/2}/\log R)$ is $\sim \log R$. Thus, by Cauchy–Schwarz,

$$\int_{Q_R} \Big|\Big(\sum_\theta |f_\theta|^2\Big)_h\Big|^2 \lesssim (\log R) \int_{Q_R} \sum_\theta |f_\theta|^4 \tag{2.4}$$

where we used another Cauchy–Schwarz to absorb the auxiliary function $\check{\eta}$ into the implicit constant.

To summarize, we have proved so far that

$$\alpha^4 |\{x \in Q_R : |f| \sim \alpha\}| \lessapprox \int_{Q_R} \Big(\sum_\theta |f_\theta|^2\Big)^2 \lessapprox \int_{Q_R} \sum_\theta |f_\theta|^4.$$

Note that for some $x \in Q_R$,

$$\alpha \sim |f(x)| \le \sum_\theta \|f_\theta\|_{L^\infty(Q_R)} \sim \sum_\theta \|f_\theta\|^2_{L^\infty(Q_R)}$$

and $\|f_\theta\|_{L^\infty(Q_R)} \sim 1$ for at least one θ. Thus,

$$\alpha^4 |\{x \in Q_R : |f| \sim \alpha\}| \lessapprox \alpha^{-2} \Big(\sum_\theta \|f_\theta\|^2_{L^\infty(Q_R)}\Big)^2 \sum_\theta \int_{Q_R} |f_\theta|^2$$

which is the L^6-decoupling result we were aiming for and which concludes the special case. ∎

Suppose that we are not in the high frequency dominating case above, but that we have a high frequency dominance at a different scale $\tilde{R}$:

$$\int_{Q_R}\Big(\sum_{\tau}|f_\tau|^2\Big)^2 \lessapprox \int_{Q_R}\Big|\Big(\sum_{\tau}|f_\tau|^2\Big)_h\Big|^2.$$

where the τ are $\tilde{R}^{-1/2}\times\tilde{R}^{-1}$-rectangles covering the $\tilde{R}^{-1}$-neighborhood of $\mathbb{P}^1$ and the high part is with respect to this new scale. If we repeat the above argument, we obtain the (ℓ^4, L^4) result

$$\alpha^4|\{x\in Q_R : |f|\sim\alpha\}| \lessapprox \int_{Q_R}\sum_{\tau}|f_\tau|^4.$$

If we try to relate the right-hand side to a sum of L^2 norms, then

$$\alpha^4|\{x\in Q_R : |f|\sim\alpha\}| \lessapprox \Big(\max_{\tau}\|f_\tau\|^2_{L^\infty(Q_R)}\Big)\sum_{\tau}\int_{Q_R}|f_\tau|^2.$$

By L^2-orthogonality, this is equivalent to

$$\alpha^4|\{x\in Q_R : |f|\sim\alpha\}| \lessapprox \Big(\max_{\tau}\|f_\tau\|^2_{L^\infty(Q_R)}\Big)\sum_{\theta}\int_{Q_R}|f_\theta|^2.$$

The issue now is that in the special case above, we had good control over $\|f_\theta\|_{L^\infty(Q_R)}$ given in (2.2), but we do not have any corresponding estimate for $\|f_\tau\|_{L^\infty(Q_R)}$. A key part of the proof is a pruning process for the wave packets of f_τ which will allow us to control $\|f_\tau\|_{L^\infty(Q_R)}$.

2.2. *Many frequency scales*

Our argument will involve many scales, and so we introduce a sequence of intermediate scales and high-low decompositions for each scale. Denote the intermediate scales by

$$1 < R_1 < \cdots < R_k < R_{k+1} < \cdots < R_N = R.$$

We will use scales which have the property that

$$R_{k+1}/R_k \sim (\log R)^c. \tag{2.5}$$

Let $\{\tau_k\}$ denote $R_k^{-1/2}\times R_k^{-1}$-rectangles which partition the R_k^{-1}-neighborhood of $\mathbb{P}^1$. Note that for each $k = 1,\ldots,N$,

$$f = \sum_{\tau_k} f_{\tau_k}. \tag{2.6}$$

We analyze the square functions

$$g_k = \sum_{\tau_k}|f_{\tau_k}|^2.$$

Intuitively, since the first scale is $R_1 \approx 1$, by Cauchy–Schwarz we have

$$|f| \lessapprox g_1^{1/2},$$

and we should think of g_1 as being close to $|f|^2$. On the other hand, g_N is our original square function $\sum_\theta |f_\theta|^2$.

We define a high-low decomposition for g_k, building on [4,5]. As in the special case above, observe that the Fourier transform of g_k is

$$\widehat{g}_k = \sum_{\tau_k} \widehat{|f_{\tau_k}|^2} = \sum_{\tau_k} \widehat{f}_{\tau_k} * \widehat{\overline{f}}_{\tau_k}.$$

By definition, $\widehat{f}_{\tau_k}$ has support on τ_k and $\widehat{\overline{f}}_{\tau_k}$ has support on $-\tau_k$. Thus

$$\operatorname{supp} \widehat{|f_{\tau_k}|^2} \subset \tau_k - \tau_k$$

and $\tau_k - \tau_k$ is the same as τ_k translated to the origin and dilated by a factor of 2. Since the $\{\tau_k\}$ formed a partition of the neighborhood of the parabola, the Fourier support of g_k is a union of $\sim R_k^{-1/2} \times R_k^{-1}$-rectangles centered at the origin oriented at $\sim R_k^{-1/2}$-separated angles. Analogous to the discussion preceding (2.4), the intersection of all of these tubes is the R_k^{-1}-ball centered at the origin, and outside of some neighborhood of the origin, the tubes look more disjoint (or at least finitely overlapping). This is the setting for a high-low frequency decomposition.

We separate out a low frequency part of g_k and a high frequency part of g_k.

Definition 2.1. Let $\eta_k(\xi)$ be a bump function associated to the ball of radius ρ_k centered at the origin, where $\rho_k = (\log R)^{-c/2} R_k^{-1/2}$ with c in (2.5). We have $\eta_k(\xi) = 1$ on B_{ρ_k} and $\eta_k(\xi) = 0$ outside of $2B_{\rho_k}$.

Define the low frequency part of g_k by

$$\widehat{g}_{k,\ell} = \eta_k \widehat{g}_k.$$

Let the high frequency part of g_k be equal to

$$g_{k,h} = g_k - g_{k,\ell}.$$

The following lemmas describe the good features of the low frequency part and the high frequency part of g_k.

Lemma 2.2 ("High lemma"). *For any ball B_{R_k},*

$$\int_{B_{R_k}} |g_{k,h}|^2 \lesssim \rho_k^{-1} R_k^{-1/2} \sum_{\tau_k} \int |f_{\tau_k}|^4 \omega_{R_{k+1}}$$

where $1_{B_{R_{k+1}}} \leq \omega_{R_{k+1}}$ and $\widehat{\omega}_{R_{k+1}}$ is supported in the ball of radius $2R_{k+1}^{-1}$ centered at the origin.

Remark. Since $\rho_k = (\log R)^{-c} R_k^{-1/2}$, the factor $\rho_k^{-1} R_k^{-1/2}$ is bounded by $(\log R)^c$. We omit the proof sketch of Lemma 2.2, which is similar to the discussion preceding (2.4).

The function g_{k+1} is roughly locally constant on balls of radius $\sim R_{k+1}^{1/2}$. An application of local L^2 orthogonality and this locally constant property leads to

Lemma 2.3 (White lie! "Low lemma").

$$g_{k,\ell}(x) \le C_{\mathrm{low}}\, g_{k+1}(x).$$

See Lemma 2.4 for more details. In order to rigorously justify this lemma, it is important to replace g_k by an averaged version of it.

The "low lemma" tells us that

$$g_k = g_{k,\ell} + g_{k,h} \le C_{\mathrm{low}} g_{k+1} + |g_{k,h}|.$$

Therefore, either $g_k(x) \le A|g_{k,h}|$ or $g_k(x) \le \frac{A}{A-1} C_{\mathrm{low}} g_{k+1}(x)$. Here A is a parameter of size $\sim \log R$. This leads to a partition of the domain into the following sets:

$$Q_R = L \sqcup \Omega_1 \sqcup \cdots \sqcup \Omega_{N-1}.$$

Define

$$\Omega_{N-1} = \{x \in Q_R : g_{N-1}(x) \le A|g_{N-1,h}(x)|\}.$$

For $k = 1, \ldots, N-2$, define

$$\begin{aligned}\Omega_k &= \{x \in Q_R \setminus (\Omega_{k+1} \cup \cdots \cup \Omega_{N-1}) : g_k(x) \le A|g_{k,h}(x)|\} \\ &\subseteq \Big\{x \in Q_R : g_k(x) \le A|g_{k,h}(x)|,\ g_\ell \le \frac{A}{A-1} C_{\mathrm{low}} g_{\ell+1} \text{ for } k+1 \le \ell \le N-1\Big\},\end{aligned}$$

and set

$$L = Q_R \setminus (\Omega_1 \sqcup \cdots \sqcup \Omega_{N-1}) \subseteq \Big\{x \in Q_R : g_\ell \le \frac{A}{A-1} C_{\mathrm{low}} g_{\ell+1} \text{ for } 1 \le \ell \le N-1\Big\}.$$

Since we have partitioned Q_R into $\sim \log R$ many sets, it suffices to consider the cases

$$\|f\|_{L^6(Q_R)} \lesssim (\log R)\|f\|_{L^6(L)} \tag{2.7}$$

or for some k,

$$\|f\|_{L^6(Q_R)} \lesssim (\log R)\|f\|_{L^6(\Omega_k)}. \tag{2.8}$$

The case where L dominates, which means (2.7) holds, is simple because for $x \in L$, $|f(x)|^2 \lesssim g_1(x) \lesssim g_N(x) = \sum_\theta |f_\theta(x)|^2$, and so

$$\int_L |f|^6 \lesssim \int_L \Big(\sum_\theta |f_\theta|^2\Big)^3 \le \Big(\sum_\theta \|f_\theta\|^2_{L^\infty(Q_R)}\Big)^2 \sum_\theta \|f_\theta\|^2_{L^2(Q_R)},$$

which implies the conclusion of Theorem 1.2.

More quantitatively, for $x \in L$, we have the bound

$$|f(x)|^2 \lesssim g_1 \le \Big(\frac{A}{A-1} C_{\mathrm{low}}\Big)^N \sum_\theta |f_\theta(x)|^2.$$

This ultimately gives us a bound for $\mathrm{Dec}(R)$ of the form $(\frac{A}{A-1} C_{\mathrm{low}})^N$. Recall that $N \sim \log R / \log\log R$ and $A \sim \log R$. But if C_{low} is a constant bigger than 1, then C_{low}^N will be much larger than $(\log R)^C$ (although still smaller than $C_\varepsilon R^\varepsilon$). We will have to work more carefully in the low lemma to make C_{low} very close to 1. We will return to this below.

But first we discuss the case where one of the Ω_k dominates. In this case we begin by applying a broad/narrow analysis. The narrow case is handled by an induction on scales argument. For the broad case, we consider the set

$$U := \left\{x \in Q_R : |f(x)| \sim \alpha,\ |f(x)| \lessapprox \max_{\tau_1,\tau_1' \text{ non-adjacent}} |f_{\tau_1} f_{\tau_1'}|^{1/2}(x)\right\}$$

where τ_1 and τ_1' are $R_1^{-1} \times R_1^{-1/2}$ rectangles.

The L^∞ norm of g_N plays an important role in our argument, so we give it a name:

$$r = \|g_N\|_{L^\infty(Q_R)}.$$

Since $r = \|\sum_\theta |f_\theta|^2\|_{L^\infty(Q_R)} \le \sum_\theta \|f_\theta\|^2_{L^\infty(Q_R)}$, the main estimate (2.1) follows from the bound

$$\alpha^6 |U \cap \Omega_k| \lessapprox r^2 \sum_\theta \|f_\theta\|^2_{L^2(Q_R)}. \tag{2.9}$$

Focusing on the broad case allows us to use bilinear restriction, which leads to the following bound:

$$\alpha^4 |U \cap \Omega_k| \lessapprox \int_{\Omega_k} g_k^2.$$

From now on, we use the fact that $g_k \lessapprox |g_{k,h}|$ on Ω_k and proceed as in the special case above, using the "high lemma", Lemma 2.2, to obtain

$$\alpha^4 |U \cap \Omega_k| \lessapprox \max_{\tau_k} \|f_{\tau_k}\|^2_{L^\infty(Q_R)} \sum_\theta \|f_\theta\|^2_{L^2(Q_R)}.$$

2.3. *Pruning the wave packets*

Recall that we must modify the function f to get a good bound for the $\|f_{\tau_k}\|_{L^\infty(Q_R)}$. Here is the idea for modifying f. If $x \in U$, then we know that $|f(x)| \sim \alpha$. By the definition of Ω_k, we know that for $x \in \Omega_k$,

$$\sum_{\tau_k} |f_{\tau_k}(x)|^2 \le Kr, \tag{2.10}$$

where $K \lessapprox 1$ if C_{low} is sufficiently close to 1. This property is immediate from the more technical definition of Ω_k (Definition 3.27) in the proof. Inequality (2.10) implies that for $x \in U \cap \Omega_k$, the f_{τ_k} with $|f_{\tau_k}(x)| > 100Kr/\alpha$ make a small contribution to $f(x)$. More precisely, if $x \in U \cap \Omega_k$, then

$$\sum_{\tau_k : |f_{\tau_k}(x)| > 100Kr/\alpha} |f_{\tau_k}(x)| \le \frac{\alpha}{100Kr} \sum_{\tau_k} |f_{\tau_k}(x)|^2 \le \frac{\alpha}{100} \le \frac{1}{10}|f(x)|. \tag{2.11}$$

We define

$$\lambda = 100Kr/\alpha. \tag{2.12}$$

Roughly speaking, the parts of f_τ with norm bigger than λ do not make a significant contribution to f on the set $U \cap \Omega_k$. To take advantage of this observation, we divide each f_{τ_k} into wave packets, and then prune the wave packets with amplitude bigger than λ. Note that in the actual proof, we will start with pruning and define the g_k with respect to a k-pruned version of f.

The pruning process goes roughly as follows (but this account is a little oversimplified). First we expand f_{τ_k} into wave packets,

$$f_{\tau_k} = \sum_T \psi_T f_{\tau_k}. \tag{2.13}$$

Here T denotes a translate of the dual convex τ_k^* (see Definition 3.12), and the sum is over a collection of translates that tile the plane. The function ψ_T is a smooth approximation of the characteristic function of T, and the ψ_T form a partition of unity. Each $\psi_T f_{\tau_k}$ is called a *wave packet*, and it has Fourier support essentially contained in τ_k. We define $\tilde{f}_{\tau_k}$ to be the result of pruning the high amplitude wave packets from f_{τ_k}:

$$\tilde{f}_{\tau_k} = \sum_{T:\,\|\psi_T f_{\tau_k}\|_{L^\infty(Q_R)} \le \lambda} \psi_T f_{\tau_k}.$$

The Fourier support of $\tilde{f}_{\tau_k}$ is still essentially contained in τ_k. Suppose for a moment that ψ_T was just χ_T, the characteristic function of T. Then because of our pruning, $\|\tilde{f}_{\tau_k}\|_\infty \le \lambda$. Next we define $\tilde{f} = \sum_{\tau_k} \tilde{f}_{\tau_k}$.

To analyze $|U \cap \Omega_k|$, we use the argument above with $\tilde{f}$ in place of f and $\tilde{f}_{\tau_k}$ in place of f_{τ_k}. Here are the key features of f_k that makes this possible:

- The function f_k is close to f on $U \cap \Omega_k$. If ψ_T was just χ_T, then the analysis in (2.11) would show that for $x \in U \cap \Omega_k$, $|f(x) - f_k(x)| \le \frac{1}{100}\alpha$. We will ultimately define f_k in a slightly more complicated way, and we will prove this bound for $|f(x) - f_k(x)|$.
- We now have the bound $\|\tilde{f}_{\tau_k}\|_{L^\infty(Q_R)} \le \lambda \approx r/\alpha$.
- The function $\tilde{f}$ has Fourier support properties similar to those of f so that we can run the argument above. For instance, the Fourier support of $\tilde{f}_{\tau_k}$ is essentially contained in τ_k.

When we run the argument above with $\tilde{f}$ in place of f, and then plug in the bound $\|\tilde{f}_{\tau_k}\|_{L^\infty(Q_R)} \lessapprox r/\alpha$, we get the estimate

$$\alpha^4 |U \cap \Omega_k| \lessapprox (r/\alpha)^2 \int_{Q_R} \sum_\theta |f_\theta|^2. \tag{2.14}$$

This is our desired estimate (2.9).

2.4. *Delicate estimates*

There are two main sources of technical difficulties that come up in implementing the sketch above. One is to prove the low lemma with very sharp control. The other has to do with pruning wave packets, which we have to do at many different scales. To make the argument work rigorously, g_k and f_k both have to be defined in a more complex way than above.

