SHARP SUPERLEVEL SET ESTIMATES FOR SMALL CAP DECOUPLINGS OF THE PARABOLA

YUQIU FU, LARRY GUTH, AND DOMINIQUE MALDAGUE

Abstract. We prove sharp bounds for the size of superlevel sets \( \{ x \in \mathbb{R}^2 : |f(x)| > \alpha \} \) where \( \alpha > 0 \) and \( f : \mathbb{R}^2 \to \mathbb{C} \) is a Schwartz function with Fourier transform supported in an \( R^{-1} \)-neighborhood of the truncated parabola \( \mathbb{P}^1 \). These estimates imply the small cap decoupling theorem for \( \mathbb{P}^1 \) from [DGW20] and the canonical decoupling theorem for \( \mathbb{P}^1 \) from [BD15]. New \((\ell^q, L^p)\) small cap decoupling inequalities also follow from our sharp level set estimates.

In this paper, we further develop the high/low frequency proof of decoupling for the parabola [GMW20] to prove sharp level set estimates which recover and refine the small cap decoupling results for the parabola in [DGW20]. We begin by describing the problem and our results in terms of exponential sums. The main results in full generality are in §1.

For \( N \geq 1, R \in [N, N^2] \), and \( 2 \leq p \leq 2p \), let \( D(N, R, p) \) denote the smallest constant so that

\[
|Q_R|^{-1} \int_{Q_R} \left| \sum_{\xi \in \Xi} a_\xi e((x,t) \cdot (\xi,\xi^2)) \right|^p dx dt \leq D(N, R, p) N^{p/2}
\]

for any collection \( \Xi \subset [-1,1] \) with \( |\Xi| \sim N \) consisting of \( \sim \frac{1}{N} \)-separated points, \( a_\xi \in \mathbb{C} \) with \( |a_\xi| \sim 1 \), and any cube \( Q_R \subset \mathbb{R}^2 \) of sidelength \( R \).

A corollary of the small cap decoupling theorem for the parabola in [DGW20] is that if \( 2 \leq p \leq 2 + 2s \) for \( R = N^s \), then

\[
D(N, R, p) \leq C_\varepsilon N^\varepsilon.
\]

This estimate is sharp, up to the \( C_\varepsilon N^\varepsilon \) factor, which may be seen by Khintchine’s inequality. The range \( 2 \leq p \leq 2 + 2s \) is the largest range of \( p \) for which \( D(N, R, p) \) may be bounded by sub-polynomial factors in \( N \). The case \( R = N^2 \) of (2) follows from the canonical \( \ell^2 \) decoupling theorem of Bourgain and Demeter for the parabola [BD15]. For \( R < N^2 \) and the subset \( \Xi = \{ k/N \}_{k=1}^{N} \), the inequality (1) is an estimate for the moments of exponential sums over subsets smaller than the full domain of periodicity (i.e. \( N^2 \) in the \( t \)-variable). Bourgain investigated examples of this type of inequality in [Bou17a, Bou17b].

By a pigeonholing argument (see Section 5 of [GMW20]), (2) follows from upper bounds for superlevel sets \( U_\alpha \) defined by

\[
U_\alpha = \{ (x,t) \in \mathbb{R}^2 : \left| \sum_{\xi \in \Xi} a_\xi e((x,t) \cdot (\xi,\xi^2)) \right| > \alpha \}.
\]

In particular, (2) is equivalent, up to a \( \log N \) factor, to proving that for any \( \alpha > 0 \) and for \( R = N^s \),

\[
\alpha^{2 + 2s} |U_\alpha \cap Q_R| \leq C_\varepsilon R^\varepsilon N^{1 + s} R^2
\]
when $\Xi$, $a_\xi$ satisfy the hypotheses following (1). In this paper, we improve the above superlevel set estimate for all $\alpha > 0$ strictly between $N^{1/2}$ and $N$.

**Theorem 1.** Let $R \in [N, N^2]$. For any $\varepsilon > 0$, there exists $C_\varepsilon < \infty$ such that

$$|U_\alpha \cap Q_R| \leq C_\varepsilon N^\varepsilon \begin{cases} \frac{N^2 R}{\alpha} \sum_{\xi \in \Xi} |a_\xi|^2 & \text{if } \alpha^2 > R \\ \frac{N^2 R^2}{\alpha} \sum_{\xi \in \Xi} |a_\xi|^2 & \text{if } N \leq \alpha^2 \leq R \\ R^2 & \text{if } \alpha^2 < N. \end{cases}$$

whenever $\Xi \subset [-1, 1]$ is a $\gtrsim \frac{1}{N}$-separated subset, $|a_\xi| \leq 1$ for each $\xi \in \Xi$, and $Q_R \subset \mathbb{R}^2$ is a cube of sidelength $R$.

Our superlevel set estimates are essentially sharp, which follows from analyzing the function $F(x,t) = \sum_{n=1}^N e((x,t) \cdot (\frac{n}{N}, \frac{n^2}{N^2}))$. It is not known whether the implicit constant in the upper bound of (2) goes to infinity with $N$ except in the case that $p = 6$ and $s = 2$, when the same example $F(x,t) = \sum_{n=1}^N e((x,t) \cdot (\frac{n}{N}, \frac{n^2}{N^2}))$ shows that $D(N, N^2, 6) \gtrsim (\log N)$ [Bon93]. Roughly, the argument is that for each dyadic value $\alpha \in [N^{3/4}, N]$, one can show by counting the “major arcs” that

$$\alpha^6 \{(x,t) \in Q_{N^2} : |F(x,t)| \sim \alpha\} \gtrsim N^4 \cdot N^3.$$ 

Since there are $\sim \log N$ values of $\alpha$, the lower bound for $\int_{Q_{N^2}} |F|^6$ follows. Theorem 1 implies that the corresponding superlevel set estimates (3) are not sharp for $1 \leq s < 2$, unless $\alpha \sim N$ or $\alpha^2 \sim N$, which leads to the following conjecture.

**Conjecture 2.** Let $s \in [1, 2)$ and $2 \leq p \leq 2 + 2s$. There exists $C(s) > 0$ so that

$$D(N, N^s, p) \leq C(s).$$

A more refined version of Theorem 1 leads to the following essentially sharp ($\ell^q$, $L^p$) small cap decoupling theorem, stated here for general exponential sums.

**Corollary 1.** Let $\frac{3}{p} + \frac{1}{q} \leq 1$, and let $R \in [N, N^2]$. Then for each $\varepsilon > 0$, there exists $C_\varepsilon < \infty$ so that

$$\| \sum_{\xi \in \Xi} a_\xi e((x,t) \cdot (\xi, \xi^2)) \|_{L^p(B_R)} \leq C_\varepsilon N^\varepsilon (N^{1 - \frac{1}{p} - \frac{1}{q}} + N^{1 - \frac{1}{p} - \frac{2}{q}} R^2) \left( \sum |a_\xi|^q \right)^{1/q}.$$ 

In the above corollary, the assumptions are that $\Xi$ is a $\gtrsim \frac{1}{N}$-separated subset of $[-1, 1]$ and that $a_\xi \in \mathbb{C}$.

### 1. Main results

We state our main results in the more general set-up for decoupling. Let $\mathbb{P}^1$ denote the truncated parabola

$$\{(t, t^2) : |t| \leq 1\}$$

and write $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ for the $R^{-1}$-neighborhood of $\mathbb{P}^1$ in $\mathbb{R}^2$, where $R \geq 2$. For a partition $\{\gamma\}$ of $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ into almost rectangular blocks, an ($\ell^2$, $L^p$) decoupling inequality is

$$\|f\|_{L^p(B_R)} \leq D(R, p) \left( \sum_{\gamma} \|f_{\gamma}\|_{L^p(\mathbb{R}^2)}^2 \right)^{1/2}$$

in which $f : \mathbb{R}^2 \to \mathbb{C}$ is a Schwartz function with $\text{supp} \hat{f} \subset \mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ and $f_{\gamma}$ means the Fourier projection onto $\gamma$, defined precisely below. When we refer to canonical caps or to
canonical decoupling, we mean that $\gamma$ are approximately $R^{-1/2} \times R^{-1}$ blocks corresponding to the $l^2$-decoupling paper of [BD15]. In this paper, we allow $\gamma$ to be approximate $R^{-\beta} \times R^{-1}$ blocks, where $\frac{1}{2} \leq \beta \leq 1$. This is the "small cap" regime studied in [DGW20]. We also consider $(\ell^q, L^p)$ decoupling for small caps, which replaces $\left(\sum_\gamma \|f_\gamma\|_p^2\right)^{1/2}$ by $\left(\sum_\gamma \|f_\gamma\|_p^q\right)^{1/q}$ in the decoupling inequality above (see Corollary 5).

To precisely discuss the collection $\{\gamma\}$, fix a $\beta \in \left[\frac{1}{2}, 1\right]$. Let $\mathcal{P} = \mathcal{P}(R, \beta) = \{\gamma\}$ be the partition of $N_{R^{-1}}(\mathbb{P}^1)$ given by

$$\bigcup_{|k| \leq \lceil R^\beta \rceil - 2} \{(x, t) \in N_{R^{-1}}(\mathbb{P}^1) : k[R^\beta]^{-1} \leq x < (k + 1)[R^\beta]^{-1}\}$$

and the two end pieces

$$\{(x, t) \in N_{R^{-1}}(\mathbb{P}^1) : x < -1 + [R^\beta]^{-1}\} \cup \{(x, t) \in N_{R^{-1}}(\mathbb{P}^1) : 1 - [R^\beta]^{-1} \leq x\}.$$ 

For a Schwartz function $f : \mathbb{R}^2 \to \mathbb{C}$ with $\text{supp} \hat{f} \subset N_{R^{-1}}(\mathbb{P}^1)$, define for each $\gamma \in \mathcal{P}(R, \beta)$

$$f_\gamma(x) := \int_{\gamma} \hat{f}(\xi)e^{2\pi i x \cdot \xi} d\xi.$$ 

For $a, b > 0$, the notation $a \lesssim b$ means that $a \leq Cb$ where $C > 0$ is a universal constant whose definition varies from line to line, but which only depends on fixed parameters of the problem. Also, $a \sim b$ means $C^{-1}b \leq a \leq Cb$ for a universal constant $C$.

Let $U_\alpha := \{x \in \mathbb{R}^2 : |f(x)| \geq \alpha\}$. In Section 5 of [GMW20], through a wave packet decomposition and series of pigeonholing steps, bounds for $D(R, p)$ in (4) follow (with an additional power of $(\log R)$) from bounds on the constant $C(R, p)$ in

$$\alpha^{|U_\alpha|} \leq C(R, p)(\#\{\gamma : f_\gamma \neq 0\})^{\beta - 1} \sum_\gamma \|f_\gamma\|_2^2$$

for any $\alpha > 0$ and under the additional assumptions that $\|f_\gamma\|_\infty \leq 1$, $\|f_\gamma\|_p \sim \|f_\gamma\|_2$ for each $\gamma$. Thus decoupling bounds follow from upper bounds on the superlevel set $|U_\alpha|$. In this paper, we consider the question: given $\alpha > 0$ and a partition $\{\gamma\}$, how large can $|U_\alpha|$ be, varying over functions $f$ satisfying $\|f_\gamma\|_\infty \lesssim 1$ for each $\gamma$? We answer this question in the following theorem.

**Theorem 3.** Let $\beta \in \left[\frac{1}{2}, 1\right]$, $R \geq 2$. Let $f : \mathbb{R}^2 \to \mathbb{C}$ be a Schwartz function with Fourier transform supported in $N_{R^{-1}}(\mathbb{P}^1)$ satisfying $\|f_\gamma\|_\infty \leq 1$ for all $\gamma \in \mathcal{P}(R, \beta)$. Then for any $\alpha > 0$,

$$|U_\alpha \cap [-R, R]^2| \leq C_\varepsilon R^\varepsilon \begin{cases} \frac{R^{2\beta - 1}}{\alpha^\beta} \sum_\gamma \|f_\gamma\|^2_{L^2(\mathbb{R}^2)} & \text{if } \alpha^2 > R \\ \frac{R^{2\beta}}{\alpha} \sum_\gamma \|f_\gamma\|^2_{L^2(\mathbb{R}^2)} & \text{if } R^\beta \leq \alpha^2 \leq R^3 \\ R^2 & \text{if } \alpha^2 < R^3. \end{cases}$$

Each bound in Theorem 3 is sharp, up to the $C_\varepsilon R^\varepsilon$ factor, which we show in 3.2. Define notation for a distribution function for the Fourier support of a Schwartz function $f$ with Fourier transform supported in $N_{R^{-1}}(\mathbb{P}^1)$ as follows. For each $0 \leq s \leq 2$, let

$$\lambda(s) = \sup_{\omega(s)} \#\{\gamma : \gamma \cap \omega(s) \neq \emptyset, f_\gamma \neq 0\}$$

where $\omega(s)$ is any arc of $\mathbb{P}^1$ with projection onto the $\xi_1$-axis equal to an interval of length $s$. The following theorem implies Theorem 3 and replaces factors of $R^3$ in the upper bounds.
Theorem 4. Let $\beta \in [\frac{1}{2}, 1]$, $R \geq 2$. For any $f$ with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ satisfying $\|f\|_{\mathcal{N}} \lesssim 1$ for each $\gamma \in \mathcal{P}(R, \beta)$,

$$|U_\alpha| \leq C_\varepsilon R^\varepsilon \left\{ \begin{array}{ll} \frac{1}{\alpha^2} \max_{s} \sum_{\gamma} \lambda(s) \|f\|_2^2 & \text{if } \alpha^2 > \frac{\lambda(1)^2}{\max_{s} \lambda(s)^{1+R^{-1}}} \\ \frac{\lambda(1)^2}{\alpha^2} \sum_{\gamma} \|f\|_2^2 & \text{if } \alpha^2 \leq \frac{\lambda(1)^2}{\max_{s} \lambda(s)^{1+R^{-1}}} \end{array} \right. $$

in which the maxima are taken over dyadic $s$, $R^{-\beta} \leq s \leq R^{-1/2}$.

