

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Mathematics

2017

Perturbed Brascamp-Lieb inequalities and application to Parsell-Vinogradov systems

Zhang, Ruixiang

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp0137720g37s>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

PERTURBED BRASCAMP-LIEB INEQUALITIES
AND APPLICATION TO PARSELL-VINOGRADOV
SYSTEMS

RUIXIANG ZHANG

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MATHEMATICS
ADVISER: PETER SARNAK

SEPTEMBER 2017

© Copyright by Ruixiang Zhang, 2017.

All Rights Reserved

Abstract

In this thesis, we study the perturbed Brascamp-Lieb inequalities and its applications in translation-dilation systems. We prove the endpoint perturbed Brascamp-Lieb inequalities using polynomial partition techniques. We also look at the Parsell-Vinogradov system and verify the Brascamp-Lieb condition holds in its decoupling approach. As a corollary of this and the work of Guo, the main conjecture about the system is true in dimension 2 and can be proved by the decoupling approach.

Acknowledgements

I would like to thank my advisor Professor Peter Sarnak, for always being so insightful and encouraging. I was lucky enough to know him from the beginning of my life at Princeton. Without a doubt, his support and encouragement have played an indispensable role for my academic progress at Princeton over the years. And I believe that I would continue being inspired by his vision in the future years.

I would like to thank Professors James Arthur, Jean Bourgain, Anthony Carbery, William Casselman, Ciprian Demeter, Manfred Einsiedler, Charles Fefferman, Zihua Guo, Larry Guth, Elon Lindenstrauss, Simon Marshall, Werner Müller, Terence Tao and Nicholas Templier, all for having very important impact on some stage of my graduate study. It is hard to imagine where I would have been without talking to them.

I would also like to sincerely thank Professors Jonathan Bennett, Jordan Ellenberg, Nets Katz, Wen-Wei Li, Elliott Lieb, Birgit Speh, Elias Stein, Gang Tian, Yijun Yao, Trevor Wooley, Pin Yu and Po-Lam Yung for their help.

Special thanks go to my following friends in math: Jun Su for many inspiring conversations on automorphic forms while being so patient with me; Bin Xu for the great help on reading Arthur's trace formula papers and being an extremely reliable friend; Chenjie Fan for a huge amount of enlightening conversations on graduate school experience and mathematics; Yu Deng for technical and mental support during my last year; Hong Wang for a great amount of fruitful discussions on harmonic analysis; Yi Hu for many inspiring discussions on discrete restriction; Shaoming Guo for being a great friend and collaborator; Kevin Hughes for the great help in my

first year; Jasmin Matz for many enlightening conversations on the trace formula and Naser Sardari for being so helpful on basic knowledge of representation and automorphic forms in my early years at Princeton.

I would like to thank all the staff members of the mathematics department at Princeton University, whose help made my days at Princeton much more enjoyable.

I would like to thank my friends Xinliang An, Yumeng Chen, Mike Fouchet, Ziyang Gao, Jiequn Han, Han Hao, Hang Huang, Peter Humphries, Marina Ilioupoulou, Long Jin, Ning Jin, Daniel Kriz, Yuming Kuang, Junchi Li, Wanlin Li, Wei Li, Xiaohan Li, Yue Li, Zane Li, Bo Lin, Yuchen Liu, Hongliang Lu, Jiarui Lu, Tong Lv, Xiaosheng Mu, Liangming Shen, Lei Shi, Xiao Shi, Lve Pan, Bo Qin, Will Sawin, Yunqing Tang, Ao Wang, Zhiyu Wang, Fan Wei, Zhentian Wei, Percy Wong, Xingye Wu, Jingyan Xu, Yuanda Xu, Jerry Xue, Ben Yang, Yucheng Yang, Jian Zhang, Jing Zhang, Zili Zhang and Bo Zhao, with whom my graduate life became considerably more delightful and less suffering.

Finally, I would like to thank my family. To my grandfather Yueren Zhang, my grandmother Lijun Zhou, my father Qingsong Zhang and my mother Cuifang Dang for their ample support and patience throughout my lifetime.

To my family.

*I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I –
I took the one less traveled by,
And that has made all the difference.*
— Robert Frost, *the Road not Taken*

Contents

Abstract	iii
Acknowledgements	iv
1 Introduction	1
1.1 Notations	1
1.2 The Endpoint Perturbed Brascamp-Lieb Inequalities	1
1.3 Application of the Perturbed Brascamp-Lieb Inequalities in the Study of Vinogradov Mean Value Theorem and Parsell-Vinogradov Systems	6
2 Recent Developments of the Polynomial Method in Incidence Ge- ometry Type Problems	10
3 Motivation and Ideas of the Proof of the Endpoint Perturbed Bracamp- Lieb Inequality	15
3.1 The Endpoint Multilinear k_j -Plane Theorem	15
3.2 The Endpoint Perturbed Brascamp-Lieb Inequalities (Continued) . .	20
3.3 Ideas of the Proofs	21

4	Polynomials with High Visibility	27
5	Wedge Product Estimate Based on Visibility	31
6	Linear Algebra Preliminaries	35
7	Integral Geometry Preliminaries	39
8	Proof of Theorem 3.1.2 and Theorem 3.1.3	44
9	An Analogue of Lemma 6.0.8	51
10	Proof of Theorem 1.2.7	57
11	Verifying the Brascamp-Lieb Condition for Parsell-Vinogradov Systems	63

Chapter 1

Introduction

1.1 Notations

In this thesis, we use C to denote positive universal constants that may differ from place to place. We use $X \lesssim Y$ or $X \ll Y$ to denote $X = O(Y)$ for $X, Y > 0$. I keep both in the thesis as they are conventions from the analysis and number theory sides, respectively. We use those notations with footnotes like C_ε or $\ll_\varepsilon$ to feature the dependence on variables.

1.2 The Endpoint Perturbed Brascamp-Lieb Inequalities

The celebrated Brascamp-Lieb inequalities, first introduced by these authors in [BL76], is a multilinear inequality that generalizes many inequalities in analysis such

as Holder's inequality, Young's inequality for convolutions and the Loomis-Whitney inequality. We start by briefly reviewing this very general inequality.

We will mostly follow the notational convention in [Lie90], [BCCT08] and [BCCT10], which are three important references in literature. Assume that in $\mathbb{R}^d$ we have n linear surjections $B_j : \mathbb{R}^d \rightarrow E_j$. Then for certain positive numbers $p_j, 1 \leq j \leq n$, the following Brascamp-Lieb Inequality holds for any measurable function f_j on $E_j (1 \leq j \leq n)$ with some finite constant $C > 0$:

$$\int_{\mathbb{R}^d} \prod_{j=1}^n |f_j \circ B_j|^{p_j} \leq C \prod_{j=1}^n \left(\int_{E_j} |f_j|^{p_j} \right). \quad (1.1)$$

If this is the case, we call the minimal constant C such that (1.1) holds to be the *Brascamp-Lieb constant* $BL(\mathbf{B}, \mathbf{p})$. Here we use $\mathbf{B}$ to denote the data $(B_1, \dots, B_n)$ and $\mathbf{p}$ to denote the data $(p_1, \dots, p_n)$. $(\mathbf{B}, \mathbf{p})$ is called the corresponding *Brascamp-Lieb datum*. If (1.1) fails for any finite C , we define $BL(\mathbf{B}, \mathbf{p}) = +\infty$.

Lieb [Lie90] showed

Theorem 1.2.1. $BL(\mathbf{B}, \mathbf{p}) = BL_g(\mathbf{B}, \mathbf{p})$ where

$$BL_g(\mathbf{B}, \mathbf{p}) = \sup_{A_j: E_j \rightarrow E_j \text{ is a positive definite linear transform}} \left(\frac{\prod_{j=1}^n (\det_{E_j} A_j)^{p_j}}{\det(\sum_{j=1}^n p_j B_j^* A_j B_j)} \right)^{\frac{1}{2}}. \quad (1.2)$$

An alternative way to state Theorem 1.2.1 is that the Brascamp-Lieb constant is what one would obtain by restricting his attention to the special case in which each f_j is a certain Gaussian function.

Subsequently, Bennett-Carbery-Christ-Tao [BCCT08][BCCT10] determined a necessary and sufficient condition for $BL(\mathbf{B}, \mathbf{p}) = BL_g(\mathbf{B}, \mathbf{p}) < +\infty$. They proved that

$BL(\mathbf{B}, \mathbf{p}) = BL_g(\mathbf{B}, \mathbf{p}) < +\infty$ is equivalent to the following two conditions:

(1) Scaling condition:

$$\sum_j p_j \dim E_j = d; \quad (1.3)$$

and (2) Dimension condition: For any linear subspace $V \subseteq \mathbf{R}^d$,

$$\dim V \leq \sum_j p_j \dim(B_j V). \quad (1.4)$$

So we know when we can have the actual Brascamp-Lieb Inequality (1.1) thanks to their work.

(1.1) has an equivalent version that is easier to understand intuitively. We refer the readers to [Ben] for this observation. Before stating this equivalent form we introduce some terminology to describe “higher dimensional” analogue of cylinders.

Definition 1.2.2. *In a space of dimension d , for any $1 \leq b < d$ define a b -slab to be the Cartesian product of a b -dimensional ball B_1 and a $(d - b)$ -dimensional ball B_2 (the spaces spanned by both balls are required to be orthogonal). The radius of B_1 will be called the size of our b -slab and the radius of B_2 will be called the radius of it. The Cartesian product of B_1 and the center of B_2 is called the core of this b -slab.*

Proposition 1.2.3 (Combinatorial Brascamp-Lieb). *Assume we have a Brascamp-Lieb datum $(\mathbf{B}, \mathbf{p})$ in $\mathbb{R}^d$. Assume $k_j = \dim \ker B_j$ and we have n families of slabs. Assume the j -th family $\mathbb{T}_j$ consists of only k_j -slabs of radius 1 whose cores are all parallel to $\ker B_j$. Also assume each $|\mathbb{T}_j|$ is finite. Then $BL(\mathbf{B}, \mathbf{p}) < +\infty$ if and only*

if we always have

$$\int_{\mathbb{R}^d} \prod_{j=1}^n \left(\sum_{T_j \in \mathbb{T}_j} \chi_{T_j} \right)^{p_j} \lesssim \prod_{j=1}^n |\mathbb{T}_j|^{p_j}. \quad (1.5)$$

This form of the inequality allows perturbation: When (1.5) holds we can perturb the direction of each slab in a non-uniform way and ask if (1.5) still holds. Pioneering pieces of work was done by Bennett-Carbery-Tao (non-endpoint case, [BCT06]) and Guth (endpoint case, [Gut10]), concluding that in the Loomis-Whitney case, (1.5) continues to hold as long as perturbation within each family is uniformly small. Their theorems are usually called multilinear Kakeya estimates and have already found remarkable applications to classical harmonic analysis problems as well[BG11][Bou13b][Bou13a][Gut14b][BD15]. See the beginning of [Gut15] for a good introduction.

Theorem 1.2.4 (Multilinear Kakeya Theorem with R^ε Loss[BCT06]). *Assume R is a large positive number. For $1 \leq j \leq d$, let $\{T_{j,a} : 1 \leq a \leq A(j)\}$ be a family of unit cylinders of size $\leq R$ in $\mathbb{R}^d$. Assume the core line of each cylinder from the j -th family form an angle $\leq \delta$ with the j -th coordinate axis.*

Then when $\delta > 0$ is sufficiently small depending on d , we have

$$\int_{\mathbb{R}^d} \left(\prod_{j=1}^d \sum_{a=1}^{A(j)} \chi_{T_{j,a}} \right)^{\frac{1}{d-1}} \lesssim_{\varepsilon,d} R^\varepsilon \prod_{j=1}^d A(j)^{\frac{1}{d-1}}. \quad (1.6)$$

Theorem 1.2.5 (Endpoint Multilinear Kakeya Theorem [Gut10]). *For $1 \leq j \leq d$, let $\{T_{j,a} : 1 \leq a \leq A(j)\}$ be a family of unit cylinders in $\mathbb{R}^d$, we denote $v_{j,a}$ to be the direction of the core line of cylinder $T_{j,a}$. Assume the core lines of cylinders from different families are “quantitatively transversal”, i.e. for any $1 \leq a_j \leq A(j)$, we*

have $v_{1,a_1} \wedge v_{2,a_2} \wedge \cdots \wedge v_{d,a_d} \geq \theta > 0$ where θ is fixed. Then we have

$$\int_{\mathbb{R}^d} \left(\prod_{j=1}^d \sum_{a=1}^{A(j)} \chi_{T_{j,a}} \right)^{\frac{1}{d-1}} \lesssim_d \theta^{-\frac{1}{d-1}} \prod_{j=1}^d A(j)^{\frac{1}{d-1}}. \quad (1.7)$$

Theorems 1.2.4 and 1.2.5 suggest that it is possible to have a perturbed version of the more general Brascamp-Lieb inequalities. This is indeed true. Recently, Bennett, Bez, Flock and Lee [BBFL15] proved the following (non-endpoint) theorem (Theorem 1.2 of [BBFL15]) via generalizations of Guth's method in [Gut15].

Theorem 1.2.6 (Perturbed Brascamp-Lieb Inequalities with R^ε -loss[BBFL15]). *Assume we have a Brascamp-Lieb datum $(\mathbf{B}, \mathbf{p})$ in $\mathbb{R}^d$ with $BL(\mathbf{B}, \mathbf{p}) < +\infty$. Let $k_j = \dim \ker B_j$. Assume we have n families of slabs and the j -th family $\mathbb{T}_j$ consists of only k_j -slabs of radius 1 and size $\leq R$. Assume each $|\mathbb{T}_j|$ is finite. Also assume that each slab in the j -th family has its core k_j -plane within a δ -neighborhood of $\ker B_j$ on the corresponding Grassmannian (with a given standard metric). Then when δ is sufficiently small depending on $(\mathbf{B}, \mathbf{p})$ we have*

$$\int_{\mathbb{R}^d} \prod_{j=1}^n \left(\sum_{T_j \in \mathbb{T}_j} \chi_{T_j} \right)^{p_j} \lesssim_{d, \mathbf{p}, BL(\mathbf{B}, \mathbf{p}), \varepsilon} R^\varepsilon \prod_{j=1}^n |\mathbb{T}_j|^{p_j}. \quad (1.8)$$

They conjectured that R^ε can be removed (see inequality (7) and (8) of [BBFL15]) and in this thesis we prove their conjecture.

Theorem 1.2.7 (Endpoint Perturbed Brascamp-Lieb Inequalities). *Assumptions are the same as Theorem 1.2.6, but no restriction on the size of slabs. When δ is*

sufficiently small depending on $(\mathbf{B}, \mathbf{p})$ we have

$$\int_{\mathbb{R}^d} \prod_{j=1}^n \left(\sum_{T_j \in \mathbb{T}_j} \chi_{T_j} \right)^{p_j} \lesssim_{d, \mathbf{p}, BL(\mathbf{B}, \mathbf{p})} \prod_{j=1}^n |\mathbb{T}_j|^{p_j}. \quad (1.9)$$

Remark 1.2.8. *Theorem 1.2.7 formally implies the stability of Brascamp-Lieb constants which was also a result of Bennett-Bez-Flock-Lee [BBFL15]. However it is worth noticing (see the proof later in Section 8) that the main result in [BBFL15] is an input rather than an output in our proof of Theorem 1.2.7. In particular, we do not have a new proof of the main result in [BBFL15] in this thesis.*

1.3 Application of the Perturbed Brascamp-Lieb Inequalities in the Study of Vinogradov Mean Value Theorem and Parsell-Vinogradov Systems

Recently, the Brascamp-Lieb inequalities got a lot of attention, as it shows up naturally in the proof of the Main Conjecture of Vinogradov's mean value theorem by the decoupling approach of Bourgain, Demeter and Guth. We first recall the conjecture.

Definition 1.3.1. *Let $s \geq 1$ and $k, N \geq 2$ be integers. Let $J_{s,k}(N)$ denote the number of integer solutions $(x_1, \dots, x_s, y_1, \dots, y_s)$ to the following system:*

$$x_1^l + x_2^l + \dots + x_s^l = y_1^l + y_2^l + \dots + y_s^l, 1 \leq l \leq k \quad (1.10)$$

with $1 \leq x_1, \dots, x_s, y_1, \dots, y_s \leq N$.

Alternatively, $J_{s,k}(N)$ has the following analytic explanation:

$$J_{s,k}(N) = \int_{[0,1]^k} |g_{s,k}(\alpha_1, \dots, \alpha_k)|^{2s} d\alpha_1 \cdots d\alpha_k \quad (1.11)$$

where the exponential sum

$$g_{s,k}(\alpha_1, \dots, \alpha_k) = \sum_{n=1}^N e(\alpha_1 n + \alpha_2 n^2 + \cdots + \alpha_k n^k). \quad (1.12)$$

Theorem 1.3.2 (Main Conjecture of VMVT, proved in [BDG16b]).

$$J_{s,k}(N) \ll N^{s+\varepsilon} + N^{2s - \frac{k(k+1)}{2} + \varepsilon} \quad (1.13)$$

Remark 1.3.3. (1.13) is sharp up to a factor of N^ε .

Besides decoupling, the Main Conjecture also has another approach called efficient congruencing introduced by Wooley. A lot of remarkable progress was made via efficient congruencing on the conjecture, and the conjecture itself now also has an efficient congruencing proof by Wooley.

The Main Conjecture has applications to a plethora of problems including the famous Waring's problem. For a good introduction to these applications, see [Woo14].

In the decoupling approach to the Main Conjecture, a key ingredient is to use Brascamp-Lieb inequalities to exploit transversality of wave packets at different stages of induction. This was verified by linear algebra in [BDG16b], enabling the authors to establish the full conjecture with their iteration machine.

Waring's problem has a strong diagonality assumption in its formulation. As suggested in [PPW13] a natural step forward is to extend Vinogradov's approach to more general equations or systems with weaker diagonality feature. In this spirit [PPW13] suggested a generalization of system (1.10), often referred as Parsell-Vinogradov system, shown as (1.14) below. It was studied in [PPW13] and they made a conjecture that generalizes the Main Conjecture of Vinogradov's mean value theorem.

Definition 1.3.4. *Let $d, s \geq 1$ and $k, N \geq 2$ be integers. Let $J_{s,k,d}(N)$ denote the number of integer solutions $(x_{j,l}, y_{j,l})_{1 \leq j \leq s, 1 \leq l \leq d}$ to the following system:*

$$\sum_{j=1}^s x_{j,1}^{i_1} \cdots x_{j,d}^{i_d} = \sum_{j=1}^s y_{j,1}^{i_1} \cdots y_{j,d}^{i_d}, \forall i_1, \dots, i_d \geq 0, 1 \leq i_1 + \cdots + i_d \leq k \quad (1.14)$$

with $1 \leq x_{j,l}, y_{j,l} \leq N$.

The d here is usually referred to as the *dimension* of the system.

The following conjecture was made in [PPW13]:

Conjecture 1.3.5. *Define $r_{d,k} = \binom{k+d}{d} - 1$ and $K_{d,k} = \frac{d}{d+1}(r_{d,k} + 1)k$. Then*

$$J_{s,k,d}(N) \ll_{\varepsilon} N^{sd+\varepsilon} + \sum_{j=1}^d N^{(2s-1)j+d-K_{j,k}+\varepsilon}. \quad (1.15)$$

Remark 1.3.6. *It was proved in [PPW13] that right hand side of (1.15) is also a lower bound up to a factor of N^{ε} . Parsell-Vinogradov systems were studied in [PPW13] by the efficient congruencing approach and they obtain the conjecture for large s close to the largest "kink point". This system has applications in the study of Diophantine systems including, for example, those that generalizes Waring's problem.*

See [PPW13] for this application.

There is a natural decoupling approach to Conjecture 1.3.5 given in [BDG16a]. Again a key issue that arises is that of transversality. For these Parsell-Vinogradov systems this is formulated as Conjecture 1.3 in [BDG16a]. This conjecture is stated in Chapter 11 as Theorem 11.0.29 where we prove it using elementary algebraic and combinatorial techniques. This coupled with the analysis in [BDG16a] place Conjecture 1.3.5 well in sight and recently [Guo] has resolved Conjecture 1.3.5 for the case $d = 2$.

Chapter 2

Recent Developments of the Polynomial Method in Incidence Geometry Type Problems

Our proof of the endpoint perturbed Brascamp-Lieb inequalities uses the so-called polynomial method. In this chapter we put a brief introduction of its recent impact on incidence geometry type problems.

The polynomial method has its roots in earlier number theory and combinatorics, and in turn inspired a lot of works. Among them the best known ones in combinatorics are probably Dvir's proof of the Kakeya set conjecture over finite fields [Dvi09] and the almost full solution to the Erdős distinct distance conjecture in $\mathbb{R}^2$ by Guth and Katz [GK15]. In our current introduction to this method, we will highlight Guth's proof of the endpoint multilinear Kakeya inequality (i.e. (1.7) of Theorem

1.2.5), which is the most relevant piece of work.

We start by recalling the Kakeya set conjecture over finite fields and Dvir’s proof.

Theorem 2.0.7 (Kakeya sets have full dimension in finite fields [Dvi09]). *For a finite field $\mathbb{F}$, if K is a Kakeya set in $\mathbb{F}^n$, i.e. for each direction, there is a line $l \subseteq K$ parallel to that direction. Then $|K| \gtrsim_n |\mathbb{F}|^n$.*

Theorem 2.0.7 was originally conjectured by Wolff. He was hoping its proof would shed some light on the Kakeya set conjecture over $\mathbb{R}$. Dvir’s proof via the polynomial method was surprising at that time even for experts. The proof was by contradiction: Assuming K does not have enough points, then there is a nonzero polynomial of low degree vanishing on K . But for any direction there is a line of that direction lying in K . Look at this line and we deduce that the highest homogeneous part of the polynomial vanishes on this direction. But since the polynomial does not have high enough degree, this could not hold for every direction. Hence we have a contradiction.

The proof of Theorem 2.0.7, as sketched above, represents the spirit of the polynomial method in incidence geometry type problems. In simple words, this method works in the following two steps:

1. When given a finite number of algebraic geometric objects (e.g. points, lines, tubes) in $\mathbb{R}^d$, we use zero sets of polynomials to “approximate” it; and
2. Use the algebraicity of the polynomial and the objects to obtain some quantitative estimates.

For example in Dvir’s proof of the finite field Kakeya, the first step was to construct a polynomial that vanishes on all the given points and the second was to gather

information from the vanishing of the polynomial on lines of different directions and then deduce a contradiction. We can see that an advantage of the polynomial method is that a polynomial or algebraic variety can relate objects of different dimensions (e.g. $\mathbb{F}^d$ and lines).

