Local smoothing type estimates on L^p for large p

Thomas Wolff
Department of Mathematics
253-37 Caltech
Pasadena, Ca 91125 USA
wolff@cco.caltech.edu

This paper is a contribution to the theory of Fourier transform estimates related to the light cone $\Gamma = \{\xi \in \mathbb{R}^{n+1} : \xi_n = \sqrt{\sum_{j=1}^{n-1} |\xi_j|^2}\}$. There has been quite a bit of recent work on such estimates (see for example [1], [8], [10], [11], [2], [18], [6], [17], [20], [16]) but a number of apparently deeper questions remain open.

The best known are the so-called cone multiplier and local smoothing conjectures, both of which would follow (cf. [10] and [11]) from a sharp estimate by the associated square function,

$$\forall \epsilon\, \exists C_\epsilon : \ \|f\|_p \leq C_\epsilon N^\epsilon \|Sf\|_p \tag{1}$$

Here $p = 4$ in the case of the $2+1$ dimensional problem, which is the case that we will consider in this paper. Also f is a function with supp$\hat{f} \subset \Gamma_N(1)$, where $\Gamma_N(1)$ is the points which are at distance about N from the origin and at distance ≤ 1 from Γ. Sf corresponds to a decomposition of Γ into sectors of angular width $N^{-\frac{1}{2}}$:

$$(Sf(x))^2 = \sum_\Theta |f_\Theta(x)|^2$$

where $f_\Theta = \Xi_\Theta * f$, with Ξ_Θ having Fourier support near the part of $\Gamma_N(1)$ corresponding to an arc of the circle Θ of angular length $N^{-\frac{1}{2}}$ and the sum runs over a covering by such arcs.

We do not know how to prove (1). However, we can prove a related estimate with a high value of p, which has some of the same consequences that (1) would have.

Let us define

$$\|f\|_{p,mic} = \left(\sum_\Theta \|f_\Theta\|_p^p\right)^{\frac{1}{p}}$$

with the obvious modification when $p = \infty$. Here $p \geq 2$; it is then clear that $\|f\|_{p,mic} \leq \|Sf\|_p$.

<u>Theorem 1</u> The following estimate is valid if $p > 74$ and $\hat{f}$ is supported in $\Gamma_N(1)$.

$$\forall \epsilon \, \exists C_\epsilon : \ \|f\|_p \leq C_\epsilon N^\epsilon N^{\frac{1}{2}-\frac{2}{p}} \|f\|_{p,mic} \tag{2}$$

We remark that for the indicated range of p this estimate is clearly sharp except for issues connected with the N^ϵ factor. Further, that the main point is to obtain some value of $p < \infty$ for which (2) holds. We have tried to arrange the argument in a reasonably efficient way, but improvements in the value of p should certainly be possible.

It is easy to see that (2) fails when p is below the Strichartz exponent 6; for this, consider a function f such that $|f_\Theta|$ is approximately equal to 1 on a proportional subset of the unit cube for each Θ. It is therefore natural to conjecture that $p \geq 6$ is the correct range.

Like other estimates on high L^p spaces, Theorem 1 means that we understand the "large" values of f, but only them. Thus we record the following (immediate) corollary:

<u>Corollary 1</u> Assume that $f \in L^2$, $\hat{f}$ is supported in $\Gamma_N(1)$ and $\|f\|_{\infty,mic} \leq 1$. Then

$$|\{|f| \geq N^{\frac{1}{2}-\eta}\}| \leq C_\eta N^{-2+\delta(\eta)} \|f\|_2^2 \tag{3}$$

where $\delta(\eta) \to 0$ as $\eta \to 0$.

In the limit as $\eta \to 0$ this estimate may easily be seen to be sharp and it is considerably stronger than what can be obtained from any of the various results in the literature (cf. [2], [18], [17], [6], [20]). If one weakens (3) a bit by replacing $\|f\|_2^2$ with $\|Sf\|_4^4$, then the resulting estimate would be a consequence of the conjectured (1), but the known partial results on (1) lose a factor of $N^{\frac{1}{2}}$ [8] or slightly less than that ([2], [17], [20]). We remark here that the assumption $\|f\|_{\infty,mic} \leq 1$ is just a normalization and scales out if one replaces $\|f\|_2^2$ with $\|Sf\|_4^4$.

Theorem 1 easily implies the optimal L^{74} local smoothing bound, and a corresponding result for the cone multipliers. Let $\|f\|_{p,\alpha}$ be the inhomogeneous Sobolev norm with α derivatives in L^p.

Corollary 2 (i) If u is a solution of $\Box u = 0$, $u(\cdot, 0) = f$, $\frac{du}{dt}(\cdot, 0) = g$ in $2+1$ dimensions then

$$\|u\|_{L^p(\mathbb{R}^2 \times [1,2])} \leq C_{p\alpha} (\|f\|_{p,\alpha} + \|g\|_{p,\alpha-1})$$

if $p > 74$ and $\alpha > \frac{1}{2} - \frac{2}{p}$.

(ii) Let ρ_1 be a C_0^∞ function of one variable supported in the interval $(1, 2)$, and let ρ_2 be a C_0^∞ function of two variables. Then the cone multiplier operators T_α defined via $\widehat{T_\alpha f} = m_\alpha \hat{f}$, where $m_\alpha(x) = |x_3 - \sqrt{x_1^2 + x_2^2}|^\alpha \rho_1(x_3)\rho_2(x_1, x_2)$, are bounded on L^p if $p > 74$ and $\alpha > \frac{1}{2} - \frac{2}{p}$.

In connection with the cone multiplier statement, we note that the usual argument ([5] and [15], p. 422) which proves the optimal Bochner-Riesz estimate on L^p assuming an L^2 restriction theorem does not work as easily for the cone. Nevertheless it seems possible that bounds like Corollary 2(ii) could be proved without going through Theorem 1.

Theorem 1 also implies the following result, answering a question which the author heard of some time ago from P. Mattila.

Corollary 3 Let E be a set in $\mathbb{R}^2 \times \mathbb{R}$ with Hausdorff dimension greater than 1, and let F be a set in $\mathbb{R}^2$ with the property that if $(x, t) \in E$ then the one dimensional Lebesgue outer measure of $C(x, t) \cap E$ is nonzero. Then F cannot be of two dimensional Lebesgue measure zero.

Part of the point here is that no set theoretic restrictions are needed, in contrast to many related situations where one has to be more careful. The reason for this is that Theorem 1 is a measure estimate and makes no reference to any notion of dimension. On the other hand, the special case of Corollary 3

($*$) If F is a compact set in $\mathbb{R}^2$ and there is a compact set $E \subset \mathbb{R}^2$ with Hausdorff dimension greater than 1 such that F contains some circle centered at each point of E, then F has positive measure

is new also, although it appears to us that ($*$) could be proved without going through the entire proof of Theorem 1. Assertion ($*$) with "positive measure" replaced by "Hausdorff dimension two" follows from [18], but the combinatorial techniques from [3] used there are known to be hard to improve. On the other hand, it has been known for a while (Schlag [13]) that ($*$) would be a consequence of (1). The argument can be set up so that it only needs the large values estimate of Corollary 1, and it proves Corollary 3 as well. In the case where E has Hausdorff dimension greater than $\frac{3}{2} - \epsilon$ for suitably small ϵ, $(*)$ is a consequence of the abovementioned partial results on (1), and when $\dim E > \frac{3}{2}$ a geometric proof was given by Mitsis [9].

In addition, we give a further corollary (see Proposition 6.1 and the subsequent discussion) which is essentially a refinement of Corollary 3 to an L^p estimate.

We now give a brief sketch of the proof of Theorem 1. It involves a combination of incidence geometry and L^2 arguments.

In section 1 we consider a certain "continuum incidence problem" involving approximately tangent pairs of circles and prove a bound in this problem which generalizes a result in [18]. The proof given here is also a lot simpler than the proof in [18], although the same techniques are used (roughly a combination of those in [8] and [3]), so this may be of some independent interest.

To motivate the subsequent sections we recall a certain general principle applicable to many problems in this area. Namely, one often wants to use an induction argument of the following type: let $X(N)$ be an estimate making reference to functions living on Fourier scale N, for present purposes the estimate of Theorem 1. Suppose it has been proved on scale $N^{\frac{1}{2}}$. If f lives on Fourier scale N then apply $X(N^{\frac{1}{2}})$ to the part of f whose Fourier transform lives on a given sector of the cone of angular width $N^{-\frac{1}{4}}$, which can be rescaled using Lorentz transformations so that it becomes a portion of $\Gamma_{N^{\frac{1}{2}}}$. The resulting question of controlling interactions between different such sectors can be localized in x-space to $N^{-\frac{1}{2}}$-discs, so after dilating by $N^{\frac{1}{2}}$ one can apply $X(N^{\frac{1}{2}})$ again.

In practice this logic does not work very well, essentially because solutions of the differential inequality $|\dot{q}| \leq q^2$ can blow up in finite time so that applying the inductive hypothesis twice leads to problems. However it would be a different matter it instead one could apply the inductive hypothesis once on scale $N^{\frac{1}{2}}$ and once on a significantly smaller scale, let us say $N^{\frac{1}{2}(1-\epsilon_0)}$. As it turns out, in our context it is possible to do this.

Namely, in section 2 we put the result proved in section 1 into a dual form where it becomes a property of light rays instead of circles. In section 3 we obtain from this a certain "localization" property of functions with Fourier support near the light cone (related to the "two ends" condition used in several recent related papers) which allows us to change scales from $N^{\frac{1}{2}}$ to $N^{\frac{1}{2}(1-\epsilon_0)}$ under certain circumstances. Once we have this property, it is possible to set up an induction argument which proves Theorem 1. This is done in section 5; section 4 is the proofs of some lemmas. Other than the localization property, we need only uncertainty principle type L^2 techniques. In particular, we use neither stationary phase nor L^4 orthogonality.

Finally, section 6 is the proofs of the corollaries.

A few remarks concerning generalizations to higher dimensions. The analogue of Theorem 1 in $n+1$ dimensions is

$$\|f\|_p \leq C_\epsilon N^\epsilon N^{\frac{n-1}{2}-\frac{n}{p}} \|f\|_{p,mic} \tag{4}$$

with the natural definition of $\|f\|_{p,mic}$, and the possible range of validity is $p \geq \frac{2n+2}{n-1}$. Estimate (4) is substantially easier to prove in higher dimensions, and much better values of p (of the form $2+\mathcal{O}(\frac{1}{n})$ as $n \to \infty$) can be obtained; this will be done in a forthcoming paper by H. Farag and the author [4]. At present it is unclear to us whether or not one should be able to obtain a sharp result.

One could also consider the analogous question for spheres instead of cones. However, it is possible here to "go down" from estimate (4) for the cone in $\mathbb{R}^{n+1}$ to the corresponding estimate for the sphere in $\mathbb{R}^n$, so the case of spheres may be of less interest.

<u>List of notation</u>

We give here only notation which will be used throughout the paper.

A t-cube is a cube in $\mathbb{R}^3$ of side t whose vertices belong to $t\mathbb{Z}^3$; thus any two are identical or disjoint. We will always tacitly assume that t is dyadic; thus the grids may be taken to be nested.

χ_E: indicator function of the set E.

$|E|$: Lebesgue measure or cardinality of E depending on the context.

$\mathcal{E}_\delta(X)$: δ-entropy of X, i.e. maximum possible cardinality for a δ-separated subset of X.

We fix once and for all a small constant $\epsilon_0 > 0$ and then a large constant M.

$\phi(x)$: the function $(1+|x|^2)^{-\frac{M}{2}}$.

a_R: an affine map taking the unit cube centered at 0 to the rectangle R.

ϕ_R: the composition $\phi \circ a_R^{-1}$

$\psi(x)$: A function mapping $\mathbb{R}^3 \to \mathbb{R}$ such that

1. $\psi = \eta^2$, where $\hat{\eta}$ is supported in a small disc centered at 0.
2. $\psi \neq 0$ on a large cube centered at the origin.
3. The $\mathbb{Z}^3$ translations of ψ form a partition of unity.

ψ_R: the composition $\psi \circ a_R^{-1}$.

Γ: the forward light cone $\{\xi \in \mathbb{R}^3 : \xi_3 = \sqrt{\xi_1^2 + \xi_2^2}\}$

Γ_N: the cone segment $\Gamma \cap \{|\xi| \in [\frac{N}{2}, 2N]\}$.

$\Gamma_N(1)$: the 1-neighborhood of the cone segment Γ_N. Similarly we let $\Gamma_N(C)$ be the C-neighborhood of the cone segment $\{\xi : |\xi| \in [\frac{N}{2^C}, 2^C N]\}$.

For fixed N, we take a partition of unity subordinate to a covering of the circle by arcs Θ of length about $N^{-\frac{1}{2}}$, and use this to form a partition of unity y_Θ on $\Gamma_N(C)$ in the natural way. We let Ξ_Θ be a function whose Fourier transform coincides with y_Θ when $|\xi| \approx N$. If the support of $\hat{f}$ is contained in $\Gamma_N(C)$ we define

$$\|f\|_{p,mic} = \left(\sum_\Theta \|\Xi_\Theta * f\|_p^p \right)^{\frac{1}{p}}$$

for $p < \infty$ and

$$\|f\|_{\infty,mic} = \sup_\Theta \|\Xi_\Theta * f\|_\infty$$

1. A bound for circle tangencies As was mentioned in the introduction, this section of the paper is essentially a generalization and simplification of the result of [18]. One can obtain the result of [18] by specializing Lemma 1.4 below to the case of families of circles with δ-separated radii and then using fairly standard arguments.

In this section we use the notation $c(x, r)$ for the circle with center x and radius r, and we will always assume that x is in the disc centered at the origin with radius $\frac{1}{4}$ and that $\frac{1}{2} \leq r \leq 2$. If $c_i = c(x_i, r_i)$ then we define $d(c_1, c_2) = |x_1 - x_2| + |r_1 - r_2|$, $\Delta(c_1, c_2) = ||x_1 - x_2| - |r_1 - r_2||$. We fix large constants C_1, C_2, C_3, C_0, which will be specified below, and are chosen in the indicated order. We also fix numbers δ and t with δ small compared with t. A δ,t-rectangle is by definition the δ-neighborhood of an arc of length $\sqrt{\frac{\delta}{t}}$ on some circle. We say that a circle c is tangent to a δ, t-rectangle R if the $C_1\delta$-neighborhood of c contains R. Two δ, t-rectangles are close if there is a $2\delta, t$-rectangle which contains them both and comparable if there is a $C_0\delta, t$-rectangle which contains them both; if they are not close they are nonclose and if they are not comparable they are incomparable.

Lemma 1.1 If R_1 and R_2 are incomparable (δ, t)-rectangles both tangent to a common circle c, and if $p_i \in R_i$, then $|p_1 - p_2| \geq C_3\sqrt{\frac{\delta}{t}}$.

Proof Both R_1 and R_2 are contained in the $C_1\delta$-neighborhood of c. If they contain points with $|p_1 - p_2| < C_3\sqrt{\frac{\delta}{t}}$ then it follows that they are contained in the $C_1\delta$-neighborhood of an arc of c of length $< (C_3 + 2)\sqrt{\frac{\delta}{t}}$, so we can take $C_0 = \max(C_1, (C_3 + 2)^2)$. $\square$

The definitions of "close" and "comparable" are almost the same, but for technical reasons we need both of them. The distinction between the two is insignificant because of the following

Lemma 1.2 For any constant A there is a constant $C(A)$ such that if R is any $(A\delta, t)$ rectangle then the cardinality of any set of pairwise nonclose (δ, t) rectangles contained in R is at most C.

Proof This is left to the reader. $\square$

With the given choice of C_0 we therefore have the following fact.

Lemma 1.3 Any set of pairwise non-close (δ, t)-rectangles has a pairwise incomparable subset with comparable cardinality.

Proof This follows from Lemma 1.2 and standard counting arguments. $\square$

We now fix two finite sets of circles, $\mathcal{W}$ and $\mathcal{B}$, which we call white and black respectively, and let $m = |\mathcal{W}|$ and $n = |\mathcal{B}|$. We make the following assumptions:

$$\mathcal{W} \text{ and } \mathcal{B} \text{ are each } \delta\text{-separated.} \tag{5}$$

$$\text{If } c_1 \text{ and } c_2 \text{ are the same color then } d(c_1, c_2) \leq t. \tag{6}$$

$$\text{If } c_1 \text{ and } c_2 \text{ are opposite colors then } d(c_1, c_2) \geq t. \tag{7}$$

$$\text{If } c_1 \text{ and } c_2 \text{ are opposite colors then } d(c_1, c_2) \leq 100t. \tag{8}$$

Define $w \in \mathcal{W}$ and $b \in \mathcal{B}$ to be incident if $\Delta(b, w) < \delta$, and define $\mathcal{I}(\mathcal{W}, \mathcal{B})$ to be the set of pairs (w, b), $w \in \mathcal{W}$, $b \in \mathcal{B}$, such that w and b are incident. In principle, one wants to make a nontrivial estimate of the cardinality $|\mathcal{I}(\mathcal{W}, \mathcal{B})|$, but the "clamshell" configuration where all the white and black circles are tangent at a point shows that this is not possible, so it is necessary to set things up a bit differently. We make the following further definitions.

A rectangle of type $(\geq \mu, \geq \nu)$ is a (δ, t)-rectangle which is tangent to at least μ white circles and to at least ν black circles. Additionally, a rectangle of type (μ, ν) is a (δ, t)-rectangle which is tangent to between μ and 2μ white circles and to between ν and 2ν black circles, a rectangle of type $(\mu, \geq \nu)$ is a (δ, t)-rectangle which is tangent to between μ and 2μ white circles and to at least ν black circles, and a rectangle of type $(\geq \mu, \nu)$ is defined analogously. The main result of this section is

Lemma 1.4 If $\epsilon > 0$ then there is a constant C_ϵ such that the cardinality of any set of pairwise incomparable (δ, t)-rectangles of type $(\geq \mu, \geq \nu)$ is bounded by

$$C_\epsilon (mn)^\epsilon \left((\frac{mn}{\mu\nu})^{\frac{3}{4}} + \frac{m}{\mu} + \frac{n}{\nu} \right)$$

The remainder of the section is the proof of Lemma 1.4 and its corollary, Lemma 1.17 below. We first recall some geometric and combinatorial facts; these are largely elementary, and most of them have been used before in similar contexts (e.g. [8], [12], [18]). However, our present notation is rather different, so we discuss the proofs.

Lemma 1.5 If w and b are incident, then there is a (δ, t)-rectangle R such that both w and b are tangent to any (δ, t)-rectangle close to R. Conversely, if w and b are tangent to a common (δ, t)-rectangle, then $\Delta(b, w) \leq C_2\delta$. If they are tangent to comparable (δ, t)-rectangles then $\Delta(b, w) \lesssim \delta$ where the implicit constant should be chosen after C_0.

Proof Let $w = c(x_1, r_1)$, $b = c(x_2, r_2)$; we can suppose that $r_1 > r_2$. Then $||x_1 - x_2| - |r_1 - r_2|| < \delta$. Take the ray emanating from x_1 and passing through x_2. It intersects w at a point p. Let R be the δ-neighborhood of the arc of w centered at p with length $\sqrt{\frac{\delta}{t}}$, and let S be the 100δ-neighborhood of the arc of w centered at p with length $10\sqrt{\frac{\delta}{t}}$. It is not difficult to verify the following two facts provided C_1 has been chosen large enough:

1. S is contained in the $C_1\delta$-neighborhoods of both w and b.
2. Any rectangle which is close to R in the sense defined above is contained in S.

The first statement of the lemma follows. For the converse statement, we use that the area of the intersection of the $C_1\delta$-neighborhoods of w and of b is $\lesssim \frac{\delta^2}{\sqrt{t\Delta(w,b)}}$ (see [19], Lemma 3.1(a)). If w and b are tangent to a common δ, t-rectangle, then the intersection of their $C_1\delta$-neighborhoods contains this rectangle, so we get

$$\delta \cdot \sqrt{\frac{\delta}{t}} \lesssim \frac{\delta^2}{\sqrt{t\Delta(w, b)}}$$

which gives $\Delta(b, w) \leq C_2\delta$. The final statement of the lemma follows in a similar way. □

Lemma 1.6 For any w and b the cardinality of a set of pairwise incomparable (δ, t)-rectangles tangent to both w and b cannot exceed a fixed constant C.

Proof If any such rectangle exists then $\Delta(w, b) \leq C_2\delta$ by Lemma 1.5. The intersection of the $C_1\delta$-neighborhoods of w and b is then contained in the $C\delta$-neighborhood of an arc of length $C\sqrt{\frac{\delta}{t}}$ ([19], Lemma 3.1b), and by Lemma 1.2 the latter set contains only a bounded number of pairwise incomparable (δ, t)-rectangles. □

We define $\tilde{\mathcal{I}}(\mathcal{B}, \mathcal{W})$ similarly to $\mathcal{I}(\mathcal{B}, \mathcal{W})$ replacing δ with $C_2\delta$.