Corollary 5 ($([l^q, l^p]$ small cap decoupling). Let $\frac{3}{p} + \frac{1}{q} \leq 1$. Then

$$\|f\|_{l^p(B_R)} \leq C_\varepsilon R^\varepsilon \left( R^\beta \left( \frac{1}{2} - \frac{\beta}{2q} \right) \right) \left( \|f\|_{l^p([\mathbb{R}^2])^1/q} \right)$$

whenever $f$ is a Schwartz function with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$.

The powers of $R$ in the upper bound come from considering two natural sharp examples for the ratio $\|f\|_{l^p(B_R)}/\left( \sum_{\gamma} \|f\|_p^p \right)^{1/q}$. The first is the square root cancellation example, where $\|f\| \sim \chi B_R$ for all $\gamma$ and $f = \sum_{\gamma} e_{\gamma} f_{\gamma}$ where $e_{\gamma}$ are $\pm 1$ signs chosen (using Khintchine’s inequality) so that $\|f\|_{l^p(B_R)} \sim R^{\beta(p/2)R^2}$.

$$\|f\|_{l^p(B_R)}^p \left( \sum_{\gamma} \|f_{\gamma}\|_{l^p}^p \right)^{1/q} \gtrsim \left( R^{\beta(p/2)R^2} \right) / \left( R^{\beta(p/q)R^2} \right) \sim R^{\beta(p/2 - \beta)}$$

The second example is the constructive interference example. Let $f_{\gamma} \sim R^{1+\beta} \tilde{\eta}_{\gamma}$ where $\tilde{\eta}_{\gamma}$ is a smooth bump function approximating $\chi_{\gamma}$. Since $|f| = \sum_{\gamma} \gamma f_{\gamma}$ is approximately constant on unit balls and $|f(0)| \sim R^\beta$, we have

$$\|f\|_{l^p(B_R)}^p \left( \sum_{\gamma} \|f_{\gamma}\|_{l^p}^p \right)^{1/q} \gtrsim \left( R^{\beta(p)R^1} \right) / \left( R^{\beta(p/q)R^{1+\beta}} \right) \sim R^{\beta(p/(1-q) - 1 - \beta)}$$

There is one more example which may dominate the ratio: The block example is $f = R^{1+\beta} \tilde{\eta}_{\gamma}$ where $\tilde{\eta}_{\gamma}$ is a canonical $R^{-1/2} \times R^{-1}$ block. Since $f = f_{\theta}$ and $|f_{\theta}|$ is approximately constant on dual $\sim R^{1/2} \times R$ blocks $\theta^*$, we have

$$\|f\|_{l^p(B_R)}^p \left( \sum_{\gamma} \|f_{\gamma}\|_{l^p}^p \right)^{1/q} \gtrsim \left( R^{\beta(p)/(1-q)} \right)$$

One may check that the constructive interference examples dominate the block example when $\frac{3}{p} + \frac{1}{q} \leq 1$. We do not investigate $([l^q, l^p]$ small cap decoupling in the range $\frac{3}{p} + \frac{1}{q} > 1$ in the present paper.

The paper is organized as follows. In §2 we demonstrate that Theorem 4 is sharp using an exponential sum example. In §3 we show how Theorem 4 follows easily from Theorem 3 and how after some pigeonholing steps, so does Corollary 5. Then in §4 we develop the multi-scale high/low frequency tools we use in the proof of Theorem 4. These tools are very similar to those developed in [GMW20]. It appears that a more careful version of the proof of Theorem 4 could also replace the $C_\varepsilon R^\varepsilon$ factor by a power of $(\log R)$, as is done for canonical decoupling in [GMW20]. Finally, in §5 we prove a bilinear version of Theorem 4 and then reduce to the bilinear case to finish the proof.

LG is supported by a Simons Investigator grant. DM is supported by the National Science Foundation under Award No. 2103249.
2. A sharp example

Because we will show that Theorem 4 implies Theorem 3, it suffices to show that Theorem 3 is sharp, which we mean up to a $C_\varepsilon R^\varepsilon$ factor. Write $N = \lfloor R^2 \rfloor$. The function achieving the sharp bounds is

$$F(x_1, x_2) = \sum_{k=1}^{N} e^{\left(\frac{k}{N} x_1 + \frac{k^2}{N^2 x_2}\right)} \eta(x_1, x_2),$$

where $\eta$ is a Schwartz function satisfying $\eta \sim 1$ on $[-R, R]^2$ and supp $\hat{\eta} \subset B_{R-1}$. We will bound the set

$$U_\alpha = \{(x_1, x_2) \in [-R, R]^2 : |F(x_1, x_2)| \geq \alpha\}.$$

Case 1: $R < \alpha^2$.

Suppose that $\alpha \sim N$ and note that $F(0,0) = N$ and $|F(0,0)| \sim N$ when $|(x_1, x_2)| \leq R/N$. Using periodicity in the $x_1$ variable, there are $\sim R/N$ many other heavy balls where $|F(x)| \sim N$ in $[-R, R]^2$. For $\alpha$ in the range suppose that $R < \alpha^2 < N^2$, we will show that $U_\alpha$ is dominated by larger neighborhoods of the heavy balls.

Let $r = N^2/\alpha^2$ and assume without loss of generality that $r$ is in the range $R^\varepsilon < r < N^2/R \sim R^{2\beta-1} \ll N$. The upper bound for $|U_\alpha|$ in Theorem 3 for this range is

$$|U_\alpha| \leq C_\varepsilon R^\varepsilon \frac{N^2}{\alpha^4 R} \sum_{\gamma} \|F_{\gamma}\|^2_2 \sim C_\varepsilon R^\varepsilon \frac{N^2}{\alpha^4 R} N R^2.$$

To demonstrate that this inequality is sharp, by periodicity in $x_1$, it suffices to show that $|U_\alpha \cap B_r| \gtrsim r^2$. Let $\phi_{r-1}$ be a nonnegative bump function supported in $B_{r-1/2}$ with $\phi_{r-1} \gtrsim 1$ on $B_{r-1/4}$. Let $\eta_r = r^4(\phi_{r-1} * \phi_{r-1})^-$ and analyze the $L^2$ norm $\|F\|_{L^2(\eta_r)}$. By Plancherel’s,

$$\|F\|^2_{L^2(\eta_r)} = \int |F|^2 \eta_r = \int \left| \sum_{k=1}^{N} e^{\left(\frac{k}{N} x_1 + \frac{k^2}{N^2 x_2}\right)} \eta_r(x_1, x_2) \right|^2 \sim N \cdot N / r \cdot r^2 = r N^2.$$

Next we bound $\|F\|_{L^4(B_{r^\varepsilon})}$ above. It follows from the local linear restriction statement (see [Dem20] Theorem 1.14, Prop 1.27, and Exercise 1.32)

$$\|f\|^4_{L^4(B_{r^\varepsilon})} \lesssim C_\varepsilon R^O(\varepsilon) r^{-3} \|\hat{f}\|^4_{L^4(R^2)}$$

that

$$\|F\|^4_{L^4(B_{r^\varepsilon})} \sim \left\| \sum_{k=1}^{N} e^{\left(\frac{k}{N} x_1 + \frac{k^2}{N^2 x_2}\right)} \eta_r(x_1, x_2) \right\|^4_{L^4(B_{r^\varepsilon})} \lesssim C_\varepsilon r^\varepsilon r^{-3} \left\| \sum_{k=1}^{N} \hat{\eta}_r(\xi - (\frac{k}{N}, \frac{k^2}{N^2})) \right\|^4_{L^4(R^2)}.$$
The $L^4$ norm on the right hand side is bounded above by
\[
\int_{B_{2}} |\sum_{k=1}^{N} \tilde{\eta}_{r}(\xi - (kN, k^2))|^{4} d\xi \lesssim (N r^{-1})^{3} \int_{B_{2}} |\tilde{\eta}_{r}(\xi - (kN, k^2))|^{4} d\xi
\]
\[
\lesssim (N r^{-1})^{3} (r^2)^{3} \int_{B_{2}} |\tilde{\eta}_{r}(\xi - (kN, k^2))| d\xi \sim N^4 r^3.
\]
This leads to the upper bound $\|F\|_{L^4(B_{R^r r^4})}^{1} \lesssim (\log R) N^4$.

Finally, by dyadic pigeonholing, there is some $\lambda \in [R^{-1000}, N]$ so that $\|F\|_{L^2(\eta_r)}^{2} \lesssim (\log R) \lambda^2 |\{x \in B_{R^r r^4} : |F(x)| \sim \lambda\}| + C_\varepsilon R^{-2000}$. The lower bound for $\|F\|_{L^2(\eta_r)}^{2}$ and the upper bound for $\|F\|_{L^4(B_{R^r r^4})}^{1}$ tell us that
\[
\lambda^2 r N^2 \sim \lambda^2 \|F\|_{L^2(\eta_r)}^{2} \|\{x \in B_{R^r r^4} : |F(x)| \sim \lambda\}| + C_\varepsilon \lambda^4 R^{-2000}
\]
\[
\lesssim (\log R) \|F\|_{L^4(B_{R^r r^4})}^{1} + C_\varepsilon \lambda^4 R^{-2000} \lesssim C_\varepsilon r^\varepsilon N^4 + C_\varepsilon \lambda^4 R^{-2000}.
\]
Conclude that $\lambda^2 \lesssim C_\varepsilon r^\varepsilon N^2/r \sim C_\varepsilon R^\varepsilon \alpha^2$. Assuming $R$ is sufficiently large depending on $\varepsilon$,
\[
r N^2 \sim (\log R) \lambda^2 |\{x \in B_{R^r r^4} : |F(x)| \sim \lambda\}| \lesssim C_\varepsilon R^\varepsilon (N^2/r) |\{x \in B_{R^r r^4} : |F(x)| \sim \lambda\}|,
\]
so $|\{x \in B_{R^r r^4} : |F(x)| \sim \lambda\}| \gtrsim C_\varepsilon^{-1} R^{-\varepsilon} r^2$ and $\lambda^2 \gtrsim C_\varepsilon^{-1} R^{-\varepsilon} N^2/r \sim C_\varepsilon^{-1} R^{-\varepsilon} \alpha^2$.

Case 2: $R^3 < \alpha^2 \leq R$. Let $q, a,$ and $b$ be integers satisfying
\[
q \text{ odd, } 1 \leq b \leq q \leq N^{2/3}, \quad (b, q) = 1, \quad \text{and} \quad 0 \leq a \leq q.
\]
Define the set $M(q, a, b)$ to be
\[
M(q, a, b) := \{(x_1, x_2) \in [0, N] \times [0, N^2] : |x_1 - \frac{a}{q} N| \leq \frac{1}{1010}, \quad |x_2 - \frac{b}{q} N^2| \leq \frac{1}{1010}\}.
\]

**Lemma 6.** For each $(q, a, b) \neq (q', a', b')$, both tuples satisfying (4), $M(q, a, b) \cap M(q', a', b') = \emptyset$.

**Proof.** If $\frac{b}{q} = \frac{b'}{q'}$, then using the relatively prime part of (4), $b = b'$ and $q = q'$. Then we must have $a \neq a'$, meaning that if $x_1$ is the first coordinate of a point in $M(q, a, b) \cap M(q', a', b')$, then
\[
\frac{2}{1010} \geq |x_1 - \frac{a}{q} N| + |x_1 - \frac{a'}{q'} N| \geq \left|\frac{a - a'}{q} N\right| \geq N^{1/3}
\]
which is clearly a contradiction. The alternative is that $\frac{b}{q} \neq \frac{b'}{q'}$ in which case for $x_2$ the second coordinate of a point in $M(q, a, b) \cap M(q', a', b')$,
\[
\frac{2}{1010} \geq |x_2 - \frac{b}{q} N^2| + |x_2 - \frac{b'}{q'} N^2| \geq \left|\frac{b'q - bq}{qq'} N^2\right| \geq \frac{N^2}{qq'} \geq N^{2/3},
\]
which is another contradiction. \hfill \Box

**Lemma 7.** For each $(x_1, x_2) \in M(q, a, b)$, $|F(x_1, x_2)| \sim \frac{N}{q^{2/3}}$, here meaning within a factor of 4.

**Proof.** This follows from Proposition 13.4 in [Dem20]. \hfill \Box

**Proposition 8.** Let $R^3 < \alpha^2 \leq R$ be given. There exists $v \in [0, N^2]$ satisfying
\[
|\{(x_1, x_2) \in [0, R]^2 : |F(x_1, x_2 + v)| \geq \alpha\}| \gtrsim \frac{R^2 N^3}{\alpha^{3/2}}.
\]
Proof. First note that by $N$-periodicity in $x_1$,
\[ |\{(x_1, x_2) \in [0, R]^2 : |F(x_1, x_2 + v)| \geq \alpha\}| \geq \frac{R}{N} |\{(x_1, x_2) \in ([0, N] \times [0, R]) : |F(x_1, x_2 + v)| \geq \alpha\}|. \]
The function $F$ is $N^2$ periodic in $x_2$, but $R < N^2$ so we need to find $v \in [0, N^2]$ making the set in the lower bound above largest.

