Rev. Mat. Iberoam. 39 (2023), no. 3, 975–1004
DOI 10.4171/RMI/1393

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 $\mathbb{P}^1$ of Demeter, Guth, and Wang (2020) and the canonical decoupling theorem for $\mathbb{P}^1$ of Bourgain and Demeter (2015). New (ℓ^q, L^p) small cap decoupling inequalities also follow from our sharp level set estimates.

1. Introduction

In this paper, we further develop the high/low frequency proof of decoupling for the parabola [9] to prove sharp level set estimates which recover and refine the small cap decoupling results for the parabola in [8]. We begin by describing the problem and our results in terms of exponential sums. The main results in full generality are in §2.

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

$$|Q_R|^{-1} \int_{Q_R} \Big| \sum_{\xi \in \Xi} a_\xi \, e((x,t) \cdot (\xi, \xi^2)) \Big|^p \, dx \, dt \leq D(N, R, p) \, N^{p/2} \tag{1.1}$$

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 [8] is that if $2 \leq p \leq 2 + 2s$ for $R = N^s$, then

$$D(N, R, p) \leq C_\varepsilon N^\varepsilon. \tag{1.2}$$

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 (1.2) follows from

2020 Mathematics Subject Classification: Primary 42B15; Secondary 42B20.
Keywords: Decoupling inequalities, superlevel set.

the canonical ℓ^2 decoupling theorem of Bourgain and Demeter for the parabola [5]. For $R < N^2$ and the subset $\Xi = \{k/N\}_{k=1}^N$, the inequality (1.1) 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 [3, 4].

By a pigeonholing argument (see Section 5 of [9]), (1.2) follows from upper bounds for superlevel sets U_α defined by

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

In particular, (1.2) is equivalent, up to a $\log N$ factor, to proving that for any $\alpha > 0$ and for $R = N^s$,

$$\alpha^{2+2s}\,|U_\alpha \cap Q_R| \le C_\varepsilon R^\varepsilon N^{1+s} R^2 \tag{1.3}$$

when Ξ and a_ξ satisfy the hypotheses following (1.1). In this paper, we improve the above superlevel set estimate for all $\alpha > 0$ strictly between $N^{1/2}$ and N.

Theorem 1.1. *Let $R \in [N, N^2]$. For any $\varepsilon > 0$, there exists $C_\varepsilon < \infty$ such that*

$$|U_\alpha \cap Q_R| \le C_\varepsilon N^\varepsilon \begin{cases} \frac{N^2 R}{\alpha^4}\sum_{\xi\in\Xi}|a_\xi|^2 & \text{if } \alpha^2 > R,\\ \frac{N^2R^2}{\alpha^6}\sum_{\xi\in\Xi}|a_\xi|^2 & \text{if } N \le \alpha^2 \le R,\\ R^2 & \text{if } \alpha^2 < N,\end{cases}$$

whenever $\Xi \subset [-1,1]$ is a $\gtrsim \frac{1}{N}$-separated subset, $|a_\xi| \le 1$ 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 $F(x,t) = \sum_{n=1}^N e((x,t)\cdot(n/N, n^2/N^2))$. It is not known whether the implicit constant in the upper bound of (1.2) goes to infinity with N except in the case that $p = 6$ and $s = 2$, when the same example $F(x,t) = \sum_{n=1}^N e((x,t)\cdot(n/N, n^2/N^2))$ shows that $D(N, N^2, 6) \gtrsim (\log N)$, see [2]. Roughly, the argument is that for each dyadic value $\alpha \in [N^{3/4}, N]$, one can show by counting the "major arcs" that

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

Since there are $\sim \log N$ values of α, the lower bound for $\int_{Q_{N^2}} |F|^6$ follows. Theorem 1.1 implies that the corresponding superlevel set estimates (1.3) are not sharp for $1 \le s < 2$, unless $\alpha \sim N$ or $\alpha^2 \sim N$, which leads to the following conjecture.

Conjecture 1.2. *Let $s \in [1,2)$ and $2 \le p \le 2 + 2s$. There exists $C(s) > 0$ so that*

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

A more refined version of Theorem 1.1 leads to the following essentially sharp (ℓ^q, L^p) small cap decoupling theorem, stated here for general exponential sums.

Corollary 1.3. *Let $3/p + 1/q \le 1$, and let $R \in [N, N^2]$. Then for each $\varepsilon > 0$, there exists $C_\varepsilon < \infty$ so that*

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

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

2. Main results

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

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

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

$$\|f\|_{L^p(B_R)} \le D(R,p)\Big(\sum_\gamma \|f_\gamma\|^2_{L^p(\mathbb{R}^2)}\Big)^{1/2}, \tag{2.1}$$

in which $f:\mathbb{R}^2 \to \mathbb{C}$ is a Schwartz function with $\operatorname{supp}\widehat{f} \subset \mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ and f_γ means the Fourier projection onto γ, defined precisely below. When we refer to canonical caps or to canonical decoupling, we mean that γ are approximately $R^{-1/2} \times R^{-1}$ blocks corresponding to the ℓ^2-decoupling of the paper [5]. In this paper, we allow γ to be approximate $R^{-\beta} \times R^{-1}$ blocks, where $1/2 \le \beta \le 1$. This is the "small cap" regime studied in [8]. We also consider (ℓ^q, L^p) decoupling for small caps, which replaces $(\sum_\gamma \|f_\gamma\|_p^2)^{1/2}$ by $(\sum_\gamma \|f_\gamma\|_p^q)^{1/q}$ in the decoupling inequality above (see Corollary 2.3).

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

$$\bigsqcup_{|k|\le\lceil R^\beta\rceil-2} \big\{(x,t) \in \mathcal{N}_{R^{-1}}(\mathbb{P}^1) : k\lceil R^\beta\rceil^{-1} \le x < (k+1)\lceil R^\beta\rceil^{-1}\big\} \tag{2.2}$$

and the two end pieces

$$\big\{(x,t) \in \mathcal{N}_{R^{-1}}(\mathbb{P}^1) : x < -1 + \lceil R^\beta\rceil^{-1}\big\} \bigsqcup \big\{(x,t) \in \mathcal{N}_{R^{-1}}(\mathbb{P}^1) : 1 - \lceil R^\beta\rceil^{-1} \le x\big\}.$$

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

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

For $a, b > 0$, the notation $a \lesssim b$ means that $a \le Cb$, 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 $C^{-1}b \le a \le Cb$ for a universal constant C.

Let $U_\alpha := \{x \in \mathbb{R}^2 : |f(x)| \geq \alpha\}$. In Section 5 of [9], through a wave packet decomposition and series of pigeonholing steps, the bounds for $D(R,p)$ in (2.1) follow (with an additional power of $(\log R)$) from bounds on the constant $C(R,p)$ in

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

for any $\alpha > 0$ and under the additional assumptions that $\|f_\gamma\|_\infty \lesssim 1$, $\|f_\gamma\|_p^p \sim \|f_\gamma\|_2^2$ for each γ. 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 γ? We answer this question in the following theorem.

Theorem 2.1. *Let $\beta \in [1/2, 1]$ and $R \geq 2$. Let $f : \mathbb{R}^2 \to \mathbb{C}$ be a Schwartz function with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ satisfying $\|f_\gamma\|_\infty \leq 1$ for all $\gamma \in \mathcal{P}(R,\beta)$. Then for any $\alpha > 0$,*

$$|U_\alpha \cap [-R,R]^2| \leq C_\varepsilon R^\varepsilon \begin{cases} \frac{R^{2\beta-1}}{\alpha^4} \sum_\gamma \|f_\gamma\|_{L^2(\mathbb{R}^2)}^2 & \text{if } \alpha^2 > R, \\ \frac{R^{2\beta}}{\alpha^6} \sum_\gamma \|f_\gamma\|_{L^2(\mathbb{R}^2)}^2 & \text{if } R^\beta \leq \alpha^2 \leq R, \\ R^2 & \text{if } \alpha^2 < R^\beta. \end{cases}$$

Each bound in Theorem 2.1 is sharp, up to the $C_\varepsilon R^\varepsilon$ factor, which we show in §3.

Define notation for a distribution function for the Fourier support of a Schwartz function f with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ 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 ξ_1-axis equal to an interval of length s. The following theorem implies Theorem 2.1 and replaces factors of R^β in the upper bounds from Theorem 2.1 by expressions involving $\lambda(\cdot)$, which see the actual Fourier support of the input function f.

Theorem 2.2. *Let $\beta \in [1/2, 1]$ and $R \geq 2$. For any f with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ satisfying $\|f_\gamma\|_\infty \lesssim 1$ for each $\gamma \in \mathcal{P}(R,\beta)$,*

$$|U_\alpha| \leq C_\varepsilon R^\varepsilon \begin{cases} \frac{1}{\alpha^4} \max_s \lambda(s^{-1}R^{-1})\lambda(s) \sum_\gamma \|f_\gamma\|_2^2 & \text{if } \alpha^2 > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}, \\ \frac{\lambda(1)^2}{\alpha^6} \sum_\gamma \|f_\gamma\|_2^2 & \text{if } \alpha^2 \leq \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}, \end{cases}$$

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

See §2.1 for a discussion of the proof of Theorem 2.2.

Corollary 2.3 ((l^q, L^p) small cap decoupling). *Let $3/p + 1/q \leq 1$. Then*

$$\|f\|_{L^p(B_R)} \leq C_\varepsilon R^\varepsilon \left(R^{\beta(1-1/q)-\frac{1}{p}(1+\beta)} + R^{\beta(1/2-1/q)}\right) \Big(\sum_\gamma \|f_\gamma\|_{L^p(\mathbb{R}^2)}^q\Big)^{1/q}$$

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 $\|f\|^p_{L^p(B_R)}/(\sum_\gamma \|f_\gamma\|_p^q)^{p/q}$. The first is the square root cancellation example, where $|f_\gamma| \sim \chi_{B_R}$ for all γ and $f = \sum_\gamma e_\gamma f_\gamma$ where e_γ are ± 1 signs chosen (using Khintchine's inequality) so that $\|f\|^p_{L^p(B_R)} \sim R^{\beta p/2} R^2$.

$$\|f\|_p^p \Big/ \Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q} \gtrsim \frac{R^{\beta p/2} R^2}{R^{\beta p/q} R^2} = R^{\beta p(1/2-1/q)}.$$

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

$$\|f\|_p^p \Big/ \Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q} \gtrsim \frac{R^{\beta p}}{R^{\beta p/q} R^{1+\beta}} = R^{\beta p(1-1/q)-1-\beta}.$$

There is one more example which may dominate the ratio: The block example is $f = R^{1+\beta}\sum_{\gamma\subset\theta}\check{\eta}_\gamma$, where θ 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 θ^*, we have

$$\|f\|_p^p \Big/ \Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q} \gtrsim \frac{R^{(\beta-1/2)p} R^{3/2}}{R^{(\beta-1/2)p/q} R^{1+\beta}} = R^{(\beta-1/2)p(1-1/q)+1/2-\beta}.$$

One may check that the constructive interference examples dominate the block example when $3/p + 1/q \le 1$. We do not investigate (l^q, L^p) small cap decoupling in the range $3/p + 1/q > 1$ in the present paper.

The paper is organized as follows. In §3, we demonstrate that Theorem 2.1 is sharp using an exponential sum example. In §4, we show how Theorem 2.1 follows easily from Theorem 2.2 and how after some pigeonholing steps, so does Corollary 2.3. Then in §5, we develop the multi-scale high/low frequency tools we use in the proof of Theorem 2.2. These tools are very similar to those developed in [9]. It appears that a more careful version of the proof of Theorem 2.2 could also replace the $C_\varepsilon R^\varepsilon$ factor by a power of $(\log R)$, as is done for canonical decoupling in [9]. Finally, in §6, we prove a bilinear version of Theorem 2.2 and then reduce to the bilinear case to finish the proof.

