Factorization of sparse polynomials of bounded individual degree

Master's Thesis report submitted to

Chennai Mathematical Institute

in partial fulfilment for the award of the degree of

Master of Science

in

Computer Science

by

Sanyam Agarwal

(MCS202017)

Under the supervision of

Prof. Nitin Saxena

Department of Computer Science

Chennai Mathematical Institute

May, 2022

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: May 19, 2022

(Sanyam Agarwal)
(MCS202017)

CERTIFICATE

This is to certify that the project report entitled "**Factorization of sparse polynomials of bounded individual degree**" submitted by **Sanyam Agarwal** (Roll No. MCS202017) to Chennai Mathematical Institute towards partial fulfilment of requirements for the award of degree of Master of Science in Computer Science is a record of bona fide work carried out by him under my supervision and guidance during Oct'21-May'22.

Date: May 19, 2022

Prof. Nitin Saxena
Dept. of Comp. Sci. & Engg.
IIT Kanpur

Abstract

Name of the student: **Sanyam Agarwal** Roll No: **MCS202017**

Thesis title: **Factorization of sparse polynomials of bounded individual degree**

Thesis supervisor: **Prof. Nitin Saxena**

The motivation of this thesis is to obtain sparsity bounds on factors of $n-$variate polynomials of constant bounded individual degree d, under a given field $\mathbb{F}$. Given a $n-$variate polynomial f, *sparsity* of f (denoted $||f||$) is defined as the number of distinct monomials in f. Mathematically, given $f = g \cdot h$, we aim to bound $||g||, ||h||$ in terms of $||f||$.

Presently, the best known upper bounds for the sparsity of factors of a polynomial f are quasipolynomial in $\|f\|$, as shown in [BSV20]. However, it is conjectured in [Vol17] that the true bound for sparsity of factors is polynomial in $||f||$. [BSV20] get their result by using an elegant convex polytope representation of a polynomial to get a upper bound the size of the polytope in terms of its vertex set. Simultaneously, they also describe a polytope for which the above stated bound is tight, which might generate a polynomial that refutes the polynomial sparsity conjecture. But in our first result, we are able to show that the limiting polytope example infact doesn't kill the polynomial sparsity conjecture completely. Further, we generalize those proof techniques to come up with sparsity bounds for factors of a few class of polynomials under certain special conditions, and show the limitations of these techniques.

Acknowledgements

I would like to begin by thanking Prof. Nitin Saxena for his guidance throughout this project. Regular interactions with him has what made this thesis possible, and he was always on hand to provide new insights and ideas whenever we were stuck somewhere. I will also like to thank Pranav Bisht, for his continued assistance during the project. He was always available for discussions and for clearing my stupid doubts. Additionally, I would like to thank Sagnik, Sayak, and all other lab mates for making my stay at IITK pleasurable.

At CMI, I want to thank Prof. Partha Mukhopadhyay for introducing me to complexity theory, and all the other professors who put in extra efforts to make our learning as enjoyable as possible, despite the course being predominantly online.

I would also like to take this opportunity to thank IITK for hosting me for a year, and the IITK administration, especially Rajesh ji for ensuring a smooth and pleasant stay.

In the end, I would like to thank my family who have always loved and supported me and I fondly dedicate my thesis to them.

Contents

List of Symbols

$\mathbb{R}$	The field of **real numbers**
$\mathbb{Z}$	The ring of **integers**
$\mathbb{Q}$	The field of **rational numbers**
$\mathbb{Z}_{\geq 0}$	The set of **non-negative integers**
$\mathbb{N}$	The set of **natural numbers**
$\mathbb{F}$	Any field
$\mathbb{F}^n$	$n-$dimensional *vector space* over a field $\mathbb{F}$
$[n]$	The set of all positive integers upto n

Chapter 1

Introduction

Polynomials have been long studied in mathematics and provide many intriguing and natural questions that remain unsolved to this day. One such interesting class of questions relates to factors and factorization of polynomials. Another notion that has long fascinated mathematicians and computer scientists is the time and space complexity of any solution that they develop for a given problem, with the general aim of devising more and more efficient solutions. In this work, we look at the questions related to the "size" of the factors of a given polynomial with respect to it's own size, and how it has implications on the required time in which we can calculate these factors. Formally, the "size" of a polynomial is measured by its **sparsity**, which can be understood as the number of unique monomials in the polynomial f, and is denoted as $\|f\|$.

Factoring of polynomials has been a long-studied problem in mathematics and has many known applications like cryptography ([CR88]), derandomization ([KI04]), and list decoding ([Sud97]). There has been a lot of work in this area with several *randomized algorithms* known for factoring of multivariate polynomials, courtesy [vzGK85], [Kal87], and [Kal89] among others. The *Polynomial Identity Testing (PIT)* problem is one of the most fundamental problems in algebraic complexity. In

the PIT problem, given a polynomial $f \in \mathbb{F}[x_1, ..., x_n]$ represented by a small arithmetic circuit, the aim is to find whether f is identically zero. In [KSS14], the authors showed that the problem of derandomizing multivariate polynomial factorization is equivalent to the problem of derandomizing PIT for *general arithmetic circuits*, in both the white-box and the black-box settings. For other interesting open circuit classes, the equivalence is still not known and left as an open question in [KSS14]. To show the equivalence for any particular class, one needs to tackle the problem of *Factor Closure*, which asks for the upper bound on the size of the factors of a polynomial $f \in \mathbb{F}[x_1, ..., x_n]$. One such interesting class of polynomials to consider are *sparse polynomials*, or polynomials with few monomials.

Factoring of sparse polynomials has attracted significant interest over the past three decades. It was initiated by the work of [vzGK85] that gives the first randomized algorithm for factorization of sparse multivariate polynomials. The time complexity of this algorithm is *polynomially dependent* on the sparsity of the factors of the underlying polynomial, thus naturally raising the question of finding efficient bounds on the sparsity of factors of a sparse polynomial. However, sometimes we also need to constrain the individual degree of the variables in the polynomial, to ensure efficiency of algorithms. Consider the example,

Example 1.1. *([BSV20]) Let $\mathbb{F} = \mathbb{F}_p$ where p is prime, and let $0 < d < p$.*
Let $f = (x_1 + x_2 + ... + x_n)^p \Rightarrow f = x_1^p + ... + x_n^p \Rightarrow \|f\| = n$.
Let $g = (x_1 + x_2 + ... + x_n)^d$ be a factor of f.
But, $\|g\| = \binom{d+n-1}{d} \approx n^d \Rightarrow \|g\| = \|f\|^d$

Thus, if d is unbounded, the size of the factors could blow-up, making it impossible to get efficient factorization algorithms, as each factorization algorithm must output the factors monomial by monomial, and hence has a time complexity atleast linear in the sparsity of factors. Thus, we restrict ourselves to the *bounded individual degree* domain.

For multilinear polynomials ($d = 1$), [SV10] were able to give a derandomized factoring algorithm. Using factor sparsity bounds for multilinear polynomials, along

with derandomization of a certain PIT problem that arises, they are able to give an efficient factoring algorithm. In [Vol17], the author was able to extend it to the case of multiquadratic polynomials ($d \leq 2$), by first showing a non-trivial factor sparsity bound (Theorem 2.3) and then derandomizing the polynomial factorization problem. These techniques fail to generalise for $d > 2$. However, [BSV20] were the first to give a **deterministic factoring algorithm** that says that if, $f \in \mathbb{F}[x_1, ..., x_n]$ is a polynomial with sparsity $\|f\| = s$ and individual degrees of its variables bounded by d, then f can be deterministically factored in time $s^{poly(d)\log n}$. They achieve this by giving the first *quasi-polynomial* sparsity bound for factors of sparse polynomials of bounded individual degree $d > 2$, by cleverly leveraging the *Newton Polytope* representation of a polynomial, which was first introduced in a now redacted paper [DdO14]. However, [Vol17] conjectures that the true bound for sparsity of factors of f is $poly(\|f\|)$. Formally,

Conjecture 1.1 ([Vol17])**.** *(Polynomial sparsity conjecture) There exists a function $\nu : \mathbb{N} \to \mathbb{N}$ such that if $f \in \mathbb{F}[x_1, x_2, ..., x_n]$ is a polynomial with individual degrees at most d, then*

$$g \mid f \Rightarrow \|g\| \leq \|f\|^{\nu(d)}$$

Remark 1.2. Notice, that $\|g\|$ has no dependence on n, and hence the conjectured sparsity is polynomial in terms of $\|f\|$ since d is a constant.

As a first step towards this goal, in [BS22] the authors were able to show a $s^{poly(d)}$ sparsity bound for factors of *symmetric polynomials* over any field. This thesis is an attempt at characterizing certain other classes of polynomials for which we can get *polynomial* sparsity bounds for their factors.

1.1 Contributions of this Thesis

The main aim of this thesis is to move closer towards getting polynomial sparsity bounds on factors of a polynomial, as per Conjecture 1.1.