By Lemma 7, it suffices to count the tuples $(q, a, b)$ satisfying (8), $q \leq N^2/(16\alpha^2)$, and $|\frac{b}{q}N^2 - v| \leq R$, where $v$ is to be determined. Begin by considering the distribution of points $\frac{b}{q}$ in $[0, 1]$, where $1 \leq b \leq q \sim \frac{N^2}{\alpha^2}$, $(b, q) = 1$. As in the proof of Lemma 6 if $\frac{b}{q} \neq \frac{b'}{q'}$, then $|\frac{b}{q} - \frac{b'}{q'}| \gtrsim \frac{1}{\alpha^2}$. Fix $b_0, q_0$ and consider the set $\{\frac{b}{q} : \frac{b}{q} = \frac{b_0}{q_0}, 1 \leq b \leq q \sim \frac{N^2}{\alpha^2}\}$. Let $q_m$ be maximal such that for some $1 \leq b_m \leq q_m \sim \frac{N^2}{\alpha^2}$ and $(b_m, q_m) = 1$, $\frac{b_0}{q_m} = \frac{b_0}{q_0}$. Then $q_0 = q_m - k$ for some integer $k$ and $b_m(q_m - k) = b_0q_m$. Rearrange to get $q_m(1 - \frac{b_0}{b_m}) = k$. Thus $q_0 = q_m\frac{b_0}{b_m} \sim \frac{N^2}{\alpha^2}$, which implies that $\frac{b_0}{b_m} \sim 1$. Conclude that there are $\gtrsim \sum_{q \sim \frac{N^2}{\alpha^2}} \varphi(q)$ many unique points $\frac{b}{q}$ in $[0, 1]$ satisfying our prescribed conditions for $\varphi$ denoting the Euler totient function. Use Theorem 3.7 in \textit{Apo76} to estimate $\sum_{q \sim \frac{N^2}{\alpha^2}} \varphi(q) \sim \frac{N^4}{\alpha^4}$, as long as $N/\alpha$ is larger than some absolute constant. By the pigeonhole principle, there exists some $R/N^2$ interval $I \subset [0, 1]$ containing $\sim \left\lceil \frac{N^4}{\alpha^4}R^2/N^2 \right\rceil$ many points $\frac{b}{q}$ with $1 \leq b \leq q \sim \frac{N^2}{\alpha^2}$, $(b, q) = 1$. There are also $\sim \frac{N^2}{\alpha^2}$ many choices for $a$ to complete the tuple $(q, a, b)$ satisfying (8). Let $c$ denote the center of $I$ and take $v = cN^2$ in the proposition statement and conclude that
\[ |\{(x_1, x_2) \in ([0, N] \times [0, R]) : |F(x_1, x_2 + v)| \geq \alpha\}| \gtrsim \frac{RN^4}{\alpha^6} \]
to finish the proof. \hfill \Box

Note that Proposition 8 shows the sharpness of Theorem 3 in the range $R^\beta < \alpha \leq R$ since
\[ \frac{R^{2\beta}}{\alpha^6} \sum_{\gamma} \|F_\gamma\|_2^2 \sim \frac{R^{2\beta}}{\alpha^6} R^\beta R^2 = \frac{N^3 R^2}{\alpha^6}. \]
The sharpness of the trivial estimate $|U_\alpha \cap [-R, R]^2| \lesssim R^2$ in the range $\alpha^2 < R^\beta$ follows from Case 2 since for $\alpha^2 < R^\beta$,
\[ |U_\alpha \cap [-R, R]^2| \geq |U_{R^{\beta/2}} \cap [-R, R]^2| \gtrsim \frac{R^{2\beta}}{(R^{\beta/2})^6} \sum_{\gamma} \|F_\gamma\|_2^2 \sim R^2. \]

3. Implications of Theorem 4

Proof of Theorem 3 from Theorem 4. First suppose that $\alpha^2 > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}$. Then
\[
\max_s \lambda(s^{-1}R^{-1})\lambda(s) \lesssim \max_s (s^{-1}R^{-1}R^3)(sR^2) = R^{2\beta-1}
\]
\[ \lesssim \begin{cases} R^{2\beta-1} & \text{if } \alpha^2 > R \\ \frac{R^{2\beta}}{\alpha^2} & \text{if } R^\beta \leq \alpha^2 \leq R \end{cases} \]

Now suppose that $\alpha^2 \leq \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}$. Then
\[ \frac{\lambda(1)^2}{\alpha^2} \lesssim \begin{cases} R^{2\beta-1} & \text{if } \alpha^2 > R \\ \frac{R^{2\beta}}{\alpha^2} & \text{if } R^\beta \leq \alpha^2 \leq R \end{cases} \]
Proof of Corollary 5 from Theorem 4. To see how this corollary follows from Theorem 4, first use an analogous series of pigeonholing steps as in Section 5 of [GMW20] to reduce to the case where \( \|f_r\|_\infty \lesssim 1 \) for all \( \gamma \) and there exists \( C > 0 \) so that \( \|f_r\|_p^p \) is either 0 or comparable to \( C \) for all \( \gamma \). Split the integral

\[
\int |f|^p = \sum_{R^{-1000} < \alpha \leq R^3} \int_{U_\alpha} |f|^p + \int_{|f|<R^{-1000}} |f|^p
\]

where \( U_\alpha = \{ x : |f(x)| \sim \alpha \} \) and assume via dyadic pigeonholing that

\[
\int |f|^p \lesssim \alpha^p |U_\alpha|
\]

(ignoring the case that the set where \( |f| \leq R^{-1000} \) dominates the integral which may be handled trivially). The result of all of the pigeonholing steps is that the statement of Corollary 5 follows from showing that

\[
\alpha^p |U_\alpha| \leq \frac{C_p}{\lambda} R^\varepsilon (R^\beta (1+\frac{1}{\varepsilon}) - (1+\beta)) \lambda (1)^{\frac{\varepsilon}{\beta}-1} \sum_{\gamma} \|f_\gamma\|_p^2
\]

where \( f \) satisfies the hypotheses of Theorem 4. The full range \( \frac{2}{\varepsilon} + \frac{1}{\beta} \leq 1 \) follows from \( p \) in the critical range \( 4 \leq p \leq 6 \), which we treat first.

4 \leq p \leq 6: There are two cases depending on which upper bound is larger in Theorem 4. First we assume the \( L^4 \) bound holds, in which case

\[
\alpha^p |U_\alpha| \leq C_p R^\varepsilon \alpha^{p-4} \max\lambda(s^{-1}R^{-1}) \lambda(s) \sum_{\gamma} \|f_\gamma\|_p^2
\]

\[
\sim C_p R^\varepsilon \frac{\alpha^{p-4}}{\lambda(1)^{\frac{\varepsilon}{\beta}-1}} \max\lambda(s^{-1}R^{-1}) \lambda(s) \left( \sum_{\gamma} \|f_\gamma\|_p^2 \right)^{\frac{\varepsilon}{\beta}}
\]

\[
\leq C_p R^\varepsilon \lambda(1)^{p-4} \max(s^{-1}R^{-1})(R^\beta s)(\sum_{\gamma} \|f_\gamma\|_p^q)^{\frac{\varepsilon}{\beta}}
\]

\[
\leq C_p R^\varepsilon \lambda(1)^{p(1-\frac{1}{\beta})-3} R^{2\beta-1} \left( \sum_{\gamma} \|f_\gamma\|_p^q \right)^{\frac{\varepsilon}{\beta}}.
\]

Since \( p(1-\frac{1}{\beta})-3 \geq 0 \), we may use the bound \( \lambda(1) \lesssim R^\beta \) to conclude that

\[
\lambda(1)^{p(1-\frac{1}{\beta})-3} R^{2\beta-1} \leq R^{3p(1-\frac{1}{\beta})-3+2\beta-1} = R^{3p(1-\frac{1}{\beta})-(1+\beta)}.
\]

The other case is that the \( L^6 \) bound holds in Theorem 4. We may also assume that \( \alpha^2 > \lambda(1) \) since otherwise we trivially have

\[
\alpha^p |U_\alpha| \leq \lambda(1)^{\frac{\varepsilon}{\beta}-1} \sum_{\gamma} \|f_\gamma\|_p^2 \sim \lambda(1)^{\frac{\varepsilon}{\beta}-1+\frac{\varepsilon}{\beta}} \left( \sum_{\gamma} \|f_\gamma\|_p^q \right)^{\frac{\varepsilon}{\beta}} \leq R^{\beta p(\frac{1}{\beta}-\frac{1}{\varepsilon})} \left( \sum_{\gamma} \|f_\gamma\|_p^q \right)^{\frac{\varepsilon}{\beta}}
\]
Suppose that $q \geq 2$ since $4 \leq p \leq 6$ and $\frac{3}{p} + \frac{1}{q} \leq 1$. Now using the assumptions $\alpha^2 \leq \lambda(1)$ and $p \leq 6$, we have

$$
\alpha^p |U_\alpha| \leq C_\varepsilon R^6 \alpha^{p-6} \lambda(1)^2 \lambda(1)^{\frac{1}{1-q}} \left( \sum_{\gamma} \|f_\gamma\|_p^{\frac{q}{q-1}} \right)^{\frac{p}{q}} \\
\sim C_\varepsilon R^6 \lambda(1)^{\beta (\frac{1}{2} - \frac{1}{q})} \left( \sum_{\gamma} \|f_\gamma\|_p^{\frac{q}{q-1}} \right)^{\frac{p}{q}} \lesssim C_\varepsilon R^6 R^{\beta (\frac{1}{2} - \frac{1}{q})} \left( \sum_{\gamma} \|f_\gamma\|_p^{\frac{q}{q-1}} \right)^{\frac{p}{q}}.
$$

**3 \leq p < 4:** Suppose that $\alpha < R^{\beta/2}$. Then using $L^2$-orthogonality,

$$
\alpha^p |U_\alpha| \leq R^6 \alpha^{(p-2)} \sum_{\gamma} \|f_\gamma\|_2^2 \sim R^6 \alpha^{(p-2)} \lambda(1)^{\frac{1}{1-q}} \left( \sum_{\gamma} \|f_\gamma\|_p^{\frac{q}{q-1}} \right)^{\frac{p}{q}}.
$$

Since in this subcase, $1 - \frac{p}{q} \geq 1 - (p-3) > 0$, we are done after noting that $R^{\beta (p-2)} \lambda(1)^{1 - \frac{p}{q}} \leq R^{\beta (\frac{1}{2} - \frac{1}{q})}$. Now assume that $\alpha \geq R^{\beta/2}$ and use the $p = 4$ case above (noting that $R^{4 \beta (\frac{1}{2} - \frac{1}{q})} \sim R^{\beta (\frac{1}{2} - \frac{1}{q})}$) to get

$$
\alpha^p |U_\alpha| \leq \frac{\alpha^4}{(R^{\beta/2})^{4-p}} |U_\alpha| \leq R^{\beta (4-p)} C_\varepsilon R^6 R^{4 \beta (\frac{1}{2} - \frac{1}{q})} \lambda(1)^{\frac{1}{q}} \left( \sum_{\gamma} \|f_\gamma\|_2^2 \right)^{\frac{p}{q}} \\
\leq C_\varepsilon R^6 R^{\beta p (\frac{1}{2} - \frac{1}{q})} \lambda(1)^{\frac{1}{q}} \left( \sum_{\gamma} \|f_\gamma\|_2^2 \right)^{\frac{p}{q}}.
$$

**6 \leq p:** In this range, we use the trivial bound $\alpha \leq \lambda(1)$ and the $p = 6$ case above (noting that $R^{6 \beta (\frac{1}{2} - \frac{1}{q})} \leq R^{6 \beta (\frac{1}{1-q}) - (1+\beta)}$) to get

$$
\alpha^p |U_\alpha| \leq \lambda(1)^{p-6} \alpha^6 |U_\alpha| \leq \lambda(1)^{p-6} C_\varepsilon R^6 R^{6 \beta (\frac{1}{2} - \frac{1}{q}) - (1+\beta)} \lambda(1)^{\frac{1}{q}} \left( \sum_{\gamma} \|f_\gamma\|_2^2 \right)^{\frac{p}{q}} \\
= \left( \frac{\lambda(1)}{R^\beta} \right)^{(p-6)(\frac{1}{2})} C_\varepsilon R^6 R^{6 \beta (\frac{1}{q}) - (1+\beta)} \lambda(1)^{\frac{1}{q}} \left( \sum_{\gamma} \|f_\gamma\|_2^2 \right)^{\frac{p}{q}} \\
\leq C_\varepsilon R^6 R^{6 \beta (\frac{1}{q}) - (1+\beta)} \lambda(1)^{\frac{1}{q}} \left( \sum_{\gamma} \|f_\gamma\|_2^2 \right)^{\frac{p}{q}}.
$$

\[\Box\]

**4. Tools to prove Theorem 4**

The proof of Theorem 4 follows the high/low frequency decomposition and pruning approach from [GMW20]. In this section, we introduce notation for different scale neighborhoods of $\mathbb{P}^1$, a pruning process for wave packets at various scales, some high/low lemmas which are used to analyze the high/low frequency parts of square functions, and a version of a bilinear restriction theorem for $\mathbb{P}^1$.

Begin by fixing some notation, as above. Let $\beta \in [\frac{1}{2}, 1]$ and $R \geq 2$. The parameter $\alpha > 0$ describes the superlevel set

$$
U_\alpha = \{ x \in \mathbb{R}^2 : |f(x)| \geq \alpha \}.
$$

For $\varepsilon > 0$, we analyze scales $R_k = R^{k\varepsilon}$, noting that $R^{-1/2} \leq R_k^{-1/2} \leq 1$. Let $N$ distinguish the index so that $R_N$ is closest to $R$. Since $R$ and $R_N$ differ at most by a factor of $R^\varepsilon$, we will ignore the distinction between $R_N$ and $R$ in the rest of the argument.
Define the following collections, each of which partitions a neighborhood of $\mathbb{P}$ into approximate rectangles.

1. $\{\gamma\}$ is a partition of $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ by approximate $R^{-\beta} \times R^{-1}$ rectangles, described explicitly in \[5\].
2. $\{\theta\}$ is a partition of $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ by approximate $R^{-1/2} \times R^{-1}$ rectangles. In particular, let each $\theta$ be a union of adjacent $\gamma$.
3. $\{\tau_k\}$ is a partition of $\mathcal{N}_{R_k^{-1}}(\mathbb{P}^1)$ by approximate $R_k^{-1/2} \times R_k^{-1}$ rectangles. Assume the additional property that $\gamma \cap \tau_k = \emptyset$ or $\gamma \subset \tau_k$.

4.1. A pruning step. We will define wave packets at each scale $\tau_k$, and prune the wave packets associated to $f_{\tau_k}$ according to their amplitudes.