2.1. Overview of the proof of Theorem 2.2

Versions of the high/low method in which we analyze high-frequency and low-frequency portions of functions separately have been used in [10] and [9]. The original small cap decoupling result from [8] also uses a high/low argument to prove a certain refinement of the planar Kakeya estimate.

The proof of Theorem 2.2 closely follows the argument from [9], which is summarized in Section 2 of [9]. We briefly recall the high/low argument in [9], and will highlight the new aspects of adapting the argument to small caps γ.

We write $g_k = \sum_{\tau_k} |f_{\tau_k}|^2$ for canonical caps τ_k of dimensions $R_k^{-1/2}\times R_k^{-1}$, where $R_k = R^{k\varepsilon} \in [1, R]$ for some fixed $\varepsilon > 0$. For $k = 1, \ldots, N$ with $g_N = \sum_\theta |f_\theta|^2$, we consider sets

$$\Omega_k = \{x : g_k(x) \text{ high-dominated},\ g_{k+1}(x), \dots, g_N(x) \text{ low-dominated}\}.$$

Here we say $g_k(x)$ is low-dominated if $g_k(x) \le 2|g_k * \check{\eta}_k|(x)$, and high-dominated otherwise. The function η_k is a smooth bump adapted to $B_{R_{k+1}^{-1/2}}(0)$ in the frequency space. Essentially, the "low lemma" (Lemma 3.24 in [9]) says that when $g_k(x)$ is low-dominated, we have $|g_k(x)| \lesssim |g_{k+1}(x)|$, and the "high lemma" (Lemma 3.25 in [9]) states that $|f_{\tau_k}|^2 - |f_{\tau_k}|^2 * \check{\eta}_k$ are essentially orthogonal, and therefore

$$\int |g_k - g_k * \check{\eta}_k|^2 \lessapprox \int \sum_{\tau_k} |f_{\tau_k}|^4.$$

Another important step in [9] was to observe that on $U_\alpha \cap \Omega_k$, one could prune the wave packets of f to arrange for the upper bound $\|f_{\tau_k}\|_\infty \lessapprox \#\theta/\alpha$ for all τ_k. These three ingredients allowed a re-proof of the canonical cap decoupling of $\mathbb{P}^1$ in [5].

Here is how we adapt the argument from [9] to the small cap case. Write $G_k = \sum_{\theta_k} |f_{\theta_k}|^2$ for small caps θ_k of dimension $R^{-1/2}R^{-k\varepsilon} \times R^{-1}$. In particular, $G_0 = \sum_\theta |f_\theta|^2$ is the square function for the canonical caps θ, and equals to g_N defined in the previous paragraph. Then for $k = 1, \dots, M$ with $G_M = \sum_\gamma |f_\gamma|^2$, we consider sets

$$\Lambda_k = \{x : G_k(x) \text{ high-dominated},\ G_{k+1}(x), \dots, G_M(x) \text{ low-dominated}\}.$$

Here we say that $G_k(x)$ is low-dominated if $G_k(x) \le 2|G_k * \check{\chi}_k|(x)$, and high-dominated otherwise. The function χ_k is a smooth bump adapted to $B_{R^{-1/2}R^{-(k+1)\varepsilon}}(0)$ in the frequency space. Adopting the argument from [9], when $G_k(x)$ is low-dominated, we would have $|G_k(x)| \lesssim |G_{k+1}(x)|$, and when $G_k(x)$ is high-dominated, we could exploit orthogonality properties of $|f_{\theta_k}|^2 - |f_{\theta_k}|^2 * \check{\chi}_k$, as the supports of their Fourier transforms have some quantitatively controlled overlap. In Theorem 2.2, we have the additional hypothesis that $\|f_\gamma\|_\infty \lesssim 1$ for all γ, which leads to a trivial upper bound of $\|f_{\theta_k}\|_\infty \lesssim \#\gamma \subset \theta_k$. To prove Theorem 2.2 in the case involving Λ_k, it turns out that this trivial L^∞ bound suffices, so we do not need to prune the wave packets to get an L^∞ bound on f_{θ_k} of the form $\#\gamma/\alpha$. This allows us to greatly simplify the cases involving square functions at intermediate small cap scales. In particular, we only need to consider the high set H on which $|G_0(x)| \lesssim |G_0 * \check{\chi}_{>R^{-\beta}}|$ and off of which we have the low-dominance inequality $G_0(x) \lesssim |G_0 * \check{\chi}_{\le R^{-\beta}}|$. On the high set, we could simply combine the orthogonality-based estimates of all intermediate scales into one estimate, which will be Lemma 5.12 below. If G_0 satisfies the low-dominance inequality, then we will have $G_0 \lesssim \lambda(1)$ (Lemma 5.8 below), and we consider more high/low cases involving canonical block square functions $\sum_{\tau_k} |f_{\tau_k}|^2$ as in the previous paragraph. The low-dominance inequality $G_0 \lesssim \lambda(1)$ for G_0 is precisely what allows us to re-initiate the pruning process from [9] to guarantee $\|f_{\tau_k}\|_\infty \lessapprox \#\gamma/\alpha$, which is more efficient to use in the cases involving canonical block square functions. Aside from this difference in the pruning process, much of the remainder of the argument resembles [9].

Compared with the argument in [9], we take a more unified approach of applying the high/low method at every scale including the small cap scales θ_k, while [9] uses the high/low method to study a Kakeya-type problem for wavepackets at the canonical scale θ

(see Section 5.2 of [9]) and combines it with a refined decoupling inequality of canonical caps to infer small cap decoupling. Having a systematic high/low argument at every scale allows us to get superlevel set estimates which are more accurate than the ones that can be deduced from the small cap decoupling inequality in [9].

3. A sharp example

Because we will show that Theorem 2.2 implies Theorem 2.1, it suffices to show that Theorem 2.1 is sharp, which we mean up to a $C_\varepsilon R^\varepsilon$ factor. Write $N = \lceil R^\beta \rceil$. The function achieving the sharp bounds is

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

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

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

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

Suppose that $\alpha \sim N$, and note that $F(0,0) = N$ and $|F(x_1,x_2)| \sim N$ when $|(x_1,x_2)| < 1/10^3$. Using the periodicity in the x_1 variable, there are $\sim R/N$ many other heavy balls where $|F(x_1, x_2)| \sim N$ in $[-R, R]^2$. For α in the range suppose that $R < \alpha^2 < N^2$, we will show that U_α 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 2.1 for this range is

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

To demonstrate that this inequality is sharp, by the 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 $\eta_r = r^4(\phi_{r^{-1}} * \phi_{r^{-1}})^\vee$ and analyze the L^2 norm $\|F\|_{L^2(\eta_r)}$. By Plancherel's theorem,

$$\begin{aligned}\|F\|_{L^2(\eta_r)}^2 &= \int |F|^2 \eta_r \sim \int \Big|\sum_{k=1}^{N} e\Big(\frac{k}{N} x_1 + \frac{k^2}{N^2} x_2\Big)\Big|^2 \eta_r(x_1, x_2)\\ &= \sum_{k=1}^{N}\sum_{k'=1}^{N} \widehat{\eta}_r\Big(\xi\Big(\frac{k-k'}{N}, \frac{k^2-(k')^2}{N^2}\Big)\Big) \sim N \cdot N/r \cdot r^2 = rN^2.\end{aligned}$$

Next we bound $\|F\|_{L^4(B_{R^\varepsilon r})}$ above. It follows from the local linear restriction statement (see Theorem 1.14, Proposition 1.27 and Exercise 1.32 in [7])

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

that

$$\|F\|_{L^4(B_{R^\varepsilon r})}^4 \sim \Big\| \sum_{k=1}^{N} e\Big(\frac{k}{N}x_1 + \frac{k^2}{N^2}x_2\Big)\,\eta_r(x_1,x_2)\Big\|_{L^4(B_{R^\varepsilon r})}^4$$
$$\lesssim C_\varepsilon R^\varepsilon r^{-3} \Big\| \sum_{k=1}^{N} \widehat{\eta}_r\Big(\xi - \Big(\frac{k}{N},\frac{k^2}{N^2}\Big)\Big)\Big\|_{L^4(\mathbb{R}^2)}^4.$$

The L^4 norm on the right-hand side is bounded above by

$$\int_{B_2}\Big|\sum_{k=1}^{N}\widehat{\eta}_r\Big(\xi-\Big(\frac{k}{N},\frac{k^2}{N^2}\Big)\Big)\Big|^4 d\xi \lesssim (Nr^{-1})^3\int_{B_2}\sum_{k=1}^{N}\Big|\widehat{\eta}_r\Big(\xi-\Big(\frac{k}{N},\frac{k^2}{N^2}\Big)\Big)\Big|^4 d\xi$$
$$\lesssim (Nr^{-1})^3(r^2)^3\int_{B_2}\sum_{k=1}^{N}\Big|\widehat{\eta}_r\Big(\xi-\Big(\frac{k}{N},\frac{k^2}{N^2}\Big)\Big)\Big| d\xi \sim N^4r^3.$$

This leads to the upper bound $\|F\|_{L^4(B_{R^\varepsilon r})}^4 \lesssim (\log R)N^4$.

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

$$\lambda^2 rN^2 \sim \lambda^2\|F\|_{L^2(\eta_r)}^2 \lesssim (\log R)\lambda^4|\{x\in B_{R^\varepsilon r} : |F(x)|\sim\lambda\}| + C_\varepsilon\lambda^4R^{-2000}$$
$$\lesssim (\log R)\|F\|_{L^4(B_{R^\varepsilon r})}^4 + C_\varepsilon\lambda^4R^{-2000} \lesssim C_\varepsilon R^\varepsilon N^4 + C_\varepsilon\lambda^4R^{-2000}.$$

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 ε,

$$rN^2 \sim (\log R)\lambda^2|\{x\in B_{R^\varepsilon r} : |F(x)|\sim\lambda\}| \lesssim C_\varepsilon R^\varepsilon(N^2/r)|\{x\in B_{R^\varepsilon r} : |F(x)|\sim\lambda\}|,$$

so $|\{x\in B_{R^\varepsilon r} : |F(x)|\sim\lambda\}| \gtrsim C_\varepsilon^{-1}R^{-\varepsilon}r^2$ and $\lambda^2 \gtrsim C_\varepsilon^{-1}R^{-\varepsilon}N^2/r \sim C_\varepsilon^{-1}R^{-\varepsilon}\alpha^2$.

Case 2. $R^\beta < \alpha^2 \le R$.

Let q, a, and b be integers satisfying

$$q \text{ odd},\quad 1\le b\le q\le N^{2/3},\quad (b,q)=1,\quad\text{and}\quad 0\le a\le q. \tag{3.1}$$

Define the set $M(q,a,b)$ to be

$$M(q,a,b) := \big\{(x_1,x_2)\in[0,N]\times[0,N^2] : |x_1-\tfrac{a}{q}N| \le \tfrac{1}{10^{10}},\quad |x_2-\tfrac{b}{q}N^2|\le\tfrac{1}{10^{10}}\big\}.$$

Lemma 3.1. *For each $(q,a,b)\neq(q',a',b')$, both tuples satisfying (3.1), $M(q,a,b)\cap M(q',a',b')=\emptyset$.*

Proof. If $b/q=b'/q'$, then using the relatively prime part of (3.1), $b=b'$ 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

$$\frac{2}{10^{10}} \ge \Big|x_1-\frac{a}{q}N\Big| + \Big|x_1-\frac{a'}{q}N\Big| \ge \frac{|a-a'|N}{q} \ge N^{1/3},$$

which is clearly a contradiction. The alternative is that $b/q \neq b'/q'$, in which case for x_2 the second coordinate of a point in $M(q,a,b) \cap M(q',a',b')$,

$$\frac{2}{10^{10}} \geq \left|x_2 - \frac{b}{q}N^2\right| + \left|x_2 - \frac{b'}{q'}N^2\right| \geq \frac{|b'q - bq'|N^2}{qq'} \geq \frac{N^2}{qq'} \geq N^{2/3},$$

which is another contradiction. ■

Lemma 3.2. *For each $(x_1, x_2) \in M(q,a,b)$, $|F(x_1,x_2)| \sim N/q^{1/2}$, here meaning within a factor of* 4.