Presently, the best known upper bounds for the sparsity of factors of a polynomial f are $||f||^{O(d^2 log\, n)}$, as shown in [BSV20]. They get their result by using the convex polytope representation of a polynomial to upper bound the size of the polytope (*denoted* E) in terms of its vertex set (*denoted* V) as $E \leq V^{O(d^2 log\, n)}$. At the same time though, they also provide an example of a polytope (called *Hadamard polytope*) for which the above stated bound is tight, *i.e.*, $E = V^{\Theta(log\, n)}$, when d is a constant. Informally, this means we can have a 'dense' polytope with a 'sparse' vertex set. This raises the possibility that the polynomial corresponding to the *Hadamard polytope* could be a factor of a 'sparse' polynomial, consequently refuting the sparsity conjecture. But in our first result Corollary 3.6, we are able to show that the product of any 'sparse' polynomial with the polynomial corresponding to the *Hadamard polytope* always gives a 'dense' polynomial, preserving hope for the sparsity conjecture. We then generalize those proof techniques to come up with improved sparsity bounds for factors of certain classes of polynomials in Theorem 3.8 and Theorem 3.12. Moreover, we provide some counterexamples to show why our proof techniques can't be generalised further, and why certain other approaches may be doomed too.

Chapter 2

Preliminaries

In this chapter, we will briefly discuss the terminology that will be used in the thesis. Along with this, we will also introduce some important results upper bounding the sparsity of factors.

2.1 Basic notations and definitions

We use $\mathbf{x}$ to denote the variable set $\{x_1, x_2, ..., x_n\}$. For a field $\mathbb{F}$, we define $\mathbb{F}[\mathbf{x}]$ as the ring of polynomials in variables $\mathbf{x}$ with the coefficients coming from $\mathbb{F}$. For a poly $f \in \mathbb{F}[\mathbf{x}]$, a monomial m in f is denoted as $m \triangleq \mathbf{x}^{\mathbf{e}} = \prod_{i=1}^{n} x_i^{e_i}$ where $\mathbf{e} \in \mathbb{Z}_{\geq 0}^n$. Also, the set of monomials of polynomial f is denoted as M_f. A polynomial $f \in \mathbb{F}[\mathbf{x}]$ is said to have *individual degree* d, if for every monomial $m = \prod_{i=1}^{n} x_i^{e_i}$ in f, $e_i \leq d$. Finally, $f \in \mathbb{F}[\mathbf{x}]$ is *multilinear* if $\forall\ x_i \in f$, $deg(x_i) \leq 1$. Similarly, f is *multiquadratic* if $\forall\ x_i \in f$, $deg(x_i) \leq 2$. Apart from this, we use $|A|$ to denote *cardinality* of a set A.

Definition 2.1 (Support of a polynomial)**.** Let $f \in \mathbb{F}[\mathbf{x}]$ such that,

$$f = \sum_{i=1}^{s} a_{i_1} a_{i_2} ... a_{i_n} \cdot x_1^{i_1} x_2^{i_2} ... x_n^{i_n}$$

the support of f (denoted as $supp(f)$) is defined as:

$$supp(f) = \{(i_1, i_2, ..., i_n) \mid a_{i_1} a_{i_2} ... a_{i_n} \neq 0\} \quad \subseteq \mathbb{R}^n$$

Definition 2.2 (Sparsity of a polynomial)**.** Let $f \in \mathbb{F}[\mathbf{x}]$. The sparsity of f (denoted as $\|f\|$) is defined as:

$$\|f\| = |supp(f)|$$

2.2 Sparsity bounds using induction

We can use a simple induction argument (courtesy [Vol17]), that helps us prove that all factors of "sparse" polynomial with individual degree ≤ 2 are also "sparse".

Theorem 2.3. *Let $f, g \in \mathbb{F}[\mathbf{x}]$ such that f is multiquadratic. Then,*

$$g \mid f \Rightarrow \|f\| \geq \|g\|$$

Proof. Proof by induction on number of variables n.

Basis: n=0. Thus, $\|f\| = 1 = \|g\|$

Hypothesis: Let it hold for number of variables $< n$.

Induction: Let f have n variables. Suppose $f = g \cdot h$ for some $h \in \mathbb{F}[\mathbf{x}]$. Let $var(g), var(h)$ be the variable sets of polynomials g, h. If $var(g) \cap var(h) = \phi$. Then, $g_i \cdot h_j$ is unique for all monomials $g_i \in g$ and $h_j \in h$. Otherwise, they would have to share atleast one variable. Thus, in this case, $\|f\| = \|g \cdot h\| = \|g\|\|h\| \Rightarrow \|f\| \geq \|g\|$. Otherwise, let $y \in var(g) \cap var(h)$. Write $g = g_1 y + g_0$, and $h = h_1 y + h_0$, where g_0, g_1, h_0, h_1 are free of y, and hence polynomials in $n-1$ variables. Also, by the above representation, $\|g\| = \|g_1\| + \|g_0\|, \|h\| = \|h_1\| + \|h_0\|$. $f = g \cdot h = g_1 h_1 y^2 + (g_1 h_0 + g_0 h_1) y + g_0 h_0 \Rightarrow \|f\| = \|g_1 h_1\| + \|g_1 h_0 + g_0 h_1\| + \|g_0 h_0\| \geq \|g_1 h_1\| + \|g_0 h_0\| \geq \|g_1\| + \|g_0\|$ (by *induction hypothesis*). Hence, $\|f\| \geq \|g_1\| + \|g_0\| = \|g\|$. □

This result proves the sparsity conjecture mentioned in [Vol17] for $ideg \leq 2$. However, it fails to generalise for the case where $ideg \geq 3$, motivating the need to look for other proof techniques.

2.3 Sparsity Bounds using Newton polytope - A convex geometry approach

2.3.1 Notations and definitions

We will use boldface for vectors, and regular font for scalars, *i.e.*, a vector $\vec{v} = (v_1, v_2, ..., v_n) \in \mathbb{R}^n$ will be represented as $\mathbf{v}$, and a scalar $s \in \mathbb{R}$ will be denoted as s. Multiplication of a vector $\mathbf{x}$ by a scalar z will be denoted as the product $z \cdot \mathbf{x}$. Dot product between two vectors $\mathbf{x}, \mathbf{y}$ will be $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$. We will now define a few important terms:

Definition 2.4 (Convex span of a set)**.** Let $P = \{p_2, p_2, ..., p_r\}$ be a set. The *convex span* of P (denoted $CS(P)$) is defined as:

$$CS(P) = \left\{ \sum_{i=1}^{r} \lambda_i p_i \;\middle|\; \lambda_i \geq 0 \; \forall i \;\; \& \;\; \sum_{i=1}^{r} \lambda_i = 1 \right\}$$

Definition 2.5 (Polytope)**.** A set $P \in \mathbb{R}^n$ is a Polytope if $\exists$ a finite set of points $\mathbf{y}_1, ..., \mathbf{y}_k \in \mathbb{R}^n$ such that $P = CS(\mathbf{y}_1, ..., \mathbf{y}_k)$

Definition 2.6. (Convex set) A convex set $C \in \mathbb{R}^n$ is a set such that $CS(\mathbf{x}, \mathbf{y}) \subseteq C \;\; \forall \, \mathbf{x}, \mathbf{y} \in C$

Definition 2.7. (Supporting hyperplane) A supporting hyperplane H of a convex set $C \in \mathbb{R}^n$ is a hyperplane denoted by $\mathbf{h} \cdot \mathbf{x} = a$, where $\mathbf{h} \in \mathbb{R}^n \setminus \{0\}, a \in \mathbb{R}$ such that H intersects the closure of C in $\mathbb{R}^n$ and $\mathbf{h} \cdot \mathbf{x} \leq a \;\; \forall \, \mathbf{x} \in C$.

Definition 2.8. (Face of a polytope) Let P be a polytope. A face of P is the intersection of P with a supporting hyperplane H. That is, $F = P \cap \{\mathbf{h} \cdot \mathbf{x} = a \mid \mathbf{x} \in P\}$ is a face of P.

Definition 2.9. (Vertex of a polytope) Let P be a polytope. A vertex v of P is a *face* of dimension 0, *i.e.*, v can't be written as a convex combination of two points $x, y \in P \setminus \{v\}$. Formally,

$$v \neq \lambda x + (1-\lambda)y \quad \forall\, x, y \in P \setminus \{v\}, \lambda \in [0,1]$$

The vertex set of P is denoted as $V(P)$.

Corollary 2.10. *By Definition 2.7, Definition 2.8, Definition 2.9, for every vertex* $\mathbf{v} \in V(P)$, $\exists \mathbf{h} \in \mathbb{R}^n \setminus \{0\}$ *such that* $\mathbf{h} \cdot \mathbf{x} < \mathbf{h} \cdot \mathbf{v} \quad \forall\, \mathbf{x} \in P$.