Lemma 1.7 (a) If $\mathcal{R}$ is any set of pairwise nonclose (δ, t)-rectangles then

$$|\tilde{\mathcal{I}}(\mathcal{B}, \mathcal{W})| \gtrsim |\{(R, b, w) : R \in \mathcal{R}, \text{ and } b \text{ and } w \text{ are both tangent to } R\}|$$

(b) There is a set $\mathcal{R}$ of pairwise incomparable (δ, t)-rectangles such that

$$|\mathcal{I}(\mathcal{B}, \mathcal{W})| \lesssim |\{(R, b, w) : R \in \mathcal{R}, \text{ and } b \text{ and } w \text{ are both tangent to } R\}|$$

Proof Consider a pair $(b, w) \in \mathcal{B} \times \mathcal{W}$. If $(b, w) \notin \tilde{\mathcal{I}}(\mathcal{B}, \mathcal{W})$ then no rectangle R as in the lemma can exist, by Lemma 1.5, and if $(b, w) \in \tilde{\mathcal{I}}(\mathcal{B}, \mathcal{W})$ then just a bounded number of pairwise nonclose such R's can exist, by Lemma 1.6. This proves (a). For (b), use Lemma 1.5 to choose for each $(b, w) \in \mathcal{I}(\mathcal{B}, \mathcal{W})$ a rectangle R such that any rectangle close to R is tangent to w and b. Take a maximal pairwise nonclose subset of these R's, and then use Lemma 1.3. □

Lemma 1.8 Let c_1, c_2, c_3 be three circles. Let $\mathcal{R}$ be a set of pairwise incomparable (δ, t)-rectangles with the following property:

$(*)$ For each $R \in \mathcal{R}$ there is a circle c such that $d(c, c_i) \geq t$ for each i, such that c and c_1 are tangent to R, and furthermore such that there are two (δ, t)-rectangles R_2 and R_3 such that c and c_i are tangent to R_i for $i = 2, 3$, and such that R, R_2 and R_3 are pairwise incomparable.

Then the cardinality of $\mathcal{R}$ cannot exceed a fixed constant.

Proof We use the following fact, which is included in the so-called Marstrand three circle lemma.

(M): Let $\mathcal{C}$ be all circles c such that $d(c, c_i) \geq t$ and $\Delta(c, c_i) \leq C_2\delta$, and such that the $C_1\delta$-neighborhood of c intersects the $C_1\delta$-neighborhood of c_i $(i = 1, 2, 3)$ at points p_i satisfying $|p_i - p_j| \geq C_3\sqrt{\frac{\delta}{t}}$. Then the set $\cup_{c \in \mathcal{C}}(C_1\delta\text{-neighborhood of } c) \cap (C_1\delta\text{-neighborhood of } c_1)$ is contained in the union of the $C_1\delta$-neighborhoods of two arcs of c_1 each of length less than $\lesssim \sqrt{\frac{\delta}{t}}$.

This fact is implicit in [8] and is stated in [18], Lemma 1.8. A detailed proof is contained in the proof of Lemma 3.2 in [19].

To pass from this to Lemma 1.8, let c be a circle as in $(*)$. Then $\Delta(c, c_i) \leq C_2\delta$ for each i by Lemma 1.5. Let p_i be a point of R_i, $i = 2, 3$, and let p_1 be a point of R. All three rectangles R, R_2 and R_3 are tangent to c, so by Lemma 1.1 we must have $|p_i - p_j| \geq C_3\sqrt{\frac{\delta}{t}}$. The definition of tangent implies that R is contained in the intersection of the $C_1\delta$-neighborhoods of c and of c_1. Accordingly, by (M), R is contained in the union of the $C_1\delta$-neighborhoods of two arcs of c_1 each of length $\lesssim \sqrt{\frac{\delta}{t}}$. Now Lemma 1.2 implies that there are at most a bounded number of pairwise incomparable R's. □

Lemma 1.9 Let $c_1 = c(x_1, r_1)$ and $c_2 = c(x_2, r_2)$ be circles with $d(c_1, c_2) \geq t$, and assume that $r_1 \geq r_2$. Let R_1 and R_2 be comparable (δ, t)-rectangles and assume that c_i is tangent to R_i for $i = 1, 2$. Then

(i) c_2 is contained in the $C\delta$-neighborhood of the interior of c_1.

(ii) For any constant A there is a constant $C = C(A)$ such that the cardinality of a set of pairwise incomparable (δ, t) rectangles each of which is tangent to c_1 and intersects the $A\delta$-neighborhood of the interior of c_2 does not exceed C.

Proof (i) Let $c_i = C(x_i, r_i)$. Lemma 1.5 implies $\Delta(c_1, c_2) \leq C\delta$. Then $z \in c_2$ implies $|z - x_1| \leq |z - x_2| + |x_2 - x_1| \leq |z - x_2| + r_1 - r_2 + C\delta \leq r_2 + \delta + r_1 - r_2 + C\delta = r_1 + (C+1)\delta$, as claimed.

(ii) $\Delta(c_1, c_2) \lesssim \delta$ by Lemma 1.5. Using that $r_1 \geq r_2$ it follows similarly to the proof of (i) that the intersection of the δ-neighborhood of c_1 with the $A\delta$-neighborhood of the interior of c_2 is contained in the $C\delta$-neighborhood of c_2 for suitable C. The lemma now follows using that the intersections of the $C\delta$-neighborhoods of c_1 and c_2 has diameter $\lesssim \sqrt{\frac{\delta}{t}}$ by [19], Lemma 3.1, and using also Lemma 1.2. □

In the proof of Lemma 1.4 we employ the "Canham bound plus divide and conquer" strategy of (for example) [3]. The next few lemmas provide the necessary combinatorial ingredients.

Lemma 1.10 (Canham Lemma) The cardinality of any set of pairwise incomparable (δ, t) rectangles of type (μ, ν) is $\lesssim \frac{mn^{\frac{2}{3}}}{\mu\nu^{\frac{2}{3}}} + \frac{n}{\nu}$. The cardinality of any set of pairwise incomparable (δ, t) rectangles of type $\geq \mu, \geq \nu$ is $\lesssim \frac{mn^{\frac{2}{3}}}{\mu\nu^{\frac{2}{3}}} + \frac{n}{\nu}\log\frac{m}{\mu}$.

Proof The second statement follows readily from the first by summing over the different dyadic levels for μ and ν.

To prove the first statement, let $\mathcal{R}$ be a set of pairwise incomparable rectangles of type (μ, ν). Observe that for each $w \in \mathcal{W}$ and $b \in \mathcal{B}$ there is at most a bounded number of rectangles $R \in \mathcal{R}$ which are tangent to both w and b, by Lemma 1.5; this fact will be used several times below.

The letter b always denotes an element of $\mathcal{B}$, and w an element of $\mathcal{W}$. We define the following sets:

$\mathcal{P}$: all pairs (w, b) such that there is $R \in \mathcal{R}$ with both w and b tangent to R.

$\mathcal{P}(b)$: all pairs $(w, b) \in \mathcal{P}$ with given second member b.

$\mathcal{Q}$: all quadruples (w_1, w_2, w_3, b) such that the following holds: there are distinct rectangles $R_i \in \mathcal{R}$ such that, for each i, w_i and b are both tangent to R_i.

$\mathcal{Q}(b)$: all quadruples $(w_1, w_2, w_3, b) \in \mathcal{Q}$ with given last member b.

Lemma 1.8 implies that, for given w_1, w_2, w_3, there are at most C rectangles $R_1 \in \mathcal{R}$ for which some $b \in \mathcal{B}$ may be found such that $(w_1, w_2, w_3, b) \in \mathcal{Q}$ and such that the choice of R_1 in the definition is the given one. For each choice of R_1, there are at most 2ν ways to choose b, since R_1 must have type (μ, ν). Accordingly

$$|\mathcal{Q}| \leq m^3 \cdot C \cdot 2\nu \tag{9}$$

On the other hand, for each $b \in \mathcal{B}$ there are exactly $|\mathcal{P}(b)|^3$ triples (w_1, w_2, w_3) such that, for some $R_1, R_2, R_3 \in \mathcal{R}$ and each $i \in \{1, 2, 3\}$, w_i and b are both tangent to R_i. Similarly, there are exactly $|\mathcal{P}(b)|^2$ pairs (w_1, w_2) such that, for some $R_1, R_2 \in \mathcal{R}$ and each $i \in \{1, 2\}$, w_i and b are both tangent to R_i. For each of these, there are at most 4μ choices of w_3 which are tangent to either R_1 or R_2. It follows that

$$|\mathcal{Q}(b)| \geq |\mathcal{P}(b)|^3 - 12\mu|\mathcal{P}(b)|^2 \tag{10}$$

Therefore

$$\begin{aligned} |\mathcal{P}| &= \sum_b |\mathcal{P}(b)| \\ &\leq \sum_{b:|\mathcal{P}(b)|<24\mu} |\mathcal{P}(b)| + \sum_{b:|\mathcal{P}(b)|\geq 24\mu} |\mathcal{P}(b)| \\ &\leq 24\mu n + \sum_{b:|\mathcal{P}(b)|\geq 24\mu} (2|\mathcal{Q}(b)|)^{\frac{1}{3}} \end{aligned}$$

$$
\begin{aligned}
&\leq 24\mu n + n^{\frac{2}{3}}(2 \sum_{b:|\mathcal{P}(b)|\geq 24\mu} |\mathcal{Q}(b)|)^{\frac{1}{3}} \\
&\lesssim \mu n + \nu^{\frac{1}{3}} n^{\frac{2}{3}} m
\end{aligned}
$$

where the second inequality follows from (10) and the last inequality follows from (9). If $R \in \mathcal{R}$ then there are at least $\mu\nu$ pairs $(w, b) \in \mathcal{W} \times \mathcal{B}$ such that w and b are both tangent to R. It follows that $|\mathcal{P}| \lesssim \frac{n}{\nu} + \frac{n^{\frac{2}{3}} m}{\nu^{\frac{2}{3}} \mu}$ as claimed. □

Lemma 1.11 Assume $\mathcal{W}$ and $\mathcal{B}$ have no (δ, t) rectangles of type $(\geq 1, \geq \nu_0)$ nor of type $(\geq \mu_0, \geq 1)$. Then

$$|\mathcal{I}(\mathcal{W}, \mathcal{B})| \lesssim \mu_0^{\frac{1}{3}} n m^{\frac{2}{3}} \log \nu_0 + \nu_0 m \log \mu_0$$

Proof Using Lemma 1.7 there is a pairwise incomparable set of (δ, t)-rectangles $\mathcal{R}$ with (sums over μ and ν run through dyadic values)

$$
\begin{aligned}
|\mathcal{I}(\mathcal{W}, \mathcal{B})| &\lesssim \sum_{R \in \mathcal{R}} |\{(b, w) : b \text{ and } w \text{ are both tangent to } R\}| \\
&\lesssim \sum_{\mu, \nu} \mu\nu \cdot \#(\text{members of } \mathcal{R} \text{ which are of type } \mu, \nu) \\
&\lesssim \sum_{\substack{\mu \leq \mu_0 \\ \nu \leq \nu_0}} \mu\nu \left(\frac{n m^{\frac{2}{3}}}{\nu \mu^{\frac{2}{3}}} + \frac{m}{\mu}\right) \\
&\lesssim \mu_0^{\frac{1}{3}} n m^{\frac{2}{3}} \log \nu_0 + \nu_0 m \log \mu_0
\end{aligned}
$$

At the next to last line, we used Lemma 1.10 with the roles of white and black reversed. □

Lemma 1.12 (Cell decomposition lemma) Suppose we randomly choose r black circles. Then with high probability there is a decomposition of the white circles

$$\mathcal{W} = \mathcal{W}^* \cup (\cup_{i=1}^{R} \mathcal{W}_i) \quad \text{(disjoint decomposition)}$$

with

(i) $R \leq r^3 a(r)$, where $a(r)$ grows extremely slowly - in particular slower than any power of $\log r$.

(ii) For each i, there are $\lesssim \frac{n\log n}{r}$ black circles b such that $\Delta(b,w) < C_2\delta$ for some $w \in \mathcal{W}_i$.

(iii) Each circle in $\mathcal{W}^*$ satisfies $\Delta(b,w) < C_2\delta$ for some circle b in the random sample.

Proof This lemma is due essentially to workers in computational geometry (see for example [3], [14]), and was adapted for analysis purposes in [18]. The version stated here differs from Proposition 2.1 in [18] in that there we had $\mathcal{W} = \mathcal{B}$, but the present bipartite version follows in exactly the same way. $\square$

Notice that the same parameter (namely $C_2\delta$) appears on the right hand side in (ii) and (iii). It would be more convenient to be able to use $C_2\delta$ in (ii) and δ in (iii), but this is not possible without additional work; see the proof of the claim in [18], p. 997. We must therefore phrase the next lemma in terms of $\tilde{\mathcal{I}}(\mathcal{B},\mathcal{W})$.

Lemma 1.13 With high probability the cardinality of $\mathcal{W}^*$ is $\lesssim \frac{r|\tilde{\mathcal{I}}(\mathcal{B},\mathcal{W})|}{n}$.

Proof Part (iii) of Lemma 1.12 implies that the probability of a given circle $w \in \mathcal{W}$ to belong to $\mathcal{W}^*$ is bounded by

$$\frac{r}{n}|\{b \in \mathcal{B} : \Delta(b,w) < C_2\delta\}|$$

Accordingly the expectation of the cardinality of $\mathcal{W}^*$ is bounded by $\frac{r|\tilde{\mathcal{I}}(\mathcal{B},\mathcal{W})|}{n}$, which suffices. $\square$

We define a cluster of white circles to be a subset $\mathcal{C} \subset \mathcal{W}$, with the property that there is a (δ,t) rectangle R such that every circle in $\mathcal{C}$ is tangent to a (δ,t) rectangle comparable to R. In a similar manner we define a cluster of black circles. The next two lemmas are related to considerations in section 1 of [20].

Lemma 1.14 Let $\mathcal{C}$ be a cluster of white circles and let b be a black circle. Then the cardinality of any set of pairwise incomparable $)\delta,t)$-rectangles each of which is tangent to some circle in $\mathcal{C}$ and to b is bounded by a constant.

Proof Consider the circles in the cluster with smaller radius than b. Choose a circle in the cluster, w_1, whose radius is as large as possible subject to this constraint. By the clustering property and Lemma 1.9(i) above, any other $w \in \mathcal{C}$ with these properties is contained within the $C\delta$-neighborhood of the interior of w_1. By Lemma 1.9(ii), b cannot be tangent to more than a bounded number of incomparable (δ, t)-rectangles which intersect the latter set. A symmetric argument for the circles with larger radius than b completes the proof. $\square$

Lemma 1.15 (Clustering Lemma) Given a value of μ_0 we can subdivide the white circles as

$$\mathcal{W} = \mathcal{W}_g \cup \mathcal{W}_b$$

where

1. W_g and $\mathcal{B}$ have no (δ, t)-rectangles of type $(\geq \mu_0, \geq 1)$, and
2. $\mathcal{W}_b$ is the union of at most $\frac{|\mathcal{W}|}{\mu_0}(\log m)(\log n)$ clusters.

Proof We fix a large constant A.

We will use a recursive argument. Accordingly, if $\mathcal{W}^i \subset \mathcal{W}$, then we let $\kappa(\mathcal{W}^i)$ be the maximum possible cardinality for a set of pairwise incomparable rectangles of type $\geq \mu_0, \geq 1$ for $\mathcal{W}^i$ and $\mathcal{B}$. Note first of all that $\kappa(\mathcal{W}) \lesssim \frac{mn}{\mu_0}$ by Lemma 1.6.

Assume now that $\kappa(\mathcal{W}^i) = k$. We will prove: $\mathcal{W}^i = \mathcal{W}^{i+1} \cup \mathcal{W}^i_b$ where $\kappa(\mathcal{W}^{i+1}) \leq \frac{k}{2}$, and there is a set $\mathcal{R}_i$ of at most $A\frac{m}{\mu_0}\log m$ rectangles such that each circle in $\mathcal{W}^i_b$ is tangent to a rectangle which is comparable to a rectangle in $\mathcal{R}_i$.

Namely, let $\mathcal{R}$ be a set of pairwise incomparable rectangles of type $(\geq \mu_0, \geq 1)$ for $\mathcal{W}_i$ and $\mathcal{B}$ with maximum possible cardinality k. There are two cases.

(i) If $k < A\frac{m}{\mu_0}\log m$ then we let $\mathcal{W}^i_b$ be all circles in $\mathcal{W}^i$ which are tangent to a rectangle comparable to a rectangle in $\mathcal{R}_i$ and $\mathcal{W}^{i+1} = \mathcal{W}^i \backslash \mathcal{W}^i_b$. Evidently $\kappa(\mathcal{W}^{i+1}) = 0$.

(ii) If $k < A\frac{m}{\mu_0}\log m$ then we choose $A\frac{m}{\mu_0}\log m$ rectangles from $\mathcal{R}$ at random. We let $\mathcal{W}^i_b$ be the circles tangent to a rectangle which is comparable to a rectangle in the random sample, and $\mathcal{W}^{i+1} = \mathcal{W}^i \backslash \mathcal{W}^i_b$. We will show that with high probability $\kappa(\mathcal{W}^{i+1}) \leq \frac{k}{2}$.

For this define, for each $c \in \mathcal{W}^i$,

$$P(w) = k^{-1}|\{R \in \mathcal{R}_i : w \text{ is tangent to a rectangle comparable to } R\}|$$

Then w is in $\mathcal{W}^{i+1}$ with probability at most $(1 - P(w))^{A\frac{m}{\mu_0}\log m}$. If $P(w) \geq \frac{1}{2}\frac{\mu_0}{m}$ then this is at most m^{-CA}. Accordingly with high probability no circles with $P(w) \geq \frac{1}{2}\frac{\mu_0}{m}$ belong to $\mathcal{W}^{i+1}$.

Now let $\mathcal{R}_{i+1}$ be a maximal set of pairwise incomparable rectangles of type $(\geq \mu_0, \geq 1)$ for $\mathcal{W}^{i+1}$ and $\mathcal{B}$. Every rectangle in $\mathcal{R}_{i+1}$ is comparable to a rectangle in $\mathcal{R}_i$, since otherwise we could have enlarged $\mathcal{R}_i$ by adjoining an incomparable member of $\mathcal{R}_{i+1}$ to it. Consider now the quantity

$$\sum_{w \in \mathcal{W}^{i+1}} |\{R \in \mathcal{R}_i : R \text{ is comparable to a rectangle tangent to } w\}| \tag{11}$$

This is equal to $\sum_{w \in \mathcal{W}^{i+1}} kP(w)$, so with high probability it is less than $\frac{k}{2}\frac{\mu_0}{m}$. On the other hand, the quantity (11) can also be described as

$$\sum_{R \in \mathcal{R}_i} |\{w \in \mathcal{W}^{i+1} : w \text{ is tangent to a rectangle comparable to } R\}|$$

and therefore dominates $\mu_0|\mathcal{R}_{i+1}|$. We conclude that $|\mathcal{R}_{i+1}| \leq \frac{k}{2}$, as was to be shown.

We now proceed recursively. Let $\mathcal{W}^0 = \mathcal{W}$ and apply the preceding to express $\mathcal{W}^0 = \mathcal{W}_b^0 \cup \mathcal{W}^1$. Then apply the preceding to express $\mathcal{W}^1 = \mathcal{W}_b^1 \cup \mathcal{W}^2$ and continue in this manner, stopping when we reach a situation where we are in case (i) above. Suppose we stop after T stages. Since $\kappa(\mathcal{W}^i)$ is initally $\leq \frac{mn}{\mu_0}$ and decreases each time by a factor of 2, we then have $T \lesssim \log n$. We now define $\mathcal{W}_g$ to be the set $\mathcal{W}^{i+1}$ defined at the last iteration. It satisfies $\kappa(\mathcal{W}_g) = 0$ as required. On the other hand we define $\mathcal{W}^b = \cup_i \mathcal{W}_b^i$. This set is the union of the $\mathcal{O}(\log n)$ sets $\mathcal{W}_b^i$, for each of which there is a collection of $A\frac{m}{\mu_0}\log m$ rectangles such that each circle in $\mathcal{W}_b^i$ is tangent to a rectangle comparable to one of these. The lemma follows. $\square$

We prove Lemma 1.4 in two steps: first we consider the case $\mu = \nu = 1$ and then the general case.

<u>Lemma 1.16</u> For any $\epsilon > 0$ the cardinality of a set of pairwise incomparable rectangles of type $(\geq 1, \geq 1)$ is

$$\leq C_\epsilon \left(((mn)^{\frac{3}{4}+\epsilon} + m \log n + n \log m \right) \tag{12}$$

for a suitable constant C_ϵ.

<u>Proof</u> We fix ϵ and will argue by induction on mn. We note first of all that the case of small mn is a tautology, and may therefore assume that (12) has been proved for families of circles $\mathcal{W}', \mathcal{B}'$ with $|\mathcal{W}'||\mathcal{B}'| \leq \frac{mn}{2}$. If $m \leq n^{\frac{1}{3}+\epsilon}$ (or vice versa) then the bound (12)

follows directly from Lemma 1.10, since in this case $mn^{\frac{2}{3}} \leq (mn)^{\frac{3}{4}+\epsilon}$. We may therefore assume that

$$m^{\frac{1}{3}+\epsilon} \leq n \leq m \tag{13}$$

Let $\mathcal{R}$ be a maximal set of pairwise incomparable (δ, t) rectangles of type $(\geq 1, \geq 1)$.

