



# Fourier restriction for smooth hyperbolic 2-surfaces

Stefan Buschenhenke<sup>1</sup> · Detlef Müller<sup>1</sup> · Ana Vargas<sup>2</sup>

Received: 8 October 2021 / Revised: 4 July 2022 / Accepted: 13 July 2022 /  
Published online: 27 August 2022  
© The Author(s) 2022

## Abstract

We prove Fourier restriction estimates by means of the polynomial partitioning method for compact subsets of any sufficiently smooth hyperbolic hypersurface in  $\mathbb{R}^3$ . Our approach exploits in a crucial way the underlying hyperbolic geometry, which leads to a novel notion of strong transversality and corresponding “exceptional” sets. For the division of these exceptional sets we make crucial and perhaps surprising use of a lemma on level sets for sufficiently smooth one-variate functions from a previous article of ours.

**Mathematics Subject Classification.** 42B25

## Contents

1 Introduction	18
2 Auxiliary results	23
3 Geometric background on strong transversality	26
3.1 Strong transversality for bilinear estimates	26
3.1.1 Null vectors for $D^2\phi$	27

Stefan Buschenhenke and Detlef Müller were partially supported by the DFG Grant MU 761/ 11-2. Ana Vargas was partially supported by Grants PID2019-105599GB-I00 / AEI / 10.13039/501100011033 and MTM2016-76566-P (Ministerio de Ciencia, Innovación y Universidades), Spain.

✉ Stefan Buschenhenke  
buschenhenke@math.uni-kiel.de  
<http://www.math.uni-kiel.de/analysis/de/buschenhenke>

Detlef Müller  
mueller@math.uni-kiel.de  
<http://www.math.uni-kiel.de/analysis/de/mueller>

Ana Vargas  
ana.vargas@uam.es  
<http://matematicas.uam.es/AFA/>

<sup>1</sup> Mathematisches Seminar, C.A.-Universität Kiel, Heinrich-Hecht-Platz 6, 24118 Kiel, Germany

<sup>2</sup> Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid, Spain

3.1.2	Back to $\Gamma_z(z_1, z_2, z'_1, z'_2)$	29
3.1.3	Caps and the basic decomposition of $S$	30
3.2	Not strongly separated caps	31
3.2.1	Level curves of $t_{z_1}^2(z_1, \cdot)$	32
3.2.2	On the direction of level curves of $t_{z_1}^2(z_1, \cdot)$ within $R_I \cap R_{II}$	32
3.3	Moving rectangular boxes into vertical position at the origin	33
3.4	The induction on scales step	35
3.4.1	Final rescaling step in the induction on scales argument	36
4	Broad points	38
4.1	Definition of broad points and the underlying family of rectangles	38
4.2	The geometric lemma	41
5	Reduction to estimates for the broad part	44
6	Proof of Theorem 1.2	45
6.1	Completing the proof of Theorem 6.1	54
	References	55

## 1 Introduction

Let  $S \subset \mathbb{R}^n$  be a sufficiently smooth hypersurface. The Fourier restriction problem, introduced by E. M. Stein in the seventies (for general submanifolds), asks for the range of exponents  $\tilde{p}$  and  $\tilde{q}$  for which an a priori estimate of the form

$$\left( \int_S |\widehat{f}|^{\tilde{q}} d\sigma \right)^{1/\tilde{q}} \leq C \|f\|_{L^{\tilde{p}}(\mathbb{R}^n)}$$

holds true for every Schwartz function  $f \in \mathcal{S}(\mathbb{R}^n)$ , with a constant  $C$  independent of  $f$ . Here,  $d\sigma$  denotes the Riemannian surface measure on  $S$ .

The sharp range in dimension  $n = 2$  for curves with non-vanishing curvature was determined through work by Fefferman et al. [15, 42]. In higher dimension, the sharp  $L^{\tilde{p}} - L^2$  result for hypersurfaces with non-vanishing Gaussian curvature was obtained by Stein and Tomas [30, 38] (see also Strichartz [32]). Some more general classes of surfaces were treated by Greenleaf [16]. In work by Ikromov et al. [21] and Ikromov and Müller [22, 23], the sharp range of Stein–Tomas type  $L^{\tilde{p}} - L^2$  restriction estimates has been determined for a large class of smooth, finite-type hypersurfaces, including all analytic hypersurfaces.

The question about general  $L^{\tilde{p}} - L^{\tilde{q}}$  restriction estimates is nevertheless still wide open. Fourier restriction to hypersurfaces with non-negative principal curvatures has been studied intensively by many authors. Major progress was due to Bourgain in the nineties [3–5]. At the end of that decade the bilinear method was introduced [26–28, 33, 35–37, 41]. A new impulse to the problem has been given with the multilinear method [2, 6]. The best results up to date have been obtained with the polynomial partitioning method, developed by Guth [18, 19] (see also [20, 40] for recent improvements).

For the case of hypersurfaces of non-vanishing Gaussian curvature but principal curvatures of different signs, besides Tomas–Stein type Fourier restriction estimates, until a few years ago the only case which had been studied successfully was the case of the hyperbolic paraboloid (or “saddle”) in  $\mathbb{R}^3$ : in 2005, independently Lee [25] and Vargas [39] established results analogous to Tao’s theorem [33] on elliptic surfaces

(such as the 2-sphere), with the exception of the end-point, by means of the bilinear method.

First results based on the bilinear approach for particular one-variate perturbations of the saddle were eventually proved by the authors in [10–12]. Furthermore, Stovall [31] was able to include also the end-point case for the hyperbolic paraboloid. Building on the papers [25, 31, 39], and by strongly making use of Lorentzian symmetries, even global restriction estimates for one-sheeted hyperboloids have been established recently by Bruce et al. [9], with extensions to higher dimensions by Bruce [8]. Results on higher dimensional hyperbolic paraboloids have been reported by Barron [1]. All these results are in the bilinear range given by [33].

Improvements over the results for the saddle by means of an adaptation of the polynomial partitioning method from Guth’s articles [18] were achieved by Cho and Lee [14], and Kim [24]. Moreover, for a particular class of one-variate perturbations of the hyperbolic paraboloid, an analogue of Guth’s result had been proved by the authors in [13], and more lately by making use of Lorentzian symmetries, Bruce [7] has established analogous results for compact subsets of the one-sheeted hyperboloid.

In this article, we shall obtain the analogous result to [18] for compact subsets of any sufficiently smooth hyperbolic surface.

More precisely, we shall study embedded  $C^m$ -hypersurfaces  $S$  in  $\mathbb{R}^3$  of sufficiently high degree of regularity  $m \geq 3$  which are *hyperbolic* in the sense that the Gaussian curvature is strictly negative at every point, i.e., that at every point of  $S$  one principal curvature is strictly positive, and the other one is strictly negative.

A result comparable to the one of the authors was reported by Guo and Oh [17], though the initial approach is different and is based on approximation of arbitrary compact hypersurfaces with negative curvature in  $\mathbb{R}^3$  by polynomial surfaces.

As usual, it will be more convenient to use duality and work in the adjoint setting. If  $\mathcal{R}$  denotes the Fourier restriction operator  $g \mapsto \mathcal{R}g := \hat{g}|_S$  to the surface  $S$ , its adjoint operator  $\mathcal{R}^*$  is given by  $\mathcal{R}^*f(\xi) = \mathcal{E}f(-\xi)$ , where  $\mathcal{E} = \mathcal{E}_S$  denotes the “Fourier extension” operator given by

$$\mathcal{E}f(\xi) := \widehat{f d\sigma}(\xi) = \int_S f(x)e^{-i\xi \cdot x} d\sigma(x),$$

with  $f \in L^q(S, \sigma)$ . The restriction problem is therefore equivalent to the question of finding the appropriate range of exponents for which the estimate

$$\|\mathcal{E}f\|_{L^p(\mathbb{R}^3)} \leq C\|f\|_{L^q(S, d\sigma)}$$

holds true with a constant  $C$  independent of the function  $f \in L^q(S, d\sigma)$ . We shall here concentrate on local estimates of this form, where  $S$  is replaced by a sufficiently small neighborhood of a given point on  $S$ . Such estimates then allow for estimates of the form

$$\|\mathcal{E}_{S_c}f\|_{L^p(\mathbb{R}^3)} \leq C_{S_c, p, q}\|f\|_{L^q(S, d\sigma)}, \tag{1.1}$$

for any compact subset  $S_c$  of  $S$ , where we have put

$$\mathcal{E}_{S_c} f(\xi) := \int_{S_c} f(x) e^{-i\xi \cdot x} d\sigma(x).$$

Our main result will be

**Theorem 1.1** *Assume that  $p > 3.25$  and  $p > 2q'$ . Then there is some sufficiently large  $M(p, q) \in \mathbb{N}$  such that for any embedded hyperbolic hypersurface  $S \subset \mathbb{R}^3$  of class  $C^{M(p,q)}$  the estimate (1.1) holds true, i.e., for any compact subset  $S_c$  of  $S$ , we have*

$$\|\mathcal{E}_{S_c} f\|_{L^p(\mathbb{R}^3)} \leq C_{S_c,p,q} \|f\|_{L^q(S,d\sigma)}.$$

For the proof of this result, we shall consider the following classes of functions: Let  $\Sigma := [-1, 1] \times [-1, 1]$ . For  $M \in \mathbb{N}$ ,  $M \geq 3$ , we denote by  $\text{Hyp}^M = \text{Hyp}^M(\Sigma) \subset C^M(\Sigma)$  the set of all functions  $\phi$  on  $\Sigma$  satisfying the following properties:

$\phi$  extends from  $\Sigma$  to a  $C^M$ -function on  $2\Sigma$ , also denoted by  $\phi$ , which satisfies the following conditions (1.2), (1.3) on  $2\Sigma$  :

$$\phi(0) = 0, \quad \nabla\phi(0) = 0, \quad D^2\phi(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{1.2}$$

and

$$\|\partial_x^\alpha \partial_y^\beta \phi\|_\infty \leq 10^{-5} \quad \text{for } 3 \leq \alpha + \beta \leq M. \tag{1.3}$$

**Note:** (i)  $\phi \in C^M(\Sigma)$  lies in  $\text{Hyp}^M$  if and only if  $\phi(x, y)$  is a small perturbation of  $xy$ , in the sense that  $\phi(x, y) = xy + \psi(x, y)$ , where  $\psi(0) = 0$ ,  $\nabla\psi(0) = 0$ , and  $D^2\psi(0) = 0$ , and  $\psi$  satisfies the estimates (1.3) (even on  $2\Sigma$ ), with  $\phi$  replaced by  $\psi$ .

(ii) If  $\phi \in \text{Hyp}^M$ , then

$$\max \{ \|\phi_{xx}\|_\infty, \|\phi_{yy}\|_\infty, \|\phi_{xy} - 1\|_\infty \} \leq 2 \cdot 10^{-5}. \tag{1.4}$$

Our key result then is

**Theorem 1.2** *Assume that  $p > 3.25$  and  $p > 2q'$ . Then there is some sufficiently large  $M(p, q) \in \mathbb{N}$  such that for any  $\phi \in \text{Hyp}^{M(p,q)}$  the Fourier extensions operator*

$$\mathcal{E}_\phi f(\xi) := \int_\Sigma f(x, y) e^{-i(\xi_1 x + \xi_2 y + \xi_3 \phi(x,y))} dx dy$$

associated to the graph  $S_\phi$  of  $\phi$  satisfies the estimate

$$\|\mathcal{E}_\phi f\|_{L^p(\mathbb{R}^3)} \leq C_{p,q} \|f\|_{L^q(\Sigma)} \quad \text{for every } f \in \mathcal{S}(\mathbb{R}^2), \tag{1.5}$$

with a constant which is independent of  $\phi$  and  $f$ .

Note that Theorem 1.1 follows easily from Theorem 1.2. Indeed, if  $S$  is an embedded hyperbolic hypersurface  $S \subset \mathbb{R}^3$  of class  $C^{M(p,q)}$ , with  $M(p, q)$  as in Theorem 1.2, and if  $S_c$  is a compact subset, then by compactness we can localized to sufficiently small neighborhoods of any points  $x^0$  in  $S_c$ . So, after permuting coordinates, we may assume that near such a point  $S$  is given as the graph of a  $C^{M(p,q)}$  function  $\phi$ , and after translation and linear change of coordinates, that  $x^0 = 0$  and that  $S$  is the graph of  $\phi$  over some sufficiently small neighborhood  $U$  of the origin in  $\mathbb{R}^2$ , where  $\phi$  satisfies (1.2). Finally, after applying a suitable isotropic scaling by putting  $\tilde{\phi}(z) := \frac{1}{r^2}\phi(rz)$ , where  $0 < r \ll 1$ , assuming that  $U$  was sufficiently small, we see that we can reduce to a function  $\tilde{\phi}$  in  $\text{Hyp}^{M(p,q)}(\Sigma)$ .

Denote by  $B_R$  the cube  $B_R := [-R, R]^3$ ,  $R \geq 0$ . In a similar way as in [13], Theorem 1.2 will be a consequence of the following local Fourier extension estimate:

**Theorem 1.3.** *Assume that  $3.25 \geq q > 2.6$ . Then, for every  $\varepsilon > 0$ , there is a sufficiently large  $M(\varepsilon) \in \mathbb{N}$  such that for any  $\phi \in \text{Hyp}^{M(\varepsilon)}$  the following holds true: there is a constant  $C_\varepsilon$  such that for any  $R \geq 1$*

$$\|\mathcal{E}_\phi f\|_{L^{3.25}(B_R)} \leq C_\varepsilon R^\varepsilon \|f\|_{L^2(\Sigma)}^{2/q} \|f\|_{L^\infty(\Sigma)}^{1-2/q}, \tag{1.6}$$

for all  $f \in L^\infty(\Sigma)$ .

Indeed, a simple interpolation argument as in [13] shows that the estimate in Theorem 1.3 implies the following:

If  $p > 3.25$  and  $p > 2q'$ , then

$$\|\mathcal{E}_\phi f\|_{L^p(B_R)} \leq C_{p,q,\varepsilon} R^\varepsilon \|f\|_{L^q(\Sigma)}, \tag{1.7}$$

with a constant which is independent of  $\phi$  and  $f$ .

Finally, as usual, we can invoke an  $\varepsilon$ -removal theorem to pass to Theorem 1.2, but we have to be a bit more precise here than usually:

From [24, Theorem 5.3] (which, as Kim observes, is an immediate extension of Tao’s  $\varepsilon$ -removal theorem [34, Theorem 1.2]), applied to the adjoints of the restriction operators, we see that the estimate in (1.7) implies the following: there is a constant  $C > 0$  such that if

$$\frac{1}{p} > \frac{1}{\tilde{p}} + \frac{C}{-\log \varepsilon}, \tag{1.8}$$

then

$$\|\mathcal{E}_\phi f\|_{L^{\tilde{p}}(\mathbb{R}^3)} \leq C_{\tilde{p},q} \|f\|_{L^q(\Sigma)}. \tag{1.9}$$

Thus, given  $\tilde{p}$  with  $\tilde{p} > 3.25$  and  $\tilde{p} > 2q'$ , we can first choose an appropriate  $p$  such that  $\tilde{p} > p > 3.25$  and  $\tilde{p} > p > 2q'$ , and then an appropriate  $\varepsilon = \varepsilon(\tilde{p}, q) > 0$  so that (1.8) holds true, hence also (1.9). This proves Theorem 1.2.

Later in our proofs we shall always assume without further mentioning that  $\varepsilon$  is sufficiently small.

OUTLINE OF THE PAPER: Since Guth’s polynomial partitioning method has been discussed in various papers by now, we shall be brief in many parts and refer to Guth [18] and our previous paper [13] whenever possible. Instead, we will focus on the novelties of our approach, so that some familiarity with the polynomial partitioning method is recommended.

A crucial property in restriction estimates is “strong” transversality (which we shall define precisely at a later stage). A major task in our paper will then be to understand the “exceptional” sets which lack this kind of transversality. In case of the unperturbed hyperbolic paraboloid, i.e., the graph of  $xy$ , any two small caps that are not strongly transversal have to be contained in the part of the surface lying over a vertical or horizontal strip in the  $(x, y)$ -plane. For our general hyperbolic surfaces, the “exceptional” sets are again certain rectangles, but their size and slope strongly depend on the geometry of the surface. Rescaling arguments for functions restricted to a rectangle are here a priori difficult, since, unlike for the unperturbed hyperbolic paraboloid, our class of phase functions  $\text{Hyp}^m$  is not closed under anisotropic scalings, nor any other suitably large group of symmetries, such as the Lorentz group in case of the one-sheeted hyperboloid. However, we will show that the lack of strong transversality eventually still does allow for a rescaling argument.

The article is organized as follows: in Sect. 2 we present two important auxiliary results: on the one hand our “sublevel lemma” (Corollary 2.2) that will allow us to control certain sub-level sets, on the other hand Lemma 2.4 which will allow to control certain derivatives in the rescaling argument.

In Sect. 3, we relate the hyperbolic geometry of our surfaces to the strong transversality property which is needed to establish the required bilinear estimates. In particular, we shall derive a “hyperbolic factorization” of the crucial transversality function  $\Gamma_z(z_1, z_2)$  in (3.14), which will be of central importance to our approach and will lead to our notation of “strong separation” of caps. Our definition of “exceptional” sets will be based on this notion. In Sects. 3.3 and 3.4 of Sect. 3 we shall show how to move such exceptional rectangles into vertical position and prepare for the subsequent rescaling argument.