Definition 2.11 (Minkowski Sum of two polytopes)**.** Let A, B be polytopes. The *Minkowski Sum* of the two polytopes (denoted as $A + B$) is defined as:

$$A + B = \{u + v \mid u \in A, v \in B\}$$

Lastly, we will define an alternate way of representing a polynomial.

Definition 2.12 (Newton Polytope of a polynomial)**.** Let $f \in \mathbb{F}[\mathbf{x}]$. The Newton polytope of f (denoted as P_f) is defined as:

$$P_f = CS(supp(f)) \subseteq \mathbb{R}^n$$

where $supp(f)$ is as defined in Definition 2.1.

This is the most important definition, and will be used frequently in the subsequent sections.

2.3.2 Important properties of Newton polytopes

We will now prove some interesting results on polytopes in general, and use them to get non-trivial properties about polynomials by leveraging the connection between polytopes and polynomial shown in Definition 2.12

Lemma 2.13. *Let V be a vector space over $\mathbb{R}$.*
Given $\mathbf{x}_i, \mathbf{y}_j \in V \ \forall i \in [n]$, and $j \in [m]$. Let $a_i, b_j \in \mathbb{R} \ \forall i \in [n]$, and $j \in [m]$ such that $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$.
Then, we can always get $s_{ij} \ \ \forall i \in [n], j \in [m]$ such that:

1. $$\sum_{i=1}^{n} a_i \cdot \mathbf{x}_i + \sum_{j=1}^{m} b_j \cdot \mathbf{y}_j = \sum_{i=1}^{n} \sum_{j=1}^{m} s_{ij} \cdot (\mathbf{x}_i + \mathbf{y}_j)$$

   and

2. $$\sum_{i=1}^{n} \sum_{j=1}^{m} s_{ij} = \sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$$

Proof. By **induction** on m.
Basis (m=1): Since, we have $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j$, it means that $\sum_{i=1}^{n} a_i = b_1$
$\Rightarrow \sum_{i=1}^{n} a_i \cdot \mathbf{x}_i + b_1 \cdot \mathbf{y}_1 = \sum_{i=1}^{n} a_i \cdot (\mathbf{x}_i + \mathbf{y}_1)$
Thus, condition holds.
Hypothesis: Let it hold for all values of $m < r$.
Induction (m=r):

$$\sum_{i=1}^{n} a_i \cdot \mathbf{x}_i + \sum_{j=1}^{r} b_j \cdot \mathbf{y}_j = \sum_{i=1}^{n} s_{ir} \cdot (\mathbf{x}_i + \mathbf{y}_r) + \sum_{i=1}^{n} (a_i - s_{ir}) \cdot \mathbf{x}_i + \sum_{j=1}^{r-1} b_j \cdot \mathbf{y}_j$$

We have,

$$\sum_{i=1}^{n} (a_i - s_{ir}) = \sum_{i=1}^{n} a_i - \sum_{i=1}^{n} s_{ir} = \sum_{i=1}^{n} a_i - b_r = \sum_{j=1}^{r-1} b_j - b_r = \sum_{j=1}^{r-1} b_j$$

Thus, we can apply induction hypothesis to get,

$$\sum_{i=1}^{n}(a_i - s_{ir}) \cdot \mathbf{x}_i + \sum_{j=1}^{r-1} b_j \cdot \mathbf{y}_j = \sum_{i=1}^{n}\sum_{j=1}^{r-1} s_{ij} \cdot (\mathbf{x}_i + \mathbf{y}_j)$$

such that $\sum_{i=1}^{n}\sum_{j=1}^{r-1} s_{ij} = \sum_{j=1}^{r-1} b_j$. Thus,

$$\sum_{i=1}^{n} a_i \cdot \mathbf{x}_i + \sum_{j=1}^{r} b_j \cdot \mathbf{y}_j = \sum_{i=1}^{n} s_{ir} \cdot (\mathbf{x}_i + \mathbf{y}_r) + \sum_{i=1}^{n}(a_i - s_{ir}) \cdot \mathbf{x}_i + \sum_{j=1}^{r-1} b_j \cdot \mathbf{y}_j$$

$$= \sum_{i=1}^{n} s_{ir} \cdot (\mathbf{x}_i + \mathbf{y}_r) + \sum_{i=1}^{n}\sum_{j=1}^{r-1} s_{ij} \cdot (\mathbf{x}_i + \mathbf{y}_j) = \sum_{i=1}^{n}\sum_{j=1}^{r} s_{ij} \cdot (\mathbf{x}_i + \mathbf{y}_j)$$

. Also, we can easily check that

$$\sum_{i=1}^{n}\sum_{j=1}^{r} s_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{r-1} s_{ij} + \sum_{i=1}^{n} s_{ir} = \sum_{j=1}^{r-1} b_j + b_r = \sum_{j=1}^{r} b_j$$

Hence, the condition holds for $m = r$, and as a result for all values of m. □

Theorem 2.14 ([Sch00])**.** *Let P_1, P_2 be two Newton polytopes. Then their Minkowski Sum $P_1 + P_2$ is also a Newton polytope.*

Proof. We will be done if we can show that the Minkowski Sum $P_1 + P_2$ is the convex span of some set by Definition 2.5.

Let $V(P_1) :=$ vertex set of $P_1 = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r\}$

Let $V(P_2) :=$ vertex set of $P_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_s\}$

Let $V(P_1) + V(P_2) = M \triangleq \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \in V(P_1), \mathbf{v} \in V(P_2)\}$. Then,

$$CS(M) = \left\{ \lambda_{11}(\mathbf{u}_1 + \mathbf{v}_1) + ... + \lambda_{1s}(\mathbf{u}_1 + \mathbf{v}_s) + ... + \lambda_{rs}(\mathbf{u}_r + \mathbf{v}_s) \;\middle|\; \lambda_{ij} \geq 0 \,\forall i \,\forall j,\ \sum\sum \lambda_{ij} = 1 \right\}$$

$$= \left\{ \sum_{j=1}^{s} \lambda_{1j}\mathbf{u}_1 + ... \sum_{j=1}^{s} \lambda_{rj}\mathbf{u}_r + \sum_{i=1}^{r} \lambda_{i1}\mathbf{v}_1 + ... + \sum_{i=1}^{r} \lambda_{is}\mathbf{v}_s \;\middle|\; \lambda_{ij} \geq 0 \,\forall i \,\forall j,\ \sum\sum \lambda_{ij} = 1 \right\}$$

Now, define the following terms:

$$\forall i \in [r] \beta_i \triangleq \sum_{j=1}^{s} \lambda_{ij} \;\; and \;\; \forall j \in [s], \gamma_j \triangleq \sum_{i=1}^{s} \lambda_{ij}$$

Also, we have $\sum_{i=1}^{r} \beta_i = \sum_{i=1}^{r} \sum_{j=1}^{s} \lambda_{ij} = 1$. Similarly, $\sum_{j=1}^{s} \gamma_j = 1$. Therefore,

$$CS(M) = \left\{ \beta_1 \mathbf{u}_1 + ... + \beta_r \mathbf{u}_r + \gamma_1 \mathbf{v}_1 + ... + \gamma_s \mathbf{v}_s \middle| \beta_i \geq 0, \gamma_j \geq 0, \sum \beta_i = 1, \sum \gamma_j = 1 \right\}$$

But $\beta_1 \mathbf{u}_1 + ... + \beta_r \mathbf{u}_r$ is a convex combination of vertices of P_1 and hence lies in the polytope P_1, and similarly $\gamma_1 \mathbf{v}_1 + ... + \gamma_s \mathbf{v}_s$ is a point in the polytope P_2. Thus,

$$CS(M) = \left\{ p_1 + p_2 \;\middle|\; p_1 \in P_1, p_2 \in P_2 \right\} \Rightarrow CS(M) \subseteq P_1 + P_2 \tag{2.1}$$

Conversely,

$$P_1 + P_2 = \left\{ \mu_1 + \mu_2 \;\middle|\; \mu_1 \in P_1, \mu_2 \in P_2 \right\}$$

But, any point in a polytope is a convex combination of it's vertices. Thus, μ_1 is convex combination of $\{\mathbf{u}_i\}$ and μ_2 is convex combination of $\{\mathbf{v}_j\}$. Using Lemma 2.13 with $\sum a_i = \sum b_j = 1$, we can write $\mu_1 + \mu_2$ as a convex combination of $\{\mathbf{u}_i + \mathbf{v}_j\}_{i=1,j=1}^{i=r,j=s}$. Thus, $P_1 + P_2 \subseteq CS(V(P_1) + V(P_2)) = CS(M)$. Combining with (2.1), we get $P_1 + P_2 = CS(V(P_1) + V(P_2))$. Hence, the *Minkowski Sum* $P_1 + P_2$ is also a Newton polytope. □

Theorem 2.15. *([Sch00]) Let P_1, P_2 be ploytopes in $\mathbb{R}^n$ and $P_1 + P_2$ be their Minkowski Sum. Then,*

1. *every vertex $\mathbf{w} \in V(P_1 + P_2)$ can be expressed uniquely as $\mathbf{w} = \mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in V(P_1), \mathbf{v} \in V(P_2)$ and,*
2. *For every vertex $\mathbf{u} \in V(P_1)$, there exists a vertex $\mathbf{v} \in V(P_2)$ such that $\mathbf{u} + \mathbf{v} \in V(P_1 + P_2)$*

Proof.

