## Mixed norm estimates for a restricted X-ray transform in $\mathbb{R}^4$ and $\mathbb{R}^5$

M. Burak Erdoğan
Department of Mathematics
253-37 Caltech
Pasadena, CA 91125 USA
mehmet@cco.caltech.edu

*Dedicated to the memory of my advisor Thomas H. Wolff.*

### 1. Introduction.

Let $\mathcal{G}_{k,d}$ be the space of all $k$-planes in $\mathbb{R}^d$. The Radon transform or the $k$-plane transform $\mathcal{R}_{k,d}$ is defined as an operator from the functions defined on $\mathbb{R}^d$ to the functions defined on $\mathcal{G}_{k,d}$ via

$$\mathcal{R}_{k,d}f(p) = \int_p f, \quad p \in \mathcal{G}_{k,d}.$$

The Radon transform found important applications in integral geometry and in the study of PDE's.

$\mathcal{R}_{1,d}$ is often called the $X$-ray transform due to its applications in radiology, we denote it by $X_{full}$. It is well-known [19], [12] that the sharp mixed norm estimates for the full $X$-ray transform implies the Kakeya conjecture and it is related to some of the main problems in the summability of Fourier transform, Fourier restriction and more generally to oscillatory integrals, non-linear P.D.E's and number theory [7], [1], [2], [20], [3], [14]. For some mapping properties of $X_{full}$, see, e.g., [6], [5], [19] and [12].

Note that $\mathcal{G}_{1,d}$ is a $2d-2$-dimensional manifold, thus $X_{full}$ is overdetermined for $d \geq 3$, and it is of interest to consider its restrictions to lower dimensional subspaces of $\mathcal{G}_{1,d}$. For the definition of the restricted $X$-ray transforms as part of a more general class of transformations and some of its properties, see [11].

One particular example is the restriction of $X_{full}$ to the space of light rays (lines in $\mathbb{R}^d$ making a 45 degree angle with the plane $x_d = 0$). Recently, Wolff [21] obtained mixed norm estimates for this operator (almost sharp in $\mathbb{R}^3$) and used this information to prove almost sharp bilinear cone restriction estimates in all dimensions.

We are interested in the restriction of $X_{full}$ to $d$ dimensional line complexes in $\mathbb{R}^d$. Let $d \geq 3$; the subspace $G_d$ of $\mathcal{G}_{1,d}$ we are interested in is defined as follows: Let $\gamma_d$ be the curve $\{\gamma_d(t) : \gamma_d(t) = (1, t, t^2, ..., t^{d-1}), t \in (-1, 1)\}$ in $\mathbb{R}^d$. Let $l(t, x)$ denote the line $\{x + s\gamma_d(t) : s \in \mathbb{R}\}$, where $x \in H_t := \{x : x \perp \gamma_d(t)\}$. We identify $G_d$ with $[-1, 1] \times \mathbb{R}^{d-1}$ via $G_d = \{l(t, x) : t \in [-1, 1], x \in H_t\}$. This line complex is a model case for a general class called *rigid well-curved line complexes* (see, e.g., [9], [10] and [11]). It is called well-curved since $\gamma_d'(t), ..., \gamma_d^{(d-1)}(t)$ are linearly independent for any $t \in [-1, 1]$, and the term rigid is used to describe the fact that for any point $\gamma_d(t)$ in the "direction set" $\gamma_d$, $G_d$ contains all the lines in $\mathbb{R}^d$ having the direction $\gamma_d(t)$. We call the lines in $G^d$ the $\gamma_d$-rays.

Now, we define the restricted X-ray transform as an operator from the functions defined on $\mathbb{R}^d$ to the functions defined on $G_d$ in the following way

$$Xf(l(t, x)) = \int_{l(t,x)} f, \quad t \in [-1, 1], x \in H_t.$$

We work with the following mixed norms for the functions defined on $G_d$:

$$\|f\|_{L^q(L^r)} = \|f\|q, r = (\int_{-1}^{1} (\int_{H_t} |f(l(t, x))|^r dx)^{q/r} dt)^{1/q}.$$

We are interested in the estimates of the following type: If $f : \mathbb{R}^d \to \mathbb{R}$ is supported in the unit cube $Q_1$ then

$$\|Xf\|_{q,r} \leq C_{pqr}\|f\|_p. \tag{1}$$

**Proposition 1.1.** *The following conditions for $p, q$ and $r$ are necessary for* (1) *to hold*

$$\frac{d}{p} \leq \frac{d-1}{r} + 1, \tag{2}$$
$$\frac{(d-1)d}{p} \leq \frac{2}{q} + \frac{(d-1)d}{r}, \tag{3}$$
$$\frac{(d-2)(d+1)}{p} \leq \frac{(d-1)d}{r}. \tag{4}$$

The following counter-examples prove Proposition 1.1, they are quite standard (see, e.g., [5], [9], [10] and [11]). The restriction (2) can be obtained by applying $X$ to the characteristic function of a $\delta$-ball. To obtain (3), let $f$ be the characteristic function of the set $|x_1| \leq 1, |x_2| \leq \delta, ..., |x_d| \leq \delta^{d-1}$. Note that $\|f\|_p \approx \delta^{\frac{d(d-1)}{2p}}$ and for all $|t| < \delta$, we have $Xf \approx 1$ on a subset of $H_t$ of measure $\gtrsim \delta^{\frac{d(d-1)}{2}}$. Hence $\|Xf\|_{q,r} \gtrsim \delta^{\frac{1}{q}}\delta^{\frac{d(d-1)}{2r}}$, which proves the necessity of (3). Finally, divide $\gamma_d$ into $M$ ($\approx 1/\delta$) segments $s_1, ..., s_M$ of length $\delta$ centered at $t_1, ..., t_M$, respectively. For any segment $s_i$, consider the parallelogram $P_i \subset \mathbb{R}^{d-1}$ with dimensions $\delta \times \delta^2 \times ... \times \delta^{d-1}$, whose longest axis is tangent to $\gamma_d$ at $\gamma_d(t_i)$ and other axes are in the directions $\gamma_d''(t_i), ..., \gamma_d^{(d-1)}(t_i)$, respectively. Let $f$ be the characteristic function of the set $\{(x_1, x_2, ..., x_d) \in \mathbb{R}^d : x_1 \in (1,2), (\frac{x_2}{x_1}, ..., \frac{x_d}{x_1}) \in \cup_i^M P_i\}$. Note that $\|f\|_p \approx \delta^{\frac{d^2-d-2}{2p}}$ and for all $t$, $Xf \approx 1$ on a subset of $H_t$ of measure $\gtrsim \delta^{\frac{d(d-1)}{2}}$. Hence $\|Xf\|_{q,r} \gtrsim \delta^{\frac{d(d-1)}{2r}}$, which proves the necessity of (4).

One may conjecture that

**Conjecture.** *If $p$, $q$ and $r$ satisfy the inequalities* (2), (3) *and* (4), *then* (1) *holds.*

We have the following theorem which contains the main result of the paper.

**Theorem 1.2.** *The conjecture is true in $\mathbb{R}^d$ for $d = 3$, 4 or 5 except the end-point issues. More explicitly, if $p$, $q$ and $r$ satisfy* (2), (3) *and* (4) *with inequalities replaced with strict inequalities, then* (1) *holds in $\mathbb{R}^d$ for $d = 3$, 4 or 5.*

The case $d = 3$ follows from Wolff's above-mentioned mixed norm estimates for the $X$-ray transform restricted to light rays [21], since in $\mathbb{R}^3$ the space of light rays is a rigid well-curved line complex.

If one considers the case $q = r$ only, the conjecture had been settled for $d = 3$ in [8] and [17], and for the case $q = r$ and $d = 4$, it had been verified except the endpoint issues in [10]. In higher dimensions, the conjecture was verified for $p = \frac{d}{d-1}$ and $q = r = \frac{d-1}{d-2}$ in [13] and for $q = r = 2$ and $p = \frac{2d^2-2d}{d^2-d+2}$ in [9]. Note that the results mentioned here are valid for all rigid well-curved line complexes whereas Theorem 1.2 is valid only in the model case.

Remarks: i) Note that (1) holds for all $q$ and $r$ if $p = \infty$, since we are interested in local estimates.
ii) Fubini's theorem implies that (1) holds for $p = q = r = 1$.
iii) It is well-known that (see, e.g., [10] and [9]) $X$ is bounded from $W^{2,-\eta}$ to $L^2$ for some positive $\eta$. Here $W^{p,\varepsilon}(Q_1)$ is the Sobolev space consisting of all functions $f$ supported in $Q_1$ such that $\|(1-\triangle)^{\varepsilon/2} f\|_p < \infty$.

In the light of these remarks, Theorem 1.2. (the cases $d = 4$ and $d = 5$) can be obtained from the following theorem by interpolation.

**Theorem 1.3.** *Let $d = 4$ or 5. Let $p = q = \frac{d+2}{d}$ and $r = \frac{d^2+d-2}{d^2-d-2}$. Then, the restricted X-ray transform $X$ is bounded from the Sobolev space $W^{p,\varepsilon}(Q_1)$ to $L^q(L^r)$ for any $\varepsilon > 0$, where $Q_1$ is the unit cube in $\mathbb{R}^d$.*

In [21], Wolff used the "bush" construction. It was introduced by Bourgain in [1] and used by several other authors (see, e.g., [15]). A bush is a family of tubes passing through a common point. The basic observation there was the following; in the case of light rays the intersection of a bush with a tube passing through a point far from the bush is at most a small ball.

As in [21], in the proof of Theorem 1.3, we use the bush construction. The basic property of the bushes in our case is the following transversality property: Let $d \geq 4$. If the basepoint of one bush is far from another bush, then their intersection is at most a finite union of small balls. This is consequence of well-curvedness. This property yields the proof in $\mathbb{R}^4$.

However, for $d = 5$, this property by itself is not enough. The reason for this is that in $\mathbb{R}^5$ two generic bushes do not intersect at all. We overcome this difficulty by collecting the bushes into groups that we denote by "bushfields". A bushfield is a set of tubes intersecting a given tube which we call the basetube. In some aspects, this object is similar to that used in [18], which came to be recognized as the "hairbrush" (see, e.g., [12]). The main difference is that a bushfield behaves like a disjoint union of bushes. This is because of the following basic properties:
i) The tubes in a bushfield are disjoint away from the base tube.
ii) If the basepoint of a given bush $\beta$ in $\mathbb{R}^5$ is far from a given bushfield $bf$, then $\beta \cap bf$ consists of at most finitely many small balls, as in the case of two bushes in $\mathbb{R}^4$.

To make use of these properties, we use a standard technique which is usually called the bilinear reduction (see, e.g., [16], [15], [12] and [21]) together with the rescaling argument in [21]. However, our exposition of the proof of Theorem 1.3 is largely self-contained and does not require the reader to be familiar with the cited works.

In section 2, we discuss the bush decomposition lemma from [21]. In section 3, we prove the geometric properties of the bushes and obtain the main estimate in $\mathbb{R}^4$. In section 4, we discuss bushfields and the main estimate in $\mathbb{R}^5$. In section 5, we obtain a bilinear estimate for the adjoint of $X$. Finally, in section 6, we convert this bilinear estimate into a mixed norm estimate for $X$.

List of notation.

$|A|$: Cardinality or the measure of the set $A$.
$\chi_A$: Characteristic function of the set $A$.
$\mathcal{N}(A, \eta)$: $\eta$ neighborhood of the set $A$.
$C$: A constant which may vary from line to line.
$A \lesssim B$: $A \leq CB$.
$A \approx B$: $A \lesssim B$ and $B \lesssim A$.