Next we focus on Kakeya type problems over $\mathbb{R}$. The key feature of such problems is that they are about small cubes and thin tubes (in $\mathbb{R}^d$) rather than points and lines. In order to apply the polynomial method, we need some changes as apparently a nonzero polynomial cannot vanish on an entire tube. Guth was able to find a version of the polynomial method and proved the endpoint multilinear estimate (Theorem 1.2.5) [Gut10].

To prove Theorem 1.2.5, Guth found a working analogy of Dvir's approach in the Kakeya setting over $\mathbb{R}$. By the polynomial ham-sandwich theorem of Stone and Tukey [ST42], if we are given a set of small cubes there is a polynomial Q of reasonable degree whose zero set $Z(Q)$ bisects each small cube. Then by isoperimetric inequality, the surface area of $Z(Q)$ inside each cube must be large. If we take the support of the function on the left hand side of (1.7) and approximately view it as a characteristic function of a union of unit cubes, then we can choose a polynomial Q in the above way so that the surface area of $Z(Q)$ inside each cube is $\gtrsim 1$.

Fix such a cube and look at it. Since it belongs to some tube in each of the d families, $Z(Q)$ inside this cube has to have large projected area after the orthogonal projection along the core line of some tube (here we use the fact that the d families of tubes are quantitatively transversal). However the total projected area cannot be too large by the following crucial *cylinder estimate*:

Lemma 2.0.8 (Cylinder estimate). *Let T be a cylinder of radius 1 and Q be a polynomial of degree D . Let $\mathbf{v}$ be a unit vector parallel to the core line of T . If we denote $Z(Q)$ to be the zero set of P , then the directed volume (see Definition 4.1)*

$$V_{Z(Q) \cap T}(\mathbf{v}) \lesssim_d D. \quad (2.1)$$

Lemma 2.0.8 can be viewed as a substitute of the fact that a line not in $Z(Q)$ can meet $Z(Q)$ at in at most $\deg Q$ points. It is indeed an easy consequence of the above statement.

Working out all the details of the argument sketched above, we actually obtain a slighter weaker result:

Theorem 2.0.9 (Endpoint Multilinear Kakeya Theorem (toy model without multiplicity) [Gut10]). *For $1 \leq j \leq d$, let $\{T_{j,a} : 1 \leq a \leq A(j)\}$ be a family of unit cylinders in $\mathbb{R}^d$. Assume the direction of the core lines of cylinders from $T_{j,a}$ are all close enough to the x_j -coordinate axis. Then we have*

$$|\bigcap_{j=1}^d \bigcup_{a=1}^{A(j)} T_{j,a}| \lesssim_d \prod_{j=1}^d A(j)^{\frac{1}{d-1}}. \quad (2.2)$$

In order to prove Theorem 1.2.5, finer analysis is needed. Guth's original treatment in [Gut10] is to introduce a more delicate concept called visibility that works well with multiplicity. This is also going to play an important role in our proof of the perturbed Brascamp-Lieb (see Chapter 4 for its definition). Later another proof was found by Carbery-Valdimarsson [CV13] that only uses Borsuk-Ulam and avoids visibility.

Our proof of the perturbed Brascamp-Lieb inequalities will be based on the above work of Guth. Next we come to a more detailed introduction of our result.

Chapter 3

Motivation and Ideas of the Proof of the Endpoint Perturbed Brascamp-Lieb Inequality

3.1 The Endpoint Multilinear k_j -Plane Theorem

As interesting and rich special cases of Brascamp-Lieb inequalities, we look at the multilinear k_j -plane problem first. The multilinear k_j -plane conjecture was implicitly proved by Bennett, Carbery and Tao in [BCT06], except for the endpoint case. In the first part of this thesis we formulate it and prove the endpoint case. In fact we will prove the endpoint multilinear k_j -variety theorem, which is more general.

The proof uses the polynomial method. We will set up the polynomial like Guth did in his proof of the endpoint multilinear Kakeya Conjecture (Theorem 1.2.5).

Then we make some crucial new observations and development of the theory, enabling us to estimate “the quantitative interaction of the polynomial with itself” in terms of its visibility. As a result, we are able to deal with the codimension difficulty and complete the proof.

The multilinear k_j -plane estimate is a natural generalization of the multilinear Keakeya estimate (Theorems 1.2.4 and 1.2.5), as mentioned in the introduction. In the Keakeya setting we have cylinders which are neighborhoods of lines. A natural analogue is to replace lines with higher dimensional affine subspaces and this will exactly be our multilinear k_j -plane setting. In Remark 5.4 of [BCT06], the authors note that their techniques can also be used to obtain non-endpoint case of multilinear k -plane transform estimates that Oberlin and Stein consider in [OS82]. There is also a k -plane version of Keakeya problem [Bou91] that could be relevant here.

They did not state the result precisely and we will state what we can get from their proof below. If we go check the proof, similar techniques in [Gut15] can also give us the result. Here we allow subspaces of different dimensions and hence call the theorem a “multilinear k_j -plane theorem”.

Before stating the theorem we introduce our terminology to describe “higher dimensional” analogue of cylinders.

We recall Definition 1.2.2. By that definition, a 1-slab is a cylinder. Its length is the size in our language. Our definitions of radius and core are consistent with familiar definitions for cylinders. As explained above, we call our theorem a k_j -plane theorem because when the size is large a k -slab looks flat and is like a “fattened” k -plane.

Theorem 3.1.1 (Multilinear k_j -Plane Theorem with R^ε Loss[BCT06]). *Assume R is a large positive number. Assume $K_1, K_2, \dots, K_n \subsetneq \{1, 2, \dots, d\}$ are disjoint and $K_1 \cup \dots \cup K_n = \{1, 2, \dots, d\}$. Let $k_j = |K_j|$.*

For $1 \leq j \leq n$, let $\{T_{j,a} : 1 \leq a \leq A(j)\}$ be a family of k_j -slabs of size $\leq R$ and radius 1. Assuming that for any $1 \leq a_j \leq A(j)$, the core of T_{j,a_j} is on a k_j -plane that forms an angle $< \delta$ against the k_j -plane spanned by all $\mathbf{e}_i, i \in K_j$.

Then when $\delta > 0$ is sufficiently small depending on d , we have

$$\int_{\mathbb{R}^d} \left(\prod_{j=1}^n \sum_{a=1}^{A(j)} \chi_{T_{j,a}} \right)^{\frac{1}{n-1}} \lesssim_{\varepsilon, d} R^\varepsilon \prod_{j=1}^n A(j)^{\frac{1}{n-1}}. \quad (3.1)$$

When $n = d$ and $K_j = \{j\}$, this theorem is the multilinear Kakeya theorem with R^ε loss, being the main theorem of [BCT06]. In [Gut15], a simpler proof of this special case is also proven. The proof in [Gut15] can also be generalized easily to prove the whole Theorem 3.1.1.

We can obtain various k_j -plane theorems by taking different n and K_j in Theorem 3.1.1. As we saw in Theorem 1.2.5, Guth [Gut10] was able to remove the R^ε in the multilinear Kakeya case. So in general we would also expect the removal of R^ε . Conceptually, this will allow us to have slabs with “size ∞ ” (that are actually 1-neighborhoods of k_j planes) in the theorem. It turns out to be true and will be proved in this thesis.

Theorem 3.1.2 (Multilinear k_j -Plane Theorem). *Assumptions are the same as The-*

orem 3.1.1, but no restriction on the size of slabs. We have

$$\int_{\mathbb{R}^d} \left(\prod_{j=1}^n \sum_{a=1}^{A(j)} \chi_{T_{j,a}} \right)^{\frac{1}{n-1}} \lesssim_d \prod_{j=1}^n A(j)^{\frac{1}{n-1}}. \quad (3.2)$$

Theorem 3.1.2 has an affine-invariant version, just like the multilinear Kakeya case which was first pointed out in [BG11]. We will actually prove this version (Theorem 3.1.3 below). Theorem 3.1.2 is a direct corollary of it.

In order to state the theorem, we introduce some notations. For any $q \leq d$ vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q$, we denote $|\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \dots \wedge \mathbf{v}_q|$ to be the volume of the parallelepiped generated by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q$. Moreover for any m (affine) subspaces $V_1, V_2, \dots, V_m$ with a total dimension d , we can define $|V_1 \wedge V_2 \wedge \dots \wedge V_m|$ to be $|\mathbf{v}_{1,1} \wedge \dots \wedge \mathbf{v}_{1,d_1} \wedge \mathbf{v}_{2,1} \wedge \dots \wedge \mathbf{v}_{2,d_2} \wedge \dots \wedge \mathbf{v}_{m,1} \wedge \dots \wedge \mathbf{v}_{m,d_m}|$, where $\{\mathbf{v}_{j,i} (1 \leq i \leq d_j)\}$ form an orthonormal basis of the linear subspace parallel to V_j .

Theorem 3.1.3 (Affine invariant Multilinear k_j -Plane Theorem). *Assume that positive integers $1 \leq k_1, \dots, k_n \leq d-1$ satisfy $\sum_{j=1}^n k_j = d$. For $1 \leq j \leq n$, let $\{T_{j,a} : 1 \leq a \leq A(j)\}$ be a family of k_j -slabs of radius 1. Assume the core k_j -plane of $T_{j,a}$ is parallel to the linear subspace $H_{j,a}$. Then for any real numbers $\rho_{j,a}$ we have*

$$\int_{\mathbb{R}^d} \left(\sum_{a_1=1}^{A(1)} \dots \sum_{a_n=1}^{A(n)} \prod_{j=1}^n \rho_{j,a_j} \chi_{T_{j,a_j}}(x) \cdot H_{1,a_1} \wedge \dots \wedge H_{n,a_n} \right)^{\frac{1}{n-1}} dx \lesssim_d \prod_{j=1}^n \left(\sum_{a_j=1}^{A(j)} |\rho_{j,a_j}| \right)^{\frac{1}{n-1}}. \quad (3.3)$$

Remark 3.1.4. *We refer the reader to [BB10] on why the exponents are as they are here in Theorem 3.1.3. Also we note that in [BB10] the authors already observed*

the affine-invariant Finner inequality, which is an “unperturbed” version of Theorem 3.1.3.

Our Theorem 3.1.3 has some application in the multilinear restriction theorem too. Assuming for each $1 \leq j \leq n$, $\Sigma_j : U_j \rightarrow \mathbb{R}^d$ is a smooth parametrization of a subset of smooth submanifold Ω_j whose closure is compact. Also assume $\sum_{j=1}^n \dim \Omega_j = d$. Here we assume U_j is a neighborhood of the origin 0. We can associate the *extension operator* to Σ_j as follows:

$$E_j g_j(\xi) = \int_{U_j} e^{2\pi i \xi \cdot \Sigma_j(x)} g_j(x) dx. \quad (3.4)$$

Assuming $T_{\Sigma_1(0)}\Omega_1 \wedge \cdots \wedge T_{\Sigma_n(0)}\Omega_n \neq 0$. Then just like the classical multilinear restriction case discussed in [BCT06], we can form the endpoint multilinear restriction conjecture:

Conjecture 3.1.5 (Endpoint Multilinear k_j -Restriction Conjecture). *Assume we have Σ_j as above such that $T_{\Sigma_1(0)}\Omega_1 \wedge \cdots \wedge T_{\Sigma_n(0)}\Omega_n \neq 0$. Then when the U_j ’s are sufficiently small we have*

$$\int_{\mathbb{R}^d} \prod_{j=1}^n |E_j g_j|^{\frac{2}{n-1}} \lesssim_d \prod_{j=1}^n \|g_j\|_{L^2(U_j)}^{\frac{2}{n-1}}. \quad (3.5)$$

The methods in [BCT06] yield the following local variant of the conjectured (3.5) with R^ε -loss :

$$\int_{B(0,R)} \prod_{j=1}^n |E_j g_j|^{\frac{2}{n-1}} \lesssim_{d,\varepsilon} R^\varepsilon \prod_{j=1}^n \|g_j\|_{L^2(U_j)}^{\frac{2}{n-1}}. \quad (3.6)$$

We can use Theorem 3.1.2 to slightly improve (3.6). Using exactly the same proof techniques as in the proof of Theorem 4.2 in [Ben13], from Theorem 3.1.2 we deduce that there exists a $\kappa = \kappa(d) > 0$ such that:

$$\int_{B(0,R)} \prod_{j=1}^n |E_j g_j|^{\frac{2}{n-1}} \lesssim_d (\log R)^\kappa \prod_{j=1}^n \|g_j\|_{L^2(U_j)}^{\frac{2}{n-1}}. \quad (3.7)$$

3.2 The Endpoint Perturbed Brascamp-Lieb Inequalities (Continued)

The Loomis-Whitney inequality and the multilinear k_j -plane Theorem are both special cases of the Brascamp-Lieb inequalities. As advertised, in this thesis we generalize the Brascamp-Lieb inequalities in the same way we do with the multilinear k_j -plane theorem, with some new combinatorial ideas. Our result is an endpoint version (Theorem 1.2.7).

Our proof of Theorem 1.2.7 will follow the same scheme as the proof of Theorem 3.1.2. Some new difficulties present themselves and we deal with them in due course.

Like what we had in the end of last subsection, our perturbed Brascamp-Lieb theorem has some impact on the endpoint Brascamp-Lieb type restriction conjecture formulated in [BBFL15]. To introduce it, we setup as we did in Conjecture 3.1.5. But this time we don't assume that $\sum_j k_j = d$ or that $T_{\Sigma_1(0)}\Omega_1 \wedge \cdots \wedge T_{\Sigma_n(0)}\Omega_n \neq 0$. Instead, we assume that there exists $\mathbf{p} = (p_1, \dots, p_n)$, $p_j > 0$, such that $BL(\mathbf{B}(\Sigma), \mathbf{p}) < \infty$ where $\mathbf{B}(\Sigma) = (T_{\Sigma_1(0)}\Omega_1, \dots, T_{\Sigma_n(0)}\Omega_n)$ (here we abuse the notation a bit and for each component we really mean the linear subspace of $\mathbb{R}^d$ parallel to it).

Conjecture 3.2.1 (Endpoint Brascamp-Lieb Type Restriction Conjecture). *With the above setup, when the U_j 's are sufficiently small, we have*

$$\int_{\mathbb{R}^d} \prod_{j=1}^n |E_j g_j|^{2p_j} \lesssim_{d, \mathbf{p}, BL(\mathbf{B}(\Sigma), \mathbf{p})} \prod_{j=1}^n \|g_j\|_{L^2(U_j)}^{2p_j}. \quad (3.8)$$

In [BBFL15] a local variant of (3.8) with R^ε -loss is proved:

$$\int_{B(0, R)} \prod_{j=1}^n |E_j g_j|^{2p_j} \lesssim_{d, \mathbf{p}, BL(\mathbf{B}(\Sigma), \mathbf{p}), \varepsilon} R^\varepsilon \prod_{j=1}^n \|g_j\|_{L^2(U_j)}^{2p_j}. \quad (3.9)$$

By Theorem 1.2.7 and again the same method as in the proof of Theorem 4.2 in [Ben13], we can slightly improve (3.9): There is a $\kappa = \kappa(BL(\mathbf{B}(\Sigma), \mathbf{p})) > 0$ such that

$$\int_{B(0, R)} \prod_{j=1}^n |E_j g_j|^{2p_j} \lesssim_{d, \mathbf{p}} (\log R)^\kappa \prod_{j=1}^n \|g_j\|_{L^2(U_j)}^{2p_j}. \quad (3.10)$$

3.3 Ideas of the Proofs

When looking for removing the factor R^ε in Theorem 3.1.1 and Theorem 1.2.6, the methods in [BCT06] or [Gut15] apparently do not apply. Instead we will follow the path led by Guth [Gut10] with a variant of the polynomial method.

However, there is a major difficulty to generalize Guth's argument: Note that the zero set of one polynomial has codimension 1. in the setting of [Gut10], because a line has dimension 1, a line will intersect the above zero set at discrete points. And the number of such points is controlled by the degree of the polynomial. Hence we

can do some counting to obtain estimates. In particular, Guth’s proof relies heavily on the cylinder estimate (Lemma 2.0.8).

In the k_j -plane setting, the zero set of a single polynomial no longer interacts well with a k_j -plane: Because the latter generally has a smaller codimension, it won’t intersect the former at discrete points in general. Due to this issue we cannot do counting and seem to lose our main weapon (Lemma 2.0.8).

In this thesis, we deal with this difficulty and obtain our Theorem 3.1.2. The main idea is the following: for a k -plane, instead of finding one single polynomial, we would like to intersect it with the zero sets of k polynomials. Because the codimensions of the k -plane and zero sets of the k polynomials add up to d , they will intersect in points and it is possible to do counting to estimate the intersection again.

Along this line, we are taking more than 1 polynomial to approximate an arbitrary set of N cubes. We would like the zero sets of all the polynomials to be “transverse”, with this requirement we can choose at most d such polynomials. Like the original polynomial method, we would like to know how low the degrees of our polynomials can be. Guth [Gut10] showed that we can always choose the first polynomial to be of degree $\lesssim_d N^{\frac{1}{d}}$. But for the second polynomial this degree bound may already be no longer valid. Think about N unit squares lining up on a line in the plane $\mathbb{R}^2$. Any polynomial with degree significantly less than N would have most of its zero set “almost parallel” to the line (see [Gut14a]) and hence two such polynomials cannot intersect transversely at most of the squares. However in this example it is possible to find two transverse polynomials with degree product being N . More generally one can look at examples of cube grids or more generally transverse intersections of

hypersurfaces and similar phenomenon happen there. Based on the above discussion, we ask the following question in the spirit of the polynomial method.

Question 3.3.1. *Given any N disjoint unit cubes in $\mathbb{R}^d$ and $A_\nu > 1$ for each given cube Q_ν , do there always exist d polynomials $P_1, P_2, \dots, P_d$ such that $\prod_{i=1}^d \deg P_i$ is roughly $\sum_\nu A_\nu$, and the zero sets of all P_i have “quantitative interaction” $\gtrsim_d A_\nu$ at each of the above cubes?*

We explain the term “quantitative interaction” in Question 3.3.1. Morally we want this to be the number of points where the underlying hypersurfaces intersect. However this definition has a technical problem: The hypersurfaces can have degenerate intersections being a set of dimension larger than 0. Hence we actually move each hypersurface around a bit and take the expected number points where the displaced hypersurfaces intersect to be the “quantitative interaction”. A formal definition would take too much space and we omit it here. The term “quantitative interaction” would never appear in any statement in this thesis and only serves purpose of getting intuition.

Question 3.3.1 looks like a “continuous version” of the inverse Bézout’s theorem (see for example [Tao]). The analogue is very difficult in algebraic geometry (see [Tao] for part of the reason), and is conceivably very hard in its current continuous version too. I believe it can be formulated as an explicit question with an affirmative answer though. One can make this question rigorous by specifying the meaning of “quantitative interaction”, see the discussion below and (8.9) for a result of this flavor.

Luckily enough, we find the full power of this hard version is not needed this time.

Instead, it will be equally useful to have a morally positive answer to the following “softer” question.

Question 3.3.2. *Given any N disjoint unit cubes in $\mathbb{R}^d$ and $A_\nu > 1$ for each given cube Q_ν , do there exist d polynomials $P_1, P_2, \dots, P_d$ and positive numbers $\alpha_\nu > 1$ such that $\prod_{i=1}^d \deg P_i$ is roughly $\sum_\nu \alpha_\nu A_\nu$, and the zero sets of all P_i have quantitative interaction $\gtrsim_d \alpha_\nu A_\nu$ at each of the above cubes?*

This question is weaker than Question 3.3.1 because there we have the additional requirement that $a_\nu = 1$. In other words, we allow polynomials of higher degree here but “with the right multiplicity”. In general, higher degree polynomials, even with the right multiplicity, do not necessarily work as well as ones with lowest possible degree, see for example some estimates in [Gut14a]. But in this application it makes no difference, as we are in a situation similar to that in [Gut10].

Surprisingly, it turns out that after some further refinement of the question, we find that we can take $P_1, \dots, P_d$ all to be the same P and that we can obtain P by the refined polynomial method of Guth involving visibility. Once this is clear we are able to prove our theorem with a great amount of help of (multi-)linear algebra and geometry.

To be more specific, we find that we can take a single nonzero polynomial (that is complicated enough to look like the product of several transverse polynomials) such that the following holds: If we denote $Z(P)$ to be the zero set of P , then for each relevant Q_ν , d copies of $Z(P) \cap Q_\nu$ interact in a sufficiently transverse manner. Since the d copies of $Z(P) \cap Q_\nu$ interact in a very transverse way, and the copies are all the same, for any j and any Q_ν we deduce that k_j copies of $Z(P) \cap Q_\nu$ interacts

sufficiently transversely with the part of the j -th family of slabs inside Q_ν . But for any j , the j -th family has a limited capacity of transverse interaction with k_j copies of $Z(P)$ by Bézout's Theorem. This gives us an estimate that leads to Theorem 3.1.2.

As we saw above, we end up taking one single polynomial for d times. Nevertheless, we choose to keep the entire thought process on “ d transverse polynomials” here because after all, it is how we eventually come up with the solution and the reader might find our thought process useful elsewhere. Also, Question 3.3.1 that remains open is still fundamental, as it's a general one concerning the polynomial approximation of any N cubes. For example, it implies the existence of the polynomial in the polynomial method. Its discrete analogue is also open, see [Tao]. But progress in various subcases has been made.

In the multilinear k_j -plane setting, our method actually proves a stronger theorem (Multilinear k_j -Variety Theorem 8.0.12) which largely generalizes Theorem 3.1.2. We will state its exact form after a bit more preparation. Here let me briefly describe it.

Let's take a new viewpoint. To know that a point belongs to a slab of radius 1 is equivalent to knowing the existence of another point on the core of the slab that lies in its 1-neighborhood. Also note that the union of all cores (k_j -planes) of the j -th family of slabs can be viewed as an algebraic variety of degree $A(j)$ and dimension k_j . This variety is a smoothly embedded k_j -manifold except some zero-volume subset. Our Theorem 3.1.2 is basically saying that the n families of k_j -planes have limited capacity of “transversally interaction”. We will prove that this is the general case for any n algebraic varieties with total dimension d in Theorem 8.0.12.

This multilinear k_j -variety theorem immediately has interesting special cases. For instance, we have a theorem about collections of sphere shells in the flavor of Theorem 3.1.2.

The proof of Theorem 1.2.7 is with almost the same technique, but we have some new difficulties: When we use this technique, we want to know how well each k_j -plane interacts with our polynomial. However the information on the Brascamp-Lieb constant seemed to be very hard to use when we try to look at things “locally” as we do in the proof of Theorem 3.1.2. We address this issue in Chapters 9 and 10 by proving a weaker “integral version” of our previous pointwise estimate. Albeit weaker, it already leads to a proof of Theorem 1.2.7.