Let $\mu_0 = \nu_0 = (mn)^{\frac{1}{4}}$, replace δ with $C_2\delta$, and apply Lemma 1.15 twice with these parameters, once as stated and once with the roles of $\mathcal{W}$ and $\mathcal{B}$ reversed. This gives decompositions $\mathcal{W} = \mathcal{W}_g \cup \mathcal{W}_b$, $\mathcal{B} = \mathcal{B}_g \cup \mathcal{B}_b$.

Any $R \in \mathcal{R}$ is a rectangle of type $(\geq 1, \geq 1)$ for one of the three pairs $(\mathcal{W}_b, \mathcal{B})$, $(\mathcal{W}, \mathcal{B}_b)$ or $(\mathcal{W}_g, \mathcal{B}_g)$. We show first that there are $\lesssim (\log m)(\log n)(mn)^{\frac{3}{4}}$ members of $\mathcal{R}$ which are rectangles of type $(\geq 1, \geq 1)$ for $(\mathcal{W}_b, \mathcal{B})$. Let R be such a rectangle. Then there is $\mathcal{C}$, one of the clusters in 2. of Lemma 1.15, such that R is a $\geq 1, \geq 1$-rectangle for $\mathcal{C}$ and $\mathcal{B}$. For given $b \in \mathcal{B}$, by Lemma 1.14, there are at most a bounded number of incomparable $(C_2\delta, t)$-rectangles which are $C_2\delta$-tangent to some circle in $\mathcal{C}$ and to b. It follows using Lemma 1.2 that there are at most a bounded number of $R \in \mathcal{R}$ which are δ-tangent to some circle in $\mathcal{C}$ and to b. Hence the number of (δ, t) rectangles $R \in \mathcal{R}$ of type $(\geq 1, \geq 1)$ for $\mathcal{C}$ and $\mathcal{B}$ is $\lesssim n$. Summing over the clusters, there are $\lesssim \frac{mn}{\mu_0}(\log m)(\log n)$ rectangles of type $(\geq 1, \geq 1)$ for $\mathcal{W}_b$ and $\mathcal{B}$, and this is $\lesssim (\log m)(\log n)(mn)^{\frac{3}{4}}$ by choice of μ_0.

Likewise there are $\lesssim (\log m)(\log n)(mn)^{\frac{3}{4}}$ members of $\mathcal{R}$ which are rectangles of type $(\geq 1, \geq 1)$ for $(\mathcal{W}, \mathcal{B}_b)$. These contributions are small compared with $C_\epsilon (mn)^{\frac{3}{4}+\epsilon}$, so it remains to consider $(\mathcal{W}_g, \mathcal{B}_g)$.

Fix an appropriate r, which should be large compared with $(\log n)^{\frac{1}{\epsilon}}$, but small compared with both $\frac{n^{\frac{3}{4}}}{m^{\frac{1}{4}}\log(mn)}$ and $\frac{m^{\frac{1}{4}}}{n^{\frac{1}{12}}\log(mn)}$; this is possible by the assumption (13). Now apply Lemma 1.12 to $\mathcal{W}_g$ and $\mathcal{B}_g$ with this r. We first show that $|\mathcal{W}_g^*| < \frac{m}{100}$.

We know by Lemma 1.13 that $|\mathcal{W}_g^*| \lesssim \frac{r|\tilde{\mathcal{I}}(\mathcal{W}_g, \mathcal{B}_g)|}{n}$. Since $\mathcal{W}_g$ and $\mathcal{B}_g$ have no $(C_2\delta, t)$-rectangles of type $(\geq \mu_0, \geq 1)$ nor of type $(\geq 1, \geq \nu_0)$, we may apply Lemma 1.11 replacing δ with $C_2\delta$. We conclude that

$$\begin{aligned} |\tilde{\mathcal{I}}(\mathcal{W}_g, \mathcal{B}_g)| &\lesssim m\nu_0 \log \mu_0 + nm^{\frac{2}{3}}\mu_0^{\frac{1}{3}} \log \nu_0 \\ &\leq m^{\frac{5}{4}}n^{\frac{1}{4}} \log m + m^{\frac{3}{4}}n^{\frac{13}{12}} \log n \end{aligned}$$

and therefore

$$|\mathcal{W}_g^*| \lesssim r(m^{\frac{5}{4}}n^{-\frac{3}{4}} \log m + m^{\frac{3}{4}}n^{\frac{1}{12}} \log n) \tag{14}$$

The right side of (14) is small compared with m by choice of r, so $|\mathcal{W}_g^*| < \frac{m}{100}$ as claimed. We may therefore apply the inductive hypothesis obtaining that $\mathcal{W}_g^*$ and $\mathcal{B}_g$ have fewer than $C_\epsilon\left((\frac{mn}{1000})^{\frac{3}{4}+\epsilon}+m\log n+n\log m\right)$ rectangles of type $(\geq 1, \geq 1)$. Using (13) it follows that this is $\leq \frac{1}{10}C_\epsilon(mn)^{\frac{3}{4}+\epsilon}$.

Now consider the individual cells $\mathcal{W}_g^i$. For each of them, we let $\mathcal{B}_g^i$ be the circles $b\in\mathcal{B}_g$ for which there exists a rectangle $R\in\mathcal{R}$ tangent to b and to some $w\in\mathcal{W}_g^i$. By Lemma 1.5 and part (ii) of Lemma 1.12 we have $|\mathcal{B}_g^i|\lesssim\frac{n\log n}{r}$. Since r is large compared with $\log n$, we conclude by the inductive hypothesis that $\mathcal{W}_g^i$ and $\mathcal{B}_g^i$ have at most

$$C_\epsilon\left((C\frac{n\log n}{r}|\mathcal{W}_g^i|)^{\frac{3}{4}+\epsilon}+|\mathcal{W}_g^i|\log(C\frac{n\log n}{r})+C\frac{n\log n}{r}\log(|\mathcal{W}_g^i)\right)$$

rectangles of type $(\geq 1,\geq 1)$. Summing up, we bound the number of rectangles of type $(\geq 1,\geq 1)$ for $\cup_i\mathcal{W}_g^i$ and $\mathcal{B}_g$ by

$$C_\epsilon\left((C\frac{n\log n}{r})^{\frac{3}{4}+\epsilon}\sum_i|\mathcal{W}_g^i|^{\frac{3}{4}+\epsilon}+\log(C\frac{n\log n}{r})\sum_i|\mathcal{W}_g^i|+C\frac{n\log n}{r}\sum_i\log|\mathcal{W}_g^i|\right)\quad(15)$$

We use (i) of Lemma 1.12 and Holder's inequality to bound the first term, and we also use (i) of Lemma 1.12 to bound the third term. For the middle term, we use that r is large compared with $\log n$. This leads to

$$\begin{aligned}(15)\quad&\leq\quad C_\epsilon\left((C\frac{n\log n}{r})^{\frac{3}{4}+\epsilon}m^{\frac{3}{4}+\epsilon}(r^3a(r))^{\frac{1}{4}-\epsilon}+m\log n+C\frac{n\log n}{r}r^3a(r)\log m\right)\\&\leq\quad C_\epsilon\left(\frac{(C\log n)^{\frac{3}{4}+\epsilon}a(r)^{\frac{1}{4}-\epsilon}}{r^{4\epsilon}}(mn)^{\frac{3}{4}+\epsilon}+m\log n+Cr^2a(r)n\log n\log m\right)\end{aligned}$$

The last term is small compared with $(mn)^{\frac{3}{4}+\epsilon}$ since $r<\frac{m^{\frac{1}{4}}}{n^{\frac{1}{12}}}$ and $n\leq m$. The first term is small compared with $(mn)^{\frac{3}{4}+\epsilon}$ since r is large compared with $(\log n)^{\frac{1}{\epsilon}}$, so that the coefficient $\frac{(\log n)^{\frac{3}{4}+\epsilon}a(r)^{\frac{1}{4}-\epsilon}}{r^{4\epsilon}}$ is small. The middle term is also small compared with $(mn)^{\frac{3}{4}+\epsilon}$ since $m\leq n^3$. We conclude that (15) is small compared with $(mn)^{\frac{3}{4}+\epsilon}$, which concludes the proof. □

Proof of Lemma 1.4 This is now a simple random sampling argument. Let A be a large constant. Randomly choose $A\frac{n}{\nu}$ black and $A\frac{m}{\mu}$ white circles, with the convention that if $\nu\leq 2A$ we choose all the black circles, and if $\mu\leq 2A$ we choose all the white ones. Each fixed rectangle of type $(\geq\mu,\geq\nu)$ for $\mathcal{W}$ and $\mathcal{B}$ is then a $(\geq 1,\geq 1)$ rectangle for the samples with probability bounded below. Hence, if $\mathcal{R}$ is a set of pairwise incomparable (δ, t)-rectangles of type $(\geq \mu, \geq \nu)$ for $\mathcal{W}$ and $\mathcal{B}$ then with high probability $\mathcal{R}$ has a subset with proportional cardinality consisting of rectangles which are of type $(\geq 1, \geq 1)$ for the samples. Hence $|\mathcal{R}| \leq C_\epsilon \left((\frac{mn}{\mu\nu})^{\frac{3}{4}+\epsilon} + \frac{m}{\mu} \log \frac{n}{\nu} + \frac{n}{\nu} \log \frac{m}{\mu} \right)$ by Lemma 1.16. □

We will apply Lemma 1.3 by means of a certain slightly weaker statement, Lemma 1.17 below, which is formulated in terms of Euclidean rectangles instead of the curvilinear "rectangles" which were convenient in the preceding proof; we remark that no further use of the curvilinear rectangles will be made in this paper. For a suitable constant C, a (Euclidean) rectangle R with dimensions $\delta \times \delta^{\frac{1}{2}}$ is δ-tangent to a circle c if it is contained in the $C\delta$-neighborhood of c. If we are given collections $\mathcal{W}$ and $\mathcal{B}$ of white and black circles then R is of type $(\geq \mu, \geq \nu)$ if it is tangent to $\geq \mu$ white circles and to $\geq \nu$ black ones. Two rectangles are incomparable if neither is contained in a suitable fixed multiple of the other.

Lemma 1.17 Let $\mathcal{W}$ and $\mathcal{B}$ satisfy (5), (6), and (7). The cardinality of a set $\mathcal{R}$ of pairwise incomparable (Euclidean) $\delta \times \delta^{\frac{1}{2}}$-rectangles of type $(\geq \mu, \geq \nu)$ obeys the bound

$$|\mathcal{R}| \leq C_\epsilon t^{-\frac{1}{2}} \cdot \delta^{-\epsilon} \left((\frac{mn}{\mu\nu})^{\frac{3}{4}} + \frac{m}{\mu} + \frac{n}{\nu} \right) \tag{16}$$

Proof Assume at first that $\mathcal{W}$ and $\mathcal{B}$ also satisfy (8). It is not difficult to see that if C_1 was chosen large enough then for any $R \in \mathcal{R}$ there must be a $(C\delta, t)$-rectangle in the sense of Lemma 1.4 which contains R and which is tangent in the sense of Lemma 1.4 to any white or black circle to which R is tangent. Each such $(C\delta, t)$-rectangle contains $\lesssim t^{-\frac{1}{2}}$ incomparable R's, and the bound (16) therefore follows from Lemma 1.4. To eliminate the assumption $d(w, b) \leq 100t$, note that by (6) there must be a number $\tau \geq t$ such that $\tau \leq d(w, b) \leq 100\tau$ for all w and b, and apply the above reasoning replacing t by τ. □

Remark We take $t = 1$ and $m = n$, $\mu = \nu$ in this remark to avoid certain technical issues. The question arises as to whether or not the exponent $\frac{3}{2}$ in the bound

$$\forall \epsilon \, \exists C_\epsilon : \; |\mathcal{R}| \leq C_\epsilon \delta^{-\epsilon} (\frac{m}{\mu})^{\frac{3}{2}} \tag{17}$$

of Lemma 1.17 is sharp or not. This appears to be a difficult question. One can show that $\frac{3}{2}$ cannot be replaced by a number less than $\frac{4}{3}$, and it is also fairly clear that $\frac{3}{2}$ is the best exponent that can be obtained by an argument based on the techniques of [3]. See [19], p. 143.

2. A dual formulation

A δ-plate is a $\delta \times \delta^{\frac{1}{2}} \times 1$-rectangle whose longest axis is a light ray and whose second longest axis is tangent to the corresponding light cone, and a δ-tube is a $\delta \times \delta \times 1$-rectangle whose longest axis is a light ray. The direction of a tube or plate is the direction of its longest axis. This direction will always be of the form $(e, 1)$ with $e \in S^1$. We remark that it is sometimes necessary to replace the length parameter 1 by a fixed constant (e.g. this is the case in Lemma 2.2 below) and that we will often ignore technicalities of this kind. Two δ-plates or δ-tubes are comparable if one is contained in the dilate of the other by a fixed constant C, and they are parallel if their axis directions fail to be $C\delta$-separated for a suitable C. A family of δ-plates or δ-tubes is separated if no more than C are comparable to any given one. We also fix a small number $\epsilon_0 > 0$ and then a much smaller number ϵ.

Lemma 2.1 Let $\mathcal{W} \subset \mathbb{R}^3$ and $\mathcal{B} \subset \mathbb{R}^3$ be δ-separated sets of points; assume that $d(w, b) \geq t$ and that $d(w_1, w_2) \leq t$, $d(b_1, b_2) \leq t$. Let $\mathcal{R}$ be a separated set of δ-plates, each containing at least μ points of $\mathcal{W}$ and ν points of $\mathcal{B}$. Then for any $\epsilon > 0$

$$|\mathcal{R}| \lesssim t^{-\frac{1}{2}}\delta^{-\epsilon}\left((\frac{mn}{\mu\nu})^{\frac{3}{4}} + \frac{m}{\mu} + \frac{n}{\nu} \right)$$

In particular, this is $\lesssim \delta^{-\epsilon} t^{-\frac{1}{2}} (\frac{\max(m,n)}{\min(\mu,\nu)})^{\frac{3}{2}}$.

Proof Since the plates have bounded diameter we can assume that all the plates and points are contained in the region $1 \leq x_3 \leq C$ for suitable C. To any plate π we can associate a rectangle R_π in $\mathbb{R}^2$ as follows: extend the plate to an infinite $\delta \times \delta^{\frac{1}{2}}$-rectangular cylinder with the same axes, and then intersect this with the plane $x_3 = 0$. It is not difficult to see that if a point $(\overline{x}, x_3) \in \mathcal{W} \cup \mathcal{B}$ belongs to π, then the circle with center $\overline{x}$ and radius x_3 must be δ-tangent to this $\delta \times \delta^{\frac{1}{2}}$-rectangle in the manner discussed at the end of section 1. The result therefore follows from Lemma 1.17. □

Lemma 2.2 (i) Let $\mathcal{P}$ be a family of δ-plates. Then there is a separated family of $(C\delta)^{\frac{1}{2}}$-tubes $\mathcal{T}$ so that

• Each $\pi \in \mathcal{P}$ is contained in some $\tau(\pi) \in \mathcal{T}$, and for fixed τ the plates with $\tau(\pi) = \tau$ are all parallel.

(ii) Let $\mathcal{T}$ be a family of $\delta^{\frac{1}{2}}$-tubes. Then there is a separated family of $(C\delta)^{\frac{1}{2}}$-plates $\mathcal{A}$ so that

• Each $\tau \in \mathcal{T}$ is contained in some $\Pi(\tau) \in \mathcal{A}$. If $\Pi \in \mathcal{A}$ then the directions of the tubes with $\Pi(\tau) = \Pi$ will all be of the form $(e, 1)$ where e belongs to an arc of S^1 of length $C\delta^{\frac{1}{4}}$.

Proof This is almost a tautology. For (i), we choose for each π a roughly $\delta^{\frac{1}{2}}$-tube τ containing π, and observe that if τ is replaced by any comparable tube, then after dilation by by a fixed constant the comparable tube will still contain π. Now pass to a maximal separated subset and let $\mathcal{T}$ be the dilations of the tubes in the subset. The fact that plates π with a given $\tau(\pi)$ are all "parallel" in our sense is a simple property of light rays. Part (ii) can be seen in a similar way. $\square$

In, say, part (i) of the above lemma it is of course possible for π to be contained in more than one tube $\tau \in \mathcal{T}$; if so we fix a definite choice of $\tau(\pi)$. In the sequel, given a family of δ-plates $\mathcal{P}$, we will always fix a family of tubes $\mathcal{T}$ as in (i), and will denote $\mathcal{T}$ by $\mathcal{T}(\mathcal{P})$. For each $\tau \in \mathcal{T}(\Pi)$ we define

$$X_{\mathcal{P}}(\tau) = \{\pi \in \mathcal{P} : \tau(\pi) = \tau\}$$

Thus $\mathcal{P}$ is the disjoint union of the $X_{\mathcal{P}}(\tau)$'s where τ runs over $\mathcal{T}(\mathcal{P})$. Similarly, given a family of $\delta^{\frac{1}{2}}$-tubes, we will fix a family of $(C\delta)^{\frac{1}{2}}$-plates $\mathcal{A}$ as in (ii), and will denote $\mathcal{A}$ by $\mathcal{A}(\mathcal{T})$, and for $\Pi \in \mathcal{A}$ we will denote

$$Y_{\mathcal{T}}(\Pi) = \{\tau \in \mathcal{T} : \Pi(\tau) = \Pi\}$$

Let $\mathcal{W}$ be a set of points (of $\mathbb{R}^3$), $\mathcal{P}$ a set either of δ-plates or of δ-tubes. We let

$$\mathcal{I}(\mathcal{W}, \mathcal{P}) = \{(w, \pi) \in \mathcal{W} \times \mathcal{P} : w \in \pi\}$$

If t is fixed and if x is a point of $\mathbb{R}^3$ then we let $Q(x)$ be the t-cube containing x. In the sequel we will be working (analogously to [20]) with a relation $\sim$ between t-cubes for suitable t and elements of $\mathcal{P}$.

If $\sim$ is such a relation, and if $w \in \mathcal{W}$, $\pi \in \mathcal{P}$, then we use the notation $w \sim \pi$ to mean that $Q \sim \pi$, where Q is the t-cube containing w, and in addition $w \in \pi$.

A good incidence is a pair (w, π), $w \sim \pi$ and a bad incidence is a pair $(w, \pi) \in \mathcal{I}(\mathcal{W}, \mathcal{P})$, $w \not\sim \pi$. We use $\mathcal{I}_b$ for the set of bad incidences; sometimes we also use $\mathcal{I}_b(\sim)$ if we want to emphasize the dependence on the relation $\sim$. In what follows, we always assume that on a logarithmic scale t is large compared with $\delta^{\frac{1}{2}}$; in practice t will be set equal to δ^{ϵ_0}.

The following property will be crucial. We note here that the value of C_0 will be adjusted numerous times during the course of the proof.

Property (R): For each $\pi \in \mathcal{P}$ there are at most $(\log \frac{1}{\delta})^{C_0}$ t-cubes Q's such that $Q \sim \Pi$.

The next few lemmas construct relations with this property. The main results which we need in subsequent sections are Lemmas 2.7 and 2.8 below.

Lemma 2.3 (i) If $\mathcal{P}$ is a δ-separated family of δ-plates and $\mathcal{W}$ is a δ-separated set of points then there is a relation satisfying (R) and with $|\mathcal{I}_b| \lesssim \delta^{-C\epsilon} t^{-6} |\mathcal{W}||\mathcal{P}|^{\frac{1}{3}}$.

(ii) If $\mathcal{T}$ is a δ-separated family of δ-tubes and $\mathcal{W}$ is a δ-separated set of points then there is a relation satisfying (R) and with $|\mathcal{I}_b| \lesssim \delta^{-C\epsilon} t^{-5} |\mathcal{W}||\mathcal{P}|^{\frac{1}{2}}$.

Remark The exponents 6 and 5 in (i) and (ii) respectively can easily be improved substantially, but this would not lead to any improvement in the results of the paper. On the other hand, in part (a) it would be interesting to decide whether the exponent $\frac{1}{3}$ in (i) can be improved to $\frac{1}{p}$ for arbitrary $p < 4$ or even to $\frac{1}{p}$ for some fixed $p > 3$; this is the same issue we have discussed in the remark at the end of section 1.

Proofs From the form of the statements we can assume that all the plates and points are contained in a disc of fixed size.

(i) Define a relation $\sim$ as follows: for each π let $Q(\pi)$ be the t-cube Q for which $|\mathcal{W} \cap \pi \cap Q|$ is maximum (if there are several such, then pick one arbitrarily), and declare π to be related to $Q(\pi)$ and to its neighbors. Property (R) is clear, so it remains to estimate $|\mathcal{I}_b(\sim)|$.

By pigeonholing, we can find $\mathcal{P}' \subset \mathcal{P}$ and a positive integer ν so that the following hold:

1. $|\mathcal{P}'|\nu \gtrsim (\log \frac{1}{\delta})^{-1}|\mathcal{I}_b|$.
2. If $\pi \in \mathcal{P}'$ then $|\{w \in \mathcal{W} : (w,\pi) \in \mathcal{I}_b(\sim)\}| \in [\nu, 2\nu]$.

If $\pi \in \mathcal{P}'$, then π intersects $\lesssim t^{-1}$ t-cubes. Hence there is a t-cube $Q'(\pi)$ with $Q'(\pi) \not\sim \pi$ such that $|\mathcal{W} \cap Q'(\pi) \cap \pi| \gtrsim t\nu$. $Q(\pi)$ is at distance at least t from $Q'(\pi)$; and by the maximality property in its definition, $Q(\pi)$ must also satisfy $|\mathcal{W} \cap Q(\pi) \cap \pi| \gtrsim t\nu$.