## 2. Bush decomposition lemma.

Fix $\delta > 0$. We work with tubes $\tau \subset \mathbb{R}^d$ such that the axis of $\tau$ is a $\gamma_d$-ray and it has dimensions $\delta \times ... \times \delta \times 1$. Two $\delta$-tubes are called $\delta$-separated if the distance between their axis with respect to a (fixed) smooth metric on $G_d$ is greater than $\delta$.

We say two segments of $\gamma_d$ are disjoint if the distance between them is positive. Fix two disjoint segments $W$ and $B$ of $\gamma_d$. We call a tube whose axis direction belongs to $W$ (resp. $B$) a white (resp. black) tube. Also fix two arbitrary $\delta$-separated families of white and black tubes, $\mathcal{W}$ and $\mathcal{B}$ respectively. Until the end of Lemma 5.1, we work with these $\delta$, $\mathcal{W}$ and $\mathcal{B}$.

Let $\Phi_S$ denotes the sum of the characteristic functions of the objects in the set $S$, e.g. $\Phi_{\mathcal{W}}$, $\Phi_{\mathcal{B}}$.

In the sections 2-5, we estimate the $L^{p'}$ norm of the function $\min(\Phi_{\mathcal{W}}, \Phi_{\mathcal{B}})$. This can be considered as a bilinear estimate for the adjoint of $X$. We begin with the following bush decomposition lemma of Wolff [21]. We give a proof for the reader's convenience. A bush [1] is a set of tubes passing through a common point $p$, which is called a base point for the bush. A white (resp. black) bush means a bush consisting of white (resp. black) tubes. Given a set $\mathcal{W}$ of $\delta$-tubes, we define a $\mu$-fold point for $\mathcal{W}$ to be a point contained in at least $\mu$ tubes from $\mathcal{W}$ or equivalently a point $x$ such that $\Phi_{\mathcal{W}}(x) \geq \mu$.

**Lemma 2.1.** *Given a set $\mathcal{W}$ of $\delta$-tubes, we have a decomposition*

$$\mathcal{W} = \cup_{j=1}^{J} \mathcal{W}_j, \qquad 2^J \approx |\mathcal{W}|,$$

*such that*
*i) $\mathcal{W}_j$ is a union of $\lesssim 2^j$ bushes $\beta_i^j$, and any tube in $\mathcal{W}$ belongs to at most one of the bushes $\beta_i^j$.*
*ii) $\mathcal{W}_g^k := \cup_{j>k} \mathcal{W}_j$ does not have any $\frac{|\mathcal{W}|}{2^k}$-fold points, i.e. $\Phi_{\mathcal{W}_g^k} \leq \frac{|\mathcal{W}|}{2^k}$, for all $k \leq J$.*
*iii) $\mathcal{W}_b^k := \cup_{j \leq k} \mathcal{W}_j$ is a union of $\lesssim 2^k$ bushes.*

**Proof.** First, we prove the following lemma:

**Lemma 2.2.** *Given a set $\mathcal{W}$ of $\delta$-tubes and a positive number $\mu \leq |\mathcal{W}|$, we can decompose $\mathcal{W}$ as*

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

*where $\mathcal{W}_b$ is a union of $\lesssim \frac{|\mathcal{W}|}{\mu}$ bushes and $\mathcal{W}_g$ does not have any $\mu$-fold points.*

**Proof.** We construct $\mathcal{W}_b$ inductively. Take any $\mu$-fold point $x_1 \in \mathbb{R}^d$ for $\mathcal{W}$. The tubes in $\mathcal{W}$ containing $x_1$ forms a bush $\beta_1$. Let $\mathcal{W}_b = \beta_1$ and $\mathcal{W}^1 = \mathcal{W} \backslash \beta_1$. Repeat this procedure with $\mathcal{W}_1$ instead of $\mathcal{W}$. This gives another bush $\beta_2$. Let $\mathcal{W}_b = \beta_1 \cup \beta_2$ and $\mathcal{W}_2 = \mathcal{W}_1 \backslash \beta_2$. Continue to repeat this procedure until there is no $\mu$-fold points. Since we subtract at least $\mu$ tubes from $\mathcal{W}$ in each step, we stop at most in $\frac{|\mathcal{W}|}{\mu}$ steps. Note that this gives $\mathcal{W}_b = \cup_{i=1}^{k} \beta_i$, $k \leq \frac{|\mathcal{W}|}{\mu}$, and $\mathcal{W}_g := \mathcal{W}^k$ has no $\mu$-fold points. $\square$

**Proof of Lemma 2.1.** Apply Lemma 2.2 to $\mathcal{W}$ with $\mu = |\mathcal{W}|/2$. This gives a set $\mathcal{W}_g^1$ with no $|\mathcal{W}|/2$-fold points and a collection $\mathcal{W}_1$ of bushes $\beta_i^1$. Then apply Lemma

2.2 to $\mathcal{W}_g^1$ with $\mu = |\mathcal{W}|/4$ to obtain $\mathcal{W}_g^2$ with no $|\mathcal{W}|/2^2$-fold points and a collection $\mathcal{W}_2$ of bushes $\beta_i^2$. Continue to repeat this procedure taking $\mu = |\mathcal{W}|/2^j$ at the $j$th step. We stop the procedure at $J$th step, where $J$ is the smallest integer such that $|\mathcal{W}|/2^J < 1$. Note that $\mathcal{W} = \cup_{j=1}^J \mathcal{W}_j$ and by Lemma 2.2, $\mathcal{W}_j$ is a union of at most $|\mathcal{W}|/(|\mathcal{W}|/2^j) = 2^j$ bushes $\beta_i^j$. This yields the part i) of the lemma. Part ii) follows from the construction and part iii) immediately follows from part i). $\square$.

Lemma 2.1 gives a decomposition of $\mathcal{W}$ into a set of bushes $\beta_i^j$. At this point, we fix $\varepsilon > 0$ and a tiling of $Q_1$ by $\delta^\varepsilon$-cubes. The letter $Q$ is reserved for these $\delta^\varepsilon$-cubes. The following definitions are from [21].

**Definition.** A tube $w$ is related to a $\delta^\varepsilon$-cube $Q$, $w \sim Q$, if $w$ belongs to a bush $\beta_i^j$ whose basepoint is in $Q$ or one of its neighbors. Similarly, a tube $w$ is related to a point $x$, $w \sim x$, if $x$ is in a cube which is related to $w$.

**Definition.**

$$\Phi^*_{\mathcal{W}}(x) := \sum_{w \sim x} \Phi_{\mathcal{W}}(x), \qquad \tilde{\Phi}_{\mathcal{W}}(x) := \sum_{w \nsim x} \Phi_{\mathcal{W}}(x) = \Phi_{\mathcal{W}}(x) - \Phi^*_{\mathcal{W}}(x).$$

We use Lemma 2.1 for $\mathcal{B}$ too and define $\tilde{\Phi}_{\mathcal{B}}$ and $\Phi^*_{\mathcal{B}}$ similarly.

## 3. Main lemma in $\mathbb{R}^4$; bushes.

The following lemma is the main lemma of the proof in $\mathbb{R}^4$. Let $m = |\mathcal{W}|$, $n = |\mathcal{B}|$.

**Lemma 3.1.** *Let $d = 4$. With the notation in section 2, for any $\mu$ and $\nu$ we have*
*i)* $|\{x \in Q_1 : \tilde{\Phi}_{\mathcal{W}}(x) \geq \mu, \Phi_{\mathcal{B}} \geq \nu\}| \lesssim \delta^{\frac{5}{2} - C\varepsilon} \frac{nm^{\frac{1}{2}}}{\nu^2 \mu^{\frac{3}{2}}}$,
*ii)* $|\{x \in Q_1 : \Phi_{\mathcal{W}}(x) \geq \mu, \tilde{\Phi}_{\mathcal{B}} \geq \nu\}| \lesssim \delta^{\frac{5}{2} - C\varepsilon} \frac{n^{\frac{1}{2}} m}{\nu^{\frac{3}{2}} \mu^2}$.

We begin the proof with the following geometric lemma about the transversality of white and black $\delta$-bushes.

**Lemma 3.2.** *Fix $\varepsilon > 0$, and let $W$ and $B$ be two disjoint segments of $\gamma_4$. Let $x$ and $y$ be two arbitrary points in $2Q_1$ and $S_W$ (resp. $S_B$) be the surface consisting of all white (resp. black) rays passing from the point $x$ (resp. $y$). Let $Q_x$ be the $\delta^\varepsilon$-cube centered at the point $x$. Then*
*i) the measure of the intersection of the $\delta$ neighborhood of $S_W$ and a black $\delta$-tube is $\lesssim \delta^4$,*
*ii) the measure of the set $Q_1 \cap (\mathcal{N}(S_W, \delta) \backslash Q_x) \cap \mathcal{N}(S_B, \delta)$ is $\lesssim \delta^{-C\varepsilon} \delta^4$.*

**Proof.** We use the following parametrizations:

$$S_B = \{y + (a, at, at^2, at^3) : a \in (-2, 2), \gamma_4(t) \in B\},$$

$$S_W = \{x + (b, bs, bs^2, bs^3) : b \in (-2, 2), \gamma_4(s) \in W\},$$

and any black $\delta$-tube is the $\delta$ neighborhood of a line

$$l(t_0, z) = \{z + (c, ct_0, ct_0^2, ct_0^3), |c| < 2\},$$

where $t_0$ is a point such that $\gamma_4(t_0) \in B$.

i) It is easy to check that the intersection of $S_W$ and $l(t_0, z)$ consists of at most 2 points. The claim follows from the observations that the tangent plane $T(b, s)$ of $S_W$ at the point corresponding to the parameter values $(b, s)$ is spanned by the vectors $e_1 = (0, 1, 2s, 3s^2)$ and $e_2 = (1, s, s^2, s^3)$, and the angle between $l(t_0, z)$ and $T(b, s)$ is greater than a fixed constant depending on the distance between $W$ and $B$. We omit the details.

ii) It is easy to check that for fixed $x$, the intersection of $S_W$ and $S_B$ consists of $\lesssim 1$ points for $y$ in a dense subset of $\mathbb{R}^4$. Therefore, by changing $y$ slightly if necessary and replacing $\delta$ with $2\delta$, we can assume that $S_W \cap S_B$ consists of $\lesssim 1$ points.

Note that if $E$ and $F$ are subsets of a metric space, then

$$\mathcal{N}(E, \delta) \cap \mathcal{N}(F, \delta) \subseteq \mathcal{N}(E \cap \mathcal{N}(F, 2\delta), \delta);$$

hence, it suffices to prove that the induced Lebesque measure of the set of points on $S_B \cap Q_1$, which are in the $4\delta$ neighborhood of $S_W \backslash Q_x$, is $\lesssim \delta^{-C\varepsilon}\delta^2$.

Let $A_\lambda = \{z : |z - y| \in [\lambda, 2\lambda]\}$. We prove that forall $\lambda \in (0, \frac{1}{2})$, the measure of the set of points on $S_B \cap A_\lambda \cap Q_1$ which are in the $4\delta$ neighborhood of $S_W \backslash Q_x$ is $\lesssim \delta^{-C\varepsilon}\delta^2$. This yields the claim since $S_B \cap Q_1$ can be covered by $\lesssim \log(\delta^{-1})$ $A_\lambda$'s.