The argument giving the bound $\mathrm{Dec}(R) \leq (\log R)^c$ is more sensitive to some constants than others. A constant that is iterated $N \sim \log R/\log\log R$ times must be very close to 1 whereas steps which are iterated $O(1)$ times can lose a power of $\log R$.

A good example is the low lemma. If we are not very careful with how we formulate the "low lemma", we will get a bound for $\mathrm{Dec}(R)$ which is much larger than $(\log R)^c$. As we discussed above, if we prove the low lemma in the form $|g_{k,\ell}(x)| \leq C_{\mathrm{low}} g_{k+1}(x)$, then we will get a bound for $\mathrm{Dec}(R)$ which is at least as big as C_{low}^N. To get our desired bound for $\mathrm{Dec}(R)$, we need C_{low} to be almost 1.

Above we gave a non-rigorous sketch of the low lemma. To get some perspective, let us now rigorously prove a version of the low lemma to get a perspective on C_{low}.

Lemma 2.4 (Baby low lemma). *Let $\eta_k(\xi)$ be a bump function defined as in Definition 2.1. Then*

$$|g_{k,\ell}| = \Big|\sum_{\tau_k} |f_{\tau_k}|^2 * \check{\eta}_k\Big| \leq 2 \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\check{\eta}_k|.$$

Proof. We write $\sum_{\tau_k} |f_{\tau_k}|^2 * \check{\eta}_k(x)$ using Fourier inversion:

$$\sum_{\tau_k} |f_{\tau_k}|^2 * \check{\eta}_k(x) = \sum_{\tau_k} \int \widehat{f}_{\tau_k} * \widehat{\bar{f}}_{\tau_k} e^{2\pi i \xi\cdot x} \eta_k(\xi)\, d\xi.$$

Now $f_{\tau_k} = \sum_{\tau_{k+1}\subset\tau_k} f_{\tau_{k+1}}$, so we can expand out the last expression to get

$$\sum_{\tau_k} \sum_{\tau_{k+1},\tau'_{k+1}\subset\tau_k} \int \widehat{f}_{\tau_{k+1}} * \widehat{\bar{f}}_{\tau'_{k+1}} e^{2\pi i \xi\cdot x} \eta_k(\xi) d\xi.$$

The point is that most of the integrals in the sum above vanish. The convolution $\widehat{f}_{\tau_{k+1}} * \widehat{\bar{f}}_{\tau'_{k+1}}$ is supported in $\tau_{k+1} - \tau'_{k+1}$, and η_k is supported in the ball of radius $2\rho_k \leq R_{k+1}^{-1/2}$ centered at the origin. Now each rectangle τ_{k+1} has dimensions $R_{k+1}^{-1} \times R_{k+1}^{-1/2}$. So $\tau_{k+1} - \tau'_{k+1}$ intersects the support of η_k only if τ'_{k+1} is equal to or adjacent to τ_{k+1}. We keep only these terms in the sum to get

$$\sum_{\tau_k} |f_{\tau_k}|^2 * \check{\eta}_k = \sum_{\tau_{k+1},\tau'_{k+1}\,\text{equal or adjacent}} (f_{\tau_{k+1}} \bar{f}_{\tau'_{k+1}}) * \check{\eta}_k.$$

For the cross terms, we note that

$$|(f_{\tau_{k+1}} \bar{f}_{\tau'_{k+1}}) * \check{\eta}_k| \leq (|f_{\tau_{k+1}}||f_{\tau'_{k+1}}|) * |\check{\eta}_k| \leq \big(\tfrac{1}{2}|f_{\tau_{k+1}}|^2 + \tfrac{1}{2}|f_{\tau'_{k+1}}|^2\big) * |\check{\eta}_k|.$$

Finally, grouping all the terms gives the desired bound:

$$\Big|\sum_{\tau_k} |f_{\tau_k}|^2 * \check{\eta}_k\Big| \leq 2 \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\check{\eta}_k|. \qquad \blacksquare$$

There are a couple of issues with this bound. One is that we have an unwanted factor of 2 on the right-hand side. A second issue is that we have a convolution on the right-hand side. If we take $g_k = \sum_{\tau_k} |f_{\tau_k}|^2$, then we have $|g_{k,\ell}| \le 2g_{k+1} * |\check{\eta}_k|$.

To deal with the factor of 2, we consider a special case when the Fourier support of f has a helpful spacing condition. Let $\Theta = \{\theta\}$ be a collection of $\sim R^{-1/2} \times R^{-1}$-rectangles contained in the R^{-1}-neighborhood of $\mathbb{P}^1$. The collection Θ has the spacing property at scale R_k if there exists a collection of $\sim R_k^{-1/2} \times R_k^{-1}$-rectangles τ_k which cover $\bigcup_{\theta\in\Theta}\theta$ and such that

$$\operatorname{dist}(\tau_k, \tau_k') \ge (\log R)^{-1} R_k^{-1/2}$$

whenever τ_k and τ_k' are distinct. If Θ has the spacing property at scales $R_1, \dots, R_{N-1}$, then say Θ is well-spaced. A well-spaced collection of rectangles θ can include most of the rectangles needed to cover the parabola, and we will be able to reduce our theorem for a general f to the case that the Fourier support of f is well-spaced. The spacing condition helps us because whenever τ_{k+1}, τ_{k+1}' are distinct, $\tau_{k+1} - \tau_{k+1}'$ is supported outside the ball of radius $(\log R)^{-1} R_{k+1}^{-1/2}$. Now we choose $\rho_k \le (\log R)^{-1} R_{k+1}^{-1/2}$, and we see that all the cross terms in Lemma 2.4 vanish. This gets rid of the factor of 2. Assuming that f obeys the spacing condition, we conclude that

$$\Big|\sum_{\tau_k} |f_{\tau_k}|^2 * \check{\eta}_k\Big| \le \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\check{\eta}_k|.$$

Next we discuss the $* |\check{\eta}_k|$ on the right-hand side. In order to deal with this factor, we define g_k in a more complicated way:

$$g_k := \sum_{\tau_k} |f_{k+1,\tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}}. \tag{2.15}$$

Here f_{k+1,τ_k} is given by pruning high amplitude wave packets from f_{τ_k}, and we will discuss it below. The function $\varphi_{\tilde{T}_{\tau_k}}$ is roughly $\frac{1}{|\tau_k^*|}\chi_{\tau_k^*}$. It is a bit bigger than this, so a more accurate model is

$$\varphi_{\tilde{T}_{\tau_k}} \sim \frac{(\log R)^c}{|\tau_k^*|}\chi_{(\log R)^c \tau_k^*}.$$

Let us see why this extra convolution helps us. In the well-spaced case, the argument above shows that

$$|g_k * \check{\eta}_k| \le \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^2 * \varphi_{\tilde{T}_{\tau_k}} * |\check{\eta}_k|. \tag{2.16}$$

On the right-hand side, τ_k denotes the parent of τ_{k+1}. We choose the functions $\varphi_{\tilde{T}_{\tau_k}}$ so that for any $\tau_{k+1} \subset \tau_k$,

$$\varphi_{\tilde{T}_{\tau_k}} * |\check{\eta}_k| \le \varphi_{\tilde{T}_{\tau_{k+1}}}. \tag{2.17}$$

With this choice, the right-hand side of (2.16) is bounded by $\sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * \varphi_{\tilde{T}_{\tau_{k+1}}} \le g_{k+1}$. So with this definition, we get $|g_{k,\ell}| \le g_{k+1}$. (We will prove this in Lemma 3.25.)

Redefining g_k in this way makes the statement of the low lemma very clean. It does have a cost though. We have to make sure that the contribution of $\varphi_{\tilde{T}_{\tau_k}}$ is not too big. To control their size, we have to choose η_k carefully, and the key bound is $\|\check{\eta}_k\|_{L^1(\mathbb{R}^2)} \leq 1 + C/\log R$, which is proved in Lemma 3.10.

Finally, let us briefly discuss pruning wave packets. Our argument involves many different scales and we have to prune wave packets at all scales. We can define f_N to be our initial function f. We decompose f_N into wave packets by combining (2.6) and (2.13),

$$f_N = \sum_{\theta} \sum_{T} \psi_T f_{N,\theta}.$$

Then we remove the wave packets with amplitude bigger than $\lambda = 100Kr/\alpha$. The resulting function is called f_{N-1}:

$$f_{N-1} = \sum_{\theta,T:\, \|\psi_T f_{N,\theta}\|_{L^\infty(\mathbb{R}^2)} \leq \lambda} \psi_T f_{N,\theta}.$$

Next, we decompose f_{N-1} into wave packets at the next scale:

$$f_{N-1} = \sum_{\tau_{N-1}} \sum_{T_{\tau_{N-1}}} \psi_{T_{\tau_{N-1}}} f_{N-1,\tau_{N-1}}.$$

Here τ_{N-1} is a rectangle of dimensions $R_{N-1}^{-1/2} \times R_{N-1}^{-1}$, and $T_{\tau_{N-1}}$ is roughly a tube which is roughly a translate of τ_{N-1}^*. We remove the wave packets with amplitude bigger than λ and call the resulting function f_{N-2}. This iterative pruning is necessary to make our argument work, but it also makes it fairly complex. In particular, since the pruning has N steps, we have to be very careful with all the estimates related to the pruning process. For example, we have to define the smooth cutoff functions ψ_T carefully.

3. Proof of Theorem 1.2: the broad, well-spaced case

The argument outlined in the above intuition section leads to the $(\log R)^c$ upper bound in Theorem 1.2 for functions which satisfy two extra properties. The function f being *broad* allows us to bound an L^4 norm of f by an L^2 norm of a square function g_k. The property that f is *well-spaced* allows us to replace Lemma 2.4 with

$$|g_{k,\ell}| \leq \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * |\check{\eta}_k|$$

(so we have no accumulated constant after iterating the inequality $\lesssim \log R$ times). Theorem 1.2 in the special case of broad, well-spaced functions f is called Proposition 3.5, which we prove in this section. In §4, we remove the assumptions on f.

3.1. Statement of Proposition 3.5

Let $f \in \mathcal{S}$ have Fourier support in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$; θ is always an approximate $R^{-1/2} \times R^{-1}$-rectangle in a neighborhood of $\mathbb{P}^1$. The property that f is broad means that $\|f\|_{L^6(Q_R)}$

is dominated by the L^6 norm of a bilinearized version of f. We state the results in terms of a parameter $\alpha > 0$ which measures this bilinearized version of f. Precisely, we make

Definition 3.1.

$$U_\alpha := \Big\{ x \in \mathbb{R}^2 : \max_{\substack{\tau,\tau' \\ \text{nonadj.}}} |f_\tau f_{\tau'}|^{1/2}(x) \sim \alpha,\ \Big(\sum_\tau |f_\tau(x)|^6\Big)^{1/6} \le (\log R)^9 \alpha \Big\}$$

where the maximum is taken over non-adjacent $\sim (\log R)^{-6} \times (\log R)^{-12}$-rectangles τ and τ'. By $\sim$ here, we mean within a factor of 2.

Our argument involves a sequence of scales R_k defined as follows:

Definition 3.2. For $k \in \mathbb{N}$, let $R_k = (\log R)^{12k}$. We analyze scales $R_1, \ldots, R_N$ where $R_N \le R < R_{N+1}$. This means that $N = \lfloor \frac{\log R}{12 \log\log R} \rfloor$.

We will not make a distinction between R_N and R since we may use Cauchy–Schwarz to trivially decouple $\sim R_N^{-1/2}$-arcs of $\mathbb{P}^1$ into $\sim R^{-1/2}$-arcs.

Definition 3.3. For each k, the notation $\Theta(R_k)$ refers to a collection of $\sim R_k^{-1/2} \times R_k^{-1}$-rectangles τ_k covering the R_k^{-1}-neighborhood of $\mathbb{P}^1$. We use Θ to denote a collection of $\sim R^{-1/2} \times R^{-1}$-rectangles partitioning the R^{-1}-neighborhood of $\mathbb{P}^1$.

Definition 3.4 (Spacing property). The collection Θ has the *spacing property* at scale R_k if there exists a collection $\Theta(R_k)$ whose union covers $\bigcup_{\theta\in\Theta} \theta$ and such that

$$\operatorname{dist}(\tau_k, \tau_k') \ge \tfrac{1}{2} R_{k+1}^{-1/2}$$

whenever $\tau_k, \tau_k' \in \Theta(R_k)$ are distinct. If Θ has the spacing property at scales $R_1, \ldots, R_{N-1}$, then say Θ is *well-spaced*. A function $f \in \mathcal{S}$ is well-spaced if $\hat{f}$ is supported in $\bigcup_{\theta\in\Theta} \theta$ for some well-spaced Θ.

See Section 4.2 for the reduction to the well-spaced case. For the rest of Section 3, τ_k will be assumed to be part of a fixed $\Theta(R_k)$ from the well-spaced definition above. Note that $\Theta(R_k)$ depends on Θ, which depends on f.

Proposition 3.5. *There exists $c \in (0, \infty)$ such that for all well-spaced $f \in \mathcal{S}$ and all $\alpha > 0$,*

$$\alpha^6 |U_\alpha \cap Q_R| \le (\log R)^c \Big(\sum_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \sum_\theta \|f_\theta\|^2_{L^2(\mathbb{R}^2)}.$$

Lemma 3.6. *For any $p \ge 1$, $\|f_\theta\|_{L^\infty(\mathbb{R}^2)} \lesssim |\theta|^{1/p} \|f_\theta\|_{L^p(\mathbb{R}^2)}$.*

Proof. Since the Fourier transform of f_θ is supported on 2θ, we can choose a smooth cutoff function ϕ_θ such that $\phi_\theta = 1$ on 2θ and $\phi_\theta = 0$ outside of 3θ. Then

$$\|f_\theta\|_{L^\infty(\mathbb{R}^2)} = \|f_\theta * \varphi_\theta\|_{L^\infty(\mathbb{R}^2)} \le \|f_\theta\|_{L^p(\mathbb{R}^2)} \|\psi_\theta\|_{L^{p'}(\mathbb{R}^2)} \lesssim |\theta|^{1/p} \|f_\theta\|_{L^p(\mathbb{R}^2)},$$

where we have used Hölder's inequality with $1/p + 1/p' = 1$. ∎

Note 3.7. For the remainder of §3, assume that we have replaced f by a constant multiple cf so that $\max_\theta \|f_\theta\|_{L^\infty(\mathbb{R}^2)} = 1$. Note that this means that α is replaced by $c\alpha$ and r is replaced by $c^2 r$, where r is defined later in Definition 3.17. The purpose of this assumption is to simplify the error terms which are often written as negative powers of R. Note for example that by Lemma 3.6,

$$R^{-50} \ll \frac{1}{\alpha^2}\Big(\max_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^3 \lesssim \frac{1}{\alpha^2}\Big(\sum_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \sum_\theta \|f_\theta\|^2_{L^2(\mathbb{R}^2)},$$

because $\alpha \sim \max_{\tau,\tau'} |f_\tau f_{\tau'}|^{1/2}$ and $|f_\tau| \lesssim R^{1/2} \max_\theta \|f_\theta\|_\infty$ for each τ. The displayed inequality is useful because we will encounter inequalities of the form

$$\alpha^4 |U_\alpha \cap Q_R| \le (\text{main term}) + R^{-50}$$

on our way to proving Proposition 3.5.

3.2. *Auxiliary functions*

There are two places described in §2 that involve auxiliary bump functions, Definition 2.1 and (2.15), which we analyze carefully in this section. To formally carry out the pruning process from f to a pruned version of f, we define $\varphi_{\tilde{T}_{\tau_k}}$, and to define the high/low decomposition of g_k, define $\check{\eta}_k$. We control the L^1 norms of these functions in Lemmas 3.10 and 3.14. This is important for achieving the $(\log R)^c$ upper bound in Proposition 3.5.

Notation 3.8. Let $\delta = \frac{1}{\log R}$.

Definition 3.9. Let $\xi \in \mathbb{R}$ and $G_0(\xi) = e^{-(\log R)^{1/2}\xi^2}$. Define the Gaussian-like function

$$G(\xi) = G_0(\xi)\chi_{[-1,1]\setminus[-\delta,\delta]}(\xi) + G_0(\delta)\chi_{[-\delta,\delta]}(\xi) - G_0(1)\chi_{[-1,1]}(\xi).$$

Define $\eta : \mathbb{R}^2 \to \mathbb{R}$ by

$$\eta(\xi_1, \xi_2) = G(0)^{-2} G(\xi_1) G(\xi_2).$$

For each k, define

$$\eta_k(\xi) = \eta(4R_{k+2}^{1/2}\xi). \tag{3.1}$$

Note that $\eta_k(\xi) = 1$ for all $|\xi_i| \le \frac{1}{4\log R} R_{k+2}^{-1/2}$ and η_k is supported in $|\xi_i| \le \frac14 R_{k+2}^{-1/2}$.