Motivated by these “exceptional” sets, we will devise a notion of  $\alpha$ -broadness adapted to our class of surfaces in Sect. 4, and prove our crucial “Geometric Lemma”, and we shall show that any two small caps that are not strongly separated have to be contained in the part of the surface lying over some rectangle (of possibly quite arbitrary direction) or be contained in a larger cap, thus ensuring that a family of pairwise not strongly separated caps is in some sense sparse.

In Sect. 5, we reduce our main result to estimates for the broad part of the extension operator, and in Sect. 6 we outline the actual polynomial partitioning argument and indicate which changes will be required compared to previous work, in particular to [13, 18]. Here, in particular, we will be brief and only highlight the steps that differ from previous work.

CONVENTION: Unless stated otherwise,  $C > 0$  will stand for an absolute constant whose value may vary from occurrence to occurrence. We will use the notation  $A \sim_C B$  to express that  $\frac{1}{C}A \leq B \leq CA$ . In some contexts where the size of  $C$  is irrelevant we shall drop the index  $C$  and simply write  $A \sim B$ . Similarly,  $A \lesssim B$  will express the fact that there is a constant  $C$  (which does not depend on the relevant quantities in the

estimate) such that  $A \leq CB$ , and we write  $A \ll B$ , if the constant  $C$  is sufficiently small.

## 2 Auxiliary results

In this section, we prove two auxiliary results which will be crucial for our analysis, but which may also be of independent interest. We begin by recalling the following theorem from our previous paper [12]:

**Theorem 2.1** *Let  $I = [a, b]$  be a compact interval and  $g \in C^r(I, \mathbb{R})$ ,  $r \geq 1$ , and put  $C_r := \|g^{(r)}\|_{L^\infty(I)}$ . Then there exists a decomposition of  $\{g \neq 0\}$  into pairwise disjoint intervals  $J_{\lambda, \iota}$ , where  $\lambda$  ranges over the set of all positive dyadic numbers  $\lambda \leq \|g\|_\infty$ , and where for any given  $\lambda$ , the index  $\iota$  is from some index set  $\mathcal{J}_\lambda$ , such that the following hold true:*

- (i)  $|\mathcal{J}_\lambda| \leq 10r(1 + |I|C_r^{1/r}\lambda^{-1/r}) \lesssim 1 + \lambda^{-1/r}$ .
- (ii) For any  $\lambda$  dyadic,  $\iota \in \mathcal{J}_\lambda$  and any  $t \in J_{\lambda, \iota}$  we have  $\frac{1}{2}\lambda < |g(t)| < 4\lambda$ .

The following corollary will later allow us to control certain sub-level sets which will be important for our definition of broad points.

**Corollary 2.2** *Let  $I = [a, b]$  be a compact interval and  $g \in C^r(I, \mathbb{R})$ ,  $r \geq 1$ , and put  $C_r := \|g^{(r)}\|_{L^\infty(I)}$ . Then for any  $0 < \lambda \leq \|g\|_\infty$ , there is a finite family of pairwise disjoint intervals  $I_{\lambda, i}$ , where the index  $i$  is from some index set  $\mathcal{I}_\lambda$ , so that the following hold true:*

- (i)  $|\mathcal{I}_\lambda| \leq 30r(1 + |I|C_r^{1/r}\lambda^{-1/r}) \lesssim 1 + \lambda^{-1/r}$ .
- (ii) If we denote by  $V_\lambda$  the union  $\bigcup_{i \in \mathcal{I}_\lambda} I_{\lambda, i}$  of all the intervals  $I_{\lambda, i}$ , then

$$\{|g| < \lambda\} \subset V_\lambda \subset \{|g| < 8\lambda\}.$$

**Proof** It is sufficient to prove this for any dyadic number  $\lambda$ , however, with the slightly stronger estimates

$$\{|g| < \lambda\} \subset V_\lambda \subset \{|g| < 4\lambda\}. \tag{2.1}$$

We shall consider  $I$  to be endowed with its relative topology from  $\mathbb{R}$ . Consider then the open subset  $U_\lambda := \{|g| < \lambda\}$  of  $I$ . We decompose it into its connected components  $U_{\lambda, \nu}$ , where  $\nu$  will be from an at most countable index set. The  $U_{\lambda, \nu}$  are then open subintervals of  $I$ . Observe that if such an interval  $U_{\lambda, \nu}$  has endpoints  $\alpha < \beta$ , then  $|g(\alpha)| = \lambda$  if  $\alpha > a$ , and  $|g(\beta)| = \lambda$  if  $\beta < b$ . Moreover, if  $\alpha = a$  and  $\beta = b$ , the case where  $|g(\alpha)| < \lambda$  and  $|g(\beta)| < \lambda$  cannot arise, since then we would have  $U_{\lambda, \nu} = I$  and  $\|g\|_\infty < \lambda$ , contradicting our assumptions. Thus we see that every  $U_{\lambda, \nu}$  will have at least one endpoint, say  $t_{\lambda, \nu}$ , such that  $|g(t_{\lambda, \nu})| = \lambda$ .

Next, according to Theorem 2.1, the point  $t_{\lambda, \nu}$  must be contained in either one of the intervals  $J_{\lambda, \iota}$ , or in one of the intervals  $J_{\lambda/2, \iota}$ . This interval is unique, and we

denote it by  $J_\lambda^v$ . Observe that then also  $U_{\lambda,v} \cup J_\lambda^v$  is an interval, and that  $|g| < 4\lambda$  on  $U_{\lambda,v} \cup J_\lambda^v$ .

Let us finally put

$$V_\lambda := \bigcup_v U_{\lambda,v} \cup J_\lambda^v.$$

Then clearly  $\{|g| < \lambda\} = U_\lambda \subset V_\lambda$ , and  $|g| < 4\lambda$  on  $V_\lambda$ . Moreover, if we decompose  $V_\lambda = \bigcup_{i \in \mathcal{I}_\lambda} I_{\lambda,i}$  into its connected components  $I_{\lambda,i}$ , then each  $I_{\lambda,i}$  is an interval. And, if  $t$  is any point in  $I_{\lambda,i}$ , then there is some  $v_i$  so that  $t \in U_{\lambda,v_i} \cup J_\lambda^{v_i}$ , so that  $U_{\lambda,v_i} \cup J_\lambda^{v_i} \subset I_{\lambda,i}$ . The mapping  $I_{\lambda,i} \mapsto J_\lambda^{v_i}$  is clearly injective, since the intervals  $I_{\lambda,i}$  are pairwise disjoint, so that  $|\mathcal{I}_\lambda| \leq |\mathcal{J}_{\lambda/2}| + |\mathcal{J}_\lambda|$ . The estimate in (i) follows thus from estimate (i) in Theorem 2.1.  $\square$

The following remark shows that we may even assume that the intervals from Corollary 2.2 are not too short. We shall not directly make use of this remark, but the same idea will be used later in Sect. 4 to show that we may choose the rectangles in our definition of  $\alpha$ -broadness sufficiently long.

**Remark 2.3** If  $\|g'\|_\infty \leq 1$ , we may even assume that all the intervals  $I_{\lambda,i}$  in Corollary 2.2 have length  $|I_{\lambda,i}| \geq C\lambda$ .

More precisely, assume that  $g \in C^r(I, \mathbb{R})$  satisfies the assumptions of Corollary 2.2, that  $C > 0$  is given, and that in addition  $\|g'\|_\infty \leq 1$ . Then for any  $0 < \lambda \leq \|g\|_\infty$  such that  $C\lambda \leq |I|$ , there is a finite family of pairwise disjoint open intervals  $I_{\lambda,i}$  of length  $|I_{\lambda,i}| \geq C\lambda$ , where the index  $i$  is from some index set  $\mathcal{I}_\lambda$ , so that the following hold true:

- (i)  $|\mathcal{I}_\lambda| \leq 30r(1 + |I|C_r^{1/r}\lambda^{-1/r}) \lesssim 1 + \lambda^{-1/r}$ .
- (ii) If we denote by  $V_\lambda$  the union  $\bigcup_{i \in \mathcal{I}_\lambda} I_{\lambda,i}$  of all the intervals  $I_{\lambda,i}$ , then

$$\{|g| < \lambda\} \subset V_\lambda \subset \{|g| < (8 + C)\lambda\}.$$

**Proof** Let us denote for  $\delta > 0$  by  $A^\delta := (A + (-\delta, \delta)) \cap I$  the  $\delta$ -thickening within the interval  $I$  of any subset  $A$  of  $I$ . For this proof, let us endow the quantities appearing in Corollary 2.2 with a superscript  $\tilde{\cdot}$ , so that for instance  $\tilde{I}_{\lambda,\tilde{i}}, \tilde{i} \in \tilde{\mathcal{I}}_\lambda$  denote the intervals devised in this corollary. Then clearly, with  $\delta := C\lambda$ ,

$$\{|g| < \lambda\} \subset \{|g| < \lambda\}^\delta \subset V_\lambda^\delta = \bigcup_{\tilde{i} \in \tilde{\mathcal{I}}_\lambda} (\tilde{I}_{\lambda,\tilde{i}})^\delta \subset \{|g| < 8\lambda\}^\delta \subset \{|g| < (8 + C)\lambda\}.$$

Let us again decompose the set  $V_\lambda^\delta = \bigcup_{i \in \mathcal{I}_\lambda} I_{\lambda,i}$  into its connected components  $I_{\lambda,i}$ . Since  $V_\lambda^\delta$  is open, the  $I_{\lambda,i}$  are open intervals. Moreover, clearly any  $I_{\lambda,i}$  must contain at least one of the intervals  $(\tilde{I}_{\lambda,\tilde{i}})^\delta$ , so that its length must at least be  $\delta = C\lambda$ , and since the mapping from  $i$  to the chosen index  $\tilde{i}$  is injective, we see that  $|\mathcal{I}_\lambda| \leq |\tilde{\mathcal{I}}_\lambda|$ .  $\square$

The next lemma will become important for certain induction on scales arguments.



**Lemma 2.4** *Let  $g \in C^k(I, \mathbb{R})$ , where  $I$  is an interval of length  $|I| = b$ , and  $k \in \mathbb{N}$ , and assume that*

$$\|g^{(m)}\|_\infty \leq c_m, \quad m = 0, \dots, k, \quad \text{where } c_0 < 1. \tag{2.2}$$

*Let  $\varepsilon > 0$ , and assume that  $k \geq 1/\varepsilon$ , and that  $b \leq (c_0)^\varepsilon$ . Then there are constants  $\tilde{C}_m(\varepsilon) \geq 0$  depending only on  $\varepsilon$  and the constants  $c_1, \dots, c_k$  (increasing with the values of the  $c_j$ ), but not on  $b$  and  $c_0$ , so that*

$$\|g^{(m)}\|_\infty \leq c_0 \tilde{C}_m(\varepsilon) b^{-m}, \quad m = 0, \dots, k. \tag{2.3}$$

**Remark 2.5** In our later application, we shall have  $c_1, \dots, c_k \sim 1$  and  $c_0 \ll 1$ , so that for  $m \geq 1$  and  $c_0$  sufficiently small the estimates in (2.3) are stronger than the a priori estimates from (2.2), at least when  $m$  is not too large.

**Proof** We scale by setting  $h(x) := g(bx)$ . Then, after translation, we may assume that  $I = [0, 1]$ , and that

$$\|h^{(j)}\|_\infty \leq c_j b^j, \quad j = 0, \dots, k, \tag{2.4}$$

and what we need to show is that there are constants  $\tilde{C}_m(\varepsilon) \geq 0$  as above such that

$$\|h^{(m)}\|_\infty \leq c_0 \tilde{C}_m(\varepsilon), \quad m = 0, \dots, k. \tag{2.5}$$

To this end, choose  $M = M(\varepsilon) \in \mathbb{N}$  minimal so that  $M \geq 1/\varepsilon$ . Then  $M \leq k$ . Assume  $m \leq k$ .

a) If  $m \geq M$ , then

$$c_m b^m \leq c_m b^M \leq c_m c_0^{M\varepsilon} \leq c_m c_0,$$

so we may choose  $\tilde{C}_m(\varepsilon) := C_m$ . Notice that in particular  $\tilde{C}_M(\varepsilon) := c_{M(\varepsilon)}$ .

b) Assume next that  $0 \leq m \leq M = M(\varepsilon)$ . We know already that

$$\|h\|_\infty \leq c_0, \tag{2.6}$$

$$\|h^{(M)}\|_\infty \leq c_0 \tilde{C}_M(\varepsilon). \tag{2.7}$$

The estimates in (2.5) for  $0 < m < M$  then follow from these two estimates by interpolation. Let us give an elementary argument for this claim. Fix  $m$  with  $0 \leq m \leq M - 1$ .

We first claim that (2.6) implies that for any  $j = 0, \dots, m$  there are points

$$t_0^j < t_1^j < \dots < t_{2^{m-j}-1}^j$$

in  $[0, 1]$  such that  $t_{i+1}^j - t_i^j \geq 2^{-m}$  and

$$|h^{(j)}(t_i^j)| \leq c_0 2^{(m+1)j} \tag{2.8}$$

for every  $i$ . This is easily proved by induction on  $j$ .

If  $j = 0$ , we may choose  $t_i^0 := i2^{-m}$ ,  $i = 0, \dots, 2^m - 1$ . And, assuming that the claim holds for  $j$ , by the induction hypothesis and the mean value theorem we can find points  $t_i^{j+1} \in (t_{2i}^j, t_{2i+1}^j)$  so that

$$|h^{(j+1)}(t_i^{j+1})| = \left| \frac{h^{(j)}(t_{2i+1}^j) - h^{(j)}(t_{2i}^j)}{t_{2i+1}^j - t_{2i}^j} \right| \leq \frac{2 \cdot c_0 2^{(m+1)j}}{2^{-m}} = c_0 2^{(m+1)(j+1)}.$$

Moreover, since

$$t_i^{j+1} < t_{2i+1}^j < t_{2i+2}^j < t_{i+1}^{j+1},$$

where  $t_{2i+2}^j - t_{2i+1}^j \geq 2^{-m}$ , we also have that  $t_{i+1}^{j+1} - t_i^{j+1} \geq 2^{-m}$ .

In particular, for  $j = m$ , we find a  $t^m := t_0^m$  so that  $|h^{(m)}(t^m)| \leq c_0 2^{(m+1)m}$ . Then, for any  $t \in [0, 1]$ , we have

$$|h^{(m)}(t)| \leq |h^{(m)}(t^m)| + |h^{(m)}(t) - h^{(m)}(t^m)| \leq c_0 2^{(m+1)m} + \|h^{(m+1)}\|_\infty \cdot 1,$$

so that

$$\|h^{(m)}\|_\infty \leq c_0 2^{(m+1)m} + \|h^{(m+1)}\|_\infty. \tag{2.9}$$

Now we can use ‘‘downward induction’’ on  $m$ , starting with  $m = M - 1$ , to prove (2.5). Indeed, when  $m = M - 1$ , then by (2.9) we have

$$\|h^{(M-1)}\|_\infty \leq c_0 2^{M(M-1)} + \|h^{(M)}\|_\infty \leq c_0 (2^{M(M-1)} + \tilde{C}_M(\varepsilon)) =: c_0 \tilde{C}_{M-1}(\varepsilon).$$

Finally, we can pass from  $m$  to  $m - 1$  by means of our induction hypothesis on  $m$  and (2.9):

$$\|h^{(m-1)}\|_\infty \leq c_0 2^{m(m-1)} + \|h^{(m)}\|_\infty \leq c_0 2^{m(m-1)} + c_0 \tilde{C}_m(\varepsilon) =: c_0 \tilde{C}_{m-1}(\varepsilon).$$

□

### 3 Geometric background on strong transversality

Assume that  $\phi \in \text{Hyp}^M(\Sigma)$ ,  $M \geq 3$ , and recall that  $S$  is the graph of  $\phi$ .

#### 3.1 Strong transversality for bilinear estimates

We begin by recalling some facts about what kind of ‘‘strong transversality’’ is required for establishing suitable bilinear estimates.

Following [25], given two open subsets  $U_1, U_2 \subset \Sigma$ , we consider the quantity

$$\Gamma_z(z_1, z_2, z'_1, z'_2) := \left\langle (D^2\phi(z))^{-1}(\nabla\phi(z_2) - \nabla\phi(z_1)), \nabla\phi(z'_2) - \nabla\phi(z'_1) \right\rangle \tag{3.1}$$

for  $z_i = (x_i, y_i), z'_i = (x'_i, y'_i) \in U_i, i = 1, 2$ , and  $z = (x, y) \in U_1 \cup U_2$ . Bilinear estimates have constants depending only on upper bounds for the derivatives of  $\phi$  and on lower bounds of (the modulus of) (3.1). As in [13], for our estimates it will be enough to have lower bounds only for  $z \in U_2$  (or only for  $z \in U_1$ ). If  $U_1$  and  $U_2$  are sufficiently small (with sizes depending on upper bounds of the first and second order derivatives of  $\phi$  and a lower bound for the Hessian determinant of  $\phi$ ) this condition reduces to the estimate

$$|\Gamma_z(z_1, z_2)| \geq c > 0, \tag{3.2}$$

for  $z_i = (x_i, y_i) \in U_i, i = 1, 2, z = (x, y) \in U_2$ , where

$$\Gamma_z(z_1, z_2) := \left\langle (D^2\phi(z))^{-1}(\nabla\phi(z_2) - \nabla\phi(z_1)), \nabla\phi(z_2) - \nabla\phi(z_1) \right\rangle. \tag{3.3}$$

In contrast to [10–12], where we had to devise quite specific “admissible pairs” of sets  $U_1, U_2$  for our bilinear estimates, as in [13] we shall here only have to consider “caps” (cf. Sect. 3.1.3)  $\tau_1, \tau_2$  for  $U_1, U_2$ , and the required bilinear estimates will be a of somewhat different nature. Nevertheless, the geometric transversality conditions that we need here will be the same.