Note that the area element on the surface $S_B$ is

$$\mathrm{d}A = f(a, t)a \, \mathrm{d}a\mathrm{d}t, \tag{5}$$

where $f$ is a bounded function. Hence, the measure of a subset of $S_B \cap A_\lambda$ of the form $\{x + (a, at, at^2, at^3) : |a - a_0| < \alpha, |t - t_0| < \frac{\alpha}{\lambda}\}$ is $\lesssim \alpha^2$. Therefore, by using part (i) of the lemma, we only need to show that the measure of the set

$$S_t := \{t \in [-1, 1] : \exists a, b, s \text{ such that } |F(a, b, t, s)| \leq 4\delta\}$$

is $\lesssim \delta^{-\varepsilon}\frac{\delta}{\lambda}$, where $F : \mathbb{R}^4 \to \mathbb{R}^4$ is the function defined via

$$F(a, b, t, s) = x - y + (a - b, at - bs, at^2 - bs^2, at^3 - bs^3).$$

Note that any derivative of $F$ of order less than two is bounded by $C$ and

$$JF = \det \begin{pmatrix} 1 & t & t^2 & t^3 \\ -1 & -s & -s^2 & -s^3 \\ 0 & a & 2at & 3at^2 \\ 0 & -b & -2bs & -3bs^2 \end{pmatrix} = ab(t - s)^4 \gtrsim \lambda\delta^\varepsilon C.$$

Hence, a quantitative version of the inverse function theorem, for example the one in [4], implies that $F^{-1}(B(0,4\delta))$ is contained in $\lesssim 1$ balls of diameter $\lesssim \delta^{-\varepsilon}\frac{\delta}{\lambda}$. This shows that the measure of the set $S_t$ is $\lesssim \delta^{-\varepsilon}\frac{\delta}{\lambda}$. $\square$

**Proof of Lemma 3.1.** We prove part i) only.

Let $j_0$ be the smallest integer so that $\frac{m}{2^{j_0}} \leq \frac{\mu}{2}$. By Lemma 2.1, we have $\Phi_{\mathcal{W}_g^{j_0}} \leq \frac{\mu}{2}$. Note that $\tilde{\Phi}_{\mathcal{W}} \leq \Phi_{\mathcal{W}_g^{j_0}} + \tilde{\Phi}_{\mathcal{W}_b^{j_0}}$. Therefore $\{\tilde{\Phi}_{\mathcal{W}} \geq \mu\} \subset \{\tilde{\Phi}_{\mathcal{W}_b^{j_0}} \geq \frac{\mu}{2}\}$ and it is enough to prove part i) with $\tilde{\Phi}_{\mathcal{W}_b^{j_0}}$ instead of $\tilde{\Phi}_{\mathcal{W}}$. Also by Lemma 2.1, $\mathcal{W}_b^{j_0}$ is a union of $\lesssim 2^{j_0} \lesssim \frac{m}{\mu}$ bushes. Similarly, let $k_0$ be the smallest integer so that $\frac{n}{2^{k_0}} \leq \frac{\nu}{2}$. Note that $\Phi_{\mathcal{B}} \leq \Phi_{\mathcal{B}_g^{k_0}} + \Phi_{\mathcal{B}_b^{k_0}}$; hence, by the same reasoning, it is enough to prove part i) with $\Phi_{\mathcal{B}_b^{k_0}}$ instead of $\Phi_{\mathcal{B}}$ and $\mathcal{B}_b^{k_0}$ is a union of $\lesssim 2^{k_0} \lesssim \frac{n}{\nu}$ bushes.

Denote the bushes in $\mathcal{W}_b^{j_0}$ (resp. $\mathcal{B}_b^{k_0}$) by $\beta_w$ (resp. $\beta_b$). We have

$$\int \Phi_{\mathcal{B}_b^{k_0}} \tilde{\Phi}_{\mathcal{W}_b^{j_0}} = \sum_{\beta_b} \sum_{\beta_w} \int_{Q_1 \setminus 2Q} \Phi_{\beta_b} \Phi_{\beta_w}, \tag{6}$$

where $Q$ is the $\delta^{\varepsilon}$-cube containing the base of $\beta_w$.

Now, we divide each black bush into $\approx \log(\delta^{-1})$ disjoint segments $\beta_b^k$. The segment $\beta_b^0$ consists of the parts of the tubes which are in the $\delta$ neighborhood of the basepoint, and for $k > 0$, $\beta_b^k$ consists of the parts of the tubes whose distance to the basepoint is between $2^{k-1}\delta$ and $2^k\delta$. We have

$$(6) \lesssim \sum_{k=0}^{\log(\delta^{-1})} \sum_{\beta_b^k} \sum_{\beta_w} \int_{Q_1 \setminus 2Q} \Phi_{\beta_b^k} \Phi_{\beta_w}. \tag{7}$$

We need the following lemma to estimate the right hand side of the inequality (7).

**Lemma 3.3.** *Fix a black bush segment $\beta_b^k$.*
*i) There are $\lesssim 2^{2k}\delta^{-1}$ white tubes which intersect $\beta_b^k$.*
*ii) For any white bush $\beta_w$ which intersects $\beta_b^k$, we have*

$$\int_{Q_1 \setminus 2Q} \Phi_{\beta_b^k} \Phi_{\beta_w} \lesssim \delta^{3-C\epsilon} 2^{-k},$$

*where $Q$ is the $\delta^{\varepsilon}$-cube containing the basepoint of the white bush $\beta_w$,*
iii) For any white tube $w$ which intersects $\beta_b^k$, we have

$$\int_{Q_1} \Phi_{\beta_b^k} \chi_w \lesssim \delta^3 2^{-k}.$$

**Proof.** i) Note that there are at most $\delta^{-1}$ tubes through a given point, and (5) implies that the maximum possible cardinality of a $\delta$-separated set of points on $\beta_b^k$ is $\lesssim 2^{2k}$. Hence, there are at most $2^{2k}\delta^{-1}$ white tubes which intersect $\beta_b^k$.

ii) Part ii) of Lemma 1 shows that the measure of the set of points which belong to both $\beta_b^k$ and $\beta_w$ is $\lesssim \delta^{4-C\varepsilon}$. The claim follows from the following pointwise inequalities:

$$\Phi_{\beta_b^k} \lesssim 2^{-k}\delta^{-1}, \tag{8}$$

$$\Phi_{\beta_w}\chi_{Q_1\setminus 2Q} \lesssim \delta^{-\varepsilon}. \tag{9}$$

To prove (8), note that the angle between the axis of the adjacent tubes is $\gtrsim \delta$. Also note that the distance between the points on $\beta_b^k$ and the basepoint of $\beta_b^k$ is at least $2^k\delta$. These show that at most $2^{-k}\delta^{-1}$ many tubes passes through a given point on $\beta_b^k$.

Proof of (9) is similar, since the points in the complement of $2Q$ are at least at a distance $\delta^{\varepsilon}$ to the basepoint of the bush.

iii) This follows from part i) of Lemma 3.3 and (8). □

We continue the proof of part i) of Lemma 3.1. Fix a black bush segment $\beta_b^k$. Using part ii) of Lemma 3.3, and remembering that there are at most $\frac{m}{\mu}$ white bushes, we obtain

$$\sum_{\beta_w}\int_{Q_1\setminus 2Q}\Phi_{\beta_b^k}\Phi_{\beta_w} \lesssim \frac{m}{\mu}\delta^{3-C\epsilon}2^{-k}. \tag{10}$$

On the other hand, parts i) and iii) of Lemma 3.3 imply that

$$\sum_{\beta_w}\int_{Q_1\setminus 2Q}\Phi_{\beta_b^k}\Phi_{\beta_w} \lesssim 2^{2k}\delta^{-1}\delta^{3}2^{-k}. \tag{11}$$

Using (10) and (11) in (7), and remembering that there are at most $\frac{n}{\nu}$ black bushes, we obtain

$$(7) \lesssim \sum_{k=0}^{\log(\delta^{-1})}\frac{n}{\nu}\min(\frac{m}{\mu}, 2^{2k}\delta^{-1})\delta^{3-C\epsilon}2^{-k} \lesssim \log(\delta^{-1})\frac{n}{\nu}(\frac{m}{\mu})^{1/2}\delta^{5/2-C\varepsilon},$$

which yields the claim of part i) using Tschebyshev's inequality. □

## 4. Main lemma in $\mathbb{R}^5$; bushfields.

The following lemma is the main lemma for the proof in $\mathbb{R}^5$. Let $m = |\mathcal{W}|$, $n = |\mathcal{B}|$.

**Lemma 4.1.** *Let $d = 5$. With the notation in section 2, for any $\mu$ and $\nu$ we have*
*i)* $|\{x \in Q_1 : \tilde{\Phi}_{\mathcal{W}}(x) \geq \mu, \Phi_{\mathcal{B}} \geq \nu\}| \lesssim \delta^{\frac{7}{2}-C\varepsilon}\frac{nm^{\frac{1}{4}}}{\nu^2\mu^{\frac{3}{4}}}$,
*ii)* $|\{x \in Q_1 : \Phi_{\mathcal{W}}(x) \geq \mu, \tilde{\Phi}_{\mathcal{B}} \geq \nu\}| \lesssim \delta^{\frac{7}{2}-C\varepsilon}\frac{n^{\frac{1}{4}}m}{\nu^{\frac{3}{4}}\mu^2}$.

In the proof of the lemma, we use a geometric construction called bushfield. A bushfield is a set of tubes intersecting a common tube $\tau$; we call $\tau$ the basetube of the bushfield. We call a bushfield consisting of white (resp. black) tubes a white (resp. black) bushfield. We begin the proof with the following lemma about the geometric properties of the bushfields.

**Lemma 4.2.** *Let $bf$ be a bushfield of white $\delta$-tubes with basetube $w$ and $\beta$ be a bush of black $\delta$-tubes with basepoint $p$. Let $A_\lambda$ be the cylinder $A_\lambda = \{y \in \mathbb{R}^5 : dist(w, y) \in [\lambda, 2\lambda]\}$. Then*
*i) If $y \in A_\lambda$, then $\Phi_{bf}(y) \lesssim \lambda^{-1}$.*
*ii) $|bf \cap A_\lambda| \lesssim \lambda^2\delta^2$; hence, there are at most $\lambda^2\delta^{-4}$ $\delta$-separated tubes intersecting $bf \cap A_\lambda$.*
*iii) $|(bf \cap A_\lambda) \cap (b\backslash\mathcal{N}(p, \delta^\varepsilon))| \lesssim \delta^{-C\varepsilon}\delta^5$.*

**Proof.** Using the maps $T_N^t$ that is defined before Lemma 6.2, it is easy to see that for all $s$ and $t$ in $[-1, 1]$, there exists a linear map $T_s^t$, which takes the curve $\gamma_5$ to itself and in particular takes $\gamma_5(s)$ to $\gamma_5(t)$, such that the entries in the matrix representation of $T_s^t$ and its inverse are bounded by a fixed constant. Because of this and translation invariance, it is enough to prove the lemma by assuming that $W$ is a segment around $\gamma_5(0)$, and the basetube of $bf$ is the $\delta$ neighborhood of the line $l(\gamma_5(0), 0)$.