Like the situation of Theorem 3.1.2, Theorem 1.2.7 has a generalization to algebraic varieties (Theorem 10.0.18) and we prove the latter which automatically implies the former. Again the current form is quite strong and interesting in its own right.

Chapter 4

Polynomials with High Visibility

In this chapter, we review the refined polynomial method by Guth [Gut10]. We review the definition and properties of visibility and state Guth's theorem that we can find a polynomial with reasonable degree and large visibility in many cubes. Along the way we define a relevant notion, namely the *fading zone*, for future convenience.

Definition 4.0.3. *In $\mathbb{R}^d$, for any compact smooth hypersurface Z (possibly with boundary) and any vector $\mathbf{v}$, define the directed volume*

$$V_Z(\mathbf{v}) = \int_Z |\mathbf{v} \cdot \mathbf{n}| d\text{Vol}_Z \tag{4.1}$$

where $\mathbf{n}$ is the normal vector at the corresponding point of Z .

If $\mathbf{v}$ is a unit vector, there is a formula for $V_Z(\mathbf{v})$ that is geometrically more meaningful. Let $\pi_{\mathbf{v}}$ be the orthogonal projection of $\mathbb{R}^d$ onto the subspace $\mathbf{v}^\perp$. Then

for almost $y \in \mathbf{v}^\perp$, $|Z \cap \pi_{\mathbf{v}}^{-1}(y)|$ is finite and we have (see [Gut10])

$$V_Z(\mathbf{v}) = \int_{\mathbf{v}^\perp} |Z \cap \pi_{\mathbf{v}}^{-1}(y)| dy. \quad (4.2)$$

Definition 4.0.4. *The fading zone $F(Z)$ is defined to be the set $\{\mathbf{v} : |\mathbf{v}| \leq 1, V_Z(\mathbf{v}) \leq 1\}$. It is a nonempty convex compact subset of the unit ball (see [Gut10]). The visibility $Vis[Z] = \frac{1}{|F(Z)|}$.*

First we explain the heuristic meaning of the two concepts. Imagine that it is in midnight and we are looking at a glittering object exactly with the same shape as Z from a fixed distance. To describe the situation mathematically, we can find a vector $\mathbf{v}$ such that its direction is the direction of the object and its length is the brightness of the object. Then we can intuitively think that Z fades away when $\mathbf{v}$ enters the fading zone. And naturally the less visible the object is, the larger we want the fading zone to be. Hence we can define the visibility to be the inverse of the volume of the fading zone. See the beginning of Section 6 in [Gut10] for how to intuitively understand visibility and a few simple examples.

It is good to keep in mind that in this thesis we will mostly deal with hypersurfaces Z with $V_Z(\mathbf{v}) \gtrsim_d 1$ for any unit vector $\mathbf{v}$. For hypersurfaces that don't satisfy this we will typically fix it by taking its union with several hyperplanes parallel to coordinate hyperplanes.

Apparently as long as Z has finite volume, $F(Z)$ has a nonempty interior.

We are interested in polynomials and want to use the notions above to study them. Recall that the space of degree D algebraic hypersurfaces in $\mathbb{R}^d$ is parameterized by

$\mathbb{RP}^K$ for $K = \binom{D+d}{d} - 1$ in the following way: any such hypersurface corresponds to a polynomial P up to a scalar. By viewing P also as the $\binom{D+d}{d}$ -tuple of its coefficients we find this parametrization [Gut10]. We want to think of the directed volume and the visibility as functions over $\mathbb{RP}^K$. However as Guth pointed out [Gut10], they are bad functions that may even be discontinuous.

Following Guth [Gut10], we get around this difficulty by looking at the mollified version of them. If we take the standard metric on $\mathbb{RP}^K$, we will mollify those functions over small balls around some $P \in \mathbb{RP}^K$. In the rest of this thesis, we take ε to be a very small positive number depending on all the constants, and in application on the set of cubes and visibility conditions. This kind of assumption is often dangerous but as we can eventually see, here it does no harm at all (mainly because all the algebraic hypersurfaces of degree D satisfy the same Intersection Estimate (7.5) uniformly), just like the case of [Gut10]. There instead of the Intersection Estimate, we have the Cylinder Estimate (2.1) as a special case counterpart.

For any $P \in \mathbb{RP}^K$, let $B(P, \varepsilon)$ be the ε -neighborhood of P . Let $Z(P)$ denote the zero set of P . Note that for any P , the set of singular points on $Z(P)$ has zero $(d-1)$ -dimensional Hausdorff measure. And the rest of $Z(P)$ is a smooth embedded hypersurface by the Implicit Function Theorem.

Definition 4.0.5. *For any bounded open set U and any vector $\mathbf{v}$, define the mollified directed volume*

$$\overline{V}_{Z(P) \cap U}(\mathbf{v}) = \frac{1}{|B(P, \varepsilon)|} \int_{B(P, \varepsilon)} V_{Z(P') \cap U}(\mathbf{v}) dP'. \quad (4.3)$$

Define the mollified fading zone and mollified visibility based on the mollified directional volumes:

$$\overline{F}(Z(P) \cap U) = \{\mathbf{v} : |\mathbf{v}| \leq 1 : \overline{V}_{Z(P) \cap U}(\mathbf{v}) \leq 1\} \quad (4.4)$$

$$\overline{Vis}[Z(P) \cap U] = \frac{1}{|\overline{F}(Z(P) \cap U)|}. \quad (4.5)$$

Like we had before for $F(Z)$, $\overline{F}(Z(P) \cap U)$ is a convex compact subset of the unit ball with an nonempty interior. By John's Ellipsoid Theorem [Joh14], for any convex set Γ with interior, there is an ellipsoid $Ell(\Gamma)$ such that $Ell(\Gamma) \subseteq \Gamma \subseteq C_d Ell(\Gamma)$ and that $|Ell(\Gamma)| \sim_d |\Gamma|$. It is easy to see that if the convex set is symmetric about the origin (which will be the case for all convex sets considered in this thesis), then we may require the ellipsoid to be symmetric about the origin too. We assume so henceforth in the thesis. We call any such $Ell(\Gamma)$ an *elliptical approximation* of Γ .

$\overline{V}_{Z(P) \cap U}(\mathbf{v})$ and $\overline{Vis}[Z(P) \cap U]$ are continuous with respect to $P \in \mathbb{RP}^M$ [Gut10]. Guth proved the following key lemma in [Gut10].

Lemma 4.0.6 (Large Visibility on Many Cubes [Gut10]). *For any finite set of cubes $Q_1, \dots, Q_N$ and non-negative integers $M(Q_i), 1 \leq i \leq N$, there exists a polynomial P of degree $\leq D$ (but viewed as a degree D polynomial when we mollify) such that $\overline{Vis}(Z(P) \cap Q_k) \geq M(Q_k)$ and that $D \lesssim_d (\sum_{i=1}^N M(Q_k))^{\frac{1}{d}}$.*

Chapter 5

Wedge Product Estimate Based on Visibility

As we are actually dealing with the mollification version of everything, it is convenient to have a generalized definition of visibility on any space of finite measure. The generalized setup here will also be cleaner and more flexible in the inductive arguments which are needed.

Assume we have a measure space (X, μ) with $\mu(X) < \infty$ and a vector-valued measurable function $f : X \rightarrow \mathbb{R}^d$. For any vector $\mathbf{v} \in \mathbb{R}^d$ define the *total absolute inner product* of $\mathbf{v}$ and f (the directed volume of last section being the example we have in mind):

$$V_{X,f}(\mathbf{v}) = \int_X |\mathbf{v} \cdot f(x)| d\mu(x). \quad (5.1)$$

Define the fading zone $F(X, f) = \{\mathbf{v} \leq 1, V_{X,f}(\mathbf{v}) \leq 1\}$ and visibility $Vis[X, f] = \frac{1}{|F(X, f)|}$. As we had in the end of last section, we have an elliptical approximation

$Ell(F(X, f))$ such that $Ell(F(X, f)) \subseteq F(X, f) \subseteq C_d Ell(F(X, f))$.

Next we obtain a lower bound of a wedge product integral in terms of visibility.

Theorem 5.0.7 (Wedge Product Estimate). *Assume that for any unit vector $\mathbf{v}$ we have $V_{X,f}(\mathbf{v}) \geq 1$. Then*

$$\int \cdots \int_{X^d} |\wedge_{i=1}^d f(x_i)| d\mu(x_1) d\mu(x_2) \cdots d\mu(x_d) \gtrsim_d Vis[X, f]. \quad (5.2)$$

Proof. We do induction on the dimension d to prove the theorem. First observe that if $Ell(F(X, f))$ is an elliptical approximation of $F(X, f)$, then for any linear subspace W of $\mathbb{R}^d$, $Ell(F(X, f)) \cap W$ (ellipsoid) is also an elliptical approximation of $F(X, f) \cap W$ by definition (this may seem problematic as the C_d will vary, but for the conclusion only finite many intermediate dimensions are involved in the whole induction process and we can set the C_d of them all being the same).

For $d = 1$, by definition we easily see

$$Vis[X, f] = \frac{1}{2} \int_X |f(x)| d\mu(x) \quad (5.3)$$

and the conclusion holds. Note that even in the argument here we are using the hypothesis to ensure we have (5.3).

Assume the conclusion holds for $d < d_0$ and $d_0 > 1$. Now we deal with the case $d = d_0$. Assume $\mathbf{v}_1, \dots, \mathbf{v}_{d_0}$ are parallel to the semi-principal axes of any elliptical approximation $Ell(F(X, f))$, respectively, and that they form an orthonormal basis (we can arbitrarily choose a set of orthogonal semi-principal axes if there is ambiguity defining the semi-principal axes). Among them we assume $\mathbf{v}_1$ is parallel to a semi-

minor axis (i.e. a shortest semi-principal axis) that has length t_1 . Taking $\mathbf{v} = \lambda \mathbf{v}_1$ where $\lambda \sim_{d_0} t_1$ in (5.1) we deduce

$$\int_X |f(x)| d\mu(x) \geq \frac{1}{t_1}. \quad (5.4)$$

Next for any unit vector $\mathbf{v} \in \mathbb{R}^{d_0}$, we prove

$$\int \cdots \int_{X^{d_0-1}} |f(x_1) \wedge \cdots \wedge f(x_{d_0-1}) \wedge \mathbf{v}| d\mu(x_1) d\mu(x_2) \cdots d\mu(x_{d_0-1}) \gtrsim_{d_0} t_1 \cdot \text{Vis}[X, f]. \quad (5.5)$$

Let $\pi_{\mathbf{v}^\perp}$ be the orthogonal projection from $\mathbb{R}^{d_0}$ to its subspace $\mathbf{v}^\perp$. Define $f_{\mathbf{v}^\perp} = \pi_{\mathbf{v}^\perp} \circ f$. If we identify $\mathbb{R}^{d_0-1}$ with $\mathbf{v}^\perp$, $f_{\mathbf{v}^\perp}$ is another (d_0-1) -dimensional-vector-valued function on X . By definition, we know for any $\mathbf{w} \in \mathbf{v}^\perp$, $V_{X,f}(\mathbf{w}) = V_{X,f_{\mathbf{v}^\perp}}(\mathbf{w})$. Hence $F(X, f_{\mathbf{v}^\perp}) = F(X, f) \cap \mathbf{v}^\perp$. By the previous discussion, we know we can choose $\text{Ell}(F(X, f_{\mathbf{v}^\perp}))$ to be $\text{Ell}(F(X, f)) \cap \mathbf{v}^\perp$. But among all the (d_0-1) -dimensional sections of $\text{Ell}(F(X, f))$ passing through the origin, the section cut by $\mathbf{v}_1^\perp$ has the largest volume (see also Lemma 9.0.16) which is $\sim_{d_0} \frac{|\text{Ell}(F(X, f))|}{t_1} = \frac{1}{t_1 \cdot \text{Vis}[X, f]}$. Hence $\text{Vis}[X, f_{\mathbf{v}^\perp}] = \frac{1}{|F(X, f_{\mathbf{v}^\perp})|} \sim_{d_0} \frac{1}{|\text{Ell}(F(X, f_{\mathbf{v}^\perp}))|} \gtrsim_{d_0} t_1 \cdot \text{Vis}[X, f]$. By induction hypothesis we have

$$\begin{aligned} & \int \cdots \int_{X^{d_0-1}} |f(x_1) \wedge \cdots \wedge f(x_{d_0-1}) \wedge \mathbf{v}| d\mu(x_1) d\mu(x_2) \cdots d\mu(x_{d_0-1}) \\ &= \int \cdots \int_{X^{d_0-1}} |f_{\mathbf{v}^\perp}(x_1) \wedge \cdots \wedge f_{\mathbf{v}^\perp}(x_{d_0-1})| d\mu(x_1) d\mu(x_2) \cdots d\mu(x_{d_0-1}) \\ &\gtrsim_{d_0} \text{Vis}[X, f_{\mathbf{v}^\perp}] \\ &\gtrsim_{d_0} t_1 \cdot \text{Vis}[X, f]. \end{aligned} \quad (5.6)$$

This is (5.5).

Combining (5.4) and (5.5), we have

$$\begin{aligned}
& \int \cdots \int_{X^d} |\wedge_{i=1}^d f(x_i)| d\mu(x_1) d\mu(x_2) \cdots d\mu(x_d) \\
&= \int_X |f(x)| \left(\int \cdots \int_{X^{d_0-1}} |f(x_1) \wedge \cdots \wedge f(x_{d_0-1}) \wedge \frac{f(x)}{|f(x)|}| d\mu(x_1) d\mu(x_2) \cdots d\mu(x_{d_0-1}) \right) d\mu(x) \\
&\stackrel{\geq_{d_0}}{\sim} t_1 \cdot Vis[X, f] \cdot \int_X |f(x)| d\mu(x) \\
&\stackrel{\geq_{d_0}}{\sim} Vis[X, f]
\end{aligned} \tag{5.7}$$

which concludes the induction. □

Chapter 6

Linear Algebra Preliminaries

Our proof relies heavily on linear algebra. In this chapter we do the linear algebraic part and prove two useful lemmas.

Lemma 6.0.8. *Assume $V_1, \dots, V_n \subseteq \mathbb{R}^d$ and $k_j = \dim V_j$ satisfies $\sum_{j=1}^n k_j = d$.*

Then for any vectors $\mathbf{w}_1, \dots, \mathbf{w}_d \in \mathbb{R}^d$, we have

$$\max_{\substack{1 \leq i_{j,h} \leq d \text{ for } 1 \leq j \leq n, 1 \leq h \leq k_j, \\ \text{each } 1 \leq i \leq d \text{ is chosen exactly once among all } i_{j,h}}} \prod_{j=1}^n |(V_j)^\perp \wedge \mathbf{w}_{i_{j,1}} \wedge \dots \wedge \mathbf{w}_{i_{j,k_j}}| \gtrsim_d |V_1 \wedge \dots \wedge V_n| \cdot |\wedge_{i=1}^d \mathbf{w}_i|. \quad (6.1)$$

Proof. Assume that $\{\mathbf{v}_{j,h}\}_{1 \leq h \leq k_j}$ is an orthonormal basis of V_j . Then by definition we have

$$|(V_1 \wedge \dots \wedge V_n) \cdot (\wedge_{i=1}^d \mathbf{w}_i)| = |(\mathbf{v}_{j,h} \cdot \mathbf{w}_i)|. \quad (6.2)$$

By the generalized Laplace cofactor expansion, the determinant on the right hand

side of (6.2) is a sum of terms in the form

$$\pm \det(\mathbf{v}_{1,h} \cdot \tilde{\mathbf{w}}_{1,h}) \det(\mathbf{v}_{2,h} \cdot \tilde{\mathbf{w}}_{2,h}) \cdots \det(\mathbf{v}_{n,h} \cdot \tilde{\mathbf{w}}_{n,h}) \quad (6.3)$$

where $\tilde{\mathbf{w}}_{1,1}, \tilde{\mathbf{w}}_{1,2}, \dots, \tilde{\mathbf{w}}_{1,k_1}, \tilde{\mathbf{w}}_{2,1}, \dots, \tilde{\mathbf{w}}_{2,k_2}, \dots, \tilde{\mathbf{w}}_{n,1}, \dots, \tilde{\mathbf{w}}_{n,k_n}$ is a rearrangement of $\mathbf{w}_1, \dots, \mathbf{w}_d$. Hence for some such rearrangement we have

$$|V_1 \wedge \cdots \wedge V_n| \cdot |\wedge_{i=1}^d \mathbf{w}_i| \lesssim_d |\det(\mathbf{v}_{1,h} \cdot \tilde{\mathbf{w}}_{1,h}) \det(\mathbf{v}_{2,h} \cdot \tilde{\mathbf{w}}_{2,h}) \cdots \det(\mathbf{v}_{n,h} \cdot \tilde{\mathbf{w}}_{n,h})|. \quad (6.4)$$

By the properties of the Hodge $*$ -operator, we then have

$$|\det(\mathbf{v}_{j,h} \cdot \tilde{\mathbf{w}}_{j,h})| = |(*\mathbf{v}_{j,1} \wedge \cdots \wedge \mathbf{v}_{j,k_j}) \wedge \tilde{\mathbf{w}}_{j,1} \wedge \cdots \wedge \tilde{\mathbf{w}}_{j,k_j}| = |V_j^\perp \wedge \tilde{\mathbf{w}}_{j,1} \wedge \cdots \wedge \tilde{\mathbf{w}}_{j,k_j}| \quad (6.5)$$

which concludes the proof. $\square$

The rest of this section is dedicated to the computation of a determinant that will be useful in the next section.

Lemma 6.0.9. *Assume that $0 \leq c_j \leq d$ are integers, $1 \leq j \leq m$, satisfying $\sum_{j=1}^m c_j = d$. For any j , assume $\mathbf{v}_{j,1}, \mathbf{v}_{j,2}, \dots, \mathbf{v}_{j,d}$ is an orthonormal basis of $\mathbb{R}^d$ (written as*

column vectors). Then we have

$$\begin{aligned}
& \left| \det \begin{pmatrix} \mathbf{v}_{1,c_1+1} & \cdots & \mathbf{v}_{1,d} & \mathbf{v}_{2,c_2+1} & \cdots & \mathbf{v}_{2,d} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \mathbf{v}_{1,c_1+1} & \cdots & \mathbf{v}_{1,d} & 0 & \cdots & 0 & \mathbf{v}_{3,c_3+1} & \cdots & \mathbf{v}_{3,d} & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{v}_{1,c_1+1} & \cdots & \mathbf{v}_{1,d} & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \mathbf{v}_{m,c_m+1} & \cdots & \mathbf{v}_{m,d} \end{pmatrix} \right| \\
&= \left| \det (\mathbf{v}_{1,1} \cdots \mathbf{v}_{1,c_1} \mathbf{v}_{2,1} \cdots \mathbf{v}_{2,c_2} \cdots \mathbf{v}_{m,1} \cdots \mathbf{v}_{m,c_m}) \right|. \tag{6.6}
\end{aligned}$$

Proof. For $2 \leq j \leq m$, let $A_j = (\mathbf{v}_{1,1} \cdots \mathbf{v}_{1,c_1} \mathbf{0} \cdots \mathbf{0} \cdots \mathbf{v}_{j,1} \cdots \mathbf{v}_{j,c_j} \cdots \mathbf{0} \cdots \mathbf{0})$. The rule here is that its first c_1 columns are $\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,c_1}$ and its $(\sum_{j' < j} c_{j'} + 1)$ -th to $(\sum_{j' < j} c_{j'})$ -th columns are $\mathbf{v}_{j,1}, \dots, \mathbf{v}_{j,c_j}$, while its other columns being zero vectors. The left hand side of (6.6) is equal to

$$\left| \det \begin{pmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A_2 & \mathbf{v}_{1,c_1+1} & \cdots & \mathbf{v}_{1,d} & \mathbf{v}_{2,c_2+1} & \cdots & \mathbf{v}_{2,d} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ A_3 & \mathbf{v}_{1,c_1+1} & \cdots & \mathbf{v}_{1,d} & 0 & \cdots & 0 & \mathbf{v}_{3,c_3+1} & \cdots & \mathbf{v}_{3,d} & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_m & \mathbf{v}_{1,c_1+1} & \cdots & \mathbf{v}_{1,d} & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \mathbf{v}_{m,c_m+1} & \cdots & \mathbf{v}_{m,d} \end{pmatrix} \right|.$$

We exchange the columns to make it look better. For simplicity let $V_j = (\mathbf{v}_{j,1} \cdots \mathbf{v}_{j,d})$. This is an orthogonal matrix. We also define a matrix $B_j = (b_j(k, l))(1 \leq k \leq d)$ such that: $b_j(k, l) = 1$ if $l \leq c_j$ and $k = l + \sum_{j' < j} c_{j'}$, $b_j(k, l) = 0$ otherwise. Then after rearranging the columns of the matrix above we find the determinant (in

absolute value) is equal to

$$\left| \det \begin{pmatrix} B_1 & B_2 & B_3 & \cdots & B_m \\ V_1 & V_2 & 0 & \cdots & 0 \\ V_1 & 0 & V_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ V_1 & 0 & 0 & \cdots & V_m \end{pmatrix} \right|.$$

We can multiply the j -th column by $V_j^{-1} = V_j^t$ on the right, then subtract all the $j(> 1)$ -th column from the first column. This keeps the determinant. Note the definition of B_j , if we denote $\Delta = (\mathbf{v}_{1,1} \cdots \mathbf{v}_{1,c_1} - \mathbf{v}_{2,1} \cdots - \mathbf{v}_{2,c_2} \cdots - \mathbf{v}_{m,1} \cdots - \mathbf{v}_{m,c_m})$, then the determinant is

$$\left| \det \begin{pmatrix} \Delta^t & B_2 V_2^t & B_3 V_3^t & \cdots & B_m V_m^t \\ 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & I \end{pmatrix} \right|.$$

(6.6) then follows directly. □

Chapter 7

Integral Geometry Preliminaries

In this chapter we prepare some integral geometry tools for our proof of Theorem 3.1.2. First we generalize (4.2) to the following lemma.