Proof. This follows from Proposition 13.4 in [7]. ■

Proposition 3.3. *Let $R^\beta < \alpha^2 \leq R$ be given. There exists $v \in [0, N^2]$ satisfying*

$$|\{(x_1,x_2) \in [0,R]^2 : |F(x_1, x_2 + v))| \geq \alpha\}| \gtrsim \frac{R^2N^3}{\alpha^6}.$$

Proof. First note that, by the N-periodicity in x_1,

$$\begin{aligned}|\{(x_1,x_2) \in [0,R]^2 &: |F(x_1,x_2+v))| \geq \alpha\}| \\ &\gtrsim \frac{R}{N}|\{(x_1,x_2) \in ([0,N]\times[0,R]) : |F(x_1,x_2+v))| \geq \alpha\}|.\end{aligned}$$

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 3.2, it suffices to count the tuples (q,a,b) satisfying (3.1), $q \leq N^2/(16\alpha^2)$, and $|\frac{b}{q}N^2 - v| \leq R$, where v is to be determined. Begin by considering the distribution of points b/q in $[0,1]$, where $1 \leq b \leq q \sim N^2/\alpha^2$, $(b,q) = 1$. As in the proof of Lemma 3.1, if $b/q \neq b'/q'$, then $|b/q - b'/q'| \gtrsim \alpha^2/N^4$. There are $\gtrsim \sum_{q\sim N^2/\alpha^2}\varphi(q)$ many unique points b/q in $[0,1]$ satisfying $1 \leq b \leq q \sim N^2/\alpha^2$, $(b,q) = 1$, with φ denoting the Euler totient function. Use Theorem 3.7 in [1] to estimate $\sum_{q\sim N^2/\alpha^2}\varphi(q) \sim N^4/\alpha^4$, as long as N/α is larger than some absolute constant. By the pigeonhole principle, there exists some R/N^2 interval $I \subset [0,1]$ containing $\sim \lceil \frac{N^4}{\alpha^4}\frac{R}{N^2}\rceil$ many points b/q with $1 \leq b \leq q \sim N^2/\alpha^2$ and $(b,q) = 1$. There are also $\sim N^2/\alpha^2$ many choices for a to complete the tuple (q,a,b) satisfying (3.1). Let c denote the center of I and take $v = cN^2$ in the proposition statement to conclude that

$$|\{(x_1,x_2) \in ([0,N]\times[0,R]) : |F(x_1,x_2+v))| \geq \alpha\}| \gtrsim \frac{RN^4}{\alpha^6},$$

which finishes the proof. ■

Note that Proposition 3.3 shows the sharpness of Theorem 2.1 in the range $R^\beta < \alpha \leq R$ since

$$\frac{R^{2\beta}}{\alpha^6}\sum_\gamma \|F_\gamma\|_2^2 \sim \frac{R^{2\beta}}{\alpha^6}R^\beta R^2 = \frac{N^3R^2}{\alpha^6}.$$

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 $\alpha^2 < R^\beta$,

$$|U_\alpha \cap [-R,R]^2| \geq |U_{R^{\beta/2}} \cap [-R,R]^2| \gtrsim \frac{R^{2\beta}}{(R^{\beta/2})^6} \sum_\gamma \|F_\gamma\|_2^2 \sim R^2.$$

4. Implications of Theorem 2.2

Proof of Theorem 2.1 from Theorem 2.2. First suppose that $\alpha^2 > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}$. Then

$$\max_s \lambda(s^{-1}R^{-1})\,\lambda(s) \lesssim \max_s (s^{-1}R^{-1}R^\beta)(sR^\beta) = R^{2\beta-1}$$
$$\leq \begin{cases} R^{2\beta-1} & \text{if } \alpha^2 > R, \\ R^{2\beta}/\alpha^2 & \text{if } R^\beta \leq \alpha^2 \leq R. \end{cases}$$

Now suppose that $\alpha^2 \leq \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}$. Then

$$\frac{\lambda(1)^2}{\alpha^2} \lesssim \begin{cases} R^{2\beta-1} & \text{if } \alpha^2 > R, \\ R^{2\beta}/\alpha^2 & \text{if } R^\beta \leq \alpha^2 \leq R. \end{cases}$$ ■

Proof of Corollary 2.3 from Theorem 2.2. To see how this corollary follows from Theorem 2.2, first use an analogous series of pigeonholing steps as in Section 5 of [9] to reduce to the case where $\|f_\gamma\|_\infty \lesssim 1$ for all γ and there exists $C > 0$ so that $\|f_\gamma\|_p^p$ is either 0 or comparable to C for all γ. Split the integral

$$\int |f|^p = \sum_{R^{-1000} \leq \alpha \lesssim R^\beta} \int_{U_\alpha} |f|^p + \int_{|f|<R^{-1000}} |f|^p,$$

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 2.3 follows from showing that

$$\alpha^p |U_\alpha| \leq C_\varepsilon R^\varepsilon (R^{\beta p(1-1/q)-(1+\beta)} + R^{\beta p(1/2-1/q)})\, \lambda(1)^{p/q-1} \sum_\gamma \|f_\gamma\|_2^2,$$

where f satisfies the hypotheses of Theorem 2.2. The full range $3/p + 1/q \leq 1$ follows from p in the critical range $4 \leq p \leq 6$, which we treat first.

Case $4 \leq p \leq 6$.

There are two cases depending on which upper bound is larger in Theorem 2.2.

First we assume the L^4 bound holds, in which case

$$\begin{aligned}\alpha^p|U_\alpha| &\le C_\varepsilon R^\varepsilon \alpha^{p-4}\max_s \lambda(s^{-1}R^{-1})\,\lambda(s)\sum_\gamma \|f_\gamma\|_2^2\\ &\sim C_\varepsilon R^\varepsilon \frac{\alpha^{p-4}}{\lambda(1)^{p/q-1}}\max_s \lambda(s^{-1}R^{-1})\,\lambda(s)\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}\\ &\lesssim C_\varepsilon R^\varepsilon \frac{\lambda(1)^{p-4}}{\lambda(1)^{p/q-1}}\max_s (R^\beta s^{-1}R^{-1})(R^\beta s)\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}\\ &\lesssim C_\varepsilon R^\varepsilon \lambda(1)^{p(1-1/q)-3}R^{2\beta-1}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}.\end{aligned}$$

Since $p(1-1/q)-3\ge 0$, we may use the bound $\lambda(1)\lesssim R^\beta$ to conclude that

$$\lambda(1)^{p(1-1/q)-3}R^{2\beta-1}\le R^{\beta p(1-1/q)-3\beta+2\beta-1}=R^{\beta p(1-1/q)-(1+\beta)}.$$

The other case is that the L^6 bound holds in Theorem 2.2. We may also assume that $\alpha^2>\lambda(1)$ since otherwise we trivially have

$$\begin{aligned}\alpha^p|U_\alpha| &\le \lambda(1)^{p/2-1}\sum_\gamma\|f_\gamma\|_2^2\sim \lambda(1)^{p/2-1+1-p/q}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}\\ &\lesssim R^{\beta p(1/2-1/q)}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}\end{aligned}$$

where we used that $q\ge 2$ since $4\le p\le 6$ and $3/p+1/q\le 1$. Now using the assumptions $\alpha^2>\lambda(1)$ and $p\le 6$, we have

$$\begin{aligned}\alpha^p|U_\alpha| &\le C_\varepsilon R^\varepsilon\alpha^{p-6}\lambda(1)^2\lambda(1)^{1-p/q}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}\\ &\sim C_\varepsilon R^\varepsilon\lambda(1)^{p(1/2-1/q)}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}\lesssim C_\varepsilon R^\varepsilon R^{\beta p(1/2-1/q)}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}.\end{aligned}$$

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

$$\alpha^p|U_\alpha|\le R^{\frac{\beta}{2}(p-2)}\sum_\gamma\|f_\gamma\|_2^2\sim R^{\frac{\beta}{2}(p-2)}\lambda(1)^{1-p/q}\Big(\sum_\gamma \|f_\gamma\|_p^q\Big)^{p/q}.$$

Since in this subcase, $1-p/q\ge 1-(p-3)>0$, we are done after noting that

$$R^{\frac{\beta}{2}(p-2)}\lambda(1)^{1-p/q}\le R^{\beta p(1/2-1/q)}.$$

Now assume that $\alpha\ge R^{\beta/2}$ and use the $p=4$ case above (noting that $R^{4\beta(1-1/q)-(1+\beta)}\le R^{4\beta(1/2-1/q)}$) to get

$$\begin{aligned}\alpha^p|U_\alpha| &\le \frac{\alpha^4}{(R^{\beta/2})^{4-p}}|U_\alpha|\le R^{-\frac{\beta}{2}(4-p)}C_\varepsilon R^\varepsilon R^{4\beta(1/2-1/q)}\lambda(1)^{4/q-1}\sum_\gamma\|f_\gamma\|_2^2\\ &\le C_\varepsilon R^\varepsilon R^{\beta p(1/2-1/q)}\lambda(1)^{p/q-1}\sum_\gamma\|f_\gamma\|_2^2.\end{aligned}$$

Case $6 < p$.

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

$$\begin{aligned}\alpha^p |U_\alpha| &\le \lambda(1)^{p-6}\,\alpha^6\,|U_\alpha| \le \lambda(1)^{p-6}\,C_\varepsilon\,R^\varepsilon\,R^{6\beta(1-1/q)-(1+\beta)}\,\lambda(1)^{6/q-1}\sum_\gamma \|f_\gamma\|_2^2\\ &= \Big(\frac{\lambda(1)}{R^\beta}\Big)^{(p-6)(1-1/q)}\,C_\varepsilon\,R^\varepsilon\,R^{p\beta(1-1/q)-(1+\beta)}\,\lambda(1)^{p/q-1}\sum_\gamma \|f_\gamma\|_2^2\\ &\le C_\varepsilon\,R^\varepsilon\,R^{p\beta(1-1/q)-(1+\beta)}\,\lambda(1)^{p/q-1}\sum_\gamma \|f_\gamma\|_2^2. \qquad \blacksquare\end{aligned}$$

5. Tools to prove Theorem 2.2

The proof of Theorem 2.2 follows the high/low frequency decomposition and the pruning approach from [9]. 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 $\beta \in [1/2, 1]$ and $R \ge 2$. The parameter $\alpha > 0$ describes the superlevel set

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

For $\varepsilon > 0$, we analyze scales $R_k = R^{k\varepsilon}$, noting that $R^{-1/2} \le R_k^{-1/2} \le 1$. 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^ε, 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 (2.2).

(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 θ be a union of adjacent γ.

(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 $\gamma \subset \tau_k$. Note that $\{\tau_N\} = \{\theta\}$.

We will repeatedly make use of the hypothesis that f is a Schwartz function with Fourier transform supported in $\mathcal{N}_{R^{-1}}(\mathbb{P}^1)$ and satisfies $\|f_\gamma\|_\infty \le 1$ for all $\gamma \in \mathcal{P}(R, \beta)$.

5.1. A pruning step

We will define wave packets at each scale τ_k, and prune the wave packets associated to f_{τ_k} according to their amplitudes.