Lemma 3.10. *Let $R > 0$ be larger than a certain absolute constant. Then for each k,*

$$\|\check{\eta}_k\|_1 \le 1 + \frac{c}{\log R}.$$

Proof. It suffices to show the claim for η since η_k is equal to η composed with an affine transformation. Since $G(0)^{-1} = (e^{\delta^{3/2}} - e^{-\delta^{-1/2}})^{-1} \le 1 + \frac{1}{\log R}$ and Fourier transforms of products factor, it suffices to show that

$$\|\check{G}\|_1 \le 1 + \frac{c}{\log R}.$$

Define the functions

$$G_1(\xi) = (G_0(\xi) - G_0(\delta))\chi_{[-\delta,\delta]}(\xi),$$
$$G_2(\xi) = G_0(\xi)\chi_{[-1,1]^c}(\xi) + G_0(1)\chi_{[-1,1]}(\xi)$$

and note that

$$G(\xi) = G_0(\xi) - G_1(\xi) - G_2(\xi).$$

Since $\check{G}_0 \geq 0$, $\|\check{G}_0\|_1 = \|G_0\|_\infty = 1$, which means it suffices to show that

$$\|\check{G}_1\|_1 \lesssim \frac{1}{\log R} \quad \text{and} \quad \|\check{G}_2\|_1 \lesssim \frac{1}{\log R}.$$

Observe that G_1 and G_2 are continuous, L^1 functions (though not differentiable at a few points). G_1 is Riemann-integrable, so for each $x \neq 0$, we can use integration by parts to compute $x\check{G}_1(x)$ as the inverse Fourier transform of a function. For G_2, the same is true after a limiting argument to approximate G_2 by Riemann-integrable functions.

First, we have

$$\begin{aligned}
\|\check{G}_1\|_1 &= \|(1+x^2)^{-1/2}(1+x^2)^{1/2}\check{G}_1\|_1 \\
&\lesssim \|(1+x^2)^{1/2}\check{G}_1\|_2 \\
&\lesssim \|G_1\|_2 + \|x\check{G}_1\|_2 \\
&\lesssim (\delta|1-G_0(\delta)|^2)^{1/2} + (\log R)^{1/2}\left(\int_{[-\delta,\delta]} |\xi e^{-(\log R)^{1/2}\xi^2}|^2\, d\xi\right)^{1/2} \\
&\lesssim \delta^{1/2}(\log R)^{1/2}\delta^2 + (\log R)^{1/2}\delta^{3/2} \sim \frac{1}{\log R}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|\check{G}_2\|_1 &\lesssim \|G_2\|_2 + \|x\check{G}_2\|_2 \\
&\lesssim G_0(1) + (\log R)^{1/2}\left(\int_{[-1,1]^c} |\xi e^{-(\log R)^{1/2}\xi^2}|^2\, d\xi\right)^{1/2} \\
&\lesssim e^{-(\log R)^{1/2}} + (\log R)^{1/8}\left(\int_{[-\delta^{-1/4},\delta^{-1/4}]^c} e^{-\xi^2/2}\, d\xi\right)^{1/2} \ll \frac{1}{\log R}. \qquad \blacksquare
\end{aligned}$$

Definition 3.11. Let $\rho : \mathbb{R}^2 \to \mathbb{R}$ be given by $\rho(\xi) = \eta(\delta\xi)$ where η is defined in Definition 3.9 and $\delta = \frac{1}{\log R}$. For each scale R_k and each $\sim R_k^{-1/2} \times R_k^{-1}$-rectangle τ_k, let

$$\rho_{\tau_k} = \rho \circ \ell_{\tau_k}$$

where ℓ_{τ_k} is an affine transformation mapping the smallest ellipse containing $2\tau_k$ to $\mathbb{B}$.

Definition 3.12. If τ is a symmetric convex set with center $C(\tau)$, then the dual of τ is defined as

$$\tau^* = \{x : |x \cdot (y - C(\tau))| \leq 1,\ \forall y \in \tau\}.$$

Definition 3.13. For each $R_1^{-1/2} \times R_1^{-1}$-rectangle τ_1, let

$$\varphi_{\tilde{T}_{\tau_1}}(x) = \sup_{y \in x + (\log R)^{11}\tau_1^*} |\check{\rho}_{\tau_1}(y)|.$$

For each $2 \le k \le N$ and each $R_k^{-1/2} \times R_k^{-1}$-rectangle τ_k, define $\varphi_{\tilde{T}_{\tau_k}}$ inductively by

$$\varphi_{\tilde{T}_{\tau_k}}(x) = \max\Big(\sup_{y \in x + (\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}(y)|, \varphi_{\tilde{T}_{\tau_{k-1}}} * |\check{\eta}_{k-1}|(x)\Big),$$

where τ_{k-1} is the $R_{k-1}^{-1/2} \times R_{k-1}^{-1}$-rectangle containing τ_k.

Lemma 3.14. *For each k and τ_k,*

$$\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)} \lesssim (\log R)^{\tilde{c}}.$$

The implicit constant is uniform in k and τ_k. The letter $\tilde{c}$ means a uniform constant and varies from place to place in later lemmas.

Proof. If $k = 1$, then $\|\varphi_{\tilde{T}_{\tau_1}}\|_{L^1(\mathbb{R}^2)} \lesssim (\log R)^{\tilde{c}}$ is clear from the definition. Recall the definition of $\varphi_{\tilde{T}_{\tau_k}}$ for $k \ge 2$ to be

$$\varphi_{\tilde{T}_{\tau_k}}(x) = \max\Big(\sup_{y \in x + (\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}(y)|, \varphi_{\tilde{T}_{\tau_{k-1}}} * |\check{\eta}_{k-1}|(x)\Big),$$

where τ_{k-1} is the $R_{k-1}^{-1/2} \times R_{k-1}^{-1}$-rectangle containing τ_k. Let A_k be the set on which

$$\sup_{y \in x + (\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}(y)| \ge \varphi_{\tilde{T}_{\tau_{k-1}}} * |\check{\eta}_{k-1}|(x)$$

and let B_k be the complement of A_k. Note that

$$\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)} = \Big\| \sup_{y \in \cdot + (\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}(y)| \Big\|_{L^1(A_k)} + \big\|\varphi_{\tilde{T}_{\tau_{k-1}}} * |\check{\eta}_{k-1}|\big\|_{L^1(B_k)}$$

If

$$\Big\| \sup_{y \in \cdot + (\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}(y)| \Big\|_{L^1(A_k)} \ge \delta \big\|\varphi_{\tilde{T}_{\tau_{k-1}}} * |\check{\eta}_{k-1}|\big\|_{L^1(B_k)},$$

then

$$\begin{aligned} \|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)} &\le 2\delta^{-1} \Big\| \sup_{y \in \cdot + (\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}(y)| \Big\|_{L^1(\mathbb{R}^2)} \\ &= 2\delta^{-1} \Big\| \sup_{y \in \cdot + B(0, (\log R)^{11})} |\check{\rho}(y)| \Big\|_{L^1(\mathbb{R}^2)} \lesssim (\log R)^c \end{aligned}$$

since $|\check{\rho}|$ is bounded by $(\log R)^c$ and is rapidly decaying outside the ball of radius $(\log R)^c$ centered at the origin.

If not, then

$$\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)} \le (1 + \delta) \big\|\varphi_{\tilde{T}_{\tau_{k-1}}} * |\check{\eta}_{k-1}|\big\|_{L^1(B_k)}.$$

By Young's convolution inequality and Lemma 3.10,

$$\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)} \le (1+\delta)^2 \|\varphi_{\tilde{T}_{\tau_{k-1}}}\|_{L^1(\mathbb{R}^2)}.$$

If we iterate this argument for $\le N \lesssim \frac{\log R}{\log\log R}$ times, then we also obtain the desired conclusion. ■

Definition 3.15 (Gaussian partition of unity). First define

$$\psi_0(x) = c\int_Q g(x-y)\,dy$$

where Q is the unit cube $[-1/2, 1/2]^2$, $g(x) = e^{-|x|^2}$, and $c = (\int g)^{-1}$. Note that for every x,

$$\sum_{n\in\mathbb{Z}^2} \psi_0(x-n) = \sum_{n\in\mathbb{Z}^2} \int_{n+Q} g(x-y)\,dy = 1.$$

Let T be any rectangle in $\mathbb{R}^2$. Let $A : \mathbb{R}^2 \to \mathbb{R}^2$ be an affine transformation mapping T to Q. Define the Gaussian bump function adapted to T ψ_T by

$$\psi_T(x) = c|T|^{-1}\int_T g(A(x-y))\,dy.$$

(There are several different affine transformations A taking T to Q, but they all give the same function ψ_T because the Gaussian g and the area form are invariant under the affine automorphisms of Q.)

If $\mathbb{T}$ is a set of congruent rectangles T tiling the plane, then

$$\sum_{T\in\mathbb{T}} \psi_T(x) = 1,$$

and so the Gaussians $\{\psi_T\}_{T\in\mathbb{T}}$ form a partition of unity.

The Fourier transform of ψ_T is

$$\widehat{\psi}_T(\xi) = c|T|^{-1}\int_T |A|^{-1}\widehat{g}((A^{-1})^t\xi)e^{-2\pi i\xi\cdot y}\,dy$$

and satisfies

$$|\widehat{\psi}_T(\xi)| \le |T|e^{-|(A^{-1})^t\xi|^2}.$$

If $x \notin 100\sqrt{\log R}\,T$, then

$$|\psi_T(x)| \le ce^{-\inf_{y\in T}|x-y|^2} \le R^{-1000}.$$

If $\xi \notin 100\sqrt{\log R}\,T^*$, then

$$|\widehat{\psi}_T(\xi)| \le c|T|R^{-10^4}.$$

In the rest of the paper, a function being *essentially supported in* S means that $|\cdot| \le R^{-1000}$ off of S and the function rapidly decays away from S. A weight function w_S being *localized to* S means that $w_S \sim 1$ on S and $|w_S| \le R^{-1000}$ off of $(\log R)S$.

3.3. Pruning wave packets

We define pruned versions of the function f and the intermediate square functions $\sum_{\tau_k}|f_{\tau_k}|^2$. The pruning process depends on the parameter $\alpha > 0$ which measures the bilinearized version of f, and a new parameter $r > 0$ (defined below) related to the final square function $\sum_\theta |f_\theta|^2$.

Definition 3.16. Define

$$g_N = \sum_\theta |f_\theta|^2 * \varphi_{\tilde{T}_\theta}$$

where $\varphi_{\tilde{T}_\theta}$ is defined in Definition 3.13 for $\theta = \tau_N$.

Definition 3.17. Define

$$r = \|g_N\|_{L^\infty(\mathbb{R}^2)}. \tag{3.2}$$

Note that

$$r \lesssim (\log R)^{\tilde{c}} \sum_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)} \tag{3.3}$$

for the constant $\tilde{c}$ in Lemma 3.14.

Notation 3.18. The parameter λ measures the ratio between r and α:

$$\lambda = (\log R)^m \frac{r}{\alpha} \tag{3.4}$$

where the exponent m is sufficiently large as required by the proof of Lemma 3.23 and Lemma 3.28 and Proposition 3.5.

Definition 3.19. For each rectangle τ_k, we write T_{τ_k} for a translate of $(\log R)^9\tau_k^*$. We let $\mathbb{T}_{\tau_k}$ be a tiling of the plane by rectangles T_{τ_k}. In Definition 3.15, we defined a Gaussian partition of unity associated to such a tiling:

$$\sum_{T_{\tau_k}\in\mathbb{T}_{\tau_k}} \psi_{T_{\tau_k}}(x) = 1.$$

Definition 3.20 (Defining f_{k,τ_k} with respect to λ). Let

$$\mathbb{T}_{\tau_{N-1},\lambda} = \{T_{\tau_{N-1}} \in \mathbb{T}_{\tau_{N-1}} : \|\psi_{T_{\tau_{N-1}}} f_{\tau_{N-1}}\|_{L^\infty(\mathbb{R}^2)} \le \lambda\}.$$

Define

$$f_{N-1,\tau_{N-1}} = \sum_{T_{\tau_{N-1}}\in\mathbb{T}_{\tau_{N-1},\lambda}} \psi_{T_{\tau_{N-1}}} f_{\tau_{N-1}}$$

and note that

$$\text{ess sup}\, \widehat{f}_{N-1,\tau_{N-1}} \subset (1+(\log R)^{-8})\tau_{N-1}.$$

Also define

$$f_{N-1,\tau_{N-2}} = \sum_{\tau_{N-1}\subset\tau_{N-2}} f_{N-1,\tau_{N-1}}.$$

Now we define f_{k,τ_k} and $f_{k,\tau_{k-1}}$, starting with $k = N-2$ and going down to $k = 1$. Let $\mathbb{T}_{\tau_k,\lambda} = \{T_{\tau_k} \in \mathbb{T}_{\tau_k} : \|\psi_{T_{\tau_k}} f_{k+1,\tau_k}\|_{L^\infty(\mathbb{R}^2)} \le \lambda\}$ and define

$$f_{k,\tau_k} = \sum_{T_{\tau_k} \in \mathbb{T}_{\tau_k,\lambda}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}, \tag{3.5}$$

$$f_{k,\tau_{k-1}} = \sum_{\tau_k \subset \tau_{k-1}} f_{k,\tau_k}. \tag{3.6}$$

Lemma 3.21 (Properties of f_{k,τ_k}). (1) $|f_{k,\tau_k}(x)| \le |f_{k+1,\tau_k}(x)|$.

(2) $\|f_{k,\tau_k}\|_{L^\infty(\mathbb{R}^2)} \le C(\log R)^2\lambda + R^{-1000}$.

(3) $\operatorname{ess\,supp} \widehat{f}_{k,\tau_k} \subset (1 + (\log R)^{-8})\tau_k$.

(4) $\operatorname{ess\,supp} \widehat{f}_{k,\tau_{k-1}} \subset (1 + (\log R)^{-10})\tau_{k-1}$.

Proof. The first property follows straight from the definition.

The second property follows because $\sum_{T_{\tau_k} \in \mathbb{T}_{\tau_k}} \psi_{T_{\tau_k}}$ is a partition of unity, and

$$f_{k,\tau_k} = \sum_{T_{\tau_k} \in \mathbb{T}_{\tau_k,\lambda} \subset \mathbb{T}_{\tau_k}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}.$$

Now consider the L^∞ bound in the third property. We write

$$f_{k,\tau_k}(x) = \sum_{T_{\tau_k} \in \mathbb{T}_{\tau_k,\lambda},\, x \in (\log R) T_{\tau_k}} \psi_{T_{\tau_k}} f_{k+1,\tau_k} + \sum_{T_{\tau_k} \in \mathbb{T}_{\tau_k,\lambda},\, x \notin (\log R) T_{\tau_k}} \psi_{T_{\tau_k}} f_{k+1,\tau_k}.$$

The first sum has at most $C(\log R)^2$ terms, and each term has norm bounded by λ by the definition of $\mathbb{T}_{\tau_k,\lambda}$. By the normalization in Note 3.7, it follows easily that

$$|f_{k+1,\tau_k}(x)| \le R^{100}. \tag{3.7}$$

But if $x \notin (\log R) T_{\tau_k}$, then $\psi_{T_{\tau_k}}(x) \le e^{-(\log R)^2} \le R^{-2000}$. Moreover, as T_{τ_k} gets further away from x, $\psi_{T_{\tau_k}}(x)$ is rapidly decaying. Therefore, the second sum has norm at most R^{-1000}.

The fourth and fifth properties depend on the essential Fourier support of $\psi_{T_{\tau_k}}$ (and on similar trivial bounds as (3.7)). Recall from Definition 3.19 that T_{τ_k} is a translate of $(\log R)^9 \tau_k^*$. Because of this factor $(\log R)^9$, the essential Fourier support of $\psi_{T_{\tau_k}}$ is contained in $100\sqrt{\log R}(\log R)^{-9}\tau_k$ (see Definition 3.15).

Initiate a 2-step induction with base case $k = N$: $f_{N,\theta}$ has essential Fourier support in $(1 + (\log R)^{-8})\theta$ because of the above definition. Then

$$f_{N,\tau_{N-1}} = \sum_{\theta \subset \tau_{N-1}} f_{N,\theta}$$

has essential Fourier support in $\bigcup_{\theta \subset \tau_{N-1}} (1 + (\log R)^{-8})\theta$, which is contained in $(1 + (\log R)^{-10})\tau_{N-1}$. Since each $\psi_{T_{\tau_{N-1}}}$ has essential Fourier support in $100\sqrt{\log R}\,(\log R)^{-9}\tau_{N-1}$,

$$f_{N-1,\tau_{N-1}} = \sum_{T_{\tau_{N-1}} \in \mathbb{T}_{\tau_{N-1},\lambda}} \psi_{\tau_{N-1}} f_{N,\tau_{N-1}}$$

has essential Fourier support in

$$\big(100\sqrt{\log R}\,(\log R)^{-9}+1+(\log R)^{-10}\big)\tau_{N-1}\subset(1+(\log R)^{-8})\tau_{N-1}.$$

Iterating this reasoning until $k=1$ gives (3) and (4). ∎

Definition 3.22 (Definition of g_k). For $k=1,\dots,N-1$, define

$$g_k:=\sum_{\tau_k}|f_{k+1,\tau_k}|^2*\varphi_{\tilde{T}_{\tau_k}}$$

where $\varphi_{\tilde{T}_{\tau_k}}$ is specified in Definition 3.13. For $k=N-1$, the notation $f_{N,\tau_{N-1}}$ means $f_{\tau_{N-1}}$.