For each $\tau_k$, fix a dual rectangle $\tau_k^*$ which is a $2R_k^{1/2} \times 2R_k$ rectangle centered at the origin and comparable to the convex set

$$\{x \in \mathbb{R}^2 : |x \cdot \xi| \leq 1 \quad \forall \xi \in \tau_k\}.$$ Let $T_{\tau_k}$ be the collection of tubes $T_{\tau_k}$ which are dual to $\tau_k$, contain $\tau_k^*$, and which tile $\mathbb{R}^2$. Next, we will define an associated partition of unity $\psi_{T_{\tau_k}}$. First let $\varphi(\xi)$ be a bump function supported in $[-\frac{1}{4}, \frac{1}{4}]^2$. For each $m \in \mathbb{Z}^2$, let

$$\psi_m(x) = c \int_{[-\frac{1}{4}, \frac{1}{4}]^2} |\varphi|^2(x - y - m) dy,$$

where $c$ is chosen so that $\sum_{m \in \mathbb{Z}^2} \psi_m(x) = c \int_{\mathbb{R}^2} |\varphi|^2 = 1$. Since $|\varphi|$ is a rapidly decaying function, for any $n \in \mathbb{N}$, there exists $C_n > 0$ such that

$$\psi_m(x) \leq c \int_{[0,1]^2} \frac{C_n}{(1 + |x - y - m|^2)^n} dy \leq \frac{\tilde{C}_n}{(1 + |x - m|^2)^n}.$$ Define the partition of unity $\psi_{T_{\tau_k}}$ associated to $\tau_k$ to be $\psi_{T_{\tau_k}}(x) = \psi_m \circ A_{\tau_k}$, where $A_{\tau_k}$ is a linear transformation taking $\tau_k^*$ to $[-\frac{1}{4}, \frac{1}{4}]^2$ and $A_{\tau_k}(T_{\tau_k}) = m + [-\frac{1}{4}, \frac{1}{4}]^2$. The important properties of $\psi_{T_{\tau_k}}$ are (1) rapid decay off of $T_{\tau_k}$ and (2) Fourier support contained in $\tau_k$.

To prove upper bounds for the size of $U_k$, we will actually bound the sizes of $\sim \varepsilon^{-1}$ many subsets which will be denoted $U_{\alpha} \cap \Omega_k$, $U_{\alpha} \cap H$, and $U_{\alpha} \cap L$. The pruning process sorts between important and unimportant wave packets on each of these subsets, as described in Lemma \[12\] below.

Partition $T_{\theta} = T_{\theta}^g \sqcup T_{\theta}^b$ into a “good” and a “bad” set as follows. Let $\delta > 0$ be a parameter to be chosen in \S 5.2 and set

$$T_{\theta} \in T_{\theta}^g \quad \text{if} \quad \|\psi_{T_{\theta}}f_{\theta}\|_{L^\infty(\mathbb{R}^2)} \leq R^M \tilde{\delta} \frac{\lambda(1)}{\alpha},$$

where $M > 0$ is a universal constant we will choose in the proof of Proposition \[4\].

**Definition 1** (Pruning with respect to $\tau_k$). For each $\theta$ and $\tau_{N-1}$, define the notation $f_{\theta}^N = \sum_{T_{\theta} \in T_{\theta}^g} \psi_{T_{\theta}}f_{\theta}$ and $f_{\tau_{N-1}}^N = \sum_{\theta \subset \tau_{N-1}} f_{\theta}^N$. For each $k < N$, let

$$T_{\tau_k}^g = \{T_{\tau_k} \in T_{\tau_k} : \|\psi_{T_{\tau_k}} f_{\tau_k}^{k+1}\|_{L^\infty(\mathbb{R}^2)} \leq R^M \tilde{\delta} \frac{\lambda(1)}{\alpha}\},$$

$$f_{\tau_k}^k = \sum_{T_{\tau_k} \in T_{\tau_k}^g} \psi_{T_{\tau_k}} f_{\tau_k}^{k+1} \quad \text{and} \quad f_{\tau_{k-1}}^k = \sum_{\tau_k \subset \tau_{k-1}} f_{\tau_k}^k.$$
For each $k$, define the $k$th version of $f$ to be $f^k = \sum_{\tau_k} f^k_{\tau_k}$.

**Lemma 9 (Properties of $f^k$).**

1. $|f^k_{\tau_k}(x)| \leq |f^k_{\tau_k}(x)| \leq \# \gamma \subset \tau_k$.
2. $\|f^k_{\tau_k}\|_{L^\infty} \leq C_{\gamma} R^{O(\epsilon)} R^{M \delta \lambda(1)}$.
3. $\text{supp} f^k_{\tau_k} \subset 2\tau_k$.
4. $\text{supp} f^k_{\tau_k-1} \subset (1 + (\log R)^{-1})\tau_{k-1}$.

**Proof.** The first property follows because $\sum_{\tau_k \in \mathbb{T}_{\tau_k}} \psi_{\tau_k} f^k_{\tau_k}$ is a partition of unity, and

$$f^k_{\tau_k} = \sum_{\tau_k \in \mathbb{T}_{\tau_k}} \psi_{\tau_k} f^k_{\tau_k}.$$ 

Furthermore, by definition of $f^{k+1}_{\tau_k}$ and iterating, we have

$$|f^k_{\tau_k}| \leq |f^{k+1}_{\tau_k}| \leq \sum_{\tau_k+1 \subset \tau_k} |f^{k+1}_{\tau_k}| \leq \cdots \leq \sum_{\tau_N \subset \tau_k} |f^N_{\tau_N}|$$

$$\leq \sum_{\theta \subset \tau_k} |f_{\theta}| \leq \sum_{\gamma \subset \tau_k} |f_{\gamma}| \approx \# \gamma \subset \tau_k$$

where we used the assumption $\|f_{\gamma}\|_\infty \lesssim 1$ for all $\gamma$. Now consider the $L^\infty$ bound in the second property. We write

$$f^k_{\tau_k}(x) = \sum_{\tau_k \in \mathbb{T}_{\tau_k}} \psi_{\tau_k} f^{k+1}_{\tau_k} + \sum_{\tau_k \in \mathbb{T}_{\tau_k}} \psi_{\tau_k} f^{k+1, \tau_k}_{\tau_k}.$$ 

The first sum has at most $C R^{2\epsilon}$ terms, and each term has norm bounded by $R^{M \delta \lambda(1)}$ by the definition of $\mathbb{T}^h_{\tau_k}$. By the first property, we may trivially bound $f^{k+1}_{\tau_k}$ by $R \max_{\gamma} \|f_{\gamma}\|_\infty$.

But if $x \notin R^c \mathbb{T}_{\tau_k}$, then $\psi_{\tau_k} f^k_{\tau_k}(x) \leq R^{-1000}$. Thus

$$|\sum_{\tau_k \in \mathbb{T}_{\tau_k}} \psi_{\tau_k} f^{k+1}_{\tau_k}| \leq \sum_{\tau_k \in \mathbb{T}_{\tau_k}} R^{-500} \psi_{\tau_k}^{1/2}(x) \|f^{k+1}_{\tau_k}\|_\infty \leq R^{-250} \max_{\gamma} \|f_{\gamma}\|_\infty.$$ 

Since $\alpha \lesssim |f(x)| \lesssim \sum_{\gamma} \|f_{\gamma}\|_\infty \approx \lambda(1)$, (recalling the assumption that each $\|f_{\gamma}\|_\infty \lesssim 1$), we note $R^{-250} \leq C R^{2\theta \lambda(1)}$.

The fourth and fifth properties depend on the Fourier support of $\psi_{\tau_k}$, which is contained in $\frac{1}{2} \tau_k$. Initiate a 2-step induction with base case $k = N$: $f^N_{\theta}$ has Fourier support in $2\theta$ because of the above definition. Then

$$f^N_{\tau_{N-1}} = \sum_{\theta \subset \tau_{N-1}} f^N_{\theta}$$

has Fourier support in $\bigcup_{\theta \subset \tau_{N-1}} 2\theta$, which is contained in $(1 + (\log R)^{-1})\tau_{N-1}$. Since each $\psi_{\tau_{N-1}}$ has Fourier support in $\frac{1}{2} \tau_{N-1}$,

$$f^{N-1}_{\tau_{N-1}} = \sum_{\tau_{N-1} \in \mathbb{T}_{\tau_{N-1}}} \psi_{\tau_{N-1}} f^N_{\tau_{N-1}}$$

has Fourier support in $\frac{1}{2} \tau_{N-1} + (1 + (\log R)^{-1})\tau_{N-1} \subset 2\tau_{N-1}$. Iterating this reasoning until $k = 1$ gives (3) and (4).
Definition 2. For each $\tau_k$, let $w_{\tau_k}$ be the weight function adapted to $\tau_k^*$ defined by

$$w_{\tau_k}(x) = w_k \circ R_{\tau_k}(x)$$

where

$$w_k(x, y) = \frac{e}{(1 + \frac{|x|^2}{R_k^2})^{10} (1 + \frac{|y|^2}{R_k^2})^{10}}, \quad \|w\|_1 = 1,$$

and $R_{\tau_k} : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation taking $\tau_k^*$ to $[-R_k^{1/2}, R_k^{1/2}] \times [-R_k, R_k]$. For each $T_{\tau_k} \in T_{\tau_k}$, let $w_{T_{\tau_k}} = w_{\tau_k}(x - c T_{\tau_k})$ where $c T_{\tau_k}$ is the center of $T_{\tau_k}$. For $s > 0$, we also use the notation $w_s$ to mean

$$w_s(x) = \frac{e'}{(1 + |x|^2/s^2)^{10}}, \quad \|w_s\|_1 = 1.$$

The weights $w_{\tau_k}$, $w_\theta = w_{\tau_N}$, and $w_s$ are useful when we invoke the locally constant property. By locally constant property, we mean generally that if a function $f$ has Fourier transform supported in a convex set $A$, then for a bump function $\varphi_A \equiv 1$ on $A$, $\hat{f} = f \ast \widehat{\varphi_A}$. Since $|\widehat{\varphi_A}|$ is an $L^1$-normalized function which is positive on a set dual to $A$, $|\hat{f}| \ast |\widehat{\varphi_A}|$ is an averaged version of $|\hat{f}|$ over a dual set $A^\ast$. We record some of the specific locally constant properties we need in the following lemma.

Lemma 10 (Locally constant property). For each $\tau_k$ and $T_{\tau_k} \in T_{\tau_k}$,

$$\|f_{T_{\tau_k}}\|_{L^\infty(T_{\tau_k})} \lesssim |f_{T_{\tau_k}}| \ast w_{\tau_k}(x) \quad \text{for any} \quad x \in T_{\tau_k}.$$

For any collection of $\sim s^{-1} \times s^{-2}$ blocks $\theta_s$ partitioning $N_{s^{-2}}(\mathbb{P}^1)$ and any $s$-ball $B$,

$$\|\sum_{\theta_s} |f_{\theta_s}|^2\|_{L^\infty(B)} \lesssim \sum_{\theta_s} |f_{\theta_s}|^2 \ast w_s(x) \quad \text{for any} \quad x \in B.$$

Because the pruned versions of $f$ and $f_{T_{\tau_k}}$ have essentially the same Fourier supports as the unpruned versions, the locally constant lemma applies to the pruned versions as well.

Proof of Lemma 10. Let $\rho_{T_{\tau_k}}$ be a bump function equal to 1 on $\tau_k$ and supported in $2\tau_k$. Then using Fourier inversion and Hölder’s inequality,

$$|f_{T_{\tau_k}}(y)|^2 = |f_{T_{\tau_k}} \ast \rho_{T_{\tau_k}}(y)|^2 \leq \|\rho_{T_{\tau_k}}\|_1 |f_{T_{\tau_k}}|^2 \ast \rho_{T_{\tau_k}}(y).$$

Since $\rho_{T_{\tau_k}}$ may be taken to be an affine transformation of a standard bump function adapted to the unit ball, $\|\rho_{T_{\tau_k}}\|_1$ is a constant. The function $\rho_{T_{\tau_k}}$ decays rapidly off of $\tau_k^*$, so $|\rho_{T_{\tau_k}}| \lesssim w_{\tau_k}$. Since for any $T_{\tau_k} \in T_{\tau_k}$, $w_{\tau_k}(y)$ is comparable for all $y \in T_{\tau_k}$, we have

$$\sup_{x \in T_{\tau_k}} |f_{T_{\tau_k}}|^2 \ast w_{\tau_k}(x) \leq \int |f_{T_{\tau_k}}|^2(y) \sup_{x \in T_{\tau_k}} w_{\tau_k}(x - y) dy \sim \int |f_{T_{\tau_k}}|^2(y) w_{\tau_k}(x - y) dy \quad \text{for all} \quad x \in T_{\tau_k}.$$

For the second part of the lemma, repeat analogous steps as above, except begin with $\rho_{\theta_s}$ which is identically 1 on a ball of radius $2s^{-1}$ containing $\theta_s$. Then

$$\sum_{\theta_s} |f_{\theta_s}(y)|^2 = \sum_{\theta_s} |f_{\theta_s} \ast \rho_{\theta_s}(y)|^2 \lesssim \sum_{\theta_s} |f_{\theta_s}|^2 \ast |\rho_{\theta_s^{-1}}|(y),$$

where we used that each $\rho_{\theta_s}$ is a translate of a single function $\rho_{s^{-1}}$. The rest of the argument is analogous to the first part. \qed
**Definition 3** (Auxiliary functions). Let $\varphi(x) : \mathbb{R}^2 \to [0, \infty)$ be a radial, smooth bump function satisfying $\varphi(x) = 1$ on $B_1$ and $\text{supp} \varphi \subset B_2$.