Lemma 7.0.10. *Assume in $\mathbb{R}^d$ we have m smooth compact submanifolds $Z_1, Z_2, \dots, Z_m$ (possibly with boundary) with codimensions $c_1, \dots, c_m$ respectively. If $\sum_{j=1}^m c_j = d$ then for any measurable subset $U \subseteq \mathbb{R}^{d(m-1)} = (\mathbb{R}^d)^{m-1}$, we have*

$$\begin{aligned} & \int_{Z_1} \int_{Z_2} \cdots \int_{Z_m} \chi_U(\overrightarrow{p_1 p_2}, \dots, \overrightarrow{p_1 p_m}) |(T_{p_1} Z_1)^\perp \wedge \cdots \wedge (T_{p_m} Z_m)^\perp| d\text{Vol}_1 \cdots d\text{Vol}_m \\ &= \int_{\mathbf{v}_2, \dots, \mathbf{v}_m \in \mathbb{R}^d, (\mathbf{v}_2, \dots, \mathbf{v}_m) \in U} |(Z_1) \cap (Z_2 + \mathbf{v}_2) \cap \cdots \cap (Z_{m-1} + \mathbf{v}_{m-1}) \cap (Z_m + \mathbf{v}_m)| d\mathbf{v}_2 \cdots d\mathbf{v}_m \end{aligned} \quad (7.1)$$

where $p_j \in Z_j$, $T_{p_j} Z_j$ is the tangent space of Z_j at p_j , $d\text{Vol}_j$ is the volume element on the j -th submanifold, and $Z_j + \mathbf{v}_j = \{p_j + \mathbf{v}_j : p_j \in Z_j\}$ is the translation of Z_j along the vector $\mathbf{v}_j$. The $|\cdot|$ on the right hand side defines cardinality.

This lemma has a lot of information so we pause a bit and go through several

examples to understand it better.

When $d = 2$, if Z_1 and Z_2 are two non-parallel line segments and U is the whole $\mathbb{R}^2$, the integrand on the right hand side of (7.1) is the characteristic function of a parallelogram generalized by Z_1 and Z_2 . Hence the right hand side is the area of the parallelogram, which is easily seen to be equal to the left hand side. When $d = 3$, if Z_1 is a line segment, Z_2 is a parallelogram in a plane and U is the whole $\mathbb{R}^3$, the situation is totally analogous.

When $d = 3$, Z_1 is a whole line and Z_2 is a smooth surface of finite area, we can take U to be the point set between two planes orthogonal to Z_1 with distance 1. It is a simple exercise to show that (7.1) then becomes (4.2). Hence it is indeed a generalization of the latter.

Finally let's look at a more complicated example. Again take $d = 3$, $U = \mathbb{R}^6$. Take a parallelepiped $\Omega = ABCD - A_1B_1C_1D_1$. Take three parallelograms $Z_1 = ABCD$, $Z_2 = ABB_1A_1$, $Z_3 = ADD_1A_1$. Denote $\mathbf{u} = \overrightarrow{AB}$, $\mathbf{v} = \overrightarrow{AD}$, $\mathbf{w} = \overrightarrow{AA_1}$. Again the integrand on the right hand side is a characteristic function. We find it is plainly equal to $Vol(\Omega)^2$. Now the left hand side is equal to

$$\begin{aligned}
& |\mathbf{u} \times \mathbf{v}| \cdot |\mathbf{v} \times \mathbf{w}| \cdot |\mathbf{w} \times \mathbf{u}| \cdot \left| \frac{\mathbf{u} \times \mathbf{v}}{|\mathbf{u} \times \mathbf{v}|} \wedge \frac{\mathbf{v} \times \mathbf{w}}{|\mathbf{v} \times \mathbf{w}|} \wedge \frac{\mathbf{w} \times \mathbf{u}}{|\mathbf{w} \times \mathbf{u}|} \right| \\
= & |(\mathbf{u} \times \mathbf{v}) \wedge (\mathbf{v} \times \mathbf{w}) \wedge (\mathbf{w} \times \mathbf{u})| \\
= & |((\mathbf{u} \times \mathbf{v}) \times (\mathbf{v} \times \mathbf{w})) \cdot (\mathbf{w} \times \mathbf{u})| \\
= & |(\mathbf{v} \cdot (\mathbf{u} \times \mathbf{w})) \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u})| \\
= & Vol(\Omega)^2.
\end{aligned} \tag{7.2}$$

Proof of Lemma 7.0.10. Without loss of generality we can assume U is open and bounded. By the multilinear feature of both sides of (7.1), we only need to consider this problem locally. Hence we can assume each Z_j is smoothly parameterized by a domain in $\mathbb{R}^{d-c_j}$. In other words we may assume $Z_j : x_i = f_{j,i}(y_{j,1}, \dots, y_{j,d-c_j})$ and that the $(d-c_j)$ vectors $\mathbf{w}_{j,l} = (\frac{\partial f_{j,i}}{\partial y_{j,l}})_{1 \leq i \leq d}$ has a nonzero wedge product at any point $p_j \in Z_j$. They span the tangent space $T_{p_j} Z_j$ and will be written as column vectors below.

Look at the cartesian product $Z = Z_1 \times Z_2 \times \dots \times Z_m \subseteq (\mathbb{R}^d)^m \cong \mathbb{R}^{dm}$. This is a smooth submanifold of dimension $\sum_{j=1}^m (d-c_j) = d(m-1)$. Use $x_{j,i}$ ($1 \leq i \leq d$) to denote the standard Euclidean coordinates in the j -th copy of $\mathbb{R}^d$ as let $v_{j,i} = x_{j,i} - x_{1,i}$ ($j > 1$). For simplicity let $\mathbf{x}_j = (x_{j,i})_{1 \leq i \leq d}$ and $\mathbf{v}_j = (v_{j,i})_{1 \leq i \leq d}$. Notice that the right hand side of (7.1) is equal to

$$\int_Z \chi_U((\mathbf{v}_j)_{2 \leq j \leq m}) |d\mathbf{v}_2 d\mathbf{v}_3 \dots d\mathbf{v}_m|.$$

Define the density form $\theta = |d\mathbf{v}_2 d\mathbf{v}_3 \dots d\mathbf{v}_m| = |dv_{2,1} \wedge dv_{2,2} \wedge \dots \wedge dv_{2,d} \wedge \dots \wedge dv_{m,1} \wedge \dots \wedge dv_{m,d}|$. On the manifold Z it is a multiple of the volume density element $|dV| = \prod_{j=1}^m |\wedge_{l=1}^{d-c_j} \mathbf{w}_{j,l}| \wedge_{1 \leq j \leq m, 1 \leq l \leq d-c_j} dy_{j,l}$. Next we find $\frac{\theta}{|dV|}$.

We have $\frac{\theta}{|dV|} = \frac{1}{\prod_{j=1}^m |\wedge_{l=1}^{d-c_j} \mathbf{w}_{j,l}|} |(\frac{\partial v_{j,i}}{\partial y_{j,l}})|$. And by change of variable we have

$$= \left| \det \begin{pmatrix} -\mathbf{w}_{1,1} & \cdots & -\mathbf{w}_{1,d-c_1} & \mathbf{w}_{2,1} & \cdots & \mathbf{w}_{2,d-c_2} & \cdots & 0 & \cdots & 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\mathbf{w}_{1,1} & \cdots & -\mathbf{w}_{1,d-c_1} & 0 & \cdots & 0 & \cdots & \mathbf{w}_{m,1} & \cdots & \mathbf{v}_{m,d-c_m} & \cdots & \cdots & \cdots \end{pmatrix} \right| \quad (7.3)$$

This looks very much like the left hand side of (6.6). Indeed, the extra negative signs does not change the determinant and can be ignored. The only essential difference here is that for each j , our $\{w_{j,l}\}_{1 \leq l \leq d-c_j}$ is not a set of orthonormal vectors. If we do a change of variable to make them orthonormal we will extract a factor of $|\wedge_{l=1}^{d-c_j} \mathbf{w}_{j,l}|$ from right hand side of (7.3) for each j . We then apply Lemma 6.0.9 and get

$$\frac{\theta}{|\mathrm{d}V|} = |\wedge_{j=1}^m (T_{p_j} Z_j)^\perp|. \quad (7.4)$$

Hence the right hand side of (7.1) is equal to $\int_Z \chi_U((\mathbf{v}_j)_{2 \leq j \leq m}) |\wedge_{j=1}^m (T_{p_j} Z_j)^\perp| \mathrm{d}V$. Note that $\mathrm{d}V$ is the product of all $\mathrm{d}Vol_j$. This can easily be recognized as the left hand side. $\square$

In application, we want to look at the case where each V_j is the zero set of an algebraic variety of codimension c_j . Such a V_j may contain singular points, but they form a subset of measure 0 when we take the $(d-c_j)$ -dimensional Hausdorff measure. Hence almost all points on V_j are smooth points and we can apply our Lemma 7.0.10 to obtain the following theorem.

Theorem 7.0.11 (Intersection Estimate). *Assume in $\mathbb{R}^d$ we have m algebraic subvarieties $Z_1, Z_2, \dots, Z_m$ with codimensions $c_1, \dots, c_m$ and degree $s_1, \dots, s_m$ respectively. If $\sum_{j=1}^m c_j = d$ then for any measurable subset $U \subseteq \mathbb{R}^{d(m-1)} = (\mathbb{R}^d)^{m-1}$, we have*

$$\begin{aligned} & \int_{Z_1} \int_{Z_2} \cdots \int_{Z_m} \chi_U(\overrightarrow{p_1 p_2}, \dots, \overrightarrow{p_1 p_m}) |(T_{p_1} Z_1)^\perp \wedge \cdots \wedge (T_{p_m} Z_m)^\perp| d\text{Vol}_1 \cdots d\text{Vol}_m \\ & \leq \text{Vol}(U) \prod_{j=1}^m s_j \end{aligned} \tag{7.5}$$

where $d\text{Vol}_j$ is the $(d-c_j)$ -dimensional volume element on the j -th subvariety. Almost all $p_j \in Z_j$ are smooth points and we define $T_{p_j} Z_j$ to be the tangent space of Z_j at p_j .

Proof. (7.5) follows directly from Lemma 7.0.10 and Bézout's Theorem. $\square$

Theorem 7.0.11 generalizes the cylinder estimate in [Gut10] that was recorded as Lemma 2.0.8 in our current thesis.

Chapter 8

Proof of Theorem 3.1.2 and Theorem 3.1.3

In this chapter, we prove Theorem 3.1.3 and deduce Theorem 3.1.2 as a corollary. As briefly described in the introduction, we actually prove a generalized theorem about any n varieties.

Basically, our multilinear k_j -variety theorem says that for n algebraic subvarieties of $\mathbb{R}^d$ with their codimension adding up to d , their tubular neighborhoods will provide us with an inequality similar to Theorem 3.1.3 if we take their “amount of interaction” into account. In particular, if we take each subvariety to be a union of k_j -planes we obtain Theorem 3.1.3 and Theorem 3.1.2 (see the end of this section).

Theorem 8.0.12 (Multilinear k_j -Variety Theorem). *Assume $1 \leq k_j \leq d - 1$ ($1 \leq j \leq n$) satisfy $\sum_{j=1}^n k_j = d$. Assume that for $1 \leq j \leq n$, $H_j \subseteq \mathbb{R}^d$ is part of a k_j -dimensional algebraic subvariety of degree $A(j)$, respectively. Let $d\sigma_j$ denote*

the k_j -dimensional (Hausdorff) volume measure of H_j . Then with respect to this measure, almost all $y_j \in H_j$ are smooth points. For a smooth point $y_j \in H_j$, let $T_{y_j}H_j$ denote the tangent space of H_j at y_j . Then

$$\int_{\mathbb{R}^d} (\int_{H_1 \times H_2 \times \dots \times H_n} \chi_{\{\text{dist}(y_j, x) \leq 1\}} |\wedge_{j=1}^n T_{y_j}H_j| d\sigma_1(y_1) \cdots d\sigma_n(y_n))^{\frac{1}{n-1}} dx \lesssim_d \prod_{j=1}^n A(j)^{\frac{1}{n-1}}. \quad (8.1)$$

We give an outline of the proof before we actually do it. For the convenience of the statement we wrote Theorem 8.0.12 in an integral form. However because of the truncation $\chi_{\{\text{dist}(y_j, x) \leq 1\}}$ it is really of discrete flavour. In other words, around any unit cube, we only take into account the part of the varieties near this cube. By Lemma 4.0.6, we can find a polynomial with large visibility around each relevant cube. In the lemma, it is possible to assign different weights to different cubes in the above movement. We assign the weights according to the cubes' “popularity” among H_j , as Guth did in [Gut10] for the multilinear Kakeya theorem.

We will see it does not matter if we multiply all the weights by a same large positive number simultaneously. As long as the weights are large enough, we can add hyperplanes to the polynomial which do not essentially increase its degree and make its zero set satisfy the assumption of Theorem 5.0.7 at each relevant cube. Then we can invoke Theorem 5.0.7 for the resulting zero set $Z(P)$ at all relevant cubes to show that d copies of $Z(P)$ have enough interaction there. Now around each relevant cube we are ready to assign some copies of $Z(P)$ to each variety H_j and use Lemma 6.0.8 to show that those “have enough interaction”. On the other

hand, we can use Theorem 7.0.11 to bound the amount of interaction from above. Hence we obtain a nontrivial inequality. All quantities there work out well as they supposed to be and we obtain our theorem.

Proof of Theorem 8.0.12. Apparently, we only need to prove the case where each H_j is compact and take a limiting argument to complete the proof. Fix a large constant N in terms of d , for example $N = 100e^d$ should be more than sufficient.

Consider the standard lattice of unit cubes in $\mathbb{R}^d$. For each cube Q_ν in the lattice, let O_ν be its center. Let

$$G(Q_\nu) = \int_{H_1 \times H_2 \times \dots \times H_n} \prod_{j=1}^n \chi_{\{\text{dist}(y_j, O_\nu) \leq N\}} | \wedge_{j=1}^n T_{y_j} H_j | d\sigma_1(y_1) \cdots d\sigma_n(y_n). \quad (8.2)$$

Obviously

$$G(Q_\nu) \geq \int_{H_1 \times H_2 \times \dots \times H_n} \prod_{j=1}^n \chi_{\{\text{dist}(y_j, x) \leq 1\}} | \wedge_{j=1}^n T_{y_j} H_j | d\sigma_1(y_1) \cdots d\sigma_n(y_n). \quad (8.3)$$

for any $x \in Q_\nu$. Hence it suffices to prove that under assumptions of Theorem 8.0.12, we have

$$\sum_{\nu} G(Q_\nu)^{\frac{1}{n-1}} \lesssim_d \prod_{j=1}^m A(j)^{\frac{1}{n-1}}. \quad (8.4)$$

We only have finite many relevant cubes Q_ν such that $G(Q_\nu) \neq 0$. Hence we can choose a huge cube of side length S containing all of the relevant cubes. By Guth's Lemma 4.0.6, we can find a polynomial P of degree $\lesssim_d S$ such that for each cube Q_ν ,

$$\overline{Vis}[Z(P) \cap Q_\nu] \geq S^d G(Q_\nu)^{\frac{1}{n-1}} \left(\sum_{\nu} G(Q_\nu)^{\frac{1}{n-1}} \right)^{-1}. \quad (8.5)$$

Adding $\lesssim_d S$ hyperplanes to P (in other words multiply P by linear equations of those hyperplanes) if necessary, we may assume that for all Q_ν where $G(Q_\nu) > 0$ we have $\bar{V}_{Z(P) \cap Q_\nu}(\mathbf{v}) \geq |\mathbf{v}|$. Hence we are in a good position to use the wedge product estimate (5.2).

Before we move on let me remark on a technical issue. If we do have to add hyperplanes at this point, we need to modify our Definition 4.0.5 a little bit: Assume all the hyperplanes we added form a zero set of a polynomial P_0 . We call the original polynomial in Guth's lemma P_{old} and our P is actually $P_{old}P_0$. Then when we are talking about the mollified directed volume, mollified visibility, etc. around P , we want to look at all $P'P_0$ where $P' \in B(P_{old}, \varepsilon)$ instead of all $P' \in B(P, \varepsilon)$. For example, the definition (4.3) should now be modified to

$$\bar{V}_{Z(P) \cap U}(\mathbf{v}) = \frac{1}{|B(P_{old}, \varepsilon)|} \int_{B(P_{old}, \varepsilon)} V_{Z(P'P_0) \cap U}(\mathbf{v}) dP'. \quad (8.6)$$

We also note that an alternative strategy to “adding hyperplanes” is given in the Carbery-Valdimarsson paper [CV13] (see Lemma 3, Lemma 6 and the argument on page 1654 there). It is a very detailed and clear account.

For the rest of the section for simplicity of the notations, we deal with the case where no hyperplanes are added. For the general case the proof is identical except for proper correction of notations.

For any $y_1 \in H_1 \cap B(O_\nu, N), \dots, y_n \in H_n \cap B(O_\nu, N)$, $P_1, \dots, P_d \in B(P, \varepsilon)$ (see section 2), $p_1 \in Z(P_1) \cap B(O_\nu, N), \dots, p_d \in Z(P_d) \cap B(O_\nu, N)$, by Lemma 6.0.8, we

can find some $i_{j,h}$ for $1 \leq j \leq n$, $1 \leq h \leq k_j$ such that

$$\prod_{j=1}^n |(T_{y_j} H_j)^\perp \wedge (T_{p_{i_{j,1}}} Z(P_{i_{j,1}}))^\perp \wedge \cdots \wedge (T_{p_{i_{j,k_j}}} Z(P_{i_{j,k_j}}))^\perp| \gtrsim_d |\wedge_{j=1}^n T_{y_j} H_j| \cdot |\wedge_{i=1}^d (T_{p_i} Z(P_i))^\perp| \quad (8.7)$$

and that all $i_{j,h}$ are distinct and form exactly the set $\{1, 2, \dots, d\}$.

Integrate over $(H_1 \cap B(O_\nu, N)) \times \cdots \times (H_n \cap B(O_\nu, N))$, we obtain

$$\begin{aligned} & G(Q_\nu) \cdot |\wedge_{i=1}^d (T_{p_i} Z(P_i))^\perp| \\ & \lesssim_d \sum_{(i_{j,h})} \int_{H_1 \cap B(O_\nu, N)} \cdots \int_{H_n \cap B(O_\nu, N)} \\ & \prod_{j=1}^n |(T_{y_j} H_j)^\perp \wedge (T_{p_{i_{j,1}}} Z(P_{i_{j,1}}))^\perp \wedge \cdots \wedge (T_{p_{i_{j,k_j}}} Z(P_{i_{j,k_j}}))^\perp| d\sigma_1(y_1) \cdots d\sigma_n(y_n) \end{aligned} \quad (8.8)$$

Here we sum over all possible choices of $\{i_{j,h}, 1 \leq j \leq n, 1 \leq h \leq k_j\}$ such that all $i_{j,h}$ are distinct and form exactly the set $\{1, 2, \dots, d\}$.

Integrate (8.8) over $P_1, \dots, P_d \in B(P, \varepsilon)$ and $p_i \in Z(P_i) \cap B(O_\nu, N)$ (we abuse the notation a bit and write $dp = d\sigma(p)$ where $d\sigma$ is the $(d-1)$ -dimensional Hausdorff volume measure on $Z(P)$). Taking Definition 4.0.5 into account, we use Wedge Product Estimate Theorem 5.0.7 and (8.5), (8.8) and deduce

$$\begin{aligned} & \sum_{(i_{j,h})} \frac{1}{|B(P, \varepsilon)|^d} \int \cdots \int_{B(P, \varepsilon)^d} \int_{H_1 \cap B(O_\nu, N)} \cdots \int_{H_n \cap B(O_\nu, N)} \int_{Z(P_1) \cap B(O_\nu, N)} \cdots \int_{Z(P_d) \cap B(O_\nu, N)} \\ & \prod_{j=1}^n |(T_{y_j} H_j)^\perp \wedge (T_{p_{i_{j,1}}} Z(P_{i_{j,1}}))^\perp \wedge \cdots \wedge (T_{p_{i_{j,k_j}}} Z(P_{i_{j,k_j}}))^\perp| \\ & dp_1 \cdots dp_d d\sigma_1(y_1) \cdots d\sigma_n(y_n) dP_1 \cdots dP_d \\ & \gtrsim_d G(Q_\nu) \cdot \overline{Vis}[Z(P) \cap Q_\nu] \\ & \gtrsim_d S^d G(Q_\nu)^{\frac{n}{n-1}} (\sum_\nu G(Q_\nu)^{\frac{1}{n-1}})^{-1}. \end{aligned} \quad (8.9)$$

Rewrite (8.9) as the following

$$\begin{aligned}
& \sum_{(i,j,h)} \prod_{j=1}^n \left(\frac{1}{|B(P,\varepsilon)|^{k_j}} \int \cdots \int_{B(P,\varepsilon)^{k_j}} \frac{1}{S^{k_j} \cdot A(j)} \int_{H_j} \cap B(O_\nu, N) \int_{Z(P_{i,j,1})} \cap B(O_\nu, N) \cdots \int_{Z(P_{i,j,k_j})} \cap B(O_\nu, N) \right. \\
& \left. |(T_{y_j} H_j)^\perp \wedge (T_{p_{i,j,1}} Z(P_{i,j,1}))^\perp \wedge \cdots \wedge (T_{p_{i,j,k_j}} Z(P_{i,j,k_j}))^\perp| dp_{i,j,1} \cdots dp_{i,j,k_j} d\sigma_j(y_j) dP_{i,j,1} \cdots dP_{i,j,k_j} \right) \\
& \gtrsim_d \frac{1}{\prod_{j=1}^n A(j)} G(Q_\nu)^{\frac{n}{n-1}} (\sum_\nu G(Q_\nu)^{\frac{1}{n-1}})^{-1}. \tag{8.10}
\end{aligned}$$

By the arithmetic-geometric mean inequality we deduce

$$\begin{aligned}
& \frac{1}{(\prod_{j=1}^n A(j))^{\frac{1}{n}}} G(Q_\nu)^{\frac{1}{n-1}} (\sum_\nu G(Q_\nu)^{\frac{1}{n-1}})^{-\frac{1}{n}} \\
& \lesssim_d \sum_{(i,j,h)} \sum_{j=1}^n \frac{1}{|B(P,\varepsilon)|^{k_j}} \int \cdots \int_{B(P,\varepsilon)^{k_j}} \frac{1}{S^{k_j} \cdot A(j)} \int_{H_j} \cap B(O_\nu, N) \int_{Z(P_{i,j,1})} \cap B(O_\nu, N) \cdots \int_{Z(P_{i,j,k_j})} \cap B(O_\nu, N) \\
& \quad |(T_{y_j} H_j)^\perp \wedge (T_{p_{i,j,1}} Z(P_{i,j,1}))^\perp \wedge \cdots \wedge (T_{p_{i,j,k_j}} Z(P_{i,j,k_j}))^\perp| dp_{i,j,1} \cdots dp_{i,j,k_j} d\sigma_j(y_j) dP_{i,j,1} \cdots dP_{i,j,k_j} \\
& \tag{8.11}
\end{aligned}$$

Sum (8.11) over ν , and then we invoke the Intersection Estimate Theorem 7.0.11 with $U = \{(u_i)_{1 \leq i \leq k_j+1} : u_i \in \mathbb{R}^d, \text{dist}(u_i, u_{i'}) < N^2\}$ (it suffices to choose U large enough). Note that $\deg P_j = S$ and $\deg H_j = A(j)$, we have

$$\begin{aligned}
& \frac{1}{(\prod_{j=1}^n A(j))^{\frac{1}{n}}} (\sum_\nu G(Q_\nu)^{\frac{1}{n-1}}) (\sum_\nu G(Q_\nu)^{\frac{1}{n-1}})^{-\frac{1}{n}} \\
& \lesssim_d \sum_{(i,j,h)} \sum_{j=1}^n \frac{1}{|B(P,\varepsilon)|^{k_j}} \int \cdots \int_{B(P,\varepsilon)^{k_j}} \frac{1}{S^{k_j} \cdot A(j)} \int_{H_j} \int_{Z(P_{i,j,1})} \cdots \int_{Z(P_{i,j,k_j})} \chi_U(y_j, p_{i,j,1}, \dots, p_{i,j,k_j}) \\
& \quad \cdot |(T_{y_j} H_j)^\perp \wedge (T_{p_{i,j,1}} Z(P_{i,j,1}))^\perp \wedge \cdots \wedge (T_{p_{i,j,k_j}} Z(P_{i,j,k_j}))^\perp| dp_{i,j,1} \cdots dp_{i,j,k_j} d\sigma_j(y_j) dP_{i,j,1} \cdots dP_{i,j,k_j} \\
& \lesssim_d \sum_{(i,j,h)} \sum_{j=1}^n \frac{1}{|B(P,\varepsilon)|^{k_j}} \int \cdots \int_{B(P,\varepsilon)^{k_j}} 1 dP_{i,j,1} \cdots dP_{i,j,k_j} \\
& \lesssim_d 1 \tag{8.12}
\end{aligned}$$

and (8.4) holds. □

Theorem 3.1.3 and Theorem 3.1.2 follow easily from Theorem 8.0.12. To prove Theorem 3.1.3 it suffices to prove the case where all ρ_{j,a_j} are rational numbers. Then without loss of generality we may assume further that they are integers. By considering multiple copies of U_{j,a_j} 's, we can further assume they are all 1. Then one just takes the j -th variety to be the union of the cores of the j -th family of slabs and apply Theorem 8.0.12 (after a scaling). Theorem 3.1.2 is a direct corollary of Theorem 3.1.3.