In view of the assumption that everything is contained in a fixed disc there are $\lesssim t^{-6}$ possible pairs of t-cubes (Q, Q'), so some pair must be $(Q(\pi), Q'(\pi))$ for $\gtrsim t^6|\mathcal{P}'|$ choices of π. Applying Lemma 2.1 to $\mathcal{W} \cap Q(\pi)$ and $\mathcal{W} \cap Q'(\pi)$, we obtain

$$\begin{aligned} t^6|\mathcal{P}'| &\lesssim \delta^{-\epsilon} t^{-\frac{1}{2}} (\frac{|\mathcal{W}|}{t\nu})^{\frac{3}{2}} \\ &\lesssim \delta^{-\epsilon} t^{-\frac{1}{2}} (\frac{|\mathcal{W}|}{t|\mathcal{I}_b|/|\mathcal{P}'|\log\frac{1}{\delta}})^{\frac{3}{2}} \end{aligned}$$

so that

$$\begin{aligned} |\mathcal{I}_b| &\lesssim \log\tfrac{1}{\delta}\delta^{-\frac{2}{3}\epsilon} t^{-\frac{16}{3}} |\mathcal{W}||\mathcal{P}'|^{\frac{1}{3}} \\ &\lesssim \delta^{-\epsilon} t^{-6} |\mathcal{W}||\mathcal{P}|^{\frac{1}{3}} \end{aligned}$$

as claimed.

(ii) This is done the same way, but we need a substitute for Lemma 2.1. For this we use the following easy bound: suppose we have two δ-separated sets of points $\mathcal{W}$ and $\mathcal{B}$ with $d(w,b) \geq t$ for each $w \in \mathcal{W}, b \in \mathcal{B}$, and a set of tubes $\mathcal{T}$ such that each $\tau \in \mathcal{T}$ is incident to μ points or $\mathcal{W}$ and to ν points of $\mathcal{B}$. Then

$$|\mathcal{T}| \lesssim t^{-1} \frac{|\mathcal{W}||\mathcal{B}|}{\mu\nu} \tag{18}$$

To prove (18) we argue as follows (see [19], p. 137-8). Consider all triples (τ, w, b) where $\tau \in \mathcal{T}$, $w \in \mathcal{W}$, $b \in \mathcal{B}$ and τ is incident to both w and b. There are at least $|\mathcal{T}|\mu\nu$ such triples but at most $t^{-1}|\mathcal{W}||\mathcal{B}|$, since two t-separated points can have at most t^{-1} tubes[1] in common. This proves (18).

The proof of (ii) is now entirely analogous to the proof of (i). We define $\sim$ via: for each $\tau \in \mathcal{T}$ let $Q(\tau)$ be the t-cube Q for which $|\mathcal{W} \cap Q \cap \tau|$ is maximum. Declare τ

[1]The bound is t^{-1} instead of t^{-2} because of the light ray structure.

to be related to $Q(\tau)$ and to its neighbors. Let $\mathcal{I}_b$ be the set of all bad incidences. We can find $\mathcal{T}' \subset \mathcal{T}$ and ν so $|\mathcal{T}'|\nu \gtrsim (\log\frac{1}{\delta})^{-1}|\mathcal{I}_b|$ and each plate $\tau \in \mathcal{T}'$ has about ν bad incidences. Hence there is a cube $Q'(\pi)$ containing $\gtrsim t\nu$ points $w \in \mathcal{W}$ such that (w,π) is a bad incidence. Some pair of t-cubes must be $(Q(\pi), Q'(\pi))$ for at least $t^6|\mathcal{T}'|$ choices of τ. Applying (18), we obtain

$$\begin{aligned} t^6|\mathcal{T}'| &\lesssim t^{-1}\left(\frac{|\mathcal{W}|}{t\nu}\right)^2 \\ &\lesssim t^{-1}\left(\frac{|\mathcal{W}|}{t|\mathcal{I}_b|/|\mathcal{T}'|\log\frac{1}{\delta}}\right)^2 \end{aligned}$$

so that

$$\begin{aligned} |\mathcal{I}_b| &\lesssim \log\tfrac{1}{\delta} t^{-\frac{9}{2}}|\mathcal{W}||\mathcal{T}'|^{\frac{1}{2}} \\ &\lesssim \delta^{-\epsilon}t^{-5}|\mathcal{W}||\mathcal{T}|^{\frac{1}{2}} \end{aligned}$$

□

Lemma 2.3 leads immediately to a certain refinement of itself. If $\sim$ is a relation as above, and h is a number then we define $\mathcal{W}^h = \{w \in \mathcal{W} : w$ belongs to between $\frac{h}{2}$ and h plates from $\mathcal{P}\}$, and $\mathcal{I}_b^h = \{(w,\pi) \in \mathcal{I}_b : w \in \mathcal{W}^h\}$. We also make the analogous definitions for tubes.

Lemma 2.4 (i) If $\mathcal{P}$ is a δ-separated family of δ-plates and $\mathcal{W}$ is a δ-separated set of points then there is a relation satisfying (R) and with $|\mathcal{I}_b^h| \lesssim \delta^{-C\epsilon}t^{-6}|\mathcal{W}^h||\mathcal{P}|^{\frac{1}{3}}$ for all h.

(ii) If $\mathcal{T}$ is a δ-separated family of δ-tubes and $\mathcal{W}$ is a δ-separated set of points then there is a relation satisfying (R) and with $|\mathcal{I}_b| \lesssim \delta^{-C\epsilon}t^{-5}|\mathcal{W}^h||\mathcal{T}|^{\frac{1}{2}}$ for all h.

Proof We do (i); (ii) is exactly the same. We may clearly assume h is a power of 2, and also that $h \le C\delta^{-\frac{1}{2}}$ since otherwise $\mathcal{W}^h = \emptyset$. For each dyadic integer $h = 2^j \le C\delta^{-\frac{1}{2}}$ we apply Lemma 2.5 to $\mathcal{P}$ and $\mathcal{W}^h$, obtaining a sequence of relations $\sim_j$. We then define $Q \sim \pi$ if $q \sim_j \pi$ for some j. Property (R) will still hold, since there are $\lesssim \log\frac{1}{\delta}$ values of h, and the rest follows from the definitions and the bound in Lemma 2.3. □

We now prove a result for tubes which is stronger than the one in Lemmas 2.3(ii) and 2.4(ii); this will follow roughly by combining Lemma 2.4(i) for the family of plates $\mathcal{A}(\mathcal{T})$ and Lemma 2.4(ii) for the families of tubes $Y_{\mathcal{T}}(\Pi)$.

Lemma 2.5 Let $\mathcal{W} \subset \mathbb{R}^3$ be a $\delta^{\frac{1}{2}}$-separated set of points; $\mathcal{T}$ a family of $\delta^{\frac{1}{2}}$-tubes with cardinality k; assume that for each plate $\Pi \in \mathcal{A}(\mathcal{T})$ we have $|Y_{\mathcal{T}}(\Pi)| \leq m$. Then there is a relation $\sim$ satisfying (R) and with

$$|\mathcal{I}_b^\mu| \lesssim \delta^{-C\epsilon} t^{-6} m^{\frac{1}{12}} k^{\frac{1}{6}} \mu^{\frac{1}{2}} |\mathcal{W}^\mu| \tag{19}$$

for each μ.

Proof We assume at first that

$$|Y_{\mathcal{T}}(\Pi)| \in [\frac{m}{2}, m]$$

for each $\Pi \in \mathcal{A}(\mathcal{T})$. Notice that this implies

$$|\mathcal{A}(\mathcal{T})| \lesssim \frac{k}{m} \tag{20}$$

We also first fix μ and construct a relation $\sim$ satisfying (19) for the given value of μ. This relation $\sim$ is defined as follows.

1. Apply Lemma 2.4(i) to the family $\mathcal{A}(\mathcal{T})$ obtaining a relation $\sim_1$ between plates from $\mathcal{A}(\Pi)$ and t-cubes.
2. Apply Lemma 2.4(ii) to the family $Y_{\mathcal{T}}(\Pi)$ for each $\Pi \in \mathcal{A}(T)$ obtaining, for each Π, a relation $\sim_\Pi$ between tubes from $Y_{\mathcal{T}}(\Pi)$ and t-cubes.
3. Define $Q \sim \tau$ if either $Q \sim_1 \Pi(\tau)$ or $Q \sim_{\Pi(\tau)} \tau$.

Property (R) is clear from the definition 3. and the corresponding property for $\sim_1$ and $\sim_\Pi$. We now prove (19).

Fix a positive integer a and define

$$\mathcal{W}^{\mu,a,\Pi} = \{w \in \mathcal{W}^\mu : \sum_{\tau \in Y_{\mathcal{T}}(\Pi)} \chi_\tau(w) \geq a\}$$

$$\mathcal{I}_b(a,\Pi) = \{(w,\tau) : \tau \in Y_{\mathcal{T}}(\Pi), w \in \mathcal{W}^{\mu,a,\Pi}, w \in \tau \text{ and } w \not\sim_\Pi \tau\}$$

By Lemma 2.4(ii) we have

$$|\mathcal{I}_b(a,\Pi)| \lesssim \delta^{-C\epsilon} t^{-5} m^{\frac{1}{2}} |W^{\mu,a,\Pi}|$$

Hence also

$$\begin{aligned}\sum_{\Pi\in\mathcal{A}(\mathcal{T})}|\mathcal{I}_b(a,\Pi)| &\lesssim \delta^{-C\epsilon}t^{-5}m^{\frac{1}{2}}\sum_{\Pi}|W^{\mu,a,\Pi}| \\ &\lesssim \delta^{-C\epsilon}t^{-5}m^{\frac{1}{2}}\frac{\mu}{a}|\mathcal{W}^{\mu}| \qquad (21)\end{aligned}$$

The last line followed since a point of $\mathcal{W}^{\mu}$ is in $W^{\mu,a,\Pi}$ for $\lesssim \frac{\mu}{a}$ choices of Π.

Now fix a and b and define $\mathcal{W}^{\mu}(a,b)$ to be the set of points $w\in\mathcal{W}^{\mu}$ such that, for at least b plates $\Pi\in\mathcal{A}(\mathcal{T})$, we have both

(i). $(w,\Pi)\in\mathcal{I}_b(\sim_1)$ and
(ii). There are at least a tubes $\tau\in Y_{\mathcal{T}}(\Pi)$ such that (w,τ) is a $\sim_{\Pi}$-bad incidence.

Notice that if w and Π are as in (i), (ii), then $w\in\mathcal{W}^{\mu,a,\Pi}$, and furthermore if τ is as in (ii), then $(w,\tau)\in\mathcal{I}_b(a,\Pi)$. We may therefore estimate $W^{\mu}(a,b)$ using (21). We obtain

$$\begin{aligned}|\mathcal{W}^{\mu}(a,b)| &\leq (ab)^{-1}\sum_{\Pi}|\mathcal{I}_b(a,\Pi)| \\ &\lesssim \delta^{-C\epsilon}t^{-5}m^{\frac{1}{2}}(a^2b)^{-1}\mu|\mathcal{W}^{\mu}| \qquad (22)\end{aligned}$$

On the other hand, using property (i), and then Lemma 2.4(i) and estimate (20) we obtain

$$\begin{aligned}|\mathcal{W}^{\mu}(a,b)| &\leq b^{-1}|\mathcal{I}_b(\sim_1)| \\ &\lesssim \delta^{-C\epsilon}t^{-6}b^{-1}(\frac{k}{m})^{\frac{1}{3}}|\mathcal{W}(\mu)| \qquad (23)\end{aligned}$$

Now define

$$\mathcal{W}^{\mu}(\nu)=\{w\in\mathcal{W}^{\mu}:|\{\tau:(w,\tau)\in\mathcal{I}_b(\sim)\}|\geq\nu\}$$

If $w\in\mathcal{W}^{\mu}(\nu)$ then by pigeonholing there are dyadic values of a and b with $ab\approx(\log\frac{1}{\delta})^{-1}\nu$ such that, for at least b plates $\Pi\in\mathcal{A}(\mathcal{T})$, there are at least a tubes $\tau\in Y_{\mathcal{T}}(\Pi)$ such that $(w,\tau)\in\mathcal{I}_b(\sim)$. Thus w belongs to $\mathcal{W}^{\mu}(a,b)$ with these values of a and b. Using (23), (22) and summing over dyadic a we get that

$$\begin{aligned}\nu|\mathcal{W}^{\mu}(\nu)| &\lesssim \sum_a \min\left(\delta^{-C\epsilon}t^{-6}a(\frac{k}{m})^{\frac{1}{3}}|\mathcal{W}(\mu)|, \delta^{-C\epsilon}t^{-5}m^{\frac{1}{2}}a^{-1}\mu|\mathcal{W}(\mu)|\right) \\ &\lesssim \delta^{-C\epsilon}t^{-6}\left(a_0(\frac{k}{m})^{\frac{1}{3}}+\frac{m^{\frac{1}{2}}\mu}{a_0}\right)|\mathcal{W}(\mu)|\end{aligned}$$

for any given a_0. Optimizing in regard to a_0 we get

$$\nu|\mathcal{W}^\mu(\nu)| \lesssim \delta^{-C\epsilon} t^{-6} m^{\frac{1}{12}} k^{\frac{1}{6}} \mu^{\frac{1}{2}} |\mathcal{W}^\mu|$$

Summing over dyadic ν we obtain

$$\begin{aligned} |\mathcal{I}_b^\mu| &\lesssim \sum_\nu \nu |\mathcal{W}^\mu(\nu)| \\ &\lesssim \delta^{-C\epsilon} t^{-6} m^{\frac{1}{12}} k^{\frac{1}{6}} \mu^{\frac{1}{2}} |\mathcal{W}^\mu| \end{aligned}$$

which is (19).

We can clearly now construct a relation $\sim$ satisfying (19) for <u>all</u> μ by the argument in the proof of Lemma 2.4, so it remains only to remove the assumption that $|Y_{\mathcal{T}}(\Pi)| \geq \frac{m}{2}$. For this, we partition $\mathcal{T}$ as $\cup_{j:2^j \leq m} \mathcal{T}_j$ where

$$\mathcal{T}_j = \cup_{\substack{\Pi \in \mathcal{A}(\mathcal{T}) \\ |Y_{\mathcal{T}}(\Pi)| \in (2^{j-1}, 2^j]}} Y_{\mathcal{T}}(\Pi)$$

Then we apply what we have already done to construct a relation $\sim_j$ between tubes from $\mathcal{T}_j$ and t-cubes satisfying (R) and satisfying (19) with m replaced by 2^j on the right hand side. Then we can define a relation $\sim$ between $\mathcal{T}$ and t-cubes by taking the union of the relations $\sim_j$. It clearly satisfies (R), and it satisfies (19) in view of the favorable dependence on m there. $\square$

Notice that the hypothesis about the cardinality of $X(\Pi)$ is always satisfied with $m \approx \delta^{-\frac{1}{2}}$, since a $C\delta^{\frac{1}{2}}$-plate contains $\mathcal{O}(\delta^{-\frac{1}{2}})$ separated $\delta^{\frac{1}{2}}$-tubes. Thus the following statement is immediate from (19):

<u>Lemma 2.6</u> Let $\mathcal{W}$ be a $\sqrt{\delta}$-separated set of points; $\mathcal{T}$ a family of $\delta^{\frac{1}{2}}$-tubes with cardinality k. Then there is a relation satisfying (R) and with

$$|\mathcal{I}_b^\mu| \lesssim \delta^{-C\epsilon} t^{-6} \delta^{-\frac{1}{24}} k^{\frac{1}{6}} \mu^{\frac{1}{2}} |\mathcal{W}^\mu| \tag{24}$$

for each μ.

What we actually need in the sequel is versions of the above lemmas incorporating "Schwartz tails", which can now be obtained in a routine way.

We make some further definitions. If $\mathcal{P}$ is a family of tubes or plates we define

$$\Phi_{\mathcal{P}} = \sum_{\pi \in \mathcal{P}} \phi_\pi \tag{25}$$

where ϕ_π was defined after the introduction. Given a relation $\sim$ between t-cubes and elements of $\mathcal{P}$, we let

$$\Phi^b_{\mathcal{P}}(x) = \sum_{\pi \not\sim Q(x)} \phi_\pi(x) \tag{26}$$

be the corresponding sum involving only tubes unrelated to the t-cube $Q(x)$ containing x. We note that if $\mathcal{P}$ is a set of δ-plates then for any δ-cube Δ we have

$$\max_\Delta \Phi_{\mathcal{P}} \leq C \min_\Delta \phi_{\mathcal{P}} \tag{27}$$

where C depends only on the choice of M. There is a corresponding fact for families of $\delta^{\frac{1}{2}}$-tubes, where Δ may then be taken to be a $\delta^{\frac{1}{2}}$-cube.

If $\mathcal{W}$ is a set of points, $\mathcal{T}$ a set of tubes or plates, $\sim$ a relation as above, $\mu > 0$ then we define

$$\mathcal{J}_b = \sum_{w \in \mathcal{W}} \Phi^b_{\mathcal{T}}(w)$$

Lemma 2.7 Let $\mathcal{W}$ be a $\sqrt{\delta}$-separated set, $\mathcal{T}$ a set of $C\sqrt{\delta}$-tubes with cardinality k, μ a number, assume that $\Phi_{\mathcal{T}} \leq \mu$ on $\mathcal{W}$. Then there is a relation between t-cubes and elements of $\mathcal{T}$ satisfying (R) and with

$$|\mathcal{J}^\mu_b| \leq \delta^{-C\epsilon} t^{-6} \delta^{-\frac{1}{24}} k^{\frac{1}{6}} \mu^{\frac{1}{2}} |\mathcal{W}| + \delta^{100} |\mathcal{W}| \tag{28}$$

Proof If $T \geq 1$ is dyadic then subdivide in $T\sqrt{\delta}$ cubes and define $\mathcal{W}^T$ to be a subset of $\mathcal{W}$ containing (exactly) one point in each such cube which contains a point of $\mathcal{W}$; thus $\mathcal{W}^T$ is essentially a $T\sqrt{\delta}$-separated set. If $w \in \mathcal{W}$ then we use below the notation w_T for the point of $\mathcal{W}^T$ which belongs to the same $T\sqrt{\delta}$ cube as w.

If $\tau \in \mathcal{T}$ then let τ^T be the dilation of τ by T and $\mathcal{T}^T$ the family $\{\tau^T\}$. $\mathcal{T}^T$ is evidently not separated, so we let $\mathcal{T}^T_0$ be a maximal separated subset. It is then clear that, for any $\tau \in \mathcal{T}$, τ^T will be contained in a suitable fixed dilate τ^{AT}_*, where $\tau^T_* \in \mathcal{T}^T_0$, and that there are only boundedly many possible choices for τ^*.

We now apply a rescaled version of Lemma 2.6 to $\mathcal{T}^T_0$ and $\mathcal{W}^T$ for each dyadic T with $T\sqrt{\delta} \leq \delta^{C\epsilon_0}$ for a suitable constant C, obtaining a family of relations which we call $\simeq_T$; thus $\simeq_T$ is a relation between tubes from $\mathcal{T}_0^T$ and t-cubes. We then define a corresponding relation $\sim_T$ between tubes from $\mathcal{T}$ and t-cubes via $\tau \sim Q$ if $\tau_*^T \simeq_T Q$ for some $\beta^T \in \mathcal{T}_0^T$ with $\tau^T \subset \tau_*^{AT}$, and we define $\tau \sim Q$ if $\tau \sim_T Q$ for some T. Since there are only logarithmically many dyadic scales we see that the relation $\sim$ has property (R). It remains to check (28).

Suppose that $w \in \mathcal{W}$ and that $\tau \not\sim Q(w)$. We express

$$\phi_\tau(w) \leq \sum_{T: T\sqrt{\delta} \leq \delta^{C\epsilon_0}} T^{-M} \chi_{\tau_T}(w)$$

plus a small error. The error is bounded by a high negative power of the quantity $\delta^{-1} + \mathrm{dist}(w, \tau)$, and is therefore negligible for present purposes since the right side of (28) is $\geq \delta^{100}|\mathcal{W}|$; we will in fact neglect this error term in the calculations below.

For each given τ and T choose τ_*^T as above. We know that τ_*^T is not $\simeq_T$-related to the t-cube containing w. Accordingly, we can bound

$$\Phi^b_{\mathcal{T}}(w) \leq \sum_T T^{-M} \sum_\tau |\{\tau_*^T : (w_T, \tau_*^T) \text{ is a } \simeq_T\text{-bad incidence}\}|$$

Any given point can arise as w_T for at most $\mathcal{O}(T^3)$ choices of w, and any given tube in $\mathcal{T}_0^T$ can arise as τ_T^* for at most $\mathcal{O}(T^3)$ choices of τ. Accordingly,

$$\mathcal{J}_b \leq \sum_T T^{6-M} |\{(x, \sigma) \in \mathcal{W}^T \times \mathcal{T}_0^T : (x, \sigma) \text{ is a } \simeq_T\text{-bad incidence}\}| \tag{29}$$

Further, if $w \in \mathcal{W}$ then w is at most a $C\mu T^M$-fold point for $\mathcal{T}^T$, since $\Phi_{\mathcal{T}}(x) \leq \mu$. *A fortiori* x is at most a $C\mu T^M$-fold point for $\mathcal{T}_0^T$. We conclude using Lemma 2.6 for the relation $\simeq_T$ (and (29)) that

$$\mathcal{J}_b \leq \sum_T T^{6-M} \delta^{-C\epsilon} t^{-6} \delta^{-\frac{1}{24}} k^{\frac{1}{6}} (T^M \mu)^{\frac{1}{2}} |\mathcal{W}| \tag{30}$$

We have overestimated in several of the factors here; for example, k bounds the cardinality of $\mathcal{T}_0^T$.