We shall next exploit the hyperbolicity assumption on  $S$  in order to gain a better understanding of  $\Gamma_z(z_1, z_2, z'_1, z'_2)$  for such surfaces. In particular, we shall derive the “hyperbolic factorization” (3.14) which will be of central importance.

### 3.1.1 Null vectors for $D^2\phi$

Recall that  $\phi \in \text{Hyp}^M(\Sigma), M \geq 3$ . We put  $H := \phi_{xy}^2 - \phi_{xx}\phi_{yy}$ , so that  $-H$  is the Hessian determinant of  $\phi$ . From (1.4) we easily deduce that  $|H(z) - 1| \leq 10^{-4}$  for every  $z \in \Sigma$ .

It is easy to check that we then explicitly have

$$\begin{aligned} -H(z)\Gamma_z(z_1, z_2) &= \phi_{yy}(z)(\phi_x(z_2) - \phi_x(z_1))^2 + \phi_{xx}(z)(\phi_y(z_2) - \phi_y(z_1))^2 \\ &\quad - 2\phi_{xy}(z)(\phi_x(z_2) - \phi_x(z_1))(\phi_y(z_2) - \phi_y(z_1)). \end{aligned} \tag{3.4}$$

Let us further introduce the functions on  $\Sigma$  defined by

$$A(z) := \frac{\phi_{yy}}{\phi_{xy} + \sqrt{H}}(z), \quad B(z) := \frac{\phi_{xx}}{\phi_{xy} + \sqrt{H}}(z).$$

Note that

$$\begin{aligned}
 1 + AB &= 2 \frac{\phi_{xy}}{\phi_{xy} + \sqrt{H}}, \quad 1 - AB = 2 \frac{\sqrt{H}}{\phi_{xy} + \sqrt{H}}, \\
 \frac{A}{1 + AB} &= \frac{\phi_{yy}}{2\phi_{xy}}, \quad \frac{B}{1 + AB} = \frac{\phi_{xx}}{2\phi_{xy}},
 \end{aligned}
 \tag{3.5}$$

and that (1.4) implies that

$$|\phi_{xy} + \sqrt{H} - 2| \leq 10^{-4}, \quad |\phi_{xx}|, |\phi_{yy}| \leq 10^{-4} \quad \text{on } \Sigma,
 \tag{3.6}$$

so that  $|A(z)|, |B(z)| \leq 10^{-3}$ .

$A$  and  $B$  are in fact closely linked with the geometry of the surface  $S$ . Indeed, consider the vectors  $\omega := (-A(z), 1)$  and  $\nu := (1, -B(z))$ . Then one checks easily that these two vectors form a *basis of null vectors* of the Hessian matrix  $D^2\phi(z)$ , i.e., for every  $z \in \Sigma$ , we have

$$(-A(z), 1)D^2\phi(z)^t(-A(z), 1) = 0 \quad \text{and} \quad (1, -B(z))D^2\phi(z)^t(1, -B(z)) = 0.
 \tag{3.7}$$

For fixed  $z$ , let us therefore set

$$T := T_z := \begin{pmatrix} 1 & -A(z) \\ -B(z) & 1 \end{pmatrix}.
 \tag{3.8}$$

Then clearly

$$\begin{aligned}
 (\xi_1, \xi_2)^t T D^2\phi(z) T^t (\eta_1, \eta_2) &= (\xi_1\nu + \xi_2\omega)D^2\phi(z)^t(\eta_1\nu + \eta_2\omega) \\
 &= q(z)(\xi_1\eta_2 + \xi_2\eta_1),
 \end{aligned}$$

where  $q(z) := \omega D^2\phi(z)^t \nu$ . This shows that

$${}^t T D^2\phi(z) T = q(z) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
 \tag{3.9}$$

Moreover, we have

$$q(z) = (-A(z), 1)D^2\phi(z)^t(1, -B(z)) = (-A\phi_{xx} + (1 + AB)\phi_{xy} - B\phi_{yy})(z).$$

And, by our definitions of  $A$  and  $B$ , and (3.5), we easily see that

$$-A\phi_{xx} + (1 + AB)\phi_{xy} - B\phi_{yy} = 2 \frac{H}{\phi_{xy} + \sqrt{H}},$$

so that

$$q(z) = 2 \frac{H}{\phi_{xy} + \sqrt{H}}(z). \tag{3.10}$$

This implies in particular that  $|q(z) - 1| \leq 10^{-3}$ . Note also that by (3.5)

$$\det T_z = 1 - A(z)B(z) = 2 \frac{\sqrt{H}}{\phi_{xy} + \sqrt{H}}(z) = \frac{q(z)}{\sqrt{H(z)}} \sim q(z) \sim 1. \tag{3.11}$$

### 3.1.2 Back to $\Gamma_z(z_1, z_2, z'_1, z'_2)$

Observe next that (3.9) implies that

$$(D^2\phi(z))^{-1} = \frac{1}{q(z)} T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} {}^t T.$$

Thus,

$$\begin{aligned} & (\xi_1, \xi_2)(D^2\phi(z))^{-1} {}^t (\eta_1, \eta_2) \\ &= \frac{1}{q(z)} \left[ (\xi_1 - B(z)\xi_2)(\eta_2 - A(z)\eta_1) + (\eta_1 - B(z)\eta_2)(\xi_2 - A(z)\xi_1) \right]. \end{aligned} \tag{3.12}$$

If we accordingly define the functions

$$\begin{aligned} t_z^1(z_1, z_2) &:= \phi_x(z_2) - \phi_x(z_1) - B(z)(\phi_y(z_2) - \phi_y(z_1)), \\ t_z^2(z_1, z_2) &:= \phi_y(z_2) - \phi_y(z_1) - A(z)(\phi_x(z_2) - \phi_x(z_1)), \end{aligned}$$

then the identity (3.12) shows that we may re-write

$$\Gamma_z(z_1, z_2, z'_1, z'_2) = \frac{1}{q(z)} \left[ t_z^1(z_1, z_2) \cdot t_z^2(z'_1, z'_2) + t_z^1(z'_1, z'_2) \cdot t_z^2(z_1, z_2) \right]. \tag{3.13}$$

In particular, we obtain the following ‘‘hyperbolic factorization’’:

$$\Gamma_z(z_1, z_2) = \frac{2}{q(z)} \cdot t_z^1(z_1, z_2) \cdot t_z^2(z_1, z_2), \tag{3.14}$$

where the first factor  $2/q(z)$  is of size 2, more precisely  $|2/q(z) - 2| \leq 10^{-3}$ , so that it is irrelevant.

Note also that, e.g.,

$$t_{z_1}^2(z_1, z_2) - t_{z_2}^2(z_1, z_2) = (A(z_1) - A(z_2))(\phi_x(z_2) - \phi_x(z_1)), \tag{3.15}$$

and that

$$t_z^i(z_1, z_2) = -t_z^i(z_2, z_1), \quad i = 1, 2. \quad (3.16)$$

### 3.1.3 Caps and the basic decomposition of $S$

**Definition 3.1** Fix  $K \gg 1$  to be a large dyadic number, and  $\mu \geq 1$  real (the reasons for this notation will be clarified in Sect. 6).

Given  $K$  and  $\mu$ , following [18] we shall consider a given covering of  $\Sigma = [-1, 1] \times [-1, 1]$  by  $K^2$  disjoint squares (called *caps*)  $\tau$  of side length  $\mu^{1/2}K^{-1}$ , whose centers are  $K^{-1}$  separated. It can then happen that such a cap  $\tau$  is no longer contained in  $\Sigma$ ; in that case, we truncate it by replacing it with its intersection with  $\Sigma$ . Note that one usually envisions caps to be subsets of the hypersurface  $S$ ; for our purposes, however, it is more convenient to work with caps  $\tau \subset \Sigma$ , which then can be identified with the corresponding caps  $\{(z, \phi(z)) : z \in \tau\}$  on  $S$ .

Observe that for  $\mu = 1$ , this includes in particular the case of the covering of  $\Sigma$  by caps  $\tau$  which are pairwise disjoint in measure - this is what we had called the *basic decomposition of  $S$  into caps* in [13]. If  $f$  is a given function on  $\Sigma$ , we had then defined  $f_\tau := f\chi_\tau$ .

For general  $\mu \geq 1$ , motivated by Guth's inductive argument in [18], we shall, however, only assume that  $f_\tau$  is a function such that  $\text{supp } f_\tau \subset \tau$ . Actually, Guth assumes more generally that  $\tau$  is a cap of side length  $r_\tau$  with  $K^{-1} \leq r_\tau \leq \mu^{1/2}K^{-1}$ , but since we are only assuming that  $f_\tau$  is supported in  $\tau$ , we can then as well replace  $\tau$  by a larger cap of side length  $\mu^{1/2}K^{-1}$ , as we did.

Assume now that  $\tau_1 \neq \tau_2$  are two different caps, with centers  $z_1^c = (x_1^c, y_1^c)$ , respectively  $z_2^c = (x_2^c, y_2^c)$ , of side length  $\mu^{1/2}K^{-1}$ . The previous definition of strong transversality motivates the following:

**Definition 3.2** We say that  $\tau_1 \neq \tau_2$  are *strongly separated* if

$$\max\{\min\{|t_{z_1^c}^1(z_1^c, z_2^c)|, |t_{z_1^c}^2(z_1^c, z_2^c)|\}, \min\{|t_{z_2^c}^1(z_1^c, z_2^c)|, |t_{z_2^c}^2(z_1^c, z_2^c)|\}\} \geq 50\mu^{1/2}K^{-1}.$$

We shall distinguish between the cases where  $|y_2^c - y_1^c| \geq |x_2^c - x_1^c|$ , and where  $|x_2^c - x_1^c| \geq |y_2^c - y_1^c|$ . Let us mostly concentrate on the first case; the other case can be treated in the same way by interchanging the roles of  $x$  and  $y$ . So, for the rest of this section, let us make the following assumption:

**Assumption 1** Assume that  $|y_2^c - y_1^c| \geq |x_2^c - x_1^c|$ .

**Remark 3.3** If the caps  $\tau_1$  and  $\tau_2$  are strongly separated, so that, say,  $|t_{z_2^c}^1(z_1^c, z_2^c)| \geq 50\mu^{1/2}K^{-1}$  and  $|t_{z_2^c}^2(z_1^c, z_2^c)| \geq 50\mu^{1/2}K^{-1}$ , then

$$|\Gamma_z(z_1, z_2, z'_1, z'_2)| \geq 4\mu K^{-2} \quad \text{for all } z_1, z'_1 \in \tau_1, z, z_2, z'_2 \in \tau_2. \quad (3.17)$$

This result generalizes the corresponding result in Remark 4.8 of [13], whose proof easily extends to our present situation by means of the identity (3.13). It will allow us to establish favorable bilinear estimates later on for the contributions by the tangential terms associated to the cells arising in Guth’s cell decomposition.

### 3.2 Not strongly separated caps

Assume now that  $\tau_1$  and  $\tau_2$  are not strongly separated.

**Case A.** Assume that  $|y_2^c - y_1^c| \leq 100\mu^{1/2}K^{-1}$ . Then, by Assumption 1, also  $|x_2^c - x_1^c| \leq 100\mu^{1/2}K^{-1}$ . Both caps are then contained in a cap of slightly bigger size  $100\mu^{1/2}K^{-1} \leq 100\mu^{1/2}K^{-1/4}$ .

Let us therefore assume from here on that  $|y_2^c - y_1^c| > 100\mu^{1/2}K^{-1}$ .

Observe that our assumptions on  $\phi$  in combination with Assumption 1 then easily imply that  $|t_z^1(z_1, z_2)| \sim |y_2 - y_1|$  for every  $z_1 \in \tau_1, z_2 \in \tau_2$  and  $z \in \tau_1 \cup \tau_2$ , and thus we may assume that

$$\min\{|y_2^c - y_1^c|, |t_{z_1^c}^2(z_1^c, z_2^c)|\} \leq 100\mu^{1/2}K^{-1} \text{ and } \min\{|y_2^c - y_1^c|, |t_{z_2^c}^2(z_1^c, z_2^c)|\} \leq 100\mu^{1/2}K^{-1}.$$

In particular, we have

$$|t_{z_1^c}^2(z_1^c, z_2^c)| \leq 100\mu^{1/2}K^{-1} \text{ and } |t_{z_2^c}^2(z_1^c, z_2^c)| \leq 100\mu^{1/2}K^{-1}. \tag{3.18}$$

Note also that by (3.15)

$$|t_{z_1^c}^2(z_1^c, z_2^c) - t_{z_2^c}^2(z_1^c, z_2^c)| \sim |A(z_1^c) - A(z_2^c)||y_2^c - y_1^c|, \tag{3.19}$$

with constants very close to 1.

We shall therefore distinguish two further cases.

**Case B.** Assume that  $|y_2^c - y_1^c| > 100\mu^{1/2}K^{-1}$  and  $|A(z_1^c) - A(z_2^c)| > \mu^{1/2}K^{-3/4}$ .

Then, by (3.18) and (3.19),  $|y_2^c - y_1^c| \leq 300K^{-1/4} \leq 300\mu^{1/2}K^{-1/4}$ , and arguing as before we see that both caps are contained in a cap of size  $400\mu^{1/2}K^{-1/4}$ .

This leaves us with the case where  $|y_2^c - y_1^c| > 100\mu^{1/2}K^{-1}$  and  $|A(z_1^c) - A(z_2^c)| \leq \mu^{1/2}K^{-3/4}$ . Actually, in what follows, we shall not really make use of the condition  $|y_2^c - y_1^c| > 100\mu^{1/2}K^{-1}$  and shall therefore henceforth concentrate on

**Case C.** Assume that

$$|t_{z_1^c}^2(z_1^c, z_2^c)| \leq 100\mu^{1/2}K^{-1} \text{ and } |A(z_1^c) - A(z_2^c)| \leq \mu^{1/2}K^{-3/4}. \tag{3.20}$$

**Notation** We fix a point  $z_1$  (which would be the point  $z_1^c$  in Case C), and set  $A_1 := A(z_1)$ . Then, we define

$$R_I := \{z \in \Sigma : |t_{z_1}^2(z_1, z)| \leq 100\mu^{1/2}K^{-1}\}, \tag{3.21}$$

$$R_{II} := \{z \in \Sigma : |A(z) - A_1| \leq \mu^{1/2}K^{-3/4}\}. \tag{3.22}$$

### 3.2.1 Level curves of $t_{z_1}^2(z_1, \cdot)$

Let us fix  $z_1 \in \Sigma$ , and let us abbreviate  $\mathbf{t}_{z_1}(x, y) := t_{z_1}^2(z_1, (x, y))$ . From our definition of  $\mathbf{t}_{z_1}$  we compute that

$$\nabla \mathbf{t}_{z_1}(z) = (\phi_{xy} - A_1 \phi_{xx}, \phi_{yy} - A_1 \phi_{xy})(z) = (1, 0) + O(10^{-5}). \tag{3.23}$$

**Lemma 3.4** *Let  $\alpha := \min_{z \in \Sigma} t_{z_1}^2(z_1, z)$ ,  $\beta := \max_{z \in \Sigma} t_{z_1}^2(z_1, z)$ . There exists a  $C^M$ -function  $h : [\alpha, \beta] \times [-1, 1] \rightarrow \mathbb{R}$  such that the curves  $\gamma_{I,v}(y) := (h(v, y), y)$ ,  $y \in [-1, 1]$ , with  $v \in [\alpha, \beta]$ , are level curves of the function  $t_{z_1}^2(z_1, \cdot)$  which fibre the set  $\Sigma_{z_1} := \{(h(v, y), y) : [(v, y) \in [\alpha, \beta] \times [-1, 1]]\}$  into (“almost vertical”) curves, and  $\Sigma \subset \Sigma_{z_1} \subset 2\Sigma$ . Moreover, the mapping  $\mathcal{H} : [\alpha, \beta] \times [-1, 1] \rightarrow \Sigma_{z_1}$ ,  $(v, y) \mapsto (h(v, y), y)$ , is a  $C^M$ -diffeomorphism, and more precisely we have that  $h_y(y, v) = O(10^{-4})$ ,  $h_v(y, v) = 1 + O(10^{-4})$ .*

**Proof** Consider the mapping  $\mathcal{G} : (x, y) \mapsto (\mathbf{t}_{z_1}(x, y), y)$ . Then, by (3.23)

$$D\mathcal{G}(x, y) = \begin{pmatrix} \phi_{xy} - A_1 \phi_{xx} & \phi_{yy} - A_1 \phi_{xy} \\ 0 & 1 \end{pmatrix} (z) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + O(10^{-5}).$$

Therefore, the results follow in a straight-forward manner from the inverse function theorem, by patching together local inverse functions. Note that the inverse function  $\mathcal{H}$  to  $\mathcal{G}$  must be of the form  $\mathcal{H}(v, y) = (h(v, y), y)$ , and that  $t_{z_1}^2(z_1, (h(y, v), y)) = \mathbf{t}_{z_1}(h(y, v), y) = v$ . This also implies that

$$0 = \partial_x \mathbf{t}_{z_1} \cdot h_y + \partial_y \mathbf{t}_{z_1}; \quad 1 = \partial_x \mathbf{t}_{z_1} \cdot h_v, \tag{3.24}$$

so that, in view of (3.23),  $h_y(y, v) = O(10^{-4})$  and  $h_v(y, v) = 1 + O(10^{-4})$ . □

Note that this result also implies that horizontal sections of  $R_I$  have length  $O(\mu^{1/2} K^{-1})$ .