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

$$\{x \in \mathbb{R}^2 : |x \cdot \xi| \le 1,\ \forall \xi \in \tau_k\}.$$

Let $\mathbb{T}_{\tau_k}$ be the collection of tubes T_{τ_k} which are dual to τ_k, contain τ_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 $[-1/4, 1/4]^2$. For each $m \in \mathbb{Z}^2$, let

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

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

$$\psi_m(x) \le c \int_{[0,1]^2} \frac{C_n}{(1 + |x - y - m|^2)^n}\, dy \le \frac{\tilde{C}_n}{(1 + |x - m|^2)^n}.$$

Define the partition of unity $\psi_{T_{\tau_k}}$ associated to τ_k to be $\psi_{T_{\tau_k}}(x) = \psi_m \circ A_{\tau_k}$, where A_{τ_k} is a linear transformation taking τ_k^* to $[-1/2, 1/2]^2$ and $A_{\tau_k}(T_{\tau_k}) = m + [-1/2, 1/2]^2$. The important properties of $\psi_{T_{\tau_k}}$ are (1) rapid decay off of T_{τ_k} and (2) Fourier support contained in $\frac{1}{2}\tau_k$.

To prove upper bounds for the size of U_α, 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 5.9 below.

Partition $\mathbb{T}_\theta = \mathbb{T}_\theta^g \sqcup \mathbb{T}_\theta^b$ into a "good" and a "bad" set as follows. Let $\delta > 0$ be a parameter to be chosen in §6.2 and set

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

where $M > 0$ is a universal constant we will choose in the proof of Proposition 6.1.

Definition 5.1 (Pruning with respect to τ_k). For each θ and τ_{N-1}, define the notation

$$f_\theta^N = \sum_{T_\theta \in \mathbb{T}_\theta^g} \psi_{T_\theta} f_\theta \quad \text{and} \quad f_{\tau_{N-1}}^N = \sum_{\theta \subset \tau_{N-1}} f_\theta^N.$$

For each $k < N$, let

$$\mathbb{T}_{\tau_k}^g = \{T_{\tau_k} \in \mathbb{T}_{\tau_k} : \|\psi_{T_{\tau_k}} f_{\tau_k}^{k+1}\|_{L^\infty(\mathbb{R}^2)} \le R^{M\delta} \lambda(1)/\alpha\},$$
$$f_{\tau_k}^k = \sum_{T_{\tau_k} \in \mathbb{T}_{\tau_k}^g} \psi_{T_{\tau_k}} f_{\tau_k}^{k+1} \quad \text{and} \quad f_{\tau_{k-1}}^k = \sum_{\tau_k \subset \tau_{k-1}} f_{\tau_k}^k.$$

For each k, define the kth version of f to be $f^k = \sum_{\tau_k} f_{\tau_k}^k$.

Lemma 5.2 (Properties of f^k). (1) $|f_{\tau_k}^k(x)| \le |f_{\tau_k}^{k+1}(x)| \le \#\gamma \subset \tau_k$.

(2) $\|f_{\tau_k}^k\|_{L^\infty} \le C_\varepsilon R^{O(\varepsilon)} R^{M\delta} \lambda(1)/\alpha$.

(3) $\operatorname{supp} \widehat{f_{\tau_k}^k} \subset 2\tau_k$.

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

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

$$f_{\tau_k}^k=\sum_{T_{\tau_k}\in\mathbb{T}_{\tau_k}^g}\psi_{T_{\tau_k}}f_{\tau_k}^{k+1}.$$

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

$$|f_{\tau_k}^k|\le|f_{\tau_k}^{k+1}|\le\sum_{\tau_{k+1}\subset\tau_k}|f_{\tau_{k+1}}^{k+1}|\le\cdots\le\sum_{\tau_N\subset\tau_k}|f_{\tau_N}^N|\le\sum_{\theta\subset\tau_k}|f_\theta|\le\sum_{\gamma\subset\tau_k}|f_\gamma|\lesssim\#\gamma\subset\tau_k,$$

where we used the assumption $\|f_\gamma\|_\infty\lesssim1$ for all γ. Now consider the L^∞ bound in the second property. We write

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

The first sum has at most $CR^{2\varepsilon}$ terms, and each term has norm bounded by $R^{M\delta}\lambda(1)/\alpha$, by the definition of $\mathbb{T}_{\tau_k}^g$. By property (1), we may trivially bound $f_{\tau_k}^{k+1}$ by $R\max_\gamma\|f_\gamma\|_\infty$. But if $x\notin R^\varepsilon T_{\tau_k}$, then $\psi_{T_{\tau_k}}(x)\le R^{-1000}$. Thus

$$\Bigg|\sum_{\substack{T_{\tau_k}\in\mathbb{T}_{\tau_k}^g,\\ x\notin R^\varepsilon T_{\tau_k}}}\psi_{T_{\tau_k}}f_{\tau_k}^{k+1}\Bigg|\le\sum_{\substack{T_{\tau_k}\in\mathbb{T}_{\tau_k}^g,\\ x\notin R^\varepsilon T_{\tau_k}}}R^{-500}\psi_{T_{\tau_k}}^{1/2}(x)\|f_{\tau_k}^{k+1}\|_\infty\le R^{-250}\max_\gamma\|f_\gamma\|_\infty.$$

Since $\alpha\lesssim|f(x)|\lesssim\sum_\gamma\|f_\gamma\|_\infty\lesssim\lambda(1)$, (recalling the assumption that each $\|f_\gamma\|_\infty\lesssim1$), we note $R^{-250}\le CR^{2\varepsilon}\lambda(1)/\alpha$.

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

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

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

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

has Fourier support in $\frac{1}{2}\tau_{N-1}+(1+(\log R)^{-1})\tau_{N-1}\subset2\tau_{N-1}$. Iterating this reasoning until $k=1$ gives (3) and (4). ∎

Definition 5.3. For each τ_k, let w_{τ_k} be the weight function adapted to τ_k^* defined by

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

where

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

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

$$w_s(x) = \frac{c'}{(1 + |x|^2/s^2)^{10}}, \qquad \|w_s\|_1 = 1. \tag{5.1}$$

The weights w_{τ_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, $f = f * \widecheck{\varphi_A}$. Since $|\widecheck{\varphi_A}|$ is an L^1-normalized function which is positive on a set dual to A, $|f| * |\widecheck{\varphi_A}|$ is an averaged version of $|f|$ over a dual set A^*. We record some of the specific locally constant properties we need in the following lemma.

Lemma 5.4 (Locally constant property). *For each τ_k and $T_{\tau_k} \in \mathbb{T}_{\tau_k}$,*

$$\|f_{\tau_k}\|^2_{L^\infty(T_{\tau_k})} \lesssim |f_{\tau_k}|^2 * w_{\tau_k}(x) \qquad \text{for any } x \in T_{\tau_k}.$$

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

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

Because the pruned versions of f and f_{τ_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 5.4. Let ρ_{τ_k} be a bump function equal to 1 on τ_k and supported in $2\tau_k$. Then using Fourier inversion and Hölder's inequality,

$$|f_{\tau_k}(y)|^2 = |f_{\tau_k} * \widecheck{\rho_{\tau_k}}(y)|^2 \le \|\widecheck{\rho_{\tau_k}}\|_1 \, |f_{\tau_k}|^2 * |\widecheck{\rho_{\tau_k}}|(y).$$

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

$$\begin{aligned} \sup_{x \in T_{\tau_k}} |f_{\tau_k}|^2 * w_{\tau_k}(x) &\le \int |f_{\tau_k}|^2(y) \sup_{x \in T_{\tau_k}} w_{\tau_k}(x - y)\, dy \\ &\sim \int |f_{\tau_k}|^2(y)\, w_{\tau_k}(x - y)\, dy, \quad \text{for all } x \in T_{\tau_k}. \end{aligned}$$

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

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

where we used that each ρ_{θ_s} is a translate of a single function $\rho_{s^{-1}}$. The rest of the argument is analogous to the first part. ∎

Definition 5.5 (Auxiliary functions). Let $\varphi(x)\colon \mathbb{R}^2 \to [0,\infty)$ be a radial, smooth bump function satisfying $\varphi(x) = 1$ on B_1 and $\operatorname{supp}\varphi \subset B_2$. Observe that

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

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

$$\eta_{\sim s}(\xi) = \varphi(2^j\xi) - \varphi(2^{j+1}\xi) \quad \text{and} \quad \eta_{<\lceil R^\beta\rceil^{-1}}(\xi) = \varphi(2^{J+1}\xi).$$

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

$$\eta_k(\xi) = \varphi(R_{k+1}^{1/2} x).$$

Definition 5.6. Let $G(x) = \sum_\theta |f_\theta|^2 * w_\theta$, $G^\ell(x) = G * \check{\eta}_{<\lceil R^\beta\rceil^{-1}}$, and $G^h(x) = G(x) - G^\ell(x)$. For $k = 1, \dots, N-1$, let

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

Definition 5.7. Define the high set as

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

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

$$\Omega_k = \{x \in B_R \setminus H : g_k \le 2|g_k^h|, g_{k+1} \le 2|g_{k+1}^\ell|, \dots, g_N \le 2|g_N^\ell|\}$$

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

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

5.2. High/low frequency lemmas

Lemma 5.8 (Low lemma). *For each x, $|G^\ell(x)| \lesssim \lambda(1)$ and $|g_k^\ell(x)| \lesssim g_{k+1}(x)$.*

Proof. For each θ, by Plancherel's theorem,

$$\begin{aligned}|f_\theta|^2 * \check{\eta}_{<\lceil R^\beta\rceil^{-1}}(x) &= \int_{\mathbb{R}^2} |f_\theta|^2(x-y)\, \check{\eta}_{<\lceil R^\beta\rceil^{-1}}(y)\, dy \\ &= \int_{\mathbb{R}^2} \widehat{f}_\theta * \widehat{\overline{f}}_\theta(\xi)\, e^{-2\pi i x\cdot\xi}\, \eta_{<\lceil R^\beta\rceil^{-1}}(\xi)\, d\xi \\ &= \sum_{\gamma,\gamma'\subset\theta} \int_{\mathbb{R}^2} e^{-2\pi i x\cdot\xi}\, \widehat{f}_\gamma * \widehat{\overline{f}}_{\gamma'}(\xi)\, \eta_{<\lceil R^\beta\rceil^{-1}}(\xi)\, d\xi.\end{aligned}$$

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

Then

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

Use Plancherel's theorem again to get back to a convolution in x and conclude that

$$|G * \check{\eta}_{<\lceil R^\beta\rceil^{-1}}(x)| = \Big|\sum_\theta \sum_{\substack{\gamma,\gamma'\subset\theta\\ \gamma\sim\gamma'}} (f_\gamma \overline{f}_{\gamma'}) * w_\theta * \check{\eta}_{<\lceil R^\beta\rceil^{-1}}(x)\Big|$$
$$\lesssim \sum_\theta\sum_{\gamma\subset\theta}|f_\gamma|^2 * w_\theta * |\check{\eta}_{<\lceil R^\beta\rceil^{-1}}|(x) \lesssim \sum_\gamma \|f_\gamma\|_\infty^2 \lesssim \lambda(1).$$

By an analogous argument as above, we have that

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

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

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

since $\tau_k^* \subset \tau_{k+1}^*$ and $\check{\eta}_k$ is an L^1-normalized function that is rapidly decaying away from $B_{R_{k+1}^{1/2}}(0)$.