$$\varphi(2^{J+1} \xi) + \sum_{j=-2}^{J} [\varphi(2^j \xi) - \varphi(2^{j+1} \xi)]$$

where $J$ is defined by $2^J \leq \lceil R^\beta \rceil < 2^{J+1}$. Then for each dyadic $s = 2^j$, let

$$\eta_{=s}(\xi) = \varphi(2^j \xi) - \varphi(2^{j+1} \xi)$$

and let

$$\eta_{< \lceil R^\beta \rceil - 1}(\xi) = \varphi(2^{j+1} \xi).$$

Finally, for $k = 1, \ldots, N - 1$, define

$$\eta_k(\xi) = \varphi(R_k^{1/2} \xi).$$

**Definition 4.** Let $G(x) = \sum_\theta |f_\theta|^2 * w_\theta$, $G^h(x) = G * \tilde{\eta}_{< \lceil R^\beta \rceil - 1}$, $G^h(x) = G(x) - G^h(x)$. For $k = 1, \ldots, N - 1$, let

$$g_k(x) = \sum_{\tau_k} |f_{\tau_k}^{k+1}|^2 * w_{\tau_k}, \quad g_k^\ell(x) = g_k * \tilde{\eta}_k, \quad \text{and} \quad g_k^\ell(x) = g_k - g_k^\ell.$$

**Definition 5.** Define the high set

$$H = \{ x \in B_R : G(x) \leq 2|G^h(x)| \}.$$

For each $k = 1, \ldots, N - 1$, let

$$\Omega_k = \{ x \in B_R \setminus H : g_k \leq 2|g_k^\ell|, g_k + 1 \leq 2|g_k^{\ell+1}|, \ldots, g_N \leq 2|g_N^\ell| \}$$

and for each $k = 1, \ldots, N$. Define the low set

$$L = \{ x \in B_R \setminus H : g_1 \leq 2|g_1^\ell|, \ldots, g_N \leq 2|g_N^\ell|, G(x) \leq 2|G^\ell(x)| \}.$$

### 4.2. High/low frequency lemmas.

**Lemma 11** (Low lemma). For each $x$, $|G^\ell(x)| \lesssim \lambda(1)$ and $|g_k^\ell(x)| \lesssim g_{k+1}(x)$.

**Proof.** For each $\theta$, by Plancherel’s theorem,

$$|f_\theta|^2 * \tilde{\eta}_{< \lceil R^\beta \rceil - 1}(x) = \int_{\mathbb{R}^2} |f_\theta|^2(\gamma - \xi) \tilde{\eta}_{< \lceil R^\beta \rceil - 1}(\xi)d\xi$$

$$= \int_{\mathbb{R}^2} \hat{f}_\theta * \hat{f}_\theta(\xi)e^{-2\pi i x \cdot \xi} \tilde{\eta}_{< \lceil R^\beta \rceil - 1}(\xi)d\xi$$

$$= \sum_{\gamma, \gamma' \subset \theta} \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} \hat{f}_\gamma * \hat{f}_\gamma(\xi) \tilde{\eta}_{< \lceil R^\beta \rceil - 1}(\xi)d\xi.$$

The integrand is supported in $(\gamma \setminus \gamma') \cap B_{2\lceil R^\beta \rceil - 1}$. This means that the integral vanishes unless $\gamma$ is within $CR^{-\beta}$ of $\gamma'$ for some constant $C > 0$, in which case we write $\gamma \sim \gamma'$. Then

$$\sum_{\gamma, \gamma' \subset \theta} \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} \hat{f}_\gamma * \hat{f}_\gamma(\xi) \tilde{\eta}_{< \lceil R^\beta \rceil - 1}(\xi)d\xi = \sum_{\gamma, \gamma' \subset \theta} \int_{\gamma \sim \gamma'} \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} \hat{f}_\gamma * \hat{f}_\gamma(\xi) \tilde{\eta}_{< \lceil R^\beta \rceil - 1}(\xi)d\xi.$$
Use Plancherel’s theorem again to get back to a convolution in $x$ and conclude that

$$|G * \tilde{\eta}_{<|R^3|^{-1}}(x)| = \left| \sum_{\theta} \sum_{\gamma, \gamma' \in \theta} (f_{\gamma} \overline{f_{\gamma'}}) * w_{\theta} * \tilde{\eta}_{<|R^3|^{-1}}(x) \right|$$

$$\lesssim \sum_{\theta} \sum_{\gamma \in \theta} |f_{\gamma}|^2 * w_{\theta} * |\tilde{\eta}_{<|R^3|^{-1}}|(x) \lesssim \sum_{\gamma} \|f_{\gamma}\|_\infty \lesssim \lambda(1).$$

By an analogous argument as above, we have that

$$|g_k^k(x)| \lesssim \sum_{\tau_{k+1}} |f_{\tau_{k+1}}|^2 * w_{\tau_k} * |\tilde{\eta}_k|(x)$$

where for each summand, $w_{\tau_k}$ corresponds to the $\tau_k$ containing $\tau_{k+1}$. By definition, $|f_{\tau_{k+1}}|^2 \lesssim |f_{\tau_{k+1}}|$. By the locally constant property, $|f_{\tau_{k+1}}|^2 \lesssim |f_{\tau_{k+1}}|^2 * w_{\tau_{k+1}}$. It remains to note that

$$w_{\tau_{k+1}} * |\tilde{\eta}_k|(x) \lesssim w_{\tau_{k+1}}(x)$$

since $\tau_{k}^* \subset \tau_{k+1}^*$ and $\tilde{\eta}_k$ is an $L^1$-normalized function that is rapidly decaying away from $B_{R_{\tau_{k+1}}}(0)$.

\[ \square \]

**Lemma 12** (Pruning lemma). For any $\tau$,

$$\left| \sum_{\tau_k \subset \tau} f_{\tau_k} - \sum_{\tau_k \subset \tau} f_{\tau_{k+1}}(x) \right| \leq C \varepsilon R^{-M\delta} \alpha \quad \text{for all } x \in \Omega_k$$

and

$$\left| \sum_{\tau_\lambda \subset \tau} f_{\tau_\lambda} - \sum_{\tau_\lambda \subset \tau} f_{\tau_\lambda}(x) \right| \leq C \varepsilon R^{-M\delta} \alpha \quad \text{for all } x \in L.$$

**Proof.** By the definition of the pruning process, we have

$$f_\tau = f_N^\tau + (f_\tau - f_N^\tau) = \cdots = f_{k+1}^\tau(x) + \sum_{m=k+1}^N (f_{m+1}^\tau - f_m^\tau)$$

with the understanding that $f_{N+1}^\tau = f$ and formally, the subscript $\tau$ means $f_\tau = \sum_{\gamma \subset \tau} f_\gamma$ and $f_{m}^\tau = \sum_{\tau_m \subset \tau} f_{\tau_m}^m$. We will show that each difference in the sum is much smaller than $\alpha$. For each $m \geq k + 1$ and $\tau_m$,

$$|f_{m}^\tau(x) - f_{m+1}^\tau(x)| = \sum_{\tau_{m} \subset \tau} \psi_{\tau_{m}}(x) \overline{f_{\tau_{m}}^m(x)} = \sum_{\tau_{m} \subset \tau} \overline{\psi_{\tau_{m}}^{1/2}(x)} f_{\tau_{m}}^m(x) \psi_{\tau_{m}}^{1/2}(x)$$

$$\lesssim \sum_{\tau_{m} \subset \tau} R^{-M\delta} \frac{\alpha}{\lambda(1)} \psi_{\tau_{m}}^{1/2}(x) \overline{f_{\tau_{m}}^m(x)} \|\psi_{\tau_{m}}^{1/2}(x)\|_{L^\infty(R^2)} \|\overline{\psi_{\tau_{m}}^{1/2}(x)}\|_{L^\infty(R^2)}$$

$$\lesssim R^{-M\delta} \frac{\alpha}{\lambda(1)} \sum_{\tau_{m} \subset \tau} \overline{\psi_{\tau_{m}}^{1/2}(x)} \|f_{\tau_{m}}^m(x)\|_{L^\infty(R^2)} \psi_{\tau_{m}}^{1/2}(x)$$

$$\lesssim R^{-M\delta} \frac{\alpha}{\lambda(1)} \sum_{\tau_{m} \subset \tau} \sum_{\tau_{m}' \subset \tau_{m}} \|\psi_{\tau_{m}}(x)\|_{L^\infty(T_{\tau_{m}})} \|f_{\tau_{m}}^m(x)\|_{L^\infty(T_{\tau_{m}})} \psi_{\tau_{m}}^{1/2}(x)$$

$$\lesssim R^{-M\delta} \frac{\alpha}{\lambda(1)} \sum_{\tau_{m} \subset \tau} \sum_{\tau_{m}' \subset \tau_{m}} \|\psi_{\tau_{m}}(x)\|_{L^\infty(T_{\tau_{m}})} \|f_{\tau_{m}}^{m+1}(x)\|_{L^\infty(T_{\tau_{m}})} \psi_{\tau_{m}}^{1/2}(x).$$
Let $c_{\tilde{T}_m}$ denote the center of $\tilde{T}_m$ and note the pointwise inequality
\[ \sum_{T_m} \| \psi_{T_m} \|_{L^\infty(\tilde{T}_m)} \psi_{T_m}^{1/2}(x) \lesssim R_m^{3/2} w_{\tau_m}(x - c_{\tilde{T}_m}), \]
which means that
\[ |f_{\tau_m}^m(x) - f_{\tau_m}^{m+1}(x)| \lesssim R_m^{-\delta} \frac{\alpha}{\lambda(1)} R_m^{3/2} \sum_{\tilde{T}_m \in T_m} w_{\tau_m}(x - c_{\tilde{T}_m}) \| f_{\tau_m}^{m+1} \|_{L^\infty(\tilde{T}_m)} ^2 \]
\[ \lesssim R_m^{-\delta} \frac{\alpha}{\lambda(1)} R_m^{3/2} \sum_{\tilde{T}_m \in T_m} w_{\tau_m}(x - c_{\tilde{T}_m}) \| f_{\tau_m}^{m+1} \|_{L^\infty(\tilde{T}_m)} ^2 \]
\[ \lesssim R_m^{-\delta} \frac{\alpha}{\lambda(1)} |f_{\tau_m}^{m+1}|^2 * w_{\tau_m}(x). \]
where we used the locally constant property in the second to last inequality and the pointwise relation $w_{\tau_m} * w_{\tau_m} \lesssim w_{\tau_m}$ for the final inequality. Then
\[ | \sum_{\tau_m \subset \tau} f_{\tau_m}^m(x) - f_{\tau_m}^{m+1}(x) | \lesssim R_m^{-\delta} \frac{\alpha}{\lambda(1)} \sum_{\tau_m \subset \tau} |f_{\tau_m}^{m+1}|^2 * w_{\tau_m}(x) \lesssim R_m^{-\delta} \frac{\alpha}{\lambda(1)} g_m(x). \]
By the definition of $\Omega_k$ and Lemma 11, $g_m(x) \leq 2 |g_m^\ell(x)| \leq 2 C g_{m+1}(x) \leq \cdots \leq (2C)^{-2} G(x) \lesssim (2C)^{\varepsilon-1} r$. Conclude that
\[ | \sum_{\tau_m \subset \tau} f_{\tau_m}^m(x) - f_{\tau_m}^{m+1}(x) | \lesssim (2C)^{\varepsilon-1} R_m^{-\delta} \alpha. \]

The claim for $L$ follows immediately from the above argument, using the low-dominance of $g_k$ for all $k$.

**Definition 6.** Call the distribution function $\lambda$ associated to a function $f$ $(R, \varepsilon)$-normalized if for any $\tau_k, \tau_m$,
\[ \# \{ \tau_k \subset \tau_m : f_{\tau_k} \neq 0 \} \leq 100 \frac{\lambda(R_m^{-1/2})}{\lambda(R_k^{-1/2})}. \]

**Lemma 13** (High lemma I). Assume that $f$ has an $(R, \varepsilon)$-normalized distribution function $\lambda(\cdot)$. For each dyadic $s$, $R^{-\beta} \leq s \leq R^{-1/2}$,
\[ \int_{\mathbb{R}^2} |G * \tilde{\eta}_s|^2 \lesssim C \varepsilon R^{2\varepsilon} \lambda(s^{-1} R^{-1}) \lambda(s) \sum_\gamma \| f_\gamma \|^2_2. \]

**Proof.** Organize the $\{\gamma\}$ into subcollections $\{\theta_s\}$ in which each $\theta_s$ is a union of $\gamma$ which intersect the same $\sim s$-arc of $\mathbb{R}^1$, where here for concreteness, $\sim s$ means within a factor of 2. Then by Plancherel’s theorem, since $\tilde{\eta}_s \equiv \tilde{\eta}_s$, we have for each $\theta$
\[ |f_\theta|^2 * \tilde{\eta}_s(x) = \int_{\mathbb{R}^2} |f_\theta|^2(x - y) \tilde{\eta}_s(y) dy = \int_{\mathbb{R}^2} \tilde{f}_\theta * \tilde{\eta}_s(x) e^{-2\pi i x \cdot \xi} \eta_s(\xi) d\xi \]
\[ = \sum_{\theta_s} \int_{\mathbb{R}^2} e^{-2\pi i x \cdot \xi} \tilde{f}_\theta(x) \eta_s(\xi) d\xi. \]

(8)
The support of \( \hat{f}_{\theta} (\xi) = \int e^{2\pi i x \cdot \xi} f_{\theta} (x) dx = \hat{f}_{\theta} (\xi) \) is contained in \(-\theta^*\). This means that the support of \( \hat{f}_{\theta} * \hat{f}_{\theta'} (\xi) \) is contained in \( \theta - \theta^* \). Since the support of \( \eta_{\sim s} (\xi) \) is contained in the ball of radius \( 2s \), for each \( \theta_s \subset \theta \), there are only finitely many \( \theta^*_s < \theta \) so that the integral in (3) is nonzero. Thus we may write

\[
G * \tilde{\eta}_{\sim s} (x) = \sum_{\theta} |f_{\theta}|^2 * \hat{w}_\theta * \tilde{\eta}_{\sim s} (x) = \sum_{\theta} \sum_{\theta_s, \theta' \subset \theta} (f_{\theta_s} \hat{f}_{\theta'} *) * w_{\theta} * \tilde{\eta}_{\sim s} (x).
\]

where the second sum is over \( \theta_s, \theta'_s \subset \theta \) with \( \text{dist}(\theta_s, \theta'_s) < 2s \). Using the above pointwise expression and then Plancherel’s theorem, we have

\[
\int_{\mathbb{R}^2} |G * \tilde{\eta}_{\sim s}|^2 = \int_{\mathbb{R}^2} \left| \sum_{\theta} \sum_{\theta_s, \theta'_s \subset \theta} (f_{\theta_s} \hat{f}_{\theta'} *) * w_{\theta} * \tilde{\eta}_{\sim s} \right|^2
\]

\[
= \int_{\mathbb{R}^2} \left| \sum_{\theta} \sum_{\theta_s, \theta'_s \subset \theta} (\hat{f}_{\theta_s} * \hat{f}_{\theta'_s} \hat{w}_\theta \eta_{\sim s}) \right|^2
\]