For large M we have a favorable dependence on T here, and may sum up to obtain the lemma. □

Lemma 2.7 was of course based on Lemma 2.6. We can do the same in the context of Lemma 2.3(i), obtaining the following lemma whose detailed proof we omit.

Lemma 2.8 Let $\mathcal{P}$ be a set of δ-plates; $\mathcal{W}$ a δ-separated set in $\mathbb{R}^3$. Then there is a relation satisfying (R) and with

$$\mathcal{J}_b \leq \delta^{-C\epsilon} t^{-6} |\mathcal{W}|\, |\mathcal{P}|^{\frac{1}{3}}$$

$\square$

In the sequel, we will apply Lemma 2.7 to $\mathcal{T}(\mathcal{P})$ where $\mathcal{P}$ is a set of δ-plates. We will be using the following terminology. If $\mathcal{P}$ is a set of δ-plates then we will say that $\mathcal{P}$ is type r if for each $\tau \in \mathcal{T}(\mathcal{P})$ the cardinality of $X_{\mathcal{P}}(\tau)$ is between r and $2r$. For a general $\mathcal{P}$, the type r component of $\mathcal{P}$ is the subset $\mathcal{P}_r \subset \mathcal{P}$ defined by

$$\mathcal{P}_r = \cup (X_{\mathcal{P}}(\tau) : r \leq |X_{\mathcal{P}}(\tau)| \leq 2r)$$

Evidently $\mathcal{P}_r$ is of type r.

We insert the following fact here for future reference.

Lemma 2.9 Assume that $\mathcal{P}$ is type r. Then for any $\delta^{\frac{1}{2}}$-cube Q we have

$$\int \Phi_{\mathcal{P}} \phi_Q \lesssim \delta^{\frac{1}{2}} r \int \Phi_{\mathcal{T}(\mathcal{P})} \phi_Q$$

Proof First consider the version of this where we ignore the tail of ϕ_Q, i.e. the estimate

$$\int_Q \Phi_{\mathcal{P}} \lesssim \delta^{\frac{1}{2}} r \int_Q \Phi_{\mathcal{T}(\mathcal{P})} \tag{31}$$

If a tube τ contains r parallel plates π_j, then we have $\sum_{j=1}^{r} \phi_{\pi_j}(x) \lesssim r\delta^{\frac{1}{2}} \phi_\tau(x)$ if x is not in the double of the infinite tube coaxial with τ. This gives (31) if Q doesn't intersect the double of the infinite tube coaxial with τ. On the other hand, if Q is a $\delta^{\frac{1}{2}}$-cube intersecting τ then $\int_Q \phi_\tau \gtrsim |Q|$ and $\int_Q \phi_{\pi_j} \lesssim \delta^{\frac{1}{2}} |Q|$ for each j, and a corresponding statement is valid if Q intersects the infinite tube coaxial with τ since ϕ_τ and ϕ_{π_j} die at the same rate as one moves away from τ in its axis direction. This gives (31) in the remaining cases. The lemma follows since ϕ_Q is essentially a sum of indicator functions of $\delta^{\frac{1}{2}}$-cubes. $\square$.

In addition, we will need to deal with certain technical issues involving the distinction between measure and entropy; the following definition and lemma give an easy way of doing this.

Definition A set E is δ-evenly covered if there is a constant a such that $|E \cap \Delta| \in [a\delta^3, 2a\delta^3]$ for each δ-cube Δ which intersects E.

Notice that if E is δ-evenly covered and if $F \subset E$ then we have

$$|F| \lesssim a\delta^3 \mathcal{E}_\delta(F) \tag{32}$$

Conversely, if F is the intersection of E with the union of some collection of δ-cubes then we have

$$|F| \gtrsim a\delta^3 \mathcal{E}_\delta(F) \tag{33}$$

Lemma 2.10 Let E be a set with the property that for each 1-cube Q, $E \cap Q$ is either empty or of measure $\geq \delta^{100000}$. Then E has a δ-evenly covered subset E' with $|E'| \gtrsim (\log \frac{1}{\delta})^{-1} |E|$.

Proof This is just a pigeonhole argument. Let S_n be the union of the δ-cubes Δ with the property that $|\Delta \cap E| \in [2^{-(n+1)}\delta^3, 2^{-n}\delta^3]$. Then $S_n = \emptyset$ if $n < 0$. Let $N = C(\log \frac{1}{\delta})^{-1}$ for a suitable C. Then $|(\cup_{n \geq N} S_n) \cap E| \leq 2^{-N} \leq \frac{1}{2}|E|$ by the hypothesis concerning the measure of the intersection of E with 1-cubes. So there must be a choice of $n \leq N$ such that $|S_n \cap E| \gtrsim (\log \frac{1}{\delta})^{-1}|E|$. We now set $E' = S_n \cap E$. $\square$

3. Localization property

In this section and subsequently, we make the convention that the parameters N and δ always satisfy $N = \delta^{-1}$. In this section we always let $t = \delta^{\epsilon_0}$ where ϵ_0 was defined after the introduction.

If π is a $1 \times \delta^{\frac{1}{2}} \times \delta$-plate, with respective axes e_1, e_2, e_3, then we let π_* be a rectangle and centered at the point Ne_3 with axes e_1, e_2, e_3 and respective axis lengths $C_1, C_1 N^{\frac{1}{2}}, N$, where C_1 is a large constant. Thus π^* is approximately dual to π and is contained in $\Gamma_N(C)$ for suitable C.

An N-function is a function f which has a decomposition

$$f = \sum_{\pi \in \mathcal{P}} f_\pi$$

where $\mathcal{P} = \mathcal{P}(f)$ is a separated family of δ-plates and

1. $|f_\pi| \le \phi_\pi$.
2. $\widehat{f_\pi}$ is supported on π^*.

Such a decomposition is of course not unique; however, given an N-function we will always fix a family of plates for f, and the associated functions f_π. The properties of plate families derived in section 2 should evidently translate into properties of N-functions and we now carry this out. We note that we will use cancellation here, via the fact that functions with disjoint Fourier supports are orthogonal.

Let f be an N-function with family of plates $\mathcal{P}$, and let $\tilde{\mathcal{P}}$ be a subset of $\mathcal{P}$. The subfunction of f corresponding to $\tilde{\mathcal{P}}$ is $f_{\tilde{\mathcal{P}}} \stackrel{def}{=} \sum_{\pi \in \tilde{\mathcal{P}}} f_\pi$.

Let $\mathcal{W}_\lambda = \{x : |f(x)| \ge \lambda\}$. We say that $(f, \mathcal{W}_\lambda)$ localizes if there are a collection of functions f_Q (Q runs over t-cubes) such that

1. Each f_Q is the restriction to Q of the subfunction of f corresponding to a certain collection of plates $\mathcal{P}_Q$, and

$$\sum_Q |\mathcal{P}_Q| \le (\log \tfrac{1}{\delta})^C |\mathcal{P}| \tag{34}$$

2. The (Lebesgue) measure of the set $\cup_Q \{x \in Q : |f_Q(x)| \ge (\log \frac{1}{\delta})^{-C} \lambda\}$ is at least $(\log \frac{1}{\delta})^{-C} |\mathcal{W}_\lambda|$.

In the subsequent arguments we will use the terminology "A $(\log \frac{1}{\delta})^{-C}$-fraction of $\mathcal{W}$". This means a subset of $\mathcal{W}$ with measure $\ge (\log \frac{1}{\delta})^{-C} |\mathcal{W}|$.

The following lemmas are what we have been leading up to.

Lemma 3.1 Let f be an N-function with plate family $\mathcal{P}$ and assume that

$$|\mathcal{P}| \le t^{300} \lambda^3 \tag{35}$$

Then $(f, \mathcal{W}_\lambda)$ localizes.

Before giving the proof we record a simple uncertainty principle estimate which will be used several times below. Namely, suppose that supp$\hat{f}$ is contained in $|\xi| \le N$ and that $|f(a)| \geq \lambda$. Then, for a suitable fixed constant C

$$\text{If } |x-a| \leq (CN\|f\|_\infty)^{-1}\lambda \text{ then } |f(x)| \geq \frac{\lambda}{2} \tag{36}$$

by Bernstein's inequality and the mean value theorem.

Proof of Lemma 3.1 Let $k = |\mathcal{P}(f)|$, and notice that since k is an integer we necessarily have $\lambda > 1$. Since $|f_\pi| \leq \phi_\pi$ and $\mathcal{P}$ is separated we also have $\lambda \leq \Phi_{\mathcal{P}}(x) \lesssim \delta^{-\frac{1}{2}}$ if $x \in \mathcal{W}_\lambda$. We now make a "pigeonhole" type reduction.

Claim There are a value of $\mu \in [\frac{\lambda}{2}, C\delta^{-\frac{1}{2}}]$ and a set $\mathcal{W} \subset \mathcal{W}_{\frac{\lambda}{2}}$ so that the following hold:

1. $|\mathcal{W}| \geq (\log\frac{1}{\delta})^{-C}|\mathcal{W}_\lambda|$.
2. If $x \in \mathcal{W}$ then $\Phi_{\mathcal{P}}(x) \in [\mu, 2\mu]$.
3. $\mathcal{W}$ is δ-evenly covered.

For this, we show first that there is a set $\mathcal{W}_0 \subset \mathcal{W}_{\frac{\lambda}{2}}$ such that $|\mathcal{W}_0| \geq (\log\frac{1}{\delta})^{-C}|\mathcal{W}_\lambda|$ and

($*$) $\mathcal{W}_0$ intersects each 1-cube either in measure $\geq \delta^5$ or not at all.

Namely, we have $\|f\|_\infty \lesssim \delta^{-\frac{1}{2}}$, so, by (36), if $\mathcal{W}_\lambda$ intersects a 1-cube Q then $|\mathcal{W}_{\frac{\lambda}{2}} \cap Q| \gtrsim \delta^{\frac{9}{2}}$. We may therefore take $\mathcal{W}_0$ to be the set $\cup(Q \cap \mathcal{W}_{\frac{\lambda}{2}} : Q \cap \mathcal{W}_\lambda \neq \emptyset)$.

Next we choose a value of $\mu \geq \lambda$ so that $\mathcal{W}_1$, a $\gtrsim (\log\frac{1}{\delta})^{-1}$ proportion of $\mathcal{W}_0$ consists of points with $\Phi_{\mathcal{P}}(x) \in [\mu, 2\mu]$. We discard from $\mathcal{W}_1$ the set $\cup(Q \cap \mathcal{W}_1 : |Q \cap \mathcal{W}_1| \leq \delta^{100}\}$, where Q runs over 1-cubes, and denote the remaining set by $\mathcal{W}_2$. It is easily seen using ($*$) that $\mathcal{W}_2$ still has measure $\gtrsim (\log\frac{1}{\delta})^{-1}|\mathcal{W}_0|$. We then apply Lemma 2.10 to obtain a δ-evenly covered subset and let $\mathcal{W}$ be this subset. That proves the claim.

We apply Lemma 2.8 to the plate family $\mathcal{P}$ and a maximal δ-separated subset of $\mathcal{W}$, obtaining a relation $\sim$ which satisfies $\int_X \Phi_{\mathcal{P}}^b \leq \delta^{-C\epsilon}t^{-6}k^{\frac{1}{3}}|X|$. Here X is the union of the δ-cubes containing points of $\mathcal{W}$, and we have used (27). By the even covering property we can replace X by $\mathcal{W}$ here; see (32), (33). Thus

$$\int_{\mathcal{W}} \Phi_{\mathcal{P}}^b \leq \delta^{-C\epsilon}t^{-6}k^{\frac{1}{3}}|\mathcal{W}|$$

It follows by the assumption (35) that

$$\int_{\mathcal{W}} \Phi^b_{\mathcal{P}} \leq t^{90}\lambda|\mathcal{W}|$$

Accordingly there is a subset $\mathcal{W}^* \subset \mathcal{W}$ with proportional measure and so that at each point $x \in \mathcal{W}^*$ we have, say,

$$\Phi^b_{\mathcal{P}}(x) \leq t\lambda \tag{37}$$

Now we define two functions f^* and h as follows: On each given t-cube Q

$$f^* = \sum_{\pi \sim Q} f_\pi$$

$$h = \sum_{\pi \not\sim Q} f_\pi$$

Notice that even though f^* is not a subfunction of f, nevertheless its restriction to any t-cube Q is the restriction to Q of a subfunction of f. Evidently $f^* + h = f$, and $|h|$ is $\lesssim \Phi^b_{\mathcal{P}}$ and therefore much less than λ on $\mathcal{W}^*$ so $|f^*|$ must be $\geq \frac{\lambda}{2}$ on $\mathcal{W}^*$. Let f_Q be f^* restricted to Q, and $\mathcal{P}_Q$ the plates related to Q. The bound (34) follows from the property (R), so we are done. $\square$

Lemma 3.1 does not suffice for what we want to do, because the assumption (35) is too strong[2]. Accordingly we now give another lemma of the same type; notice that the hypothesis (38) below is considerably weaker than (35) when λ is close to its maximum possible value $\delta^{-\frac{1}{2}}$.

<u>Lemma 3.2</u> Let f be an N-function with plate family $\mathcal{P}$ and assume that

$$|\mathcal{P}| \leq t^{10000}\delta^{\frac{11}{4}}\lambda^9 \tag{38}$$

Then either $(f, \mathcal{W}_\lambda)$ localizes, or else there are a subfunction f^* of f, and a subset $\mathcal{W} \subset \mathcal{W}_{\frac{\lambda}{2}}$ with $|\mathcal{W}| \geq (\log \frac{1}{\delta})^{-C}|\mathcal{W}_\lambda|$, so that $|f^*| \geq (\log \frac{1}{\delta})^{-C}\lambda$ on $\mathcal{W}$, and so that

$$\|\psi_\Delta f^*\|_2^2 \leq t^{1000}\delta^{\frac{5}{2}}\lambda^3 \tag{39}$$

[2]We note though that any improvement on the exponent $\frac{1}{3}$ in Lemma 2.3(i) would lead to a version of Lemma 3.1 where the exponent 3 is replaced by some $q > 3$, and that this would allow us to prove Theorem 1 for some p without using Lemma 3.2.

for each $\delta^{\frac{1}{2}}$-cube Δ.

Proof This is related to the proof of Lemma 3.1, but we use Lemma 2.7 instead of Lemma 2.8. We note that necessarily $\lambda > 1$, and as with Lemma 3.1 we begin the proof with some pigeonhole reductions. Let $k = |\mathcal{P}|$. Let $\mathcal{P}_r$ be the type r component of $\mathcal{P}$. For some r we must have $|\{x \in \mathcal{W}_{\lambda} : |f_{\mathcal{P}_r}(x)| \geq (\log \frac{1}{\delta})^{-1}\lambda\}| \gtrsim (\log \frac{1}{\delta})^{-1}|\mathcal{W}_{\lambda}|$. With this value of r, let $\mathcal{T} = \mathcal{T}(\mathcal{P}_r)$ $(= \{\tau \in \mathcal{T}(\mathcal{P}) : r \leq |X_{\mathcal{P}}(\tau)| \leq 2r\})$. We clearly have

$$|\mathcal{T}| \leq \frac{k}{r} \tag{40}$$

A minor variant on the proof of the claim in the proof of Lemma 3.1 now shows the following:

There are a value of r, a value of $\mu \lesssim \delta^{-\frac{1}{2}}$, and a subset $\mathcal{W} \subset \mathcal{W}_{\frac{\lambda}{2}} \cap \{x : |f_{\mathcal{P}_r}(x)| \geq \frac{1}{2}(\log \frac{1}{\delta})^{-1}\lambda\}$ so that the following hold:

1. $|\mathcal{W}| \geq (\log \frac{1}{\delta})^{-C}|\mathcal{W}_{\lambda}|$.
2. If $x \in \mathcal{W}$ then $\Phi_{\mathcal{T}}(x) \in [\mu, 2\mu]$.
3. $\mathcal{W}$ is $\sqrt{\delta}$-evenly covered.

Now we consider cases.

Case 1: $\lambda \geq t^{-1500}\delta^{-\frac{1}{24}}(\frac{k}{r})^{\frac{1}{6}}\mu^{\frac{1}{2}}$.
Case 2: $\lambda \leq t^{-1500}\delta^{-\frac{1}{24}}(\frac{k}{r})^{\frac{1}{6}}\mu^{\frac{1}{2}}$.

In case 1., let S be the set of all $\sqrt{\delta}$-cubes which intersect $\mathcal{W}$, and let $X = \cup_{Q \in S} Q$. Then $\Phi_{\mathcal{T}}(x)$ is approximately equal to μ for any $x \in X$ by (27) and the subsequent discussion. It follows using Lemma 2.8 (and (40)) that we have a relation between tubes from $\mathcal{T}$ and t-cubes satisfying (R) and

$$\int_X \Phi^b_{\mathcal{T}} \leq \delta^{-C\epsilon} t^{-6} \delta^{-\frac{1}{24}} (\frac{k}{r})^{\frac{1}{6}} \mu^{\frac{1}{2}} |X| + \delta^{100}|X| \lesssim t^{1000}\lambda|X|$$

The second inequality here follows from the hypothesis of case 1 (and the fact that $\lambda \geq 1$). We conclude that there is a subset S_1 of S with proportional cardinality so that $\Phi^b_{\mathcal{T}} \leq t\lambda$ on the union of the S_1-cubes. Let $\mathcal{W}_1 = \mathcal{W} \cap \cup_{\Delta \in S_1} \Delta$. Then $|\mathcal{W}_1|$ is comparable to $|\mathcal{W}|$ in view of the even covering property (see (32), (33)). On the other hand, we can define a relation between plates $\pi \in \mathcal{P}_r$ and t-cubes via $\pi \sim Q$ if $\tau(\pi) \sim Q$. Now we proceed as in the proof of Lemma 3.1: define two functions f^* and h as follows. On each given t-cube Q, define f^* and h via

$$f^* = \sum_{\substack{\pi\in\mathcal{P}_r \\ \pi\sim Q}} f_\pi$$

$$h = \sum_{\substack{\pi\in\mathcal{P}_r \\ \pi\not\sim Q}} f_\pi$$

Then $f^*+h = f_{\mathcal{P}_r}$, and $|h|$ is $\lesssim \Phi^b_{\mathcal{T}}$ and is therefore much less than $(\log\frac{1}{\delta})^{-1}\lambda$ on $\mathcal{W}_1$. So $|f^*|$ must be $\gtrsim (\log\frac{1}{\delta})^{-1}\lambda$ on $\mathcal{W}_1$. The bound (34) follows from property (R), so $(f, \mathcal{W}_\lambda)$ localizes.

In case 2. we will show that $f_{\mathcal{P}_r}$ satisfies (39). The functions $\psi_\Delta f_\pi$ (for fixed Δ and variable $\pi \in \mathcal{P}_r$) are essentially orthogonal[3]. Accordingly

$$\begin{aligned} \|\psi_\Delta f_{\mathcal{P}_r}\|_2^2 &= \|\sum_{\pi\in\mathcal{P}_r} \psi_\Delta f_\pi\|_2^2 \\ &\lesssim \sum_{\pi\in\mathcal{P}_r} \|\psi_\Delta f_\pi\|_2^2 \qquad (41) \\ &\lesssim \int \sum_{\pi\in\mathcal{P}_r} |f_\pi|^2 \phi_\Delta \\ &\lesssim \int \Phi_{\mathcal{P}_r}\phi_\Delta \qquad (42) \end{aligned}$$

On the other hand the hypothesis of case 2 says that

$$\begin{aligned} r &\leq t^{-9000}k\delta^{-\frac{1}{4}}\mu^3\lambda^{-6} \\ &\lesssim t^{-9000}k\delta^{-\frac{1}{4}}\delta^{-\frac{3}{2}}\lambda^{-6} \qquad (43) \end{aligned}$$

For the four inequalities below, we use (42) and Lemma 2.9, then a simple calculation based on the fact that $\Phi_{\mathcal{T}}$ is pointwise $\lesssim \delta^{-\frac{1}{2}}$, then (43) and finally the hypothesis (38):

$$\|\psi_\Delta f_{\mathcal{P}_r}\|_2^2 \lesssim \delta^{\frac{1}{2}} r \int \Phi_{\mathcal{T}}\phi_\Delta$$

[3]By this we mean that (41) is valid, which follows in a standard way: roughly, for any pair of incomparable π's, either the Fourier supports are disjoint or their x-supports are disjoint up to tails, since the support of $\widehat{\psi_\Delta}$ has diameter $\leq N^{\frac{1}{2}}$. See also section 4 below.

$$\begin{aligned} &\lesssim r\delta^{\frac{3}{2}} \\ &\lesssim t^{-9000}\delta^{-\frac{1}{4}}k\lambda^{-6} \\ &\lesssim t^{1000}\delta^{\frac{5}{2}}\lambda^{3} \end{aligned}$$

The lemma is proved. $\Box$

4. Properties of $\| \ \|_{p,mic}$

The purpose of this section is to derive certain properties of the norms $\| \ \|_{p,mic}$ which we need below. The arguments here are quite standard. What we do is (1) to track down the behavior of $\| \ \|_{p,mic}$ under various kinds of rescaling and (2) to relate $\|f\|_{p,mic}$ to N-functions. In this section t is arbitrary, i.e. may not be equal to δ^{ϵ_0}.