### 3.2.2 On the direction of level curves of $t_{z_1}^2(z_1, \cdot)$ within $R_I \cap R_{II}$

The following lemma gives us an important geometric information.

**Lemma 3.5** *The set  $R_I$  fibers into the level curves  $\gamma_{I,v}(y) := (h(v, y), y)$ ,  $y \in [-1, 1]$ , with  $|v| \leq 100\mu^{1/2} K^{-1}$ . If  $z = (h(v, y), y)$  lies on such a curve, denote by  $X_z := (h_y(v, y), 1)$  the corresponding tangent vector at the point  $z$ . Then, if  $z$  lies also in  $R_{II}$ , i.e., if  $z \in R_I \cap R_{II}$ , we have that*

$$|X_z - (-A_1, 1)| \leq 3\mu^{1/2} K^{-3/4}.$$

Thus, up to an error of order  $O(\mu^{1/2} K^{-3/4})$ , for all points  $z$  in  $R_I \cap R_{II}$  the tangent vectors to the level curves of  $t_{z_1}^2(z_1, \cdot)$  point in the same direction given by  $(-A_1, 1)$ .



**Proof** By Lemma 3.4 the set  $R_I$  fibers into the level curves  $\gamma_{I,v}(y) := (h(v, y), y)$ ,  $y \in [-1, 1]$ , with  $|v| \leq 100\mu^{1/2}K^{-1}$ . Fix any such  $v$ . Then  $X_z = (h_y(v, y), 1)$  is a tangent vector of length  $1 + O(10^{-4})$  to the corresponding curve, if  $z := (h(v, y), y)$ . Note that, by (3.23) and (3.24),

$$h_y(v, y) = -\frac{\partial_y \mathbf{t}_{z_1}}{\partial_x \mathbf{t}_{z_1}}(z) = -\frac{\phi_{yy}(z) - A_1 \phi_{xy}(z)}{\phi_{xy}(z) - A_1 \phi_{xx}(z)}.$$

Let us compare this quantity with the one where  $A_1$  is replaced by  $A(z)$ , i.e., with

$$-\frac{\phi_{yy}(z) - A(z)\phi_{xy}(z)}{\phi_{xy}(z) - A(z)\phi_{xx}(z)}.$$

Our definition of  $A(z)$  implies that  $\phi_{yy} - A\phi_{xy} = A\sqrt{H}$  and that

$$\phi_{xy}(z) - A\phi_{xx} = \frac{H + \phi_{xy}\sqrt{H}}{\phi_{xy} + \sqrt{H}} = \sqrt{H},$$

so that

$$-\frac{\phi_{yy}(z) - A\phi_{xy}(z)}{\phi_{xy}(z) - A\phi_{xx}(z)} = -A(z). \tag{3.25}$$

Therefore, if  $z = (h(v, y), y) \in R_{II}$ , i.e., if  $|A(z) - A_1| \leq \mu^{1/2}K^{-3/4}$ , by (1.2), (1.3), we see that  $|h_y(v, y) - (-A(z))| \leq 2\mu^{1/2}K^{-3/4}$ , hence  $|h_y(v, y) - (-A_1)| \leq 3\mu^{1/2}K^{-3/4}$ , if  $K$  is supposed to be sufficiently large. This implies that  $|X_z - (-A_1, 1)| \leq \mu^{1/2}3K^{-3/4}$ .  $\square$

### 3.3 Moving rectangular boxes into vertical position at the origin

Suppose that  $I$  is a subinterval of  $[-1, 1]$  of length  $b := |I|$  so that for any  $y \in I$  there is some  $x_y$  so that the point  $z_y := (x_y, y)$  lies in  $R_I \cap R_{II}$ . Lemma 3.5 then shows that the set  $(R_I \cap R_{II}) \cap ([-1, 1] \times I)$  is essentially contained in a rectangular box  $L$  of dimension  $100\mu^{1/2}K^{-3/4} \times b$ , pointing in the direction of the vector  $\omega := (-A_1, 1)$ . Moreover, up to an error of order  $O(\mu^{1/2}K^{-3/4})$ , we may replace  $A_1$  by  $A(z)$ , for any choice of point  $z \in L$ .

Indeed, in Sect. 4, based on Lemma 3.5 and Corollary 2.2, we shall devise in a systematic way such kind of rectangular boxes  $L$ , whose lengths will in addition satisfy the following condition:

$$100\mu^{1/2}K^{-3/4} \leq b \leq K^{-\varepsilon'}, \tag{3.26}$$

where  $\varepsilon' \in (0, \varepsilon)$  will be a fixed, but sufficiently small constant to be defined later. Moreover, in view of Remark 6.2, we may also assume that  $\mu \leq K^\varepsilon$ , so that  $100\mu^{1/2}K^{-3/4} \leq K^{-\varepsilon} \leq K^{-\varepsilon'}$ .

Here comes another **crucial observation**: let  $z = z_y$  for  $y \in I$ . Then, by (3.7), we know that  $(-A(z), 1)D^2\phi(z)^t(-A(z), 1) = 0$ . But, since  $z \in R_{II}$ , we also have that  $|A(z) - A_1| \leq \mu^{1/2}K^{-3/4}$ , and thus  $\omega D^2\phi(z_y)^t\omega = (-A_1, 1)D^2\phi(z_y)^t(-A_1, 1) = O(\mu^{1/2}K^{-3/4})$ . And, since the box  $L$  is of horizontal width  $100\mu^{1/2}K^{-3/4}$ , the same estimate holds throughout  $L$ . Consequently, we see that we may assume (with  $\omega = (-A_1, 1)$ ) that, say,

$$|\langle \omega, \nabla \rangle^2 \phi(z)| \leq C\mu^{1/2}K^{-3/4} \quad \text{for all } z \in L. \tag{3.27}$$

Moreover, by our assumptions on  $\phi$ , clearly we also have

$$|\langle \omega, \nabla \rangle^m \phi(z)| \leq c_m \quad \text{for all } z \in L, \tag{3.28}$$

for all  $m = 3, \dots, M$ , with constants  $c_m \lesssim 10^{-5}$ . Restricting these estimates to lines parallel to  $\mathbb{R}\omega$ , and applying Lemma 2.4 to  $\langle \omega, \nabla \rangle^2 \phi$  along these lines, we see that these two estimates imply that

$$|\langle \omega, \nabla \rangle^m \phi(z)| \leq \mu^{1/2}\tilde{C}_m(\varepsilon)K^{-3/4}b^{2-m} \quad \text{for all } z \in L, m = 2, \dots, M. \tag{3.29}$$

Let us now denote by  $z_0 = (x_0, y_0)$  the center of our rectangular box  $L$ . For simplicity, we may and shall assume that  $A_1 = A(z_0)$ .

Our goal will be to find an affine-linear transformation  $z = z_0 + T\tilde{z}$  so that for the accordingly transformed function  $\tilde{\phi}(\tilde{z}) := \phi(z_0 + T\tilde{z})$  we have that

$$D^2\tilde{\phi}(0) = q(z_0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where  $q(z_0) \in \mathbb{R}$ . Note that in the coordinates  $\tilde{z}$ , the point  $z_0$  then corresponds to  $\tilde{z}_0 = 0$ .

To this end, let us put  $B_1 := B(z_0)$ , and choose for  $T$  the matrix  $T_{z_0}$  defined by (3.8), i.e.,

$$T := T_{z_0} := \begin{pmatrix} 1 & -A_1 \\ -B_1 & 1 \end{pmatrix}.$$

Then, by (3.9), we have indeed that

$$D^2\tilde{\phi}(0) = {}^tT D^2\phi(z_0)T = q(z_0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

with  $q(z)$  defined as in Sect. 3.1.1. Note also that by (3.11) the Jacobian determinant of our change of coordinates is given by

$$\det T_{z_0} = \frac{q(z_0)}{\sqrt{H(z_0)}} \sim q(z_0) \sim 1.$$

In the new affine-linear coordinates  $\tilde{z}$ , denote quantities like  $A, L$ , etc., by  $\tilde{A}, \tilde{L}$ , etc.. Then  $\tilde{A}(\tilde{z}_0) = \tilde{A}(0) = 0$ , so that  $\tilde{\omega} = (0, 1)$ . This corresponds to the following observation: we have  $\tilde{\phi}(\tilde{x}, \tilde{y}) = \phi(x_0 + \tilde{x} - A_1\tilde{y}, y_0 - B_1\tilde{x} + \tilde{y})$ , so that  $\frac{\partial}{\partial \tilde{y}}\tilde{\phi}(\tilde{z}) = \langle \omega, \nabla \rangle \phi(z_0 + T\tilde{z})$ . This shows that indeed the directional derivative  $\langle \omega, \nabla \rangle$  corresponds to the partial derivative with respect to  $\tilde{y}$  in the coordinates  $\tilde{z} = (\tilde{x}, \tilde{y})$ . Thus, by (3.29), we have that

$$|\partial_{\tilde{y}}^m \tilde{\phi}(\tilde{z})| \leq \tilde{C}_m(\varepsilon)\mu^{1/2}K^{-3/4}b^{2-m} \quad \text{for all } \tilde{z} \in \tilde{L}, m = 2, \dots, M, \quad (3.30)$$

if  $\tilde{L}$  corresponds to  $L$  in the  $\tilde{z}$ -coordinates, i.e.,  $\tilde{L} = T^{-1}(-z_0 + L)$ . Note that  $\tilde{L}$  is essentially again a rectangular box of dimension  $100\mu^{1/2}K^{-3/4} \times b$ , but centered at the origin and vertical, so that we may assume that  $\tilde{L}$  is contained in  $\tilde{\tilde{L}} := [-100\mu^{1/2}K^{-3/4}, 100\mu^{1/2}K^{-3/4}] \times [-b, b]$ . We may and shall assume that (3.30) holds even on  $\tilde{\tilde{L}}$ .

From (1.3), we also get the following estimates on the box  $\tilde{\tilde{L}}$  :

$$\|\partial_{\tilde{x}}^\alpha \partial_{\tilde{y}}^\beta \tilde{\phi}\|_\infty \leq C_m 10^{-5} \quad \text{for } 3 \leq \alpha + \beta \leq M, \quad (3.31)$$

for constants  $C_m > 0$  which may possibly be much bigger than 1 if  $m$  is very large.

Finally we note that by replacing  $\tilde{\phi}$  with  $q(z_0)^{-1}\tilde{\phi}$ , and subtracting the first order Taylor polynomial at the origin from  $\tilde{\phi}$ , we may even assume that

$$\tilde{\phi}(0) = 0, \quad \nabla \tilde{\phi}(0) = 0, \quad D^2 \tilde{\phi}(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.32)$$

We remark that our affine-linear coordinate change for passing to the  $\tilde{z}$ -coordinates, as well as the further adjustments to  $\tilde{\phi}$  that we just have explained, have no essential effect on the associated Fourier extension estimates, except that the operator norms may be bigger by a factor  $C \sim 1$ .

Last, but not least, observe also that we may assume that the same type of estimates (3.29) and (3.31) will also hold on the double  $2\tilde{L}$  of  $\tilde{L}$ , with constants  $\tilde{C}_m(\varepsilon)$  respectively  $C_m$  that possibly increase by yet another factor  $C \sim 1$ .

In the next subsection, we shall show how our previous results allow also the carry out later the induction on scales step.

### 3.4 The induction on scales step

For this subsection, we may and shall assume that  $\mu = 1$ . Moreover, to simplify the subsequent discussions, let us drop the factor 100 from the horizontal length of our box  $\tilde{L}$ , i.e., let us assume that  $\tilde{L}$  is contained in  $\tilde{\tilde{L}}$  given by  $[-K^{-3/4}, K^{-3/4}] \times [-b, b]$ , and let us correspondingly drop this factor also from (3.26).

In a second step, let us then scale the  $\tilde{z}$ -coordinates by writing  $\tilde{x} = K^{-3/4}x'$ ,  $\tilde{y} = by'$ ,  $z' = (x', y')$ , and

$$\phi^s(z') := \frac{K^{3/4}}{b} \tilde{\phi}(K^{-3/4}x', by').$$

Note that in the new coordinates  $z'$ ,  $\tilde{L}$  corresponds to our standard square  $\Sigma$ . Then, on  $\Sigma$ , we have

$$\partial_{x'}^\alpha \partial_{y'}^\beta \phi^s(x', y') = (K^{-3/4})^{\alpha-1} b^{\beta-1} \partial_{\tilde{x}}^\alpha \partial_{\tilde{y}}^\beta \tilde{\phi}(K^{-3/4}x', by'),$$

so that in view of (3.32)

$$\phi^s(0) = 0, \nabla \phi^s(0) = 0, D^2 \phi^s(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.33}$$

And, if  $\alpha \geq 1$  and  $\alpha + \beta \geq 3$ , by (3.31) and (3.26), we have

$$\|\partial_{x'}^\alpha \partial_{y'}^\beta \phi^s\|_\infty \leq (K^{-3/4})^{\alpha-1} b^{\beta-1} C_m 10^{-5} \leq b^{\alpha-1} b^{\beta-1} C_m 10^{-5} \leq C_m 10^{-5} b \leq 10^{-5}.$$

For the remaining case  $\alpha = 0$ , we use the improved estimate (3.29), which implies that

$$\|\partial_{y'}^\beta \phi^s\|_\infty \leq (K^{-3/4})^{-1} b^{\beta-1} \tilde{C}_\beta(\varepsilon) K^{-3/4} b^{2-\beta} = \tilde{C}_\beta(\varepsilon) b \leq 10^{-5},$$

since  $b \leq K^{-\varepsilon'}$  and we can choose  $K$  sufficiently large.

We thus see that

$$\|\partial_{x'}^\alpha \partial_{y'}^\beta \phi^s\|_\infty \leq 10^{-5} \quad \text{for } 3 \leq |\alpha| + |\beta| \leq M,$$

if we assume  $K \gg 1$  to be sufficiently large.

Actually, if we denote by  $2L$  the doubling of  $L$  which keeps the center of  $L$  fixed, we may even assume that the estimates (3.28), (3.29) hold true on  $2L$ , with constants bigger by some fixed factor only, and so the same arguments used before show that we may even assume that  $\phi^s$  is defined on  $2\Sigma$ , and that the previous estimates hold true even on  $2\Sigma$ .

Thus we see that the function  $\phi^s$  lies again in  $\text{Hyp}^M(\Sigma)$ .

### 3.4.1 Final rescaling step in the induction on scales argument

Recall that we assume that  $\mu = 1$ . Explicitly, our construction of  $\phi^s$  shows that

$$\begin{aligned} \phi^s(x', y') &= \frac{1}{q(z_0)} b K^{-3/4} \left[ \phi(x_0 + K^{-3/4}x' - A_1 by', y_0 - B_1 K^{-3/4}x' + by') - \phi_x(z_0) \right. \\ &\quad \left. \times (K^{-3/4}x' - A_1 by') - \phi_y(z_0)(-B_1 K^{-3/4}x' + by') \right] + \text{constant}. \end{aligned} \tag{3.34}$$

Thus, if  $f_L := f\chi_L$ , and if  $f^L$  denotes the corresponding function in the  $z'$ -coordinates, then changing coordinates we obtain

$$\begin{aligned} \mathcal{E}_\phi f_L(\xi) &= \int f_L(x, y) e^{-i(\xi_1 x + \xi_2 y + \xi_3 \phi(x, y))} dx dy \\ &= C b K^{-3/4} \int f^L(x', y') e^{-i\Phi(x', y'; \xi)} dx' dy', \end{aligned}$$

$C \sim 1$ , and where the phase is given by

$$\begin{aligned} \Phi(x', y'; \xi) &:= \xi_3 \left[ b K^{-3/4} \phi^s(x', y') + \phi_x(z_0)(K^{-3/4} x' - A_1 b y') \right. \\ &\quad \left. + \phi_y(z_0)(-B_1 K^{-3/4} x' + b y') \right] + \xi_1(x_0 + K^{-3/4} x' - A_1 b y') \\ &\quad + \xi_2(y_0 - B_1 K^{-3/4} x' + b y') \\ &\quad + \text{constant} \cdot \xi_3. \end{aligned}$$

Thus, up to a fixed linear function in  $\xi$ , which is irrelevant, we may assume that

$$\begin{aligned} \Phi(x', y'; \xi) &= \xi_3 b K^{-3/4} \phi^s(x', y') + x' K^{-3/4} (\xi_1 - B_1 \xi_2 + (\phi_x(z_0) - B_1 \phi_y(z_0)) \xi_3) \\ &\quad + y' b (\xi_2 - A_1 \xi_1 + (\phi_y(z_0) - A_1 \phi_x(z_0)) \xi_3). \end{aligned}$$

This implies that

$$|\mathcal{E}_\phi f_L(\xi)| = b K^{-3/4} |\mathcal{E}_{\phi^s} f^L(S\xi)|, \tag{3.35}$$

where  $S\xi$  is defines by

$$\begin{aligned} S\xi &:= \left( K^{-3/4} (\xi_1 - B_1 \xi_2 + (\phi_x(z_0) - B_1 \phi_y(z_0)) \xi_3), \right. \\ &\quad \left. b (\xi_2 - A_1 \xi_1 + (\phi_y(z_0) - A_1 \phi_x(z_0)) \xi_3), b K^{-3/4} \xi_3 \right). \end{aligned}$$

We shall be interested in estimating  $\|\mathcal{E}_\phi f_L\|_{L^p(B_R)}$ , where  $B_R$  denotes the Euclidean ball of radius  $R$  centered at the origin. Note that if  $\xi \in B_R$ , then  $\xi' = S\xi$  lies in the set  $B'_R$  defined by

$$|\xi'_1| \leq 3K^{-3/4}R, \quad |\xi'_2| \leq 2bR, \quad |\xi'_3| \leq bK^{-3/4}R.$$

What is important to us is the estimate for the third component  $\xi'$ , which is bounded by  $R' := K^{-3/4}R \ll R$ . This will allow to go from scale  $R'$  to scale  $R$  as in [13].