Note that $bf$ is contained in the $2\delta$ neighborhood of the set

$$S_{bf} = \{(u, 0, 0, 0, 0) + a(1, t, t^2, t^3, t^4) | u \in (-1, 1), a \in (-2, 2), \gamma_5(t) \in W\}. \tag{12}$$

It is easy to see that $S_{bf}$ can also be parametrized as

$$S_{bf} = \{(u, a, at, at^2, at^3) : u \in (-2, 2), a \in (-2, 2), \gamma_5(t) \in W\}. \tag{13}$$

Using this parametrization, we see that $bf \cap A_\lambda$ is contained in the $2\delta$ neighborhood of

$$S_{bf}^\lambda = \{(u, a, at, at^2, at^3) : u \in (-2, 2), |a| \in [\lambda/2, 2\lambda], \gamma_5(t) \in W\}. \tag{14}$$

Also as before, we define

$$S_\beta = \{(b, bs, bs^2, bs^3, bs^4) : |b| \in (\delta^{-\varepsilon}, 2), \gamma_5(s) \in B\}. \tag{15}$$

Note that $\beta\backslash\mathcal{N}(p, \delta^\varepsilon)$ is contained in the $2\delta$ neighborhood of the set $p + S_\beta$.

i) Let $bf_i$ be the set of tubes in $bf$ whose direction is $\gamma_5(t)$ for some $t \in [i\delta, (i+1)\delta]$. Note that because of $\delta$-separatedness every point in $\mathbb{R}^5$ is contained in $\lesssim 1$ of the tubes in $bf_i$. Let $P_i^\delta$ be the $2\delta$ neighborhood of the 2-plane $P_i$ through the origin which is spanned by the vectors $\gamma_5(0)$ and $\gamma_5(i\delta)$. Note that all of the tubes in $bf_i$ are contained in $P_i^\delta$. Also note that the angle between the planes $P_i$ and $P_j$ is

$$\begin{aligned}\measuredangle(P_i,P_j) &\approx \measuredangle(\gamma_5(i\delta)-\gamma_5(0),\gamma_5(j\delta)-\gamma_5(0))\\ &\approx \measuredangle((0,1,i\delta,(i\delta)^2,(i\delta)^3),(0,1,j\delta,(j\delta)^2,(j\delta)^3))\approx |i-j|\delta.\end{aligned}$$

This and the observation that the distance between the points in $A_\lambda$ and the basetube is approximately $\lambda$ show that any point in $A_\lambda$ is contained in $\lesssim \frac{1}{\lambda}$ $P_i^\delta$'s, which is the claim of part i).

ii) Note that the volume element on $S_{bf}$ with respect to the parametrization (13) is

$$\mathrm{d}W = f(a,t)a\,\mathrm{d}u\mathrm{d}a\mathrm{d}t, \tag{16}$$

where $f$ is a bounded function. This and (14) prove the first part. The second part follows from the observations that there are at most $\lambda^2\delta^{-3}$ $\delta$-separated points on $bf \cap A_\lambda$, and at most $\delta^{-1}$ $\delta$-separated tubes pass through a given point.

iii) This is similar to the proof of Lemma 3.2. Using the parametrization (12), it is easy to check that for $p$ in a dense subset of $\mathbb{R}^5$, the intersection of $S_{bf}$ and $S_\beta$ consists of $\lesssim 1$ points. Hence, by changing $p$ slightly if necessary and replacing $\delta$ with $2\delta$, we can assume that $S_{bf}\cap S_\beta$ consists of $\lesssim 1$ points.

As in the proof of Lemma 3.2 ii), it suffices to prove that the induced Lebesque measure of the set of points on $S_{bf}\cap A_\lambda$, which are in the $4\delta$ neighborhood of $S_\beta\backslash Q_x$, is $\lesssim \delta^{-C\varepsilon}\delta^3$.

(16) implies that the measure of a subset of $S_{bf}\cap A_\lambda$ of the form $\{(u,a,at,at^2,at^3) : |u-u_0|<\alpha, |a-a_0|<\alpha, |t-t_0|<\frac{\alpha}{\lambda}\}$ is $\lesssim \alpha^3$. Also note that for fixed $t$, the intersection of the 2-plane $\{(u,a,at,at^2,at^3) : |u|<2, |a|<2\}$ with the $4\delta$ neighborhood of $S_\beta$ is of measure $\lesssim \delta^2$. This is because of the transversality as in the proof of Lemma 3.2 i). Hence, it suffices to prove that the measure of the set

$$S_t = \{t : \exists u,a,b,s \text{ such that } |F(u,a,b,t,s)| \le 4\delta\}$$

is $\lesssim \delta^{-\varepsilon}\frac{\delta}{\lambda}$, where $F:\mathbb{R}^5\to\mathbb{R}^5$ is the function defined via

$$F(u,a,b,t,s) = p + (u-b, a-bs, at-bs^2, at^2-bs^3, at^3-bs^4).$$

Note that any derivative of $F$ of order less than two is bounded by $C$ and

$$JF = \det\begin{pmatrix} 1 & 0 & 0 & 0 & 0\\ 0 & 1 & t & t^2 & t^3\\ -1 & -s & -s^2 & -s^3 & -s^4\\ 0 & 0 & a & 2at & 3at^2\\ 0 & -b & -2bs & -3bs^2 & -4bs^3\end{pmatrix} = abC_{W,B} \gtrsim \lambda\delta^\varepsilon,$$

where $C_{W,B}$ is a constant which depends on the distance between $W$ and $B$ only. Hence, $F^{-1}(B(0,4\delta))$ is contained in $\lesssim 1$ balls of diameter $\lesssim \delta^{-\varepsilon}\frac{\delta}{\lambda}$. This shows that the measure of the set $S_t$ is $\lesssim \delta^{-\varepsilon}\frac{\delta}{\lambda}$. $\square$

**Proof of Lemma 4.1.** We prove part i) only.

Let $j_0$ be the smallest integer so that $\frac{m}{2^{j_0}} \leq \frac{\mu}{2}$. Using Lemma 2.1 as in the proof of Lemma 3.1, we note that it is enough to prove part i) with $\tilde{\Phi}_{\mathcal{W}_b^{j_0}}$ instead of $\tilde{\Phi}_{\mathcal{W}}$, and $\mathcal{W}_b^{j_0}$ is a union of $\lesssim 2^{j_0} \lesssim \frac{m}{\mu}$ bushes.

Now, we decompose the black tubes into bushfields. Let $\Omega$ be the set $\{\Phi_{\mathcal{B}} > \nu/2\}$. Fix a number $\eta \in (0,1)$ which is determined later. We need the following lemmas.

**Lemma 4.3.** *Let $\tau$ be a black tube. If $|\tau \cap \Omega| \geq \eta|\tau| \approx \eta\delta^4$, then $\tau$ intersects $\gtrsim \eta\nu^2$ tubes from $\mathcal{B}$.*

**Proof.** Without loss of generality, we can assume that $\tau$ is the $\delta$-tube with axis $l(\gamma_5(0),0)$. Note that the set $\tau \cap \Omega$ is covered by the black tubes $\approx \nu/2$ times. Since we are trying to find a lower bound for the number of tubes required to cover $\tau \cap \Omega$ $\nu/2$ times, we can assume that all the tubes which intersect $\tau$ in a small angle are in the covering.

Let $\mathcal{B}_i$ be the set of tubes $b$ in $\mathcal{B}$ which intersect $\tau$ and such that the direction of $b$ is $\gamma_5(t)$ for some $t \in [i\delta,(i+1)\delta]$. Note that using the tubes in $\mathcal{B}_i$, one can cover the set $\tau \cap \Omega$ at most once. This is because of $\delta$-separatedness.

The angle between $\tau$ and the tubes in $\mathcal{B}_i$ is approximately $i\delta$. This shows that $|\tau \cap b| \lesssim \frac{\delta^4}{|i|+1}$; hence, to cover the set $\tau \cap \Omega$ with the tubes in $\mathcal{B}_i$, we need at least $\eta(|i|+1)$ tubes from $\mathcal{B}_i$. This yields the claim of the lemma, since we have to cover the set $\tau \cap \Omega$ approximately $\nu/2$ times and $\sum_{i=0}^{\nu/2}\eta(i+1) \approx \eta\nu^2$. $\square$

**Lemma 4.4.** *Given $\eta > 0$ we can decompose $\mathcal{B}$ as*

$$\mathcal{B} = \mathcal{B}_r \cup \mathcal{B}_s,$$

*where each tube $b$ in $\mathcal{B}_r$ satisfies $|b \cap \Omega| \leq \eta|b|$, and $\mathcal{B}_s$ is a union of $\lesssim \frac{n}{\eta\nu^2}\log(\delta^{-1})$ bushfields.*

**Proof.** Let $A$ be a large enough constant. Choose $\frac{An}{\eta\nu^2}\log(\delta^{-1})$ tubes from $\mathcal{B}$ randomly. The following claim yields the lemma.

**Claim.** With high probability all the tubes $b$ in $\mathcal{B}$ with $|b \cap \omega| \geq \eta|b|$ intersect at least one of the tubes from the random sample.

**Proof of the claim.** Lemma 4.3 implies that $b$ intersects at least $\eta\nu^2$ tubes; hence, $b$ intersects none of the tubes from the random sample with probability $\lesssim (1-\frac{\eta\nu^2}{n})^{\frac{An}{\eta\nu^2}\log(\delta^{-1})} \approx \delta^A$. This shows that the above-mentioned probability is $\geq 1 - Cn\delta^A$, which is $\geq \frac{1}{2}$ if $A$ is large enough. $\square$

Choose such a sample. Let $\mathcal{B}_s$ be the set of tubes which intersect one of the tubes in the sample and $\mathcal{B}_r$ be the set of remaining tubes. Obviously, $\mathcal{B}_s$ is a union of $\lesssim \frac{n}{\eta\nu^2}\log(\delta^{-1})$ bushfields, and any tube $b \in \mathcal{B}_r$ satisfies $|b \cap \Omega| \leq \eta|b|$. $\square$

We continue the proof of Lemma 4.1. Note that

$$\{\tilde{\Phi}_{\mathcal{W}_b^{j_0}} \geq \mu, \Phi_{\mathcal{B}} \geq \nu\} \subseteq \{\tilde{\Phi}_{\mathcal{W}_b^{j_0}} \geq \mu, \Phi_{\mathcal{B}_s} \geq \frac{\nu}{2}\} \cup \{\Phi_{\mathcal{B}_r} \geq \frac{\nu}{2}\}. \tag{17}$$

Using Lemma 4.4, we obtain

$$\|\Phi_{\mathcal{B}_r}\chi_{\{\Phi_{\mathcal{B}_r}\geq\frac{\nu}{2}\}}\|_1 \lesssim \sum_{b\in\mathcal{B}_r} |b \cap \{\Phi_{\mathcal{B}_r} \geq \frac{\nu}{2}\}| \lesssim \eta\delta^4 n.$$

Thus,

$$|\{\Phi_{\mathcal{B}_r} \geq \frac{\nu}{2}\}| \lesssim \frac{\eta n}{\nu}\delta^4. \tag{18}$$

Now, we estimate the measure of the set $\{\tilde{\Phi}_{\mathcal{W}_b^{j_0}} \geq \mu, \Phi_{\mathcal{B}_s} \geq \frac{\nu}{2}\}$ as in the proof of Lemma 3.1. Denote the bushfields in $\mathcal{B}_s$ by $bf$ and white bushes by $\beta_w$. We have

$$\int \Phi_{\mathcal{B}_s}\tilde{\Phi}_{\mathcal{W}_b^{j_0}} = \sum_{bf}\sum_{\beta_w}\int_{Q_1\setminus 2Q} \Phi_{bf}\Phi_{\beta_w}, \tag{19}$$

where $Q$ is the $\delta^{\varepsilon}$-cube containing the base of $\beta_w$.