1. Let $\mathbf{w} \in V(P_1 + P_2)$. It follows that there exists a hyperplane $\mathbf{h} \in \mathbb{R}^n$, such that $\mathbf{h} \cdot \mathbf{r} < \mathbf{h} \cdot \mathbf{w}$ for all $\mathbf{r} \in P_1 + P_2 \setminus \{\mathbf{w}\}$. Since $\mathbf{w} \in P_1 + P_2$, it can be written as $\mathbf{w} = \mathbf{x}_1 + \mathbf{x}_2$ where $\mathbf{x}_i \in P_i$. Suppose there exists $\mathbf{y}_1 \in P_1, \mathbf{y}_2 \in P_2$ such that $(\mathbf{y}_1, \mathbf{y}_2) \neq (\mathbf{x}_1, \mathbf{x}_2)$ and $\mathbf{w} = \mathbf{y}_1 + \mathbf{y}_2$. Now, $\mathbf{y}_1 + \mathbf{x}_2, \mathbf{x}_1 + \mathbf{y}_2 \in P_1 + P_2$ implies that $\mathbf{h} \cdot (\mathbf{y}_1 + \mathbf{x}_2) < \mathbf{h} \cdot \mathbf{w} = \mathbf{h} \cdot (\mathbf{x}_1 + \mathbf{x}_2)$ giving us $\mathbf{h} \cdot \mathbf{y}_1 < \mathbf{h} \cdot \mathbf{x}_1$. Similarly, $\mathbf{h} \cdot \mathbf{y}_2 < \mathbf{h} \cdot \mathbf{x}_2$. Now $\mathbf{h} \cdot \mathbf{w} = h \cdot (\mathbf{y}_1 + \mathbf{y}_2) < \mathbf{h} \cdot (\mathbf{x}_1 + \mathbf{x}_2) < \mathbf{h} \cdot \mathbf{w}$. Thus, we get a contradiction. Hence, $\mathbf{w} \neq \mathbf{y}_1 + \mathbf{y}_2$, and is uniquely represented as $\mathbf{w} = \mathbf{x}_1 + \mathbf{x}_2$. Further, consider $\mathbf{r} = \mathbf{z}_1 + \mathbf{x}_2$ where $\mathbf{z}_1 \in P_1 \setminus \{\mathbf{x}_1\}$. Since, $\mathbf{r} \in P_1 + P_2$ it means that $\mathbf{h} \cdot \mathbf{r} < \mathbf{h} \cdot \mathbf{w} \Rightarrow \mathbf{h} \cdot \mathbf{z}_1 < \mathbf{h} \cdot \mathbf{x}_2 \ \ \forall \mathbf{z}_1 \in P_1 \setminus \{\mathbf{x}_1\}$ Thus, we have $\mathbf{x}_1 \in V(P_1)$ (by Corollary 2.10). Similarly, $\mathbf{x}_2 \in V(P_2)$.

2. Proof from [BSV20].

   Let $\mathbf{u} \in V(P_1)$. By definition of a vertex, there exists a hyperplane $\mathbf{h} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, such that $\forall \mathbf{w} \in P_1 \setminus \{\mathbf{u}\}$, $\mathbf{h} \cdot \mathbf{w} < \mathbf{h} \cdot \mathbf{u}$. Define a *degree 1* polynomial $p_h(\mathbf{x}) = \mathbf{h} \cdot \mathbf{x} - \mathbf{h} \cdot \mathbf{u}$. Thus, $p_h(\mathbf{u}) = 0$ and $\forall \mathbf{w} \in P_1 \setminus \{\mathbf{u}\}, p_h(\mathbf{w}) < 0$. Consider polynomial $p_h^{'}(\mathbf{x}) = p_h(\mathbf{x}) - d$, where $d \in \mathbb{R}$. For large enough d, $\forall \mathbf{y} \in P_2$, $p_h^{'}(\mathbf{y}) < 0$. Start decreasing the value of d, until $\exists \mathbf{v} \in P_2$ such that $p_h^{'}(\mathbf{v}) = 0$. We can ensure that only one such point in P_2 has this property by carefully choosing the initial hyperplane $\mathbf{h}$ from the family of hyperplanes that satisfy the initial condition for $\mathbf{u}$. Thus, we have that $p_h^{'}(\mathbf{v}) = 0$, and $\forall \mathbf{y} \in P_2 \setminus \{\mathbf{v}\}, p_h^{'}(\mathbf{y}) < 0$. Thus, $\mathbf{v} \in V(P_2)$ by Corollary 2.10. Now, we will show that $\mathbf{u} + \mathbf{v} \in V(P_1 + P_2)$. Define $\overline{p}(\mathbf{x}) = \mathbf{h} \cdot \mathbf{x} - 2\mathbf{h} \cdot \mathbf{u} - d$. It follows that, $\overline{p}(\mathbf{u} + \mathbf{v}) = \mathbf{h} \cdot (\mathbf{u} + \mathbf{v}) - 2\mathbf{h} \cdot \mathbf{u} - d = \mathbf{h} \cdot \mathbf{v} - \mathbf{h} \cdot \mathbf{u} - d = p_h^{'}(\mathbf{v}) = 0$. Also, for any $\mathbf{w} \in P_1 \setminus \{\mathbf{u}\}, \mathbf{y} \in \setminus\{\mathbf{v}\}, \overline{p}(\mathbf{w} + \mathbf{y}) = \mathbf{h} \cdot (\mathbf{w} + \mathbf{y}) - 2\mathbf{h} \cdot \mathbf{u} - d = (\mathbf{h} \cdot \mathbf{w} - \mathbf{h} \cdot \mathbf{u}) + (\mathbf{h} \cdot \mathbf{y} - \mathbf{h} \cdot \mathbf{u} - d) = p_h(\mathbf{w}) + p_h^{'}(\mathbf{y}) < 0$. Thus, the hyperplane represented as $\mathbf{h} \cdot \mathbf{x} = b$, where $b = 2\mathbf{h} \cdot \mathbf{u} + d$ acts as the supporting hyperplane for $P_1 + P_2$ such that $\mathbf{h} \cdot \mathbf{w} < b \ \ \forall w \in P_1 + P_2 \setminus \{\mathbf{u} + \mathbf{v}\}$, and $\mathbf{h} \cdot (\mathbf{u} + \mathbf{v}) = b$. Hence, $\mathbf{u} + \mathbf{v}$ is a vertex of $P_1 + P_2$. □

Corollary 2.16. *Let P_1, P_2 be ploytopes in $\mathbb{R}^n$. Then*

$$|V(P_1 + P_2)| \geq max\ \{|V(P_1)|, |V(P_2)|\}$$

Proof. By item 2 of Theorem 2.15, for each vertex of P_1, there exists a vertex of $P_1 + P_2$. Thus, $|V(P_1 + P_2)| \geq |V(P_1)|$. Similarly, for each vertex of P_2, there exists a vertex of $P_1 + P_2$. Thus, $|V(P_1 + P_2)| \geq |V(P_2)|$. □

The following is a result observed by Ostrowski ([Ost21]) in 1921, which will play an important part in proving an lower bound on the size of vertex set of polynomial f in terms of the sizes of the vertex set of its factors g and h.