Indeed, observe the following estimates, which follow easily from our definition of  $f^L$  and (3.35):

$$\begin{aligned} \|f^L\|_2 &\leq (bK^{-3/4})^{-1/2} \|f_L\|_2, \quad \|f^L\|_\infty \leq \|f\|_\infty, \\ \|\mathcal{E}_\phi f_L\|_{L^p(B_R)} &\leq (bK^{-3/4})^{1-\frac{2}{p}} \|\mathcal{E}_{\phi^s} f^L\|_{L^p(B'_R)}. \end{aligned} \tag{3.36}$$

Now assume by induction hypothesis that

$$\|\mathcal{E}_{\phi^s} f^L\|_{L^p(B_{R'})} \leq C_\varepsilon R'^\varepsilon \|f^L\|_{L^2(\Sigma)}^{2/q} \|f^L\|_{L^\infty(\Sigma)}^{1-2/q},$$

where  $\varepsilon > 0$  is as in Theorem 1.3. By means of Lemma 5.1 in [13] we can replace the ball  $B_{R'}$  on the left-hand side by  $\mathbb{R}^2 \times [-R', R']$  and keep the same estimate, with a possibly slightly larger constant  $C'_\varepsilon$ . In particular, we see that

$$\|\mathcal{E}_{\phi^s} f^L\|_{L^p(B'_{R'})} \leq C'_\varepsilon R'^\varepsilon \|f^L\|_{L^2(\Sigma)}^{2/q} \|f^L\|_{L^\infty(\Sigma)}^{1-2/q}.$$

Combining this with (3.36), we see that

$$\begin{aligned} \|\mathcal{E}_\phi f_L\|_{L^p(B_R)} &\leq (bK^{-3/4})^{1-\frac{2}{p}} C'_\varepsilon R'^\varepsilon \|f^L\|_{L^2(\Sigma)}^{2/q} \|f^L\|_{L^\infty(\Sigma)}^{1-2/q} \\ &\leq C'_\varepsilon R'^\varepsilon (bK^{-3/4})^{\frac{1}{q'}-\frac{2}{p}} \|f_L\|_2^{2/q} \|f\|_\infty^{1-2/q} \\ &\leq C'_\varepsilon R^\varepsilon (K^{-3/4})^{\frac{1}{q'}-\frac{2}{p}+\varepsilon} \|f_L\|_2^{2/q} \|f\|_\infty^{1-2/q}, \end{aligned} \tag{3.37}$$

since we assume that  $p > 2q'$ . It is important that the last estimate does not depend on the length  $b$  of  $L$ , which may vary with  $L$ . By means of (3.37), we can now proceed similarly as in Section 1 of [13] and sum these estimates over all boxes  $L$  in an appropriate way—for the details of this, we refer to Sect. 5.

## 4 Broad points

### 4.1 Definition of broad points and the underlying family of rectangles

For the unperturbed hyperbolic paraboloid, the definition of broadness is based on horizontal and vertical strips, since these are the sets which lack strong transversality. Here, we will devise a family  $\tilde{\mathcal{L}}$  of rectangles (and their intersections) adapted to our perturbed hyperbolic paraboloid.

Let us assume as in Definition 3.1 that  $K \gg 1$ . Moreover, in view of Remark 6.2, let us also assume that  $1 \leq \mu \leq K^{\varepsilon/2}$ , where  $\varepsilon > 0$  is as in Theorem 1.3. Let us further fix an according family of caps  $\tau$  of side length  $\mu^{1/2} K^{-1}$ . Note that as in [18, Theorem 2.4] and [13, Theorem 2.1], our main goal later will be to prove Theorem 5.1, in which we are indeed assuming that the caps  $\tau$  form the basic decomposition of  $\Sigma$ , with  $\mu = 1$ . However, for the inductive argument which is used to prove this theorem, we shall be forced to consider also cases where  $\mu > 1$ .

As usual, for  $0 < \alpha < 1$ , we define a point  $\xi$  to be  $\alpha$ -broad for  $\mathcal{E}_\phi f$ , if

$$\max_{L \in \tilde{\mathcal{L}}} |\mathcal{E}_\phi f_L(\xi)| \leq \alpha |\mathcal{E}_\phi f(\xi)|,$$

where  $f_L := \sum_{\tau \subset L} f_\tau$ . We define  $Br_\alpha \mathcal{E}_\phi f(\xi)$  to be  $|\mathcal{E}_\phi f(\xi)|$  if  $\xi$  is  $\alpha$ -broad, and zero otherwise.

We explain now how to construct the family  $\tilde{\mathcal{L}}$ .

As we saw in Sect. 3.2, Case C, the most troublesome sets lacking strong transversality are the intersection of the sets  $R_I$  and  $R_{II}$  from (3.21). We recall the functions

$$A(z) = \frac{\phi_{yy}}{\phi_{xy} + \sqrt{H}}(z), \quad B(z) = \frac{\phi_{xx}}{\phi_{xy} + \sqrt{H}}(z).$$

Note that for  $\phi \in \text{Hyp}^3$ , we have

$$|\nabla A(z)|, |\nabla B(z)| \leq 1 \tag{4.1}$$

for all  $z \in \Sigma$ .

We shall focus the discussion on  $A$ ; there is an analogous construction with  $B$ . Let  $\{A_k\}_k$  be a equidistant decomposition of the interval  $A(\Sigma)$  of distance  $C\mu^{1/2}K^{-3/4}$ , that is,  $A_{k+1} = A_k + C\mu^{1/2}K^{-3/4}$ , where we will choose the constant  $C$  later, and put

$$R_{II}^k := \{z \in \Sigma : |A(z) - A_k| < C\mu^{1/2}K^{-3/4}\}.$$

The sets  $R_{II}^k$  are not pairwise disjoint, but  $R_{II}^k$  and  $R_{II}^{k'}$  may only overlap if  $|k - k'| \leq 1$ . As we saw in Lemma 3.5, on  $R_I \cap R_{II}$ , the tangents to  $R_I$  point essentially in a fixed direction. This suggests to denote for any given  $k$  by  $\tilde{\omega}_k$  the unit vector pointing in the direction of  $\omega_k := (-A_k, 1)$ , and decompose  $\Sigma$  into strips

$$S_{k,j} := [(j - 1)C'\mu^{1/2}K^{-3/4}, (j + 1)C'\mu^{1/2}K^{-3/4}] \tilde{\omega}_k^\perp + \mathbb{R}\tilde{\omega}_k$$

of thickness  $2C'\mu^{1/2}K^{-3/4}$  and direction  $\omega_k$ , indexed by suitable integers  $j \in \mathbb{Z}$ . Here,  $\tilde{\omega}_k^\perp$  denotes a unit vector orthogonal to  $\omega_k$ , and  $C'$  denotes yet another suitable constant. For fixed  $k$ ,  $S_{k,j}$  and  $S_{k,j'}$  do not overlap unless  $|j - j'| \leq 1$ . Of course they can overlap quite a lot for different values of  $k$ , but we want to consider only the part of  $S_{k,j}$  that intersects  $R_{II}^k$ . More precise, we define the following subset  $S_{k,j}^0$  of  $S_{k,j}$  :

$$S_{k,j}^0 := \{z \in S_{k,j} : (z + \mathbb{R}\tilde{\omega}_k^\perp) \cap S_{k,j} \cap R_{II}^k \neq \emptyset\}.$$

It is clear that the connected components of  $S_{k,j}^0$  are all rectangles inside  $S_{k,j}$  of full width  $2C'\mu^{1/2}K^{-3/4}$ , but unknown length. Since for technical reasons that will become clear later, we do not want the rectangles to be too short, we set

$$S_{k,j}^1 := S_{k,j}^0 + [-\mu^{1/2}K^{-3/4}, \mu^{1/2}K^{-3/4}] \tilde{\omega}_k.$$

Then the connected components of  $S_{k,j}^1$  are rectangles inside  $S_{k,j}$  of full width and length at least  $2\mu^{1/2}K^{-3/4}$ .

On the other hand, we want the lengths of these rectangles not to be too long either, and therefore divide  $S_{k,j}^1$  into rectangles  $\tilde{L}_{k,j}^i$  of lengths at most  $\frac{1}{2}K^{-\epsilon'}$ , but at least

$\mu^{1/2}K^{-3/4}$  ( $\varepsilon' \ll \varepsilon$  to be determined later), by artificially chopping any connected component that is too long. Finally, since that artificial chopping may split a small cap  $\tau$  of size  $\mu^{1/2}K^{-3/4}$  into two, we set

$$L_{k,j}^i := \tilde{L}_{k,j}^i + [-\mu^{1/2}K^{-3/4}, \mu^{1/2}K^{-3/4}]\tilde{\omega}_k,$$

so that for fixed  $k$  and  $j$ , the sets  $L_{k,j}^i$  may intersect, but only two can overlap at any given point. Since  $2\mu^{1/2}K^{-3/4} < \frac{1}{2}K^{-\varepsilon'}$  for sufficiently small  $\varepsilon$ , every  $L_{k,j}^i$  has length between  $\mu^{1/2}K^{-3/4}$  and  $K^{-\varepsilon'}$ .

Let  $\mathcal{L}_1 := \{L_{k,j}^i\}_{k,j,i}$  denote the set of all these rectangles (which we often simply shall call “strips”).

By construction,  $\text{dist}(z, R_{II}^k) \leq 2C'\mu^{1/2}K^{-3/4}$  for all  $z \in S_{k,j}^0$ ; for  $z \in L_{k,j}^i$ , we still have

$$\text{dist}(z, R_{II}^k) \leq (2C' + 2)\mu^{1/2}K^{-3/4},$$

so that

$$|A(z) - A_k| \leq C\mu^{1/2}K^{-3/4} + \|\nabla A\|_\infty(2C' + 2)\mu^{1/2}K^{-3/4} \lesssim \mu^{1/2}K^{-3/4} \quad (4.2)$$

for all  $z \in L_{k,j}^i$ .

We summarize the most important properties of our family of rectangles, which follow immediately from their definition:

**Remarks 4.1** There exist absolute constants  $C_1, N_1 \geq 1$  such that the following hold true:

(i) For all  $k, j, i$  and all  $z \in L_{k,j}^i$ ,

$$|A(z) - A_k| \leq C_1\mu^{1/2}K^{-3/4}.$$

(ii) At any point, at most  $N_1$  sets from  $\mathcal{L}_1$  overlap.

(iii) Every  $L \in \mathcal{L}_1$  has length between  $\mu^{1/2}K^{-3/4}$  and  $K^{-\varepsilon'}$ .

Of course there is the symmetric situation, where the coordinates are interchanged, and  $A$  is replaced by  $B$ , which will give us a similar set of rectangles  $\mathcal{L}_2$ .

To cover also the Cases A and B from Sect. 3.2, we furthermore need caps of size  $\sim \mu^{1/2}K^{-1/4}$ . Let  $\mathcal{L}_3$  be a collection of squares of side length  $C\mu^{1/2}K^{-1/4}$  covering  $\Sigma$ , whose centers are  $\frac{C}{2}\mu^{1/2}K^{-1/4}$  separated.

Finally, we set  $\mathcal{L} := \mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3$ . This family is already well suited for proving our geometric Lemma 4.3. However, later on the application of polynomial partitioning method will even require a family which is closed under intersections. We therefore define

$$\tilde{\mathcal{L}} := \{L_1 \cap \dots \cap L_m : L_1, \dots, L_m \in \mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3, m \in \mathbb{N}\}.$$



Note that due to Remark 4.1(ii), the number  $m$  of possible intersections is uniformly bounded, and each  $\Delta \in \bar{\mathcal{L}}$  is either contained in a strip  $L \in \mathcal{L}_1 \cup \mathcal{L}_2$  of dimensions  $\mu^{1/2}K^{-3/4} \times b$  (respectively  $b \times \mu^{1/2}K^{-3/4}$ ) for some  $b$  with  $\mu^{1/2}K^{-3/4} \leq b \leq K^{-\varepsilon'}$ , or is contained in a large cap  $L \in \mathcal{L}_3$  of side length  $C\mu^{1/2}K^{-1/4}$ . It is then easy to verify the following remark.

**Remark 4.2** There exists an absolute constant  $\bar{N} \in \mathbb{N}$  such that at most  $\bar{N}$  sets  $\Delta \in \bar{\mathcal{L}}$  overlap at any given point  $z \in \Sigma$ .

### 4.2 The geometric lemma

A key observation which is important to adapt the polynomial partitioning method is that caps which are mutually not strongly transversal are somewhat sparse.

Let us fix a parameter  $\varepsilon' > 0$  depending on  $\varepsilon$  and  $\bar{N}$  and then  $M = M(\varepsilon) \in \mathbb{N}$  so that

$$10\bar{N}\varepsilon' = \varepsilon^8 \quad \text{and} \quad \frac{3}{4M(\varepsilon)} \leq \varepsilon' \leq \frac{3}{2M(\varepsilon)}, \tag{4.3}$$

where  $\bar{N}$  is the constant from Remark 4.2.

**Lemma 4.3** (The Geometric Lemma). *There is some constant  $K_1(\varepsilon)$  such that, for every  $K \geq K_1(\varepsilon)$  and every  $\phi \in \text{Hyp}^{M(\varepsilon)}$ , the following holds true: if  $\mathcal{F}$  is any family of caps of side length  $\mu^{1/2}K^{-1}$  which does not contain two strongly separated caps, then there is subcollection  $\mathcal{L}_0 \subset \mathcal{L}$  of cardinality  $|\mathcal{L}_0| \leq K^{3\varepsilon'}$ , such that each cap in  $\mathcal{F}$  is contained in at least one element of  $\mathcal{L}_0$ .*

**Proof** Fix a cap  $\tau_1 \in \mathcal{F}$ , and recall that we denoted its center by  $z_1^c = (x_1^c, y_1^c)$ . If  $\tau_2 \in \mathcal{F}$ , then

$$\min\{|y_2^c - y_1^c|, |t_{z_1^c}^2(z_1^c, z_2^c)|\} \leq 100\mu^{1/2}K^{-1} \quad \text{and} \quad \min\{|y_2^c - y_1^c|, |t_{z_2^c}^2(z_1^c, z_2^c)|\} \leq 100\mu^{1/2}K^{-1}.$$

We will follow the discussion in Sect. 3.2, and the division into cases we devised there. In cases A and B, both caps  $\tau_1, \tau_2$  are clearly contained in a cap of size  $100\mu^{1/2}K^{-1/4}$ , which is contained in a cap of size  $C\mu^{1/2}K^{-1/4}$  from our collection  $\mathcal{L}_3$ , if we assume that  $C$  is sufficiently large.

This leaves us with **Case C**, where

$$|t_{z_1^c}^2(z_1^c, z_2^c)| \leq 100\mu^{1/2}K^{-1} \quad \text{and} \quad |A(z_1^c) - A(z_2^c)| \leq \mu^{1/2}K^{-3/4}. \tag{4.4}$$

Recall from (3.23) that  $\partial_x t_{z_1^c}^2(z_1^c, z) \geq 1/2$ . In Lemma 3.4 we used (3.23) in order to show that we may parametrize the zero set of  $t_{z_1^c}^2(z_1^c, \cdot)$  by a curve  $\gamma = (\gamma_1, \gamma_2) : [-1, 1] \rightarrow \Sigma$  such that  $\gamma_2(t) = t$ . In particular,  $\gamma_2(y_2^c) = y_2^c$ . But then  $z_2^c - \gamma(y_2^c) =$

$(x_2^c - \gamma_1(y_2^c), 0)$ , and thus

$$|z_2^c - \gamma(y_2^c)| \leq 2|t_{z_1^c}^2(z_1^c, z_2^c) - t_{z_1^c}^2(z_1^c, \gamma(y_2^c))| = 2|t_{z_1^c}^2(z_1^c, z_2^c)| \leq 200\mu^{1/2}K^{-1}. \tag{4.5}$$

Consider the  $C^M$  function

$$g(t) := A(\gamma(t)) - A(z_1^c), \quad t \in [-1, 1].$$

Applying Corollary 2.2 to  $g$  and level  $\lambda := 2\mu^{1/2}K^{-3/4}$ , by our choice of  $M = M(\varepsilon)$  there exists a family  $\mathcal{I}$  of subintervals of  $[-1, 1]$  such that

- (i)  $|\mathcal{I}| \leq 60M(1 + \|g^{(M)}\|_\infty^{1/M} K^{3/4M}) \leq G(M) K^{\varepsilon'}$ ,
- (ii) for all  $I \in \mathcal{I}$  and all  $t \in I$  we have  $|g(t)| < 16\mu^{1/2}K^{-3/4}$ ,
- (iii) and for all  $t \in I_0 \setminus \bigcup_{I \in \mathcal{I}} I$  we have  $|g(t)| \geq 2\mu^{1/2}K^{-3/4}$ ,

where  $G(M) \geq 1$  for any integer  $M \geq 3$  is a constant such that the last estimate in (i) holds uniformly for any  $\phi \in \text{Hyp}^M$ . Indeed, by our definition of the function  $g$  and the uniform estimates that are assumed to hold for all functions  $\phi$  in  $\text{Hyp}^M$  in combination with Faà di Bruno’s theorem [29] on derivatives of compositions of functions we easily see that a uniform estimate

$$60M(1 + \|g^{(M)}\|_\infty^{1/M}) \leq G(M)$$

holds true for all  $\phi \in \text{Hyp}^M$ .