Now, we divide each black bushfield into $\approx \log(\delta^{-1})$ disjoint segments $bf^k$. The segment $bf^0$ consists of the parts of the tubes which are in the $\delta$ neighborhood of the basetube, and for $k > 0$, $bf^k$ consists of the parts of the tubes whose distance to the basetube is between $2^{k-1}\delta$ and $2^k\delta$. We have

$$(19) \lesssim \sum_{k=0}^{\log(\delta^{-1})}\sum_{bf^k}\sum_{\beta_w}\int_{Q_1\setminus 2Q} \Phi_{bf^k}\Phi_{\beta_w}. \tag{20}$$

Fix a black bushfield segment $bf^k$. Note that, as in the case $d = 4$, $\chi_{Q_1\setminus 2Q}\Phi_{\beta_w} \lesssim \delta^{-\varepsilon}$. Using this and parts i) and iii) of Lemma 4.2, and remembering that there are at most $\frac{m}{\mu}$ white bushes, we obtain

$$\sum_{\beta_w}\int_{Q_1\setminus 2Q} \Phi_{bf^k}\Phi_{\beta_w} \lesssim \frac{m}{\mu}\delta^{4-C\epsilon}2^{-k}. \tag{21}$$

On the other hand, part ii) of Lemma 4.2 shows that there are at most $2^{2k}\delta^{-2}$ white tubes which intersect $bf^k$. Using this and parts i) and iii) of Lemma 4.2, we obtain

$$\sum_{\beta_w}\int_{Q_1\setminus 2Q} \Phi_{bf^k}\Phi_{\beta_w} \lesssim 2^k\delta^{2-C\varepsilon}. \tag{22}$$

Using (21) and (22) in (20), and remembering that there are at most $\frac{n}{\eta\nu^2}(\log(\delta^{-1}))^2$ black bushfields, we obtain

$$
\begin{aligned}
(20) \quad &\lesssim \sum_{k=0}^{\log(\delta^{-1})} \frac{n}{\eta\nu^2}(\log(\delta^{-1}))^2 \min(\frac{m}{\mu}, 2^{2k}\delta^{-2})\delta^{4-C\epsilon}2^{-k} \\
&\lesssim (\log(\delta^{-1}))^3 \frac{n}{\eta\nu^2}(\frac{m}{\mu})^{1/2}\delta^{3-C\varepsilon}.
\end{aligned}
$$

Thus, using Tschebyshev's inequality, we obtain

$$
|\{\tilde{\Phi}_{\mathcal{W}_b^{j_0}} \geq \mu, \Phi_{\mathcal{B}_s} \geq \frac{\nu}{2}\}| \lesssim \frac{nm^{\frac{1}{2}}}{\eta\nu^3\mu^{\frac{3}{2}}}\delta^{3-C\varepsilon}. \tag{23}
$$

Using (18) and (23) in (17), we obtain

$$
|\{\tilde{\Phi}_{\mathcal{W}_b^{j_0}} \geq \mu, \Phi_{\mathcal{B}} \geq \nu\}| \lesssim \frac{\eta n}{\nu}\delta^4 + \frac{nm^{\frac{1}{2}}}{\eta\nu^3\mu^{\frac{3}{2}}}\delta^{3-C\varepsilon}. \tag{24}
$$

Minimizing the right hand side of the inequality (24) by choosing a suitable $\eta$ yields the claim of the lemma. □

## 5. Bilinear Estimate.

In this subsection, we estimate the $L^{p'}$ norm of the function $\min(\Phi_{\mathcal{W}}, \Phi_{\mathcal{B}})$. We need the following numerical inequalities. For proofs see [21]. Let $\theta \in [\frac{1}{2}, 1]$ and $a_j, b_k, a, b, x$ and $y$ be nonnegative real numbers. Then

$$
\min(ax, by)^\theta \max(ax, by)^{1-\theta} \leq \min(x, y)^\theta \max(x, y)^{1-\theta} \max(a, b)^\theta \min(a, b)^{1-\theta}, \tag{25}
$$

$$
\min(\sum_j a_j, \sum_k b_k)^\theta \max(\sum_j a_j, \sum_k b_k)^{1-\theta} \leq \sum_{j,k} \min(a_j, b_k)^\theta \max(a_j, b_k)^{1-\theta}. \tag{26}
$$

The following inequality is an immediate corollary of (26). Let $a, b, c$ and $d$ be nonnegative real numbers. Then

$$
\min(a+b, c+d)^\theta \max(a+b, c+d)^{1-\theta} \lesssim a^{1-\theta}(b+c)^\theta + c^{1-\theta}(c+d)^{1-\theta}. \tag{27}
$$

For technical reasons, we work with the function $\Psi_\theta$ defined below instead of $\min(\Phi_{\mathcal{W}}, \Phi_{\mathcal{B}})$. This is because of the asymmetry of the bounds in Lemmas 3.1 and 4.1. Here, $\theta$ is a dimension dependent parameter in $[\frac{1}{2}, 1]$.

**Definition.**

$$\Psi_\theta := \chi_{Q_1}.\min(\Phi_{\mathcal{W}}, \Phi_{\mathcal{B}})^\theta \max(\Phi_{\mathcal{W}}, \Phi_{\mathcal{B}})^{1-\theta},$$
$$S_\theta := \chi_{Q_1}.\min(\Phi^*_{\mathcal{W}}, \Phi^*_{\mathcal{B}})^\theta \max(\Phi^*_{\mathcal{W}}, \Phi^*_{\mathcal{B}})^{1-\theta},$$
$$T_\theta := \chi_{Q_1}.(\tilde{\Phi}_{\mathcal{W}}^{1-\theta}\Phi_{\mathcal{B}}^\theta + \Phi_{\mathcal{W}}^\theta\tilde{\Phi}_{\mathcal{B}}^{1-\theta}).$$

Note that the inequality (27) implies that

$$\Psi_\theta \le S_\theta + T_\theta. \tag{28}$$

By using the estimates in Lemmas 3.1 and 4.1, we obtain an estimate for $T_\theta$, and using the rescaling and induction arguments from [21], we prove the same estimate for $\Psi_\theta$. In some sense, the estimates in Lemmas 3.1 and 4.1 are stronger than the estimates we need; in the following lemma we bring them into the relevant form using trivial estimates.

**Lemma 5.1.** *Let $\theta = \frac{1}{2}$ for $d = 4$ and $\theta = \frac{4}{7}$ for $d = 5$. Let $p = \frac{d+2}{d}$ and $\frac{1}{p} + \frac{1}{p'} = 1$. Then*

$$\|\delta T_\theta\|_{p'}^{p'} \lesssim \delta^{-C\varepsilon}(\delta^d \max(|\mathcal{B}|, |\mathcal{W}|))^{\frac{d}{d-1}}. \tag{29}$$

**Proof.** First note that there are $\lesssim \delta^{-1}$ same colored tubes containing a given point. Hence,

$$\|\delta T_\theta\|_\infty \lesssim 1. \tag{30}$$

Also note that $\|\Phi_{\mathcal{W}}\|_1 \lesssim |\mathcal{W}|\delta^{d-1}$. This and the similar estimate for $\Phi_{\mathcal{B}}$ imply via Tschebyshev's inequality that

$$\{\Phi_{\mathcal{W}} \ge \mu, \Phi_{\mathcal{B}} \ge \nu\} \lesssim \delta^{d-1}\min(\frac{|\mathcal{W}|}{\mu}, \frac{|\mathcal{B}|}{\nu}). \tag{31}$$

i) The case $d = 4$: Let

$$Y(\mu,\nu) = |\{x \in Q_1 : \tilde{\Phi}_{\mathcal{W}}(x) \ge \mu, \Phi_{\mathcal{B}} \ge \nu\}|$$

and $m = |\mathcal{W}|$, $n = |\mathcal{B}|$. Using part i) of Lemma 3.1, we obtain

$$\begin{aligned} Y(\mu,\nu) &\lesssim \delta^{\frac{5}{2}-C\varepsilon}\min(\frac{nm^{\frac{1}{2}}}{\nu^2\mu^{\frac{3}{2}}}, \frac{m}{\mu}\delta^{\frac{1}{2}}) \\ &\le \delta^{\frac{5}{2}-C\varepsilon}(\frac{nm^{\frac{1}{2}}}{\nu^2\mu^{\frac{3}{2}}})^{\frac{2}{3}}(\frac{m}{\mu}\delta^{\frac{1}{2}})^{\frac{1}{3}} \\ &= \delta^{-C\varepsilon}\frac{(\delta^4 n\delta^4 m)^{\frac{2}{3}}}{(\nu\mu)^{\frac{4}{3}}}\delta^{-\frac{8}{3}}. \end{aligned} \tag{32}$$

Summing over the dyadic values of $\mu$ and $\nu$ between 1 and $\delta^{-1}$ gives

$$\|\chi_{Q_1}\sqrt{\tilde{\Phi}_{\mathcal{W}}\Phi_{\mathcal{B}}}\|_{\frac{8}{3}}^{\frac{8}{3}} \lesssim \delta^{-C\varepsilon}(\delta^4 n\delta^4 m)^{\frac{2}{3}}\delta^{-\frac{8}{3}}.$$

Estimating $\|\chi_{Q_1}\sqrt{\tilde{\Phi}_{\mathcal{B}}\Phi_{\mathcal{W}}}\|_{\frac{8}{3}}$ in the same way gives

$$\|\delta T_{\frac{1}{2}}\|_{\frac{8}{3}} \lesssim \delta^{-C\varepsilon}(\delta^4 n\delta^4 m)^{\frac{1}{4}}. \tag{33}$$

Interpolating (33) with (30) yields the claim of the lemma for $d=4$.

ii) The case $d=5$:

Define $Y(\mu,\nu)$ in the same way. Using part i) of Lemma 4.1, we obtain

$$Y(\mu,\nu) \lesssim \delta^{-C\varepsilon}\frac{nm^{\frac{1}{4}}}{\nu^2\mu^{\frac{3}{4}}}\delta^{\frac{7}{2}}.$$

Using $\mu \lesssim \delta^{-1}$, we obtain

$$Y(\mu,\nu) \lesssim \delta^{-C\varepsilon}\frac{nm^{\frac{1}{4}}}{\nu^2\mu^{\frac{3}{2}}}\delta^{\frac{11}{4}} \lesssim \delta^{-C\varepsilon}\frac{(\delta^5\max(m,n))^{\frac{5}{4}}}{\nu^2\mu^{\frac{3}{2}}}\delta^{-\frac{7}{2}}.$$

summing over the dyadic values of $\mu$ and $\nu$ between 1 and $\delta^{-1}$ gives

$$\|\chi_{Q_1}\tilde{\Phi}_{\mathcal{W}}^{\frac{3}{7}}\Phi_{\mathcal{B}}^{\frac{4}{7}}\|_{\frac{7}{2}}^{\frac{7}{2}} \lesssim \delta^{-C\varepsilon}(\delta^5\max(m,n))^{\frac{5}{4}}\delta^{-\frac{7}{2}}.$$

Estimating $\|\chi_{Q_1}\tilde{\Phi}_{\mathcal{B}}^{\frac{3}{7}}\Phi_{\mathcal{W}}^{\frac{4}{7}}\|_{\frac{7}{2}}$ in the same way gives

$$\|\delta T_{\frac{4}{7}}\|_{\frac{7}{2}}^{\frac{7}{2}} \lesssim \delta^{-C\varepsilon}(\delta^5\max(m,n))^{\frac{5}{4}}. \quad \square$$