The following lemma shows that the difference between the kth and $(k+1)$st versions of f_τ is controlled by $\lambda^{-1}g_k$. We eventually apply this lemma for $x\in\Omega_k$, defined in Definition 3.27, where we know that $g_k\sim r$. We will see that on this set, the differences between the different versions of f_τ are negligible.

Lemma 3.23. *Suppose τ is a $\rho^{-1/2}\times\rho^{-1}$-rectangle in the ρ^{-1}-neighborhood of $\mathbb{P}^1$, at any scale $1\le\rho\le R$. For $k=1,\dots,N-1$, if $R_k\ge\rho$,*

$$\Big|\sum_{\tau_k\subset\tau}f_{k+1,\tau_k}(x)-\sum_{\tau_k\subset\tau}f_{k,\tau_k}(x)\Big|\lesssim(\log R)^{\tilde{c}}\lambda^{-1}g_k(x)+R^{-1000}.$$

Proof. In the following proof, $\|\cdot\|_\infty$ means $\|\cdot\|_{L^\infty(\mathbb{R}^2)}$. By (3.5),

$$\sum_{\tau_k\subset\tau}f_{k+1,\tau_k}(x)-\sum_{\tau_k\subset\tau}f_{k,\tau_k}(x)=\sum_{\tau_k\subset\tau}\sum_{\substack{T\in\mathbb{T}_{\tau_k}\\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda}}\psi_T(x)f_{k+1,\tau_k}(x).$$

Recall that ψ_T has Gaussian decay off of T. It follows from (3.7) that

$$\Big|\sum_{\tau_k\subset\tau}\sum_{\substack{T\in\mathbb{T}_{\tau_k}\\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda\\ x\notin(\log R)T}}\psi_T(x)f_{k+1,\tau_k}(x)\Big|\le R^{-1000}.$$

Then we have

$$\Big|\sum_{\tau_k\subset\tau}\sum_{\substack{T\in\mathbb{T}_{\tau_k}\\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda\\ x\in(\log R)T}}\psi_T(x)f_{k+1,\tau_k}(x)\Big|\le\lambda^{-1}\sum_{\tau_k\subset\tau}\sum_{\substack{T\in\mathbb{T}_{\tau_k}\\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda\\ x\in(\log R)T}}\|\psi_T f_{k+1,\tau_k}\|_\infty^2.$$

The number of terms in the inner sum is $\lesssim(\log R)^2$. Let $T_{\tau_k}(x)$ be the unique rectangle in the tiling $\mathbb{T}_{\tau_k}$ that includes x. If $x\in(\log R)T$, then ψ_T is essentially supported in $10(\log R)T_{\tau_k}(x)$, and so we have

$$\sum_{\tau_k\subset\tau}\sum_{\substack{T\in\mathbb{T}_{\tau_k}\\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda\\ x\in(\log R)T}}\|\psi_T f_{k+1,\tau_k}\|_\infty^2\lesssim(\log R)^2\sum_{\tau_k\subset\tau}\|f_{k+1,\tau_k}\|^2_{L^\infty(10(\log R)T_{\tau_k}(x))}$$
$$+R^{-1000}.$$

Recall by Lemma 3.21 that the Fourier transform of f_{k+1,τ_k} is essentially supported in $(1+(\log R)^{-10})\tau_k$, so for ρ_{τ_k} from Definition 3.11,

$$|f_{k+1,\tau_k}(y)| \le |f_{k+1,\tau_k} * \check{\rho}_{\tau_k}(y)| + R^{-1000}.$$

Now we defined $\varphi_{\tilde{T}_{\tau_k}}(z)$ to be at least $\sup_{z+(\log R)^{11}\tau_k^*} |\check{\rho}_{\tau_k}|$, so

$$\|f_{k+1,\tau_k}\|_{L^\infty(10(\log R)T_{\tau_k}(x))} \le |f_{k+1,\tau_k}| * \varphi_{\tilde{T}_{\tau_k}}(x) + R^{-1000}. \tag{3.8}$$

Therefore,

$$\Big|\sum_{\tau_k\subset\tau} \sum_{\substack{T\in\mathbb{T}_{\tau_k} \\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda \\ x\in(\log R)T}} \psi_T(x) f_{k+1,\tau_k}(x)\Big| \lesssim (\log R)^2\lambda^{-1}\sum_{\tau_k\subset\tau}\big(|f_{k+1,\tau_k}| * \varphi_{\tilde{T}_{\tau_k}}(x) + R^{-1000}\big)^2 + R^{-1000}.$$

Since $\lambda^{-1} = \frac{\alpha}{r}(\log R)^{-m} \lesssim R^{1/2}$ (see the last line of Note 3.7),

$$\Big|\sum_{\tau_k\subset\tau} \sum_{\substack{T\in\mathbb{T}_{\tau_k} \\ \|\psi_T f_{k+1,\tau_k}\|_\infty>\lambda \\ x\in(\log R)T}} \psi_T(x) f_{k+1,\tau_k}(x)\Big| \lesssim (\log R)^2\lambda^{-1}\sum_{\tau_k\subset\tau}\big(|f_{k+1,\tau_k}| * \varphi_{\tilde{T}_{\tau_k}}(x)\big)^2 + R^{-1000}.$$

Applying Cauchy–Schwarz to the integral in the convolution shows that the above is

$$\le (\log R)^2\lambda^{-1}\sum_{\tau_k\subset\tau} |f_{k+1,\tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}}(x)\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1} + R^{-1000}.$$

Finally, note that $\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1} \le (\log R)^{\tilde{c}}$ by Lemma 3.14 and that

$$\sum_{\tau_k\subset\tau} |f_{k+1,\tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}}(x) \le g_k(x) \tag{3.9}$$

by Definition 3.22. ∎

3.4. *High/low lemmas for g_k*

Definition 3.24 (Definition of g_k^ℓ and g_k^h). For $k = 1,\dots,N-1$ and η_k from (3.1), define

$$g_k^\ell := g_k * \check{\eta}_k \quad \text{and} \quad g_k^h := g_k - g_k^\ell.$$

Lemma 3.25 (Low lemma).

$$|g_k^\ell| \le g_{k+1} + R^{-1000}.$$

Proof. Write

$$|g_k^\ell(x)| = \Big|\sum_{\tau_k} |f_{k+1,\tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}} * \check{\eta}_k(x)\Big|$$

where η_k is defined in Definition 3.9. For each τ_k we have

$$\begin{aligned} |f_{k+1,\tau_k}|^2 * \check{\eta}_k(x) &= \int \widehat{f}_{k+1,\tau_k} * \widehat{\bar{f}}_{k+1,\tau_k}(\xi) e^{2\pi i x\cdot\xi} \eta_k(\xi)\, d\xi \\ &= \sum_{\tau_{k+1},\tau'_{k+1}\subset\tau_k} \int \widehat{f}_{k+1,\tau_{k+1}} * \widehat{\bar{f}}_{k+1,\tau'_{k+1}}(\xi) e^{2\pi i x\cdot\xi} \eta_k(\xi)\, d\xi. \end{aligned}$$

Now, $\widehat{f}_{k+1,\tau_{k+1}} * \widehat{\bar{f}}_{k+1,\tau'_{k+1}}$ is essentially supported in $(1+(\log R)^{-8})(\tau_{k+1}-\tau'_{k+1})$. By the well-spaced property, we know that this set does not intersect the ball of radius $\frac{1}{2}R_{k+2}^{-1/2}$ (the support of η_k) unless $\tau_k = \tau'_k$. Therefore, up to errors of size R^{-1000}, we have

$$\begin{aligned} |f_{k+1,\tau_k}|^2 * \check{\eta}_k(x) &= \sum_{\tau_{k+1}\subset\tau_k} \int \widehat{f}_{k+1,\tau_{k+1}} * \widehat{\bar{f}}_{k+1,\tau_{k+1}}(\xi) e^{2\pi i x\cdot\xi} \eta_k(\xi)\, d\xi \\ &= \sum_{\tau_{k+1}\subset\tau_k} |f_{k+1,\tau_{k+1}}|^2 * \check{\eta}_k(x). \end{aligned}$$

By Lemma 3.21, $|f_{k+1,\tau_{k+1}}| \le |f_{k+2,\tau_{k+1}}|$, and so

$$\big||f_{k+1,\tau_k}|^2 * \check{\eta}_k(x)\big| \le \sum_{\tau_{k+1}\subset\tau_k} |f_{k+2,\tau_{k+1}}|^2 * |\check{\eta}_k|(x).$$

Plug this back into the definition of g_k^ℓ to get

$$|g_k^\ell(x)| \le \sum_{\tau_k}\big||f_{k+1,\tau_k}|^2 * \check{\eta}_k\big| * \varphi_{\tilde{T}_{\tau_k}} \le \sum_{\tau_{k+1}} |f_{k+2,\tau_{k+1}}|^2 * |\check{\eta}_k| * \varphi_{\tilde{T}_{\tau_k}}.$$

Now, $\varphi_{\tilde{T}_{\tau_{k+1}}}$ was defined in Definition 3.13 so that when $\tau_{k+1}\subset\tau_k$,

$$|\check{\eta}_k| * \varphi_{\tilde{T}_{\tau_k}} \le \varphi_{\tilde{T}_{\tau_{k+1}}}.$$

Plugging that in again, we get

$$|g_k^\ell| \le \sum_{\tau_{k+1}} |f_{k+2,\tau_{k+1}}|^2 * \varphi_{\tilde{T}_{\tau_{k+1}}} = g_{k+1} + R^{-1000}.$$

■

Lemma 3.26 (High lemma).

$$\int |g_k^h|^2 \lesssim (\log R)^{\tilde{c}} \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 + R^{-1000}.$$

Proof. By Definitions 3.22 and 3.24,

$$\int |g_k^h|^2 = \sum_{\tau_k}\sum_{\tau'_k} \int (|f_{k+1,\tau_k}|^2)^\wedge\, \widehat{\varphi}_{\tilde{T}_{\tau_k}}(1-\eta_k)(|f_{k+1,\tau'_k}|^2)^\wedge\, \widehat{\varphi}_{\tilde{T}_{\tau'_k}}(1-\eta_k).$$

The Fourier transform of $|f_{k+1,\tau_k}|^2$ is essentially supported in $2(\tau_k - \tau_k)$. Recall also from Definition 3.9 that $1-\eta_k$ is supported where $|\xi| \geq \frac{1}{4\log R} R_{k+2}^{-1/2}$. The set $2(\tau_k - \tau_k) \setminus B(\frac{1}{4\log R} R_{k+2}^{-1/2})$ overlaps at most $\sim (\log R) R_{k+2}^{1/2}/R_k^{1/2} = (\log R)^{13}$ many of the sets $2(\tau_k' - \tau_k') \setminus B(\frac{1}{4\log R} R_{k+2}^{-1/2})$. Thus applying Cauchy–Schwarz to the integral in the convolution, we get

$$\begin{aligned}\int |g_k^h|^2 &\lesssim (\log R)^{\tilde{c}'} \int \sum_{\tau_k} \big| |f_{k+1,\tau_k}|^2 * \varphi_{\tilde{T}_{\tau_k}} \big|^2 + R^{-1000} \\ &\lesssim (\log R)^{\tilde{c}'} \int \sum_{\tau_k} \|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)}^2 |f_{k+1,\tau_k}|^4 + R^{-1000} \\ &\lesssim (\log R)^{\tilde{c}''} \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 + R^{-1000}.\end{aligned}$$

where we use $\|\varphi_{\tilde{T}_{\tau_k}}\|_{L^1(\mathbb{R}^2)} \lesssim (\log R)^{\tilde{c}}$ (by Lemma 3.14). ∎

3.5. *The sets Ω_k*

In this subsection, we will decompose the starting set Q_R into $(Q_R \cap L) \cup (Q_R \cap \Omega_1) \cup \cdots \cup (Q_R \cap \Omega_{N-1})$. On the set Ω_k, the bilinearized version of f is basically the same as for the kth pruned version of f (see Lemma 3.28) and g_k is high-dominated (see Lemma 3.29).

Definition 3.27 (Definition of Ω_k). Recall the parameter $r > 0$ defined in (3.2). Let Ω_{N-1} be the union of pairwise disjoint $R_{N-1}^{1/2}$-cubes Q_{N-1} with non-empty intersection with Q_R and satisfying

$$(1+\delta)r + R^{-500} < \|g_{N-1}\|_{L^\infty((\log R)^9 Q_{N-1})}.$$

We define Ω_k for $k = N-2$, then $k = N-3$, down to $k = 1$. To define Ω_k, partition $Q_R \setminus (\Omega_{N-1} \sqcup \cdots \sqcup \Omega_{k+1})$ into $R_k^{1/2}$-cubes Q_k. Define Ω_k to be the union of Q_k in the partition which satisfy

$$(1+\delta)^{N-k} r + (N-k) R^{-500} < \|g_k\|_{L^\infty((\log R)^9 Q_k)}.$$

Also define

$$L := Q_R \setminus (\Omega_1 \sqcup \cdots \sqcup \Omega_{N-1}).$$

Recall that $\delta = \frac{1}{\log R}$.

Lemma 3.28. *Suppose τ is a $\rho^{-1/2} \times \rho^{-1}$-rectangle in the ρ^{-1}-neighborhood of $\mathbb{P}^1$, at any scale $1 \leq \rho \leq R$. For $k = 1, \ldots, N-1$, if $R_k \geq \rho$,*

$$\Big| f_\tau(x) - \sum_{\tau_k \subset \tau} f_{k+1,\tau_k}(x) \Big| \leq (\log R)^{-10} \alpha + R^{-500} \quad \forall x \in U_\alpha \cap Q_R \cap \Omega_k.$$

Recall that $f_\tau = (\widehat{f} \chi_\tau)^\vee$.

Proof. First note that for each $l \in \{1, \ldots, N-1\}$, by (3.6),

$$\sum_{\tau_l \subset \tau} f_{l,\tau_l} = \sum_{\tau_{l-1} \subset \tau} \sum_{\tau_l \subset \tau_{l-1}} f_{l,\tau_l} = \sum_{\tau_{l-1} \subset \tau} f_{l,\tau_{l-1}}.$$

Then we may decompose the difference as

$$\Big| f_\tau - \sum_{\tau_k \subset \tau} f_{k+1,\tau_k} \Big| = \Big| f_\tau - \sum_{\theta \subset \tau} f_{N,\theta} \Big| + \Big| \sum_{\tau_{N-1} \subset \tau} f_{N,\tau_{N-1}} - \sum_{\tau_{N-1} \subset \tau} f_{N-1,\tau_{N-1}} \Big| + \cdots + \Big| \sum_{\tau_{k+1} \subset \tau} f_{k+2,\tau_{k+1}} - \sum_{\tau_{k+1} \subset \tau} f_{k+1,\tau_{k+1}} \Big|.$$

By Lemma 3.23, this is bounded by

$$(\log R)^{\tilde{c}} \lambda^{-1} (g_N(x) + g_{N-1}(x) + \cdots + g_{k+1}(x)) + (N-k) R^{-1000}.$$

Finally, use the definition of Ω_k for $k \le N-1$ to find that

$$\Big| f_\tau(x) - \sum_{\tau_k \subset \tau} f_{k+1,\tau_k} \Big| \lesssim (\log R)^{\tilde{c}} (N-k)(1+\delta)^{N-k} \lambda^{-1} r + (N-k)^2 R^{-500}. \tag{3.10}$$

Recall that $N \le \log R$, $(1+\delta)^N \sim 1$, and recall that λ was defined in Notation 3.18 by

$$\lambda = (\log R)^m r \alpha^{-1}.$$

By choosing m sufficiently large, we can guarantee that the main term on the right-hand side of (3.10) is bounded by $(\log R)^{-10}\alpha$. ∎

Lemma 3.29 (g_k is high-dominated on Ω_k). *Let $k = 1, \ldots, N-1$. For each $R_k^{1/2}$-cube $Q_k \subset \Omega_k$,*

$$\|g_k\|_{L^\infty((\log R)^9 Q_k)} \le 2(\log R) \|g_k^h\|_{L^\infty((\log R)^9 Q_k)}.$$

Proof. Let $k \in \{1, \ldots, N-1\}$. By definition of Ω_k,

$$(1+\delta)^{N-k} r + (N-k) R^{-500} < \|g_k\|_{L^\infty((\log R)^9 Q_k)}. \tag{3.11}$$

Note that

$$\|g_k\|_{L^\infty((\log R)^9 Q_k)} \le \|g_k^\ell\|_{L^\infty((\log R)^9 Q_k)} + \|g_k^h\|_{L^\infty((\log R)^9 Q_k)}$$

and suppose that

$$\|g_k^h\|_{L^\infty((\log R)^9 Q_k)} \le \delta \|g_k^\ell\|_{L^\infty((\log R)^9 Q_k)}. \tag{3.12}$$

Then by Lemma 3.25,

$$\begin{aligned} \|g_k\|_{L^\infty((\log R)^9 Q_k)} &\le (1+\delta) \|g_k^\ell\|_{L^\infty((\log R)^9 Q_k)} \\ &\le (1+\delta) \|g_{k+1}\|_{L^\infty((\log R)^9 Q_k)} + (1+\delta) R^{-1000}. \end{aligned} \tag{3.13}$$

By the definition of Ω_k and of r,

$$\|g_{k+1}\|_{L^\infty((\log R)^9 Q_k)} \le (1+\delta)^{N-k-1} r + (N-k-1)R^{-500}.$$

This combined with (3.13) gives

$$\|g_k\|_{L^\infty((\log R)^9 Q_k)} \le (1+\delta)^{N-k} r + (N-k)R^{-500},$$

where we have used $(1+\delta)(N-k-1)R^{-500} + (1+\delta)R^{-1000} \le (N-k)R^{-500}$ since $N \lesssim \log R/\log\log R$. This contradicts (3.11) and means that (3.12) must be false, so the conclusion follows. ∎

3.6. Proof of Proposition 3.5

Recall from Note 3.7 that we have replaced f by a constant multiple cf so that $\max_\theta \|f_\theta\|_{L^\infty(\mathbb{R}^2)} = 1$.