Theorem 2.17. *Let $f \in \mathbb{F}[\mathbf{x}]$, such that $f = g \cdot h$. Then, P_f is the* Minkowski Sum *of P_g and P_h, i.e., $P_f = P_g + P_h$*

Proof. Let $supp(g) = \{g_1, ..., g_n\}, supp(h) = \{h_1, ..., h_m\}$. We know that $P_f = CS(supp(f))$, $P_h = CS(supp(h))$. Since $f = g \cdot h \Rightarrow supp(f) = \cup_{i\in[n],j\in[m]}\{g_i + h_j\}$. Also, $P_g = CS(supp(g))$.

$$P_g + P_h = \left\{ \sum_{i=1}^{n} \lambda_i g_i + \sum_{j=1}^{m} \beta_j h_j \;\middle|\; \lambda_i \geq 0\ \forall i, \sum_{i=1}^{n} \lambda_i = 1, \beta_j \geq 0\ \forall j, \sum_{j=1}^{m} \beta_j = 1 \right\}$$

Using Lemma 2.13 with $\sum \lambda_i = \sum \beta_j = 1$, we can write it as a convex combination of $\{g_i + h_j\}_{i=1,j=1}^{i=n,j=m}$. Thus, $P_g + P_h \subseteq CS(supp(f)) = P_f$.
Conversely, every monomial $\mathbf{x}^{\mathbf{e}} \in f$ is generated via product of some monomials $\mathbf{x}^{\mathbf{e}_g} \in g$ and $\mathbf{x}^{\mathbf{e}_h} \in h$. Thus, $supp(f) \subseteq supp(g) + supp(h) \Rightarrow P_f = CS(supp(f)) \subseteq CS(supp(g) + supp(h)) = P_g + P_h$. □

In [DdO14], the authors cleverly exploited the connection between a polynomial and a polytope, and the above properties to give an elegant approach for bounding the sparsity of factors of a general polynomial of some given *ideg d*. Now, we present the sparsity bound provided by them:

Corollary 2.18. *(Sparsity bound of factors) Let $f, g, h \in \mathbb{F}[\mathbf{x}]$ s.t. $f = g \cdot h$. Then,*

$$\|f\| \geq |V(P_f)| \geq max\{|V(P_g)|, |V(P_h)|\}$$

Proof. Using Theorem 2.17, and Corollary 2.16, we get $|V(P_f)| \geq max\{|V(P_g)|, |V(P_h)|\}$. Also, $V(P_f) = V(CS(supp(f)) \subseteq supp(f) \Rightarrow |V(P_f)| \leq |supp(f)| = \|f\|$. Hence, we get the desired result. □

However, we can see that this bound is not tight. Consider the example,

Example 2.1. *Let $\mathbb{F} = \mathbb{R}[x, y]$. Define, $g = \sum_{a=0}^{d} \sum_{b=0}^{d} x^a y^b$ and $h = 1$. Here, $\|g\| = d^2$ and $V(P_g) = \{(0,0), (0,d), (d,0), (d,d)\} = 2^2$.*
Let $f = g \cdot h$. By Corollary 2.18, we get that $\|f\| \geq max\{2^2, 1\} = 2^2 = |V(P_g)|$.
But, $\|f\| = \|g\| = d^2 = (2^{log_2 d})^2 = |V(P_g)|^{log_2 d}$

Thus, even though the approach did not eventually lead to an efficient sparsity bound, it did inspire further work in the direction of using convex geometry to bound the factors of sparse polynomials, notably in [BSV20].

2.3.3 Improved sparsity bounds for factors of sparse polynomials

A well known result in convex geometry is the *Catathéodory theoreom* (Theorem A.1), which states that any point in the convex hull of a $n - dimensional$ set U can be expressed as a convex combination of atmost $n + 1$ points. There also exists an "approximate" version of the Carathéodory theorem, which says that we can "reasonably" *uniformly* approximate any point in the convex hull of a $n - dimensional$ set U by only $\mathcal{O}(log\ n)$ points of U. It was first proposed in [Bar15]. Before presenting the proof, we will first introduce some notation.

Definition 2.19. (k-uniformity of a vector) Let $M = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\} \in \mathbb{R}^n$. A vector $\mathbf{u} \in M$ is defined to be $k - uniform$ with respect to M if there exists a multiset S of $[m]$ such that $|S| \leq k$ and $\mathbf{u} = \frac{1}{k}\sum_{i \in S} \mathbf{x}_i$

Also, for any given vector $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{y}_j$ denotes its $j - th$ coordinate $\forall\, j \in [n]$.

Theorem 2.20. *(Approximate Carathéodory theorem) Given a set of vectors* $U = \mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_m \in \mathbb{R}^n$ *with* $max_{\mathbf{u} \in U} \|\mathbf{u}\|_\infty \leq 1$ *and* $\epsilon > 0$. *For every* $\mathbf{z} \in CS(U)$, *there exists an* $\mathcal{O}(\frac{log\, n}{\epsilon^2})$ *uniform vector* $\mathbf{z}' \in CS(U)$ *such that* $\|\mathbf{z} - \mathbf{z}'\|_\infty \leq \epsilon$.

Proof. Since $\mathbf{z} \in CS(U) \Rightarrow \mathbf{z} = \sum_{i=1}^m a_i \mathbf{u}_i$ where $a_i \geq 0\ \forall i \in [m]$ and $\sum_{i=1}^m a_i = 1$. Consider this as a probability distribution, with the probability of sampling each $\mathbf{u}_i$ being a_i. Pick $t = \frac{log\, n}{\epsilon^2}$ samples independently from this distribution, with the resulting vectors being $\mathbf{y}_1, ..., y_t$.Now, each $\mathbf{y}_i$ is independently sampled from this distribution and $\mathbb{E}(\mathbf{y}_i) = \sum_{i=1}^m a_i \mathbf{u}_i = \mathbf{z}$. Define

$$\mathbf{z}' = \frac{\mathbf{y}_1 + ... + \mathbf{y}_t}{t}$$

Therefore, $\mathbb{E}(\mathbf{z}') = \frac{1}{t}\sum_{j=1}^t \mathbb{E}(\mathbf{y}_j) = \frac{1}{t}\sum_{j=1}^t \mathbf{z} = \mathbf{z}$. Consider the t independent samples $Y_1, Y_2, ..., Y_t$ of the random variable Y, such that for each $i \in [m]$, $Y = (\mathbf{u}_i)_j$ with probability a_i, *i.e.*, we are sampling the vectors coordinate-wise. Pick any coordinate $k \in [n]$. Then, clearly $\mathbb{E}(Y) = \frac{1}{t}\sum_{j=1}^t \mathbb{E}(Y_j) = \frac{1}{t}\sum_{j=1}^t \sum_{i=1}^m a_i (\mathbf{u}_i)_k = \frac{1}{t}\sum_{j=1}^t \mathbf{z}_k = \mathbf{z}_k$. Also, $\mathbf{z}'_k = \frac{Y_1 + ... + Y_t}{t}$. Using Theorem A.2, we get

$$\mathbb{P}\left(\left|\mathbf{z}'_k - \mathbf{z}_k\right| > \epsilon\right) \leq 2e^{-2\epsilon^2 t} = 2e^{-2\epsilon^2 \frac{log\, n}{\epsilon^2}} = 2e^{-2log\, n} = \frac{2}{n^2}$$

Now, $\|\mathbf{z}' - \mathbf{z}\|_\infty > \epsilon$ if $|\mathbf{z}'_k - \mathbf{z}_k| > \epsilon$ for atleast some $k \in [n]$. Using union bound,

$$\mathbb{P}\left(\left\|\mathbf{z}' - \mathbf{z}\right\|_\infty > \epsilon\right) = \cup_{k=1}^n \mathbb{P}\left(\left|\mathbf{z}'_k - \mathbf{z}_k\right| > \epsilon\right) \leq \sum_{k=1}^n \mathbb{P}\left(\left|\mathbf{z}'_k - \mathbf{z}_k\right| > \epsilon\right) \leq \sum_{k=1}^n \frac{2}{n^2} = \frac{2}{n}$$

Thus, $\mathbb{P}\left(\left\|\mathbf{z}' - \mathbf{z}\right\|_\infty \leq \epsilon\right) > 1 - \frac{2}{n} > 0 \quad \forall n \geq 2$ and hence a suitable $\frac{log\, n}{\epsilon^2}$ uniform vector $\mathbf{z}'$ exists that $\epsilon - approximates\ \mathbf{z}$. □

In [BSV20], the authors were able to cleverly use the "approximate" Carathéodory theorem, to get strong sparsity bounds on the factors of the polynomials of a sparse polynomial. We will now present the proof below.

Theorem 2.21. *([BSV20]) Let $E \subseteq \{0, 1, ..., d\}^n$. Let $U = V(CS(E))$. Then, there exists an absolute constant $\alpha \in \mathbb{R}$, such that, $|U|^{\alpha d^2 log\, n} \geq |E|$.*

Proof Idea. We will show that every vector $\mathbf{u} \in E$ is $\epsilon - approximated$ uniquely by the elements of U. We will then compare the size of E with the maximum number of $\epsilon - approximations$ possible from U to get the bound.