**Lemma 5.2.** *Let $\theta=\frac{1}{2}$ for $d=4$ and $\theta=\frac{4}{7}$ for $d=5$. Let $p=\frac{d+2}{d}$ and $\frac{1}{p}+\frac{1}{p'}=1$. Fix two disjoint segments $W$ and $B$ of $\gamma_d$. For any $\delta>0$, we have: For any $\delta$-separated $\mathcal{W}$ and $\mathcal{B}$ the following inequality is valid*

$$\|\delta\Psi_\theta\|_{p'}^{p'} \lesssim \delta^{-C\varepsilon}(\delta^d\max(|\mathcal{B}|,|\mathcal{W}|))^{\frac{d}{d-1}}. \tag{34}$$

**Proof.** We begin with the following rescaling lemma.

**Lemma 5.3.** *Fix $\delta\in(0,\delta_0)$ ($\delta_0$ is determined in the proof) and assume that the claim of Lemma 5.2 has been proved for $\delta^{1-\varepsilon}$. Then, we have*

$$\|\delta\Psi_\theta\|_{L^{p'}(Q)}^{p'} \leq A_\varepsilon\delta^{\frac{C\varepsilon^2}{2}-C\varepsilon}((\delta^d\max(|\mathcal{B}|,|\mathcal{W}|))^{\frac{d}{d-1}}, \tag{35}$$

*where $Q$ is a $\delta^{\varepsilon}$-cube and $\mathcal{W}$ and $\mathcal{B}$ are $\delta$-separated sets of tubes.*

**Proof.** Fix a $\delta^{\varepsilon}$-cube $Q$. For each $w \in \mathcal{W}$, let $k(w)$ be the cardinality of the set of white tubes $w_1$ such that $w_1 \cap Q$ is contained in the double of $w$. Let $\mathcal{W}_\mu$ be the set of white tubes $w$ with $k(w) \in [\mu, 2\mu]$. Define $k(b)$ and $\mathcal{B}_\nu$ analogously. Note that $k(w)$ and $k(b)$ are restricted to values between 1 and $\delta^{-\varepsilon}$. Let

$$\Psi_\theta^{\mu\nu} := \chi_Q. \min(\Phi_{\mathcal{W}_\mu}, \Phi_{\mathcal{B}_\nu})^\theta \max(\Phi_{\mathcal{W}_\mu}, \Phi_{\mathcal{B}_\nu})^{1-\theta}.$$

Note that (26) implies that

$$\Psi_\theta \le \sum_{\mu\nu} \Psi_\theta^{\mu\nu}, \tag{36}$$

pointwise on $Q$, where the sum is over the dyadic values of $\mu$ and $\nu$. We estimate the $L^{p'}$ norm of the functions $\Psi_\theta^{\mu\nu}$. We can assume that $\mu \ge \nu$.

Let $\bar{\mathcal{W}}_\mu$ be a maximal subset of $\mathcal{W}_\mu$ which satisfies the property:

(*): *If $w_1, w_2 \in \bar{\mathcal{W}}_\mu$, then $w_1 \cap Q$ is not contained in the double of $w_2$.*

Define $\bar{\mathcal{B}}_\nu$ analogously. Replace the tubes in $\bar{\mathcal{W}}_\mu$ (resp. $\bar{\mathcal{B}}_\nu$) with their doubles and let

$$\bar{\Psi}_\theta^{\mu\nu} := \chi_Q. \min(\Phi_{\bar{\mathcal{W}}_\mu}, \Phi_{\bar{\mathcal{B}}_\nu})^\theta \max(\Phi_{\bar{\mathcal{W}}_\mu}, \Phi_{\bar{\mathcal{B}}_\nu})^{1-\theta}.$$

Note that the maximality of $\bar{\mathcal{W}}_\mu$ (resp. $\bar{\mathcal{B}}_\nu$) implies that $\Phi_{\mathcal{W}_\mu} \lesssim \mu \Phi_{\bar{\mathcal{W}}_\mu}$ (resp. $\Phi_{\mathcal{B}_\nu} \lesssim \nu \Phi_{\bar{\mathcal{B}}_\nu}$), which implies via (25) that

$$\Psi_\theta^{\mu\nu} \lesssim \mu^\theta \nu^{1-\theta} \bar{\Psi}_\theta^{\mu\nu}.$$

Taking the $L^{p'}$ norms, we obtain

$$\|\Psi_\theta^{\mu\nu}\|_{p'} \lesssim \mu^\theta \nu^{1-\theta} \|\bar{\Psi}_\theta^{\mu\nu}\|_{p'}. \tag{37}$$

Finally, note that the property (*) implies that

$$|\bar{\mathcal{W}}_\mu| \lesssim \mu^{-1}|\mathcal{W}|, \quad |\bar{\mathcal{B}}_\nu| \lesssim \nu^{-1}|\mathcal{B}|. \tag{38}$$

Dilating the cube $Q$ by $\delta^{-\varepsilon}$, we obtain a cube $Q'$ of side 1 and $\delta^{1-\varepsilon}$-separated sets $\bar{\mathcal{W}}_\mu$, $\bar{\mathcal{B}}_\nu$ of $2\delta^{1-\varepsilon}$-tubes. Hence, we can apply the hypothesis to obtain

$$\|\delta^{1-\varepsilon} \bar{\Psi}_\theta^{\mu\nu}(\delta^\varepsilon x)\|_{L^{p'}(Q')}^{p'} \le A_\varepsilon \delta^{-(1-\varepsilon)C\varepsilon} (\delta^{(1-\varepsilon)d} \max(|\bar{\mathcal{W}}_\mu|, |\bar{\mathcal{B}}_\nu|))^{\frac{d}{d-1}}.$$

Making the change of variables $x \to \delta^\varepsilon x$, we obtain

$$\delta^{-p'\varepsilon} \delta^{-d\varepsilon} \|\delta \bar{\Psi}_\theta^{\mu\nu}\|_{L^{p'}(Q)}^{p'} \le A_\varepsilon \delta^{-C\varepsilon} (\delta^d \max(|\bar{\mathcal{W}}_\mu|, |\bar{\mathcal{B}}_\nu|))^{\frac{d}{d-1}} \delta^{C\varepsilon^2} \delta^{-\frac{d^2}{d-1}\varepsilon}. \tag{39}$$

Using estimates in (38), we have

$$\max(|\bar{\mathcal{W}}_\mu|, |\bar{\mathcal{B}}_\nu|) \lesssim \max(\frac{|\mathcal{W}|}{\mu}, \frac{|\mathcal{B}|}{\nu}) \leq \frac{1}{\nu}\max(|\mathcal{W}|, |\mathcal{B}|). \tag{40}$$

Using estimate (39) and then estimate (40) in (37) and making the necessary cancellations, we get

$$\begin{aligned}
\|\delta\Psi_\theta^{\mu\nu}\|_{L^{p'}(Q)}^{p'} &\lesssim \mu^{\theta p'}\nu^{(1-\theta)p'}\|\delta\bar{\Psi}_\theta^{\mu\nu}\|_{L^{p'}(Q)}^{p'} \\
&\lesssim A_\varepsilon\delta^{-C\varepsilon}(\delta^d\max(|\bar{\mathcal{W}}_\mu|, |\bar{\mathcal{B}}_\nu|))^{\frac{d}{d-1}}\delta^{C\varepsilon^2}\mu^{\theta p'}\nu^{(1-\theta)p'}\delta^{p'\varepsilon}\delta^{-\frac{d}{d-1}\varepsilon} \\
&\lesssim A_\varepsilon\delta^{-C\varepsilon}(\delta^d\max(|\mathcal{W}|, |\mathcal{B}|))^{\frac{d}{d-1}}\delta^{C\varepsilon^2}\mu^{\theta p'}\nu^{(1-\theta)p'-\frac{d}{d-1}}\delta^{p'\varepsilon}\delta^{-\frac{d}{d-1}\varepsilon} \\
&\lesssim A_\varepsilon\delta^{-C\varepsilon}(\delta^d\max(|\mathcal{W}|, |\mathcal{B}|))^{\frac{d}{d-1}}\delta^{C\varepsilon^2};
\end{aligned} \tag{41}$$

the last inequality follows from the fact that $\mu, \nu \lesssim \delta^{-\varepsilon}$ when we note that $(1-\theta)p' > \frac{d}{d-1}$. Using (41) in (36), we have

$$\begin{aligned}
\|\delta\Psi_\theta\|_{p'} &\leq \sum_{\mu\nu}\|\delta\Psi_\theta^{\mu\nu}\|_{p'} \lesssim \sum_{\mu\nu}A_\varepsilon\delta^{-C\varepsilon}(\delta^d\max(|\mathcal{W}|, |\mathcal{B}|))^{\frac{d}{d-1}}\delta^{C\varepsilon^2} \\
&\lesssim A_\varepsilon\delta^{-C\varepsilon}(\delta^d\max(|\mathcal{W}|, |\mathcal{B}|))^{\frac{d}{d-1}}\delta^{C\varepsilon^2}\log(\delta^{-1})^2,
\end{aligned}$$

since there are $\lesssim \log(\delta^{-1})^2$ terms in the summation. This yields the claim of the lemma given that $\delta_0$ is small enough. □

We continue the proof of Lemma 5.2. Note that the lemma is obvious for $\delta \geq \delta_0$, and we prove the lemma for the values of $\delta$ such that $\delta^{1-\varepsilon} > \delta_0$. An obvious induction argument yields the claim of the lemma.

We estimate $T_\theta$ using Lemma 5.1 and estimate $S_\theta$ using Lemma 5.3 in the following way. For each $\delta^\varepsilon$-cube $Q$, applying Lemma 5.3 to the sets $n_{\mathcal{W}}(Q) := \{w \in \mathcal{W} : w \sim Q\}$ and $n_{\mathcal{B}}(Q) := \{b \in \mathcal{B} : b \sim Q\}$, we obtain

$$\|\delta S_\theta\|_{L^{p'}(Q)}^{p'} \leq \delta^{\frac{C\varepsilon^2}{2}-C\varepsilon}A_\varepsilon\left(\delta^d\max(n_{\mathcal{W}}(Q), n_{\mathcal{B}}(Q))\right)^{\frac{d}{d-1}}.$$

Summing over $Q$, we obtain

$$\begin{aligned}
\|\delta S_\theta\|_{p'}^{p'} &\leq \delta^{\frac{C\varepsilon^2}{2}-C\varepsilon}A_\varepsilon\sum_Q\left(\delta^d\max(n_{\mathcal{W}}(Q), n_{\mathcal{B}}(Q))\right)^{\frac{d}{d-1}} \\
&\leq \delta^{\frac{C\varepsilon^2}{2}-C\varepsilon}A_\varepsilon\left(\sum_Q\delta^d\max(n_{\mathcal{W}}(Q), n_{\mathcal{B}}(Q))\right)^{\frac{d}{d-1}} \\
&\lesssim \delta^{\frac{C\varepsilon^2}{2}-C\varepsilon}A_\varepsilon\left(\delta^d\max(\sum_Q n_{\mathcal{W}}(Q), \sum_Q n_{\mathcal{B}}(Q))\right)^{\frac{d}{d-1}}.
\end{aligned} \tag{42}$$

Note that $\sum_Q n_{\mathcal{W}}(Q) \lesssim |\mathcal{W}|$ and $\sum_Q n_{\mathcal{B}}(Q) \lesssim |\mathcal{B}|$. This is because by Lemma 2.1 each tube belongs to at most one bush; hence, each tube is related to $\lesssim 1$ $\delta^{\varepsilon}$-cubes. Using these bounds in (42), we obtain

$$\|\delta S_\theta\|_{p'}^{p'} \lesssim \delta^{\frac{C\varepsilon^2}{2}} A_\varepsilon \delta^{-C\varepsilon} (\delta^d \max(|\mathcal{W}|, |\mathcal{B}|))^{\frac{d}{d-1}},$$

which yields the claim of the lemma. □

## 6. Proof of Theorem 1.3.

For possible future usage, we do the rest of the proof in general dimensions. Lemma 5.2 and the following theorem yield the claim of Theorem 1.3.