For each \( \theta, \theta_s, \theta'_s \subset \theta \), \( \hat{f}_{\theta_s} * \hat{f}_{\theta'_s} \) is supported in \( \theta - \theta \), since each summand is supported in \( \theta_s - \theta^* \) and \( \theta_s, \theta'_s \subset \theta \). For each \( \xi \in \mathbb{R}^2, |\xi| > \frac{1}{2} r \), the maximum number of \( \theta - \theta \) containing \( \xi \) is bounded by the maximum number of \( \theta \) intersecting an \( R^{-1/2} \cdot s^{-1} R^{-1/2} \)-arc of the parabola. Using that \( \lambda(\cdot) \) is \((R, \varepsilon)\)-normalized, this number is bounded above by \( C \varepsilon R^2 \frac{\lambda(s^{-1}R^{-1})}{\lambda(R^{-1/2})} \). Since \( \eta_{\sim s} \) is supported in the region \( |\xi| > \frac{1}{2} r \), by Cauchy-Schwarz

\[
\int_{\mathbb{R}^2} \left| \sum_{\theta} \sum_{\theta_s, \theta'_s \subset \theta} (\hat{f}_{\theta_s} * \hat{f}_{\theta'_s} \hat{w}_\theta \eta_{\sim s}) \right|^2 \lesssim C \varepsilon R^2 \frac{\lambda(r^{-1}R^{-1})}{\lambda(R^{-1/2})} \sum_{\theta} \int_{\mathbb{R}^2} \left| \sum_{\theta_s, \theta'_s \subset \theta} (\hat{f}_{\theta_s} * \hat{f}_{\theta'_s} \hat{w}_\theta \eta_{\sim s}) \right|^2
\]

\[
= C \varepsilon R^2 \frac{\lambda(r^{-1}R^{-1})}{\lambda(R^{-1/2})} \sum_{\theta} \int_{\mathbb{R}^2} \left| \sum_{\theta_s, \theta'_s \subset \theta} (f_{\theta_s} \hat{f}_{\theta'_s} *) * w_{\theta} * \tilde{\eta}_{\sim s} \right|^2
\]

\[
\lesssim C \varepsilon R^2 \frac{\lambda(r^{-1}R^{-1})}{\lambda(R^{-1/2})} \sum_{\theta} \int_{\mathbb{R}^2} \left| \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 * w_{\theta} * |\tilde{\eta}_{\sim s}| \right|^2.
\]

It remains to analyze each of the integrals above:

\[
\int_{\mathbb{R}^2} \left| \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 * w_{\theta} * |\tilde{\eta}_{\sim s}| \right|^2 \lesssim \left\| \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 * w_{\theta} * |\tilde{\eta}_{\sim s}| \right\| \| f_{\gamma} \|_\infty \int_{\mathbb{R}^2} \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 * w_{\theta} * |\tilde{\eta}_{\sim s}|.
\]

Bound the \( L^\infty \) norms using the assumption that \( \| f_{\gamma} \|_\infty \lesssim 1 \) for all \( \gamma \):

\[
\| \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 * w_{\theta} * |\tilde{\eta}_{\sim s}| \|_\infty \lesssim \sum_{\theta_s \subset \theta} \| f_{\theta_s} \|_\infty \lesssim \sum_{\theta_s \subset \theta} \sum_{\gamma \subset \theta_s} \| f_{\gamma} \|_\infty \lesssim \lambda(R^{-1/2}) \lambda(s).
\]

Finally, using Young’s convolution inequality and the \( L^2 \)-orthogonality of the \( f_{\gamma} \), we have

\[
\int_{\mathbb{R}^2} \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 * w_{\theta} * |\tilde{\eta}_{\sim s}| \leq \int_{\mathbb{R}^2} \sum_{\theta_s \subset \theta} |f_{\theta_s}|^2 = \sum_{\gamma \subset \theta} \| f_{\gamma} \|_2^2.
\]

\[\square\]
Lemma 14 (High lemma II). For each $k$,

$$\int_{\mathbb{R}^2} |g_k^h|^2 \lesssim R^{3e} \sum_{\tau_k} \int_{\mathbb{R}^2} |f_{\tau_k+1}^{k+1}|^4.$$ 

Proof. By Plancherel’s theorem, we have

$$\int_{\mathbb{R}^2} |g_k^h|^2 = \int_{\mathbb{R}^2} |g_k - g_k'|^2$$

$$= \int_{\mathbb{R}^2} \left| \sum_{\tau_k} (\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}) \hat{w}_{\tau_k} - \sum_{\tau_k} (\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}) \hat{w}_{\tau_k} \eta_k \right|^2$$

$$\leq \int |\xi| > cR_{k+1}^{-1/2} \left| \sum_{\tau_k} (\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}) \hat{w}_{\tau_k} \eta_k \right|^2$$

since $(1 - \eta_k)$ is supported in the region $|\xi| > cR_{k+1}^{-1/2}$ for some constant $c > 0$. For each $\tau_k$, $\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}$ is supported in $2\tau_k - 2\tau_k$, using property (1) of Lemma 9. The maximum overlap of the sets $\{2\tau_k - 2\tau_k\}$ in the region $|\xi| \geq cR_{k+1}^{-1/2}$ is bounded by $\sim \frac{R_{k+1}^{-1/2}}{R_{k+1}} \lesssim R^e$. Thus using Cauchy-Schwarz,

$$\int |\xi| > cR_{k+1}^{-1/2} \left| \sum_{\tau_k} (\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}) \hat{w}_{\tau_k} \right|^2 \lesssim R^e \sum_{\tau_k} \int_{|\xi| > cR_{k+1}^{-1/2}} \left| (\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}) \hat{w}_{\tau_k} \right|^2$$

$$\leq R^e \sum_{\tau_k} \int_{\mathbb{R}^2} \left| (\hat{f}_{\tau_k}^{k+1} \ast \hat{f}_{\tau_k}^{k+1}) \hat{w}_{\tau_k} \right|^2$$

$$= R^e \sum_{\tau_k} \int_{\mathbb{R}^2} \left| f_{\tau_k}^{k+1} \right|^2 \left| w_{\tau_k} \right|^2 \leq R^{3e} \sum_{\tau_{k+1}} \int_{\mathbb{R}^2} |f_{\tau_k}^{k+1}|^4$$

where we used Young’s inequality with $\|w_{\tau_k}\|_1 \lesssim 1$ and $f_{\tau_k}^{k+1} = \sum_{\tau_{k+1} \subset \tau_k} f_{\tau_{k+1}}^{k+1}$ with Cauchy-Schwarz again in the last line. \qed

4.3. Bilinear restriction. We will use the following version of a local bilinear restriction theorem, which follows from a standard Córdoba argument [Cor77] included here for completeness.

Theorem 15. Let $S \geq 4$, $\frac{1}{2} \geq D \geq S^{-1/2}$, and $X \subset \mathbb{R}^2$ be any Lebesgue measurable set. Suppose that $\tau$ and $\tau'$ are $D$-separated subsets of $\mathcal{N}_{S-1}(\mathbb{P}^1)$. Then for a partition $\{\theta_S\}$ of $\mathcal{N}_{S-1}(\mathbb{P}^1)$ into $S^{-1/2} \times S^{-1}$-blocks, we have

$$\int_X |f_{\tau}|^2 |f_{\tau'}|^2 (x) dx \lesssim D^{-2} \int_{\mathcal{N}_{S-1/2}(X)} \left| \sum_{\theta_S} |f_{\theta_S}|^2 \ast w_{S1/2} (x) \right|^2 dx.$$ 

In the following proof, the exact definition of the $S^{-1} \times S^{-1}$ blocks $\theta_S$ is not important. However, by $f_{\tau}$ and $f_{\tau'}$, we mean more formally $f_{\tau} = \sum_{\theta_S \cap \tau \neq \emptyset} f_{\theta_S}$ and $f_{\tau'} = \sum_{\theta_S \cap \tau' \neq \emptyset} f_{\theta_S}$.

Proof. Let $B$ be a ball of radius $S^{1/2}$ centered at a point in $X$. Let $\varphi_B$ be a smooth function satisfying $\varphi_B \gtrsim 1$ in $B$, $\varphi_B$ decays rapidly away from $B$, and $\varphi_B$ is supported in the $S^{-1/2}$ neighborhood of the origin. Then

$$\int_{X \cap B} |f_{\tau}|^2 |f_{\tau'}|^2 \lesssim \int_{\mathbb{R}^2} |f_{\tau}|^2 |f_{\tau'}|^2 \varphi_B.$$
Since $S$ is a fixed parameter and $\theta_S$ are fixed $\sim S^{-1/2} \times S^{-1}$ blocks, simplify notation by dropping the $S$. Expand the squared terms in the integral above to obtain

$$\int_{\mathbb{R}^2} |f_{\tau}|^2 |f_{\tau'}|^2 \varphi_B = \sum_{\theta_1 \cap \tau \neq \emptyset} \sum_{\theta_2' \cap \tau' \neq \emptyset} \int_{\mathbb{R}^2} f_{\theta_1} \overline{f}_{\theta_2} f_{\theta_2'} \overline{f}_{\theta_1'} \varphi_B.$$ 

By Plancherel’s theorem, each integral vanishes unless

$$\tag{9} (\theta_1 - \theta_2) \cap N_{S^{-1/2}}(\theta_1' - \theta_2') \neq \emptyset.$$

Next we check that the number of tuples $(\theta_1, \theta_2, \theta_1', \theta_2')$ (with $\theta_1, \theta_2$ having nonempty intersection with $\tau$ and $\theta_1', \theta_2'$ having nonempty intersection with $\tau'$) satisfying \text{(9)} is $O(D^{-1})$. Indeed, suppose that $\xi < \xi' < \xi'' < \xi'''$ satisfy

$$(\xi, \xi^2) \in \theta_1, \quad (\xi', (\xi')^2) \in \theta_2, \quad (\xi'', (\xi'')^2) \in \theta_1', \quad (\xi''', (\xi''')^2) \in \theta_2'$$

and

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

Then by the mean value theorem, $\xi^2 - (\xi')^2 = 2\xi_1 (\xi - \xi')$ for some $\xi < \xi_1 < \xi'$ and $(\xi'')^2 - (\xi''')^2 = 2\xi_2 (\xi'' - \xi''')$ for some $\xi'' < \xi_2 < \xi'''$. Since $(\xi_1, \xi_2^2) \in \tau$ and $(\xi_2, \xi_3^2) \in \tau'$, we also know that $|\xi_1 - \xi_2| \geq D$. Putting everything together, we have

$$|\xi^2 - (\xi')^2 - ((\xi'')^2 - (\xi''')^2)| = 2|\xi_1 (\xi - \xi') - \xi_2 (\xi'' - \xi''')| \geq 2|\xi_1 - \xi_2||\xi - \xi'| - cS^{-1/2} \geq (2C - c)S^{-1/2}$$

if either $\text{dist}((\xi, \xi^2), (\xi', (\xi')^2))$ or $\text{dist}((\xi'', (\xi'')^2), (\xi''', (\xi''')^2))$ is larger than $CD^{-1}S^{-1/2}$. Thus for a suitably large $C$, the heights will have difference larger than the allowed $O(S^{-1/2})$-neighborhood imposed by \text{[9]}. The conclusion is that

$$\sum_{\theta_1 \cap \tau \neq \emptyset} \sum_{\theta_2' \cap \tau' \neq \emptyset} \int_{\mathbb{R}^2} f_{\theta_1} \overline{f}_{\theta_2} f_{\theta_2'} \overline{f}_{\theta_1'} \varphi_B = \sum_{\theta_1 \cap \tau \neq \emptyset} \sum_{\theta_2' \cap \tau' \neq \emptyset} \int_{\mathbb{R}^2} f_{\theta_1} \overline{f}_{\theta_2} f_{\theta_2'} \overline{f}_{\theta_1'} \varphi_B \lesssim D^{-2} \int_{\mathbb{R}^2} (\sum_{\theta} |f_{\theta}|^2)^2 \varphi_B.$$
Using the locally constant property and summing over a finitely overlapping cover of \(\mathbb{R}^2\) by \(S^{1/2}\)-balls \(B'\) with centers \(c_{B'}\), we have

\[
\int_{\mathbb{R}^2} \left( \sum_{\theta} |f_{\theta}|^2 \right)^2 \varphi_B \leq \sum_{B'} |B| \left( \sum_{\theta} \|f_{\theta}\|_{L^\infty(B')}^2 \|\varphi_B\|_{L^\infty(B')} \right)^2 \leq |B| \left( \sum_{B'} \sum_{\theta} |f_{\theta}|^2 \|\varphi_B\|_{L^\infty(B')} \right)^2 \leq |B|^{-1} \left( \int_{\mathbb{R}^2} \sum_{\theta} |f_{\theta}|^2 \varphi_{S^{1/2}}(c_{B'}) \right)^2 \leq |B|^{-1} \left( \int_B \sum_{\theta} |f_{\theta}|^2 \varphi_{S^{1/2}}(y) dy \right)^2 \leq \int_B \left( \sum_{\theta} |f_{\theta}|^2 \varphi_{S^{1/2}}(y) \right)^2
\]

where we used that \(\varphi_{S^{1/2}}(y) \lesssim \varphi_{S^{1/2}}(c_B)\) in the second to last line.