∎

Lemma 5.9 (Pruning lemma). *For any τ,*

$$\Big|\sum_{\tau_k\subset\tau} f_{\tau_k} - \sum_{\tau_k\subset\tau} f_{\tau_k}^{k+1}(x)\Big| \le C_\varepsilon R^{-M\delta}\alpha \quad \textit{for all } x\in\Omega_k,$$
$$\Big|\sum_{\tau_1\subset\tau} f_{\tau_1} - \sum_{\tau_1\subset\tau} f_{\tau_1}^{1}(x)\Big| \le C_\varepsilon R^{-M\delta}\alpha \quad \textit{for all } x\in L.$$

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

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

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

For each $m \geq k+1$ and τ_m,

$$|f_{\tau_m}^m(x)-f_{\tau_m}^{m+1}(x)|=\Big|\sum_{T_{\tau_m}\in\mathbb{T}_{\tau_m}^b}\psi_{T_{\tau_m}}(x)f_{\tau_m}^{m+1}(x)\Big|=\sum_{T_{\tau_m}\in T_{\tau_m}^b}|\psi_{T_{\tau_m}}^{1/2}(x)f_{\tau_m}^{m+1}(x)|\psi_{T_{\tau_m}}^{1/2}(x)$$
$$\lesssim \sum_{T_{\tau_m}\in\mathbb{T}_{\tau_m}^b} R^{-M\delta}\frac{\alpha}{\lambda(1)\lambda}\|\psi_{T_{\tau_m}}f_{\tau_m}^{m+1}\|_{L^\infty(\mathbb{R}^2)}\|\psi_{T_{\tau_m}}^{1/2}f_{\tau_m}^{m+1}\|_{L^\infty(\mathbb{R}^2)}\psi_{T_{\tau_m}}^{1/2}(x)$$
$$\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}\sum_{T_{\tau_m}\in\mathbb{T}_{\tau_m}^b}\|\psi_{T_{\tau_m}}^{1/2}f_{\tau_m}^{m+1}\|_{L^\infty(\mathbb{R}^2)}^2\psi_{T_{\tau_m}}^{1/2}(x)$$
$$\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}\sum_{T_{\tau_m}\in\mathbb{T}_{\tau_m}^b}\sum_{\tilde{T}_{\tau_m}}\|\psi_{T_{\tau_m}}|f_{\tau_m}^{m+1}|^2\|_{L^\infty(\tilde{T}_{\tau_m})}\psi_{T_{\tau_m}}^{1/2}(x)$$
$$\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}\sum_{T_{\tau_m},\tilde{T}_{\tau_m}\in\mathbb{T}_{\tau_m}}\|\psi_{T_{\tau_m}}\|_{L^\infty(\tilde{T}_{\tau_m})}\||f_{\tau_m}^{m+1}|^2\|_{L^\infty(\tilde{T}_{\tau_m})}\psi_{T_{\tau_m}}^{1/2}(x).$$

Let $c_{\tilde{T}_{\tau_m}}$ denote the center of $\tilde{T}_{\tau_m}$, and note the pointwise inequality

$$\sum_{T_{\tau_m}}\|\psi_{T_{\tau_m}}\|_{L^\infty(\tilde{T}_{\tau_m})}\psi_{T_{\tau_m}}^{1/2}(x)\lesssim R_m^{3/2}\,w_{\tau_m}(x-c_{\tilde{T}_{\tau_m}}),$$

which means that

$$|f_{\tau_m}^m(x)-f_{\tau_m}^{m+1}(x)|\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}R_m^{3/2}\sum_{\tilde{T}_{\tau_m}\in T_{\tau_m}}w_{\tau_m}(x-c_{\tilde{T}_{\tau_m}})\||f_{\tau_m}^{m+1}|^2\|_{L^\infty(\tilde{T}_{\tau_m})}$$
$$\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}R_m^{3/2}\sum_{\tilde{T}_{\tau_m}\in T_{\tau_m}}w_{\tau_m}(x-c_{\tilde{T}_{\tau_m}})|f_{\tau_m}^{m+1}|^2*w_{\tau_m}(c_{\tilde{T}_{\tau_m}})$$
$$\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}|f_{\tau_m}^{m+1}|^2*w_{\tau_m}(x),$$

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

$$\Big|\sum_{\tau_m\subset\tau}f_{\tau_m}^m(x)-f_{\tau_m}^{m+1}(x)\Big|\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}\sum_{\tau_m\subset\tau}|f_{\tau_m}^{m+1}|^2*w_{\tau_m}(x)\lesssim R^{-M\delta}\frac{\alpha}{\lambda(1)}g_m(x).$$

By the definition of Ω_k and Lemma 5.8, we have that $g_m(x)\leq 2|g_m^\ell(x)|\leq 2Cg_{m+1}(x)\leq\cdots\leq(2C)^{\varepsilon^{-1}}G(x)\lesssim(2C)^{\varepsilon^{-1}}\lambda(1)$. We conclude that

$$\Big|\sum_{\tau_m\subset\tau}f_{\tau_m}^m(x)-f_{\tau_m}^{m+1}(x)\Big|\lesssim(2C)^{\varepsilon^{-1}}R^{-M\delta}\alpha.$$

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

Definition 5.10. Call the distribution function λ associated to a function f (R,ε)-normalized if for any k and τ_m,

$$\#\{\tau_k\subset\tau_m : f_{\tau_k}\neq 0\}\leq 100\,\frac{\lambda(R_m^{-1/2})}{\lambda(R_k^{-1/2})}.$$

Remark 5.11. The role of (R, ε)-normalized distribution functions is to simplify notation. It allows us to write all combinatorial quantities which arise in the high lemmas in terms of the maximal number of γ intersecting larger arcs, rather than counting the number of intermediate-scale blocks intersecting larger arcs.

Lemma 5.12 (High lemma I). *Assume that f has an (R, ε)-normalized distribution function $\lambda(\cdot)$. For each dyadic s, $R^{-\beta} \le s \le R^{-1/2}$,*

$$\int_{\mathbb{R}^2} |G * \check{\eta}_{\sim s}|^2 \lesssim C_\varepsilon R^{2\varepsilon} \lambda(s^{-1}R^{-1})\lambda(s) \sum_\gamma \|f_\gamma\|_2^2.$$

Proof. Organize the $\{\gamma\}$ into subcollections $\{\theta_s\}$ in which each θ_s is a union of γ which intersect the same $\sim s$-arc of $\mathbb{P}^1$, where here, for concreteness, $\sim s$ means within a factor of 2. Then by Plancherel's theorem, since $\overline{\check{\eta}}_{\sim s} = \check{\eta}_{\sim s}$, we have for each θ,

$$\begin{aligned} |f_\theta|^2 * \check{\eta}_{\sim s}(x) &= \int_{\mathbb{R}^2} |f_\theta|^2(x-y)\check{\eta}_{\sim s}(y)\,dy = \int_{\mathbb{R}^2} \widehat{f_\theta} * \widehat{\overline{f}}_\theta(\xi)\, e^{-2\pi i x\cdot\xi}\eta_{\sim s}(\xi)\,d\xi \\ &= \sum_{\theta_s, \theta_s' \subset \theta} \int_{\mathbb{R}^2} e^{-2\pi i x\cdot\xi}\, \widehat{f}_{\theta_s} * \widehat{\overline{f}}_{\theta_s'}(\xi)\,\eta_{\sim s}(\xi)\,d\xi. \end{aligned} \tag{5.2}$$

The support of $\widehat{\overline{f}}_{\theta_s'}(\xi) = \int e^{-2\pi i x\cdot\xi}\overline{f}_{\theta_s'}(x)\,dx = \overline{\widehat{f}}_{\theta_s'}(-\xi)$ is contained in $-\theta_s'$. This means that the support of $\widehat{f}_{\theta_s} * \widehat{\overline{f}}_{\theta_s'}(\xi)$ is contained in $\theta_s - \theta_s'$. 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 $\theta_s' \subset \theta$ so that the integral in (5.2) is nonzero. Thus we may write

$$G * \check{\eta}_{\sim s}(x) = \sum_\theta |f_\theta|^2 * w_\theta * \check{\eta}_{\sim s}(x) = \sum_\theta \sum_{\substack{\theta_s, \theta_s' \subset \theta \\ \theta_s \sim \theta_s'}} (f_{\theta_s}\overline{f}_{\theta_s'}) * w_\theta * \check{\eta}_{\sim s}(x),$$

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

$$\begin{aligned} \int_{\mathbb{R}^2} |G * \check{\eta}_{\sim s}|^2 &= \int_{\mathbb{R}^2} \Big| \sum_\theta \sum_{\substack{\theta_s, \theta_s' \subset \theta \\ \theta_s \sim \theta_s'}} (f_{\theta_s}\overline{f}_{\theta_s'}) * w_\theta * \check{\eta}_{\sim s} \Big|^2 \\ &= \int_{\mathbb{R}^2} \Big| \sum_\theta \sum_{\substack{\theta_s, \theta_s' \subset \theta \\ \theta_s \sim \theta_s'}} (\widehat{f_{\theta_s}} * \widehat{\overline{f}}_{\theta_s'})\,\widehat{w}_\theta\,\eta_{\sim s} \Big|^2. \end{aligned}$$

For each θ, $\sum_{\theta_s, \theta_s' \subset \theta, \theta_s \sim \theta_s'} (\widehat{f_{\theta_s}} * \widehat{\overline{f}}_{\theta_s'})$ is supported in $\theta - \theta$, since each summand is supported in $\theta_s - \theta_s'$ and $\theta_s, \theta_s' \subset \theta$. For each $\xi \in \mathbb{R}^2$, $|\xi| > s/2$, the maximum number of $\theta - \theta$ containing ξ is bounded by the maximum number of θ intersecting an $R^{-1/2} \cdot s^{-1} R^{-1/2}$-arc of the parabola. Using that $\lambda(\cdot)$ is (R, ε)-normalized, this number is bounded above by $C_\varepsilon R^\varepsilon \lambda(s^{-1}R^{-1})/\lambda(R^{-1/2})$.

Since $\eta_{\sim s}$ is supported in the region $|\xi| > s/2$, by Cauchy–Schwarz,

$$\int_{\mathbb{R}^2} \Big| \sum_{\theta} \sum_{\substack{\theta_s,\theta_s'\subset\theta \\ \theta_s\sim\theta_s'}} (\widehat{f_{\theta_s}} * \widehat{\overline{f}_{\theta_s'}})\, \widehat{w}_\theta\, \eta_{\sim s}\Big|^2$$

$$\lesssim C_\varepsilon R^\varepsilon \frac{\lambda(r^{-1}R^{-1})}{\lambda(R^{-1/2})} \sum_\theta \int_{\mathbb{R}^2} \Big| \sum_{\substack{\theta_s,\theta_s'\subset\theta \\ \theta_s\sim\theta_s'}} (\widehat{f_{\theta_s}} * \widehat{\overline{f}_{\theta_s'}})\, \widehat{w}_\theta \eta_{\sim s}\Big|^2$$

$$= C_\varepsilon R^\varepsilon \frac{\lambda(r^{-1}R^{-1})}{\lambda(R^{-1/2})} \sum_\theta \int_{\mathbb{R}^2} \Big| \sum_{\substack{\theta_s,\theta_s'\subset\theta \\ \theta_s\sim\theta_s'}} (f_{\theta_s} \overline{f}_{\theta_s'}) * w_\theta * \check{\eta}_{\sim s}\Big|^2$$

$$\lesssim C_\varepsilon R^\varepsilon \frac{\lambda(r^{-1}R^{-1})}{\lambda(R^{-1/2})} \sum_\theta \int_{\mathbb{R}^2} \Big| \sum_{\theta_s\subset\theta} |f_{\theta_s}|^2 * w_\theta * |\check{\eta}_{\sim s}|\Big|^2.$$

It remains to analyze each of the integrals above:

$$\int_{\mathbb{R}^2} \Big| \sum_{\theta_s\subset\theta} |f_{\theta_s}|^2 * w_\theta * |\check{\eta}_{\sim s}|\Big|^2 \lesssim \Big\| \sum_{\theta_s\subset\theta} |f_{\theta_s}|^2 * w_\theta * |\check{\eta}_{\sim s}|\Big\|_\infty \int_{\mathbb{R}^2} \sum_{\theta_s\subset\theta} |f_{\theta_s}|^2 * w_\theta * |\check{\eta}_{\sim s}|.$$

Bound the L^∞ norms using the assumption that $\|f_\gamma\|_\infty \lesssim 1$ for all γ:

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

Finally, using Young's convolution inequality and the L^2-orthogonality of the f_γ, we have

$$\int_{\mathbb{R}^2} \sum_{\theta_s\subset\theta} |f_{\theta_s}|^2 * w_\theta * |\check{\eta}_{\sim s}| \lesssim \int_{\mathbb{R}^2} \sum_{\theta_s\subset\theta} |f_{\theta_s}|^2 = \sum_{\gamma\subset\theta} \|f_\gamma\|_2^2. \qquad \blacksquare$$

Lemma 5.13 (High lemma II). *For each k,*

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

Proof. By Plancherel's theorem, we have

$$\int_{\mathbb{R}^2} |g_k^h|^2 = \int_{\mathbb{R}^2} |g_k - g_k^\ell|^2 = \int_{\mathbb{R}^2} \Big| \sum_{\tau_k} (\widehat{f_{\tau_k}^{k+1}} * \widehat{f_{\tau_k}^{k+1}}) \widehat{w}_{\tau_k} - \sum_{\tau_k} (\widehat{f_{\tau_k}^{k+1}} * \widehat{f_{\tau_k}^{k+1}}) \widehat{w}_{\tau_k} \eta_k \Big|^2$$

$$\le \int_{|\xi|>cR_{k+1}^{-1/2}} \Big| \sum_{\tau_k} (\widehat{f_{\tau_k}^{k+1}} * \widehat{f_{\tau_k}^{k+1}}) \widehat{w}_{\tau_k}\Big|^2$$

since $(1-\eta_k)$ is supported in the region $|\xi| > cR_{k+1}^{-1/2}$ for some constant $c > 0$. For each τ_k, $\widehat{f_{\tau_k}^{k+1}} * \widehat{f_{\tau_k}^{k+1}}$ is supported in $2\tau_k - 2\tau_k$, using property (4) of Lemma 5.2, and the maximum overlap of the sets $\{2\tau_k - 2\tau_k\}$ in the region $|\xi| \ge cR_{k+1}^{-1/2}$ is bounded by $\sim R_k^{-1/2}/R_{k+1}^{-1/2} \lesssim R^\varepsilon$.

Thus, using Cauchy–Schwarz,

$$\int_{|\xi|>cR_{k+1}^{-1/2}}\Big|\sum_{\tau_k}(\widehat{f_{\tau_k}^{k+1}}*\widehat{\overline{f_{\tau_k}^{k+1}}})\,\widehat{w}_{\tau_k}\Big|^2\lesssim R^\varepsilon\sum_{\tau_k}\int_{|\xi|>cR_{k+1}^{-1/2}}|(\widehat{f_{\tau_k}^{k+1}}*\widehat{\overline{f_{\tau_k}^{k+1}}})\,\widehat{w}_{\tau_k}|^2$$

$$\le R^\varepsilon\sum_{\tau_k}\int_{\mathbb{R}^2}|(\widehat{f_{\tau_k}^{k+1}}*\widehat{\overline{f_{\tau_k}^{k+1}}})\,\widehat{w}_{\tau_k}\big|^2=R^\varepsilon\sum_{\tau_k}\int_{\mathbb{R}^2}||f_{\tau_k}^{k+1}|^2*w_{\tau_k}|^2\le R^{3\varepsilon}\sum_{\tau_{k+1}}\int_{\mathbb{R}^2}|f_{\tau_{k+1}}^{k+1}|^4,$$

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 inequality. ∎

5.3. Bilinear restriction

We will use the following version of a local bilinear restriction theorem, which follows from a standard Córdoba argument [6], included here for completeness.

Theorem 5.14. *Let $S\ge 4$, $1/2\ge D\ge S^{-1/2}$, and let $X\subset\mathbb{R}^2$ be any Lebesgue measurable set. Suppose that τ and τ' 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 $\sim S^{-1/2}\times S^{-1}$-blocks, we have*

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

In the following proof, the exact definition of the $\sim S^{-1}\times S^{-1}$ blocks θ_S is not important. However, by f_τ 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 φ_B be a smooth function satisfying $\varphi_B\gtrsim 1$ in B, φ_B decays rapidly away from B, and $\widehat{\varphi_B}$ 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 θ_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

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

By Placherel's theorem, each integral vanishes unless

$$(\theta_1-\theta_2)\cap\mathcal{N}_{S^{-1/2}}(\theta_1'-\theta_2')\neq\emptyset. \tag{5.3}$$

Next we check that the number of tuples $(\theta_1,\theta_2,\theta_1',\theta_2')$ (with θ_1,θ_2 having nonempty intersection with τ and θ_1',θ_2' having nonempty intersection with τ') satisfying (5.3) 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,

$$\begin{aligned} \xi^2 - (\xi')^2 &= 2\xi_1(\xi - \xi') \quad \text{for some } \xi < \xi_1 < \xi', \text{ and} \\ (\xi'')^2 - (\xi''')^2 &= 2\xi_2(\xi'' - \xi''') \quad \text{for some } \xi'' < \xi_2 < \xi'''. \end{aligned}$$

Since $(\xi_1, \xi_1^2) \in \tau$ and $(\xi_2, \xi_2^2) \in \tau'$, we also know that $|\xi_1 - \xi_2| \geq D$. Putting everything together, we have

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

if either $\operatorname{dist}((\xi, \xi^2), (\xi', (\xi')^2))$ or $\operatorname{dist}((\xi'', (\xi'')^2), (\xi''', (\xi''')^2))$ 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 (5.3). The conclusion is that

$$\begin{aligned} \sum_{\substack{\theta_i \cap \tau \neq \emptyset \\ \theta_i' \cap \tau' \neq \emptyset}} \int_{\mathbb{R}^2} f_{\theta_1} \overline{f}_{\theta_2} f_{\theta_2'} \overline{f}_{\theta_1'} \varphi_B &= \sum_{\substack{\theta_1 \cap \tau \neq \emptyset \\ \theta_1' \cap \tau' \neq \emptyset}} \sum_{\substack{d(\theta_1, \theta_2) \leq CD^{-1}S^{-1/2} \\ d(\theta_1', \theta_2') \leq CD^{-1}S^{-1/2}}} \int_{\mathbb{R}^2} f_{\theta_1} \overline{f}_{\theta_2} f_{\theta_2'} \overline{f}_{\theta_1'} \varphi_B \\ &\lesssim D^{-2} \int_{\mathbb{R}^2} \Big(\sum_\theta |f_\theta|^2 \Big)^2 \varphi_B. \end{aligned}$$

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

$$\begin{aligned} \int_{\mathbb{R}^2} \Big(\sum_\theta |f_\theta|^2 \Big)^2 \varphi_B &\leq \sum_{B'} |B| \Big\| \sum_\theta |f_\theta|^2 \Big\|_{L^\infty(B')}^2 \|\varphi_B\|_{L^\infty(B')} \\ &\leq |B| \Big(\sum_{B'} \Big\| \sum_\theta |f_\theta|^2 \Big\|_{L^\infty(B')} \|\varphi_B^{1/2}\|_{L^\infty(B')} \Big)^2 \\ &\lesssim |B| \Big(\sum_{B'} \sum_\theta |f_\theta|^2 * w_{S^{1/2}}(c_{B'}) \|\varphi_B^{1/2}\|_{L^\infty(B')} \Big)^2 \\ &\lesssim |B|^{-1} \Big(\int_{\mathbb{R}^2} \sum_\theta |f_\theta|^2 * w_{S^{1/2}}(y)\, \varphi_B^{1/2}(y)\, dy \Big)^2 \\ &\lesssim |B|^{-1} \Big(\int_B \sum_\theta |f_\theta|^2 * w_{S^{1/2}}(y)\, dy \Big)^2 \leq \int_B \Big(\sum_\theta |f_\theta|^2 * w_{S^{1/2}} \Big)^2, \end{aligned}$$

where we used that $w_{S^{1/2}} * \varphi_B^{1/2}(y) \lesssim w_{S^{1/2}} * \chi_B(y)$ in the second to last inequality. ∎

6. Proof of Theorem 2.2

Theorem 2.2 follows from the following proposition and a broad-narrow argument in §6.2. First we prove a version of Theorem 2.2 where U_α is replaced by a "broad" version of U_α.

6.1. The broad version of Theorem 2.2

Let $\delta > 0$ be a parameter we will choose in the broad/narrow analysis. With the notation $\ell(\tau) = s$ we mean that τ is an approximate $s \times s^2$ block which is part of a partition of $\mathcal{N}_{s^2}(\mathbb{P}^1)$. For two non-adjacent blocks τ and τ' satisfying $\ell(\tau) = \ell(\tau') = R^{-\delta}$, define the broad version of U_α to be

$$\mathrm{Br}_\alpha(\tau,\tau') = \{x \in \mathbb{R}^2 : \alpha \sim |f_\tau(x) f_{\tau'}(x)|^{1/2},\ \ (|f_\tau(x)| + |f_{\tau'}(x)|) \le R^{O(\delta)}\alpha\}. \tag{6.1}$$

Proposition 6.1. *Suppose that f satisfies the hypotheses of Theorem 2.2 and that has an (R,ε)-normalized distribution function $\lambda(\cdot)$. Then*

$$|\mathrm{Br}_\alpha(\tau,\tau')| \le C_{\varepsilon,\delta} R^\varepsilon R^{O(\delta)} \begin{cases} \frac{1}{\alpha^4} \max_s \lambda(s^{-1}R^{-1})\lambda(s) \sum_\gamma \|f_\gamma\|_2^2 & \text{if } \alpha^2 > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}, \\ \frac{\lambda(1)^2}{\alpha^6} \sum_\gamma \|f_\gamma\|_2^2 & \text{if } \alpha^2 \le \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}. \end{cases}$$

Proof of Proposition 6.1. (1) *Bounding $|Br_\alpha(\tau,\tau') \cap H|$.* Using bilinear restriction, given here by Theorem 5.14, we have

$$\begin{aligned} \alpha^4 |\mathrm{Br}_\alpha(\tau,\tau') \cap H| &\lesssim \sum_{\substack{\ell(\tau)=\ell(\tau)=R^{-\delta} \\ d(\tau,\tau') \gtrsim R^{-\delta}}} \int_{U_\alpha \cap H} |f_\tau|^2 |f_{\tau'}|^2 \\ &\lesssim R^{O(\delta)} \int_{\mathcal{N}_{R^{1/2}}(\mathrm{Br}_\alpha(\tau,\tau')\cap H)} \Big(\sum_\theta |f_\theta|^2 * w_{R^{1/2}} \Big)^2. \end{aligned}$$

By the locally constant property and the pointwise inequality $w_{R^{1/2}} * w_\theta \lesssim w_\theta$ for each θ, we have that

$$\sum_\theta |f_\theta|^2 * w_{R^{1/2}} \lesssim G(x).$$

Then

$$\begin{aligned} &\int_{\mathcal{N}_{R^{1/2}}(\mathrm{Br}_\alpha(\tau,\tau')\cap H)} |G(x)|^2\,dx \\ &\le \sum_{\substack{Q_{R^{1/2}}: \\ Q_{R^{1/2}} \cap (\mathrm{Br}_\alpha(\tau,\tau')\cap H) \neq \emptyset}} |Q_{R^{1/2}}| \, \|G\|^2_{L^\infty(Q_{R^{1/2}} \cap (\mathrm{Br}_\alpha(\tau,\tau')\cap H))}. \end{aligned} \tag{6.2}$$

For each $x \in H$, $G(x) \le 2|G^h(x)|$. Also note the equality $G^h(x) = \sum_s G * \check{\eta}_{\sim s}(x)$, where the sum is over dyadic s in the range $\lceil R^\beta \rceil^{-1} \lesssim s \lesssim R^{-1/2}$. This is because the Fourier support of G^h is contained in $\cup_\theta(\theta - \theta) \setminus B_{c\lceil R^\beta \rceil^{-1}}$ for a sufficiently small $c > 0$.