In particular, if we set  $K_1(\varepsilon) := G(M(\varepsilon))^{8M(\varepsilon)/3}$ , then in combination with (4.3) we see that

$$(i') \quad |\mathcal{I}| \leq K^{3\varepsilon'/2} \text{ if } K \geq K_1(\varepsilon).$$

By (4.1), (4.4) and (4.5), we see that

$$|g(y_2^c)| \leq |A(\gamma(y_2^c)) - A(z_2^c)| + |A(z_2^c) - A(z_1^c)| < 2\mu^{1/2}K^{-3/4},$$

and hence by (iii),  $y_2^c \in I$  for some  $I \in \mathcal{I}$ . Again by (4.5), we see that  $\tau_2$  is contained in the neighborhood

$$\mathcal{N}(I) := \gamma(I) + B(0, \mu^{1/2}K^{-3/4})$$

of the curve  $\gamma(I)$ .

Choose  $k$  so that  $|A_k - A(z_1^c)| \leq \frac{1}{2}C\mu^{1/2}K^{-3/4}$ . Fix any  $t_0 \in I$ , and choose  $S_{k,j}$  to be one of our strips of direction  $\omega_k$  which contains  $\gamma(t_0)$ . Since these strips slightly overlap, we may and shall choose  $j$  so that  $\text{dist}(\gamma(t_0), \partial S_{k,j}) \geq \frac{1}{2}C'\mu^{1/2}K^{-3/4}$ .

Recall also that by (ii), for any  $t \in I$  we have

$$|A(\gamma(t)) - A(z_1^c)| = |g(t)| \leq 16\mu^{1/2}K^{-3/4},$$

hence  $\gamma(t)$  is in the intersections of the regions  $R_I$  and  $R_{II}$  as defined in (3.21) (modified by a harmless factor 16, which we will ignore). Using Lemma 3.5, we see that the tangential of  $\gamma$  on  $I$  points essentially in direction  $\omega_k$ , and more precisely we obtain that

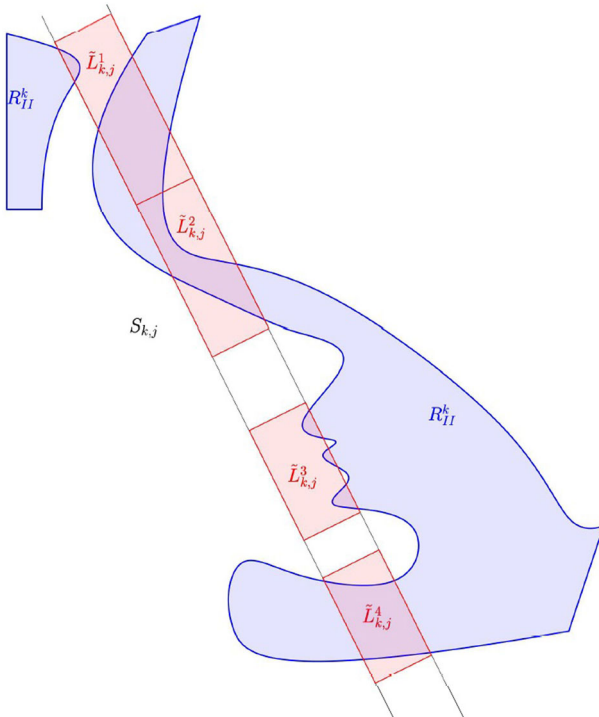
$$|\langle \gamma(t) - \gamma(t_0), \tilde{\omega}_k^\perp \rangle| \lesssim \mu^{1/2} K^{-3/4}.$$

This means that  $\mathcal{N}(I) \subset S_{k,j}$ , provided we choose the constant  $C'$  from the construction of these strips big enough.

For  $z \in \mathcal{N}(I)$ , we find a  $t \in I$  with  $|\gamma(t) - z| < \mu^{1/2} K^{-3/4}$ , so that

$$\begin{aligned} |A(z) - A_k| &\leq |A(z) - A(\gamma(t))| + |g(t)| + |A(z_1^c) - A_k| \\ &\leq (1 + 16 + C/2)\mu^{1/2} K^{-3/4} < C\mu^{1/2} K^{-3/4}, \end{aligned}$$

i.e.,  $\mathcal{N}(I) \subset R_{II}^k$ , provided  $C > 34$ . This shows that  $\mathcal{N}(I) \subset S_{k,j}^1$ , and, since  $\mathcal{N}(I)$  is clearly connected, even that  $\mathcal{N}(I)$  is contained in a connected component of  $S_{k,j}^1$ . Recall that we had divided  $S_{k,j}^1$  into rectangles  $\tilde{L}_{k,j}^i$  of lengths at most  $\frac{1}{2}K^{-\epsilon'}$ , by artificially chopping any connected component that is too long. But, for any  $\tau_2 \subset \mathcal{N}(I)$  with  $\tau_2 \cap \tilde{L}_{k,j}^i \neq \emptyset$ ,  $\tau_2 \subset L_{k,j}^i \in \mathcal{L}_1$ , and there are at most  $6K^{\epsilon'}$  sets  $\tilde{L}_{k,j}^i$  in any connected component of  $S_{k,j}^1$ . In combination with (i'), this proves the claim of the lemma also in Case C. □



### 5 Reduction to estimates for the broad part

In this section, we fix  $\mu = 1$  and consider our basic decomposition of  $\Sigma$  into caps  $\tau$  of side length  $K^{-1}$  which have pairwise disjoint interiors. Consider the families  $\mathcal{L}$  and  $\overline{\mathcal{L}}$  defined in the previous section for  $\mu = 1$ . Recall that each  $\Delta \in \overline{\mathcal{L}}$  is either contained in a strip of dimensions  $K^{-3/4} \times b$  for some  $K^{-3/4} \leq b \leq K^{-\varepsilon'}$ , or is contained in a large cap of side length  $100K^{-1/4}$ .

As in the previous section (for  $\mu = 1$ ), given the function  $f$ ,  $\alpha \in (0, 1)$  and  $K$ , we say that the point  $\xi \in \mathbb{R}^3$  is  $\alpha$ -broad for  $\mathcal{E}_\phi f$  if

$$\max_{\Delta \in \overline{\mathcal{L}}} |\mathcal{E}_\phi f_\Delta(\xi)| \leq \alpha |\mathcal{E}_\phi f(\xi)|,$$

where  $f_\Delta := \sum_{\tau \subset \Delta} f_\tau$ .

We define  $Br_\alpha \mathcal{E}_\phi f(\xi)$  to be  $|\mathcal{E}_\phi f(\xi)|$  if  $\xi$  is  $\alpha$ -broad, and zero otherwise.

We shall prove the following analogue to [18, Theorem 2.4] and [13, Theorem 2.1]:

**Theorem 5.1** *Let  $0 < \varepsilon < 10^{-10}$ , and choose  $M = M(\varepsilon) \in \mathbb{N}$  sufficiently large so that (4.3) is satisfied. Then there are constants  $K = K(\varepsilon) \gg 1$  and  $C_\varepsilon$  such that for any  $\phi \in \text{Hyp}^M$  and any radius  $R \geq 1$  the following hold true:*

$$\|Br_{K^{-\varepsilon}} \mathcal{E}_\phi f\|_{L^{3.25}(B_R)} \leq C_\varepsilon R^\varepsilon \|f\|_{L^2(\Sigma)}^{12/13} \|f\|_{L^\infty(\Sigma)}^{1/13}$$

for every  $f \in L^\infty(\Sigma)$ , and moreover  $K(\varepsilon) \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ .

To show that Theorem 1.3 follows from this result, let us note that it is enough to consider  $q$  close to 2.6 and put  $p := 3.25$ . We divide the domain of integration  $B_R$  in (1.6) into three subsets:

- $A := \{\xi \in B_R : \xi \text{ is } K^{-\varepsilon} \text{ - broad for } \mathcal{E}_\phi f\},$
- $B := \{\xi \in B_R : |\mathcal{E}_\phi f_\Delta(\xi)| > K^{-\varepsilon} |\mathcal{E}_\phi f(\xi)| \text{ for some } \Delta \in \overline{\mathcal{L}},$   
with  $\Delta$  contained in a strip  $L \in \mathcal{L}_1 \cup \mathcal{L}_2\},$
- $C := \{\xi \in B_R \setminus B : |\mathcal{E}_\phi f_\Delta(\xi)| > K^{-\varepsilon} |\mathcal{E}_\phi f(\xi)| \text{ for some } \Delta \in \overline{\mathcal{L}},$   
with  $\Delta$  contained in a large cap  $L \in \mathcal{L}_3\}.$

If  $\xi \in A$ , then  $|\mathcal{E}_\phi f(\xi)| = Br_{K^{-\varepsilon}} \mathcal{E}_\phi f(\xi)$ , so that the contribution of  $A$  can be controlled using Theorem 5.1. Notice that

$$\|f\|_{L^2(\Sigma)}^{12/13} \|f\|_{L^\infty(\Sigma)}^{1/13} \leq \|f\|_{L^2(\Sigma)}^{2/q} \|f\|_{L^\infty(\Sigma)}^{1-2/q},$$

since  $q > 2.6 > 13/6$ .

If  $\xi \in B$ , then there is some  $\Delta \in \overline{\mathcal{L}}$  which is contained in a strip of dimensions  $K^{-3/4} \times b$  for some  $b \leq K^{-\varepsilon'}$ , so that  $|\mathcal{E}_\phi f_\Delta(\xi)| > K^{-\varepsilon} |\mathcal{E}_\phi f(\xi)|$ . Then we may

estimate

$$|\mathcal{E}_\phi f(\xi)| < K^\varepsilon \sup_{\Delta} |\mathcal{E}_\phi f_\Delta(\xi)| \leq K^\varepsilon \left( \sum_{\Delta} |\mathcal{E}_\phi f_\Delta(\xi)|^p \right)^{1/p},$$

where the supremum and sum are taken over all  $\Delta \in \bar{\mathcal{L}}$  contained in a strip  $L \in \mathcal{L}$  of dimensions  $K^{-3/4} \times b$  for some  $b \leq K^{-\varepsilon'}$  (which may depend on  $\Delta$ ). Thus, we can apply to any such  $f_\Delta$  the scaling (associated to the corresponding strip  $L$ ) described in Sect. 3.4, more precisely estimate (3.37), and obtain

$$\|\mathcal{E}_\phi f_\Delta\|_{L^p(B_R)} \leq C'_\varepsilon R^\varepsilon (K^{-3/4})^{\frac{1}{q'} - \frac{2}{p} + \varepsilon} \|f_\Delta\|_2^{2/q} \|f\|_\infty^{1-2/q}.$$

Therefore,

$$\begin{aligned} \|\mathcal{E}_\phi f\|_{L^p(B)} &\leq K^\varepsilon \left( \sum_{\Delta} \|\mathcal{E}_\phi f_\Delta\|_{L^p}^p \right)^{1/p} \\ &\leq C'_\varepsilon K^\varepsilon R^\varepsilon (K^{-3/4})^{\frac{1}{q'} - \frac{2}{p} + \varepsilon} \left( \sum_{\Delta} \|f_\Delta\|_2^{2p/q} \|f\|_\infty^{p(1-2/q)} \right)^{1/p}. \end{aligned}$$

Since  $2p/q > 2$ , taking into account the overlap of the elements of  $\bar{\mathcal{L}}$  (see Remark 4.2) we estimate

$$\sum_{\Delta} \|f_\Delta\|_2^{2p/q} \leq \left( \sum_{\Delta} \|f_\Delta\|_2^2 \right)^{p/q} \leq \bar{N}^{p/q} \|f\|_2^{2p/q}.$$

Hence,

$$\begin{aligned} \|\mathcal{E}_\phi f\|_{L^p(B)} &\leq C'_\varepsilon \bar{N}^{1/q} K^\varepsilon R^\varepsilon (K^{-3/4})^{\frac{1}{q'} - \frac{2}{p} + \varepsilon} \|f\|_2^{2/q} \|f\|_\infty^{(1-2/q)} \\ &\leq \frac{1}{10} C_\varepsilon R^\varepsilon \|f\|_2^{2/q} \|f\|_\infty^{1-2/q}, \end{aligned}$$

since  $p > 2q'$ .

For  $\xi \in C$ , isotropic scaling gives the same result.

### 6 Proof of Theorem 1.2

Following Section 3 in [18] and [13], we shall next formulate a more general statement in Theorem 5.1 which will become amenable to inductive arguments. We have to consider  $\mu \geq 1$ .

We assume that we are given  $\mu \geq 1$ , a dyadic natural number  $K \gg 1$  and a family of caps  $\tau$  of side length  $\mu^{1/2} K^{-1}$ , covering  $\Sigma = [0, 1] \times [0, 1]$ , such that their centers are  $K^{-1}$ -separated. Hence, at any point there will be at most  $\mu$  of these caps which

overlap at that point. Notice also that there are at most  $K^2$  caps  $\tau$  in the family. We also assume that we have a decomposition

$$f = \sum_{\tau} f_{\tau},$$

where  $\text{supp } f_{\tau} \subset \tau$ .

We adapt the notion of broadness to the modified family of caps  $\tau$ : For each  $\Delta \in \bar{\mathcal{L}}$ , we define  $f_{\Delta} := \sum_{\tau \subset \Delta} f_{\tau}$ .

Let  $\alpha \in (0, 1)$ . Given the function  $f$  and  $K$ , we say that the point  $\xi \in \mathbb{R}^3$  is  $\alpha$ -broad for  $\mathcal{E}_{\phi} f$  if

$$\max_{\Delta \in \bar{\mathcal{L}}} |\mathcal{E}_{\phi} f_{\Delta}(\xi)| \leq \alpha |\mathcal{E}_{\phi} f(\xi)|.$$

We define  $Br_{\alpha} \mathcal{E}_{\phi} f(\xi)$  to be  $|\mathcal{E}_{\phi} f(\xi)|$  if  $\xi$  is  $\alpha$ -broad, and zero otherwise.

Theorem 5.1 will be a consequence of the following

**Theorem 6.1** *Let  $0 < \varepsilon < 10^{-10}$ . Then there are constants  $M = M(\varepsilon) \in \mathbb{N}$ ,  $K = K(\varepsilon) \gg 1$ , and  $C_{\varepsilon}$  such that for any  $\phi \in \text{Hyp}^M$ , any family of caps  $\tau$  with multiplicity at most  $\mu$  covering  $\Sigma$  as above and associated functions  $f_{\tau}$  which decompose  $f$ , any radius  $R \geq 1$ , and any  $\alpha \geq K^{-\varepsilon}$ , the following hold true:*

*If for every  $\omega \in \Sigma$ , and every cap  $\tau$  as above,*

$$\oint_{B(\omega, R^{-1/2})} |f_{\tau}|^2 \leq 1, \tag{6.1}$$

then

$$\int_{B_R} (Br_{\alpha} \mathcal{E}_{\phi} f)^{3.25} \leq C_{\varepsilon} R^{\varepsilon} \left( \sum_{\tau} \int |f_{\tau}|^2 \right)^{3/2+\varepsilon} R^{\delta_{trans} \log(K^{\varepsilon} \alpha \mu)}, \tag{6.2}$$

where  $\delta_{trans} := \varepsilon^6$ . Moreover,  $K(\varepsilon) \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ .

Here, in  $\mathbb{R}^n$ , by  $B(\omega, r)$  we denote the Euclidean ball of radius  $r > 0$  and center  $\omega$ , and by  $\oint_A f := \frac{1}{|A|} \int_A f$  we denote the mean value of  $f$  over the measurable set  $A$  of volume  $|A| > 0$ .

From this, as in [13, 18], Theorem 5.1 follows taking  $\mu = 1$ .

As in these papers, let us also choose

$$\delta_{trans} := \varepsilon^6, \quad \delta_{deg} := \varepsilon^4, \quad \delta := \varepsilon^2, \tag{6.3}$$

so that in particular

$$\delta_{trans} \ll \delta_{deg} \ll \delta \ll \varepsilon < 10^{-10}.$$

We next choose  $M = M(\varepsilon) \in \mathbb{N}$  and  $\varepsilon' > 0$  according to (4.3), i.e.,

$$10\bar{N}\varepsilon' = \varepsilon^8 \quad \text{and} \quad \frac{3}{4M(\varepsilon)} \leq \varepsilon' \leq \frac{3}{2M(\varepsilon)},$$

and assume that  $\varepsilon > 0$  is so small that also  $M\delta \gg 1000$  holds true (compare Proposition 6.3).

We also set

$$K = K(\varepsilon) := \left\lfloor K_1(\varepsilon) + e^{\varepsilon^{-10}} + 1 \right\rfloor, \quad D = D(\varepsilon) := R^{\delta_{deg}} = R^{\varepsilon^4}, \quad (6.4)$$

where  $K_1(\varepsilon)$  is the constant from the Geometric Lemma 4.3 and  $\lfloor x \rfloor$  denotes the integer part of  $x$ , and where we are assuming that  $R \geq 1$  is given. In particular, we then have  $K \geq K_1(\varepsilon)$ , as assumed in the Geometric Lemma.