Chapter 9

An Analogue of Lemma 6.0.8

In the next two chapters, we prove Theorem 1.2.7. In this chapter we prove a lemma (Lemma 9.0.17) analogous to Lemma 6.0.8 which will be used in the proof the same way as Lemma 6.0.8 was used in the proof of Theorem 3.1.2. This lemma is weaker in appearance than Lemma 6.0.8, but it turns out that it serves our purpose.

Definition 9.0.13. *In $\mathbb{R}^d$, given a convex body Γ centered at the origin, define its dual body Γ^* to be $\{\mathbf{v} \in \mathbb{R}^d : |(\mathbf{v}, \mathbf{x})| \leq 1 \text{ for all } \mathbf{x} \in \Gamma\}$, where $(\cdot, \cdot)$ is the Euclidean inner product on $\mathbb{R}^d$.*

It is trivial by definition that if two convex bodies $\Gamma_1 \subseteq \Gamma_2$ then $\Gamma_1^* \supseteq \Gamma_2^*$.

By John's Ellipsoid Theorem, we need to mainly consider ellipsoids as examples of convex bodies. Next we develop several lemmas concerning ellipsoids. From now on, when we talk about an ellipsoid in an Euclidean space, we always assume the ellipsoid has the same dimension as the background space.

Lemma 9.0.14. *If the Γ in Definition 9.0.13 is a (closed) ellipsoid centered at O (the origin), then Γ^* is also an ellipsoid centered at O . We call Γ^* the dual ellipsoid of Γ . Choose a set of principal axes of Γ (the wording is because the choices might not be unique), we have them to be again the principal axes of Γ^* . Moreover, the lengths of the corresponding principal axes of Γ and Γ^* , when divided by 2, are reciprocal to each other. Hence $(\Gamma^*)^* = \Gamma$ and $\text{Vol}(\Gamma) \cdot \text{Vol}(\Gamma^*) = C_d > 0$ is a constant depending only on d .*

Proof. Trivially the dual body of the unit ball is again the unit ball. Assume Γ_0 has a dual body Γ_0^* . Then for any positive definite linear transform A , we have by definition

$$\begin{aligned}
(A\Gamma_0)^* &= \{ \mathbf{v} \in \mathbb{R}^d : |(\mathbf{v}, A\mathbf{x})| \leq 1 \text{ for all } x \in \Gamma_0 \} \\
&= \{ \mathbf{v} \in \mathbb{R}^d : |(A^*\mathbf{v}, \mathbf{x})| \leq 1 \text{ for all } x \in \Gamma_0 \} \\
&= (A^*)^{-1}\Gamma_0^* \\
&= A^{-1}\Gamma_0^*.
\end{aligned} \tag{9.1}$$

Now we can use a positive definite linear transformation A to transform the closed unit ball to our Γ , by the computation above, Γ^* is A^{-1} acting on the unit ball, so it is an ellipsoid. Also the principal axes of Γ correspond to an orthonormal basis that diagonalizes A . This basis also diagonalizes A^{-1} . Hence the principal axes of Γ are also principal axes of Γ^* . The rest of the lemma is obvious. $\square$

Lemma 9.0.15. *Assume a subspace $V \subseteq \mathbb{R}^d$ and $\Gamma \in \mathbb{R}^d$ is an ellipsoid centered at O . Let $\pi_V(\cdot)$ be the orthogonal projection onto V . Then $\pi_V(\Gamma^*)$ and $\Gamma \cap V$ are dual*

to each other (in V with respect of the induced inner product). Note these two are both ellipsoids.

Proof. If V has dimension 1, then the lemma is true by definition of the dual body (note by Lemma 9.0.14, the two ellipsoids are dual to each other).

For general V , by the last paragraph for any $V' \subseteq V$ of dimension 1 we have $\pi_{V'}(\pi_V(\Gamma^*))$ and $(\Gamma \cap V) \cap V'$ are dual to each other. But given the ellipsoid $\Gamma_V = \Gamma \cap V \subseteq V$, apparently there is only one possible set $Y \subseteq V$ such that for any $V' \subseteq V$ of dimension 1, $\pi_{V'}(\Gamma_V)$ and $Y \cap V'$ are dual to each other (since $Y \cap V'$ is determined by Γ_V via this property). Now by last paragraph again, the dual of Γ_V in V is this unique Y . Hence $\pi_V(\Gamma^*)$ has to be this dual. $\square$

Lemma 9.0.16. *For any subspace $V \subseteq \mathbb{R}^d$ of dimension d' , we define π_V to be the orthogonal projection onto V as usual. Then for any (closed) ellipsoid $\Gamma \subseteq \mathbb{R}^d$, we have*

$$|\pi_V(\Gamma)| |\Gamma \cap V^\perp| = C_{d,d'} |\Gamma| \quad (9.2)$$

where $C_{d,d'} > 0$ only depends on d and d' . Here we use the corresponding standard Lebesgue measure on V , $V^\perp$ and $\mathbb{R}^d$, respectively.

Proof. It is well known that in $\mathbb{R}^d$, an ellipsoid defined by $\{\mathbf{x} : (\mathbf{x}, A\mathbf{x}) \leq 1\}$ has volume $\frac{C_d}{(\det A)^{\frac{1}{2}}}$, where A is a positive definite linear transform and $(\cdot, \cdot)$ is the Euclidean inner product. Assume $\Gamma = \{\mathbf{x} : \|T\mathbf{x}\|^2 \leq 1\}$ where T is a non-degenerate linear transform. Since we can multiply T by any orthogonal transform on the left,

we may assume $TV^\perp = V^\perp$. Then by last paragraph,

$$|\Gamma| = \frac{C_d}{|\det T|} \quad (9.3)$$

$$|\Gamma \cap V^\perp| = \frac{C_{d,d'}}{|\det T|_{V^\perp}}. \quad (9.4)$$

Meanwhile, $\mathbf{x} \in V$ belongs to $\pi_V(\Gamma)$ if and only if $\inf_{\mathbf{v} \in V^\perp} \|T(\mathbf{x} + \mathbf{v})\| \leq 1$. By the method of least squares, $\inf_{\mathbf{v} \in V^\perp} \|T(\mathbf{x} + \mathbf{v})\| = \|\pi_{(TV^\perp)^\perp}(T\mathbf{x})\| = \|\pi_V(T\mathbf{x})\|$. Hence

$$|\pi_V(\Gamma)| = \frac{C_{d'}}{|\det(\pi_V \circ T|_V)|}. \quad (9.5)$$

Now notice $\pi_V V^\perp = V^\perp$. Hence when written in matrix form it is easy to verify $\det T|_{V^\perp} \cdot \det(\pi_V \circ T|_V) = \det T$. This together with (9.3), (9.4), (9.5) implies (9.2). $\square$

Now we are ready to develop an analogue of Lemma 6.0.8. We recall that in Section 3 we defined the total absolute inner product $V_{X,f}(\mathbf{v})$, the fading zone $F(X, f)$, visibility $Vis[X, f]$, and chose an elliptical approximation $Ell(F(X, f))$ for any measurable vector-valued function $f : X \rightarrow \mathbb{R}^d$.

Lemma 9.0.17. *Fix positive integers d and $1 \leq k_1, \dots, k_n < d$. Let $\mathbb{R}^d$ be the standard Euclidean space.*

Assume a Brascamp-Lieb datum $(\mathbf{B}, \mathbf{p})$ that all B_j are orthogonal projections from $\mathbb{R}^d$ to a subspace and that $\dim \ker B_j = k_j$. Assume $E_j = B_j(\mathbb{R}^d) = (\ker B_j)^\perp$. Assume we have the scaling condition $\sum_{j=1}^n p_j \dim E_j = d$.

For any measurable vector valued function $f : X \rightarrow \mathbb{R}^d$ on some measure space

satisfying $V_{X,f}(\mathbf{v}) \geq 1$ for all $\mathbf{v} \in \mathbb{R}^d$, we have

$$\prod_{j=1}^n \left(\int_{X^{k_j}} |E_j \wedge f(x_1) \wedge \cdots \wedge f(x_{k_j})| dx_1 \cdots dx_{k_j} \right)^{p_j} \gtrsim_{d,\mathbf{p}} (BL(\mathbf{B}, \mathbf{p}))^{-1} (Vis[X, f])^{\sum_{j=1}^n p_j - 1}. \quad (9.6)$$

Proof. Similar to the proof of Theorem 5.0.7, we define $\pi_{\ker B_j}$ to be the orthogonal projection onto $\ker B_j$ as before and $f_{\ker B_j} = \pi_{\ker B_j} \circ f$. Then

$$\int_{X^{k_j}} |E_j \wedge f(x_1) \wedge \cdots \wedge f(x_{k_j})| dx_1 \cdots dx_{k_j} = \int_{X^{k_j}} |f_{\ker B_j}(x_1) \wedge \cdots \wedge f_{\ker B_j}(x_{k_j})| dx_1 \cdots dx_{k_j} \quad (9.7)$$

Similar to the proof of Theorem 5.0.7, we know $F(X, f_{\ker B_j}) = F(X, f) \cap \ker B_j$. Hence we can take $Ell(F(X, f_{\ker B_j}))$ to be $Ell(F(X, f)) \cap \ker B_j$. By (9.7), Theorem 5.0.7, Lemma 9.0.14 and Lemma 9.0.15,

$$\begin{aligned} & \int_{X^{k_j}} |E_j \wedge f(x_1) \wedge \cdots \wedge f(x_{k_j})| dx_1 \cdots dx_{k_j} \\ &= \int_{X^{k_j}} |f_{\ker B_j}(x_1) \wedge \cdots \wedge f_{\ker B_j}(x_{k_j})| dx_1 \cdots dx_{k_j} \\ &\gtrsim_d \frac{1}{|Ell(F(X, f)) \cap \ker B_j|} \\ &\gtrsim_d |(Ell(F(X, f)) \cap \ker B_j)^*| \\ &= |\pi_{\ker B_j}(Ell(F(X, f))^*)|. \end{aligned} \quad (9.8)$$

Hence it suffices to prove

$$\prod_{j=1}^n |\pi_{\ker B_j}(Ell(F(X, f))^*)|^{p_j} \gtrsim_{d,\mathbf{p}} (BL(\mathbf{B}, \mathbf{p}))^{-1} (Vis[X, f])^{\sum_{j=1}^n p_j - 1}. \quad (9.9)$$

At this point we invoke the definition of $BL(\mathbf{B}, \mathbf{p})$. For any ellipsoid Γ , we choose

$f_j = \chi_{\pi_{E_j}(\Gamma^*)}$ in (1.1). Then by definition $\prod_{j=1}^n (f_j \circ B_j)^{p_j} \geq \chi_{\Gamma^*}$. Hence

$$\begin{aligned}
& |\Gamma^*| \\
& \leq \int_{\mathbb{R}^d} \prod_{j=1}^n (f_j \circ B_j)^{p_j} \\
& \leq BL(\mathbf{B}, \mathbf{p}) \prod_{j=1}^n (\int_{E_j} f_j)^{p_j} \\
& = BL(\mathbf{B}, \mathbf{p}) \prod_{j=1}^n |\pi_{E_j}(\Gamma^*)|^{p_j}.
\end{aligned} \tag{9.10}$$

In other words,

$$BL(\mathbf{B}, \mathbf{p}) \cdot |\Gamma| \cdot \prod_{j=1}^n |\pi_{E_j}(\Gamma^*)|^{p_j} \gtrsim_d 1. \tag{9.11}$$

By Lemma 9.0.14, Lemma 9.0.15 and Lemma 9.0.16, we have

$$|\pi_{E_j}(\Gamma^*)| \sim_{k_j} \frac{1}{|\Gamma \cap E_j|} \sim_{k_j, d} \frac{|\pi_{\ker B_j}(\Gamma)|}{|\Gamma|}. \tag{9.12}$$

Hence

$$BL(\mathbf{B}, \mathbf{p}) \cdot |\Gamma| \cdot \prod_{j=1}^n \left(\frac{|\pi_{\ker B_j}(\Gamma)|}{|\Gamma|} \right)^{p_j} \gtrsim_{d, \mathbf{p}} 1. \tag{9.13}$$

Take $\Gamma = Ell(F(X, f))^*$. By Lemma 9.0.14 again, we have $|\Gamma| = |Ell(F(X, f))^*| \sim_d \frac{1}{|Ell(F(X, f))|} = Vis[X, f]$. This fact and (9.13) implies (9.9), which in turn implies (9.6). $\square$

Chapter 10

Proof of Theorem 1.2.7

We are ready to prove Theorem 1.2.7. Just like the proof of Theorem 3.1.2, we prove a stronger theorem concerning algebraic varieties. This theorem can also be considered as an analogue of Theorem 8.0.12.

Theorem 10.0.18 (Variety Version of Brascamp-Lieb). *Assume we have positive integers $k_1, \dots, k_n \leq d$ and rational numbers $p_1, \dots, p_n > 0$. Choose a common denominator τ of all p_j and assume $p_j = \frac{\tau_j}{\tau}$, $\tau_j \in \mathbb{Z}^+$ satisfying the scaling condition $\sum_j p_j(d - k_j) = d$.*

Assume that for $1 \leq j \leq n$, $H_j \subseteq \mathbb{R}^d$ is part of a k_j -dimensional algebraic subvariety of degree $A(j)$, respectively. Let $d\sigma_j$ denote the k_j -dimensional (Hausdorff) volume measure of H_j . Then under this measure, almost all $y_j \in H_j$ are smooth points. For a smooth point $y_j \in H_j$, let $T_{y_j}H_j$ denote the tangent space of H_j at y_j .

For $(\sum_j \tau_j)$ smooth points $\mathbf{y} = (y_{1,1}, \dots, y_{1,\tau_1}, y_{2,1}, \dots, y_{2,\tau_2}, \dots, y_{n,\tau_n})$, $y_{j,l} \in H_j$, there exists a unique Brascamp-Lieb datum $(\mathbf{B}(\mathbf{y}), \mathbf{p}(\mathbf{y}))$ with $(\sum_j \tau_j)$ projections B_j

all being orthogonal projections within $\mathbb{R}^d$ as the following: Define $(\mathbf{B}(\mathbf{y}), \mathbf{p}(\mathbf{y})) = (B_{1,1}, \dots, B_{1,\tau_1}, B_{2,1}, \dots, B_{2,\tau_2}, \dots, B_{n,\tau_n}, \frac{1}{\tau}, \dots, \frac{1}{\tau})$ such that $\ker B_{j,l} = T_{y_{j,l}} H_j$ and all components of $\mathbf{p}$ are $\frac{1}{\tau}$. Then

$$\begin{aligned} & \int_{\mathbb{R}^d} \left(\int_{H_1^{\tau_1} \times \dots \times H_n^{\tau_n}} \prod_{j=1}^n \prod_{k=1}^{\tau_j} \chi_{\{\text{dist}(y_{j,k}, x) \leq 1\}} BL(\mathbf{B}(\mathbf{y}), \mathbf{p}(\mathbf{y}))^{-\tau} \right. \\ & \quad \left. d\sigma_1(y_{1,1}) \cdots d\sigma_1(y_{1,\tau_1}) \cdots d\sigma_n(y_{n,\tau_n}) \right)^{\frac{1}{\tau}} dx \\ & \lesssim_{d, \tau_1, \dots, \tau_n, \tau} \prod_{j=1}^n A(j)^{p_j}. \end{aligned} \tag{10.1}$$

Let us explain the motivation of Theorem 10.0.18 before proving it. If we want to naturally generalize Theorem 8.0.12 to the Brascamp-Lieb setting, first of all we have to come up with a reasonable integral like the left hand side of (8.1) to put on the left hand side. However the fact that in (8.1) all $p_j = \frac{1}{n-1}$ no longer holds in our situation. In fact, p_j 's might even all be irrational numbers. A natural way would be approximating (p_j) by rational tuples. this works (see below) but eventually we need all p_j 's to be the same to get an analogous quantity to left hand of (8.1).

Another remark before we move on. It's good to keep in mind that we may assume $\tau_1 = \dots = \tau_n = 1$ in this theorem without loss of generality. This is trivial to see. But we keep the theorem in its current form here so it would be more straightforward to apply.

Proof that Theorem 10.0.18 implies Theorem 1.2.7. Note that the conditions (1.3) and (1.4) only have rational coefficients. Hence it is possible to choose $(n+1)$ different rational $\mathbf{p}'$ close enough to $\mathbf{p}$ such that the conditions (1.3) and (1.4) are satisfied (that is, $BL(\mathbf{B}, \mathbf{p}') < +\infty$), and that $\mathbf{p}$ lies in the convex hull of those $\mathbf{p}'$.

By interpolation we only need to prove the case when $\mathbf{p}$ is a rational vector.

Next in order to apply the result of Theorem 10.0.18 to prove Theorem 1.2.7, we claim that if a Brascamp-Lieb datum $(\mathbf{B}, \mathbf{p})$ is such that $p_j = \frac{\tau_j}{\tau}$ where τ all τ_j are positive integers, then $BL(\mathbf{B}, \mathbf{p}) = BL(\mathbf{B}', \mathbf{p}')$, where $\mathbf{B}' = (B_1, \dots, B_1, \dots, B_n, \dots, B_n)$ containing τ_j copies of B_j , and $\mathbf{p}' = (\frac{1}{\tau}, \dots, \frac{1}{\tau})$. In fact, look at the definition (1.1) of $BL(\mathbf{B}, \mathbf{p})$, we have

$$BL(\mathbf{B}', \mathbf{p}') = \sup_{\{f_{j,l}\}} \frac{\int_{\mathbb{R}^d} \prod_{j=1}^n \prod_{l=1}^{\tau_j} (f_{j,l} \circ B_j)^{\frac{1}{\tau}}}{\prod_{j=1}^n \prod_{l=1}^{\tau_j} (\int_{H_j} f_{j,l})^{\frac{1}{\tau}}}. \quad (10.2)$$

Since we can always take $f_{j,l} = f_j$ for all l , we deduce $BL(\mathbf{B}', \mathbf{p}') \geq BL(\mathbf{B}, \mathbf{p})$. On the other hand, in the definition of $BL(\mathbf{B}, \mathbf{p})$ we can take $f_j = f_{j,l_j}$ for every possible tuple $(l_1, \dots, l_n)$ satisfying $1 \leq l_j \leq \tau_j$ to deduce

$$\int_{\mathbb{R}^d} \prod_{j=1}^n (f_{j,l_j} \circ B_j)^{\frac{\tau_j}{\tau}} \leq BL(\mathbf{B}, \mathbf{p}) \prod_{j=1}^n \left(\int_{H_j} f_{j,l_j} \right)^{\frac{\tau_j}{\tau}} \quad (10.3)$$

Then we let (l_j) run through all possible tuples and invoke Hölder to conclude that

$$\int_{\mathbb{R}^d} \prod_{j=1}^n \prod_{l=1}^{\tau_j} (f_{j,l} \circ B_j)^{\frac{1}{\tau}} \leq BL(\mathbf{B}, \mathbf{p}) \prod_{j=1}^n \prod_{l=1}^{\tau_j} \left(\int_{H_j} f_{j,l} \right)^{\frac{1}{\tau}}. \quad (10.4)$$

Hence $BL(\mathbf{B}', \mathbf{p}') \leq BL(\mathbf{B}, \mathbf{p})$. Therefore $BL(\mathbf{B}', \mathbf{p}') = BL(\mathbf{B}, \mathbf{p})$.

By Theorem 1.1 in [BBFL15], BL is a locally bounded function. It is then not hard to derive Theorem 1.2.7 from Theorem 10.0.18 when $\mathbf{p}'$ is a fixed rational number. $\square$

Proof of Theorem 10.0.18. It's plain that we may assume $\tau_1 = \dots = \tau_n = 1$. For

short we write $B_j = B_{j,1}$ and $y_j = y_{j,1}$.

The proof will be almost identical to that of Theorem 8.0.12. In the current proof, we omit some details for familiar manipulations in that proof to reduce redundancy and refer the reader to it.