5. Proof of Theorem 4

Theorem 4 follows from the following proposition and a broad-narrow argument in §5.2. First we prove a version of Theorem 4 where \(U_\alpha\) is replaced by a “broad” version of \(U_\alpha\).

5.1. The broad version of Theorem 4

Let \(\delta > 0\) be a parameter we will choose in the broad/narrow analysis. The notation \(\ell(\tau) = s\) means that \(\tau\) is an approximate \(s \times s^2\) block which is part of a partition of \(N_{s/2}(\mathbb{P}^1)\). For two non-adjacent blocks \(\tau, \tau'\) satisfying \(\ell(\tau) = \ell(\tau') = R^{-\delta}\), define the broad version of \(U_\alpha\) to be

\[
\operatorname{Br}_\alpha(\tau, \tau') = \{ x \in \mathbb{R}^2 : \alpha \sim |f_\tau(x) f_{\tau'}(x)|^{1/2}, \ (|f_\tau(x)| + |f_{\tau'}(x)|) \leq R^{O(\delta)} \alpha \}.
\]

**Proposition 1.** Suppose that \(f\) satisfies the hypotheses of Theorem 4 and has an \((R, \varepsilon)\)-normalized distribution function \(\lambda(\cdot)\). Then

\[
|\operatorname{Br}_\alpha(\tau, \tau')| \leq C_{\varepsilon, \delta} R^4 R^{O(\delta)} \left\{ \sum_{s} \frac{\lambda(s)}{\alpha^4} \sum_{\gamma} \|f_\gamma\|_2^2 \quad \text{if} \quad \alpha^2 > \frac{\lambda(1)}{\max \lambda(s^{-1} R^{-1}) \lambda(s)} \right\}
\]

**Proof of Proposition 4**

Bounding \(|\operatorname{Br}_\alpha(\tau, \tau') \cap H|\): Using bilinear restriction, given here by Theorem 15, we have

\[
\alpha^4 |\operatorname{Br}_\alpha(\tau, \tau') \cap H| \lesssim \sum_{\ell(\tau) = \ell(\tau') = R^{-\delta}} \int_{d(\tau, \tau') \geq R^{-\delta}} |f_\tau|^2 |f_{\tau'}|^2 \lesssim R^{O(\delta)} \int_{N_{R^{1/2}}(\operatorname{Br}_\alpha(\tau, \tau') \cap H)} \sum_{\theta} |f_\theta|^2 w_{R^{1/2}}(x) \lesssim \sum_{\theta} |f_\theta|^2 w_{R^{1/2}} \lesssim R^4 R^{O(\delta)}.
\]
By the locally constant property and the pointwise inequality $w_{R^{1/2}} \ast w_{R^{1/2}} \lesssim w_{R^{1/2}}$ for each $\theta$, we have that $\sum_{\theta} |f_{\theta}|^2 \ast w_{R^{1/2}} \lesssim G(x)$. Then

\begin{equation}
\int_{N_{R^{1/2}}(B_{R^{1/2}}(\tau, \tau')) \cap H} |G(x)|^2 dx \leq \sum_{Q_{R^{1/2}}: Q_{R^{1/2}} \cap (B_{R^{1/2}}(\tau, \tau') \cap H) \neq \emptyset} |Q_{R^{1/2}}|^2 \left\| G \right\|_{L^\infty}^2 (Q_{R^{1/2}} \cap (B_{R^{1/2}}(\tau, \tau') \cap H)) \end{equation}

For each $x \in H$, $G(x) \leq 2|G^h(x)|$. Also note the equality $G^h(x) = \sum_{s} G \ast \tilde{\eta}_{s}(x)$ where the sum is over dyadic $s$ in the range $[R^\delta]^{-1} \lesssim s \lesssim R^{-1/2}$. This is because the Fourier support of $G^h$ is contained in $\cup_\theta (\theta - \theta) \setminus B_{c[R^\delta]^{-1}}$ for a sufficiently small $c > 0$. By dyadic pigeonholing, there is some dyadic $s$, $[R^\delta]^{-1} \lesssim s \lesssim R^{-1/2}$, so that the upper bound in (11) is bounded by

\[ (\log R) \sum_{Q_{R^{1/2}}: Q_{R^{1/2}} \cap (B_{R^{1/2}}(\tau, \tau') \cap H) \neq \emptyset} |Q_{R^{1/2}}|^2 \left\| G \ast \tilde{\eta}_{s} \right\|_{L^\infty}^2 (Q_{R^{1/2}} \cap (B_{R^{1/2}}(\tau, \tau') \cap H)) \]

By the locally constant property, the above displayed expression is bounded by

\[ (\log R) \sum_{Q_{R^{1/2}}: Q_{R^{1/2}} \cap (B_{R^{1/2}}(\tau, \tau') \cap H) \neq \emptyset} \int_{\mathbb{R}^2} |G \ast \tilde{\eta}_{s}|^2 w_{Q_{R^{1/2}}} \lesssim (\log R) \int_{\mathbb{R}^2} |G \ast \tilde{\eta}_{s}|^2. \]

Use Lemma 13 to upper bound the above integral to finish bounding $|B_{R^{1/2}}(\tau, \tau') \cap H|$. 

Bounding $|B_{R^{1/2}}(\tau, \tau') \cap \Omega_k|$: First write the trivial inequality

\[ \alpha^4 |B_{R^{1/2}}(\tau, \tau') \cap \Omega_k| \leq \sum_{\ell(\tau') = \ell(\tau) = R^{-\delta} \atop d(\tau, \tau') \geq R^{-\delta}} \int_{B_{R^{1/2}}(\tau, \tau') \cap \Omega_k \cap \{|f_{\tau} f_{\tau'}|^{1/2} \sim \alpha\}} |f_{\tau}|^2 |f_{\tau'}|^2. \]

By the definition of $B_{R^{1/2}}(\tau, \tau') \cap \Omega_k$ and Lemma 12 for each $x \in B_{R^{1/2}}(\tau, \tau') \cap \Omega_k$ we have

\[ |f_{\tau}(x) f_{\tau'}(x)| \leq |f_{\tau}(x)||f_{\tau'}(x) - f_{\tau'}^{k+1}(x)| + |f_{\tau}(x) - f_{\tau}^{k+1}(x)||f_{\tau'}^{k+1}(x)| + |f_{\tau}^{k+1}(x)| |f_{\tau'}^{k+1}(x)| \lesssim C_\varepsilon R^{O(\delta)} R^{-M\delta} \alpha^2 + |f_{\tau}^{k+1}(x) f_{\tau'}^{k+1}(x)|. \]

For $M$ large enough in the definition of pruning (depending on the implicit universal constant from the broad/narrow analysis which determines the set $B_{R^{1/2}}(\tau, \tau')$) so that $R^{O(\delta)} R^{-M\delta} \leq R^{-\delta}$ and for $R$ large enough depending on $\varepsilon$ and $\delta$, we may bound each integral by

\[ \int_{\{B_{R^{1/2}}(\tau, \tau') \cap \Omega_k \cap \{|f_{\tau} f_{\tau'}|^{1/2} \sim \alpha\}} |f_{\tau}|^2 |f_{\tau'}|^2 \lesssim \int_{B_{R^{1/2}}(\tau, \tau') \cap \Omega_k} |f_{\tau}^{k+1}|^2 |f_{\tau'}^{k+1}|^2 \]

Repeat analogous bilinear restriction, high-dominated from the definition of $\Omega_k$, and locally-constant steps from the argument bounding $B_{R^{1/2}}(\tau, \tau') \cap H$ to obtain

\[ \alpha^4 |B_{R^{1/2}}(\tau, \tau') \cap \Omega_k| \lesssim R^{O(\delta)} \int_{\mathbb{R}^2} |g_k^h|^2. \]
Use Lemma 14 and Lemma 9 to bound the above integral, obtaining
\[
\alpha^4 |Br_\alpha(\tau, \tau') \cap \Omega_k| \lesssim (\log R)^4 \int_{\mathbb{R}^2} |g_k^b|^2 \lesssim R^{O(\delta)} R^{O(\epsilon)} \frac{\lambda(1)^2}{\alpha^2} \sum_{\tau_{k+1}} \int_{\mathbb{R}^2} |f_{\tau_{k+1}}^b|^2.
\]

Use \(L^2\)-orthogonality and that \(|f_{\tau_{m+1}}^b| \leq |f_{\tau_{m}}^b|\) for each \(m\) to bound each integral above:
\[
\int_{\mathbb{R}^2} |f_{\tau_{k+1}}^b|^2 \leq \int_{\mathbb{R}^2} |f_{\tau_{k+1}}^b|^2 \lesssim C \sum_{\tau_{k+2} \subset \tau_{k+1}} \int_{\mathbb{R}^2} |f_{\tau_{k+2}}^b|^2 \lesssim \cdots \lesssim C^{-1} \sum_{\gamma_{\subset \tau_{k+1}}} \int_{\mathbb{R}^2} |f_{\gamma}|^2.
\]

We are done with this case because
\[
\frac{\lambda(1)^2}{\alpha^2} \leq \begin{cases} 
\max \lambda(s^{-1}R^{-1})\lambda(s) & \text{if } \alpha^2 > \frac{\lambda(1)^2}{\alpha^2}, \\
\lambda(1)^2 \alpha^2 & \text{if } \alpha^2 \leq \frac{\lambda(1)^2}{\alpha^2}.
\end{cases}
\]

Bounding \(|Br_\alpha(\tau, \tau') \cap L|\): Repeat the pruning step from the previous case to get
\[
\alpha^6 |Br_\alpha(\tau, \tau') \cap L| \lesssim \sum_{\ell(\tau) = \ell(\tau') = R^{-\delta}} \int_{Br_\alpha(\tau, \tau') \cap L} |f_{\tau}^1 f_{\tau'}^1|^2 |f_{\tau} f_{\tau'}|.
\]

Use Cauchy-Schwartz and the locally constant lemma for the bound \(|f_{\tau}^1 f_{\tau'}^1| \lesssim R^{O(\epsilon)} G_0\) and recall that by Lemma 11 \(G_0 \leq C R^\delta \lambda(1)\). Then
\[
R^{O(\epsilon)} \sum_{\ell(\tau) = \ell(\tau') = R^{-\delta}} \int_{Br_\alpha(\tau, \tau') \cap L} |G_0|^2 |f_{\tau} f_{\tau'}| \leq R^{O(\epsilon)} \lambda(1)^2 \sum_{\ell(\tau) = R^{-\delta}} \int_{\mathbb{R}^2} |f_{\tau}|^2 \lesssim R^{O(\epsilon)} \lambda(1)^2 \sum_{\gamma} \|f_{\gamma}\|_2^2.
\]

Using the same upper bound for \(\frac{\lambda(1)^2}{\alpha^2}\) as in the previous case finishes the proof.

\(\square\)

### 5.2. Bilinear reduction.

We will present a broad/narrow analysis to show that Proposition 11 implies Theorem 4. In order to apply Proposition 11 we must reduce to the case that \(f\) has an \((R, \varepsilon)\)-normalized distribution function \(\lambda(\cdot)\). We demonstrate this through a series of pigeonholing steps.

**Proposition 7 implies Theorem 4**

We will pigeonhole the \(f_\gamma\) so that, roughly, for any \(s\)-arc \(\omega\) of the parabola, the number
\[
\# \{\gamma : \gamma \cap \omega \neq \emptyset, \; f_\gamma \neq 0\}
\]
is either 0 or relatively constant among \(s\)-arcs \(\omega\). For the initial step, write
\[
\{\tau_N : \exists \gamma \text{ s.t. } f_\gamma \neq 0, \; \gamma \subset \tau_N\} = \sum_{1 \leq \lambda \leq R^\delta R^{-\varepsilon}} \Lambda_N(\lambda)
\]
where \(\lambda\) is a dyadic number, \(\{\tau_N : \# \gamma \subset \tau_N \sim \lambda\}, \# \gamma \subset \tau_N \text{ means } \# \{\gamma \subset \tau_N : f_\gamma \neq 0\},\) and \(\# \gamma \subset \tau_N \sim \lambda\) means \(\lambda \leq \# \gamma \subset \tau_N < 2\lambda\). Since there are \(\lesssim \log R\) many \(\lambda\) in the sum, there exists some \(\lambda_N\) such that
\[
|\{x : |f(x)| > \alpha\}| \leq C(\log R)|\{x : C(\log R)| \sum_{\tau_N \in \Lambda_N(\lambda_N)} f_{\tau_N}(x)| > \alpha\}|.
\]
Continuing in this manner, we have

$$\{ \tau_k : \exists \tau_{k+1} \in \Lambda_{k+1}(\lambda_{k+1}) \text{ s.t. } \tau_{k+1} \subset \tau_k \} = \sum_{1 \leq \lambda \leq \tau_k} \Lambda_k(\lambda)$$

where \( \Lambda_k(\lambda) = \{ \tau_k : \exists \tau_{k+1} \in \Lambda_{k+1}(\lambda_{k+1}) \text{ s.t. } \tau_{k+1} \subset \tau_k \} \) and \( \# \gamma \subset \tau_k \sim \lambda \) and for some \( \lambda_k \),

$$\{ x : (C(\log R))^{N-k} \sum_{\tau_{k+1} \in \Lambda_{k+1}(\lambda_{k+1})} f_{\tau_{k+1}}(x) > \alpha \} \leq C(\log R)\{ x : (C(\log R))^{N-k+1} \sum_{\tau_k \in \Lambda_k(\lambda_k)} f_{\tau_k}(x) > \alpha \}.$$ 

Continue this process until we have found \( \tau_1, \lambda_1 \) so that

$$\| x : |f(x)| > \alpha \| \leq C^{\varepsilon-1}(\log R)^{O((\varepsilon^{-1})}) \{ x : C^{\varepsilon-1}(\log R)^{O((\varepsilon^{-1})}} \sum_{\tau_1 \in \Lambda_1(\lambda_1)} f_{\tau_1}(x) > \alpha \}.$$ 

The function \( \sum_{\tau_1 \in \Lambda_1(\lambda_1)} f_{\tau_1} \) now satisfies the hypotheses of Theorem 4 and the property that \( \# \gamma \subset \tau_k \sim \lambda_k \) or \( \# \gamma \subset \tau_k = 0 \) for all \( k, \tau_k \). It follows that the associated distribution function \( \lambda(\cdot) \) of \( \sum_{\tau_1 \in \Lambda_1(\lambda_1)} f_{\tau_1} \) is \((R, \varepsilon)-normalized\) since

$$\lambda_m \sim \# \gamma \subset \tau_m = \sum_{\tau_k \subset \tau_m} \# \gamma \subset \tau_k \sim (\# \tau_k \subset \tau_m)(\lambda_k)$$

where we only count the \( \gamma \) or \( \tau_k \) for which \( f_\gamma \) or \( f_{\tau_k} \) is nonzero. Now we may apply Proposition 1. Note that since \( \log R \leq \varepsilon^{-1} R^\delta \) for all \( R \geq 1 \), the accumulated constant from this pigeonholing process satisfies \( C^{\varepsilon-1}(\log R)^{O(\varepsilon^{-1})} \leq C R^\varepsilon \). It thus suffices to prove Theorem 4 assuming that \( f \) is \((R, \varepsilon)-normalized\).