By dyadic pigeonholing, there is some dyadic s, $\lceil R^\beta\rceil^{-1}\lesssim s\lesssim R^{-1/2}$, so that the upper bound in (6.2) is bounded by

$$(\log R)\sum_{\substack{Q_{R^{1/2}}:\\ Q_{R^{1/2}}\cap(\mathrm{Br}_\alpha(\tau,\tau')\cap H)\neq\emptyset}}|Q_{R^{1/2}}|\,\|G*\check{\eta}_{\sim s}\|^2_{L^\infty(Q_{R^{1/2}}\cap(\mathrm{Br}_\alpha(\tau,\tau')\cap H))}.$$

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

$$(\log R)\sum_{\substack{Q_{R^{1/2}}:\\ Q_{R^{1/2}}\cap(\mathrm{Br}_\alpha(\tau,\tau')\cap H)}}\int_{\mathbb{R}^2}|G*\check{\eta}_{\sim s}|^2\,w_{Q_{R^{1/2}}}\lesssim(\log R)\int_{\mathbb{R}^2}|G*\check{\eta}_{\sim s}|^2.$$

Use Lemma 5.12 to upper bound the above integral to finish bounding $|\mathrm{Br}_\alpha(\tau,\tau')\cap H|$.

(2) *Bounding* $|Br_\alpha(\tau,\tau')\cap\Omega_k|$. First write the trivial inequality

$$\alpha^4|\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k|\leq\sum_{\substack{\ell(\tau)=\ell(\tau)=R^{-\delta}\\ d(\tau,\tau')\gtrsim R^{-\delta}}}\int_{\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k\cap\{|f_\tau f_{\tau'}|^{1/2}\sim\alpha\}}|f_\tau|^2|f_{\tau'}|^2.$$

By the definition of $\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k$ and Lemma 5.9, for each $x\in\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k$,

$$\begin{aligned}&|f_\tau(x)f_{\tau'}(x)|\\ &\leq|f_\tau(x)||f_{\tau'}(x)-f_{\tau'}^{k+1}(x)|+|f_\tau(x)-f_\tau^{k+1}(x)||f_{\tau'}^{k+1}(x)|+|f_\tau^{k+1}(x)f_{\tau'}^{k+1}(x)|\\ &\lesssim C_\varepsilon R^{O(\delta)}R^{-M\delta}\alpha^2+|f_\tau^{k+1}(x)f_{\tau'}^{k+1}(x)|.\end{aligned}$$

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

$$\int_{\{\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k\cap\{|f_\tau f_{\tau'}|^{1/2}\sim\alpha\}}|f_\tau|^2|f_{\tau'}|^2\lesssim\int_{\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k}|f_\tau^{k+1}|^2|f_{\tau'}^{k+1}|^2.$$

Repeat analogous bilinear restriction, high-dominated from the definition of Ω_k, and locally-constant steps from the argument bounding $\mathrm{Br}_\alpha(\tau,\tau')\cap H$ to obtain

$$\alpha^4|\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k|\lesssim R^{O(\delta)}\int_{\mathbb{R}^2}|g_k^h|^2.$$

Use Lemma 5.13 and Lemma 5.2 to bound the above integral, obtaining

$$\alpha^4|\mathrm{Br}_\alpha(\tau,\tau')\cap\Omega_k|\lesssim(\log R)^4\int_{\mathbb{R}^2}|g_k^h|^2\lesssim R^{O(\delta)}R^{O(\varepsilon)}\frac{\lambda(1)^2}{\alpha^2}\sum_{\tau_{k+1}}\int_{\mathbb{R}^2}|f_{\tau_{k+1}}^{k+1}|^2.$$

Use L^2-orthogonality and that $|f_{\tau_m}^m|\leq|f_{\tau_m}^{m+1}|$ for each m to bound each integral above:

$$\int_{\mathbb{R}^2}|f_{\tau_{k+1}}^{k+1}|^2\leq\int_{\mathbb{R}^2}|f_{\tau_{k+1}}^{k+2}|^2\leq C\sum_{\tau_{k+2}\subset\tau_{k+1}}\int_{\mathbb{R}^2}|f_{\tau_{k+2}}^{k+2}|^2\leq\cdots\leq C^{\varepsilon^{-1}}\sum_{\gamma\subset\tau_{k+1}}\int_{\mathbb{R}^2}|f_\gamma|^2.$$

We are done with this case because

$$\frac{\lambda(1)^2}{\alpha^2} \le \begin{cases} \max_s \lambda(s^{-1}R^{-1})\,\lambda(s) & \text{if } \alpha^2 > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\,\lambda(s)}, \\ \lambda(1)^2/\alpha^2 & \text{if } \alpha^2 \le \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\,\lambda(s)}. \end{cases}$$

(3) *Bounding* $|Br_\alpha(\tau,\tau')\cap L|$. Repeat the pruning step from the previous case to get

$$\alpha^6 |\mathrm{Br}_\alpha(\tau,\tau')\cap L| \lesssim \sum_{\substack{\ell(\tau)=\ell(\tau)=R^{-\delta}\\ d(\tau,\tau')\gtrsim R^{-\delta}}} \int_{\mathrm{Br}_\alpha(\tau,\tau')\cap L\cap\{|f_\tau f_{\tau'}|^{1/2}\sim\alpha\}} |f_\tau^1 f_{\tau'}^1|^2 |f_\tau f_{\tau'}|.$$

Use Cauchy–Schwarz and the locally constant lemma for the bound $|f_\tau^1 f_{\tau'}^1| \lesssim R^{O(\varepsilon)} g_1$, and recall that by Lemma 5.8, $g_1 \le C_\varepsilon R^\varepsilon \lambda(1)$. Then

$$\begin{aligned} R^{O(\varepsilon)} \sum_{\substack{\ell(\tau)=\ell(\tau)=R^{-\delta}\\ d(\tau,\tau')\gtrsim R^{-\delta}}} \int_{\mathrm{Br}_\alpha(\tau,\tau')\cap L} |g_1|^2 |f_\tau f_{\tau'}| &\le R^{O(\varepsilon)} \lambda(1)^2 \sum_{\ell(\tau)=R^{-\delta}} \int_{\mathbb{R}^2} |f_\tau|^2 \\ &\lesssim R^{O(\varepsilon)} \lambda(1)^2 \sum_\gamma \|f_\gamma\|_2^2. \end{aligned}$$

Using the same upper bound for $\lambda(1)^2/\alpha^2$ as in the previous case finishes the proof. ∎

6.2. Bilinear reduction

We will present a broad/narrow analysis to show that Proposition 6.1 implies Theorem 2.2. In order to apply Proposition 6.1, we must reduce to the case that f has an (R,ε)-normalized distribution function $\lambda(\cdot)$. We demonstrate this through a series of pigeonholing steps.

Proposition 6.1 implies Theorem 2.2. We will pigeonhole the f_γ so that roughly, for any s-arc ω of the parabola, the number

$$\#\{\gamma : \gamma\cap\omega \ne \emptyset,\ f_\gamma \ne 0\}$$

is either 0 or relatively constant among s-arcs ω. For the initial step, write

$$\{\tau_N : \exists\gamma \text{ such that } f_\gamma \ne 0,\ \gamma\subset\tau_N\} = \sum_{1\le\lambda\lesssim R^\beta R^{-\varepsilon}} \Lambda_N(\lambda),$$

where λ is a dyadic number, $\Lambda_N(\lambda) = \{\tau_N : \#\gamma\subset\tau_N \sim \lambda\}$, $\#\gamma\subset\tau_N$ means $\#\{\gamma\subset\tau_N : f_\gamma \ne 0\}$, and $\#\gamma\subset\tau_N\sim\lambda$ means $\lambda \le \#\gamma\subset\tau_N < 2\lambda$. Since there are $\lesssim \log R$ many λ in the sum, there exists some λ_N such that

$$|\{x : |f(x)| > \alpha\}| \le C(\log R)\Big|\Big\{x : C(\log R)\Big|\sum_{\tau_N\in\Lambda_N(\lambda_N)} f_{\tau_N}(x)\Big| > \alpha\Big\}\Big|.$$

Write

$$f^N = \sum_{\tau_N \in \Lambda_N(\lambda_N)} f_{\tau_N}.$$

Continuing in this manner, we have

$$\{\tau_k : \exists \tau_{k+1} \in \Lambda_{k+1}(\lambda_{k+1}) \text{ such that } \tau_{k+1} \subset \tau_k\} = \sum_{1 \le \lambda \le r_k} \Lambda_k(\lambda),$$

where

$$\Lambda_k(\lambda) = \big\{\tau_k : \exists \tau_{k+1} \in \Lambda_{k+1}(\lambda_{k+1}) \text{ s.t. } \tau_{k+1} \subset \tau_k \text{ and } \#\{\gamma : f_\gamma^{k+1} \neq 0,\ \ \gamma \subset \tau_k\} \sim \lambda\big\}$$

and for some λ_k,

$$\begin{aligned} &|\{x : (C(\log R))^{N-k}|f^{k+1}(x)| \ge |f(x)| > \alpha\}| \\ &\qquad \le C(\log R)|\{x : (C(\log R))^{N-k+1}|f^k(x)| \ge |f(x)| > \alpha\}|, \end{aligned}$$

where

$$f_k = \sum_{\tau_k \in \Lambda_k(\lambda_k)} f_{\tau_k}^{k+1}.$$

Continue this process until we have found τ_1 and λ_1 so that

$$|\{x : |f(x)| > \alpha\}| \le C^{\varepsilon^{-1}}(\log R)^{O(\varepsilon^{-1})}|\{x : C^{\varepsilon^{-1}}(\log R)^{O(\varepsilon^{-1})}|f^1(x)| > \alpha\}|.$$

The function f^1 now satisfies the hypotheses of Theorem 2.2 and the property that $\#\gamma \subset \tau_k \sim \lambda_k$ or $\#\gamma \subset \tau_k = 0$ for all k, τ_k. It follows that the associated distribution function $\lambda(\cdot)$ of f^1 is (R, ε)-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 γ or τ_k for which f_γ^1 or $f_{\tau_k}^1$ is nonzero. Now we may apply Proposition 6.1. Note that since $\log R \le \varepsilon^{-1}R^\varepsilon$ for all $R \ge 1$, the accumulated constant from this pigeonholing process satisfies $C^{\varepsilon^{-1}}(\log R)^{O(\varepsilon^{-1})} \le C_\varepsilon R^\varepsilon$. It thus suffices to prove Theorem 2.2 assuming that f is (R, ε)-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

$$|f(x)| \le \sum_{\ell(\tau)=K^{-1}} |f_\tau(x)|,$$

there is a universal constant $C > 0$ so that

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

implies

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

If

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

and

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

then we have

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

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

$$|f(x)| \le C \max_{\ell(\tau)=K^{-1}} |f_\tau(x)| + K^C \max_{\substack{\ell(\tau)=\ell(\tau')=K^{-1} \\ \tau,\tau' \text{ nonadj.} \\ C \max_{\ell(\tau_0)=K^{-1}} |f_{\tau_0}(x)| \le K^C |f_\tau(x) f_{\tau'}(x)|^{1/2}}} |f_\tau(x) f_{\tau'}(x)|^{1/2}.$$

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

$$\begin{aligned} |f(x)| &\lesssim C^m \max_{\ell(\tau)=R^{-1/2}} |f_\tau(x)| \\ &\quad + C^m K^C \sum_{\substack{R^{-1/2}<\Delta<1 \\ \Delta \in K^{\mathbb{N}}}} \max_{\ell(\tilde{\tau})\sim\Delta} \max_{\substack{\ell(\tau)=\ell(\tau')\sim K^{-1}\Delta \\ \tau,\tau' \subset \tilde{\tau}, \text{ nonadj.} \\ C \max_{\substack{\ell(\tau_0)=K^{-1}\Delta \\ \tau_0 \subset \tilde{\tau}}} |f_{\tau_0}(x)| \le K^C |f_\tau(x) f_{\tau'}(x)|^{1/2}}} |f_\tau(x) f_{\tau'}(x)|^{1/2}. \end{aligned}$$