The following "interpolation inequality" is essentially obvious:

$$\|f\|_{p,mic} \lesssim \|f\|_2^{\frac{2}{p}} \|f\|_{\infty,mic}^{1-\frac{2}{p}}, \ p \geq 2 \tag{44}$$

Next we estimate $\|f\|_{p,mic}$ for N-functions; we will prove a converse to Lemma 4.1 in Lemma 4.4 below.

Lemma 4.1 If f is an N-function with plate family $\mathcal{P}$ then (for $p \geq 2$)

$$\|f\|_{p,mic} \lesssim (N^{-\frac{3}{2}}|\mathcal{P}|)^{\frac{1}{p}} \tag{45}$$

Proof By (44) it suffices to do the cases $p = 2$ and $p = \infty$. The case $p = 2$ follows from the essential orthogonality of the functions f_π and the case $p = \infty$ follows by writing down the equation

$$\Xi_\Theta * f = \sum_\pi \Xi_\Theta * f_\pi$$

and observing that for fixed Θ the π's with $\Xi_\Theta * f_\pi \neq 0$ must all be roughly parallel, and therefore have bounded overlap. $\Box$

In Lemma 4.4 below we will prove a type of converse to Lemma 4.1.

If $\hat{f}$ is supported in $\Gamma_N(C)$ then (for cubes Q of side $t \leq 1$) we let $f_Q = (\psi_Q f) \circ a_Q$.

Lemma 4.2 $\widehat{f_Q}$ is supported in $\Gamma_{tN}(C')$ where C' depends on C, and

$$\|f_Q\|_{p,mic} \lesssim t^{-(\frac{1}{2}+\frac{1}{p})} \|f\|_2^{\frac{2}{p}} \|f\|_{\infty,mic}^{1-\frac{2}{p}} \tag{46}$$

provided $p \geq 2$. In addition, if $t \geq N^{-\frac{1}{2}}$ then

$$\|f_Q\|_2 \lesssim N^{\frac{1}{4}} \|f\|_{\infty,mic} \tag{47}$$

Proof The Fourier support statement is clear using the compact Fourier support of ψ. To prove (46), first consider the case $p = 2$, where (46) is basically the uncertainty principle. Namely, $f_Q = \psi \cdot (f \circ a_Q)$, with $\widehat{f \circ a_Q}$ supported in the Ct-neighborhood of Γ_{tN}. A standard argument with Schur's test shows that $\|f_Q\|_2 \lesssim t^{\frac{1}{2}} \|f \circ a_Q\|_2 = t^{-1}\|f\|_2$.

We also have $\|f_Q\|_{\infty,mic} \lesssim t^{-\frac{1}{2}} \|f\|_{\infty,mic}$ since each sector of angular length $(tN)^{-\frac{1}{2}}$ intersects $\mathcal{O}(t^{-\frac{1}{2}})$ sectors of angular length $N^{-\frac{1}{2}}$. Using (44) we get

$$\begin{aligned} \|f_Q\|_{p,mic} &\lesssim \|f_Q\|_2^{\frac{2}{p}} \|f_Q\|_{\infty,mic}^{1-\frac{2}{p}} \\ &\lesssim t^{-(\frac{1}{2}+\frac{1}{p})} \|f\|_2^{\frac{2}{p}} \|f\|_{\infty,mic}^{1-\frac{2}{p}} \end{aligned}$$

as claimed.

To prove (47) we use that the Fourier supports of the functions $\psi_Q \cdot (\Xi_\Theta * f)$ for fixed Q and varying Θ (Θ is an arc of angular length $N^{-\frac{1}{2}}$) have bounded overlap in view of the assumption $t \geq N^{-\frac{1}{2}}$. Thus

$$\begin{aligned} \|f_Q\|_2^2 &\lesssim \sum_\Theta \|\psi \cdot (\Xi_\Theta * f) \circ a_Q)\|_2^2 \\ &\lesssim \sum_\Theta \|f\|_{\infty,mic}^2 \cdot \|\psi\|_2^2 \\ &\lesssim N^{\frac{1}{2}} \|f\|_{\infty,mic}^2 \end{aligned}$$

□

We will also need to use a rescaling argument with suitable Lorentz transformations which we record next. Suppose that Θ is an arc of length $\rho > N^{-\frac{1}{2}}$ centered at a point $e \in S^1$. Let ω be a vector orthogonal to $(e, 1)$ and $(e, -1)$ and let T_Θ be the transformation (a scalar multiple of a Lorentz transformation) mapping $(e, 1)$ to $\rho^{-2}(e, 1)$, $(e, -1)$ to $(e, -1)$, and ω to $\rho^{-1}\omega$. Let T_Θ^{-t} be the inverse (transpose).

Lemma 4.3 Let Θ be an arc of S^1 with angular length ρ, Ξ_Θ an angular frequency cutoff[4] corresponding to Θ. Suppose that $\hat{u}$ is supported on the set $\{\xi \in \Gamma_N(1) : \widehat{\Xi_\Theta}(\xi) \neq 0\}$. Let $v = u \circ \mathcal{T}_\Theta^{-t}$. Then $\hat{v}$ is supported on $\Gamma_{\rho^2 N}(C)$ for suitable C and

$$\|v\|_{\infty,mic} \lesssim \|u\|_{\infty,mic}$$

Proof This follows from the fact that sectors of Γ_N of angular length $N^{-\frac{1}{2}}$ contained in Θ correspond to sectors of $\Gamma_{\rho^2 N}$ of angular width $(\rho^2 N)^{\frac{1}{2}}$ under T_Θ; see for example [17] or [20] for further details regarding this sort of calculation. □

Lemma 4.4 Suppose that $\hat{f}$ is supported on $\Gamma_N(A)$. Then f is the sum of at most $C = C_A$ functions $\tilde{f}$ each of which has a decomposition $\tilde{f} = \sum_\lambda \lambda f_\lambda$, where the sum is over dyadic λ satisfying

$$\lambda \lesssim \|f\|_{\infty,mic} \tag{48}$$

Each f_λ is an $\tilde{N}$-function with $\tilde{N} \in [C^{-1}N, CN]$, and

$$\sum_\lambda \lambda^p \delta^{\frac{3}{2}} |\mathcal{P}(f_\lambda)| \lesssim \|f\|_{p,mic}^p \tag{49}$$

for each fixed $p \in [2, \infty)$.

Remark We have to allow several $\tilde{f}$'s, due to the fact that we have defined N-function using a fixed choice of the constant C_1. However, it is clear that this does not cause any difficulties and we will ignore this technicality when Lemma 4.4 is applied, namely in the proof of Lemma 4.6 and in the final step of the proof of Theorem 1.

Proof By the form of the statement we can assume that $\Xi_\Theta * f$ is nonzero for only one choice of Θ (thus $\|f\|_{p,mic} = \|f\|_p$), and since we are allowing several $\tilde{f}$'s it is also easy to see that we can assume the support of $\hat{f}$ is contained in (say) $N \leq |\xi| \leq \frac{11}{10}N$. We fix a plate π so that the middle half of π^* contains the intersection of the region $N \leq |\xi| \leq \frac{11}{10}N$ with the Fourier support of Ξ_Θ. This is possible if the constant in the definition of π^* was chosen large enough. Taking $\psi_j = \psi_{\pi_j}$ we obtain a family of functions $\{\psi_j\}$ with the following properties:

[4]Of the usual type as discussed in the introduction. Our notation here is a little inconsistent since Θ does not have length $N^{-\frac{1}{2}}$.

1. For each j $\psi_j = \eta_j^2$ where $\widehat{\eta_j}$ is supported in the dilation by a small factor $\alpha << 1$ of the rectangle dual to π and centered at the origin.

2. The functions $\{\psi_j\}$ form a partition of unity.

3. $|\eta_j| \leq C\phi_{\pi_j}$; here $\{\pi_j\}$ runs through a tiling of $\mathbb{R}^3$ by translates of π.

We decompose $f = \sum_j \psi_j f$. Then we define

$$f_\lambda = \lambda^{-1} \sum_{j:\|\eta_j f\|_\infty \in [\lambda, 2\lambda]} \psi_j f \tag{50}$$

It is clear that f_λ vanishes identically if λ is large compared with $\|f\|_\infty = \|f\|_{\infty,mic}$. Further, we evidently have $|\psi_j f| \lesssim \lambda\phi_{\pi_j}$, and the support of $\widehat{\psi_j f}$ is contained in the dilation of π^* by a factor $\frac{1}{2} + C\alpha$, so the definition (50) exibits f_λ as an N-function. We now make the estimate (49). By Bernstein's inequality

$$\|\eta_j f\|_\infty^p \lesssim |\pi^*| \|\eta_j f\|_p^p = \delta^{-\frac{3}{2}} \|\eta_j f\|_p^p$$

Hence

$$\sum_\lambda \lambda^p \delta^{\frac{3}{2}} |\mathcal{P}_\lambda| \lesssim \sum_j \|\eta_j f\|_p^p \lesssim \|f\|_p^p$$

as claimed. $\square$

In practice we will need a "localized" version of Lemma 4.4. We first prove a sublemma. Fix $t \in [CN^{-1}, 1]$. Let $\{\psi_Q\}$ be the partition of unity associated to the covering by t-cubes obtained from ψ by scaling as previously defined. In the next lemma we abuse notation a bit by letting Ξ_Θ be functions whose Fourier transforms agree on $\Gamma_N(A)$ with an angular partition of unity subordinate to a covering $\mathcal{J}$ involving Γ_N by arcs of angular length $(tN)^{-\frac{1}{2}}$ (rather than $N^{-\frac{1}{2}}$).

Lemma 4.5 With this notation, if $\hat{f}$ is supported in $\Gamma_N(1)$ then the Fourier support of $(\psi_Q f) \circ a_Q$ is contained in the C-neighborhood of Γ_{tN} and

$$\sum_Q \|(\psi_Q f) \circ a_Q\|_{p,mic}^p \lesssim t^{-3} \sum_{\Theta \in \mathcal{J}} \|\Xi_\Theta * f\|_p^p \tag{51}$$

Proof The Fourier support of ψ_Q is contained in a disc of size $t^{-1} \leq (\frac{N}{t})^{\frac{1}{2}}$. Accordingly, for each $\Theta \in \mathcal{J}$ the cardinality of the set of $\mathcal{J}(\Theta) = \{\Theta' \in \mathcal{J} : \exists Q : \Xi_\Theta * (\psi_Q \cdot \Xi_{\Theta'} * f) \neq 0\}$ is bounded and similarly $|\{\Theta : \Theta' \in \mathcal{J}(\Theta)\}|$ is bounded. We therefore have

$$\begin{aligned}\sum_{\Theta\in\mathcal{J}}\sum_Q \|\Xi_\Theta * (\psi_Q f)\|_p^p &\lesssim \sum_\Theta \sum_Q \sum_{\Theta'\in\mathcal{J}(\Theta)} \|\Xi_\Theta * (\psi_Q \cdot \Xi_{\Theta'} * f)\|_p^p \\ &\lesssim \sum_Q \sum_{\Theta'} \|\psi_Q \Xi_{\Theta'} * f\|_p^p \\ &\lesssim \sum_{\Theta'} \|\Xi_{\Theta'} * f\|_p^p\end{aligned}$$

which gives (51) after rescaling. □

Lemma 4.6 Suppose that $\hat{f}$ is supported in $\Gamma_N(1)$ and $\|u\|_{\infty,mic} = 1$. Fix $\lambda \geq \delta^{10}$. Then for $t \in [CN^{-1}, 1]$ we have the following[5]. Consider a decomposition of $\mathbb{R}^3$ in t-cubes Q as above. There is a value of $\lambda_* \in [C^{-1}\delta^{2\epsilon}\lambda t^{\frac{1}{2}}, C(tN)^{\frac{1}{2}}]$ and a family of tN-functions f_Q with plate families $\mathcal{P}_Q$, so that

1) A $(\log\frac{1}{\delta})^{-C}$ fraction of the set $\{|f| \geq \lambda\}$ is contained in $\cup_Q a_Q^{-1}(\{|f_Q| \geq \lambda^*\})$.

2) $\sum_Q |\mathcal{P}_Q| \leq \delta^{-C\epsilon}(\frac{\lambda_*}{\lambda})^p(\frac{N}{t})^{\frac{3}{2}} \sum_\Theta \|\Xi_\Theta * f\|_p^p$ for any given $p \geq 2$; here Θ runs over a family of arcs of angular length $(tN)^{-\frac{1}{2}}$ as in Lemma 4.5.

3) $|\mathcal{P}_Q| \leq \delta^{-C\epsilon}(\frac{\lambda_*}{\lambda})^2(\frac{N}{t})^{\frac{3}{2}}\|\psi_Q f\|_2^2$ for each Q.

Proof Decompose $f = \sum_Q \psi_Q f$ where Q runs over t-cubes. We have a bound

$$|(\psi_Q f)(x)| \lesssim \delta^{-\frac{1}{2}}(1 + t^{-1}\text{dist}(x, Q))^{-M} \tag{52}$$

for each fixed M and all x, in view of the Schwartz decay of ψ and the fact that $\|f\|_\infty \lesssim \delta^{-\frac{1}{2}}\|f\|_{\infty,mic}$. It follows in a standard way that for any given ϵ

$$\{|f| \geq \lambda\} \subset \cup_Q \{|\psi_Q f| \geq C_\epsilon^{-1}\delta^\epsilon\lambda\} \tag{53}$$

For each Q apply Lemma 4.4 at scale tN to $(\psi_Q f) \circ a_Q$, obtaining a decomposition[6]

$$(\psi_Q f) \circ a_Q = \sum_h h g_h^Q$$

[5] Lemma 4.6 will be applied with $t = N^{-\frac{1}{2}}$.

[6] See the remark after the statement of Lemma 4.4.

Here the g_h^Q are tN-functions and

$$\begin{aligned} h &\lesssim \|(\psi_Q f)\circ a_Q\|_{\infty,mic} \\ &\lesssim t^{-\frac{1}{2}}\|f\|_{\infty,mic} \\ &= t^{-\frac{1}{2}} \end{aligned} \tag{54}$$

using (48) and then (46). For each Q, since $\lambda \geq \delta^{10}$ there are $\lesssim \log\frac{1}{\delta}$ values of h such that $|g_h^Q(x)| \geq \delta^{100}\lambda$ for some x. It follows that there is a value of $h = h(Q)$ so that a $\gtrsim (\log\frac{1}{\delta})^{-C}$ proportion of the set where $|(\psi_Q f)\circ a_Q| \geq \delta^{\epsilon}\lambda$ is contained in the set where $|hg_h^Q| \geq \delta^{2\epsilon}\lambda$. We have $\|g_h^Q\|_\infty \lesssim (tN)^{\frac{1}{2}}\|g_h^Q\|_{\infty,mic} \lesssim (tN)^{\frac{1}{2}}$ and therefore

$$h(tN)^{\frac{1}{2}} \gtrsim \delta^{2\epsilon}\lambda \tag{55}$$

We can pigeonhole to get a fixed value of h so that part 1) of the lemma holds with $f_Q = g_h^Q$ and $\lambda_* = \delta^{2\epsilon}\frac{\lambda}{h}$. Inequalities (54) and (55) imply that $\delta^{2\epsilon}t^{\frac{1}{2}}\lambda \lesssim \lambda^* \lesssim (tN)^{\frac{1}{2}}$. It remains to estimate the cardinality of $\mathcal{P}_Q$, the plate family for g_h^Q.

Applying (49) on scale tN and using that $h = \delta^{2\epsilon}\frac{\lambda}{\lambda_*}$ we have

$$|\mathcal{P}_Q| \lesssim (tN)^{\frac{3}{2}}(\delta^{2\epsilon}\frac{\lambda}{\lambda_*})^{-p}\|(\psi_Q f)\circ a_Q\|_{p,mic}^p \tag{56}$$

Summing over Q and using Lemma 4.5 we get

$$\sum_Q |\mathcal{P}_Q| \lesssim t^{-3}(tN)^{\frac{3}{2}}(\delta^{2\epsilon}\frac{\lambda}{\lambda_*})^{-p}\sum_\Theta \|\Xi_\Theta * f\|_p^p$$

which is part 2. of the lemma. For 3., we just apply (56) with $p = 2$ and then use that $\|(\psi_Q f)\circ a_Q\|_2^2 = t^{-3}\|\psi_Q f\|_2^2$. □

5. Proof of Theorem 1

In this section we always let $t = \delta^{\epsilon_0}$.

We will carry out an induction argument of the following type. Fix $p > 74$ and suppose that for functions with Fourier support in $\Gamma_N(1)$ and with $\|f\|_{\infty,mic} \leq 1$ we have

$$|\{|f| > \lambda\}| \leq \delta^{-\alpha}\frac{\|f\|_2^2}{\lambda^4(\lambda\sqrt{\delta})^{p-4}} \tag{57}$$

provided δ is small enough. Then, provided δ is small enough, for functions with Fourier support in $\Gamma_N(1)$ and with $\|f\|_{\infty,mic} \leq 1$ we also have

$$|\{|f| > \lambda\}| \leq \delta^{-\beta} \frac{\|f\|_2^2}{\lambda^4(\lambda\sqrt{\delta})^{p-4}} \tag{58}$$

for any fixed $\beta > (1 - \frac{\epsilon_0}{4})\alpha$.

We will see at the end of the section that once we establish (57)⇒(58) it is easy to obtain Theorem 1.

As a preliminary to the proof of (57)⇒(58) we make the following remarks.

1. Suppose that (57) holds for functions with Fourier support in $\Gamma_N(1)$ and with $\|f\|_{\infty,mic} \leq 1$. Then, for any fixed C (57) holds also for functions with Fourier support in $\Gamma_N(C)$ and with $\|f\|_{\infty,mic} \leq 1$, provided we include a constant factor $A(C)$ on the right hand side. This is because the dilation of $\Gamma_N(C)$ by a small fixed constant factor will be covered by a bounded number of sets of the form $\Gamma_{\tilde{N}}(1)$. This remark allows us to ignore the constant factors arising in some of the previous lemmas, e.g. Lemmas 4.3 and 4.4.

2. If (57) holds, then (for the same class of f) the corresponding strong type estimates

$$\|f\|_p^p \lesssim \log\frac{1}{\delta}\delta^{-\alpha}\delta^{2-\frac{p}{2}}\|f\|_2^2 \tag{59}$$

also hold. This follows since there are only logarithmically many relevant dyadic values for λ; $|f(x)|$ is necessarily less than $\delta^{-\frac{1}{2}}$, and on the other hand (59) with f replaced by (say) $\min(|f|, \delta^{-\frac{1}{4}})$ is trivial.

We now prove a lemma.

Lemma 5.1 Fix p and α and assume we know (57). Let f be Fourier-supported in $\Gamma_N(1)$ and such that $\|f\|_{\infty,mic} = 1$. Then for any $\lambda \geq \delta^{10}$ there is a value $\lambda_* \in (\lambda\delta^{\frac{1}{4}+\epsilon}, \delta^{-\frac{1}{4}})$, and a collection of $N^{\frac{1}{2}}$-functions $\{f_\Delta\}$ so that

1. A $(\log\frac{1}{\delta})^{-C}$ fraction of $\{|f| \geq \lambda\}$ is contained in $\cup_\Delta a_\Delta^{-1}(\{|f_\Delta| \geq \lambda^*\})$.
2. The following estimates hold for any given $\epsilon > 0$:

$$|\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}(\frac{\lambda_*}{\lambda})^2\delta^{-\frac{9}{4}}\|\psi_\Delta f\|_2^2 \tag{60}$$

$$\sum_{\Delta} |\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}\delta^{-\frac{9}{4}} \cdot \delta^{-\frac{\alpha}{2}} \frac{\|f\|_2^2}{(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}} \tag{61}$$

<u>Proof</u> We apply Lemma 4.6 with $t = N^{-\frac{1}{2}}$. The decomposition there satisfies 1., and it satisfies the estimate (60) (for this, apply 3. of Lemma 4.6 with $t = N^{-\frac{1}{2}}$) and also, by 2. of Lemma 4.5, the estimate

$$\sum_{\Delta} |\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}\delta^{-\frac{9}{4}}(\frac{\lambda_*}{\lambda})^p \sum_{\Theta} \|\Xi_\Theta * f\|_p^p \tag{62}$$

where Θ runs over a family of $N^{-\frac{1}{4}}$-arcs. We now show that (62) and the inductive hypothesis (57) imply (61).