Proof. Let $E_d = \{\frac{1}{d}\mathbf{u} \mid \mathbf{u} \in E\} \subseteq [0, 1]^n$. Let $\epsilon = \frac{1}{3d}$ and $U_d = V(CS(E_d))$. Clearly, $|U| = |U_d|$. Using Theorem 2.20, for every $\mathbf{u}_d \in E_d$, $\exists$ an $\mathcal{O}(\frac{log\, n}{\epsilon^2}) = \mathcal{O}(d^2 log\ n)$ uniform vector $\mathbf{u}'_d \in U_d$ such that

$$\|\mathbf{u}_d - \mathbf{u}'_d\|_\infty \leq \epsilon = \frac{1}{3d} \tag{2.2}$$

Let $\mathbf{u}, \mathbf{v}$ be two distinct vectors in E. Then they must differ by atleast one coordinate, *i.e.*, $\exists j \in [n]$ such that $\mathbf{u}_j \neq \mathbf{v}_j \Rightarrow |(\mathbf{u})_j - (\mathbf{v})_j| \geq 1$. Hence, for the corresponding vectors $\mathbf{u}_d, \mathbf{v}_d \in E_d$, we have $|(\mathbf{u}_d)_j - (\mathbf{v}_d)_j| \geq \frac{1}{d}$. Thus,

$$\|\mathbf{u}_d - \mathbf{v}_d\|_\infty \geq \frac{1}{d} \tag{2.3}$$

Let $\mathbf{u}'_d, \mathbf{v}'_d$ be the $\mathcal{O}(d^2 log\ n)$ uniform vectors for any two vectors $\mathbf{u}_d, \mathbf{v}_d \in E_d$ respectively. We claim that $\mathbf{u}'_d \neq \mathbf{v}'_d$.
Suppose $\mathbf{u}'_d = \mathbf{v}'_d$. Using (2.2) and *triangle inequality*, we get

$$\|\mathbf{u}_d - \mathbf{v}_d\|_\infty = \|(\mathbf{u}_d - \mathbf{u}'_d) + (\mathbf{u}'_d - \mathbf{v}_d)\|_\infty = \|(\mathbf{u}_d - \mathbf{u}'_d) + (\mathbf{v}'_d - \mathbf{v}_d)\|_\infty \leq \|\mathbf{u}_d - \mathbf{u}'_d\|_\infty + \|\mathbf{v}'_d - \mathbf{v}_d\|_\infty \leq \frac{2}{3d}$$

which contradicts (2.3).
Now, total number of $\mathcal{O}(d^2 log\ n)$ approximations that can be generated by U_d is given by $|U_d|^{\mathcal{O}(d^2 log\, n)}$. Also, $\mathbf{u}'_d \neq \mathbf{v}'_d \Rightarrow$ for every vector $u_d \in E_d$ there is an unique

approximation. It follows that $|U_d|^{\mathcal{O}(d^2 log\ n)} \geq |E_d|$.
Thus, we get $|U|^{\alpha d^2 log\ n} \geq |E|$ for some $\alpha \in \mathbb{R}$. □

Corollary 2.22. *([BSV20]: Sparsity bounds for factors of sparse polynomials of bounded individual degree) Let $f, g \in \mathbb{F}[\mathbf{x}]$ such that f has individual degree d. Then,*

$$g \mid f \Rightarrow \|g\| \leq \|f\|^{\mathcal{O}(d^2\ \log\ n)}$$

Proof. Let $U_g = V(CS(supp(g)), U_f = V(CS(supp(f))$. By definition, $\|g\| = |supp(g)|$. By Theorem 2.21, $|U_g|^{\mathcal{O}(d^2\ \log\ n)} \geq \|g\|$. Also, by Corollary 2.16, $|U_f| \geq |U_g|$. Also, we know that $\|f\| \geq |U_f|$. Thus,

$$\|g\| \leq |U_g|^{\mathcal{O}(d^2\ \log\ n)} \leq |U_f|^{\mathcal{O}(d^2\ \log\ n)} \leq \|f\|^{\mathcal{O}(d^2\ \log\ n)}$$

□

Thus, we now have a *quasipolynomial* bound on the sparsity of any factor of a sparse polynomial with bounded individual degree. However, it still doesn't help us settle the *polynomial* sparsity of factors, conjectured in [Vol17]. At the same time, it is the best known bound currently, that works for any $s - sparse$ polynomial of individual degree d. By Theorem 2.3, we already know of a much stronger bound for factors of $s - sparse$ polynomial of individual degree $d \leq 2$.

The bound in Corollary 2.22 has been obtained by bounding the size of the vertex set of a polytope in relation to the size of the whole polytope, while looking only at the intergral points. Thus a natural next step would be to obtain better bounds for this relative size. However, along with the quasipolynomial bound, [BSV20] also provide an example of a polytope for which the bound is *tight*. Thus, we cannot hope to obtain universally better bounds using the convex geometry approach alone.

Claim 2.23. *([BSV20]: Hadamard Polytope) There is a set $E \subseteq \{-1, 0, 1\}^n$ such that $|V(CS(E))| = n$ and $|E| = n^{\Omega(log\ n)}$.*

Proof. Let $m \in \mathbb{N}$. Let $n = 2^m$, and let H be a $n \times n$ **Hadamard matrix**. Formally, $H \in \mathbb{R}^{n \times n}$ with $H_{ij} \in \{\pm 1\}$, such that index both rows and columns by vectors in $\mathbb{F}_2^m$, then the $(\mathbf{u}, \mathbf{v})$ entry of H is $(-1)^{\langle \mathbf{u}, \mathbf{v} \rangle}$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^m$ where $\langle \mathbf{u}, \mathbf{v} \rangle = (\sum_{i=1}^{m} \mathbf{u}_i \mathbf{v}_i)$.

Let $V \subseteq \{\pm 1\}$ be the set of column vectors of H. Then, we will show that the $|CS(V)|$ has atleast $n^{\Omega(log\ n)}$ elements $\in \{-1, 0, 1\}$. Thus, any polytope with vertex set as V will have the desired properties, and suffice for our claim. Recall that each element of V is indexed by an element of $\mathbb{F}_2^m$. Also, for each subspace $S \subseteq \mathbb{F}_2^m$, we will show that on taking the *uniform* convex span of elements of V that correspond to those columns that were indexed by vectors in S, we will get an **unique** element in $\{0, 1\}^n$. By Corollary A.4, total number of subspaces of $\mathbb{F}_2^m$ over $\mathbb{F}_2$ is $2^{\Omega(m^2)} = n^{\Omega(log\ n)}$. Also, $|V| = n$. Thus, this will prove our claim.

Let $\mathbf{c}_i, \mathbf{r}_j \in \mathbb{F}_2^m$ be the vectors indexing column i and row j of H respectively. For any subspace $S \subseteq \mathbb{F}_2^m$, let $\mathbf{u}_S \in \mathbb{R}^n$ be it's characteristic vector. If $\mathbf{c}_i \in S$ then $(\mathbf{u}_S)_i = 1$, else $(\mathbf{u}_S)_i = 0$. For showing that the *uniform* convex span of elements of V gives a vector in $\{0, 1\}^n$, we need to show that $\frac{1}{|S|}(H \cdot \mathbf{u}_S) \in \{0, 1\}^n$.

Consider $T = \{\mathbf{v} \in \mathbb{F}_2^m \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0 \ \ \forall\, \mathbf{u} \in S\}$. Now,

$$(H \cdot \mathbf{u}_S)_i = \sum_{j=1}^{n} H_{ij} (\mathbf{u}_S)_j = \sum_{\substack{j=1 \\ \mathbf{c}_j \in S}}^{n} H_{ij} = \sum_{\substack{j=1 \\ \mathbf{c}_j \in S}}^{n} (-1)^{\langle \mathbf{r}_i, \mathbf{c}_j \rangle}$$

Case 1: For rows indexed by vectors in T, we have $\langle \mathbf{r}_i, \mathbf{c}_j \rangle = 0\ \forall \mathbf{c}_j \in S$. Thus, $(H \cdot \mathbf{u}_S)_i = \sum_{\substack{j=1 \\ \mathbf{c}_j \in S}}^{n} 1 = |S|$.

Case 2: For rows **not** indexed by vectors in T. Therefore, there exists $\overline{\mathbf{c}_j} \in S$ such that $\langle \mathbf{r}_i, \overline{\mathbf{c}_j} \rangle = 1$. Consider any vector $\mathbf{c}_j \in S$, then $\langle \mathbf{r}_i, \mathbf{c}_j + \overline{\mathbf{c}_j} \rangle = \langle \mathbf{r}_i, \mathbf{c}_j \rangle + \langle \mathbf{r}_i, \overline{\mathbf{c}_j} \rangle = \langle \mathbf{r}_i, \mathbf{c}_j \rangle + 1$. Thus, $(-1)^{\langle \mathbf{r}_i, \mathbf{c}_j \rangle}$ and $(-1)^{\langle \mathbf{r}_i, \mathbf{c}_j + \overline{\mathbf{c}_j} \rangle}$ have different signs. Hence $(H \cdot \mathbf{u}_S)_i = 0$ in this case.

Thus, we have $\frac{1}{|S|}(H \cdot \mathbf{u}_S) \in \{0, 1\}^n$ with ones in rows indexed by elements of T, and zeros otherwise. Now, we are just left to show that unique subspaces generate

unique elements in $\{0,1\}^n$.

Notice that $\frac{1}{|S|}(H \cdot \mathbf{u}_S)$ has ones only in rows indexed by T. But by definition, T is orthogonal complement of S. It can be easily shown that T is also a subspace. By Theorem A.6 we see that unique subspaces have unique orthogonal complements. Hence each subspace S generates a unique T, and hence a unique vector $H \cdot \mathbf{u}_S$ $\square$

Remark 2.24. In order to obtain a polytope with non-negative coordinates, just translate the whole polytope in $\mathbb{R}^n$ by 1.