Now we present a broad-narrow argument adapted to our set-up. Write \( K = R^\delta \) for some \( \delta > 0 \) which will be chosen later. Since \( |f(x)| \leq \sum_{\tau \subset \tau_K} |f_{\tau}(x)| \), there is a universal constant \( C > 0 \) so that \( |f(x)| > K^C \max_{\ell(\tau) = \ell(\tau') = K^{-1}} |f_{\tau}(x)| f_{\tau'}(x)|^{1/2} \) implies \( |f(x)| \leq C \max_{\ell(\tau) = K^{-1}} |f_{\tau}(x)| \). If \( |f(x)| \leq K^C \max_{\ell(\tau) = K^{-1}} \max_{\tau, \tau' \text{ nonadj.}} |f_{\tau}(x)| f_{\tau'}(x)|^{1/2} \) and

$$K^C \max_{\ell(\tau) = K^{-1}} \max_{\tau, \tau' \text{ nonadj.}} |f_{\tau}(x)| f_{\tau'}(x)|^{1/2} \leq C \max_{\ell(\tau) = K^{-1}} |f_{\tau}(x)|$$

then \( |f(x)| \leq C \max_{\ell(\tau) = K^{-1}} |f_{\tau}(x)| \).

Using this reasoning, we obtain the first step in the broad-narrow inequality

$$|f(x)| \leq C \max_{\ell(\tau) = K^{-1}} |f_{\tau}(x)| + K^C \max_{\ell(\tau) = K^{-1}} \max_{\tau, \tau' \text{ nonadj.}} |f_{\tau}(x)| f_{\tau'}(x)|^{1/2}.$$ 

Iterate the inequality \( m \) times (for the first term) where \( K^m \sim R^{1/2} \) to bound \( |f(x)| \) by

$$|f(x)| \leq C^m \max_{\ell(\tau) = R^{-1/2}} |f_{\tau}(x)| + C^m K^C \sum_{R^{-1/2} \leq \Delta < 1} \max_{\ell(\tau) = \Delta} \max_{\tau, \tau' \text{ nonadj.}} |f_{\tau}(x)| f_{\tau'}(x)|^{1/2}.$$
Recall that our goal is to bound the size of the set
\[ U_\alpha = \{ x \in \mathbb{R}^2 : \alpha \leq |f(x)| \}. \]

By the triangle inequality and using the notation \( \theta \) for blocks \( \tau \) with \( \ell(\tau) = R^{-1/2} \)
\[
|U_\alpha| \leq \left| \{ x \in \mathbb{R}^2 : \alpha \leq C^m \max_\theta |f_\theta(x)| \} \right| + \sum_{R^{1/2} < \Delta < 1} \sum_{\Delta \in \mathcal{K}^n} |U_\alpha(\tau, \tau')| \]

where \( U_\alpha(\tau, \tau') \) is the set
\[
\{ x \in \mathbb{R}^2 : \alpha \lesssim (\log R)^m K^C |f_\tau(x) f_{\tau'}(x)|^{1/2}, \ C(|f_\tau(x)| + |f_{\tau'}(x)|) \leq K^C |f_\tau(x) f_{\tau'}(x)|^{1/2} \}.
\]

The first term in the upper bound from (12) is bounded trivially by \( \frac{\lambda(R^{-1/2})^2}{\alpha^2} \sum_\gamma \|f_\gamma\|^2 \). By the assumption that \( \|f_\gamma\|_\infty \lesssim \alpha \) for every \( \gamma \), we know that \( |f_\tau| \lesssim R^\beta \) for any \( \tau \). Also assume without loss of generality that \( \alpha > 1 \) (otherwise Theorem 4 follows from \( L^2 \)-orthogonality). This means that there are \( \sim \log R \) dyadic values of \( \alpha' \) between \( \alpha \) and \( R^\beta \) so by pigeonholing, there exists \( \alpha' \in [\alpha/(C^m K^C), R^\beta] \) so that
\[
|U_\alpha(\tau, \tau')| \lesssim (\log R + \log(C^m K^C))|B_{\alpha'}(\tau, \tau')|
\]

where the set \( B_{\alpha'}(\tau, \tau') \) is defined in (10). By parabolic rescaling, there exists an affine transformation \( T \) so that \( f_\tau \circ T = g_v \) and \( f_{\tau'} \circ T = g_{v'} \) where \( \tau \) and \( \tau' \) are \( \sim K^{-1} \)-separated blocks in \( \mathcal{N}_{\Delta - 2R^{-1}}(P) \). Note that the functions \( g_v \) and \( g_{v'} \) inherit the property of being \( (\Delta^2 R, \varepsilon) \)-normalized in the sense required to apply Proposition 1 in each of the following cases.

**Case 1:** Suppose that for some \( \beta' \in \left[ \frac{1}{2}, 1 \right] \), \( \Delta^{-1} R^{-\beta} = (\Delta^2 R)^{-\beta'} \). Then for each \( \gamma \in \mathcal{P}(R, \beta) \), \( f_\gamma \circ T = g_v \) for some \( \gamma \in \mathcal{P}(\Delta^2 R, \beta') \). Applying Proposition 1 with functions \( g_v \) and \( g_{v'} \) and level set parameter \( \alpha' \) leads to the inequality
\[
|B_{\alpha'}(\tau, \tau')| \lesssim K^C \alpha'
\]
\[
\leq C_{\varepsilon, \delta} R^\varepsilon C^m K^{O(1)} \times \left\{ \begin{array}{ll}
\frac{1}{(\alpha')^2} & \text{if } (\alpha')^2 > \frac{\lambda(\Delta)}{\max_s \lambda(s^{-1} R^{-1} - 1) \lambda(s)} \\
\frac{\lambda(\Delta^2)}{(\alpha')^2} & \text{if } (\alpha')^2 > \frac{\lambda(\Delta^2)}{\max_s \lambda(s^{-1} R^{-1} - 1) \lambda(s)}
\end{array} \right.
\]

**Case 2:** Now suppose that \( \Delta^{-1} R^{-\beta} < (\Delta^2 R)^{-1} \). Let \( \bar{\theta} \) be \( \Delta^{-1} R^{-1} \times R^{-1} \) blocks and let \( \bar{\theta} \) be \( \Delta^{-1} R^{-1} \times (\Delta^2 R)^{-1} \) blocks so that \( f_\bar{\theta} \circ T = g_v \). Let \( B = \max_\bar{\theta} |f_\bar{\theta}| \) and divide everything by \( B \) in order to satisfy the hypotheses \( ||g_\bar{\theta}||_\infty/B \leq 1 \) for all \( \bar{\theta} \). Let \( \bar{\lambda}(s) := \lambda(\Delta s)/\lambda(\Delta^{-1} R^{-1}) \) count the number of \( \bar{\theta} \) intersecting an \( s \)-arc. In the case \( (\alpha')^2 > \frac{\bar{\lambda}(\Delta^2)^2}{\max_s \lambda(s^{-1} (\Delta^2 R)^{-1} - 1) \lambda(s)} \) (with the maximum taken over \( (\Delta^2 R)^{-1} < s < (\Delta^2 R)^{-1/2} \)), use Proposition 1 with functions \( g_v/B \) and \( g_{v'}/B \) and level set parameter \( \alpha'/B \) to get the inequality
\[
|B_{\alpha'}(\tau, \tau')| \leq C_{\varepsilon, \delta} R^\varepsilon C^m K^{O(1)} \frac{B^4}{(\alpha')^4} \Delta^2 R^{-1} \lesssim (\Delta^2 R)^{-1} \bar{\lambda}(s) \sum_{\bar{\theta} \subset \bar{\tau}} ||f_{\bar{\theta}}||_2^2/B^2.
\]

Note that since \( B \leq \lambda(\Delta^{-1} R^{-1}) \),
\[
B^2 \Delta^{-1} R^{-1} \lesssim (\Delta^2 R)^{-1/2} \bar{\lambda}(s) \leq \max_{\Delta^{-1} R^{-1} < s < R^{-1/2}} \lambda(s^{-1} R^{-1}) \lambda(s)
\]
and
\[
\frac{\tilde{\lambda}(1)^2 B^2}{\max_s \tilde{\lambda}(s/(\Delta R)) \tilde{\lambda}(s)} \leq \frac{\lambda(\Delta)^2 \lambda(\Delta^{-1} R^{-1})^2}{\max_{\Delta^{-1} R^{-1} < s < R^{-1/2}} \lambda(s^{-1} R^{-1}) \lambda(s)} \leq \lambda(\Delta^{-1} R^{-1}) \lambda(\Delta).
\]

Then in the case \( (\alpha')^2 \leq \frac{\tilde{\lambda}(1)^2 B^2}{\max_s \lambda(s/(\Delta^2 R)) \lambda(s)} \), compute directly that
\[
(\alpha')^4 \{ x \in \mathbb{R}^2 : \alpha' \sim |f_\tau(x)| \}, \ (|f_\tau(x)| + |f_\tau'(x)|) \leq K^{\alpha'} \}
\[
\lesssim \lambda(\Delta^{-1} R^{-1}) \lambda(\Delta) \int_{\mathbb{R}^2} (|f_\tau|^2 + |f_\tau'|^2) \lesssim \max_{\Delta^{-1} R^{-1} < s < R^{-1/2}} \lambda(s^{-1} R^{-1}) \lambda(s) \sum_{\gamma \in \tilde{\gamma}} \|f_\tau\|^2_2.
\]

Using also that \( \sum_{\delta \in \tilde{\delta}} \|f_\delta\|^2_2 \leq \sum_{\gamma \in \tilde{\gamma}} \|f_\gamma\|^2_2 \), the bound for Case 2 is
\[
\{ x \in \mathbb{R}^2 : \alpha' \sim |f_\tau(x)| \}, \ (|f_\tau(x)| + |f_\tau'(x)|) \leq K^{\alpha'} \}
\[
\lesssim C_{\varepsilon, \delta} R^\alpha C^m \sum_{\gamma \in \tilde{\gamma}} \lambda(s^{-1} R^{-1} \lambda(s)) \sum_{\gamma \in \tilde{\gamma}} \|f_\gamma\|^2_2.
\]

It follows from [12] and the combined Case 1 and Case 2 arguments above that
\[
|U_\alpha| \leq C_{\varepsilon, \delta} R^\alpha C^m \sum_{\gamma \in \tilde{\gamma}} \lambda(s^{-1} R^{-1} \lambda(s)) \sum_{\gamma \in \tilde{\gamma}} \|f_\gamma\|^2_2
\]
\[
\begin{cases}
\frac{1}{\alpha^s} \max_{\Delta^{-1} R^{-1} < s < R^{-1/2}} \lambda(s^{-1} R^{-1}) \lambda(s) \sum_{\gamma \in \tilde{\gamma}} \|f_\gamma\|^2_2 & \text{if } \alpha > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1} R^{-1}) \lambda(s)} \\
\frac{1}{\alpha^s} \sum_{\gamma \in \tilde{\gamma}} \|f_\gamma\|^2_2 & \text{if } \alpha^2 \leq \frac{\lambda(1)^2}{\max_s \lambda(s^{-1} R^{-1}) \lambda(s)}
\end{cases}
\]

Recall that \( K^m \sim R^{-1/2} \) and \( K = R^\delta \) so that \( C_{\varepsilon, \delta} R^\alpha C^m \sum_{\gamma \in \tilde{\gamma}} \lambda(s^{-1} R^{-1} \lambda(s)) \sum_{\gamma \in \tilde{\gamma}} \|f_\gamma\|^2_2 \leq C_{\varepsilon, \delta} R^\alpha C^{O(1)} R^{O(1) \delta} \). Choosing \( \delta \) small enough so that \( R^{O(1) \delta} \leq R^\varepsilon \) finishes the proof.

\[\square\]

**REFERENCES**

[1] Tom M. Apostol. *Introduction to analytic number theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1976.

[2] Jean Bourgain and Ciprian Demeter. The proof of the \( l^2 \) decoupling conjecture. *Ann. of Math.* (2), 182(1):351–389, 2015.

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

[4] J. Bourgain. Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.*, 30(1):205–224, 2017.

[5] J. Bourgain. Decoupling inequalities and some mean-value theorems. *J. Anal. Math.*, 133:313–334, 2017.

[6] Antonio Córdoba. The Kakeya maximal function and the spherical summation multipliers. *American Journal of Mathematics*, 99(1):1–22, 1977.

[7] Ciprian Demeter. *Fourier restriction, decoupling, and applications*, volume 184 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2020.

[8] Ciprian Demeter, Larry Guth, and Hong Wang. Small cap decouplings. *Geom. Funct. Anal.*, 30(4):989–1062, 2020. With an appendix by D. R. Heath-Brown.

[9] Larry Guth, Dominique Maldague, and Hong Wang. Improved decoupling for the parabola. *arXiv preprint arXiv:2009.07953*, 2020.