Take the N and set up the unit cube lattice in $\mathbb{R}^d$ as in the proof of Theorem 8.0.12. Again let O_ν be the center of any cube Q_ν in the lattice. This time we define

$$G(Q_\nu) = \int_{H_1 \times \dots \times H_n} \prod_{j=1}^n \chi_{\text{dist}(y_j, O_\nu) \leq N} BL(\mathbf{B}(\mathbf{y}), \mathbf{p}(\mathbf{y}))^{-\tau} d\sigma_1(y_1) \cdots d\sigma_n(y_n). \quad (10.5)$$

Similar to the proof of Theorem 8.0.12, it suffices to show

$$\sum_{\nu} G(Q_\nu)^{\frac{1}{\tau}} \lesssim_{d,n} \prod_{j=1}^n A(j)^{\frac{1}{\tau}}. \quad (10.6)$$

Again we may assume for the moment that each H_j is compact and use a limiting argument. Then we can again choose a large cube of side length S that contains all the relevant cubes. Finally we can find a polynomial P of degree $\lesssim_d S$ such that for each Q_ν ,

$$\overline{Vis}[Z(P) \cap Q_\nu] \geq S^d G(Q_\nu)^{\frac{1}{\tau}} \left(\sum_{\nu} G(Q_\nu)^{\frac{1}{\tau}} \right)^{-1}. \quad (10.7)$$

As before we have to make the technical comment that after adding some hyperplanes and changing the definition of $\overline{Vis}$ accordingly, we may assume for all Q_ν with $G(Q_\nu) > 0$ we have $\overline{V}_{Z(P) \cap Q_\nu}(\mathbf{v}) \geq |\mathbf{v}|$ (so that we are allowed to apply (9.6)). We only deal with the case where no hyperplanes are added so the notation would be simpler.

Similar to what we did in the proof of Theorem 8.0.12, we choose $B_j = T_{y_j} H_j$, all $p_j = \frac{1}{\tau}$ and integrate (9.6) over $y_j \in H_j \cap B(O_\nu, N)$. Then we choose the measure space X in (9.6) to be $\{p \in Z(P') \cap B(O_\nu, N) : P' \in B(P, \varepsilon)\}$ (the measure is just the surface measure on each $Z(P')$ joint with the standard measure on $B(P, \varepsilon)$, which is $\mathrm{d}p \mathrm{d}P'$ where $P' \in B(P, \varepsilon)$ and $p \in Z(P') \cap B(O_\nu, N)$) and deduce

$$\begin{aligned}
& \frac{1}{|B(P, \varepsilon)|^{(n-\tau)d}} \int \cdots \int_{B(P, \varepsilon)^{(n-\tau)d}} \int_{H_1 \cap B(O_\nu, N)} \cdots \int_{H_n \cap B(O_\nu, N)} \int_{Z(P_1) \cap B(O_\nu, N)} \cdots \int_{Z(P_{(n-\tau)d}) \cap B(O_\nu, N)} \\
& \quad \prod_{j=1}^n |(T_{y_j} H_j)^\perp \wedge (T_{p_{k_1+\dots+k_{j-1}+1}} Z(P_{k_1+\dots+k_{j-1}+1}))^\perp \wedge \cdots \wedge (T_{p_{k_1+\dots+k_j}} Z(P_{k_1+\dots+k_j}))^\perp| \\
& \quad \mathrm{d}p_1 \cdots \mathrm{d}p_{(n-\tau)d} \mathrm{d}\sigma_1(y_1) \cdots \mathrm{d}\sigma_n(y_n) \mathrm{d}P_1 \cdots \mathrm{d}P_{(n-\tau)d} \\
& \gtrsim_{d,n} G(Q_\nu) \cdot \overline{Vis}[Z(P) \cap Q_\nu]^{n-\tau} \\
& \gtrsim_{d,n} S^{(n-\tau)d} G(Q_\nu)^{\frac{n}{\tau}} (\sum_\nu G(Q_\nu)^{\frac{1}{\tau}})^{-(n-\tau)}. \tag{10.8}
\end{aligned}$$

As before we rewrite it as

$$\begin{aligned}
& \prod_{j=1}^n \left(\frac{1}{|B(P, \varepsilon)|^{k_j}} \int \cdots \int_{B(P, \varepsilon)^{k_j}} \frac{1}{S^{k_j \cdot A(j)}} \int_{H_j \cap B(O_\nu, N)} \int_{Z(P_1) \cap B(O_\nu, N)} \cdots \int_{Z(P_{k_j}) \cap B(O_\nu, N)} \right. \\
& \quad \left. |(T_{y_j} H_j)^\perp \wedge (T_{p_1} Z(P_1))^\perp \wedge \cdots \wedge (T_{p_{k_j}} Z(P_{k_j}))^\perp| \mathrm{d}p_1 \cdots \mathrm{d}p_{k_j} \mathrm{d}\sigma_j(y_j) \mathrm{d}P_1 \cdots \mathrm{d}P_{k_j} \right) \\
& \gtrsim_{d,n} \frac{1}{\prod_{j=1}^n A(j)} G(Q_\nu)^{\frac{n}{\tau}} (\sum_\nu G(Q_\nu)^{\frac{1}{\tau}})^{-(n-\tau)}. \tag{10.9}
\end{aligned}$$

Here note that since $\sum_{j=1}^n (d - k_j) = \tau d$ by assumption, we have $\sum_{j=1}^n k_j = (n - \tau)d$.

We have used this fact in the above inequality chain (10.9).

By arithmetic-geometric mean inequality we have

$$\begin{aligned}
& \sum_{j=1}^n \left(\frac{1}{|B(P, \varepsilon)|^{k_j}} \int \cdots \int_{B(P, \varepsilon)^{k_j}} \frac{1}{S^{k_j} \cdot A(j)} \int_{H_j} \int_{B(O_\nu, N)} \int_{Z(P_1) \cap B(O_\nu, N)} \cdots \int_{Z(P_{k_j}) \cap B(O_\nu, N)} \right. \\
& \quad \left. |(T_{y_j} H_j)^\perp \wedge (T_{p_1} Z(P_1))^\perp \wedge \cdots \wedge (T_{p_{k_j}} Z(P_{k_j}))^\perp| dp_1 \cdots dp_{k_j} d\sigma_j(y_j) dP_1 \cdots dP_{k_j} \right) \\
& \gtrsim_{d,n} \frac{1}{(\prod_{j=1}^n A(j))^{\frac{1}{n}}} G(Q_\nu)^{\frac{1}{\tau}} (\sum_\nu G(Q_\nu)^{\frac{1}{\tau}})^{-\frac{n-\tau}{n}}. \tag{10.10}
\end{aligned}$$

Like we did in the proof of Theorem 8.0.12, summing over ν and applying Intersection Estimate Theorem 7.0.11 with $U = \{(u_i)_{1 \leq i \leq k_j+1} : u_i \in \mathbb{R}^d, \text{dist}(u_i, u'_i) < N^2\}$, we deduce

$$\frac{1}{(\prod_{j=1}^n A(j))^{\frac{1}{n}}} (\sum_\nu G(Q_\nu)^{\frac{1}{\tau}}) (\sum_\nu G(Q_\nu)^{\frac{1}{\tau}})^{-\frac{n-\tau}{n}} \lesssim_{d,n} 1 \tag{10.11}$$

which implies (10.6) and concludes the proof. □

Remark 10.0.19. *For the perturbed Brascamp-Lieb Theorem 1.2.7 itself, it is conceivable that one can directly work with the framework of arguments in [CV13], without applying a rational approximation argument as we did in this section. Nevertheless, we still decide to keep the current approach as we feel that Theorem 10.0.18 here may be of independent interest, and that rationality seems indispensable for us to state the theorem (and prove it).*

Chapter 11

Verifying the Brascamp-Lieb Condition for Parsell-Vinogradov Systems

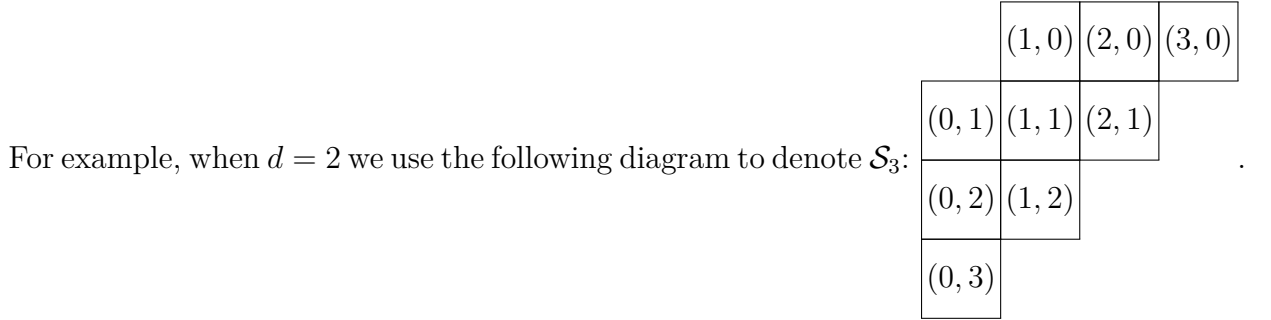
In this final chapter, we prove a conjecture of Bourgain-Demeter-Guo in [BDG16a] about Brascamp-Lieb conditions we need to check in the decoupling approach to the higher dimensional Parsell-Vinogradov systems. As advertised in the introduction, our result (Theorem 11.0.29) proves Conjecture 1.3 in [BDG16a].

Despite the technical nature of the result, intuitively by Theorem 11.0.29 we verify here that different pieces of the underlying submanifold are sufficiently transverse. It seems hopeful to the author that our approach here should readily be generalized to verify Brascamp-Lieb conditions in similar problems on translation-dilation systems.

Let a positive integer d be the dimension in our question. We usually fix it in this

section unless otherwise stated. For any positive integer l we define $[l] = \{0, 1, \dots, l\}$ for short. Note this is slightly different from the convention in combinatorics by also including 0.

Definition 11.0.20. For a d -tuple $a = (a_1, \dots, a_d)$, we define $|a| = \sum_{i=1}^d |a_i|$. For any positive integer l , define $\mathcal{S}_l = \mathcal{S}_l^d = \{a = (a_1, \dots, a_d) \in \mathbb{Z}^d : a_1, \dots, a_d \geq 0, 1 \leq |a| \leq l\} = \{(1, 0, 0, \dots, 0), (0, 1, 0, \dots, 0), (0, 0, \dots, 0, 1), \dots, (l, 0, 0, \dots, 0), (l-1, 1, 0, \dots, 0), \dots, (0, 0, \dots, 0, l)\}$. By elementary counting we have $|\mathcal{S}_l| = \binom{d+l}{l} - 1$. Denote $G_l = |\mathcal{S}_l| = \binom{d+l}{l} - 1$.



We are interested in subsets of $\mathcal{S}_l$. Hence we introduce a bit more notations.

Definition 11.0.21. We introduce a partial order on d -tuples of real numbers. For any $(b_1, b_2, \dots, b_d)$ and $(b'_1, b'_2, \dots, b'_d)$ satisfying $b_1 \leq b'_1, b_2 \leq b'_2, \dots$, and $b_d \leq b'_d$ we say $(b_1, b_2, \dots, b_d) \preceq (b'_1, b'_2, \dots, b'_d)$. And those are the only partial order relations in our definition of “ $\preceq$ ”.

Definition 11.0.22. For positive integers $l \leq l'$ and a subset $T \subseteq \mathcal{S}_l$, we denote the “positive extension” $T_{l'}^+ \subseteq \mathcal{S}_{l'}$ of T at level l' to be

$$T_{l'}^+ = \{(a_1, a_2, \dots, a_d) \in \mathcal{S}_{l'} : \exists (b_1, b_2, \dots, b_d) \in T \text{ s.t. } (b_1, b_2, \dots, b_d) \preceq (a_1, a_2, \dots, a_d)\}. \quad (11.1)$$

For example, when $d = 2$, $l = 3$, $l' = 4$, if $T = \{(1, 1), (3, 0)\}$ is the following red

colored subset in $\mathcal{S}_3$:

	(1, 0)	(2, 0)	(3, 0)
(0, 1)	(1, 1)	(2, 1)	
(0, 2)	(1, 2)		
(0, 3)			

, then T_4^+ is the following red colored

subset in $\mathcal{S}_4$:

	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	
(0, 2)	(1, 2)	(2, 2)		
(0, 3)	(1, 3)			
(0, 4)				

We sometimes want to study cube-like objects before looking at the more strange-looking and more complicated $\mathcal{S}_l$. Hence we introduce the following definition.

Definition 11.0.23. For any positive integer l , define $\mathcal{C}_l = \mathcal{C}_l^d = \{(a_1, \dots, a_d) \in \mathbb{Z}^d : 0 \leq a_i \leq l\}$. For positive integers $l \leq l'$ and a subset $T \subseteq \mathcal{C}_l$, we denote the “positive extension” $T_{l'}^+ \subseteq \mathcal{C}_{l'}$ to be

$$T_{l'}^+ = \{(a_1, a_2, \dots, a_d) \in \mathcal{C}_{l'} : \exists (b_1, b_2, \dots, b_d) \in T \text{ s.t. } (b_1, b_2, \dots, b_d) \preceq (a_1, a_2, \dots, a_d)\}. \quad (11.2)$$

We look at an example similar to the one above. If $d = 2$, assuming $T =$

$\{(1, 1), (3, 0)\}$ is the following red colored subset in $\mathcal{C}_3$:

(0, 0)	(1, 0)	(2, 0)	(3, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)
(0, 2)	(1, 2)	(2, 2)	(3, 2)
(0, 3)	(1, 3)	(2, 3)	(3, 3)

, then

T_4^+ is the following red colored subset in $\mathcal{C}_4$:

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	(4, 1)
(0, 2)	(1, 2)	(2, 2)	(3, 2)	(4, 2)
(0, 3)	(1, 3)	(2, 3)	(3, 3)	(4, 3)
(0, 4)	(1, 4)	(2, 4)	(3, 4)	(4, 4)

Lemma 11.0.24. *Let $k > 0$ be a positive integer. Assuming subsets $A, B \subseteq \mathcal{S}_k \cup \{(0, 0, \dots, 0)\}$ satisfy: For any $b = (b_1, \dots, b_d) \in B$, there is a family of inductively defined subsets as follows:*

$$R_{d;b} \subseteq [k], |R_{d;b}| = b_d.$$

$$\text{For } n_d \notin R_{d;b} \text{ we have } R_{d-1;n_d;b} \subseteq [k], |R_{d-1;n_d;b}| = b_{d-1}.$$

For $n_{d-1} \notin R_{d-1;n_d;b}$ and $n_d \notin R_{d;b}$ we have $R_{d-2;n_{d-1},n_d;b} \subseteq [k]$, $|R_{d-2;n_{d-1},n_d;b}| = b_{d-2}$, etc.

Finally for $n_2 \notin R_{2;n_3,\dots,n_d;b}$, $n_3 \notin R_{3;n_4,\dots,n_d;b}$, $\dots$, and $n_d \notin R_{d;b}$ we have $R_{1;n_2,\dots,n_d;b} \subseteq [k]$, $|R_{1;n_2,\dots,n_d;b}| = b_1$.

Moreover the above sets satisfy the following property: if some $a = (a_1, \dots, a_d) \in A$ and some $b \in B$ satisfy $a_d \notin R_{d;b}$, $a_{d-1} \notin R_{d-1;a_d;b}$, $\dots$, $a_2 \notin R_{2;a_3,\dots,a_d;b}$, then $a_1 \in R_{1;a_2,a_3,\dots,a_d;b}$.

Then

$$|A| \leq G_k + 1 - |B_k^+|. \quad (11.3)$$

Remark 11.0.25. We claim it is possible to take A to be $(\mathcal{S}_k \cup \{(0, 0, \dots, 0)\}) \setminus B_k^+$. In fact if $(a_1, \dots, a_d)$ is in such an A then by definition for any $(b_1, \dots, b_d) \in B$, one of the inequalities $0 \leq a_i < b_i$ has to hold. Therefore taking all $R_{i;*,b} = [b_i - 1]$ suffices (as a convention $[-1] = \emptyset$). Hence the above set is a possible candidate of A with $(G_k + 1 - |B_k^+|)$ elements. Thus (11.3) is actually sharp.

Still taking the previous example, when $T = 2$, $B = \{(1, 1), (3, 0)\} \in \mathcal{S}_4 \cup \{(0, 0)\}$

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	
(0, 2)	(1, 2)	(2, 2)		
(0, 3)	(1, 3)			
(0, 4)				

being the red colored pairs as the following: , then we

can take A to be the set of all blue colored pairs shown in the following diagram (where

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	
(0, 2)	(1, 2)	(2, 2)		
(0, 3)	(1, 3)			
(0, 4)				

we have already colored all pairs in B_4^+ to be red): .

Lemma 11.0.24 says we cannot have a larger A .

Lemma 11.0.24 on $\mathcal{S}_l$ and $\mathcal{S}_k$ can be deduced from the following similar Lemma

11.0.26 on $\mathcal{C}_l$ and $\mathcal{C}_k$.

Lemma 11.0.26. *Let $k > 0$ be a positive integer. Assuming subsets $A, B \subseteq \mathcal{C}_k$ satisfy that: For any $b = (b_1, \dots, b_d) \in B$, there is a family of inductively defined subsets as follows:*

$$R_{d;b} \subseteq [k], |R_{d;b}| = b_d.$$

$$\text{For } n_d \notin R_{d;b} \text{ we have } R_{d-1;n_d;b} \subseteq [k], |R_{d-1;n_d;b}| = b_{d-1}.$$

For $n_{d-1} \notin R_{d-1;n_d;b}$ and $n_d \notin R_{d;b}$ we have $R_{d-2;n_{d-1},n_d;b} \subseteq [k]$, $|R_{d-2;n_{d-1},n_d;b}| = b_{d-2}$, etc.

Finally for $n_2 \notin R_{2;n_3,\dots,n_d;b}$, $n_3 \notin R_{3;n_4,\dots,n_d;b}$, $\dots$, and $n_d \notin R_{d;b}$ we have $R_{1;n_2,\dots,n_d;b} \subseteq [k]$, $|R_{1;n_2,\dots,n_d;b}| = b_1$.

Moreover the above defined sets have the following property: if some $a = (a_1, \dots, a_d) \in A$ and some $b \in B$ satisfy $a_d \notin R_{d;b}$, $a_{d-1} \notin R_{d-1;a_d;b}$, $\dots$, $a_2 \notin R_{2;a_3,\dots,a_d;b}$, then $a_1 \in R_{1;a_2,a_3,\dots,a_d;b}$.

Then

$$|A| \leq (k+1)^d - |B_k^+|. \quad (11.4)$$

Remark 11.0.27. *As in Remark 11.0.25, Lemma 11.0.26 is also sharp. Taking a previous example when $d = 2$, $B = \{(1, 1), (3, 0)\} \in \mathcal{C}_4$ colored red as the following:*

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	(4, 1)
(0, 2)	(1, 2)	(2, 2)	(3, 2)	(4, 2)
(0, 3)	(1, 3)	(2, 3)	(3, 3)	(4, 3)
(0, 4)	(1, 4)	(2, 4)	(3, 4)	(4, 4)

, then it is possible to take A to be the blue colored subset

in $\mathcal{C}_4$ shown in the following diagram (where we have already colored all pairs in B_4^+

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	(4, 1)
(0, 2)	(1, 2)	(2, 2)	(3, 2)	(4, 2)
(0, 3)	(1, 3)	(2, 3)	(3, 3)	(4, 3)
(0, 4)	(1, 4)	(2, 4)	(3, 4)	(4, 4)

to be red):

a larger A .

Proof that Lemma 11.0.26 implies Lemma 11.0.24. Assuming Lemma 11.0.26 holds.

We prove Lemma 11.0.24.

Take A in Lemma 11.0.26 to be the A we have in Lemma 11.0.24. We would like to enlarge B and apply Lemma 11.0.26. To achieve this we exploit the constraint $A \subseteq \mathcal{S}_k \cup \{(0, 0, \dots, 0)\}$.

We add all the elements $b = (b_1, \dots, b_d) \in \mathcal{C}_k$ s.t. $|b| > k$ into the set B in Lemma 11.0.24 and form a new set $\tilde{B}$. We next check that we can apply Lemma 11.0.26 to A and $\tilde{B}$. Since $A \subseteq \mathcal{S}_k \cup \{(0, 0, \dots, 0)\} \subseteq \mathcal{C}_k$, it suffices to check the assumption of Lemma 11.0.26 for any newly added $b \in \tilde{B}$. Such b satisfies $|b| > k > |a|$ for any

$a \in A$. Hence just like we have noticed in Remark 11.0.25, one of the inequalities $0 \leq a_i < b_i$ has to hold. Therefore taking all $R_{i;*,b} = [b_i - 1]$ suffices.

Apply Lemma 11.0.26 to A and $\tilde{B}$, we deduce

$$|A| \leq (d+1)^k - |\tilde{B}_k^+|. \quad (11.5)$$

We now determine the elements of $\tilde{B}_k^+$ by definition of $\cdot_k^+$. For any $b' \in \mathcal{C}_k \setminus (\mathcal{S}_k \cup (0, 0, \dots, 0))$, $|b'| > k$. Hence $b' \in \tilde{B}$. By $b' \preceq b'$ we have $b' \in \tilde{B}_k^+$. For any $b' \in \mathcal{S}_k \cup (0, 0, \dots, 0)$, $|b'| \leq k$. Hence if $b'' \prec b'$ holds for some $b'' \in \tilde{B}$ then $|b''| \leq k$ hence $b'' \in B$. Therefore such $b' \in \tilde{B}_k^+$ if and only if $b' \in B_k^+$. As a conclusion, $\tilde{B}_k^+ = B_k^+ \sqcup (\mathcal{C}_k \setminus (\mathcal{S}_k \cup (0, 0, \dots, 0)))$.

Hence

$$\begin{aligned} |A| &\leq (k+1)^d - |\tilde{B}_k^+| \\ &= (k+1)^d - |\mathcal{C}_k \setminus (\mathcal{S}_k \cup (0, 0, \dots, 0))| - |B_k^+| \\ &= (k+1)^d - |(k+1)^d - G_k - 1| - |B_k^+| \\ &= G_k + 1 - |B_k^+| \end{aligned} \quad (11.6)$$

and (11.3) was proved. $\square$

Proof of Lemma 11.0.26. We prove (11.4) by induction on d and then on $|B|$. For convenience, we denote the “deficient function” $\Delta(d, B) = \min |\mathcal{C}_k \setminus A| = \min((k+1)^d - |A|)$ where the minimum is taken over all A s.t. (A, B) satisfies the assumption

of Lemma 11.0.26. Then (11.4) is equivalent to the statement

$$\Delta(d, B) \geq |B_k^{\tilde{+}}|. \quad (11.7)$$

We verify the induction basis. When $d = 1$ the assumption requires all $a \in A$ in a set of cardinality b for any $b \in B$. Hence $|A| \leq \min_{b \in B} b = |[b - 1]| = |\mathcal{C}_k \setminus B_k^{\tilde{+}}| = k + 1 - |B_k^{\tilde{+}}|$ and the conclusion holds. When $|B| = 1$, we count the number of possibilities of $a \in \mathcal{C}_k \setminus A$. By assumption, the last component of such an a can take $(k + 1 - b_d)$ different possible values, and after fixing it, the second last component can take $(k + 1 - b_{d-1})$ different possible values, ..., finally the first component can take $(k + 1 - b_1)$ different possible values to ensure $a \notin A$. Hence $|A| \leq |\mathcal{C}_k| - \prod_{i=1}^d (k + 1 - b_i) = (k + 1)^d - |B_k^{\tilde{+}}|$. The conclusion also holds in this case.