The first step of the proof of Proposition 3.5 involves an application of a local bilinear restriction theorem. We will use the following version.

Theorem 3.30 (Local bilinear restriction). *Let τ_k and τ_k' be non-adjacent $\sim R_k^{-1/2} \times R_k^{-1}$-rectangles in the R_k^{-1}-neighborhood of $\mathbb{P}^1$. Suppose $j \ge k$ and $f \in \mathcal{S}$ has Fourier support in $\mathcal{N}_{R_j^{-1}}(\mathbb{P}^1)$. Suppose T is in the range $R_j \ge T > 10(\log R)R_j^{1/2}/\mathrm{dist}(\tau_k, \tau_k')$, and that Q_T is a cube of sidelength T. Then*

$$\int_{Q_T} |f_{\tau_k} f_{\tau_k'}|^2 \lesssim \frac{(\log R)^4 T^{-2}}{\mathrm{dist}(\tau_k, \tau_k')} \int \sum_{\tau_j \subset \tau_k} |f_{\tau_j}|^2 \omega_{Q_T} \cdot \int \sum_{\tau_j' \subset \tau_k'} |f_{\tau_j}|^2 \omega_{Q_T} + R^{-1000}$$

for a Gaussian weight function ω_{Q_T} localized to $(\log R)Q_T$ and with Fourier transform essentially supported in the ball of radius $2(\log R)T^{-1}$ centered at the origin.

Proof. Let ϕ_{Q_T} be a Gaussian bump function adapted to Q_T as in Definition 3.15, so the Fourier transform $\widehat{\phi}_{Q_T}$ is essentially supported in a ball of radius $2(\log R)T^{-1}$. Then

$$\begin{aligned}\int_{Q_T} |f_{\tau_k} f_{\tau_k'}|^2 &\lesssim \int \Big| \sum_{\tau_j \subset \tau_k,\, \tau_j' \subset \tau_k'} f_{\tau_j} f_{\tau_j'} \Big|^2 \phi_{Q_T} \\ &\lesssim \int \sum_{\tau_j \subset \tau_k,\, \tau_j' \subset \tau_k'} |f_{\tau_j} f_{\tau_j'}|^2 \phi_{Q_T} + R^{-1000}.\end{aligned}$$

The reason for the last inequality is that for a fixed pair (τ_j, τ_j'), the number of pairs (τ_j'', τ_j''') such that

$$(\tau_j + \tau_j' + B_{2(\log R)T^{-1}}) \cap (\tau_j'' + \tau_j''' + B_{2(\log R)T^{-1}}) \ne \emptyset \tag{3.14}$$

is at most $O(1)$. Here we use the fact that $T \ge 10(\log R)R_j^{1/2}/\mathrm{dist}(\tau_k, \tau_k')$. For more details of checking (3.14), see the appendix.

It suffices to show that

$$\int |f_{\tau_j} f_{\tau_j'}|^2 \phi_{Q_T} \lesssim \frac{(\log R)^4 T^{-2}}{\operatorname{dist}(\tau_k, \tau_k')} \int |f_{\tau_j}|^2 \omega_{Q_T} \int |f_{\tau_j}|^2 \omega_{Q_T}.$$

For a translate T_{τ_j} of τ_j^* and a translate $T_{\tau_j'}$ of $\tau_j'^*$, let $c_{T_{\tau_j}} = \max_{x\in T_{\tau_j}} |f_{\tau_j}(x)|^2$ and $c_{T_{\tau_j'}} = \max_{x\in T_{\tau_j'}} |f_{\tau_j'}|^2(x)$. We consider only those T_{τ_j} and $T_{\tau_j'}$ intersecting $(\log R)Q_T$.

Since T_{τ_j} and $T_{\tau_j'}$ have angle $\sim \operatorname{dist}(\tau_k, \tau_k')$, we have

$$|T_{\tau_j} \cap T_{\tau_j'} \cap (\log R)Q_T| \le R_j/\operatorname{dist}(\tau_k, \tau_k'),$$

while $|T_{\tau_j} \cap (\log R)Q_T| \gtrsim R_j^{1/2} T$. Consequently,

$$\begin{aligned}
\int |f_{\tau_j} f_{\tau_j'}|^2 \phi_{Q_T} &\le \int_{(\log R)Q_T} |f_{\tau_j} f_{\tau_j'}|^2 + R^{-1000} \\
&\le \int_{(\log R)Q_T} \sum_{T_{\tau_j}} \sum_{T_{\tau_j'}} c_{T_{\tau_j}} \chi_{T_{\tau_j}} \cdot c_{T_{\tau_j'}} \chi_{T_{\tau_j'}} + R^{-1000} \\
&\lesssim \frac{T^{-2}}{\operatorname{dist}(\tau_k, \tau_k')} \int_{(\log R)Q_T} \sum_{T_{\tau_j}} c_{T_{\tau_j}} \chi_{T_{\tau_j}} \cdot \int_{(\log R)Q_T} \sum_{T_{\tau_j'}} c_{T_{\tau_j'}} \chi_{T_{\tau_j'}} + R^{-1000}.
\end{aligned}$$

The next step is to show that

$$\int_{(\log R)Q_T} \sum_{T_{\tau_j}} c_{T_{\tau_j}} \chi_{T_{\tau_j}} \lesssim (\log R)^2 \int |f_{\tau_j}|^2 w_{Q_T} + R^{-1000}. \tag{3.15}$$

The Fourier transform of $|f_{\tau_j}|^2$ is supported on $\tilde\tau_{j,0} = \tau_j + (-\tau_j)$, which is approximately an $R_j^{-1/2} \times R_j^{-1}$-rectangle. Let $\sum_{\tilde\tau_j} \phi_{\tilde\tau_j}$ be a Gaussian partition of unity for $\{\tilde\tau_j\}_{\tilde\tau_j \| \tilde\tau_{j,0}}$. Let $\phi_{\tau_j} = \sum_{\tilde\tau_j \subset (\log R)\tilde\tau_{j,0}} \phi_{\tilde\tau_j}$. Then $|\phi_{\tau_j} - \chi_{\tilde\tau_{j,0}}|(\xi) \le R^{-1000}$ for $\xi \in \tilde\tau_{j,0}$ and $|f_{\tau_j}|^2 = |f_{\tau_j}|^2 * \widehat{\phi}_{\tau_j} + O(R^{-1000})$.

Let $\psi_{\tau_j}(x) = \max_{y\in x+T_{\tau_j}-T_{\tau_j}} |\widehat{\phi}_{\tau_j}|$. Then

$$\sum_{T_{\tau_j}} c_{T_{\tau_j}} \chi_{T_{\tau_j}} \le |f_{\tau_j}|^2 * \psi_{\tau_j}.$$

We finish the proof since $\psi_{\tau_j} * \chi_{(\log R)Q_T} \lesssim (\log R)^2 w_{Q_T} + R^{-1000}$ for a Gaussian bump function w_{Q_T} localized at $(\log R)Q_T$. ■

Proof of Proposition 3.5. It suffices to bound $|U_\alpha \cap \Omega_k|$ for $k = 1, \dots, N-1$ and $|U_\alpha \cap L|$ since there are $\lesssim \log R$ of these sets and $|U_\alpha \cap Q_R| \le \sum_{k=1}^{N-1} |U_\alpha \cap \Omega_k| + |U_\alpha \cap L|$.

For $k = 1, \dots, N-1$, by Lemma 3.28, if $x \in U_\alpha \cap \Omega_k$, then

$$\max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)| \le \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x) f_{\tau'}(x)| + (\log R)^{-10}\alpha |f_{\tau'}(x)| + R^{-500},$$

where $f_{k+1,\tau} = \sum_{\tau_k \subset \tau} f_{k+1,\tau_k}$. By the definition of U_α (Definition 3.1), $|f_{\tau'}(x)| \le (\log R)^9 \alpha$, and so

$$\begin{aligned}\max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)| &\le \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x) f_{\tau'}(x)| + (\log R)^{-1}\alpha^2 + R^{-500} \\ &\le \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x) f_{k+1,\tau'}(x)| + (\log R)^{-1}\alpha^2 \\ &\quad + (\log R)^{-10}\alpha |f_{k+1,\tau}(x)| + 2R^{-500}.\end{aligned}$$

Using the definition of U_α as above as well as Lemma 3.28, it follows that $|f_{k+1,\tau}(x)| \le 2(\log R)^9\alpha$, and so altogether

$$\max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)| \le \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x) f_{k+1,\tau'}(x)| + 3(\log R)^{-1}\alpha^2 + 2R^{-500}.$$

The R^{-500} error term is negligible given our normalization, as explained in Note 3.7. As $x \in U_\alpha$, $\max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)| \sim \alpha^2$, and so $\max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau}(x) f_{k+1,\tau'}(x)| \sim \alpha^2$ as well. Therefore

$$\alpha^4 |U_\alpha \cap \Omega_k| \lesssim \Big\| \max_{\tau,\tau' \text{ non-adj}} |f_{k+1,\tau} f_{k+1,\tau'}|^{1/2} \Big\|^4_{L^4(U_\alpha \cap \Omega_k)} \tag{3.16}$$

where τ and τ' are non-adjacent $\sim (\log R)^{-6} \times (\log R)^{-12}$-rectangles. We wish to apply the bilinear restriction theorem above, but the functions f_{k+1,τ_k} are only essentially supported in $\sim \tau_k$. This just means that we have an error term of R^{-1000} which is negligible given our normalization, as explained in Note 3.7.

For each $Q_k \subset \Omega_k$,

$$\begin{aligned}\Big\| \max_{\tau,\tau'} |f_{k+1,\tau} f_{k+1,\tau'}|^{1/2} \Big\|^4_{L^4(Q_k)} &\le \sum_{\substack{\tau,\tau' \\ \text{non-adj}}} \int_{Q_k} |f_{k+1,\tau} f_{k+1,\tau'}|^2 \\ \text{(Theorem 3.30)} \qquad &\lesssim |Q_k|^{-1} \Big(\int \sum_{\tau_k} |f_{k+1,\tau_k}|^2 \omega_{\tilde{Q}_k} \Big)^2 + R^{-1000} \\ &\lesssim |Q_k|^{-1} \Big(\int g_k \omega_{\tilde{Q}_k} \Big)^2 + R^{-1000}|Q_k| + R^{-1000}\end{aligned}$$

where $w_{\tilde{Q}_k}$ is a weight function localized to $(\log R)^8 Q_k$ and the final inequality follows from (3.8) and (3.9) in the proof of Lemma 3.23.

By Lemma 3.29 and the decay properties of $\omega_{\tilde{Q}_k}$,

$$\begin{aligned}\int g_k \omega_{\tilde{Q}_k} &\lesssim (\log R)^{18} \|g_k\|_{L^\infty((\log R)^9 Q_k)} |Q_k| + R^{-500} \\ &\lesssim (\log R)^{19} \|g_k^h\|_{L^\infty((\log R)^9 Q_k)} |Q_k| + R^{-500} \\ &\lesssim (\log R)^{20} \|g_k^h\|_{L^1(W_{Q_k})} + R^{-500} \\ &\lesssim (\log R)^{30} \|g_k^h\|_{L^2(W_{Q_k})} |Q_k|^{1/2} + R^{-500}\end{aligned}$$

where W_{Q_k} is a Gaussian weight function localized to $\sim (\log R)^9 Q_k$ coming from the locally constant property (see (3.15)). Use this in the previous displayed math and add up the contributions from each Q_k to obtain

$$\Big\| \max_{\tau,\tau'} |f_{k,\tau} f_{k,\tau'}|^{1/2} \Big\|^4_{L^4(\Omega_k)} \lesssim (\log R)^{60} \int |g_k^h|^2 \omega_{\Omega_k} \tag{3.17}$$

where $\omega_{\Omega_k} = \sum_{Q_k} W_{Q_k}$. Note that $\omega_{\Omega_k} \lesssim 1$ and by the high lemma (Lemma 3.26),

$$\int |g_k^h|^2 \omega_{\Omega_k} \lesssim (\log R)^{\tilde{c}''} \int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 + R^{-1000}.$$

Then $f_{k+1,\tau_k} = \sum_{\tau_{k+1} \subset \tau_k} f_{k+1,\tau_{k+1}}$, so

$$\int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 \le \frac{R_{k+1}^{3/2}}{R_k^{3/2}} \int \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^4.$$

By Lemma 3.21, we have $\| f_{k+1,\tau_{k+1}} \|_{L^\infty} \le C(\log R)^2 \lambda + R^{-1000}$. The R^{-1000} error term is negligible as explained in Note 3.7, and so we essentially have

$$\int \sum_{\tau_k} |f_{k+1,\tau_k}|^4 \le (\log R)^4 \lambda^2 \int \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^2.$$

Finally, we have to carefully unwind the definition of f_{k,τ_k} and $f_{k,\tau_{k-1}}$ to relate this last quantity to the original f_θ:

$$\int \sum_{\tau_{k+1}} |f_{k+1,\tau_{k+1}}|^2 \le \sum_\theta \int |f_\theta|^2 + R^{-500}. \tag{3.18}$$

First we recall by Lemma 3.21 that $|f_{k+1,\tau_{k+1}}(x)| \le |f_{k+2,\tau_{k+1}}(x)|$, and so

$$\sum_{\tau_{k+1}} \int |f_{k+1,\tau_{k+1}}|^2 \le \sum_{\tau_{k+1}} \int |f_{k+2,\tau_{k+1}}|^2 \tag{3.19}$$

Next, by Definition 3.20, $f_{k+2,\tau_{k+1}} = \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}}$, and so

$$\sum_{\tau_{k+1}} \int |f_{k+1,\tau_{k+1}}|^2 \le \sum_{\tau_{k+1}} \int \Big| \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}} \Big|^2. \tag{3.20}$$

By Lemma 3.21, the Fourier transform of $f_{k+2,\tau_{k+2}}$ is essentially supported in the set $(1 + (\log R)^{-8})\tau_{k+2}$. Since distinct τ_{k+2} and τ'_{k+2} are $\ge \frac{1}{2} R_{k+3}^{-1/2}$-separated, these sets are disjoint. By orthogonality, we get

$$\sum_{\tau_{k+1}} \int \Big| \sum_{\tau_{k+2} \subset \tau_{k+1}} f_{k+2,\tau_{k+2}} \Big|^2 \le \sum_{\tau_{k+2}} \int |f_{k+2,\tau_{k+2}}|^2 + R^{-500}. \tag{3.21}$$

Now we repeat the reasoning in inequalities (3.19)–(3.21) at many scales to conclude that

$$\sum_{\tau_{k+1}}\int\Big|\sum_{\tau_{k+2}\subset\tau_{k+1}} f_{k+2,\tau_{k+2}}\Big|^2 \le \sum_{\tau_{k+2}}\int|f_{k+2,\tau_{k+2}}|^2$$
$$\le \sum_{\tau_{k+2}}\int|f_{k+3,\tau_{k+2}}|^2$$
$$\le \sum_{\tau_{k+3}}\int|f_{k+3,\tau_{k+3}}|^2$$
$$\dots$$
$$\le \sum_{\theta}\int|f_{N,\theta}|^2 \le \sum_{\theta}\int|f_\theta|^2.$$

In the above sequence of inequalities, we neglected to include an R^{-500} added error term in each step due to the difference between “essential support” and “actual support.” These error terms are all negligible according to Note 3.7.