Claim 2.23 gives us a polytope with a "small" vertex set but a "large" number of integral points in it. In the next chapter we will look at the implications of this result, as well as build towards obtaining better sparsity bounds for factors of certain classes of polynomials.

Chapter 3

Sparsity bounds for factors of certain classes of polynomials

As we saw in chapter 2, we have a quasipolynomial bound on sparsity of factors of a polynomial of bounded individual degree, obtained by analyzing the convex polytope representation of a polynomial. The result stated that the number of integral points in any polytope $P \subseteq \{0, 1, ..., d\}^n$ for some constant d is upper bounded by $|U_P|^{\mathcal{O}(d^2 log\, n)}$, where U_P is the vertex set of the polytope. At the same time, Claim 2.23 gives us a polytope, called the *Hadamard polytope*, for which this bound is tight. In this chapter, we will analyse how the existence of *Hadamard Polytope* has the potential for negatively impacting our quest to prove Conjecture 1.1. However, in our first result (Corollary 3.6) we will show that in fact, it doesn't act as an impediment towards proving Conjecture 1.1. We will then generalise these proof techniques to prove polynomial sparsity bounds for certain classes of polynomials in Theorem 3.8, and Theorem 3.12. Lastly, we will look at the limitations of our proof techniques by describing some counter examples, and rule out certain approaches towards attaining polynomial sparsity bounds.

3.1 Limitations of the polytope approach

Firstly, we define the polynomial corresponding to the Hadamard polytope.

Definition 3.1. (Hadamard polynomial)
Let g_H be the polynomial with vertex set as the columns of H, where H is the *Hadamard matrix* as per Claim 2.23. Every internal point in g_H corresponds to some subspace of $\mathbb{F}_2^m$. Formally,

$$V_g = \{\mathbf{v} \mid \mathbf{v} \textit{ is a column vector of } H\}$$

For every subspace $S \subseteq \mathbb{F}_2^m$, let $\mathbf{v}_S$ denote the vector generated by taking the uniform convex span of those columns of H that were indexed by elements in S.

$$E_g = \{\mathbf{v}_S \mid S \textit{ is subspace of } \mathbb{F}_2^m\}$$

Now we define

$$supp(g_H) = V_g \cup E_g$$

Then g_H is the *Hadamard polytope.*

Remark 3.2. g_H is apolynomial with $|V_g| = n$ and $\|g_H\| = |supp(g_H)| = n + n^{\Omega(log\ n)} \approx |V_g|^{\Omega(log\ n)}$. Also by Theorem 2.21, $\|g_H\| = |V_g|^{\mathcal{O}(d^2\ log\ n)}$. Therefore when d is some constant, we get $\|g_H\| = |V_g|^{\Theta(log\ n)}$.

Therefore, a possibility arises that there exists some polynomial $h \in \mathbb{F}[\mathbf{x}]$, such that $f = g_H \cdot h$ gives us f with $\|f\| = \mathcal{O}(n)$, in accordance with Corollary 2.18. However such a f would have g_H as a factor with $\|g_H\| = n^{\Theta(log\ n)}$. This would imply that the sparsity bound obtained in Corollary 2.22 is tight for some polynomials, and would *refute* the polynomial sparsity conjecture mentioned in [Vol17]. Thus, a natural first step is to either find such a polynomial h which gives f satisfying the above conditions, or show that the Hadamard polynomial g_H, with superpolynomial sparsity in n, can't be a factor of any polynomial f with polynomial sparsity.

In the following sections we will show that the product of the Hadamard polynomial g_H with any polynomial h where $\|h\|$ is not too "large", then the resulting polynomial $f = g_h \cdot h$ has $\|f\| \geq \|g\|$. In other words, product of g_H with any "sparse" polynomial h, gives us a "dense" f. Before proceeding, we highlight some interesting properties of the Hadamard polytope which will be used in the proofs.

Lemma 3.3 (Orthogonality of rows, columns of Hadamard matrix)**.** *Any two rows of the Hadamard matrix are orthogonal. Similarly, any two columns of Hadamard matrix are orthogonal.*

Proof. : Consider any two rows $\mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^n$ indexed by $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{F}_2^m$ respectively in H where $\mathbf{y}_1 \neq \mathbf{y}_2$. Let columns of H be indexed by $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n \in \mathbb{F}_2^m$. Let $(\mathbf{r}_1)_i$ denote i^{th} elt. of $\mathbf{r}_1$. Then, $(\mathbf{r}_1)_i = (-1)^{\langle \mathbf{y}_1, \mathbf{z}_i \rangle}$.
$\langle \mathbf{r}_1, \mathbf{r}_2 \rangle = \sum_{i=1}^n (-1)^{\langle \mathbf{y}_1, \mathbf{z}_i \rangle} \cdot (-1)^{\langle \mathbf{y}_2, \mathbf{z}_i \rangle} = \sum_{i=1}^n (-1)^{\langle \mathbf{y}_1 + \mathbf{y}_2, \mathbf{z}_i \rangle}$
$= \sum_{i=1}^n (-1)^{\langle \mathbf{y}_3, \mathbf{z}_i \rangle}$ $(\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{F}_2^m \Rightarrow \mathbf{y}_3 = \mathbf{y}_1 + \mathbf{y}_2 \in \mathbb{F}_2^m)$
$= \sum_{\mathbf{z} \in \mathbb{F}_2^m} (-1)^{\langle \mathbf{y}_3, \mathbf{z} \rangle} = 0$ $(by\ \ Claim\ A.7, \text{ if } \mathbf{y}_3 \neq \mathbf{0})$.
Since $2 \cdot \mathbf{y} = 0$ when $\mathbf{y} \in \mathbb{F}_2^m$, we have $\mathbf{y}_3 = 0 \Rightarrow \mathbf{y}_1 + \mathbf{y}_2 = 0 \Rightarrow \mathbf{y}_1 = \mathbf{y}_2$
As we had distinct rows, $\mathbf{y}_1 \neq \mathbf{y}_2$. Hence, $\langle \mathbf{r}_1, \mathbf{r}_2 \rangle = 0$.
Thus, any two rows in H are mutually orthogonal. Symmetrically, any two **columns are orthogonal** in H. □

Corollary 3.4 (Linear Independence of rows, columns in Hadamard matrix)**.** *Rows in H are linearly independent. Similarly, columns in H are linearly independent.*

Proof. : Let the rows of H be $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n$. Let $c_i \in \mathbb{R}$ be scalars. Then $\sum_{i=1}^n c_i \cdot \mathbf{r}_i = \mathbf{0}$
$\Rightarrow \sum_{i=1}^n c_i \cdot \langle \mathbf{r}_i, \mathbf{r}_j \rangle = \langle \mathbf{0}, \mathbf{r}_j \rangle \Rightarrow c_j \cdot \langle \mathbf{r}_j, \mathbf{r}_j \rangle = 0 \Rightarrow c_j = 0$.
Since j was arbitrary, it follows that $c_j = 0\ \forall j \in [n]$. □

We now prove an important lemma which helps us get the sparsity bounds for the product of g_H with any "sparse" polynomial h.

Lemma 3.5. *Let g_H be the Hadamard polynomial. Let $h \in \mathbb{F}[\mathbf{x}]$ with $\mathbf{h}_1, \mathbf{h}_2$ as two distinct vectors in supp(h). Let $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} \in$ supp(g_H) such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$. Then, there exists **at most** one other distinct pair $\mathbf{g}_{i_3}, \mathbf{g}_{i_4} \in$ supp(g_H) such that $\mathbf{h}_1 + \mathbf{g}_{i_3} = \mathbf{h}_2 + \mathbf{g}_{i_4}$.*

Proof Idea. We use a few important properties of the Hadamard polynomial g_H in our proof. Crucially, we use the linear independence of vertices, as shown above in Corollary 3.4. It helps us to get constraints on the coefficients of each vertex, which are subsequently used to rule out possibilities of more than two pairs of vectors in *supp*(g_H) cancelling. Along with this, we use the fact that all internal points correspond to some subspace of $\mathbb{F}_2^m$, and that each such point is generated by a *uniform convex span* of the vertices of g_H. Additionally, we use the fact that the $\mathbf{0}$ is present in every subspace S of $\mathbb{F}_2^m$. Using these features we are able to show that for any two monomials in h, there can be *atmost four* cancellations in the product $f = g_H \cdot h$.