From now on we assume that (11.4) and hence (11.7) hold for all dimension $d' < d$ and $|B'| < |B|$ in the dimension d case. We can assume $d > 1$ and $|B| > 1$.

For $j \in [k]$, call the set $\mathcal{U}_j = \{(\cdot, \cdot, \dots, \cdot, j)\} \subseteq \mathcal{C}_k = \mathcal{C}_k^d$ to be the j -th slice of $\mathcal{C}_k = \mathcal{C}_k^d$. It is isomorphic to $\mathcal{C}_k^{d-1}$. For any subset $B_1 \in B$, define the projection $\mathbf{P}B_1 = \{(b_1, \dots, b_{d-1}) : \exists b_d \text{ s.t. } (b_1, \dots, b_d) \in B_1\} \subseteq \mathcal{C}_k^{d-1}$. We have $|\mathbf{P}B_1| \leq |B_1|$.

We choose an element $b_0 = (b_{0,1}, \dots, b_{0,d}) \in B$ such that $b_{0,d}$ is the largest possible. Let $B' = B \setminus \{b_0\}$. We deal with the problem “slicewise”. For each $j \in [k]$, define $B_j \subseteq B$ to be the subset $\{b \in B : j \notin R_{d,b}\}$. We similarly define B'_j from B' . By definition, $A \cap \mathcal{U}_j \subseteq \mathcal{U}_j \simeq \mathcal{C}_k^{d-1}$ and $\mathbf{P}B_j$ have to satisfy the assumption of Lemma 11.0.26 in dimension $d - 1$ for each $j \in [k]$. Moreover, if some $A \subseteq \mathcal{C}_k^d$ such that $A \cap \mathcal{U}_j$ and $\mathbf{P}B_j$ satisfy the assumption of Lemma 11.0.26 in dimension $d - 1$, then

A satisfy the assumption of Lemma 11.0.26.

We learn from the last paragraph by the induction hypothesis for $d - 1$ that

$$\Delta(d - 1, A \cap \mathcal{U}_j) \geq |(\mathbf{P}B_j)_k^\tilde+|. \quad (11.8)$$

But from the analysis we have done and as in Remark 11.0.25, it is possible to construct A slicewise for each (11.8) (on the underlying slice) to actually take equality (no matter how the $R_{d;b}$'s are given). Such an A would have $\Delta(d, A) = \sum_{j=0}^d |(\mathbf{P}B_j)_k^\tilde+|$. Replace B by B' and run the entire set of reasoning, we get a set A' such that A' and B' satisfy the assumption of Lemma 11.0.26 and

$$\Delta(d, A') = \sum_{j=0}^d |(\mathbf{P}B'_j)_k^\tilde+|. \quad (11.9)$$

Note that we have the induction hypothesis for B' , we deduce from (11.9) that

$$\sum_{j=0}^d |(\mathbf{P}B'_j)_k^\tilde+| \geq |(B')_k^\tilde+|. \quad (11.10)$$

Note that $j \in R_{d;b_0}$ except for $(d+1-b_{0,d})$ different values of j . Hence for exactly $b_{0,d}$ different values of j we have $B_j = B'_j$. For all other j we have $B_j = B'_j \cup \{b_0\}$. When $B_j = B'_j$ we surely have $(\mathbf{P}B_j)_k^\tilde+ = (\mathbf{P}B'_j)_k^\tilde+$. When $B_j = B'_j \cup \{b_0\}$ we have

$$(\mathbf{P}B_j)_k^\tilde+ = (\mathbf{P}B'_j)_k^\tilde+ \cup (\mathbf{P}\{b_0\})_k^\tilde+ = (\mathbf{P}B'_j)_k^\tilde+ \sqcup ((\mathbf{P}\{b_0\})_k^\tilde+ \setminus (\mathbf{P}B'_j)_k^\tilde+). \quad (11.11)$$

Hence in such cases

$$|(\mathbf{P}B_j)_k^\pm| = |(\mathbf{P}B'_j)_k^\pm| + |((\mathbf{P}\{b_0\})_k^\pm \setminus (\mathbf{P}B'_j)_k^\pm)| \geq |(\mathbf{P}B'_j)_k^\pm| + |((\mathbf{P}\{b_0\})_k^\pm \setminus (\mathbf{P}B'_j)_k^\pm)|. \quad (11.12)$$

Sum (11.12) over all j such that $B_j = B'_j \cup \{b_0\}$, and invoke (11.10) and (11.8).

Using the fact

$$\Delta(d, A) = \sum_{j=0}^d \Delta(d-1, A \cap \mathcal{U}_j), \quad (11.13)$$

we deduce

$$\Delta(d, A) \geq |(B')_k^\pm| + (d+1-b_{0,d})|((\mathbf{P}\{b_0\})_k^\pm \setminus (\mathbf{P}B')_k^\pm)| \quad (11.14)$$

Since $b_{0,d}$ is the largest possible, we feel comfortable comparing $(B')_k^\pm$ and $B_k^\pm$. When we look slicewisely on each $\mathcal{U}_j$, we find that on the first $b_{0,d}$ slices the two set coincide. While on the last $(d+1-b_{0,d})$ slices the second set is exactly equal to the union of the first set and $((\mathbf{P}\{b_0\})_k^\pm \setminus (\mathbf{P}B')_k^\pm)$. Hence the right hand side of (11.14) is exactly $|B_k^\pm|$. We have proved (11.7) and hence (11.4) for d and B , thus closing the induction. $\square$

We will naturally have such an A as in Lemma 11.0.24 arising from the verification of Conjecture 1.3 in [BDG16a]. We have done a great job understanding its size by the strong Lemma 11.0.24. The expression contains B_k^+ . We study it in the next lemma and prove a key inequality.

Lemma 11.0.28. *Assuming $l < k$ are positive integers. For any nonempty $B \subseteq \mathcal{S}_l$,*

as long as $B_l^+ \neq S_l$, we have

$$\frac{|B_l^+|}{|B_k^+|} < \frac{|S_l|}{|S_k|} = \frac{G_l}{G_k}. \quad (11.15)$$

Proof. We do some preliminary reductions. First we claim that to prove (11.15) it suffices to do the case $k = l + 1$, i.e. proving

$$\frac{|B_l^+|}{|B_{l+1}^+|} < \frac{|S_l|}{|S_{l+1}|} = \frac{G_l}{G_{l+1}} \quad (11.16)$$

when $B_l^+ \neq S_l$.

In fact once we have (11.16) we always have $\frac{|B_l^+|}{|B_{l+1}^+|} \leq \frac{|S_l|}{|S_{l+1}|}$ and the equality holds only when $B_l^+ = S_l$. Similarly we have $\frac{|B_{l+1}^+|}{|B_{l+2}^+|} \leq \frac{|S_{l+1}|}{|S_{l+2}|}, \dots, \frac{|B_{k-1}^+|}{|B_k^+|} \leq \frac{|S_{k-1}|}{|S_k|}$. Taking the product of everything above we have (11.15).

In the rest of the proof we prove (11.16). In other words, the global maximum of $\frac{|B_l^+|}{|B_{l+1}^+|}$ over all possible $\emptyset \subsetneq B \subseteq S_l$ is achieved if and only if $B = S_l$.

It seems hard to control $|B_l^+|$ or $|B_{l+1}^+|$. But we notice that the difference of the two is a simpler object (a “layer” in $|B_{l+1}^+|$). Moreover, we find that B_l^+ can be decomposed into l layers of such form. This inspires us to do the following decomposition: For any positive integer m , define $B_m^\uparrow$ to be $B_k^+ \cap \{b : |b| = m\}$ for any sufficiently large k . We can alternatively use the straightforward definition

$$B_m^\uparrow = \{b \in \mathbb{Z}_{\geq 0}^d : |b| = m, \exists b' \in B \text{ s.t. } b' \preceq b\}. \quad (11.17)$$

Hence $B_l^+ = \sqcup_{m=1}^l B_m^\uparrow$, $B_{l+1}^+ = \sqcup_{m=1}^{l+1} B_m^\uparrow$. $|B_l^+| = \sum_{m=1}^l |B_m^\uparrow|$, $B_{l+1}^+ = \sum_{m=1}^{l+1} |B_m^\uparrow|$.

Next we explore the relationship between $B_m^\uparrow$'s.

For $m \geq 1$ define $\mathcal{V}_m = \mathcal{V}_m^{d-1} = \{b \in \mathbb{Z}_{\geq 0}^d, |b| = m\}$. We use the superscript $d-1$ to emphasize that $\mathcal{V}_m^{d-1}$ is a $(d-1)$ -dimensional object. As before when there is no ambiguity about the dimension we suppress this superscript.

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	
(0, 2)	(1, 2)	(2, 2)		
(0, 3)	(1, 3)			
(0, 4)				

For example, $\mathcal{V}_4^1$ is the following set in green:

For $m > 1$ and any set $T \subseteq \mathcal{V}_m$ we define its predecessor $T^- = \{b = (b_1, \dots, b_d) \in \mathcal{V}_{m-1} : (b_1, \dots, b_{i-1}, b_i + 1, b_{i+1}, \dots, b_d) \in T, \forall 1 \leq i \leq d\}$.

For example if $d = 2$ and $T \subseteq \mathcal{V}_4^1$ is the following set in green, then $T^- \in \mathcal{V}_3^1$ is

(0, 0)	(1, 0)	(2, 0)	(3, 0)	(4, 0)
(0, 1)	(1, 1)	(2, 1)	(3, 1)	
(0, 2)	(1, 2)	(2, 2)		
(0, 3)	(1, 3)			
(0, 4)				

the set in yellow:

We claim: for any $m > 1$,

$$B_{m-1}^\uparrow \subseteq (B_m^\uparrow)^-. \quad (11.18)$$

In fact assuming $b = (b_1, \dots, b_d) \in B_{m-1}^\uparrow$. Then there exists $B \ni b' \preceq b$. Hence $b' \preceq b \preceq (b_1, \dots, b_{i-1}, b_i + 1, b_{i+1}, \dots, b_d)$ for all $1 \leq i \leq d$. Thus $(b_1, \dots, b_{i-1}, b_i +$

$1, b_{i+1}, \dots, b_d) \in B_m^\uparrow$ for all $1 \leq i \leq d$. By definition (11.18) holds.

We will prove the following inequality:

$$|T^-| \leq \frac{|\mathcal{V}_{m-1}^{d-1}|}{|\mathcal{V}_m^{d-1}|} |T|, \forall m > 1, T \subseteq \mathcal{V}_m^{d-1} \quad (11.19)$$

and prove that equality holds only when $T = \mathcal{V}_m^{d-1}$ or $T = \emptyset$.

We claim that (11.19) along with its equality condition together imply (11.16). Assuming we already have (11.19) and know that equality holds only when $T = \mathcal{V}_m$ for all $m > 1$. We now prove (11.16). In fact by (11.19) and (11.18) we inductively deduce that

$$|B_l^\uparrow| \leq \frac{|\mathcal{V}_l|}{|\mathcal{V}_{l+1}|} |B_{l+1}^\uparrow|, \quad (11.20)$$

$$|B_{l-1}^\uparrow| \leq \frac{|\mathcal{V}_{l-1}|}{|\mathcal{V}_l|} |B_l^\uparrow| \leq \frac{|\mathcal{V}_{l-1}|}{|\mathcal{V}_l|} \cdot \frac{|\mathcal{V}_l|}{|\mathcal{V}_{l+1}|} |B_{l+1}^\uparrow| = \frac{|\mathcal{V}_{l-1}|}{|\mathcal{V}_{l+1}|} |B_{l+1}^\uparrow|, \quad (11.21)$$

$$\dots, \quad (11.22)$$

$$|B_1^\uparrow| \leq \frac{|\mathcal{V}_1|}{|\mathcal{V}_{l+1}|} |B_{l+1}^\uparrow|. \quad (11.23)$$

Summing over all the inequalities above we deduce

$$|B_l^+| = \sum_{m=1}^l |B_m^\uparrow| \leq \frac{\sum_{m=1}^l |\mathcal{V}_m|}{|\mathcal{V}_{l+1}|} |B_{l+1}^\uparrow| = \frac{|\mathcal{S}_l|}{|\mathcal{V}_{l+1}|} |B_{l+1}^\uparrow| = \frac{|\mathcal{S}_l|}{|\mathcal{S}_{l+1}| - |\mathcal{S}_l|} (|B_{l+1}^+| - |B_l^+|) \quad (11.24)$$

Elementary manipulations show that this is equivalent to (11.16) when we replace “ $<$ ” by “ $\leq$ ” there. However when equality holds, all the above equalities involving (11.19) and (11.18) have to hold. Note that $B \neq \emptyset$ hence $B_{l+1}^\uparrow \neq \emptyset$. By the equality condition of (11.19), we inductively see that each $B_m^\uparrow$ has to be the whole $\mathcal{V}_m$ for

$m = l, l-1, \dots, 1$. Hence $B_l^+ = \mathcal{S}_l$. A contradiction with the assumption we have for (11.16). Hence (11.16) holds (without equality).

Therefore it suffices to prove (11.19) and prove that its equality holds only when $T = \emptyset$ or $\mathcal{V}_m^{d-1}$. We do this in the rest of the proof.

We perform an induction on dimension d . When $d = 1$, each $\mathcal{V}_m^0$ is just one point. In this case $T^- = \emptyset$ when $T = \emptyset$, and $T^- = \mathcal{V}_{m-1}^0$ when $T = \mathcal{V}_m^0$. (11.19) holds in either case.

Assuming we already proved (11.19) for all dimensions $< d$ with full knowledge on the possible situations when we have equality (T being a trivial subset). We now handle the dimension $d(\geq 2)$ case.

We further decompose $\mathcal{V}_m^{d-1}$ into $(m+1)$ pieces $\mathcal{W}_{j,m}^{d-2}, 0 \leq j \leq m$ as the following (again its superscript $d-2$ is so chosen to emphasize the dimension):

$$\begin{aligned} \mathcal{W}_{j,m}^{d-2} &= \{b = (b_1, \dots, b_d) \in \mathbb{Z}_{\geq 0}^d : |b| = m, b_d = m - j\} \\ &= \{b = (b_1, \dots, b_{d-1}, m - j) \in \mathbb{Z}_{\geq 0}^d : |(b_1, \dots, b_{d-1})| = j\} \\ &\simeq \mathcal{V}_j^{d-2}. \end{aligned} \tag{11.25}$$

We denote σ_d to be the isomorphism map shown in the last line of (11.25). Hence $\mathcal{W}_{j,m}^{d-2} \xrightarrow{\sigma_d} \mathcal{V}_j^{d-2}$. It is realized simply by removing the last component.

We assume accordingly that $T = \sqcup_{j=0}^m T_j$ where $T_j = T \cap \mathcal{W}_{j,m}^{d-2}$ and $T^- = \sqcup_{j=0}^{m-1} (T^-)_j$ where $(T^-)_j = T^- \cap \mathcal{W}_{j,m-1}^{d-2}$. In this way we can study each $(T^-)_j$ separately.

For a fixed $0 \leq j \leq m-1$ and any $(d-1)$ -dimensional vector $b^* = (b_1, \dots, b_{d-1}) \in$

$\mathcal{V}_j^{d-2}$, by definition of T^- we see that $(b^*, m-1-j) \in T^- \Leftrightarrow (b^*, m-1-j) \in (T^-)_j$

holds if and only if the following two conditions both hold:

- (i) $(b^*, m-j) \in T$, i.e. $(b^*, m-j) \in T_j$. Alternatively we can say $b^* \in \sigma_d(T_j)$;
- (ii) $b^* \in (\sigma_d(T_{j+1}))^-$. Note that here $(\cdot)^-$ is a map from $\mathcal{V}_{j+1}^{d-2}$ to $\mathcal{V}_j^{d-2}$.

By (i) we have

$$|(T^-)_j| \leq |T_j| \quad (11.26)$$

where equality holds only when $\sigma_d((T^-)_j) = \sigma_d(T_j)$.

By (ii) and the induction hypothesis we have

$$|(T^-)_j| \leq \frac{|\mathcal{V}_j^{d-2}|}{|\mathcal{V}_{j+1}^{d-2}|} |T_{j+1}| \quad (11.27)$$

where equality only holds when both (a) $\sigma_d((T^-)_j) = (\sigma_d(T_{j+1}))^-$ and (b) $T_{j+1} = \emptyset$ or $T_{j+1} = \mathcal{W}_{j+1,m}^{d-2}$.

We now have the two key inequalities (11.26) and (11.27). We need a bit of numerical preparation before proving (11.19) with the above two inequalities.

For $d \geq 1$ and $q \geq 0$, define $\Lambda_{q,d-1} = |\mathcal{V}_q^{d-1}|$. These are $(d-1)$ -dimensional generalizations of triangular numbers. As before for $d > 1$ we know $\mathcal{V}_q^{d-1} = \bigsqcup_{j=0}^m \mathcal{W}_{j,q}^{d-2}$ and each $\mathcal{W}_{j,q}^{d-2} \simeq \mathcal{V}_j^{d-2}$. Hence

$$\Lambda_{q,d-1} = \sum_{j=0}^q \Lambda_{j,d-2}, \forall d > 1, q \geq 0. \quad (11.28)$$

We prove that for any $d \geq 1, q \geq 0$,

$$\frac{\sum_{j=0}^q \Lambda_{j,d-1}}{\Lambda_{q+1,d-1}} < \frac{\sum_{j=0}^{q+1} \Lambda_{j,d-1}}{\Lambda_{q+2,d-1}} \quad (11.29)$$

or equivalently

$$\frac{\sum_{j=0}^q \Lambda_{j,d-1}}{\sum_{j=0}^{q+1} \Lambda_{j,d-1}} < \frac{\sum_{j=0}^{q+1} \Lambda_{j,d-1}}{\sum_{j=0}^{q+2} \Lambda_{j,d-1}}. \quad (11.30)$$

We prove (11.30) by induction on d . For $d = 1$, (11.30) becomes $\frac{q+1}{q+2} < \frac{q+2}{q+3}$ which is trivially true. We assume (11.30) is true for $d < d_0$ and prove it for $d = d_0 > 1$.

(11.30) is equivalent to

$$\frac{\sum_{j=0}^q \Lambda_{j,d_0-1}}{\sum_{j=0}^{q+1} \Lambda_{j,d_0-1}} < \frac{\Lambda_{q+1,d_0-1}}{\Lambda_{q+2,d_0-1}} \quad (11.31)$$

which was further implied by

$$\frac{0}{\Lambda_{0,d_0-1}} < \frac{\Lambda_{j,d_0-1}}{\Lambda_{j+1,d_0-1}} < \frac{\Lambda_{j+1,d_0-1}}{\Lambda_{j+2,d_0-1}}, \forall j \geq 0. \quad (11.32)$$

But by (11.28), (11.32) is equivalent to the case $d = d_0 - 1$ which was already proved.

This closes the induction and thus (11.30) (and the equivalent (11.29)) holds.

By (11.29), we have for each $0 \leq j \leq m-1$, $\sum_{q=0}^{m-1} \Lambda_{q,d-2} - \frac{\sum_{p=0}^{j-1} \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \Lambda_{m,d-2} > 0$.

With this in mind, by (11.26) and (11.27) we have

$$\begin{aligned}
|T^-| &= \sum_{j=0}^{m-1} |(T^-)_j| \\
&= \sum_{j=0}^{m-1} \left(\frac{\sum_{q=0}^{m-1} \Lambda_{q,d-2} - \frac{\sum_{p=0}^{j-1} \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \Lambda_{m,d-2}}{\sum_{q=0}^m \Lambda_{q,d-2}} |(T^-)_j| + \frac{\sum_{p=0}^j \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \frac{\Lambda_{m,d-2}}{\sum_{q=0}^m \Lambda_{q,d-2}} |(T^-)_j| \right) \\
&\leq \sum_{j=0}^{m-1} \frac{\sum_{q=0}^{m-1} \Lambda_{q,d-2} - \frac{\sum_{p=0}^{j-1} \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \Lambda_{m,d-2}}{\sum_{q=0}^m \Lambda_{q,d-2}} |T_j| + \sum_{j=0}^{m-1} \frac{\sum_{p=0}^j \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \cdot \frac{\Lambda_{j,d-2}}{\Lambda_{j+1,d-2}} |T_{j+1}| \\
&= \sum_{j=0}^{m-1} \frac{\sum_{q=0}^{m-1} \Lambda_{q,d-2} - \frac{\sum_{p=0}^{j-1} \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \Lambda_{m,d-2}}{\sum_{q=0}^m \Lambda_{q,d-2}} |T_j| + \sum_{j=1}^m \frac{\sum_{p=0}^{j-1} \Lambda_{p,d-2}}{\Lambda_{j,d-2}} \frac{\Lambda_{m,d-2}}{\sum_{q=0}^m \Lambda_{q,d-2}} |T_j| \\
&= \frac{\sum_{q=0}^{m-1} \Lambda_{q,d-2}}{\sum_{q=0}^m \Lambda_{q,d-2}} \sum_{j=0}^m |T_j| \\
&= \frac{\Lambda_{m-1,d-1}}{\Lambda_{m,d-1}} |T|. \tag{11.33}
\end{aligned}$$

Hence (11.19) holds. Moreover, the equality there implies the equality in (11.33), which implies that the equalities of (11.26) and (11.27) both hold for all $0 \leq j \leq m-1$. We thus either have inductively $T_m = T_{m-1} = \dots = T_0 = \emptyset$, or $T_m = \mathcal{W}_{m,m}^{d-2}$, $T_{m-1} = \mathcal{W}_{m-1,m}^{d-2}, \dots, T_0 = \mathcal{W}_{0,m}^{d-2}$. The first case would imply $T = \emptyset$ and the second would imply $T = \mathcal{V}_m^{d-1}$. These together verify the desired equality condition for (11.19).