The conclusion of this argument is that for $k = 1,\dots,N-1$,

$$\alpha^4|U_\alpha\cap\Omega_k| \lesssim (\log R)^{\tilde{c}'}\frac{r^2}{\alpha^2}\sum_\theta\|f_\theta\|^2_{L^2(\mathbb{R}^2)}.$$

Finally, we check that this indeed gives the conclusion of Proposition 3.5. Recall that $g_N(x) = \sum_\theta|f_\theta|^2 * \varphi_{\tilde{T}_\theta}$. By Lemma 3.14, $\|\varphi_{\tilde{T}_\theta}\|_{L^1} \le (\log R)^c$. Thus for each θ and $x \in (\log R)^2 Q_N$ (where $Q_N \cap \Omega_N \ne \emptyset$),

$$|f_\theta|^2 * \varphi_{\tilde{T}_\theta}(x) \lesssim (\log R)^{\tilde{c}}\|f_\theta\|^2_{L^\infty(\mathbb{R}^2)} + R^{-1000}.$$

It follows that $r \lesssim (\log R)^{\tilde{c}}\sum_\theta\|f_\theta\|^2_{L^\infty(\mathbb{R}^2)} + R^{-1000}$. Plugging this in gives the conclusion of Proposition 3.5.

Finally, it remains to bound $|U_\alpha\cap L|$. The first step is going from f to f_1 using Lemma 3.28 (the argument for Ω_1 in (3.16) holds for L as well):

$$\alpha^6|U_\alpha\cap L| \lesssim \int_{U_\alpha\cap L}\max_{\tau,\tau'}|f_{1,\tau}f_{1,\tau'}|^3 + R^{-1000}$$
$$\lesssim (\log R)^c\int_{U_\alpha\cap L}\Big(\sum_{\tau_1}|f_{1,\tau_1}|^2\Big)^3 + R^{-1000}$$
$$\text{(Lemma 3.21)}\quad \lesssim (\log R)^c\int_{U_\alpha\cap L}\Big(\sum_{\tau_1}|f_{2,\tau_1}|^2\Big)^3 + R^{-1000}$$
$$\le (\log R)^c\int_{U_\alpha\cap L}g_1^2\Big(\sum_{\tau_1}|f_{2,\tau_1}|^2\Big) + R^{-1000}$$

where the last inequality is due to (3.8). Then by the definition of L,

$$\|g_1\|_{L^\infty(U_\alpha\cap L)} \lesssim Cr + NR^{-500}.$$

Finally, by (3.18),

$$\int\sum_{\tau_1}|f_{2,\tau_1}|^2 \le \int\sum_\theta|f_\theta|^2 + R^{-500}. \qquad \blacksquare$$

4. Proof of Theorem 1.2: the general case

In the last section, we proved Proposition 3.5, which establishes our main theorem in the broad, well-spaced case. In this section, we prove Theorem 1.2 in full generality. We use Proposition 3.5 as a black box, and then we remove the broad hypothesis by using a broad/narrow analysis, and we remove the well-spaced hypothesis by a random sampling argument.

4.1. Removing the broad hypothesis

The following proposition uses a broad/narrow analysis to prove an upper bound for $\|f\|_{L^6(Q_R)}$ using Proposition 3.5.

Proposition 4.1. *There exist $c, C \in (0,\infty)$ such that for all well-spaced collections Θ and $f \in \mathcal{S}$ with Fourier support in $\bigcup_{\theta\subset\Theta}\theta$,*

$$\|f\|_{L^6(Q_R)}^6 \le C(\log R)^c\Big(\sum_\theta \|f_\theta\|_{L^\infty(\mathbb{R}^2)}^2\Big)^2\sum_\theta\|f_\theta\|_{L^2(\mathbb{R}^2)}^2.$$

First we prove a few technical lemmas.

Lemma 4.2 (Narrow lemma). *Suppose that $\tilde\tau_k$ is an arc of length $R_k^{-1/2}\le \ell(\tilde\tau_k)\le 3R_k^{-1/2}$. Let $\{\tau_{k+1}\}$ be a partition of $\tilde\tau_k$ into $R_{k+1}^{-1/2}$-arcs. If x satisfies*

$$|f_{\tilde\tau_k}(x)| > \frac{(\log R)^2 R_{k+1}^{1/2}}{R_k^{1/2}}\max_{\substack{\tau_{k+1},\tau'_{k+1}\\ \text{non-adj}}}|f_{\tau_{k+1}}(x)f_{\tau'_{k+1}}(x)|^{1/2}, \tag{4.1}$$

then there exists an arc $\tilde\tau_{k+1}$ such that $\ell(\tilde\tau_{k+1}) = 3R_{k+1}^{-1/2}$ and

$$|f_{\tilde\tau_k}(x)| \le \Big(1+\frac{1}{\log R}\Big)|f_{\tilde\tau_{k+1}}(x)|.$$

Proof. Write $f_{\tilde\tau_k} = \sum_{\tau_{k+1}} f_{\tau_{k+1}}$ and let τ^*_{k+1} index a summand satisfying

$$\max_{\tau_{k+1}\subseteq\tilde\tau_k}|f_{\tau_{k+1}}(x)| = |f_{\tau^*_{k+1}}(x)|.$$

For each τ_{k+1} that is non-adjacent to τ^*_{k+1},

$$|f_{\tau_{k+1}}(x)| \le |f_{\tau_{k+1}}(x)f_{\tau^*_{k+1}}(x)|^{1/2} < \frac{R_k^{1/2}}{(\log R)^2 R_{k+1}^{1/2}}|f_{\tilde\tau_k}(x)|$$

using the hypothesis (4.1) about x. Then

$$\Big|f_{\tilde\tau_k}(x) - \sum_{\substack{\tau_{k+1}\text{ non-adj}\\ \text{to } \tau^*_{k+1}}} f_{\tau_{k+1}}(x)\Big| > \Big(1 - \#\tau_{k+1}\frac{R_k^{1/2}}{(\log R)^2 R_{k+1}^{1/2}}\Big)|f_{\tilde\tau_k}(x)|.$$

The number of τ_{k+1} is bounded by $3R_{k+1}^{1/2}/R_k^{1/2}$. Define $\tilde{\tau}_{k+1}$ to be $(\tau_{k+1}^*)_L \cup \tau_{k+1}^* \cup (\tau_{k+1}^*)_R$ where $(\tau_{k+1}^*)_L$ is the left neighbor of τ_{k+1}^* and $(\tau_{k+1}^*)_R$ is the right neighbor of τ_{k+1}^*. ■

Lemma 4.3 (Case 2 in the proof of Proposition 4.1). *Suppose τ_k^* is an $\sim R_k^{-1/2} \times R_k^{-1}$-rectangle in the R_k^{-1}-neighborhood of $\mathbb{P}^1$. Then*

$$\int_{H_{\tau_k^*}} |f_{\tau_k^*}|^6 \lesssim (\log R)^c \Big(\sum_{\theta \subset \tau_k^*} \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)} \Big)^2 \sum_{\theta \subset \tau_k^*} \|f_\theta\|^2_{L^2(\mathbb{R}^2)}$$

where

$$H_{\tau_k^*} = \Big\{ x \in Q_R : |f_{\tau_k^*}(x)| \le (\log R)^8 \max_{\substack{\tau_{k+1}, \tau'_{k+1} \\ \text{non-adj}}} |f_{\tau_{k+1}}(x) f_{\tau'_{k+1}}(x)|^{1/2},$$

$$(\log R)^{-9} \Big(\sum_{\tau_{k+1} \subset \tau_k^*} |f_{\tau_{k+1}}(x)|^6 \Big)^{1/6} \le \max_{\substack{\tau_{k+1}, \tau'_{k+1} \subset \tau_k^* \\ \text{non-adj}}} |f_{\tau_{k+1}}(x) f_{\tau'_{k+1}}(x)|^{1/2} \Big\}.$$

Here τ_{k+1} are $R_{k+1}^{-1/2} \times R_{k+1}^{-1}$-rectangles.

Proof. Let (c, c^2) be the center of $\tau_k^* \cap \mathbb{P}^1$. Define the affine map

$$\ell(\xi_1, \xi_2) = \big(R_k^{1/2}(\xi_1 - c), R_k(\xi_2 - 2\xi_1 c + c^2) \big). \tag{4.2}$$

Then $\ell(\tau_k^*)$ is contained in the $(R/R_k)^{-1}$-neighborhood of $\mathbb{P}^1$. The images $\{\ell(\theta)\}_{\theta \subset \tau_k^*}$ have the spacing property at scales $R_{k+1}/R_k, \dots, R/R_k$. Define the function h as

$$\widehat{h} = \widehat{f} \circ \ell^{-1}$$

and note that for each $R_l^{-1/2} \times R_l^{-1}$-rectangle $\tau_l \subset \tau_k^*$,

$$R_k^{-3/2} h_{\ell(\tau_l)}(\gamma(x)) e^{2\pi i x \cdot (c, c^2)} = f_{\tau_l}(x)$$

where $\ell(\tau_l)$ is approximately an $(R_l/R_k)^{-1/2} \times (R_l/R_k)^{-1}$-rectangle and

$$\gamma(x) = \Big(\frac{x_1 + 2c x_2}{R_k^{1/2}}, \frac{x_2}{R_k} \Big).$$

In particular,

$$\int_{H_{\tau_k^*}} |f_{\tau_k^*}|^6 = R_k^{-9+3/2} \int_{\gamma(H_{\tau_k^*})} |h_{\ell(\tau_k^*)}(x)|^6. \tag{4.3}$$

By dyadic pigeonholing, there exists $\alpha_{\tau_k^*} > 0$ such that

$$\|h_{\ell(\tau_k^*)}\|^6_{L^6(\gamma(H_{\tau_k^*}))} \lessapprox \alpha_{\tau_k^*}^6 \Big| \Big\{ x \in \gamma(H_{\tau_k^*}) : \max_{\substack{\tau_{k+1}, \tau'_{k+1} \subset \tau_k^* \\ \text{non-adj}}} |h_{\ell(\tau_{k+1})}(x) h_{\ell(\tau'_{k+1})}(x)|^{1/2} \sim \alpha_{\tau_k^*} \Big\} \Big|.$$

Repeat the proof of Proposition 3.5 to obtain

$$\|h_{\ell(\tau_k^*)}\|^6_{L^6(\gamma(H_{\tau_k^*}))} \lesssim (\log R)^c \Big(\sum_{\theta\subset\tau_k^*} \|h_{\ell(\theta)}\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \sum_{\theta\subset\tau_k^*} \|h_{\ell(\theta)}\|^2_{L^2(\mathbb{R}^2)}. \tag{4.4}$$

First observe that

$$\sum_{\theta\subset\tau_k^*} \|h_{\ell(\theta)}\|^2_{L^2(\mathbb{R}^2)} \lesssim R_k^{3-3/2} \sum_{\theta\subset\tau_k^*} \|f_\theta\|^2_{L^2(\mathbb{R}^2)}.$$

Next, note that for each $\theta \subset \tau_k$,

$$\|h_{\ell(\theta)}\|^2_{L^\infty(\mathbb{R}^2)} \le R_k^3 \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}.$$

These observations combined with (4.3) and (4.4) give the desired conclusion. ∎

Proof of Proposition 4.1. Define an iteration using a broad/narrow argument.

Initial step: Define

$$S_1 := \Big\{x \in X : |f(x)| \le (\log R)^8 \max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2},$$
$$\Big(\sum_\tau |f_\tau(x)|^6\Big)^{1/6} \le (\log R)^9 \max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2}\Big\}. \tag{4.5}$$

Define $B_1 = X \setminus S_1$. Split the integral into

$$\int_X |f|^6 = \int_{S_1} |f|^6 + \int_{B_1} |f|^6. \tag{4.6}$$

By the narrow lemma, if $x \in B_1$ satisfies $|f(x)| > (\log R)^8 \max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2}$, then for a collection $\{\tau^{**}\}$ of pairwise disjoint unions of three consecutive τ,

$$|f(x)| \le \Big(1 + \frac{1}{\log R}\Big)\Big(\sum_{\tau^{**}} |f_{\tau^{**}}(x)|^6\Big)^{1/6}.$$

Alternatively, $x \in B_1$ satisfies

$$|f(x)| \le (\log R)^8 \max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2}$$

but

$$(\log R)^{-9}\Big(\sum_\tau |f_\tau(x)|^6\Big)^{1/6} > \max_{\tau,\tau' \text{ non-adj}} |f_\tau(x) f_{\tau'}(x)|^{1/2}.$$

Putting this together means that

$$\int_{B_1} |f|^6 \le \Big(1 + \frac{1}{\log R}\Big)^6 \int_{B_1} \sum_{\tau^{**}} |f_{\tau^{**}}|^6 + \frac{1}{(\log R)^6} \int_{B_1} \sum_\tau |f_\tau|^6.$$

Let $\{\tau^*\}$ denote the collection $\{\tau\}$ if

$$\int_{B_1} \sum_{\tau} |f_\tau|^6 \geq \int_{B_1} \sum_{\tau^{**}} |f_{\tau^{**}}|^6$$

and let it equal $\{\tau^{**}\}$ otherwise. Then

$$\int_{B_1} |f_\tau|^6 \leq \left(1 + \frac{2}{\log R}\right)^6 \int_{B_1} \sum_{\tau^*} |f_{\tau^*}|^6$$

(this just means we only have one finer scale to keep track of rather than two almost equivalent scales). Summarizing all of the inequalities, we conclude that

$$\|f\|_{L^6(X)} \leq \int_{S_1} |f|^6 + \left(1 + \frac{2}{\log R}\right)^6 \sum_{\tau^*} \int_{B_1} |f_{\tau^*}|^6.$$

For each τ^*, further decompose B_1 into

$$S_{\tau^*} = \Big\{x \in B_1 : |f_{\tau^*}(x)| \leq (\log R)^8 \max_{\tau_2, \tau_2' \subset \tau^* \text{ non-adj}} |f_{\tau_2}(x) f_{\tau_2'}(x)|^{1/2},$$
$$\Big(\sum_{\tau_2 \subset \tau^*} |f_{\tau_2}(x)|^6\Big)^{1/6} \leq (\log R)^9 \max_{\tau_2, \tau_2' \subset \tau^* \text{ non-adj}} |f_{\tau_2}(x) f_{\tau_2'}(x)|^{1/2}\Big\}$$

where $\ell(\tau_2) = R_2^{-1/2}$. By analogous reasoning, we conclude this case with the inequality

$$\|f\|^6_{L^6(X)} \leq \int_{S_1} |f|^6 + \left(1 + \frac{2}{\log R}\right)^6 \sum_{\tau^*} \int_{S_{\tau^*}} |f_{\tau^*}|^6$$
$$+ \left(1 + \frac{2}{\log R}\right)^{12} \sum_{\tau^*} \int_{B_{\tau^*}} \sum_{\tau_2^* \subset \tau} |f_{\tau_2}|^6$$

where $B_{\tau^*} = B_1 \setminus S_{\tau^*}$.

Step *k* ($k \geq 2$). The conclusion of the previous step is

$$\|f\|^6_{L^6(X)} \leq \int_{S_1} |f|^6 + \left(1 + \frac{2}{\log R}\right)^6 \sum_{\tau^*} \int_{S_{\tau^*}} |f_{\tau^*}|^6 + \cdots$$
$$+ \left(1 + \frac{2}{\log R}\right)^{6(k-1)} \sum_{\tau^*_{k-1}} \int_{S_{\tau^*_{k-1}}} |f_{\tau^*_{k-1}}|^6$$
$$+ \left(1 + \frac{2}{\log R}\right)^{6k} \sum_{\tau^*_{k-1}} \int_{B_{\tau^*_{k-1}}} \sum_{\tau^*_k \subset \tau^*_{k-1}} |f_{\tau^*_k}|^6 \tag{4.7}$$

where for each τ^*_{k-1}, if $\tau^*_{k-1} \subset \tau^*_{k-2} \subset \cdots \subset \tau^*_2 \subset \tau$ then

$$B_{\tau^*_{k-1}} = B_1 \setminus (S_1 \cup S_{\tau^*} \cup S_{\tau^*_2} \cup \cdots \cup S_{\tau^*_{k-1}}).$$

For each $\tau_k^* \subset \tau_{k-1}^*$, define $S_{\tau_k^*}$ to be the set

$$\Big\{x \in B_{\tau_{k-1}^*} : |f_{\tau_k^*}(x)| \le (\log R)^8 \max_{\substack{\tau_{k+1},\tau'_{k+1}\subset\tau_k^* \\ \text{non-adj}}} |f_{\tau_{k+1}}(x) f_{\tau'_{k+1}}(x)|^{1/2},$$
$$\Big(\sum_{\tau_{k+1}\subset\tau_k^*} |f_{\tau_{k+1}}(x)|^6\Big)^{1/6} \le (\log R)^9 \max_{\substack{\tau_{k+1},\tau'_{k+1}\subset\tau_k^* \\ \text{non-adj}}} |f_{\tau_{k+1}}(x) f_{\tau'_{k+1}}(x)|^{1/2}\Big\}$$

where $\ell(\tau_{k+1}) = R_{k+1}^{-1/2}$. Define $B_{\tau_k^*} = B_{\tau_{k-1}^*} \setminus S_{\tau_k^*}$. By analogous arguments, we conclude that (4.7) holds with k replaced by $k+1$.