Namely, for each Θ a suitable Lorentz rescaling of $\Xi_\Theta * f$ has Fourier support in $\Gamma_{N^{\frac{1}{2}}}(C)$ and behaves as described by Lemma 4.3. So we can apply the inductive hypothesis (57) on scale $N^{\frac{1}{2}}$. We conclude by (59) that

$$\|\Xi_\Theta * f\|_p^p \leq \delta^{-C\epsilon}\delta^{-\frac{\alpha}{2}}\delta^{1-\frac{p}{4}}\|\Xi_\Theta * f\|_2^2$$

Inserting this bound into (62) we obtain

$$\begin{aligned}\sum_{\Delta} |\mathcal{P}(f_\Delta)| &\leq \delta^{-C\epsilon}\delta^{-\frac{\alpha}{2}}\delta^{-\frac{9}{4}} \frac{\sum_\Theta \|\Xi_\Theta * f\|_2^2}{(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}} \\ &\approx \delta^{-C\epsilon}\delta^{-\frac{\alpha}{2}}\delta^{-\frac{9}{4}} \frac{\|f\|_2^2}{(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}}\end{aligned}$$

as claimed. □

<u>Lemma 5.2</u> Assume (57) for δ small enough and that f is an N-function with associated family of plates $\mathcal{P}$ satisfying (35). Then (58) holds (for small δ) if $\beta > (1-\epsilon_0)\alpha$, provided we replace $\|f\|_2^2$ on the right hand side with $\delta^{\frac{3}{2}}|\mathcal{P}|$.

<u>Proof</u> Let $\mathcal{W}$ be the set where $|f| \geq \lambda$. Apply Lemma 3.1 with $t = \delta^{\epsilon_0}$. Thus $|\mathcal{W}| \leq (\log\frac{1}{\delta})^C|\cup_Q \mathcal{W}_Q|$, with $\mathcal{W}_Q = \{x \in Q : |f_Q(x)| \geq (\log\frac{1}{\delta})^{-C}\lambda\}$, where f_Q are subfunctions of f and (34) holds.

Now, for each Q, apply the inductive hypothesis (57) to $g_Q \stackrel{def}{=} (\psi_Q f_Q) \circ a_Q$ with N replaced by tN, and λ replaced by $(\log\frac{1}{\delta})^{-C}\lambda$. We have $\|g_Q\|_{\infty,mic} \lesssim t^{-\frac{1}{2}}$ by Lemma 4.2

so we obtain

$$
\begin{aligned}
t^{-3}|\mathcal{W}_Q| &\le |\{|g_Q| > (\log\tfrac{1}{\delta})^{-C}\lambda\}| \\
&\le \left(\frac{t}{\delta}\right)^{\frac{p}{2}-2}\left(\frac{t}{\delta}\right)^{\alpha}\|g_Q\|_{\infty,mic}^{p-2}\|g_Q\|_2^2\left(\lambda(\log\tfrac{1}{\delta})^{-C}\right)^{-p} \\
&\le (\frac{t}{\delta})^{\frac{p}{2}-2}(\frac{t}{\delta})^{\alpha}t^{-(\frac{p}{2}-1)}\|g_Q\|_2^2(\lambda(\log\tfrac{1}{\delta})^{-C})^{-p} \\
&= (\log\tfrac{1}{\delta})^{Cp}\delta^{-\alpha}\delta^{-(\frac{p}{2}-2)}t^{\alpha-1}\lambda^{-p}\|g_Q\|_2^2
\end{aligned}
$$

On the other hand, we have $\|g_Q\|_2^2 \lesssim t^{-2}\|f_Q\|_2^2$ by (46) with $p = 2$, and $\|f_Q\|_2^2 \lesssim \delta^{\frac{3}{2}}|\mathcal{P}_Q|$ by (45). Accordingly

$$\sum_Q \|g_Q\|_2^2 \le t^{-2}\delta^{\frac{3}{2}}(\log\tfrac{1}{\delta})^C|\mathcal{P}|$$

by (34). It follows that

$$
\begin{aligned}
|\mathcal{W}| &\le (\log\tfrac{1}{\delta})^C\sum_Q|\mathcal{W}_Q| \\
&\le (\log\tfrac{1}{\delta})^C\delta^{-\alpha}\delta^{-(\frac{p}{2}-2)}t^{\alpha+2}\lambda^{-p}\sum_Q\|g_Q\|_2^2 \\
&\le (\log\tfrac{1}{\delta})^C\delta^{-\alpha}\delta^{-(\frac{p}{2}-2)}t^{\alpha}\lambda^{-p}\cdot\delta^{\frac{3}{2}}|\mathcal{P}|
\end{aligned}
$$

The lemma follows since $t = \delta^{\epsilon_0}$. □

The next step is

Lemma 5.3 Assume (57) for δ small enough and that f is an N-function satisfying (38). Then (58) holds (for small δ) if $\beta > (1 - \frac{\epsilon_0}{2})\alpha$, provided we replace $\|f\|_2^2$ on the right hand side with $\delta^{\frac{3}{2}}|\mathcal{P}(f)|$.

The proof of this is similar to the proof of Lemma 5.2, except we use Lemma 3.2 instead of Lemma 3.1, and obtain the localization effect on one of two scales, corresponding to the two possibilities in Lemma 3.2.

Proof of Lemma 5.3 By assumption Lemma 3.2 is applicable to the set $\mathcal{W} = \{|f| \ge \lambda\}$. If $(f, \mathcal{W})$ localizes, then we obtain (58) for any $\beta > (1 - \epsilon_0)\alpha$ by the proof of Lemma 5.2, since the hypothesis (35) there was used only to guarantee the localization property.

On the other hand suppose we are in the second case of Lemma 3.2. Thus we have a subfunction f^* of f and a subset $\mathcal{W}^* \subset \mathcal{W}$ with $|\mathcal{W}^*| \geq (\log \frac{1}{\delta})^{-C}|\mathcal{W}|$ so that $|f^*| \geq (\log \frac{1}{\delta})^{-C}\lambda$ on $\mathcal{W}^*$, and such that

$$\|\psi_\Delta f^*\|_2^2 \leq t^{1000}\delta^{\frac{5}{2}}\lambda^3 \tag{63}$$

for each $\delta^{\frac{1}{2}}$-cube Δ.

Now apply Lemma 5.1 to f_*. The resulting functions f_Δ and parameter λ_* satisfy

1. $\{|f| \geq \lambda\}| \lesssim (\log \frac{1}{\delta})^C| \cup_\Delta a_\Delta^{-1}(\{|f_\Delta| \geq \lambda^*\})|$.
2. $|\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}t^{1000}\lambda_*^3$
3. $\sum_\Delta |\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}\delta^{-\frac{9}{4}} \cdot \delta^{-\frac{\alpha}{2}}\frac{\|f\|_2^2}{(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}}$.

Namely, 1. and 3. are immediate from 1. of lemma 5.1 and from (61). To obtain 2., we start from (60) and then substitute in (63); thus,

$$\begin{aligned}
|\mathcal{P}(f_\Delta)| &\leq \delta^{-C\epsilon}(\frac{\lambda_*}{\lambda})^2\delta^{-\frac{9}{4}}\|\psi_\Delta f^*\|_2^2 \\
&\leq \delta^{-C\epsilon}(\frac{\lambda_*}{\lambda})^2\delta^{-\frac{9}{4}} \cdot t^{1000}\delta^{\frac{5}{2}}\lambda^3 \\
&\leq \delta^{-C\epsilon}t^{1000}\lambda_*^3
\end{aligned}$$

since $\lambda \leq \delta^{-\frac{1}{4}-\epsilon}\lambda_*$.

The f_Δ's are $\sqrt{N}$-functions, and property 2. implies that they satisfy (35) with λ_* and $\sqrt{N}$ replacing λ and N. Accordingly by Lemma 5.2

$$|\{|f_\Delta| \geq \lambda_*\}| \leq \delta^{-C\epsilon}\delta^{-\frac{1}{2}(1-\epsilon_0)\alpha}\frac{\delta^{\frac{3}{4}}|\mathcal{P}(f_\Delta)|}{\lambda_*^4(\lambda_*\delta^{\frac{1}{4}})^{p-4}} \tag{64}$$

Scale back down and sum over Δ. This gives

$$\begin{aligned}
|\{|f| \geq \lambda\}| &\leq (\log \tfrac{1}{\delta})^C \sum_\Delta \delta^{\frac{3}{2}}|\{|f_\Delta| \geq \lambda_*\}| \\
&\leq \delta^{-C\epsilon}\delta^{\frac{3}{2}}\sum_\Delta \delta^{-\frac{1}{2}(1-\epsilon_0)\alpha}\frac{\delta^{\frac{3}{4}}|\mathcal{P}(f_\Delta)|}{\lambda_*^4(\lambda_*\delta^{\frac{1}{4}})^{p-4}}
\end{aligned}$$

$$\leq \delta^{-C\epsilon}\delta^{\frac{3}{2}}\delta^{-\frac{1}{2}(1-\epsilon_0)\alpha}\frac{\delta^{\frac{3}{4}}\delta^{-C\epsilon}\delta^{-\frac{9}{4}}\cdot\delta^{-\frac{\alpha}{2}}\|f\|_2^2}{\lambda_*^4(\lambda_*\delta^{\frac{1}{4}})^{p-4}(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}}$$

$$\leq \delta^{-C\epsilon}\delta^{-(\frac{1}{2}(1-\epsilon_0)\alpha+\frac{\alpha}{2})}\cdot\frac{\delta^{\frac{3}{2}}|\mathcal{P}|}{\lambda^4(\lambda\sqrt{\delta})^{p-4}}$$

The four inequalities followed respectively from property 1., from (64), from property 3., and from Lemma 4.1 with $p = 2$. The lemma is proved. □

Proof of Theorem 1 As discussed at the beginning of the section the main step is to show that if (57) is true for functions Fourier supported in some[7] $\Gamma_N(1)$ with $\|f\|_{\infty,mic} \leq 1$ then so is (58). Fix then a function f Fourier supported in some $\Gamma_N(1)$ and with $\|f\|_{\infty,mic} \leq 1$.

We observe to begin with that (58) follows from Tchebyshev's inequality if

$$\lambda^2(\lambda\sqrt{\delta})^{p-4} < 1$$

or equivalently if we do not have

$$\lambda \geq \delta^{-\frac{1}{2}+\frac{1}{p-2}} \tag{65}$$

We will therefore assume that (65) holds.

Apply Lemma 5.1 to f. The resulting functions f_Δ satisfy

1. $|\{|f| \geq \lambda\}| \lesssim (\log\frac{1}{\delta})^C|\cup_\Delta a_\Delta^{-1}(\{|f_\Delta| \geq \lambda^*\})|$.
2. $|\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}t^{10001}\delta^{\frac{11}{8}}\lambda_*^9$.
3. $\sum_\Delta |\mathcal{P}(f_\Delta)| \leq \delta^{-C\epsilon}\delta^{-\frac{9}{4}}\cdot\delta^{-\frac{\alpha}{2}}\frac{\|f\|_2^2}{(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}}$.

Namely, 1. and 3. follow immediately from 1. in Lemma 5.1 and from (61). We now verify 2. Namely, if Δ is a $\sqrt{\delta}$-disc then by (47) and rescaling we have

$$\|\psi_\Delta f\|_2^2 \lesssim \delta$$

Hence by (60) and then (65) we have

$$\begin{aligned}|\mathcal{P}(f_\Delta)| &\leq \delta^{-C\epsilon}(\frac{\lambda_*}{\lambda})^2\delta^{-\frac{5}{4}}\\ &\leq \delta^{-C\epsilon}\lambda_*^2\delta^{-\frac{1}{4}-\frac{2}{p-2}}\end{aligned} \tag{66}$$

[7] N is always taken sufficiently large.

We also have
$$\lambda_* \geq \delta^{C\epsilon}\delta^{-\frac{1}{4}+\frac{1}{p-2}} \tag{67}$$
because $\lambda_* \geq \delta^{C\epsilon}\delta^{\frac{1}{4}}\lambda$ and λ satisfies (65). Since $p > 74$, (67) implies that
$$\delta^{-\frac{1}{4}-\frac{2}{p-2}}\lambda_*^2 \leq \delta^{\kappa}\cdot\delta^{\frac{11}{8}}\lambda_*^9$$
for a fixed $\kappa = \kappa_p > 0$. Combining this inequality with (66) proves 2. provided ϵ_0 (and of course ϵ) have been chosen small enough.

We may therefore apply Lemma 5.3 to the f_Δ's replacing δ with $\sqrt{\delta}$ and λ with λ_*, obtaining that if $\gamma > (1-\frac{\epsilon_0}{2})\alpha$ then
$$|\{|f_\Delta| \geq \lambda^*\}| \leq \delta^{-\frac{\gamma}{2}}\frac{\delta^{\frac{3}{4}}|\mathcal{P}(f_\Delta)|}{\lambda_*^4(\lambda_*\delta^{\frac{1}{4}})^{p-4}}$$
We now repeat the last part of the proof of Lemma 5.3. Scaling back down and summing over Δ gives
$$\begin{aligned}
|\{|f| \geq \lambda\}| &\leq (\log\frac{1}{\delta})^C\sum_\Delta \delta^{\frac{3}{2}}|\{|f_\Delta| \geq \lambda_*\}| \\
&\leq \delta^{-C\epsilon}\delta^{\frac{3}{2}}\sum_\Delta \delta^{-\frac{\gamma}{2}}\frac{\delta^{\frac{3}{4}}|\mathcal{P}(f_\Delta)|}{\lambda_*^4(\lambda_*\delta^{\frac{1}{4}})^{p-4}} \\
&\leq \delta^{-C\epsilon}\delta^{\frac{3}{2}}\delta^{-\frac{\gamma}{2}}\frac{\delta^{\frac{3}{4}}\delta^{-C\epsilon}\delta^{-\frac{9}{4}}\cdot\delta^{-\frac{\alpha}{2}}\|f\|_2^2}{\lambda_*^4(\lambda_*\delta^{\frac{1}{4}})^{p-4}(\frac{\lambda}{\lambda_*})^4(\frac{\lambda}{\lambda_*}\delta^{\frac{1}{4}})^{p-4}} \\
&\leq \delta^{-C\epsilon}\delta^{-\frac{1}{2}(\gamma+\alpha)}\cdot\frac{\|f\|_2^2}{\lambda^4(\lambda\sqrt{\delta})^{p-4}}
\end{aligned}$$
Thus (58) holds for any $\beta > (1-\frac{\epsilon_0}{4})\alpha$, as was to be proved.

We conclude by iterating (57)$\Rightarrow$(58) that (57) holds for arbitrary $\alpha > 0$ provided $\|f\|_{\infty,mic} \leq 1$. It remains to pass from this estimate to Theorem 1. It is not hard to see using the partition of unity $\{\psi(x-j)\}_{j\in\mathbb{Z}^3}$ and estimates like (52) that it suffices to prove Theorem 1 locally; we omit the details of this reduction. Assume then that f is Fourier-supported in $\Gamma_N(1)$ and that $\|f\|_{p,mic} \leq 1$, and let Q_0 be the unit cube; we need to show that $\|f\|_{L^p(Q_0)} \lesssim \delta^{-(\frac{1}{2}-\frac{2}{p}+\epsilon)}$.

Using Lemma 4.4 we decompose $f = \sum_\lambda \lambda f_\lambda$ where $\|f_\lambda\|_{\infty,mic} \lesssim 1$ and (by Lemma 4.1) $\|f_\lambda\|_2^2 \lesssim \lambda^{-p}$; the sum is over dyadic values of λ bounded above by a negative power of δ. By (57), for any fixed $\epsilon > 0$ we have

$$\|f_\lambda\|_p \lesssim \delta^{-\epsilon}\delta^{-(\frac{1}{2}-\frac{2}{p})}\lambda^{-1}$$

for each λ. In view of the bound on λ, we can sum over (dyadic) $\lambda \geq \delta^{100}$, obtaining

$$\|\sum_{\lambda > \delta^{100}} \lambda f_\lambda\|_p \lesssim \delta^{-\epsilon}\log(\frac{1}{\delta})\delta^{-(\frac{1}{2}-\frac{2}{p})}$$

On the other hand, the function

$$\sum_{\lambda \leq \delta^{100}} \lambda f_\lambda$$

is clearly bounded pointwise by δ^{50}, and therefore has a similar bound in L^p norm on the unit cube. The proof is complete. □

Remark In the preceding argument we have changed scales from δ to $\sqrt{\delta}$ twice; once in the final argument, and once in the proof of Lemma 5.3. This was necessary in order to obtain Theorem 1 for some p. The reader will be able to see this by going through the logic setting $\lambda = \delta^{-\frac{1}{2}}$, since in this case it is only the condition (38) that can be obtained by a single rescaling, and not the condition (35).

6. Proofs of the corollaries

In this section Γ is redefined to be the full (rather than forward) light cone and similarly with Γ_N etc. It is easy to see that Theorem 1 remains true, since the statement is invariant under complex conjugation which interchanges the forward and backward cones on the Fourier side. Let $\chi : \mathbb{R} \to \mathbb{R}$ be a fixed Schwartz function which is nonzero on the interval $[1, 2]$ and whose Fourier transform has compact support. Let η be a radial Schwartz function in $\mathbb{R}^2$ whose Fourier transform is supported in $\frac{1}{2} \leq |\xi| \leq 2$ and such that

$$\sum_{n \geq 0} \hat{\eta}(\frac{\xi}{2^n})$$

is equal to 1 outside a compact set. Define $\eta_n(x) = 2^{2n}\eta(2^n x)$. We start with the following observation.

Lemma 6.1 Assume that u is a solution of the wave equation,

$$\Box u = 0,\ u(\cdot,0) = f,\ \frac{\partial u}{\partial t}(\cdot,0) = g \tag{68}$$

Assume that the supports of $\hat{f}$ and $\hat{g}$ are contained in the annulus $N \leq |\xi| \leq 2N$. Let $V(x,t) = \eta_n * (\chi(t)u(x,t))$, where the convolution is in the space variables. Then the support of $\hat{G}$ is contained in $\Gamma_N(C)$, $N = 2^n$, and (for $p \geq 2$)

$$\|V\|_{p,mic} \lesssim \|f\|_p + N^{-1}\|g\|_p \tag{69}$$

Proof The Fourier support statement follows since the Fourier support of u is contained in Γ_N and $\hat{\chi}$ has compact support. The bound (69) also differs only notationally from known bounds. We give the argument assuming $g = 0$. It suffices to prove (69) when $p = \infty$, since it is obvious when $p = 2$. Let A be a smooth cutoff to a sector of angular width $N^{-\frac{1}{2}}$ and Ξ the inverse Fourier transform of A. Then

$$\Xi * V(x,t) = \int \chi(t)A(\xi)\hat{\eta}(\frac{\xi}{2^n}) \int e^{2\pi i(x-y)\cdot\xi}\cos(2\pi t|\xi|)f(y)dyd\xi$$

so it suffices to show that for fixed t and y the $L^1(dx)$ norm of the function

$$B(x) = \int A(\xi)\hat{\eta}(\frac{\xi}{2^n})e^{2\pi i((x-y)\cdot\xi + t|\xi|)}d\xi$$

is bounded by a constant. However, the latter statement is well known (cf. [11]) and follows from stationary phase. $\Box$

Proof of Corollary 2 For the local smoothing statement, we apply Theorem 1 to the function V in Lemma 6.1 obtaining

$$\|\eta_n * (\chi(t)u(x,t))\|_p \leq C_\epsilon 2^{n(\frac{1}{2}-\frac{2}{p}+\epsilon)}(\|f\|_p + 2^{-n}\|g\|_p)$$

hence by the Fourier support of V, if $\alpha > \frac{1}{2} - \frac{2}{p}$ then for some ϵ we also have

$$\|\eta_n * (\chi(t)u(x,t))\|_p \lesssim 2^{-n\epsilon}(\|f\|_{p,\alpha} + \|g\|_{p,\alpha-1})$$

We then sum over n to get the result.

For the cone multiplier statement, it suffices to show that if $\widehat{\rho_\delta}$ is a smooth cutoff (with the natural bounds) to the δ-neighborhood of Γ_1 then $\|\rho_\delta * f\|_p \lesssim \delta^{-\alpha}\|f\|_p$ assuming the indicated conditions on p and α. Rescaling, it suffices to show that if $\widehat{R_N}$ is a smooth cutoff to the 1-neighborhood of Γ_N then

$$\|R_N * f\|_p \lesssim N^\alpha \|f\|_p \tag{70}$$

However, we have $\|R_N * f\|_{p,mic} \lesssim \|f\|_p$, since $\Xi_\Theta * f$ for fixed Θ is obtained from f by convolution with a function with bounded L^1 norm, and one can interpolate with L^2. So (70) follows from Theorem 1. □

In principle Corollary 3 follows from Theorem 1 using the relation between circular means and cone Fourier transforms or solutions of the wave equation; see [11] and [13]. However, there is no reference for the exact statement that we need so we give the argument. Let σ_t be normalized length measure on the circle of radius t. We will use the asymptotics

$$\widehat{\sigma}_t(\xi) = 2\sqrt{2\pi}\left(\frac{t}{|\xi|}\right)^{\frac{1}{2}} \cos(2\pi t|\xi| - \frac{\pi}{4}) + \mathcal{O}(|\xi|^{-\frac{3}{2}}) \tag{71}$$

We always assume $1 \leq t \leq 2$; thus this asymptotics is uniform in t as $|\xi| \to \infty$. We define $\sigma_t^n : \mathbb{R}^2 \to \mathbb{R}$ by

$$\sigma_t^n = (2\sqrt{2\pi})^{-1} t^{-\frac{1}{2}} \eta_n * \sigma_t$$

<u>Lemma 6.2</u> Let $n \in \mathbb{Z}^+$ and let $N = 2^n$. Let f be a function in $\mathbb{R}^2$ with $\|f\|_\infty = 1$, and define $u : \mathbb{R}^3 \to \mathbb{R}$ by

$$u(x,t) = \chi(t)\sigma_t^n * f(x)$$

Then, when $\frac{1}{2} \leq t \leq 2$, we have $u = v + w$ where

1. $\hat{v}$ is supported in $\Gamma_N(C)$ and v satisfies the estimates

$$\|v\|_{\infty,mic} \lesssim N^{-\frac{1}{2}} \tag{72}$$

$$\|v\|_2^2 \lesssim N^{-1}\|f\|_2^2 \tag{73}$$

2. $\|w\|_\infty \lesssim N^{-\frac{1}{2}}\|f\|_2$

Proof Using the asymptotics (71) we express

$$\widehat{\sigma_t^n}(\xi) = \hat{\eta}(\frac{\xi}{N})|\xi|^{-\frac{1}{2}} \cos(2\pi t|\xi| - \frac{\pi}{4}) + g_t(\xi)$$

where

$$\|\widehat{g_t}\|_{L^2(d\xi)} \lesssim N^{-\frac{1}{2}} \tag{74}$$

for each fixed t. Define $w(x,t) = \chi(t) g_t * f(x)$ and $v = u - w$. Part 2. of the lemma follows from (74) by Plancherel and Cauchy-Schwartz.