The following remark is similar to Remark 4.1 in [13] and holds (with the same proof) also here:

**Remarks 6.2** a) It is enough to consider the case where  $\alpha\mu \leq K^{-\varepsilon/2}$ , because in the other case, the exponent  $\delta_{trans} \log(K^\varepsilon \alpha\mu)$  is very large and the estimate (6.2) trivially holds true. Note that since in Theorem 6.1 we are also assuming that  $\alpha \geq K^{-\varepsilon}$ , this implies that  $\mu \leq K^{\varepsilon/2}$ . Henceforth, we shall therefore always assume that  $\alpha\mu \leq K^{-\varepsilon/2} \leq 10^{-5}$ .

b) It is then also enough to consider the case where  $R$  is tremendously bigger than  $K$ , say  $R \geq 1000 e^{K e^{\varepsilon^{-1000}}}$ .

As usual, we will work with wave packet decompositions of the functions  $f$  defined on  $S_\phi := \{(x, y, \phi(x, y)) : (x, y) \in \Sigma\}$ . Following [18], we decompose  $\Sigma$  into squares (“caps”)  $\theta$  of side length  $R^{-1/2}$ . By  $\omega_\theta$  we shall denote the center of  $\theta$ , and by  $\nu(\theta)$  the “outer” unit normal to  $S_\phi$  at the point  $(\omega_\theta, \phi(\omega_\theta)) \in S_\phi$ , which points into the direction of  $(-\nabla\phi(\omega_\theta), -1)$ .  $\mathbb{T}(\theta)$  will denote a set of  $R^{1/2}$ -separated tubes  $T$  of radius  $R^{1/2+\delta}$  and length  $R$ , which are all parallel to  $\nu(\theta)$  and for which the corresponding thinner tubes of radius  $R^{1/2}$  with the same axes cover  $B_R$ . We will write  $\nu(T) := \nu(\theta)$  when  $T \in \mathbb{T}(\theta)$ .

Note that for each  $\theta$ , every point  $\xi \in B_R$  lies in  $O(R^{2\delta})$  tubes  $T \in \mathbb{T}(\theta)$ . We put  $\mathbb{T} := \bigcup_\theta \mathbb{T}(\theta)$ . Arguing in the same way as in [18, Proposition 2.6], and observing that it is not necessary for the arguments in the proof, which are based on integrations by parts, that the phase  $\phi$  is  $C^\infty$ , but merely  $C^M$  for sufficiently large  $M$  (more precisely,  $M\delta \gg 1000$ ), we arrive at the following approximate wave packet decomposition.

**Proposition 6.3** *Assume that  $R$  is sufficiently large (depending on  $\delta$ ). Then, for any  $\phi \in \text{Hyp}^M$  (with  $M = M(\varepsilon)$  as before), given  $f \in L^2(\Sigma)$ , we may associate to each tube  $T \in \mathbb{T}$  a function  $f_T$  such that the following hold true:*

- a) *If  $T \in \mathbb{T}(\theta)$ , then  $\text{supp } f_T \subset 3\theta$ .*
- b) *If  $\xi \in B_R \setminus T$ , then  $|\mathcal{E}_\phi f_T(\xi)| \leq R^{-1000} \|f\|_2$ .*
- c) *For any  $x \in B_R$ , we have  $|\mathcal{E}_\phi f(x) - \sum_{T \in \mathbb{T}} \mathcal{E}_\phi f_T(x)| \leq R^{-1000} \|f\|_2$ .*

d) (*Essential orthogonality*) If  $T_1, T_2 \in \mathbb{T}(\theta)$  are disjoint, then

$$|\int f_{T_1} \overline{f_{T_2}}| \leq R^{-1000} \int_{3\theta} |f|^2.$$

e)  $\sum_{T \in \mathbb{T}(\theta)} \int_{\Sigma} |f_T|^2 \leq C \int_{3\theta} |f|^2.$

We next recall the version of the polynomial ham sandwich theorem with non-singular polynomials from [18]. If  $P$  is a real polynomial on  $\mathbb{R}^n$ , we denote by  $Z(P) := \{\xi \in \mathbb{R}^n : P(\xi) = 0\}$  its null variety.  $P$  is said to be *non-singular* if  $\nabla P(\xi) \neq 0$  for every point  $\xi \in Z(P)$ .

Then, by in [18, Corollary 1.7], there is a non-zero polynomial  $P$  of degree at most  $D$  which is a product of non-singular polynomials such that the set  $\mathbb{R}^3 \setminus Z(P)$  is a disjoint union of  $\sim D^3$  cells  $O_i$  such that, for every  $i$ ,

$$\int_{O_i \cap B_R} (Br_{\alpha} \mathcal{E}_{\phi} f)^{3.25} \sim D^{-3} \int_{B_R} (Br_{\alpha} \mathcal{E}_{\phi} f)^{3.25}. \tag{6.5}$$

We next define  $W$  as the  $R^{1/2+\delta}$  neighborhood of  $Z(P)$  and put  $O'_i := (O_i \cap B_R) \setminus W$ .

Moreover, note that if we apply Proposition 6.3 to  $f_{\tau}$  in place of  $f$  (what we shall usually do), then by property (a) in Proposition 6.3, for every tube  $T \in \mathbb{T}$  the function  $f_{\tau,T}$  is supported in an  $O(R^{-1/2})$  neighborhood of  $\tau$ . Following Guth, we define

$$\mathbb{T}_i := \{T \in \mathbb{T} : T \cap O'_i \neq \emptyset\}, \quad f_{\tau,i} := \sum_{T \in \mathbb{T}_i} f_{\tau,T}, \quad f_{\Delta,i} := \sum_{\tau \subset \Delta} f_{\tau,i}$$

$$\text{and } f_i := \sum_{\tau} f_{\tau,i}.$$

Then we can use the following analogue to [18, Lemma 3.2]:

**Lemma 6.4** *Each tube  $T \in \mathbb{T}$  lies in at most  $D + 1$  of the sets  $\mathbb{T}_i$ .*

We cover  $B_R$  with  $\sim R^{3\delta}$  balls  $B_j$  of radius  $R^{1-\delta}$ . Recall Definitions 3.3 and 3.4 from [18]:

**Definitions 6.5** a) We define  $\mathbb{T}_{j,tang}$  as the set of all tubes  $T \in \mathbb{T}$  that satisfy the following conditions:

$$T \cap W \cap B_j \neq \emptyset,$$

and if  $\xi \in Z(P)$  is any nonsingular point (i.e.,  $\nabla P(\xi) \neq 0$ ) lying in  $2B_j \cap 10T$ , then

$$\text{angle}(\nu(T), T_{\xi} Z(P)) \leq R^{-1/2+2\delta}.$$

Here,  $T_{\xi} Z(P)$  denotes the tangent space to  $Z(P)$  at  $\xi$ , and we recall that  $\nu(T)$  denotes the unit vector in direction of  $T$ . Accordingly, we define

$$f_{\tau,j,tang} := \sum_{T \in \mathbb{T}_{j,tang}} f_{\tau,T} \quad \text{and} \quad f_{j,tang} := \sum_{\tau} f_{\tau,j,tang}.$$



b) We define  $\mathbb{T}_{j,trans}$  as the set of all tubes  $T \in \mathbb{T}$  that satisfy the following conditions:

$$T \cap W \cap B_j \neq \emptyset,$$

and there exists a nonsingular point  $\zeta \in Z(P)$  lying in  $2B_j \cap 10T$ , so that

$$\text{angle}(v(T), T_\zeta Z(P)) > R^{-1/2+2\delta}.$$

Accordingly, we define

$$f_{\tau,j,trans} := \sum_{T \in \mathbb{T}_{j,trans}} f_{\tau,T} \quad \text{and} \quad f_{j,trans} := \sum_{\tau} f_{\tau,j,trans}.$$

We also recall Lemmas 3.5 and 3.6 in [18]:

**Lemma 6.6** *Each tube  $T \in \mathbb{T}$  belongs to at most  $\text{Poly}(D) = R^{O(\delta_{deg})}$  different sets  $\mathbb{T}_{j,trans}$ .*

**Lemma 6.7** *For each  $j$ , the number of different  $\theta$  so that  $\mathbb{T}_{j,tang} \cap \mathbb{T}(\theta) \neq \emptyset$  is at most  $R^{1/2+O(\delta)}$ .*

Note that the previous lemma makes use of the fact that the Gaussian curvature does not vanish on the surface  $\Sigma$ , so that the Gauß map is a diffeomorphism onto its image.

**Lemma 6.8** *Let  $\xi \in O'_i$ . Then, given our assumptions on  $R$  from Remarks 6.2, we have*

$$Br_\alpha \mathcal{E}_\phi f(\xi) \leq Br_{2\alpha} \mathcal{E}_\phi f_i(\xi) + R^{-900} \sum_{\tau} \|f_\tau\|_2.$$

**Proof** This is analogous to [13, Lemma 4.7]. The proof of that lemma, without any changes, also gives us Lemma 6.8. □

Following [18], we next define

$$\text{Bil}(\mathcal{E}_\phi f_{j,tang}) := \sum_{\tau_1, \tau_2 \text{ strongly separated}} |\mathcal{E}_\phi f_{\tau_1,j,tang}|^{1/2} |\mathcal{E}_\phi f_{\tau_2,j,tang}|^{1/2}.$$

The remaining part of this subsection will be devoted to the proof of the following crucial analogue to the key Lemma 3.8 in [18]:

**Lemma 6.9** *If  $\xi \in B_j \cap W$  and  $\alpha\mu \leq 10^{-5}$ , then*

$$Br_\alpha \mathcal{E}_\phi f(\xi) \leq 2 \left( \sum_I Br_{K^{4\tilde{n}_{e'}\alpha}} \mathcal{E}_\phi f_{I,j,trans}(\xi) + K^{100} \text{Bil}(\mathcal{E}_\phi f_{j,tang})(\xi) + R^{-900} \sum_{\tau} \|f_\tau\|_2 \right), \tag{6.6}$$

where the first sum is over all possible subsets  $I$  of the given family of caps  $\tau$ .

**Proof** Let  $\xi \in B_j \cap W$ . We may assume that  $\xi$  is  $\alpha$ -broad for  $\mathcal{E}_\phi f$  and that  $|\mathcal{E}_\phi f(\xi)| \geq R^{-900} \sum_\tau \|f\|_2$ . Let

$$I := \{\tau : |\mathcal{E}_\phi f_{\tau,j,tang}(\xi)| \leq K^{-100} |\mathcal{E}_\phi f(\xi)|\}. \tag{6.7}$$

We consider two possible cases:

**Case 1:**  $I^c$  contains two strongly separated caps  $\tau_1$  and  $\tau_2$ . Then trivially

$$|\mathcal{E}_\phi f(\xi)| \leq K^{100} |\mathcal{E}_\phi f_{\tau_1,j,tang}(\xi)|^{1/2} |\mathcal{E}_\phi f_{\tau_2,j,tang}(\xi)|^{1/2} \leq K^{100} \text{Bil}(\mathcal{E}_\phi f_{j,tan})(\xi),$$

hence (6.6).

**Case 2:**  $I^c$  does not contain two strongly separated caps.

We denote by  $\mathcal{L}(\xi) \subset \mathcal{L}$  the family of at most  $K^{3e'}$  strips respectively large caps given by the Geometric Lemma for the family of caps  $\mathcal{F} := I^c$ . By

$$J := \{\tau : \tau \subset L \text{ for some } L \in \mathcal{L}(\xi)\}$$

we denote the corresponding subset of caps  $\tau$ . Then  $I^c \subset J$ , i.e.,  $J^c \subset I$ . We write

$$f = \sum_{\tau \in J} f_\tau + \sum_{\tau \in J^c} f_\tau.$$

Hence,

$$|\mathcal{E}_\phi f(\xi)| \leq \left| \sum_{\tau \in J} \mathcal{E}_\phi f_\tau(\xi) \right| + \left| \sum_{\tau \in J^c} \mathcal{E}_\phi f_\tau(\xi) \right|.$$

For  $L \in \mathcal{L}$ , we denote by  $\tilde{L} := \{\tau : \tau \subset L\}$ . Note that  $f_L = \sum_{\tau \in \tilde{L}} f_\tau$ , and

$$J = \bigcup_{L \in \mathcal{L}(\xi)} \tilde{L}.$$

Thus, by the inclusion-exclusion principle and Remark 4.2,

$$\chi_J = \sum_{k=1}^{\overline{N}} (-1)^{k+1} \sum_{L_1, \dots, L_k \in \mathcal{L}(\xi)} \chi_{\tilde{L}_1 \cap \dots \cap \tilde{L}_k}. \tag{6.8}$$

Therefore,

$$\begin{aligned} \left| \sum_{\tau \in J} \mathcal{E}_\phi f_\tau(\xi) \right| &\leq \sum_{k=1}^{\bar{N}} \sum_{L_1, \dots, L_k \in \mathcal{L}(\xi)} \left| \sum_{\tau \in \tilde{L}_1 \cap \dots \cap \tilde{L}_k} \mathcal{E}_\phi f_\tau(\xi) \right| \\ &= \sum_{k=1}^{\bar{N}} \sum_{L_1, \dots, L_k \in \mathcal{L}(\xi)} |\mathcal{E}_\phi f_{L_1 \cap \dots \cap L_k}(\xi)|. \end{aligned}$$

Since  $\xi$  is  $\alpha$ -broad, and since according to Remark 6.2 we may assume that  $\alpha \leq K^{-\varepsilon/2}$ , this can be further estimated by

$$\begin{aligned} &\leq \sum_{k=1}^{\bar{N}} |\mathcal{L}(\xi)|^k \alpha |\mathcal{E}_\phi f(\xi)| \leq \bar{N} (K^{3\varepsilon'})^{\bar{N}} \alpha |\mathcal{E}_\phi f(\xi)| \\ &\leq \bar{N} K^{3\varepsilon' \bar{N}} K^{-\varepsilon/2} |\mathcal{E}_\phi f(\xi)| \leq \frac{1}{10} |\mathcal{E}_\phi f(\xi)|, \end{aligned}$$

as one can easily see by our choices of  $K$  in (6.4) and  $\varepsilon'$  in (4.3), provided  $\varepsilon > 0$  is assumed to be sufficiently small. Thus,

$$|\mathcal{E}_\phi f(\xi)| \leq \frac{1}{10} |\mathcal{E}_\phi f(\xi)| + \left| \sum_{\tau \in J^c} \mathcal{E}_\phi f_\tau(\xi) \right|,$$

and therefore

$$|\mathcal{E}_\phi f(\xi)| \leq \frac{10}{9} \left| \sum_{\tau \in J^c} \mathcal{E}_\phi f_\tau(\xi) \right|.$$

Since  $\xi \in B_j \cap W$ , by Proposition 6.3,

$$\mathcal{E}_\phi f_\tau(\xi) = \mathcal{E}_\phi f_{\tau, j, trans}(\xi) + \mathcal{E}_\phi f_{\tau, j, tang}(\xi) + O(R^{-1000}) \|f_\tau\|_2. \tag{6.9}$$

Moreover, since  $J^c \subset I$ , and since there are at most  $K^2$  caps  $\tau$ ,

$$\begin{aligned} \sum_{\tau \in J^c} |\mathcal{E}_\phi f_{\tau, j, tang}(\xi)| &\leq \sum_{\tau \in I} |\mathcal{E}_\phi f_{\tau, j, tang}(\xi)| \leq K^{-100} \sum_{\tau \in I} |\mathcal{E}_\phi f(\xi)| \\ &\leq K^{-98} |\mathcal{E}_\phi f(\xi)|, \end{aligned} \tag{6.10}$$

where the second inequality is a consequence of the definition of  $I$ . Thus,

$$\begin{aligned} \frac{9}{10} |\mathcal{E}_\phi f(\xi)| &\leq \left| \sum_{\tau \in J^c} \mathcal{E}_\phi f_{\tau, j, trans}(\xi) \right| + K^{-98} |\mathcal{E}_\phi f(\xi)| + \sum_{\tau} R^{-1000} \|f_\tau\|_2 \\ &= |\mathcal{E}_\phi f_{J^c, j, trans}(\xi)| + K^{-98} |\mathcal{E}_\phi f(\xi)| + \sum_{\tau} R^{-1000} \|f_\tau\|_2, \end{aligned}$$

and hence, since  $|\mathcal{E}_\phi f(\xi)| \geq R^{-900} \sum_\tau \|f_\tau\|_2$ ,

$$|\mathcal{E}_\phi f(\xi)| \leq \frac{11}{9} |\mathcal{E}_\phi f_{J^c, j, trans}(\xi)|. \quad (6.11)$$

It will then finally suffice to show that  $\xi$  is  $K^{4\tilde{N}\varepsilon'}$ - $\alpha$ -broad for  $\mathcal{E}_\phi g$ , where  $g := f_{J^c, j, trans}$ . To this end let us set  $g_\tau := f_{\tau, j, trans}$ , if  $\tau \in J^c$ , and zero otherwise, so that

$$g = \sum g_\tau.$$

In what follows, we shall use the short hand notation “neglig” for terms which are much smaller than  $R^{-940} \sum_\tau \|f_\tau\|_2$ .