**Theorem 6.1.** *Let $d \geq 3$ and $\varepsilon > 0$. Assume that $p$, $q$ and $r$ satisfy the inequalities* (2), (3) *and* (4), *and $r \geq q \geq p$. Let $W$ and $B$ be disjoint segments of $\gamma_d$. Assume that for any $\delta > 0$, and for any $\delta$-separated $\mathcal{W}$ and $\mathcal{B}$, we have*

$$\|\delta \min(\Phi_{\mathcal{W}}, \Phi_{\mathcal{B}})\|_{L^{p'}(Q_1)}^{p'} \lesssim \delta^{-\varepsilon} (\delta^d \max(|\mathcal{W}|, |\mathcal{B}|))^{\frac{p'}{r'}},$$

*where $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{r} + \frac{1}{r'} = 1$. Then the restricted X-ray transform $X$ is bounded from the Sobolev space $W^{p,C\varepsilon}(Q_1)$ to $L^q(L^r)$.*

In the proof of Theorem 6.1, we work with the operator

$$X_\delta f(l) = \frac{1}{\delta^{d-1}} \int_{l_\delta} f(x) dx,$$

where $l_\delta$ is the $\delta$ neighborhood of $l$ in $\mathbb{R}^d$. $X_\delta$ is simply the operator $X$ thickened by $\delta$. It is easy to see that the adjoint map $X_\delta^*$ of $X_\delta$ which takes functions defined on $G_d$ to functions defined on $\mathbb{R}^d$ is defined via

$$X_\delta^* f(u) = \int_{-1}^{1} \int_{H_t} \chi_{l_\delta(t,x)}(u) f(l(t,x)) dx dt.$$

The hypothesis of Theorem 6.1 is essentially a bilinear estimate for $X_\delta^*$; in the proof of Theorem 6.1, we convert it to a linear estimate. The argument is quite standard and we omit some details; the proof below is a variation of the one in [21].

We need the following rescaling map for the curve $\gamma_d$: Fix a point $\gamma_d(t_0)$ and consider the basis $\{\gamma_d(t_0), \gamma_d'(t_0), ..., \gamma_d^{(d-1)}(t_0)\}$ for $\mathbb{R}^d$. Define $T_N^{t_0}$ via $T_N^{t_0}(\gamma_d^{(j)}(t_0)) = N^j \gamma_d^{(j)}(t_0)$, $j = 0, 1, 2, ..., d-1$.

**Lemma 6.2.** *i) $T_N^{t_0}$ takes the curve $\gamma_d$ to itself, thus taking the $\gamma_d$-rays to $\gamma_d$-rays. Moreover, we have the following formula:*

$$T_N^{t_0}(\gamma_d(t)) = \gamma_d(N(t - t_0) + t_0).$$

*ii)* $T_N^{t_0}$ *takes a segment of length* $N^{-1}$ *centered at* $\gamma_d(t_0)$ *of the curve* $\gamma_d$ *to a segment of length* $\approx 1$*.*

**Proof.** We prove that $T_N^{t_0}(\gamma_d(t)) = \gamma_d(N(t-t_0)+t_0)$, which yields the claims of the lemma. Let $A$ be the $d \times d$ matrix whose $i$th column is $\gamma_d^{i-1}(t_0)$, i.e. $A = [\gamma_d(t_0)\ \gamma_d'(t_0)\ ...\ \gamma_d^{(d-1)}(t_0)]$ and $B = diag(1, N, ..., N^{d-1})$. Note that

$$T_N^{t_0}(\gamma_d(t)) = ABA^{-1}[1, t, t^2, ..., t^{d-1}]^T. \tag{43}$$

Let $f(t) = t^j$. Using the equality

$$f(t) = f(t_0) + f'(t_0)(t-t_0) + ... + f^{(j)}(t_0)\frac{(t-t_0)^j}{j!}, \tag{44}$$

and the definition of $\gamma_d(t)$, we have

$$[1, t, ..., t^{d-1}]^T = A[1, t-t_0, \frac{(t-t_0)^2}{2!}, ..., \frac{(t-t_0)^{d-1}}{(d-1)!}]^T. \tag{45}$$

Using (44) and (45) in (43), we have

$$\begin{aligned} T_N^{t_0}(\gamma_d(t)) &= AB[1, t-t_0, \frac{(t-t_0)^2}{2!}, ..., \frac{(t-t_0)^{d-1}}{(d-1)!}]^T \\ &= A[1, N(t-t_0), \frac{(N(t-t_0))^2}{2!}, ..., \frac{(N(t-t_0))^{d-1}}{(d-1)!}]^T. \end{aligned} \tag{46}$$

Using (45) by replacing $t$ with $N(t-t_0)+t_0$, we obtain

$$(46) = [1, N(t-t_0)+t_0, ..., (N(t-t_0)+t_0)^{d-1}] = \gamma_d(N(t-t_0)+t_0). \qquad \Box$$

Let $s$ be a segment of the curve $\gamma_d$ of length $N^{-1}$ centered at $\gamma_d(t_0)$. We denote $T_N^{t_0}$ by $T_s$ and the subset of $G_d$ consisting of all lines whose directions are in $s$ by $G_d^s$. Since $T_s$ takes $\gamma_d$-rays to $\gamma_d$-rays, there is an action $T_s : G_d \to G_d$. We give some more definitions:

**Definitions.** Let $Y$ be a subset of a metric space. We denote the characteristic function of $\mathcal{N}(Y, \eta)$ by $\chi_{Y,\eta}$.

Let $Y$ be a subset of $G_d^s$, then we have

(i) $\|\chi_{T_sY}\|_{p,r} \approx N^{\frac{1}{p}+\frac{d(d-1)}{2r}}\|\chi_Y\|_{p,r}$,

(ii) $\|\chi_{T_sY,\delta}\|_{p,r} \lesssim N^{\frac{1}{p}+\frac{d(d-1)}{2r}}\|\chi_{Y,\delta/N}\|_{p,r}$,

(iii) $X_\delta^*\chi_Y(x) \approx N^{-1}X_\delta^*\chi_{T_sY}(T_sx)$.

To prove (i), note that $T_s$ expands $s$ by a factor $\approx N$ by Lemma 6.2, and for any $\gamma_d(t) \in s$, $T_s$ expands volumes in $H_t$ by $\approx N^{\frac{d(d-1)}{2}}$. This follows from the observations that

$$\det(T_s) \approx N.N^2...N^{d-1} = N^{\frac{d(d-1)}{2}}, \tag{47}$$

and $T_s$ essentially preserves the lengths in $\gamma_d(t)$ direction. (ii) follows from (i) and the observation that $\mathcal{N}(T_sY, \delta) \subseteq T_s\mathcal{N}(Y, C\delta/N)$. Finally, (iii) follows from the fact that $T_s$ expands $s$ by a factor $\approx N$.

**Lemma 6.3.** *Fix a large constant $C$. Let $\varepsilon > 0$, $d \geq 3$ and $p, q, r$ be as in Theorem 6.1. Let $Z \subset G_d$ and $R$ be a subset of $\mathbb{R}^d$ such that for any $\gamma_d$-ray $l$, $\mathcal{N}(l, \delta) \cap R$ is contained in a cube of side 1. Let $S$ be a subset of $R$ satisfying:*
*If $x \in S$, then there are two segments $s_1$ and $s_2$ of $\gamma_d$ such that*
*i) $s_1$ and $s_2$ are of length $C^{-1}$,*
*ii) The distance between $s_1$ and $s_2$ is at least $C^{-1}$,*
*iii)* $\min(X^*_\delta(\chi_{Z\cap G_{s_1}}), X^*_\delta(\chi_{Z\cap G_{s_2}})) \geq \eta$.

*Then,*

$$|S| \lesssim \delta^{-\varepsilon}\eta^{-p'}\|\chi_{Z,\delta}\|^{p'}_{q',r'}.$$

**Proof.** First note that it suffices to prove the lemma with $R$ replaced with $Q_1$. To see this, assume that we have proved the lemma for cubes of side 1. Tile $R$ by cubes of side 1, $R = \cup_i Q^i$ say. Let $Z^i$ be the $\delta$ neighborhood of the set $\{l \in Z : \mathcal{N}(l, \delta) \cap Q^i \neq \emptyset\}$. Note that

$$\begin{aligned} |S| &\lesssim \textstyle\sum_i |S \cap Q^i| \lesssim \delta^{-\varepsilon}\eta^{-p'} \sum_i \|\chi_{Z^i}\|^{p'}_{q',r'} \\ &\lesssim \delta^{-\varepsilon}\|\textstyle\sum_i \chi_{Z^i}\|^{p'}_{q',r'} \lesssim \delta^{-\varepsilon}\|\chi_{Z,\delta}\|^{p'}_{q',r'}; \end{aligned}$$

the third inequality follows from the fact that $p' \geq q' \geq r'$, and the last inequality can be obtained by noting that for any $\gamma_d$-ray $l$, $\mathcal{N}(l, \delta)$ intersects $\lesssim 1$ of the cubes $Q^i$.

Also note that $\gamma_d$ can be covered with $\lesssim 1$ segments of length slightly larger than $C^{-1}$ so that any segment of length $C^{-1}$ is contained in one of the segments in the covering. The set $\mathcal{C}$ of pairs of the segments in the covering has $\lesssim 1$ members and for any pair of segments $s_1$ and $s_2$ as in the lemma there is a pair $(c_1, c_2) \in \mathcal{C}$ so that $s_i \subset c_i$, i=1,2. Hence, it suffices to prove the lemma assuming that the segments $s_1$ and $s_2$ are independent of $x$.