Recall that our goal is to bound the size of the set

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

By the triangle inequality and using the notation θ for blocks τ with $\ell(\tau) = R^{-1/2}$,

$$|U_\alpha| \le |\{x \in \mathbb{R}^2 : \alpha \lesssim C^m \max_\theta |f_\theta(x)|\}| + \sum_{\substack{R^{-1/2}<\Delta<1 \\ \Delta \in K^{\mathbb{N}}}} \sum_{\substack{\ell(\tilde{\tau})\sim\Delta \\ \ell(\tau)=\ell(\tau')\sim K^{-1}\Delta \\ \tau,\tau' \subset \tilde{\tau}, \text{ nonadj.}}} |U_\alpha(\tau,\tau')|, \tag{6.3}$$

where $U_\alpha(\tau,\tau')$ is the set

$$\begin{aligned} \big\{x \in \mathbb{R}^2 : \alpha \lesssim (\log R)\, C^m K^C |f_\tau(x) f_{\tau'}(x)|^{1/2}, \\ C(|f_\tau(x)| + |f_{\tau'}(x)|) \le K^C |f_\tau(x) f_{\tau'}(x)|^{1/2}\big\}. \end{aligned}$$

The first term in the upper bound from (6.3) is bounded trivially by $\frac{\lambda(R^{-1/2})^2}{\alpha^4} \sum_\gamma \|f_\gamma\|_2^2$. By the assumption that $\|f_\gamma\|_\infty \lesssim 1$ for every γ, we know that $|f_\tau| \lesssim R^\beta$ for any τ. Also assume without loss of generality that $\alpha > 1$ (otherwise Theorem 2.2 follows from L^2-orthogonality). This means that there are $\sim \log R$ dyadic values of α' between α and R^β so by pigeonholing, there exists $\alpha' \in [\alpha/(C^m K^C), R^\beta]$ so that

$$|U_\alpha(\tau,\tau')| \lesssim (\log R + \log(C^m K^C))\, |\mathrm{Br}_{\alpha'}(\tau,\tau')|,$$

where the set $\mathrm{Br}_{\alpha'}(\tau,\tau')$ is defined in (6.1). By parabolic rescaling, there exists an affine transformation T so that $f_\tau\circ T=g_{\underline{\tau}}$ and $f_{\tau'}\circ T=g_{\underline{\tau}'}$ where $\underline{\tau}$ and $\underline{\tau}'$ are $\sim K^{-1}$-separated blocks in $\mathcal{N}_{\Delta^{-2}R^{-1}}(\mathbb{P}^1)$. Note that the functions $g_{\underline{\tau}}$ and $g_{\underline{\tau}'}$ inherit the property of being (Δ^2R,ε)-normalized in the sense required to apply Proposition 6.1 in each of the following cases.

Case 1. Suppose that for some $\beta'\in[1/2,1]$, $\Delta^{-1}R^{-\beta}=(\Delta^2R)^{-\beta'}$.

Then for each $\gamma\in\mathcal{P}(R,\beta)$, $f_\gamma\circ T=g_{\underline{\gamma}}$ for some $\underline{\gamma}\in\mathcal{P}(\Delta^2R,\beta')$. Applying Proposition 6.1 with functions $g_{\underline{\tau}}$ and $g_{\underline{\tau}'}$ and level set parameter α' leads to the inequality

$$|\mathrm{Br}_{\alpha'}(\tau,\tau')|\le K^C\alpha'\}|\le C_{\varepsilon,\delta}R^\varepsilon C^mK^{O(1)}$$
$$\times\begin{cases}\frac{1}{(\alpha')^4}\max\limits_{R^{-\beta}<s<R^{-1/2}}\lambda(s^{-1}R^{-1})\lambda(s)\sum_{\gamma\subset\tilde\tau}\|f_\gamma\|_2^2 & \text{if}\quad(\alpha')^2>\frac{\lambda(\Delta)^2}{\max\limits_s\lambda(s^{-1}R^{-1})\lambda(s)},\\ \frac{\lambda(\Delta)^2}{(\alpha')^6}\sum_{\gamma\subset\tilde\tau}\|f_\gamma\|_2^2 & \text{if}\quad(\alpha')^2\le\frac{\lambda(\Delta)^2}{\max\limits_s\lambda(s^{-1}R^{-1})\lambda(s)}.\end{cases}$$

Case 2. Now suppose that $\Delta^{-1}R^{-\beta}<(\Delta^2R)^{-1}$.

Let $\tilde\theta$ be $\Delta^{-1}R^{-1}\times R^{-1}$ blocks, and let $\underline{\tilde\theta}$ be $(\Delta^2R)^{-1}\times(\Delta^2R)^{-1}$ blocks so that $f_{\tilde\theta}\circ T=g_{\underline{\tilde\theta}}$. Let $B=\max_{\tilde\theta}|f_{\tilde\theta}|$ and divide everything by B in order to satisfy the hypotheses $\|g_{\underline{\tilde\theta}}\|_\infty/B\le1$ for all $\underline{\tilde\theta}$. Let

$$\tilde\lambda(s):=\frac{\lambda(\Delta s)}{\lambda(\Delta^{-1}R^{-1})}$$

count the number of $\underline{\tilde\theta}$ intersecting an s-arc. In the case $(\alpha')^2>\frac{\tilde\lambda(1)B^2}{\max_s\tilde\lambda(s^{-1}(\Delta^2R)^{-1})\tilde\lambda(s)}$ (with the maximum taken over $(\Delta^2R)^{-1}<s<(\Delta^2R)^{-1/2}$), use Proposition 6.1 with functions $g_{\underline{\tau}}/B$ and $g_{\underline{\tau}'}/B$ and level set parameter α'/B to get the inequality

$$|\mathrm{Br}_{\alpha'}(\tau,\tau')|$$
$$\le C_{\varepsilon,\delta}R^\varepsilon C^mK^{O(1)}\frac{B^4}{(\alpha')^4}\max_{(\Delta^2R)^{-1}<s<(\Delta^2R)^{-1/2}}\tilde\lambda(s^{-1}(\Delta^2R)^{-1})\tilde\lambda(s)\sum_{\tilde\theta\subset\tilde\tau}\|f_{\tilde\theta}\|_2^2/B^2.$$

Note that since $B\le\lambda(\Delta^{-1}R^{-1})$,

$$B^2\max_{(\Delta^2R)^{-1}<s<(\Delta^2R)^{-1/2}}\tilde\lambda(s^{-1}(\Delta^2R)^{-1})\,\tilde\lambda(s)\le\max_{\Delta^{-1}R^{-1}<s<R^{-1/2}}\lambda(s^{-1}R^{-1})\,\lambda(s)$$

and

$$\frac{\tilde\lambda(1)^2B^2}{\max\limits_s\tilde\lambda(s^{-1}(\Delta^2R)^{-1})\,\tilde\lambda(s)}\le\frac{\lambda(\Delta)^2\lambda(\Delta^{-1}R^{-1})^2}{\max\limits_{\Delta^{-1}R^{-1}<s<R^{-1/2}}\lambda(s^{-1}R^{-1})\,\lambda(s)}\le\lambda(\Delta^{-1}R^{-1})\,\lambda(\Delta).$$

Then in the case $(\alpha')^2\le\frac{\tilde\lambda(1)B^2}{\max_s\tilde\lambda(s^{-1}(\Delta^2R)^{-1})\tilde\lambda(s)}$, compute directly that

$$(\alpha')^4|\{x\in\mathbb{R}^2:\alpha'\sim|f_\tau(x)f_{\tau'}(x)|^{1/2},\ \ (|f_\tau(x)|+|f_{\tau'}(x)|)\le K^C\alpha'\}|$$
$$\lesssim\lambda(\Delta^{-1}R^{-1})\lambda(\Delta)\int_{\mathbb{R}^2}(|f_\tau|^2+|f_{\tau'}|^2)\lesssim\max_{\Delta^{-1}R^{-1}<s<R^{-1/2}}\lambda(s^{-1}R^{-1})\lambda(s)\sum_{\gamma\subset\tilde\tau}\|f_\gamma\|_2^2.$$

Using also that

$$\sum_{\tilde{\theta}\subset\tilde{\tau}} \| f_{\tilde{\theta}} \|_2^2 \le \sum_{\gamma\subset\tilde{\tau}} \| f_\gamma \|_2^2,$$

the bound for Case 2 is

$$\begin{aligned} &|\{x \in \mathbb{R}^2 : \alpha' \sim |f_\tau(x) f_{\tau'}(x)|^{1/2},\ \ (|f_\tau(x)| + |f_{\tau'}(x)|) \le K^C \alpha'\}| \\ &\qquad \le C_{\varepsilon,\delta}\, R^\varepsilon\, C^m\, K^{O(1)} \frac{1}{(\alpha')^4} \max_{R^{-\beta}<s<R^{-1/2}} \lambda(s^{-1}(\Delta^2 R)^{-1})\, \lambda(s) \sum_{\gamma\subset\tilde{\tau}} \| f_\gamma \|_2^2. \end{aligned}$$

It follows from (6.3) and the combined Case 1 and Case 2 arguments above that

$$\begin{aligned} |U_\alpha| &\le C_{\varepsilon,\delta}\, R^\varepsilon\, C^m\, K^{O(1)} \\ &\quad \times \begin{cases} \frac{1}{\alpha^4} \max\limits_{R^{-\beta}<s<R^{-1/2}} \lambda(s^{-1}R^{-1})\lambda(s) \sum_\gamma \| f_\gamma \|_2^2 & \text{if } \alpha > \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}, \\ \frac{\lambda(1)^2}{\alpha^6} \sum_\gamma \| f_\gamma \|_2^2 & \text{if } \alpha^2 \le \frac{\lambda(1)^2}{\max_s \lambda(s^{-1}R^{-1})\lambda(s)}. \end{cases} \end{aligned}$$

Recall that $K^m \sim R^{-1/2}$ and $K = R^\delta$ so that

$$C_{\varepsilon,\delta}\, R^\varepsilon\, C^m\, K^{O(1)} \le C_{\varepsilon,\delta}\, R^\varepsilon\, C^{O(\delta^{-1})}\, R^{O(1)\delta}.$$

Choosing δ small enough so that $R^{O(1)\delta} \le R^\varepsilon$ finishes the proof. ■

Funding. L. G. is supported by a Simons Investigator grant. D. M. is supported by the National Science Foundation under Award No. 2103249.

References

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

[2] Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.* **3** (1993), no. 2, 107–156.

[3] Bourgain, J.: Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.* **30** (2017), no. 1, 205–224.

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

[5] Bourgain, J. and Demeter, C.: The proof of the l^2 decoupling conjecture. *Ann. of Math. (2)* **182** (2015), no. 1, 351–389.

[6] Córdoba, A.: The Kakeya maximal function and the spherical summation multipliers. *Amer. J. Math.* **99** (1977), no. 1, 1–22.

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

[8] Demeter, C., Guth, L. and Wang, H.: Small cap decouplings. *Geom. Funct. Anal.* **30** (2020), no. 4, 989–1062.

[9] Guth, L., Maldague, D. and Wang, H.: Improved decoupling for the parabola. To appear in *J. Eur. Math. Soc. (JEMS)*, doi: 10.4171/JEMS/1295.

[10] Guth, L., Solomon, N. and Wang, H.: Incidence estimates for well spaced tubes. *Geom. Funct. Anal.* **29** (2019), no. 6, 1844–1863.

Received July 28, 2021; revised August 15, 2022. Published online November 16, 2022.

Yuqiu Fu
Department of Mathematics, Massachusetts Institute of Technology,
Cambridge, MA 02142-4307, USA;
yuqiufu@mit.edu

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