Observe first that by (6.9)

$$|\mathcal{E}_\phi f_{\tau, j, trans}(\xi)| \leq |\mathcal{E}_\phi f_\tau(\xi)| + |\mathcal{E}_\phi f_{\tau, j, tang}(\xi)| + \text{neglig},$$

so that if  $\tau \in J^c \subset I$ , then by the definition of  $I$ ,

$$|\mathcal{E}_\phi f_{\tau, j, trans}(\xi)| \leq |\mathcal{E}_\phi f_\tau(\xi)| + K^{-100} |\mathcal{E}_\phi f(\xi)| + \text{neglig}.$$

We have to show that

$$|\mathcal{E}_\phi g_\Delta(\xi)| \leq K^{4\tilde{N}\varepsilon'} \alpha |\mathcal{E}_\phi g(\xi)|$$

for all  $\Delta \in \bar{\mathcal{L}}$ . Write  $\Delta = L_1 \cap \dots \cap L_r$ , where  $L_i \in \mathcal{L}$ , and set  $\tilde{\Delta} := \tilde{L}_1 \cap \dots \cap \tilde{L}_r$ , so that  $g_\Delta = \sum_{\tau \in \tilde{\Delta}} g_\tau = \sum_{\tau \in \tilde{\Delta} \cap J^c} f_{\tau, j, trans}$ . Therefore the following two cases can arise:

- (i) There is some  $i = 1, \dots, r$ , such that  $L_i \in \mathcal{L}(\xi)$ . Then,  $\tilde{\Delta} \subset \tilde{L}_i \subset J$ , hence,  $\tilde{\Delta} \cap J^c = \emptyset$ .
- (ii) For all  $i = 1, \dots, r$ ,  $L_i \notin \mathcal{L}(\xi)$ .

Observe first that by summing (6.9) over all  $\tau \in \tilde{\Delta} \cap J^c$  we obtain

$$|\mathcal{E}_\phi g_\Delta(\xi)| = \left| \sum_{\tau \in \tilde{\Delta}} \mathcal{E}_\phi g_\tau \right| \leq \left| \sum_{\tau \in \tilde{\Delta} \cap J^c} \mathcal{E}_\phi f_\tau(\xi) \right| + \sum_{\tau \in \tilde{\Delta} \cap J^c} |\mathcal{E}_\phi f_{\tau, j, tang}(\xi)| + \text{neglig}. \quad (6.12)$$

By (6.10), the second term can again be estimated by

$$\sum_{\tau \in \tilde{\Delta} \cap J^c} |\mathcal{E}_\phi f_{\tau, j, tang}(\xi)| \leq K^{-98} |\mathcal{E}_\phi f(\xi)|.$$

Case (i) is thus trivial. In case (ii), we write

$$\sum_{\tau \in \tilde{\Delta} \cap J^c} \mathcal{E}_\phi f_\tau(\xi) = \mathcal{E}_\phi f_\Delta(\xi) - \sum_{\tau \in \tilde{\Delta} \cap J} \mathcal{E}_\phi f_\tau(\xi).$$

The first term is estimated using broadness. For the second term, again by (6.8),

$$\begin{aligned} \left| \sum_{\tau \in \tilde{\Delta} \cap J} \mathcal{E}_\phi f_\tau(\xi) \right| &\leq \sum_{k=1}^{\tilde{N}} \sum_{L'_1, \dots, L'_k \in \mathcal{L}(\xi)} \left| \sum_{\tau \in \tilde{\Delta} \cap \tilde{L}'_1 \cap \dots \cap \tilde{L}'_k} \mathcal{E}_\phi f_\tau(\xi) \right| \\ &= \sum_{k=1}^{\tilde{N}} \sum_{L'_1, \dots, L'_k \in \mathcal{L}(\xi)} |\mathcal{E}_\phi f_{\Delta \cap L'_1 \cap \dots \cap L'_k}(\xi)|. \end{aligned}$$

Note that  $\Delta \cap L'_1 \cap \dots \cap L'_k \in \bar{\mathcal{L}}$ , so that, since  $\xi$  is  $\alpha$ -broad for  $\mathcal{E}_\phi f$ ,

$$\left| \sum_{\tau \in \tilde{\Delta} \cap J} \mathcal{E}_\phi f_\tau(\xi) \right| \leq \sum_{k=1}^{\tilde{N}} |\mathcal{L}(\xi)|^k \alpha |\mathcal{E}_\phi f(\xi)| \leq \tilde{N} K^{3\tilde{N}} \alpha |\mathcal{E}_\phi f(\xi)|.$$

Since  $\alpha \geq K^{-\varepsilon} \gg 10K^{-98}$ , in combination with (6.12), and in the last step with (6.11), we conclude that

$$|\mathcal{E}_\phi g_\Delta(\xi)| \leq (\tilde{N} K^{3\tilde{N}\varepsilon'} + 1) \alpha |\mathcal{E}_\phi f(\xi)| + K^{-98} |\mathcal{E}_\phi f(\xi)| + \text{neglig} \leq K^{4\tilde{N}\varepsilon'} \alpha |\mathcal{E}_\phi g(\xi)|.$$

This completes the proof of Lemma 6.9. □

The contribution by the bilinear term in (6.6) will be controlled by means of the following analogue to [13, Proposition 4.13 ] (or [18, Proposition 3.9 ]):

**Proposition 6.10** *We have*

$$\int_{B_j \cap W} \text{Bil}(\mathcal{E}_\phi f_{j,tang})^{3.25} \leq C_\varepsilon R^{O(\delta)+\varepsilon/2} \left( \sum_\tau \int |f_\tau|^2 \right)^{3/2+\varepsilon}.$$

With Proposition 6.10 at hand, the rest of the proof of Theorem 6.1, which we shall detail in Sect. 6.1, will follow the arguments in Section 4.2 in [13] (which in return are an adaptation of the arguments in pages 396–398 of [18]).

The proof of Proposition 6.10 reduces to the following analogue to [13, Lemma 4.14] and [18, Lemma 3.10]. Suppose we have covered  $B_j \cap W$  with a minimal number of cubes  $Q$  of side length  $R^{1/2}$ , and denote by  $\mathbb{T}_{j,tang,Q}$  the set of all tubes  $T$  in  $\mathbb{T}_{j,tang}$  such that  $10T$  intersects  $Q$ .

**Lemma 6.11** Fix  $j$ , i.e., a ball  $B_j$ . If  $\tau_1, \tau_2$  are strongly separated caps, then for any of the cubes  $Q$  we have

$$\int_Q |\mathcal{E}_\phi f_{\tau_1, j, \text{tang}}|^2 |\mathcal{E}_\phi f_{\tau_2, j, \text{tang}}|^2 \leq R^{O(\delta)} R^{-1/2} \left( \sum_{T_1 \in \mathbb{T}_{j, \text{tang}, Q}} \|f_{\tau_1, T_1}\|_2^2 \right) \left( \sum_{T_2 \in \mathbb{T}_{j, \text{tang}, Q}} \|f_{\tau_2, T_2}\|_2^2 \right) + \text{neglig.}$$

**Proof** Using Remark 3.3, the proof of Lemma 4.15 in [13] can be repeated word by word, giving the result. □

### 6.1 Completing the proof of Theorem 6.1

If we compare with Subsection 4.2 in [13], which deals with the induction arguments with respect to the size of the radius  $R$ , and the size of  $\sum_\tau \int |f_\tau|^2$ , we can see that Lemma 6.9, which was the only result whose proof required substantial new arguments compared to the corresponding result in [13], is needed only for the discussion of Case 2 in this subsection, i.e., the case where the dominating term in

$$\int_{B_R} (Br_\alpha \mathcal{E}_\phi f)^{3.25} = \sum_i \int_{B_R \cap O'_i} (Br_\alpha \mathcal{E}_\phi f)^{3.25} + \int_{B_R \cap W} (Br_\alpha \mathcal{E}_\phi f)^{3.25}$$

is the second term, the ‘‘wall’’ term. The estimation of this term could then be reduced to controlling  $\sum_j \int_{B_j \cap W} \sum_I (Br_{150\alpha} \mathcal{E}_\phi f_{I, j, \text{trans}})^{3.25}$ , which in view of our Lemma 6.9 here has to be modified to

$$\sum_j \int_{B_j \cap W} \sum_I (Br_{K^{4\bar{N}\epsilon'} \alpha} \mathcal{E}_\phi f_{I, j, \text{trans}})^{3.25}. \tag{6.13}$$

Dealing with this term by induction as in that paper, we arrive at

$$\begin{aligned} & \int_{B_R} (Br_{K^{4\bar{N}\epsilon'} \alpha} E f)^{3.25} \\ & \leq M_\epsilon C_\epsilon \text{Poly}(D) R^{\epsilon(1-\delta)} \left( \sum_\tau \int |f_\tau|^2 \right)^{3/2+\epsilon} R^{\delta_{\text{trans}}(1-\delta) \log(K^{4\bar{N}\epsilon'} K^\epsilon \alpha \mu)} \\ & \leq \left( M_\epsilon \text{Poly}(R^{\delta_{\text{deg}}}) R^{\delta_{\text{trans}} 4\bar{N}\epsilon' \log(K) - \epsilon \delta} \right) \\ & \quad \times C_\epsilon R^\epsilon \left( \sum_\tau \int |f_\tau|^2 \right)^{3/2+\epsilon} R^{\delta_{\text{trans}} \log(K^\epsilon \alpha \mu)} \\ & \leq C_\epsilon R^\epsilon \left( \sum_\tau \int |f_\tau|^2 \right)^{3/2+\epsilon} R^{\delta_{\text{trans}} \log(K^\epsilon \alpha \mu)}. \end{aligned}$$

For the last inequality, note that the choices of  $\delta$ ,  $\delta_{deg}$ ,  $\delta_{trans}$  in (6.3) and  $K$  in (6.4), in combination with (4.3), ensure that for  $\varepsilon$  sufficiently small

$$C\delta_{deg} + \delta_{trans}4\bar{N}\varepsilon' \log(K) - \delta\varepsilon = C\varepsilon^4 + \varepsilon^64\bar{N}\varepsilon'\varepsilon^{-10} - \varepsilon^3 < -\varepsilon^3/2.$$

Here,  $M_\varepsilon = 2K^2$ . Then, by Remark 6.2,  $M_\varepsilon R^{-\varepsilon^3/2} \ll 1$ . This completes the proof of Theorem 6.1.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Barron, A.: Restriction estimates for hyperboloids in higher dimensions via bilinear estimates. [arXiv:2002.09001](https://arxiv.org/abs/2002.09001) Preprint (2020)
2. Bennett, J., Carbery, A., Tao, T.: On the multilinear restriction and Kakeya conjectures. *Acta Math.* **196**(2), 261–302 (2006)
3. Bourgain, J.: Besicovitch-type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.* **22**, 147–187 (1991)
4. Bourgain, J.: Some New Estimates on Oscillatory Integrals. *Essays in Fourier Analysis in Honor of E. M. Stein*. Princeton Mathematical Series, vol. 42, pp. 83–112. Princeton University Press, Princeton (1995)
5. Bourgain, J.: Estimates for cone multipliers. *Oper. Theory Adv. Appl.* **77**, 1–16 (1995)
6. Bourgain, J., Guth, L.: Bounds on oscillatory integral operators based on multilinear estimates. *Geom. Funct. Anal.* **21**, 1239–1295 (2011)
7. Bruce, B. B.: Local extension estimates for the hyperbolic hyperboloid in three dimensions. [arXiv: 2005.12467v2](https://arxiv.org/abs/2005.12467v2) Preprint (2020)
8. Bruce, B. B.: Global restriction estimates for elliptic hyperboloids. [arXiv: 2008.01005v1](https://arxiv.org/abs/2008.01005v1) Preprint (2020)
9. Bruce, B. B., Oliveira e Silva, D., Stovall, B.: Restriction inequalities for the hyperbolic hyperboloid. [arXiv: 2007.06990v1](https://arxiv.org/abs/2007.06990v1) Preprint (2020)
10. Buschenhenke, S., Müller, D., Vargas, A.: A Fourier restriction theorem for a perturbed hyperbolic paraboloid. *Proc. Lond. Math. Soc.* (3) **120**(1), 124–154 (2020)
11. Buschenhenke, S., Müller, D., Vargas, A.: On Fourier restriction for finite-type perturbations of the hyperbolic paraboloid. *Proceedings of the conference “Geometric Aspects of Harmonic Analysis”, Cortona, June 25–29; Springer INdAM Series.* [arXiv:1902.05442v2](https://arxiv.org/abs/1902.05442v2) (2018)
12. Buschenhenke, S., Müller, D., Vargas, A.: Partitions of flat one-variate functions and a Fourier restriction theorem for related perturbations of the hyperbolic paraboloid. *J. Geom. Anal.* (to appear). [arXiv:2002.08726](https://arxiv.org/abs/2002.08726)
13. Buschenhenke, S., Müller, D., Vargas, A.: A Fourier restriction theorem for a perturbed hyperbolic paraboloid: polynomial partitioning. [arXiv:2003.01619v1](https://arxiv.org/abs/2003.01619v1) (to appear *Mathematische Zeitschrift*) (2020)
14. Cho, C.-H., Lee, J.: Improved restriction estimate for hyperbolic surfaces in  $\mathbb{R}^3$ . *J. Funct. Anal.* **273**(3), 917–945 (2017)
15. Fefferman, C.: Inequalities for strongly singular convolution operators. *Acta Math.* **25**, 9–36 (1970)

16. Greenleaf, A.: Principal curvature and harmonic analysis. *Indiana Univ. Math. J.* **30**, 4 (1981)
17. Guo, S., Oh, C.: A restriction estimate for polynomial surfaces with negative Gaussian curvatures. [arXiv:2005.12431v1](https://arxiv.org/abs/2005.12431v1) Preprint (2020)
18. Guth, L.: A restriction estimate using polynomial partitioning. *J. Am. Math. Soc.* **29**(2), 371–413 (2016)
19. Guth, L.: Restriction estimates using polynomial partitioning II. *Acta Math.* **221**(1), 81–142 (2018)
20. Hickman, J., Rogers, K.: Improved Fourier restriction estimates in higher dimensions. *Camb. J. Math.* **7**(3), 219–282 (2019)
21. Ikromov, I.A., Kempe, M., Müller, D.: Estimates for maximal functions associated with hypersurfaces in  $\mathbb{R}^3$  and related problems in harmonic analysis. *Acta Math.* **204**, 151–271 (2010)
22. Ikromov, I.A., Müller, D.: Uniform estimates for the Fourier transform of surface carried measures in  $\mathbb{R}^3$  and an application to Fourier restriction. *J. Fourier Anal. Appl.* **17**(6), 1292–1332 (2011)
23. Ikromov, I.A., Müller, D.: *Fourier Restriction for Hypersurfaces in Three Dimensions and Newton Polyhedra*. *Annals of Mathematics Studies*, vol. 194. Princeton University Press, Princeton (2016)
24. Kim, J.: Some remarks on Fourier restriction estimates. [arXiv:1702.01231](https://arxiv.org/abs/1702.01231) Preprint (2017)
25. Lee, S.: Bilinear restriction estimates for surfaces with curvatures of different signs. *Trans. Am. Math. Soc.* **358**(8), 3511–3533 (2005)
26. Lee, S., Vargas, A.: Restriction estimates for some surfaces with vanishing curvatures. *J. Funct. Anal.* **258**(9), 2884–2909 (2010)
27. Moyua, A., Vargas, A., Vega, L.: Schrödinger maximal function and restriction properties of the Fourier transform. *Int. Math. Res. Not.* **16**, 793–815 (1996)
28. Moyua, A., Vargas, A., Vega, L.: Restriction theorems and maximal operators related to oscillatory integrals in  $\mathbb{R}^3$ . *Duke Math. J.* **96**(3), 547–574 (1999)
29. Spindler, K.: A Short Proof of the Formula of Faj Bruno, *Elemente der Mathematik*, pp. 33–35. Swiss Mathematical Society, Basel (2005)
30. Stein, E.M.: *Oscillatory integrals in Fourier analysis*. Beijing lectures in harmonic analysis. Princeton Univ. Press 1986. Symmetric hypersurfaces. *Trans. Am. Math. Soc.* **369**(6), 4093–4117 (2017)
31. Stovall, B.: Scale invariant Fourier restriction to a hyperbolic surface. *Anal. PDE* **12**(5), 1215–1224 (2019)
32. Strichartz, R.S.: Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.* **44**(3), 705–714 (1977)
33. Tao, T.: A Sharp bilinear restriction estimate for paraboloids. *Geom. Funct. Anal.* **13**, 1359–1384 (2003)
34. Tao, T.: The Bochner–Riesz conjecture implies the restriction conjecture. *Duke Math. J.* **96**(2), 363–375 (1999)
35. Tao, T., Vargas, A., Vega, L.: A bilinear approach to the restriction and Keakeya conjectures. *J. Am. Math. Soc.* **11**(4), 967–1000 (1998)
36. Tao, T., Vargas, A.: A bilinear approach to cone multipliers I. Restriction estimates. *Geom. Funct. Anal.* **10**, 185–215 (2000)
37. Tao, T., Vargas, A.: A bilinear approach to cone multipliers II. Applications. *Geom. Funct. Anal.* **10**, 216–258 (2000)
38. Tomas, P.A.: A restriction theorem for the Fourier transform. *Bull. Am. Math. Soc.* **81**, 477–478 (1975)
39. Vargas, A.: Restriction theorems for a surface with negative curvature. *Math. Z.* **249**, 97–111 (2005)
40. Wang, H.: A restriction estimate in  $R^3$  using brooms. [arXiv:1802.04312](https://arxiv.org/abs/1802.04312)
41. Wolff, T.: A sharp bilinear cone restriction estimate. *Ann. Math. Sec. Ser.* **153**(3), 661–698 (2001)
42. Zygmund, A.: On Fourier coefficients and transforms of functions of two variables. *Stud. Math.* **50**, 189–201 (1974)