Iterate this procedure until Step $N-1$ where $N \sim \frac{\log R}{\log\log R}$ to obtain inequality (4.7) for $k = N$. Since there are $\sim \log R/\log\log R$ terms in the right-hand side, it suffices to consider cases where $\|f\|^6_{L^6(Q_R)}$ is bounded by $\log R$ times one of the terms on the right-hand side. Note that the factors $(1+\frac{2}{\log R})^{6k}$ are $\lesssim 1$ for all $k \le N$.

Case 1:

$$\|f\|^6_{L^6(Q_R)} \lesssim (\log R)\int_{S_1} |f|^6.$$

By the definition of S_1 in (4.5),

$$\int_{S_1} |f|^6 \le (\log R)^{48} \int_{S_1} \max_{\tau,\tau'} |f_\tau f_{\tau'}|^3.$$

Let $U_s = \{x \in S_1 : \max_{\tau,\tau'} |f_\tau(x) f_{\tau'}(x)|^{1/2} \le R^{-10}\max_\theta \|f_\theta\|_{L^\infty}\}$. Note that

$$\begin{aligned}\int_{U_s} \max_{\tau,\tau'} |f_\tau f_{\tau'}|^3 &\le R^{-55} \max_\theta \|f_\theta\|^6_{L^\infty(\mathbb{R}^2)} \\ \text{(Lemma 3.6)} \quad &\le R^{-55}\Big(\sum_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \int \sum_\theta |f_\theta|^2,\end{aligned}$$

which is the right-hand side in Proposition 4.1.

Then $S_1 \setminus U_s$ can be partitioned into $\lesssim \log R$ sets U_α on which $\max_{\tau,\tau'} |f_\tau f_{\tau'}|^{1/2} \sim \alpha$ with $R^{-10} \le \alpha/\max_\theta \|f_\theta\|_{L^\infty(\mathbb{R}^2)} \le R$. By pigeonholing,

$$\int_{S_1\setminus U_s} \max_{\tau,\tau'} |f_\tau f_{\tau'}|^3 \lesssim (\log R)\int_{S_1\cap U_\alpha} \max_{\tau,\tau'} |f_\tau f_{\tau'}|^3 \sim (\log R)\alpha^6 |S_1 \cap U_\alpha|.$$

Then Proposition 3.5 applies to bound $\alpha^6|S_1\cap U_\alpha|$.

Case 2:

$$\|f\|^6_{L^6(Q_R)} \lesssim (\log R)\sum_{\tau_k^*}\int_{S_{\tau_k^*}} |f_{\tau_k^*}|^6.$$

The sets $S_{\tau_k^*}$ are contained in $H_{\tau_k^*}$ from Lemma 4.3. Using Lemma 4.3, we get

$$\begin{aligned}\sum_{\tau_k^*}\int_{S_{\tau_k^*}} |f_{\tau_k^*}|^6 &\lesssim \sum_{\tau_k^*} (\log R)^c \Big(\sum_{\theta\subset\tau_k^*} \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \sum_{\theta\subset\tau_k^*} \|f_\theta\|^2_{L^2(\mathbb{R}^2)} \\ &\le (\log R)^c \Big(\sum_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \sum_\theta \|f_\theta\|^2_{L^2(\mathbb{R}^2)},\end{aligned}$$

so we have the desired conclusion.

Case 3:

$$\|f\|^6_{L^6(Q_R)} \lesssim (\log R) \sum_{\tau^*_{N-1}} \int_{B_{\tau^*_{N-1}}} \sum_{\tau^*_N \subset \tau^*_{N-1}} |f_{\tau^*_N}|^6.$$

The size of τ^*_N is approximately the size of θ (up to a factor of 3), so we may assume $\tau^*_N = \theta$. Then since the ℓ^6 norm is bounded by the ℓ^2 norm, the above inequality implies that

$$\begin{aligned}\alpha^6|U_\alpha \cap Q_R| &\lesssim (\log R) \int_{U_\alpha \cap X} \Big(\sum_\theta |f_\theta|^2\Big)^3 \\ &\lesssim (\log R) \Big(\sum_\theta \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2 \sum_\theta \|f_\theta\|^2_{L^2(\mathbb{R}^2)}. \qquad \blacksquare\end{aligned}$$

4.2. *Removing the well-spaced hypothesis*

This section has been simplified following the suggestions of a helpful reviewer. We will use the following lemmas to prove Theorem 1.2 in the following section.

Lemma 4.4. *For each $f \in \mathcal{S}$ with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$, there exists a collection Θ that is well-spaced at scale R_{N-1} such that*

$$\Big(1 - \frac{1}{\log R}\Big)\|f\|_{L^6(Q_R)} \le \Big\|\sum_{\theta\in\Theta} f_\theta\Big\|_{L^6(Q_R)}.$$

Proof. Partition $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ into a collection of $\sim R^{-1/2} \times R^{-1}$-rectangles $\{\theta\}$. Label the $R^{-1/2}$-arcs from left to right by $\theta_1, \dots, \theta_n$, so $n \sim R^{1/2}$. For each $i = 0, \dots, \lfloor(\log R)^6\rfloor - 1$, define H^i by

$$\sum_{m\equiv i \bmod \lfloor(\log R)^6\rfloor} f_{\theta_m}.$$

Then $f = \sum_i H^i$ where there are $\sim R^{1/2}/R^{1/2}_{N-1} = (\log R)^6$ terms in the sum. Therefore,

$$f = \frac{1}{\#i - 1} \sum_i \sum_{\tilde{i}\neq i} H^{\tilde{i}}$$

and so

$$\|f\|_{L^6(Q_R)} \le \frac{\#i}{\#i - 1} \max_i \Big\|\sum_{\tilde{i}\neq i} H^{\tilde{i}}\Big\|_{L^6(Q_R)}.$$

Let Θ be the collection of θ_m with $m \not\equiv i \bmod \lfloor(\log R)^6\rfloor$ for i achieving the maximum. $\blacksquare$

Lemma 4.5. *For each $f \in \mathcal{S}$ with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$, there exists a well-spaced collection $\tilde{\Theta}$ such that*

$$\|f\|_{L^6(Q_R)} \lesssim \Big\|\sum_{\theta\in\tilde{\Theta}} f_\theta\Big\|_{L^6(Q_R)}.$$

Proof. Define an iterative procedure. Let $F_N = f$. Apply Lemma 4.4 to F_N to obtain F_{N-1}, which is well-spaced at scale R_{N-1}.

Obtaining F_{N-k-1} (where $k \geq 1$) from F_{N-k}: We have F_{N-k} from the previous step; it has the spacing property at scales $R_{N-1}, \dots, R_{N-k}$. Furthermore,

$$\left(1 - \frac{1}{\log R}\right)^k \|f\|_{L^6(Q_R)} \leq \|F_{N-k}\|_{L^6(Q_R)}. \tag{4.8}$$

The $\sim R_{N-k}^{-1/2}$-arcs made up of unions of the θ are defined by the previous steps. Label them from left to right by $\tau_1, \dots, \tau_{n_k}$. For $i = 0, \dots, \lfloor(\log R)^6\rfloor - 1$, define H^i_{N-k-1} by

$$\sum_{m \equiv i \bmod \lfloor(\log R)^6\rfloor} (F_{N-k})_{\tau_m}.$$

Note that $F_{N-k} = \sum_i H^i_{N-k-1}$, where there are $\sim (\log R)^6$ terms in the sum, and for each θ, $(F_{N-k})_\theta$ equals $(H^i_{N-k-1})_\theta$ for exactly one i. Then

$$F_{N-k} = \frac{1}{\#i - 1} \sum_i \sum_{\tilde{i} \neq i} H^{\tilde{i}}_{N-k-1}$$

and

$$\|F_{N-k}\|_{L^6(Q_R)} \leq \frac{\#i}{\#i - 1} \max_i \Big\| \sum_{\tilde{i} \neq i} H^{\tilde{i}}_{N-k-1} \Big\|_{L^6(Q_R)}.$$

Define $F_{N-k-1} = \sum_{\tilde{i} \neq i} H^{\tilde{i}}_{N-k-1}$ where i achieves the maximum.

Iterate this procedure for $N - 1$ steps, until we obtain F_1, which is well-spaced, along with inequality (4.8) for $k = N - 1$. Since $N \sim \frac{\log R}{\log\log R}$, $\left(1 - \frac{1}{\log R}\right)^{-N} \leq e^{\frac{C}{\log\log R}} \lesssim 1$. ■

4.3. *Proof of Theorem* 1.2

We prove Theorem 1.2 using Proposition 4.1 and Lemma 4.5.

Proof of Theorem 1.2. By Lemma 4.5,

$$\frac{\|f\|^6_{L^6(Q_R)}}{\sum_\theta \|f_\theta\|^2_{L^2(\mathbb{R}^2)}} \lesssim \frac{\|\tilde{f}\|^6_{L^6(Q_R)}}{\sum_\theta \|(\tilde{f})_\theta\|^2_{L^2(\mathbb{R}^2)}}$$

where $\tilde{f} = \sum_{\theta \in \tilde{\Theta}} f_\theta$ for a well-spaced $\tilde{\Theta}$. Then by Proposition 4.1,

$$\frac{\|\tilde{f}\|^6_{L^6(Q_R)}}{\sum_\theta \|(\tilde{f})_\theta\|^2_{L^2(\mathbb{R}^2)}} \lesssim (\log R)^c \Big(\sum_{\theta \in \tilde{\Theta}} \|f_\theta\|^2_{L^\infty(\mathbb{R}^2)}\Big)^2.$$

Since $\tilde{\Theta} \subset \Theta$, we are done. ■

5. Showing $\mathrm{Dec}_6(R) \lesssim (\log R)^c$ from Theorem 1.2

5.1. *Wave packet decomposition and pigeonholing*

We will consider the following form of the decoupling inequality:

$$\|f\|_{L^6(Q_R)} \leq \mathrm{Dec}_6(R)\Big(\sum_\theta \|f_\theta\|^2_{L^6(\mathbb{R}^2)}\Big)^{1/2}. \tag{5.1}$$

The constant $\mathrm{Dec}_6(R)$ associated to this inequality is comparable to the constant where the L^6 norms are both taken over $\mathbb{R}^2$ and to the constant obtained from the L^6 norm in the upper bound being some weight function ω_{Q_R}.

Our goal is to begin with $f \in \mathcal{S}$ with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ and show it suffices to prove the decoupling inequality for a version of f which has relatively constant amplitudes and number of wave packets in each direction. In the following definitions, $\sim$ means within a factor of 2.

Write

$$f = \sum_\theta \sum_{T\in\mathbb{T}_\theta} \psi_T f_\theta \tag{5.2}$$

where for each θ, $\{\psi_T\}_{T\in\mathbb{T}_\theta}$ is a Gaussian partition of unity (meaning adds up to 1) adapted to $(\log R)^9(R\times R^{1/2})$-tubes T. Note that this implies that $|\widehat{\psi}_T| \lesssim R^{-1000}$ off of $(\log R)^{-3}(\theta - c_\theta)$, where c_θ is the center of θ.

Proposition 5.1 (Wave packet decomposition). *There exist subsets $\tilde{\Theta} \subset \Theta$ and $\tilde{\mathbb{T}}_\theta \subset \mathbb{T}_\theta$ as well as a constant $C \in [R^{-10^3}, 1]$ with the following properties:*

$$\Big\|\sum_\theta f_\theta\Big\|_{L^6(Q_R)} \lesssim (\log R)^2 \Big\|\sum_{\theta\in\tilde{\Theta}}\sum_{T\in\tilde{\mathbb{T}}_\theta} \psi_T f_\theta\Big\|_{L^6(Q_R)} + R^{-9.5}\Big(\sum_\theta \|f_\theta\|^2_{L^6(\mathbb{R}^2)}\Big)^{1/2}, \tag{5.3}$$

$$\#\tilde{\mathbb{T}}_\theta \sim \#\tilde{\mathbb{T}}_{\theta'} \quad \textit{for all } \theta,\theta' \in \tilde{\Theta}, \tag{5.4}$$

$$\|\psi_T f_\theta\|_{L^\infty(\mathbb{R}^2)} \sim CM \textit{ with } M \textit{ defined in (5.7) and for all } \theta \in \tilde{\Theta} \textit{ and } T \in \tilde{\mathbb{T}}_\theta. \tag{5.5}$$

Proof. Split the sum (5.2) into

$$f = \sum_\theta \sum_{T\in\mathbb{T}^c_\theta} \psi_T f_\theta + \sum_\theta \sum_{T\in\mathbb{T}^f_\theta} \psi_T f_\theta \tag{5.6}$$

where the close set is

$$\mathbb{T}^c_\theta := \{T \in \mathbb{T}_\theta : T \cap R^{10} Q_R \neq \emptyset\}$$

and the far set is

$$\mathbb{T}^f_\theta := \{T \in \mathbb{T}_\theta : T \cap R^{10} Q_R = \emptyset\}.$$

By Cauchy–Schwarz,

$$\begin{aligned}\Big\|\sum_\theta\sum_{T\in\mathbb{T}_\theta^f}\psi_T f_\theta\Big\|_{L^6(Q_R)} &\le R^{1/2}\Big\|\Big(\sum_\theta\Big|\sum_{T\in\mathbb{T}_\theta^f}\psi_T f_\theta\Big|^2\Big)^{1/2}\Big\|_{L^6(Q_R)}\\ &\le R^{1/2}\Big(\sum_\theta\Big\|\sum_{T\in\mathbb{T}_\theta^f}\psi_T f_\theta\Big\|^2_{L^6(Q_R)}\Big)^{1/2}\\ &\le R^{1/2}\max_\theta\Big\|\sum_{T\in\mathbb{T}_\theta^f}\psi_T\Big\|_{L^\infty(Q_R)}\Big(\sum_\theta\|f_\theta\|^2_{L^6(\mathbb{R}^2)}\Big)^{1/2}\\ &\le\frac{1}{R^{9.5}}\Big(\sum_\theta\|f_\theta\|^2_{L^6(\mathbb{R}^2)}\Big)^{1/2}\end{aligned}$$

where we have used the fact that one can bound the L^∞ norm by $C_N R^{-10N}$ for any $N\in\mathbb{N}$. This takes care of the *far* portion of f (i.e. the second term on the right-hand side of (5.6)).

The close set has cardinality $|\mathbb{T}_\theta^c|\le R^{22}$. Let

$$M=\max_\theta\max_{T\in\mathbb{T}_\theta^c}\|\psi_T f_\theta\|_{L^\infty(\mathbb{R}^2)}. \tag{5.7}$$

By Lemma 3.6,

$$M\le\max_\theta\|f_\theta\|_{L^\infty}\lesssim\Big(\sum_\theta\|f_\theta\|^2_{L^6(\mathbb{R}^2)}\Big)^{1/2}. \tag{5.8}$$

Split the remaining term as

$$\sum_\theta\sum_{T\in\mathbb{T}_\theta^c}\psi_T f_\theta=\sum_\theta\sum_{R^{-10^3}\le\lambda\le1}\sum_{T\in\mathbb{T}_{\theta,\lambda}^c}\psi_T f_\theta+\sum_\theta\sum_{T\in\mathbb{T}_{\theta,s}^c}\psi_T f_\theta \tag{5.9}$$

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

$$\begin{aligned}\mathbb{T}_{\theta,\lambda}^c&:=\{T\in\mathbb{T}_\theta^c:\|\psi_T f_\theta\|_{L^\infty(\mathbb{R}^2)}\sim\lambda M\},\\ \mathbb{T}_{\theta,s}^c&:=\big\{T\in\mathbb{T}_\theta^c:\|\psi_T f_\theta\|_{L^\infty(\mathbb{R}^2)}\le\tfrac12 R^{-10^3}M\big\}.\end{aligned}$$

Handle the *small* term from (5.9) by

$$\begin{aligned}\Big\|\sum_\theta\sum_{T\in\mathbb{T}_{\theta,s}^c}\psi_T f_\theta\Big\|_{L^6(Q_R)}&\le R^{1/2}\Big(\sum_\theta\Big\|\sum_{T\in\mathbb{T}_{\theta,s}^c}\psi_T f_\theta\Big\|^2_{L^6(Q_R)}\Big)^{1/2}\\ &\le R^{-10}M\le R^{-10}\Big(\sum_\theta\|f_\theta\|^2_{L^6(\mathbb{R}^2)}\Big)^{1/2}.\end{aligned}$$

Next decompose the remaining term from (5.9) as

$$\sum_{R^{-10^3}\le\lambda\le1}\sum_\theta\sum_{T\in\mathbb{T}_{\theta,\lambda}^c}\psi_T f_\theta=\sum_{R^{-10^3}\le\lambda\le1}\sum_{1\le j\le R^{22}}\sum_{\theta\in\Theta_j(\lambda)}\sum_{T\in\mathbb{T}_{\theta,\lambda}^c}\psi_T f_\theta \tag{5.10}$$

where j is a dyadic number in the range $[1,R^{22}]$ and $\Theta_j(\lambda)=\{\theta:|\mathbb{T}_{\theta,\lambda}^c|\sim j\}$.