To prove 1. we express v as the real part of the function

$$V(x,t) = e^{-i\frac{\pi}{4}} \chi(t) \int \hat{\eta}(\frac{\xi}{N})|\xi|^{-\frac{1}{2}} e^{2\pi i(x\cdot\xi + t|\xi|)} \hat{f}(\xi) d\xi$$

therefore

$$\hat{V}(\xi,\tau) = e^{-i\frac{\pi}{4}} \hat{\chi}(\tau - |\xi|) \hat{\eta}(\frac{\xi}{N})|\xi|^{-\frac{1}{2}} \hat{f}(\xi)$$

which is clearly supported in $\Gamma_N(C)$, and is also clearly bounded by $N^{-\frac{1}{2}}|\hat{f}(\xi)|$, which proves the estimate (73). It remains to prove (72). Let A be a smooth cutoff to a sector of angular width $N^{-\frac{1}{2}}$ and Ξ the inverse Fourier transform of A. Then

$$\Xi * V(x,t) = \int e^{-i\frac{\pi}{4}} \chi(t) A(\xi) \int \hat{\eta}(\frac{\xi}{N})|\xi|^{-\frac{1}{2}} e^{2\pi i((x-y)\cdot\xi + t|\xi|)} f(y) dy d\xi$$

so it suffices to show that for fixed t and y the $L^1(dx)$ norm of the function

$$B(x) = \int A(\xi) \hat{\eta}(\frac{\xi}{N})|\xi|^{-\frac{1}{2}} e^{2\pi i((x-y)\cdot\xi + t|\xi|)} d\xi$$

is $\lesssim N^{-\frac{1}{2}}$. However, the latter estimate is the same estimate from [11] that was used in the proof of Lemma 6.1. The lemma follows. □

Proof of Corollary 3 We can assume that the set F of Corollary 2 is contained in the region $1 \le t \le 2$ and then also that the set E is contained in a large fixed disc.

We will use the notation in Lemma 6.1. Let f be a (measurable) function in $\mathbb{R}^2$ with $\|f\|_\infty \le 1$. Assume $\{x : f(x) \neq 0\} \subset E$. We will show first that the three dimensional Lebesgue measure of

$$Y_\lambda \overset{def}{=} \{(x,t) \in \mathbb{R}^2 \times [1,2] : |\sigma_t^n * f(x)| \ge \lambda\}$$

is $\leq C_\epsilon \lambda^{-75} N^{-2+\epsilon}|E|$ provided that λ is large compared with $N^{-\frac{1}{2}}$.

Namely, there is a function $F : \mathbb{R}^3 \to \mathbb{R}$ with the following properties:

1. $\mathrm{supp}\hat{F} \subset \Gamma_N(C)$
2. $\|F\|_{\infty,mic} \lesssim N^{-\frac{1}{2}}$ and $\|F\|_2 \lesssim N^{-\frac{1}{2}}|E|^{\frac{1}{2}}$.
3. If $(x,t) \in Y_\lambda$ then $|F(x,t)| \geq \frac{\lambda}{2}$.

This is clear by setting F equal to the function v in Lemma 6.2; we just note that 3. follows from the estimate in 2. of the lemma since $\lambda >> N^{-\frac{1}{2}}$ and E is contained in a large fixed disc. It follows by (44) that

$$\|F\|_{p,mic} \lesssim N^{-\frac{1}{2}}|E|^{\frac{1}{p}}$$

hence by Theorem 1, if $p = 75$ and $\epsilon > 0$ then

$$\|F\|_p \lesssim N^{-\frac{2}{p}+\epsilon}|E|^{\frac{1}{p}}$$

Using Tchebyshev's inequality and 3. we get

$$|Y_\lambda| \lesssim \lambda^{-p} N^{-2+\epsilon}|E|$$

as claimed. Taking now $f = \chi_E$ we obtain the following:

For any $\epsilon > 0$ there is $\eta > 0$ such that

$$\left|\{(x,t) \in \mathbb{R}^2 \times [\frac{1}{2}, 2] : |\eta_n * \sigma_t * \chi_E(x)| \geq \frac{1}{2}N^{-\eta}\}\right| \leq N^{-2+\epsilon}|E| \tag{75}$$

We can convert this to an entropy estimate on scale $\delta = N^{-1}$:

For any $\epsilon > 0$ there is $\eta > 0$ such that

$$\mathcal{E}_\delta\left(\{(x,t) \in \mathbb{R}^2 \times [\frac{1}{2}, 2] : |\eta_n * \sigma_t * \chi_E(x)| \geq N^{-\eta}\}\right) \leq N^{1+\epsilon}|E| \tag{76}$$

Namely, since $\|\eta_n * \sigma_t * \chi_E\|_\infty \lesssim 1$, (36) shows that if $|\eta_n * \sigma_t * \chi_E(x)| \geq N^{-\eta}$ then $|\eta_n * \sigma_t * \chi_E(y)| \geq \frac{1}{2}N^{-\eta}$ for all y belonging to the $\frac{1}{2}N^{-(1+\eta)}$-disc centered at x. This shows that (75) implies (76).

Now let α be a small positive number, let U be an open set with measure $< \alpha$, $K \subset \mathbb{R}^2 \times [1,2]$ a set such that each circle $C(x,t)$, $(x,t) \in K$, intersects U in measure $\geq C > 0$. If $(x,t) \in K$ then $\sum_n |\eta_n * \sigma_t * \chi_U(x)| \gtrsim 1$, so for some n we have

$$|\eta_n * \sigma_t * \chi_U(x)| \gtrsim n^{-2} \tag{77}$$

Notice that (77) cannot hold for any x if n is too small, i.e. if $n^{-2}2^{-n} >> \alpha$. For the remaining values of n, we have shown above that the set where (77) holds has 2^{-n}-entropy $\lesssim 2^{(1+\epsilon)n}|E| < 2^{(1+\epsilon)n}$. We now have a covering of K by a family of discs of dyadic radius 2^{-n} satisfying $2^{-n}n^{-2} \lesssim \alpha$, and involving less than $2^{(1+\epsilon)n}$ discs of any given radius 2^{-n}.

If E has Lebesgue measure zero then we can apply the preceding to an arbitrarily small open neighborhood U of E. We conclude that any set in $\mathbb{R}^2 \times \mathbb{R}$ consisting of circles which intersect E in outer measure $\geq C$ must have Hausdorff $1+\epsilon$-dimensional measure zero. Letting $C \to 0$ we obtain the result. $\square$

Theorem 1 and its proof also imply various "$L^p \to L^q$ inequalities for the wave equation relative to fractal measures" which generalize the Strichartz inequality as well as the L^3 estimate proved in [18]. To motivate this recall for example the well-known Strichartz inequality into $L^2_t(L^\infty_x)$: for solutions of (68), locally in time we have

$$\|u\|_{L^q_t(L^\infty_x)} \lesssim \|f\|_{W^{2,s}} + \|g\|_{W^{2,s-1}} \tag{78}$$

provided $\frac{3}{4} < s \leq 1$, $q < \frac{1}{1-s}$. This can be considered as an L^q estimate

$$\|u\|_{L^q(d\mu)} \lesssim \|f\|_{W^{2,s}} + \|g\|_{W^{2,s-1}} \tag{79}$$

for u with respect to a measure μ whose t-projections are Lebesgue and whose marginals are Dirac measures, which is a one dimensional measure. A similar interpretation is valid for the result of [18]. One can ask for analogues of all these inequalities in general dimension α, i.e. estimates from L^p Sobolev spaces to $L^q(\mu)$ for the solution of $\Box$ with the right dependence on the dimensionality of μ. For general α and p this appears hard; when $\alpha = 3$ and $p \neq 2$ it is essentially the local smoothing conjecture. The following result is satisfactory only when $\alpha \leq 1$; when α is large (i.e. in case (i)) it is a minor variant on Theorem 1 and the value of p is far from sharp. We refer also to [9] for a discussion concerning estimates of this type.

Proposition 6.1 Suppose that μ be a measure in $\mathbb{R}^3$ supported in the unit disc and $\mu(D(x,r)) \leq r^\alpha$ for all discs $D(x,r)$ of radius $r \leq 1$. Then for functions with Fourier support in $\Gamma_N(1)$

$$\|u^*\|_{L^p(\mu)} \lesssim N^{\frac{1}{2}+\sigma}\|u\|_{p,mic} \tag{80}$$

provided that $\sigma > \frac{1-\alpha}{p}$ and one of the following is true: (i) $p > 77 - \alpha$ or (ii) $p > 3$ and $\alpha < 1 + \epsilon_{p,\sigma}$, where $\epsilon_{p,\sigma} > 0$ is a small (computable) constant.

We will use induction on N to show that for some fixed positive κ

$$\mu(\{x : u(x) \geq \lambda\} \leq \delta^{\kappa} \frac{\delta^{-p\sigma} \|u\|_{p,mic}^p}{(\lambda\sqrt{\delta})^p} \tag{81}$$

for functions with Fourier support in $\Gamma_N(1)$. Since there are only logarithmically many relevant dyadic values for λ this weak type estimate implies (80).

Assume that (81) has been proved on scales much less than N and fix u. Suppose at first that it u an N-function with plate family $\mathcal{P}$. In this case, we will show that

$$\mu(\{x : |u(x)| \geq \lambda\} \leq \delta^{\zeta} \frac{\delta^{-p\sigma} |\mathcal{P}| \delta^{\frac{3}{2}}}{(\lambda\sqrt{\delta})^p} \tag{82}$$

Here ζ is a fixed positive number depending on α and p.

We first look at what follows formally from Theorem 1. Namely, if $q > 74$ then Theorem 1 and Lemma 4.1 imply for any given $\tau > 0$ that

$$|\{x : |u(x)| \geq \lambda\}| \lesssim \delta^{-\tau} \delta^2 (\lambda\sqrt{\delta})^{-q} |\mathcal{P}| \delta^{\frac{3}{2}}$$

The bound (36) implies that $\{|u| \geq \lambda\}$ behaves essentially like a collection of $\delta^{\frac{3}{2}}\lambda$-discs. Using the dimension assumption on μ it therefore follows that

$$\begin{aligned} \mu(\{x : u(x) \geq \lambda\}) &\lesssim \delta^{\tau} (\delta^{\frac{3}{2}}\lambda)^{\alpha-3} \delta^2 (\lambda\sqrt{\delta})^{-q} |\mathcal{P}| \delta^{\frac{3}{2}} \\ &= \delta^{\tau} \delta^{\alpha-1} (\lambda\sqrt{\delta})^{\alpha-q-3} |\mathcal{P}| \delta^{\frac{3}{2}} \end{aligned} \tag{83}$$

In case (i), this is clearly smaller that the right hand side of (82), so (82) is proved (without an induction argument) in that case.

In case (ii), we need to combine (83) with an induction argument based on the localization property in Lemma 3.1. There is a technical issue that arises, namely that we must redefine "localization" so that the measure in 2. of the definition is μ rather than Lebesgue measure. However, it is easy to see that Lemma 3.1 remains valid with this change. We now distinguish two cases.

1. $(u, \mathcal{W})$ doesn't localize (to scale $t = \delta^{\epsilon_0}$).
2. $(u, \mathcal{W})$ localizes.

In case 1., by Lemma 3.1 we have $|\mathcal{P}| \geq t^C \lambda^3$. Fix a sufficiently small number $\eta > 0$. If $\lambda < \delta^{\eta - \frac{1}{2}}$ then provided the numbers $\epsilon_{p,\sigma}$ and ζ have been chosen small enough the estimate (82) follows since the right hand side of (82) is greater than 1. On the other hand, if $\lambda > \delta^{\eta - \frac{1}{2}}$ then the bound (83) implies (82).

It remains to consider case 2. In this case we let u_Q and $\mathcal{P}_Q$ be as in the definition[8] of localization. We consider the rescalings

$$(\psi_Q u_Q) \circ a_Q$$

Fix Q. Using Lemma 4.2 and then Lemma 4.1, we see that $(\psi_Q u_Q) \circ a_Q$ is Fourier supported in $\Gamma_{tN}(C)$ and

$$\begin{aligned} \|(\psi_Q u_Q) \circ a_Q\|_{p,mic}^p &\lesssim t^{-(\frac{p}{2}+1)} \|u_Q\|_2^2 \|u_Q\|_{\infty,mic}^{p-2} \\ &\lesssim t^{-(\frac{p}{2}+1)} \delta^{\frac{3}{2}} |\mathcal{P}_Q| \end{aligned} \tag{84}$$

Define a new measure ν by rescaling the restriction of μ to Q and then multiplying by $t^{-\alpha}$, i.e.

$$d\nu(x) = t^{-\alpha} \chi_0(x) d\mu(a_Q x)$$

where χ_0 is the indicator function of the unit cube. Then ν is supported in the unit cube and satisfies $\nu(D(x,r)) \lesssim r^\alpha$. We apply the inductive hypothesis on scale tN, and then substitute in the bound (84). This gives

$$\begin{aligned} \nu(\{x : |(\psi_Q u_Q) \circ a_Q(x)| \geq \lambda\}) &\lesssim \frac{(\frac{\delta}{t})^{\kappa - p\sigma} \|(\psi_Q u_Q) \circ a_Q(x)\|_{p,mic}^p}{(\lambda \sqrt{\frac{\delta}{t}})^p} \\ &\lesssim \frac{(\frac{\delta}{t})^{\kappa - p\sigma} t^{-(\frac{p}{2}+1)} \delta^{\frac{3}{2}} |\mathcal{P}_Q|}{(\lambda \sqrt{\frac{\delta}{t}})^p} \end{aligned}$$

Using that ψ_Q is bounded away from zero on Q and undoing the rescaling we arrive at

$$\mu(\{x \in Q : |u_Q(x)| \geq \lambda\}) \lesssim t^\alpha \frac{(\frac{\delta}{t})^{\kappa - p\sigma} t^{-(\frac{p}{2}+1)} \delta^{\frac{3}{2}} |\mathcal{P}_Q|}{(\lambda \sqrt{\frac{\delta}{t}})^p}$$

[8] To avoid any possible confusion we note that u here plays the role of f in that definition.

Summing over Q gives

$$\sum_Q \mu(\{x \in Q : |u_Q(x)| \geq \lambda\}) \leq (\log \tfrac{1}{\delta})^C t^\alpha \frac{(\frac{\delta}{t})^{\kappa - p\sigma} t^{-(\frac{p}{2}+1)} \delta^{\frac{3}{2}} |\mathcal{P}_Q|}{(\lambda \sqrt{\frac{\delta}{t}})^p}$$

By the (current) definition of the localization property we then have

$$\begin{aligned} \mu(\{x : u(x) \geq \lambda\}) &\lesssim (\log \tfrac{1}{\delta})^C \sum_Q \mu(\{x \in Q : |u_Q(x)| \geq \lambda\}) \\ &\leq (\log \tfrac{1}{\delta})^C t^\alpha \frac{(\frac{\delta}{t})^{\kappa - p\sigma} t^{-(\frac{p}{2}+1)} \delta^{\frac{3}{2}} |\mathcal{P}_Q|}{(\lambda \sqrt{\frac{\delta}{t}})^p} \end{aligned}$$

The exponent of t here is $\alpha + p\sigma - 1$, which is positive. Since $t = \delta^{\epsilon_0}$, it follows that (82) holds provided ζ has been chosen less than $(\alpha + p\sigma - 1)\epsilon_0$.

Now we can simply repeat the last steps in the proof of Theorem 1, decomposing a general function as a sum of N-functions and applying (82) to each. The conclusion (81) follows provided κ has been chosen less than ζ. □

Proposition 6.1 can be used to obtain various more explicit estimates. Let u be a solution of the wave equation (68). Using Lemma 6.1 and Proposition 6.1 one easily gets

$$\|u\|_{L^p(\mu)} \lesssim \|f\|_{p,\frac{1}{2}+\sigma} + \|g\|_{p,-\frac{1}{2}+\sigma} \tag{85}$$

with the same conditions on σ, p and μ. In particular this gives

Corollary 6.1 Assume $p > 3$. Consider a solution of (68). Then for measures with compact support and satisfying $\mu(D(x,r)) \leq r^\alpha$ we have

$$\|u\|_{L^p(\mu)} \lesssim \|f\|_{p,\frac{1}{2}} + \|g\|_{p,-\frac{1}{2}} \tag{86}$$

provided $\alpha > 1$, and

$$\|u\|_{L^p(\mu)} \lesssim \|f\|_{p,\frac{1}{2}+\sigma} + \|g\|_{p,-\frac{1}{2}+\sigma}$$

provided $\alpha \leq 1$ and $\sigma > \frac{1-\alpha}{p}$.

The proof is as indicated above. In particular, we have

$$\|u\|_{L^p_t(L^\infty_x)} \lesssim \|f\|_{p,\frac{1}{2}+\epsilon} + \|g\|_{p,-\frac{1}{2}+\epsilon}$$

for any $\epsilon > 0$ by the case $\alpha = 1$ and finite propagation speed, so Corollary 6.1 includes the result of [18]. Note also that (86) gives an L^p version of Corollary 3.

Remark It is also possible to derive Strichartz type estimates from Proposition 6.1, e.g. the following statement: suppose $s \in (\frac{3}{4}, 1]$ and $\alpha > 4(1-s)$. Then, for any $q < \frac{\alpha}{1-s}$,

$$\|u\|_{L^q(\mu)} \lesssim \|f\|_{W^{2,s}} + \|g\|_{W^{2,s-1}} \tag{87}$$

provided μ is a measure with fixed compact support and with $\mu(D(x,r)) \leq r^\alpha$. This follows by decomposing u in N-functions as in Lemma 4.4 and applying Proposition 6.1 in an appropriate way. We omit the argument since it appears to us that (87) would also follow from (78) using a suitable version of the Marstrand projection theorem. However, it is possible that issues of this type could be of some interest when $s < \frac{3}{4}$.

References

[1] J. Bourgain, *Averages in the plane over convex curves and maximal operators*, J. Analyse Math. 47(1986), 69-85.

[2] J. Bourgain, *Estimates for cone multipliers*, Operator Theory: Advances and Applications, 77 (1995), 41-60.

[3] K. L. Clarkson, H. Edelsbrunner, L. J. Guibas, M. Sharir, E. Welzl, *Combinatorial complexity bounds for arrangements of curves and spheres*, Discrete Comput. Geom. 5(1990), 99-160.

[4] H. Farag, T. Wolff, in preparation.

[5] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. 124(1970), 9-36.

[6] D. Foschi, S. Klainerman, *Bilinear space-time estimates for homogeneous wave equations*, preprint 1998.

[7] L. Hormander, *The Analysis of Linear Partial Differential Operators*, volume 1, 2nd edition, Springer-Verlag, 1990.

[8] J. M. Marstrand, *Packing circles in the plane*, Proc. London Math. Soc. 55 (1987), 37-58.

[9] T. Mitsis, to appear in Proc. London Math. Soc.

[10] G. Mockenhaupt, *A note on the cone multiplier*, Proc. Amer. Math. Soc. 117(1993), 145-152.

[11] G. Mockenhoupt, A. Seeger, C. Sogge, *Wave front sets and Bourgain's circular maximal theorem*, Ann. Math. 134 (1992), 207-218.

[12] W. Schlag, *A generalization of Bourgain's circular maximal theorem*, J. Amer. Math. Soc. 10(1997), 103-122.

[13] W. Schlag, Caltech thesis, 1996.

[14] M. Sharir, P. Agarwal, *Davenport-Schinzel Sequences and their Geometric Applications*, Cambridge University Press, 1995.

[15] E. M. Stein, Harmonic Analysis, Princeton University press, 1993.

[16] T. Tao, *Endpoint bilinear restriction estimates for the cone, and some sharp null form estimates*, preprint 1999.

[17] T. Tao, A. Vargas, *A bilinear approach to cone multipliers*, I and II, preprints 1998, to appear in Geometric and Functional Analysis.

[18] T. Wolff, *A Kakeya type problem for circles*, Amer. J. Math. 119(1997), 985-1026.

[19] T. Wolff, *Recent work connected with the Kakeya problem*, in Prospects in Mathematics (Princeton, N. J. 1996), ed. H. Rossi, American Mathematical Society, 1998.

[20] T. Wolff, *A sharp bilinear cone restriction estimate*, preprint 1999, to appear in Ann. Math.