By the arguments in this proof, (11.16) and (11.15) hold. $\square$

Theorem 11.0.29. *Let $l < k$ be positive integers. For any vector $v = (v_{(i_1, \dots, i_d)})_{i_1, \dots, i_d \geq 0, 1 \leq i_1 + \dots + i_d \leq k} \in \mathbb{R}^{G_k}$, define a d -variate polynomial $f_v = \sum_{(i_1, \dots, i_d)} v_{(i_1, \dots, i_d)} x_1^{i_1} x_2^{i_2} \dots x_d^{i_d}$.*

For an arbitrary nonzero subspace $V \subseteq \mathbb{R}^{G_k}$ spanned by $\{v_h\}_{1 \leq h \leq H = \dim V}$, define $r(V)$ to be the rank of the following matrix over $\mathbb{R}(x_1, \dots, x_d)$ (that is obviously

independent of the choice of the basis $\{v_h\}$):

$$M_{d,k,l}(V) = \begin{pmatrix} \partial_{x_1} f_{v_1} & \partial_{x_1} f_{v_2} & \cdots & \partial_{x_1} f_{v_H} \\ \partial_{x_2} f_{v_1} & \partial_{x_2} f_{v_2} & \cdots & \partial_{x_2} f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_d} f_{v_1} & \partial_{x_d} f_{v_2} & \cdots & \partial_{x_d} f_{v_H} \\ \partial_{x_1}^2 f_{v_1} & \partial_{x_1}^2 f_{v_2} & \cdots & \partial_{x_1}^2 f_{v_H} \\ \partial_{x_1} \partial_{x_2} f_{v_1} & \partial_{x_1} \partial_{x_2} f_{v_2} & \cdots & \partial_{x_1} \partial_{x_2} f_{v_H} \\ \partial_{x_1} \partial_{x_3} f_{v_1} & \partial_{x_1} \partial_{x_3} f_{v_2} & \cdots & \partial_{x_1} \partial_{x_3} f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_d}^2 f_{v_1} & \partial_{x_d}^2 f_{v_2} & \cdots & \partial_{x_d}^2 f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_1}^l f_{v_1} & \partial_{x_1}^l f_{v_2} & \cdots & \partial_{x_1}^l f_{v_H} \\ \partial_{x_1}^{l-1} \partial_{x_2} f_{v_1} & \partial_{x_1}^{l-1} \partial_{x_2} f_{v_2} & \cdots & \partial_{x_1}^{l-1} \partial_{x_2} f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_d}^l f_{v_1} & \partial_{x_d}^l f_{v_2} & \cdots & \partial_{x_d}^l f_{v_H} \end{pmatrix}. \quad (11.34)$$

Then when $H = \dim V < G_k$, we always have

$$\frac{r(V)}{H} > \frac{G_l}{G_k}. \quad (11.35)$$

Proof. We fix a lexicographical order on all d -variate monomials. We say that the monomial $x_1^{i_1} \cdots x_d^{i_d}$ is a lower order term than $x_1^{i'_1} \cdots x_d^{i'_d}$ if and only if there is some $i_q < i'_q$ and $i_j = i'_j$ for all $1 \leq j < q$. The lexicographical order is a total order

on the set of all monomials. Moreover it is well-known that if nonzero monomials g_1 and g_2 are of lower or equal order than nonzero monomials g'_1 and g'_2 , respectively, then g_1g_2 is of lower or equal order than $g'_1g'_2$. The equality holds if and only if g_i is a scalar multiple of g'_i for both $i = 1, 2$.

We first observe that it is possible to choose a basis $\{v_h\}_{1 \leq h \leq H}$ of V such that every f_{v_h} has a different highest order term from every other $f_{v_{h'}}$ ($h \neq h'$). This can be done by choosing v_1 s.t. f_{v_1} has the highest possible highest order term, then choosing v_2 s.t. f_{v_2} has the highest order term being (a) different from the highest order term of f_{v_1} and (b) of highest possible order, and then choosing $v_3, v_4, \dots$ in a similar way. As a remark, the set of highest order terms of $\{f_{v_h}\}$ satisfying the said condition is unique, but we do not need this fact. We always take $\{v_h\}$ to be such a basis in the following discussion.

Note that all f_{v_h} do not have constant terms, hence their highest order terms do not contain constant either. For a nonzero polynomial f denote $\widetilde{f}$ to be its highest order term. Replace each f_{v_h} by its highest order term $\widetilde{f_{v_h}}$ in the expression of

$M_{d,k,l}(V)$, we obtain a matrix

$$N_{d,k,l}(V) = \begin{pmatrix} \partial_{x_1} \widetilde{f_{v_1}} & \partial_{x_1} \widetilde{f_{v_2}} & \cdots & \partial_{x_1} \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_d} \widetilde{f_{v_1}} & \partial_{x_d} \widetilde{f_{v_2}} & \cdots & \partial_{x_d} \widetilde{f_{v_H}} \\ \partial_{x_1}^2 \widetilde{f_{v_1}} & \partial_{x_1}^2 \widetilde{f_{v_2}} & \cdots & \partial_{x_1}^2 \widetilde{f_{v_H}} \\ \partial_{x_1} \partial_{x_2} \widetilde{f_{v_1}} & \partial_{x_1} \partial_{x_2} \widetilde{f_{v_2}} & \cdots & \partial_{x_1} \partial_{x_2} \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_d}^2 \widetilde{f_{v_1}} & \partial_{x_d}^2 \widetilde{f_{v_2}} & \cdots & \partial_{x_d}^2 \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_1}^l \widetilde{f_{v_1}} & \partial_{x_1}^l \widetilde{f_{v_2}} & \cdots & \partial_{x_1}^l \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ \partial_{x_d}^l \widetilde{f_{v_1}} & \partial_{x_d}^l \widetilde{f_{v_2}} & \cdots & \partial_{x_d}^l \widetilde{f_{v_H}} \end{pmatrix} \quad (11.36)$$

over $\mathbb{R}(x_1, \dots, x_d)$. Let $r_1(V)$ be its rank. Then we claim $r_1(V) \leq r(V)$.

In fact, let

$$\widetilde{M}_{d,k,l}(V) = \begin{pmatrix} x_1 \partial_{x_1} f_{v_1} & x_1 \partial_{x_1} f_{v_2} & \cdots & x_1 \partial_{x_1} f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ x_d \partial_{x_d} f_{v_1} & x_d \partial_{x_d} f_{v_2} & \cdots & x_d \partial_{x_d} f_{v_H} \\ x_1^2 \partial_{x_1}^2 f_{v_1} & x_1^2 \partial_{x_1}^2 f_{v_2} & \cdots & x_1^2 \partial_{x_1}^2 f_{v_H} \\ x_1 x_2 \partial_{x_1} \partial_{x_2} f_{v_1} & x_1 x_2 \partial_{x_1} \partial_{x_2} f_{v_2} & \cdots & x_1 x_2 \partial_{x_1} \partial_{x_2} f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ x_d^2 \partial_{x_d}^2 f_{v_1} & x_d^2 \partial_{x_d}^2 f_{v_2} & \cdots & x_d^2 \partial_{x_d}^2 f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ x_1^l \partial_{x_1}^l f_{v_1} & x_1^l \partial_{x_1}^l f_{v_2} & \cdots & x_1^l \partial_{x_1}^l f_{v_H} \\ \cdots & \cdots & \cdots & \cdots \\ x_d^l \partial_{x_d}^l f_{v_1} & x_d^l \partial_{x_d}^l f_{v_2} & \cdots & x_d^l \partial_{x_d}^l f_{v_H} \end{pmatrix} \quad (11.37)$$

and

$$\widetilde{N}_{d,k,l}(V) = \begin{pmatrix} x_1 \partial_{x_1} \widetilde{f_{v_1}} & x_1 \partial_{x_1} \widetilde{f_{v_2}} & \cdots & x_1 \partial_{x_1} \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ x_d \partial_{x_d} \widetilde{f_{v_1}} & x_d \partial_{x_d} \widetilde{f_{v_2}} & \cdots & x_d \partial_{x_d} \widetilde{f_{v_H}} \\ x_1^2 \partial_{x_1}^2 \widetilde{f_{v_1}} & x_1^2 \partial_{x_1}^2 \widetilde{f_{v_2}} & \cdots & x_1^2 \partial_{x_1}^2 \widetilde{f_{v_H}} \\ x_1 x_2 \partial_{x_1} \partial_{x_2} \widetilde{f_{v_1}} & x_1 x_2 \partial_{x_1} \partial_{x_2} \widetilde{f_{v_2}} & \cdots & x_1 x_2 \partial_{x_1} \partial_{x_2} \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ x_d^2 \partial_{x_d}^2 \widetilde{f_{v_1}} & x_d^2 \partial_{x_d}^2 \widetilde{f_{v_2}} & \cdots & x_d^2 \partial_{x_d}^2 \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ x_1^l \partial_{x_1}^l \widetilde{f_{v_1}} & x_1^l \partial_{x_1}^l \widetilde{f_{v_2}} & \cdots & x_1^l \partial_{x_1}^l \widetilde{f_{v_H}} \\ \cdots & \cdots & \cdots & \cdots \\ x_d^l \partial_{x_d}^l \widetilde{f_{v_1}} & x_d^l \partial_{x_d}^l \widetilde{f_{v_2}} & \cdots & x_d^l \partial_{x_d}^l \widetilde{f_{v_H}} \end{pmatrix}. \quad (11.38)$$

It is immediate that $r(V)$ is the rank of $\widetilde{M}_{d,k,l}(V)$ and $r_1(V)$ is the rank of $\widetilde{N}_{d,k,l}(V)$. Note that for any d -variate monomial g in $x_1, \dots, x_d$ and any differential operator $\partial_{x_1}^{i_1} \cdots \partial_{x_d}^{i_d}$, $x_1^{i_1} \cdots x_d^{i_d} \partial_{x_1}^{i_1} \cdots \partial_{x_d}^{i_d} g$ is equal to a constant multiple (depending only on $i_1, \dots, i_d$ and the powers of $x_1, \dots, x_d$ in g) of g . We see that if we arbitrarily fix a column (say the h -th column) in $\widetilde{N}_{d,k,l}(V)$, then all entries in this column are the same up to a scalar. Indeed they are all scalar multiples of $\widetilde{f_{v_h}}$. For each such entry (note that it is possible to be zero), its difference from the corresponding entry of $\widetilde{M}_{d,k,l}(V)$ has to be a sum of monomials of strictly lower order than $\widetilde{f_{v_h}}$. Hence if some subdeterminant of $\widetilde{N}_{d,k,l}$ is nonzero, it will be a monomial and has to be the highest order term of the corresponding subdeterminant of $\widetilde{M}_{d,k,l}$. This implies the

corresponding subdeterminant of $\widetilde{M}_{d,k,l}$ is nonzero. Hence the rank $(r_1(V))$ of the former matrix is not more than the rank $(r(V))$ of the latter.

Therefore, it suffices to prove

$$\frac{r_1(V)}{H} > \frac{G_l}{G_k}. \quad (11.39)$$

We remark that since it is possible for all f_{v_h} to be different monomials, we do not lose any information by the reduction in the last paragraph.

Without loss of generality we may assume each $\widetilde{f_{v_h}}$ is a monic monomial $x_1^{a_{1,h}} \cdots x_d^{a_{d,h}}$ for some distinct $(a_{1,h}, \dots, a_{d,h}) \in \mathcal{S}_k$.

We use some elementary transforms that we just did on $N_{d,k,l}(V)$. We multiply the “ $(i_1, \dots, i_d)$ -th” row of $N_{d,k,l}(V)$, i.e. $(\partial_{x_1}^{i_1} \cdots \partial_{x_d}^{i_d} \widetilde{f_{v_1}}, \dots, \partial_{x_1}^{i_1} \cdots \partial_{x_d}^{i_d} \widetilde{f_{v_H}})$, by $x_1^{i_1} \cdots x_d^{i_d}$ and then divide the h -th column by $\widetilde{f_{v_h}}$. This would not change the rank of the

matrix and will result in a scalar matrix $S_{d,k,l}(V)$ of the following form:

$$S_{d,k,l}(V) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,H} \\ \cdots & \cdots & \cdots \\ a_{d,1} & \cdots & a_{d,H} \\ a_{1,1}(a_{1,1} - 1) & \cdots & a_{1,H}(a_{1,H} - 1) \\ a_{1,1}a_{2,1} & \cdots & a_{1,H}a_{2,H} \\ \cdots & \cdots & \cdots \\ a_{d,1}(a_{d,1} - 1) & \cdots & a_{d,H}(a_{d,H} - 1) \\ \cdots & \cdots & \cdots \\ a_{1,1}(a_{1,1} - 1) \cdots (a_{1,1} - l + 1) & \cdots & a_{1,H}(a_{1,H} - 1) \cdots (a_{1,H} - l + 1) \\ \cdots & \cdots & \cdots \\ a_{d,1}(a_{d,1} - 1) \cdots (a_{d,1} - l + 1) & \cdots & a_{d,H}(a_{d,H} - 1) \cdots (a_{d,H} - l + 1) \end{pmatrix}. \quad (11.40)$$

Using a bunch of elementary row transforms we obtain the following cleaner $J_{d,k,l}(V)$ from $S_{d,k,l}(V)$ (this step is not necessary but it is good to have things

cleaner):

$$J_{d,k,l}(V) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,H} \\ \cdots & \cdots & \cdots & \cdots \\ a_{d,1} & a_{d,2} & \cdots & a_{d,H} \\ a_{1,1}^2 & a_{1,2}^2 & \cdots & a_{1,H}^2 \\ a_{1,1}a_{2,1} & a_{1,2}a_{2,2} & \cdots & a_{1,H}a_{2,H} \\ \cdots & \cdots & \cdots & \cdots \\ a_{d,1}^2 & a_{d,2}^2 & \cdots & a_{d,H}^2 \\ \cdots & \cdots & \cdots & \cdots \\ a_{1,1}^l & a_{1,2}^l & \cdots & a_{1,H}^l \\ \cdots & \cdots & \cdots & \cdots \\ a_{d,1}^l & a_{d,2}^l & \cdots & a_{d,H}^l \end{pmatrix}. \quad (11.41)$$

Let $r_2(V)$ be the rank of $J_{d,k,l}(V)$. Then $r_2(V) = r_1(V)$. It suffices to show the following equivalent form of (11.39):

$$\frac{r_2(V)}{H} > \frac{G_l}{G_k}. \quad (11.42)$$

Assuming $Q = G_l - r_2(V)$. We may assume $Q > 0$ since otherwise by the assumption $H < G_k$, (11.42) holds. Then there is a nontrivial subspace $W \subseteq \mathbb{R}^{G_l}$ such that $\dim W = Q$ and that any vector $w \in W$ is orthogonal to all columns of $J_{d,k,l}(V)$. As we did before, we can take a basis $\{w_q\}_{1 \leq q \leq Q}$ such that the highest order term of f_{w_q} are mutually distinct. By our assumption, we have

$$f_{w_q}(a_{1,h}, \dots, a_{d,h}) = 0, \forall 1 \leq q \leq Q, 1 \leq h \leq H. \quad (11.43)$$

Moreover we always have $f_{w_q}(0, 0, \dots, 0) = 0$ since f_{w_q} does not have a constant term.

We are changing our problem towards one with purely combinatorial nature in order to apply Lemma 11.0.24. We introduce a bit more notations. Let $\text{const} \cdot x_1^{b_{1,q}} \dots x_d^{b_{d,q}} \neq 0$ be the highest order term of f_{w_q} . Hence when we fix x_d to be a real number, f_{w_q} is a polynomial in $x_1, \dots, x_{d-1}$ whose highest order term is $\text{const} \cdot x_1^{b_{1,q}} \dots x_{d-1}^{b_{d-1,q}} \neq 0$ except $\leq b_{d,q}$ possible values of x_d . In any non-exceptional case for the fixed x_d , when we further fix x_{d-1} after fixing x_d , f_{w_q} becomes a polynomial in $x_1, \dots, x_{d-2}$ whose highest order term is $\text{const} \cdot x_1^{b_{1,q}} \dots x_{d-2}^{b_{d-2,q}} \neq 0$ except $\leq b_{d-1,q}$ possible values of x_{d-1} (which depend on the x_d we choose earlier). We can continue to do similar reasonings and finally, in any non-exceptional case for the fixed $x_2, \dots, x_d$, f_{w_q} is a polynomial in x_1 whose highest order term is $\text{const} \cdot x_1^{b_{1,q}} \neq 0$. Hence there are no more than $b_{1,q}$ different possible values of x_1 (which can depend jointly on the fixed $x_2, \dots, x_d$) that can make $f_{w_q}(x_1, \dots, x_d) = 0$.

We collect the constraints we got in the last paragraph for all possible q . Take $A = \{(a_{1,h}, \dots, a_{d,h})\}_{1 \leq h \leq H} \cup \{(0, \dots, 0)\}$ and $B = \{(b_{1,q}, \dots, b_{d,q})\}_{1 \leq q \leq Q}$ and we are in a position to apply Lemma 11.0.24. (11.3) in Lemma 11.0.24 implies

$$H + 1 = |A| \leq G_k + 1 - |B_k^+|. \quad (11.44)$$

By (11.44) and (11.15) in Lemma 11.0.28, we have

$$\frac{Q}{G_k - H} = \frac{|B|}{G_k - H} \leq \frac{|B|}{|B_k^+|} \leq \frac{|B_l^+|}{|B_k^+|} \leq \frac{G_l}{G_k}. \quad (11.45)$$

If the last two inequalities of (11.45) are both actually equalities, then by Lemma 11.0.28 we have $B = B_l^+ = \mathcal{S}_l$ (note that $Q > 0$ means $B \neq \emptyset$). This and (11.44) in turn imply $0 = H = \dim V$. A contradiction. Hence we actually have

$$\frac{G_l - r_2(V)}{G_k - H} = \frac{Q}{G_k - H} < \frac{G_l}{G_k} \quad (11.46)$$

which is equivalent to (11.42). By the discussion above we see (11.39) and (11.35) hold. □

Bibliography

- [BB10] Jonathan Bennett and Neal Bez, *Some nonlinear Brascamp-Lieb inequalities and applications to harmonic analysis*, Journal of Functional Analysis **259** (2010), no. 10, 2520–2556.
- [BBFL15] Jonathan Bennett, Neal Bez, Taryn C Flock, and Sanghyuk Lee, *Stability of the Brascamp-Lieb constant and applications*, arXiv preprint arXiv:1508.07502 (2015).
- [BCCT08] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals*, Geometric and Functional Analysis **17** (2008), no. 5, 1343–1415.
- [BCCT10] ———, *Finite bounds for Hölder-Brascamp-Lieb multilinear inequalities*, Math. Res. Lett. **17** (2010), no. 4, 647–666.
- [BCT06] Jonathan Bennett, Anthony Carbery, and Terence Tao, *On the multilinear restriction and Kakeya conjectures*, Acta mathematica **196** (2006), no. 2, 261–302.

- [BD15] Jean Bourgain and Ciprian Demeter, *The proof of the l^2 Decoupling Conjecture*, Annals of mathematics **182** (2015), no. 1, 351–389.
- [BDG16a] Jean Bourgain, Ciprian Demeter, and Shaoming Guo, *Sharp bounds for the cubic Parsell-Vinogradov system in two dimensions*, arXiv preprint arXiv:1608.06346 (2016).
- [BDG16b] Jean Bourgain, Ciprian Demeter, and Larry Guth, *Proof of the main conjecture in Vinogradov’s mean value theorem for degrees higher than three*, Annals of Mathematics **184** (2016), no. 2, 633–682.
- [Ben] Jonathan Bennett, *Transversal multilinear harmonic analysis*, <http://www.uam.es/departamentos/ciencias/matematicas/AFA/Escorial2012/abstracts/sl>
- [Ben13] ———, *Aspects of multilinear harmonic analysis related to transversality*, Harmonic Analysis and Partial Differential Equations **612** (2013), 1.
- [BG11] Jean Bourgain and Larry Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geometric and Functional Analysis **21** (2011), no. 6, 1239–1295.
- [BL76] Herm Jan Brascamp and Elliott H Lieb, *Best constants in Young’s inequality, its converse, and its generalization to more than three functions*, Advances in Mathematics **20** (1976), no. 2, 151–173.
- [Bou91] Jean Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, Geometric and Functional analysis **1** (1991), no. 2, 147–187.

- [Bou13a] ———, *Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces*, Israel Journal of Mathematics **193** (2013), no. 1, 441–458.
- [Bou13b] ———, *On the Schrödinger maximal function in higher dimension*, Proceedings of the Steklov Institute of Mathematics **280** (2013), no. 1, 46–60.
- [CV13] Anthony Carbery and Stefán Ingi Valdimarsson, *The endpoint multilinear Keakeya theorem via the Borsuk–Ulam theorem*, Journal of Functional Analysis **264** (2013), no. 7, 1643–1663.
- [Dvi09] Zeev Dvir, *On the size of Keakeya sets in finite fields*, Journal of the American Mathematical Society **22** (2009), no. 4, 1093–1097.
- [GK15] Larry Guth and Nets Hawk Katz, *On the Erdős distinct distance problem in the plane*, Annals of Mathematics **181** (2015), no. 1, 155–190.
- [Guo] Shaoming Guo, private communication.
- [Gut10] Larry Guth, *The endpoint case of the Bennett–Carbery–Tao multilinear Keakeya conjecture*, Acta mathematica **205** (2010), no. 2, 263–286.
- [Gut14a] ———, *Degree reduction and graininess for Keakeya-type sets in $\mathbb{R}^3$* , arXiv preprint arXiv:1402.0518 (2014).
- [Gut14b] ———, *A restriction estimate using polynomial partitioning*, arXiv preprint arXiv:1407.1916 (2014).

- [Gut15] ———, *A short proof of the multilinear Kakeya inequality*, Mathematical Proceedings of the Cambridge Philosophical Society, vol. 158, Cambridge Univ Press, 2015, pp. 147–153.
- [Joh14] Fritz John, *Extremum problems with inequalities as subsidiary conditions*, *Traces and Emergence of Nonlinear Programming*, Springer, 2014, pp. 197–215.
- [Lie90] Elliott H Lieb, *Gaussian kernels have only Gaussian maximizers*, *Inventiones mathematicae* **102** (1990), no. 1, 179–208.
- [OS82] DM Oberlin and EM Stein, *Mapping properties of the Radon-transform*, *Indiana University Mathematics Journal* **31** (1982), no. 5, 641–650.
- [PPW13] Scott T Parsell, Sean M Prendiville, and Trevor D Wooley, *Near-optimal mean value estimates for multidimensional Weyl sums*, *Geometric and Functional Analysis* **23** (2013), no. 6, 1962–2024.
- [ST42] A H Stone and John W Tukey, *Generalized “sandwich” theorems*, *Duke Mathematical Journal* **9** (1942), no. 2, 356–359.
- [Tao] Terence Tao, *A partial converse to Bézout’s theorem*, <https://terrytao.wordpress.com/2012/09/25/a-partial-converse-to-bezouts-theorem/>.
- [Woo14] Trevor D Wooley, *Translation invariance, exponential sums, and Waring’s problem*, arXiv preprint arXiv:1404.3508 (2014).