Because j and λ are dyadic numbers, there is a choice of (λ, j) such that

$$\Big\| \sum_{R^{-10^3} \le \lambda \le 1} \sum_{1 \le j \le R^{22}} \sum_{\theta \in \Theta_j(\lambda)} \sum_{T \in \mathbb{T}^c_{\theta,\lambda}} \psi_T f_\theta \Big\|_{L^6(Q_R)} \lesssim (\log R)^2 \Big\| \sum_{\theta \in \Theta_j(\lambda)} \sum_{T \in \mathbb{T}^c_{\theta,\lambda}} \psi_T f_\theta \Big\|_{L^6(Q_R)}.$$

Take $\tilde{\Theta} = \Theta_j(\lambda)$ and for each $\theta \in \tilde{\Theta}$, take $\tilde{\mathbb{T}}_\theta = \mathbb{T}^c_{\theta,\lambda}$. ∎

5.2. Proof of Theorem 1.1

By Proposition 5.1, we have

$$\Big\| \sum_\theta f_\theta \Big\|_{L^6(Q_R)} \lesssim (\log R)^3 \Big\| \sum_{\theta \in \tilde{\Theta}} \sum_{T \in \tilde{\mathbb{T}}_\theta} \psi_T f_\theta \Big\|_{L^6(Q_R)} + R^{-3} \Big(\sum_\theta \| f_\theta \|^2_{L^6(\mathbb{R}^2)} \Big)^{1/2}$$

where $\#\tilde{\mathbb{T}}_\theta \sim \#\tilde{\mathbb{T}}_{\theta'}$ for all $\theta, \theta' \in \tilde{\Theta}$ and

$$\| \psi_T f_\theta \|_{L^\infty(\mathbb{R}^2)} \sim A := \max_{\theta' \in \tilde{\Theta}} \max_{T' \in \tilde{\mathbb{T}}_\theta} \| \psi_{T'} f_{\theta'} \|_{L^\infty(\mathbb{R}^2)}. \tag{5.11}$$

Since the Fourier transform of $\sum_{T \in \tilde{\mathbb{T}}_\theta} \psi_T f_\theta$ is essentially supported in $(1 + (\log R)^{-3})\theta$, there exists a function f'_θ with Fourier transform supported in $(1 + (\log R)^{-3})\theta$ such that

$$\sum_{T \in \tilde{\mathbb{T}}_\theta} \psi_T f_\theta(x) = f'_\theta(x) + O(R^{-998})A.$$

Thus

$$\Big\| \sum_{\theta \in \tilde{\Theta}} \sum_{T \in \tilde{T}_\theta} \psi_T f_\theta \Big\|_{L^6(Q_R)} \le \Big\| \sum_{\theta \in \tilde{\Theta}} f'_\theta \Big\|_{L^6(Q_R)} + R^{-997} A. \tag{5.12}$$

The functions f'_θ have Fourier support in $(1 + (\log R)^{-3})\theta$. We may split $\tilde{\Theta}$ into ~ 1 sets $\tilde{\Theta}_i$ where for distinct $\theta, \theta' \in \tilde{\Theta}_i$,

$$(2\theta) \cap (2\theta') = \emptyset.$$

Then for some i,

$$\Big\| \sum_{\theta \in \tilde{\Theta}} f'_\theta \Big\|^6_{L^6(Q_R)} \lesssim \Big\| \sum_{\theta \in \tilde{\Theta}_i} f'_\theta \Big\|^6_{L^6(Q_R)},$$

and it follows from Theorem 1.2 that

$$\Big\| \sum_{\theta \in \tilde{\Theta}} f'_\theta \Big\|^6_{L^6(Q_R)} \lesssim (\log R)^c \Big(\sum_{\theta \in \tilde{\Theta}_i} \| f'_\theta \|^2_{L^\infty(\mathbb{R}^2)} \Big)^2 \sum_{\theta \in \tilde{\Theta}_i} \| f'_\theta \|^2_{L^2(\mathbb{R}^2)}. \tag{5.13}$$

Note that

$$\Big(\sum_{\theta\in\tilde{\Theta}_i}\|f_\theta'\|_{L^\infty(\mathbb{R}^2)}^2\Big)^{1/2}\le\Big(\sum_{\theta\in\tilde{\Theta}}\Big\|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta\Big\|_{L^\infty(\mathbb{R}^2)}^2\Big)^{1/2}+R^{-500}A,$$

$$\Big(\sum_{\theta\in\tilde{\Theta}_i}\|f_\theta'\|_{L^2(\mathbb{R}^2)}^2\Big)^{1/2}\lesssim\Big(\sum_{\theta\in\tilde{\Theta}}\Big\|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta\Big\|_{L^2(\mathbb{R}^2)}^2\Big)^{1/2}+R^{-500}A.$$

Combining these observations with (5.13) gives

$$\Big\|\sum_{\theta\in\tilde{\Theta}}f_\theta'\Big\|_{L^6(Q_R)}^6\lesssim(\log R)^c\Big(\sum_{\theta\in\tilde{\Theta}}\Big\|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta\Big\|_{L^\infty(\mathbb{R}^2)}^2\Big)^2\sum_{\theta\in\tilde{\Theta}}\Big\|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta\Big\|_{L^2(\mathbb{R}^2)}^2 + R^{-2000}A^6. \tag{5.14}$$

The second term is bounded by

$$R^{-2000}A^6\le R^{-2000}\Big(\sum_\theta\|f_\theta\|_{L^6(\mathbb{R}^2)}^2\Big)^3 \tag{5.15}$$

using Lemma 3.6 and the fact that $0\le\psi_T\le 1$. It remains to analyze the first term in the upper bound in (5.14). For each $\theta\in\tilde{\Theta}$, we have

$$\Big|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta(x)\Big|\lesssim\Big|\sum_{\substack{T\in\tilde{\mathbb{T}}_\theta\\ x\in(\log R)T}}\psi_T f_\theta(x)\Big|+R^{-1000}A\le(\log R)^2A+R^{-1000}A.$$

This leads to the following upper bound for the first term on the right-hand side of (5.14):

$$\Big(\sum_{\theta\in\tilde{\Theta}}\Big\|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta\Big\|_{L^\infty(\mathbb{R}^2)}^2\Big)^2\le(\log R)^8(\#\tilde{\Theta}A^2)^2+R^{-1000}A^4. \tag{5.16}$$

For each $\theta\in\tilde{\Theta}$, we also have

$$\begin{aligned}\int\Big|\sum_{T\in\tilde{\mathbb{T}}_\theta}\psi_T f_\theta\Big|^2dx&\lesssim(\log R)^2\sum_{T\in\tilde{\mathbb{T}}_\theta}\int_{(\log R)T}|\psi_T f_\theta|^2dx+R^{-1998}A^2\\&\lesssim(\log R)^4\#\tilde{\mathbb{T}}_\theta A^2|T|+R^{-1998}A^2.\end{aligned}$$

Combining this with (5.16) leads to the upper bound

$$\Big\|\sum_{\theta\in\tilde{\Theta}}f_\theta'\Big\|_{L^6(Q_R)}^6\lesssim(\log R)^{c+12}\#\tilde{\Theta}^3\#\tilde{\mathbb{T}}_\theta A^6|T|+R^{-1000}A^6. \tag{5.17}$$

Finally, note that for each $\theta\in\tilde{\Theta}$,

$$\#\tilde{\mathbb{T}}_\theta|T|A^6\lesssim(\log R)^2\sum_{T\in\tilde{\mathbb{T}}_\theta}\int|\psi_T f_\theta|^6\omega_T+R^{-2000}A^6$$

where we use the locally constant property. Finally, since the ℓ^6 norm is bounded by the ℓ^1 norm and $0 \le \psi_T \le 1$,

$$\sum_{T \in \tilde{\mathbb{T}}_\theta} \int |\psi_T f_\theta|^6 \le \int \Big(\sum_{T \in \tilde{\mathbb{T}}_\theta} |\psi_T f_\theta| \Big)^6 \le \|f_\theta\|^6_{L^6(\mathbb{R}^2)}.$$

In summary, we finish the proof by combining (5.17) with (5.11) and

$$\#\tilde{\Theta}^3 \# \tilde{\mathbb{T}}_\theta |T| A^6 \lesssim (\log R)^2 \Big(\sum_{\theta \in \tilde{\Theta}} \|f_\theta\|^2_{L^6(\mathbb{R}^2)} \Big)^3. \qquad \blacksquare$$

Appendix

In the appendix, we show that a slight modification of the proof gives Theorem 1.2 and Theorem 1.1 for a set of curves

$$\mathcal{K} := \{(\xi_1, h(\xi_1)) : |\xi_1| \le 1\},$$

where h are C^2 functions satisfying $h(0) = h'(0) = 0$, and $1/2 \le h''(\xi_1) \le 2$ for $|\xi| \le 1$. In particular, one can cut the unit circle into $O(1)$ arcs, such that each arc is part of a curve in $\mathcal{K}$ after translation and rotation. The truncated parabola $\mathbb{P}^1$ is also in $\mathcal{K}$.

To prove Theorems 1.2 and 1.1 for the curves in $\mathcal{K}$, replace the affine map (4.2) in Lemma 4.3 by

$$\ell(\xi_1, \xi_2) = \big(\nu(\xi_1 - c), \nu^2(\xi_2 - h(c) - h'(c)(\xi_1 - c))\big) \tag{5.18}$$

with $\nu = R_k^{1/2}$. Since the curve $\{(\nu(\xi_1 - c), \nu^2(h(\xi_1) - h(c) - h'(c)(\xi_1 - c))) : |\xi_1 - c| \le \nu^{-1}\}$ is also in $\mathcal{K}$, the proof of Lemma 4.3 remains unchanged provided that Proposition 3.5 holds for all the curves in $\mathcal{K}$ for a smaller R.

Then it suffices to check (3.14) for the curves in $\mathcal{K}$. Let $(\xi, h(\xi))$ be the center of τ_j. Assume that

$$\xi, \xi'' \in \tau_k, \quad \xi', \xi''' \in \tau_k',$$

and

$$\xi - \xi'' = \xi''' - \xi' + O(R_j^{-1/2}).$$

Then $h(\xi) - h(\xi'') = h'(\xi_1)(\xi - \xi'')$ for some ξ_1 between ξ and ξ'', and $h(\xi''') - h(\xi') = h'(\xi_2)(\xi''' - \xi')$ for some ξ_2 between ξ''' and ξ'. We have $h'(\xi_1) + (\xi_2 - \xi_1)/2 \le h'(\xi_2)$ since $1/2 \le h''(\xi) \le 1$. Since $\xi_2 - \xi_1 \ge \operatorname{dist}(\tau_k, \tau_k')$, we obtain

$$h(\xi) - h(\xi'') - (h(\xi''') - h(\xi')) \gtrsim \operatorname{dist}(\tau_k, \tau_k') R_j^{-1/2}$$

if $\tau_j, \tau_j', \tau_j'', \tau_j'''$ have pairwise distances $\gtrsim R_j^{-1/2}$. So (3.14) is verified.

Decoupling for the circle with explicit decoupling constant $(\log R)^c$ has an application to a problem about sums of two squares. The problem arises in the study of Laplace eigenfunctions for the standard two-dimensional torus.

Let Λ_m be the set of Gaussian integers $\lambda = x + \sqrt{-1}\, y$, $x, y \in \mathbb{Z}$, with norm $\lambda\overline{\lambda} = m$.

Problem. Give a non-trivial upper bound for the number of solutions of

$$\lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6, \quad \lambda_j \in \Lambda_m.$$

Corollary 5.2. *For any $\Lambda \subset \Lambda_m$, if $N = |\Lambda| > (\log m)^{(c+6)/\epsilon}$ for some $\epsilon > 0$ and the constant c as in Theorem 1.2, then*

$$N_m(\Lambda) := \#\{\lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6 : \lambda_j \in \Lambda\} \lesssim N^{3+\epsilon}.$$

Proof. Consider the function

$$g(z) = \sum_{\lambda \in \Lambda} e^{2\pi i \frac{\lambda}{\sqrt{m}} \cdot z}$$

for $z \in \mathbb{R}^2$. Let $Q_0 = [0, m(\log m)]^2$ and $\{Q\}$ be a tiling of $\mathbb{R}^2$ with translates of Q_0. Let $\{\varphi_Q\}$ be the Gaussian partition of unity defined as in Definition 3.15. Define the weight function

$$\psi = \sum_{Q:\, \mathrm{dist}(Q, Q_0) \le m(\log m)^2} \varphi_Q.$$

Then $|\psi - 1| \le m^{-1000}$ on Q_0 and $|\psi| \le m^{-1000}$ outside of $(\log m)^2 Q_0$ and rapidly decays away from it. Moreover, the Fourier transform $\widehat{\psi}$ is essentially supported on $B(0, m^{-1})$

We apply Theorem 1.1 (for the circle) to the function $f(z) = g(z)\psi$ with $R = m$ and $Q_R = [0, m]^2$. Then $f_\theta = e^{2\pi i \lambda \cdot z/\sqrt{m}}$ for the (unique) $\lambda/\sqrt{m} \in \theta$. Note that $g(z)$ is periodic: $g(z + \sqrt{m} v) = g(z)$ for any $v \in \mathbb{Z}^2$. Since $|f - g| \le m^{-1000}$ on Q_R, we have

$$N_m(\Lambda) \lesssim (\log m)^{c+6} |\Lambda|^3$$

where the $(\log m)^6$ comes from a weight function ψ essentially supported in $[0, (\log m)^3 m]^2$. ∎

This problem was studied by Bombieri and Bourgain [1] using various methods. In particular, they obtained the bound $O(|\Lambda_m|^{3+\epsilon})$ assuming the Riemann hypothesis and the Birch and Swinnerton-Dyer conjecture for the L-functions of elliptic curves over $\mathbb{Q}$ and for a random m with $|\Lambda_m| \sim 2^{\omega(m)}$, $\omega(m) \sim \frac{\log m}{A \log\log m}$ for some constant A. Based on the Bourgain–Demeter decoupling, Zane Li [8] obtained the result of Bombieri and Bourgain unconditionally for all m with $|\Lambda_m| > \exp((\log m)^{1-o(1)})$. Corollary 5.2 proves the result for a larger range of $|\Lambda|$: $|\Lambda| > (\log m)^{c/o(1)}$.

In [3], it was conjectured that for any $\Lambda \subset \Lambda_m$, and any $\epsilon > 0$, there exists C_ϵ independent of m such that

$$\#\{\lambda_1 + \lambda_2 + \lambda_3 = \lambda_4 + \lambda_5 + \lambda_6 : \lambda_j \in \Lambda\} \le C_\epsilon |\Lambda|^{3+\epsilon}.$$

Corollary 5.2 confirms this for $|\Lambda| > (\log m)^{(c+6)/\epsilon}$.

References

[1] Bombieri, E., Bourgain, J.: A problem on sums of two squares. Int. Math. Res. Notices **2015**, 3343–3407 Zbl 1364.11091 MR 3373053

[2] Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. Geom. Funct. Anal. **3**, 107–156 (1993) Zbl 0787.35097 MR 1209299

[3] Bourgain, J., Demeter, C.: The proof of the l^2 decoupling conjecture. Ann. of Math. (2) **182**, 351–389 (2015) Zbl 1322.42014 MR 3374964

[4] Demeter, C., Guth, L., Wang, H.: Small cap decouplings. Geom. Funct. Anal. **30**, 989–1062 (2020) Zbl 1454.42008 MR 4153908

[5] Guth, L., Wang, H., Zhang, R.: A sharp square function estimate for the cone in $\mathbb{R}^3$. Ann. of Math. (2) **192**, 551–581 (2020) MR 4151084

[6] Krishnapur, M., Kurlberg, P., Wigman, I.: Nodal length fluctuations for arithmetic random waves. Ann. of Math. (2) **177**, 699–737 (2013) Zbl 1314.60101 MR 3010810

[7] Li, Z. K.: Decoupling for the parabola and connections to efficient congruencing. Ph.D. thesis, UCLA (2019) MR 4024357

[8] Li, Z. K.: Effective l^2 decoupling for the parabola. Mathematika **66**, 681–712 (2020) Zbl 1442.42032 MR 4134240

[9] Li, Z. K.: An l^2 decoupling interpretation of efficient congruencing: the parabola. Rev. Mat. Iberoamer. **37**, 1761–1802 (2021) Zbl 1467.42019 MR 4276295

[10] Rudnick, Z., Wigman, I.: On the volume of nodal sets for eigenfunctions of the Laplacian on the torus. Ann. Henri Poincaré **9**, 109–130 (2008) Zbl 1142.60029 MR 2389892