Let $Z_i = Z \cap G_{s_i}$, $i = 1, 2$. Let $\mathcal{W}$ (resp. $\mathcal{B}$) be $\delta$-separated subsets of $Z_1$ (resp. $Z_2$). Denote the characteristic function of the $\delta$ neighborhood of $w \in \mathcal{W}$ in $G_d$ by $D_w$ and the characteristic function of the $C\delta$-tube whose axis is $w$ by $\chi_w$. Note that

$X_\delta^* D_w \lesssim \delta\chi_w$. Hence, $X_\delta^* Z_1 \lesssim \delta \sum_w \chi_w = \delta\Phi_{\mathcal{W}}$. Similarly, we have $X_\delta^* Z_2 \lesssim \delta\Phi_{\mathcal{B}}$. Using these and the hypothesis of Theorem 6.1, we obtain

$$\|\min(X_\delta^*\chi_{Z_1}, X_\delta^*\chi_{Z_2})\|_{L^{p'}(Q_1)}^{p'} \lesssim \|\delta\min(\Phi_{\mathcal{W}},\Phi_{\mathcal{B}})\|_{L^{p'}(Q_1)}^{p'} \lesssim \delta^{-\varepsilon}(\delta^d\max(|\mathcal{W}|,|\mathcal{B}|))^{\frac{p'}{r'}}$$
$$\lesssim \delta^{-\varepsilon}\max(|\mathcal{N}(Z_1,\delta)|,|\mathcal{N}(Z_2,\delta)|)^{\frac{p'}{r'}} \lesssim \delta^{-\varepsilon}\|\chi_{Z,\delta}\|_{r',r'}^{p'} \lesssim \delta^{-\varepsilon}\|\chi_{Z,\delta}\|_{q',r'}^{p'};$$

we used the fact that $q' \geq r'$ in the last inequality. This yields the claim of the lemma using Tschebyshev's inequality. $\square$

**Lemma 6.4.** *With the hypothesis of Theorem 6.1, we have*

$$\|X_\delta^*\chi_Y\|_{L^{p'}(Q_1)} \lesssim \delta^{-C\varepsilon}\|\chi_{Y,\delta}\|_{q',r'},$$

*for any $Y \subset G_d$.*

**Proof.** Below, we prove that

$$|\{x \in Q_1 : X_\delta^*\chi_Y(x) \geq \lambda\}| \lesssim \delta^{-C\varepsilon}\lambda^{-p'}\|\chi_{Y,\delta}\|_{q',r'}^{p'}. \tag{48}$$

This yields the claim of the lemma as in the proof of Lemma 5.1. Note that (48) is obvious for $\lambda < \delta^B$, where $B$ is a large enough constant. The reason for this is that the left-hand side is bounded by 1 and the right-hand side is $\gtrsim 1$ if $\lambda$ is small and $Y$ is non-empty. Therefore we assume that $\lambda > \delta^B$.

Now, we prove (48). Fix a sufficiently large constant $C$ which depends on $\varepsilon$ and $B$. Let $A = \{x \in Q_1 : X_\delta^*\chi_Y(x) \geq \lambda\}$ and $A_\sigma$ be the set of all points $x \in Q_1$ such that

i) there are two segments $s_1$ and $s_2$ of length $\sigma$ of $\gamma_d$,
ii) the distance between $s_1$ and $s_2$ is between $\sigma$ and $C\sigma$,
iii) $X_\delta^*(\chi_{Y\cap G_{s_i}}) \geq C^{-1}\delta^\varepsilon\lambda$ for $i = 1, 2$.

We claim that $\cup_\sigma A_\sigma \supseteq A$, where the union is over dyadic $\sigma > \delta^K$, where $K$ is a constant which depends on $B$.

Let $x \in A$. Let $\sigma$ be the smallest number such that $X_\delta^*(\chi_{Y\cap G_s})(x) \geq (C\sigma)^{\frac{\varepsilon}{K}}\lambda$ for some segment $s$ of length $C\sigma$. Note that the lower bound for $\lambda$ implies that $\sigma \geq \delta^K$. Divide $s$ into $\approx C$ segments $s_i$ of length $\sigma$. Since $\sigma$ is minimal, for any segment $s_i$, $X_\delta^*(\chi_{Y\cap G_{s_i}})(x) < \sigma^{\frac{\varepsilon}{K}}\lambda$. On the other hand, $\sum_i X_\delta^*(\chi_{Y\cap G_{s_i}})(x) \geq X_\delta^*(\chi_{Y\cap G_s})(x) \geq (C\sigma)^{\frac{\varepsilon}{K}}\lambda$. Hence, there should be at least 3 segments $s_i$ such that $X_\delta^*(\chi_{Y\cap G_{s_i}}) \geq C^{-1}\sigma^{\frac{\varepsilon}{K}}\lambda$, which proves the claim since $\sigma > \delta^K$.

By pigeonholing, there is a $\sigma$ such that $|A_\sigma| \gtrsim \delta^\varepsilon|A|$. Using the rescaling maps and Lemma 6.3, we find a bound for $|A_\sigma|$, which is independent of $\sigma$.

To do this, consider a covering of $\gamma_d$ with $C\sigma$-segments $s_i$ with bounded overlap. Let $A_\sigma^i$ be the set of points $x \in A_\sigma$ such that the two $\sigma$-segments in the definition of $A_\sigma$ are contained in $s_i$. Note that $A_\sigma = \cup_i A_\sigma^i$.

Fix one of the $s_i$'s. Note that the sets $R = T_{s_i}(Q_1)$, $Z = T_{s_i}(Y \cap G_{s_i})$ and $S = T_{s_i}(A^i_\sigma)$ satisfy the hypothesis of Lemma 6.3 with $\eta = \delta^{-1}\lambda$. $R$ satisfies the hypothesis since $T_{s_i}$ essentially preserves distances in $\gamma_d$ direction. Thus, using Lemma 6.3, we obtain

$$|T_{s_i} A^i_\sigma| \quad \lesssim \quad \delta^{-\varepsilon}(\sigma^{-1}\lambda)^{-p'} \|\chi_{T_{s_i}(Y \cap G_{s_i}),\delta}\|^{p'}_{q',r'} \tag{49}$$

Using property (ii) of the map $T_{s_i}$, we have

$$(49) \lesssim \delta^{-\varepsilon}(\sigma^{-1}\lambda)^{-p'} \sigma^{-p'(\frac{1}{q'}+\frac{d(d-1)}{2r'})} \|\chi_{Y \cap G_{s_i},\sigma\delta}\|^{p'}_{q',r'}.$$

Using (47), we have

$$|A^i_\sigma| \lesssim \delta^{-\varepsilon} \sigma^{\frac{d(d-1)}{2}} (\sigma^{-1}\lambda)^{-p'} \sigma^{-p'(\frac{1}{q'}+\frac{d(d-1)}{2r'})} \|\chi_{Y \cap G_{s_i},\sigma\delta}\|^{p'}_{q',r'}. \tag{50}$$

(3) implies that $\sigma^{\frac{d(d-1)}{2}} \sigma^{p'} \sigma^{-p'(\frac{1}{q'}+\frac{d(d-1)}{2r'})} \lesssim 1$. Using this in (50), we obtain

$$|A^i_\sigma| \lesssim \delta^{-\varepsilon}\lambda^{-p'} \|\chi_{Y \cap G_{s_i},\sigma\delta}\|^{p'}_{q',r'}. \tag{51}$$

Now, note that the sets $\mathcal{N}(Y \cap G_{s_i}, \sigma\delta)$ have bounded overlap. Thus, using (51), we get

$$\begin{aligned} |A_\sigma| \le \sum_{s_i} |A^i_\sigma| \quad &\lesssim \quad \delta^{-\varepsilon}\lambda^{-p'} \sum_{s_i} \|\chi_{Y \cap G_{s_i},\sigma\delta}\|^{p'}_{q',r'} \\ &\lesssim \quad \delta^{-\varepsilon}\lambda^{-p'} \|\chi_{Y,\sigma\delta}\|^{p'}_{q',r'} \lesssim \delta^{-\varepsilon}\lambda^{-p'} \|\chi_{Y,\delta}\|^{p'}_{q',r'}; \end{aligned}$$

the last inequality follows from the observation that $\mathcal{N}(Y, \sigma\delta) \subset \mathcal{N}(Y, \delta)$. $\Box$

**Proof of Theorem 6.1.**

Using duality, Lemma 6.4 implies that

$$\|X_\delta f\|_{L^q(L^r)} \lesssim \delta^{-\varepsilon} \|f\|_{L^p(Q_1)}. \tag{52}$$

Now, we trade $\varepsilon$ derivatives for the $\delta^{-\varepsilon}$ factors. This argument is standard, we follow [19] and omit the details. We can assume that $\|f\|_{W^{p,\varepsilon}}=1$. Using a suitable partition of unity (see, e.g., [19], p.597), one can find functions $f_j$, $j = 1, 2, ...$ with Fourier support in $\{\xi : |\xi| \approx 2^j\}$ such that $\sum_j 2^{\eta j}\|f_j\|_p \lesssim \|f\|_{W^{p,\varepsilon}} = 1$ for small $\eta$ and

$$|Xf| \lesssim 1 + \sum_j |X_{2^{-j}} f_j|. \tag{53}$$

Using (53) and (52) with $\varepsilon = \eta$, we have

$$\|Xf\|_{q,r} \lesssim 1 + \sum_j \|X_{2^{-j}} f_j\|_{q,r} \lesssim 1 + \sum_j 2^{\eta j} \|f_j\|_p \lesssim 1,$$

which is the claim of Theorem 6.1. $\Box$

# References


[1] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, GAFA **1** (1991), 147-187.

[2] J. Bourgain, *Global solutions of nonlinear Schrödinger Equations*, AMS Colloquium Publications, V. 46, AMS, Providence, Rhode Island, 1999.

[3] J. Bourgain, *Harmonic Analysis and combinatorics: How much may they contribute to each other?*, Mathematics: Frontiers and perspectives, IMU/Amer. Math. Society 2000, 13-32.

[4] M. Christ, *Hilbert transforms along curves, I: Nilpotent groups*, Annals of Math. **122** (1985), 575-596.

[5] M. Christ, *Estimates for the k-plane transform*, Indiana Univ. Math. J. **33** (1984), 891-910.

[6] S. Drury, *$L^p$ estimates for the $X$-ray transform*, Illinois J. Math. **27** (1983), 125-129.

[7] C. Fefferman, *The multiplier problem for the ball*, Annals of Math. **94** (1971), 330-336.

[8] A. Greenleaf, A. Seeger, *Fourier integral operators with fold singularities*, J. Reine Angew. Math. **455** (1994), 35-56.

[9] A. Greenleaf, A. Seeger, *Fourier integral operators with cusp singularities*, Amer. Jour. Math. **120** (1998), 1077-1119.

[10] A. Greenleaf, A. Seeger, S. Wainger, *On X-ray transforms for rigid line complexes and integrals over curves in $R^4$*, Proc. Amer. Math. Soc. **127** (1999), no. 12, 3533-3545.

[11] A. Greenleaf, A. Seeger, S. Wainger, *Estimates for generalized Radon transforms in three and four dimensions*, preprint (1999), to appear in Contemporary Mathematics, Proc. of the Ehrenpreis Conf.

[12] I. Laba, T. Tao, *An X-ray transform estimate in $\mathbb{R}^n$*, preprint (1999), to appear in Revista Math. Iberoamericana.

[13] D. Oberlin, *An estimate for a restricted $X$-ray transform*, preprint (1997).

[14] T. Tao, *From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE*, preprint (2000), (arXiv:math.CA/0008098 Aug. 2000).

[15] T. Tao, A. Vargas, *A bilinear approach to cone multipliers I and II*, GAFA **10** (2000), 185-258.

[16] T.Tao, A. Vargas, L. Vega, *A bilinear approach to the restriction and Kakeya conjectures*, J. Amer. Math. Soc. **11** (1998), 967-1000.

[17] H.T. Wang, *$L^p$ estimates for the $X$-ray transform restricted to line complexes of Kirillov type*, Trans. Amer. Math. Soc. **332** (1992), 793-821.

[18] T. Wolff, *An improved bound for Kakeya type maximal functions*, Revista Math. Iberoamericana **11** (1995), 651-674.

[19] T. Wolff, *A mixed norm estimate for the X-ray transform*, Revista Math. Iberoamericana **14** (1998), 561-601.

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

[21] T. Wolff, *A sharp bilinear cone restriction estimate*, preprint (1999) (to appear in Annals of Math.).