Proof. Let *supp*$(g_H) = \{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_T\}$. Let $V_g = V(CS(supp(g_H))) = \{\mathbf{g}_1, ..., \mathbf{g}_n\}$ and $E_g = supp(g) \setminus V_g$. Without loss of generality, we assume that the first column of H is indexed by the $\mathbf{0} \in \mathbb{F}_2^m$, and hence the vector corresponding to it (*i.e.* $\mathbf{g}_1$) is present in every internal point in E_g.
Let us assume that there exists two pairs of vectors in *supp*(g_H) such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$ and $\mathbf{h}_1 + \mathbf{g}_{i_3} = \mathbf{h}_2 + \mathbf{g}_{i_4}$. Trivially, $\mathbf{g}_{i_1} \neq \mathbf{g}_{i_2}$ and $\mathbf{g}_{i_3} \neq \mathbf{g}_{i_4}$ as otherwise $\mathbf{h}_1 = \mathbf{h}_2$. On rearranging we get,

$$\mathbf{h}_2 - \mathbf{h}_1 = \mathbf{g}_{i_1} - \mathbf{g}_{i_2} = \mathbf{g}_{i_3} - \mathbf{g}_{i_4} \tag{3.1}$$

We can see that, $\mathbf{g}_{i_1} = \mathbf{g}_{i_3} \iff \mathbf{g}_{i_2} = \mathbf{g}_{i_4}$ giving us the same pair. Thus $\mathbf{g}_{i_1} \neq \mathbf{g}_{i_3}$, and $\mathbf{g}_{i_2} \neq \mathbf{g}_{i_4}$. Manipulating (3.1) we get,

$$\mathbf{g}_{i_1} - \mathbf{g}_{i_2} - \mathbf{g}_{i_3} + \mathbf{g}_{i_4} = 0 \tag{3.2}$$

Recall that $\mathbf{g}_1, ..., \mathbf{g}_n$ are vertices in g_H and any internal point is convex combination of these vertices. Therefore, for all $k \in [4]$, $\mathbf{g}_{i_k} = \sum_{l=1}^{n} a_{l_k} \cdot \mathbf{g}_l$ where $a_{l_k} \geq 0 \ \forall l$ and $\sum_{l=1}^{n} a_{l_k} = 1$. Rearranging and using independence of vertices in (3.2), we get:

$$a_{l_1} - a_{l_2} = a_{l_3} - a_{l_4} \ \forall l \in [n] \tag{3.3}$$

In general, if $\mathbf{g}_{i_k}$ is a vertex then, $\mathbf{g}_{i_k} = \mathbf{g}_r$ for some $r \in [n]$. Then,

$$a_{l_k} = \begin{cases} 0 \ , \ l \neq r \\ 1 \ , \ l = r \ . \end{cases} \tag{3.4}$$

Otherwise, if $\mathbf{g}_{i_k}$ is an internal point in $supp(g_H)$ generated by t_i-dimensional subspace S_i of $\mathbb{F}_2^m$ for any $1 \leq t_i \leq m$. Then,

$$a_{l_k} = \begin{cases} 0 \ , \text{ if } \mathbf{g}_l \text{ doesn't contribute in generating } \mathbf{g}_{i_k} \\ \frac{1}{2^{t_i}} \ , \ \mathbf{g}_l \text{ contributes in generating } \mathbf{g}_{i_k}. \end{cases} \tag{3.5}$$

Case 1: $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \mathbf{g}_{i_3}, \mathbf{g}_{i_4} \in V_g$.

Using (3.2) and (3.4), this means that their linear combination sums to 0, implying that they are linearly dependent. But, by Corollary 3.4 all vertices in g are linearly independent. Hence we have a $\Rightarrow\Leftarrow$ to our assumption. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

Case 2: Three of $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \mathbf{g}_{i_3}, \mathbf{g}_{i_4}$ are vertices in g_H and the other is an internal point. Wlog, let $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \mathbf{g}_{i_3} \in V_g$, and $\mathbf{g}_{i_4} \in E_g$.

Using (3.3) for $l = 1$, coefficient of LHS (denoted c_{LHS}) $\in \{-1, 0, 1\}$, as either none of $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} = \mathbf{g}_1$, or one of them equals $\mathbf{g}_1$. Meanwhile, $c_{RHS} \in \{\frac{-1}{2^t}, 1-\frac{1}{2^t}\}$, as $\mathbf{g}_{i_3}$ may or may not be $\mathbf{g}_1$. Thus, $c_{LHS} \neq c_{RHS}$ in any case. Hence, we have a $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

Case 3: Two of $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \mathbf{g}_{i_3}, \mathbf{g}_{i_4}$ are vertices in g_H and the other two are internal points.

Subcase 1: Wlog, let $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} \in V_g$ and $\mathbf{g}_{i_3}, \mathbf{g}_{i_4} \in E_g$.

Subsubcase 1: $\mathbf{g}_{i_1} = \mathbf{g}_1$. This gives us $\mathbf{g}_{i_2} \neq \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = 1$ and $c_{RHS} = \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}} < 1 = c_{LHS}$ always. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

The case for $\mathbf{g}_{i_2} = \mathbf{g}_1$ and consequently $\mathbf{g}_{i_1} \neq \mathbf{g}_1$ is analogus.

Subsubcase 2: $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} \neq \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = 0$ and $c_{RHS} = \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}}$. Thus, only possibility is $t_3 = t_4$. Suppose now, that $\mathbf{g}_{i_1} = \mathbf{g}_z$ for some $z \in [n]$. Using (3.3) for $l = z$, $c_{LHS} = 1$ and $c_{RHS} \in \{0, \frac{1}{2^{t_3}}, \frac{-1}{2^{t_4}}\} \neq c_{LHS}$ always. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

The case for $\mathbf{g}_{i_3}, \mathbf{g}_{i_4}$ as vertices, and $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ as internal points is analogus.

Subcase 2: Wlog, let $\mathbf{g}_{i_1}, \mathbf{g}_{i_3} \in V_g$ and $\mathbf{g}_{i_2}, \mathbf{g}_{i_4} \in E_g$.

Subsubcase 1: $\mathbf{g}_{i_1} = \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = 1 - \frac{1}{2^{t_2}}$ and $c_{RHS} = \frac{-1}{2^{t_4}} \neq c_{LHS}$ always. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

The case for $\mathbf{g}_{i_3} = \mathbf{g}_1$ is analogus.

Subsubcase 2: $\mathbf{g}_{i_1}, \mathbf{g}_{i_3} \neq \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = \frac{-1}{2^{t_2}}$ and $c_{RHS} = \frac{-1}{2^{t_4}}$. Thus, only possibility is $t_2 = t_4$. Suppose now, that $\mathbf{g}_{i_1} = \mathbf{g}_z$ for some $z \in [n]$. Using (3.3) for $l = z$, $c_{LHS} \in \{1, 1 - \frac{1}{2^{t_2}}\}$ and $c_{RHS} \in \{0, \frac{-1}{2^{t_4}}\} \neq c_{LHS}$ always. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

The case for $\mathbf{g}_{i_2}, \mathbf{g}_{i_4}$ as vertices, and $\mathbf{g}_{i_1}, \mathbf{g}_{i_3}$ as internal points is analogus.

Subcase 3: Wlog, let $\mathbf{g}_{i_1}, \mathbf{g}_{i_4} \in V_g$ and $\mathbf{g}_{i_2}, \mathbf{g}_{i_3} \in E_g$.

Subsubcase 1: $\mathbf{g}_{i_1} = \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = 1 - \frac{1}{2^{t_2}}$ and $c_{RHS} \in \{\frac{1}{2^{t_3}}, \frac{1}{2^{t_3}} - 1\}$. Only possibility of $c_{LHS} = c_{RHS}$ is $1 - \frac{1}{2^{t_2}} = \frac{1}{2^{t_3}} \Rightarrow t_2 = t_3 = 1$. Let, $\mathbf{g}_{i_4} = \mathbf{g}_z$ for some $z \in [n]$. Using (3.3) for $l = z$, $c_{LHS} \in \{0, \frac{1}{2}\}$ and $c_{RHS} \in \{-1, \frac{-1}{2}\}$. Thus, we can get equality when $c_{LHS} = c_{RHS} = \frac{-1}{2}$. Thus we can see that the condition in (3.1) is satisfied by

$$\mathbf{g}_{i_1} = \mathbf{g}_1; \;\; \mathbf{g}_{i_4} = \mathbf{g}_z; \;\; \mathbf{g}_{i_2} = \mathbf{g}_{i_3} = \frac{\mathbf{g}_1 + \mathbf{g}_z}{2}$$

The case for $\mathbf{g}_{i_4} = \mathbf{g}_1$ is analogus.

Subsubcase 2: $\mathbf{g}_{i_1}, \mathbf{g}_{i_4} \neq \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = \frac{-1}{2^{t_2}}$ and $c_{RHS} = \frac{1}{2^{t_3}}$. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

The case for $\mathbf{g}_{i_2}, \mathbf{g}_{i_4}$ as vertices, and $\mathbf{g}_{i_1}, \mathbf{g}_{i_3}$ as internal points is analogus.

Case 4: One of $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \mathbf{g}_{i_3}, \mathbf{g}_{i_4}$ is a vertex in g_H and the other three are internal points. Wlog, let $\mathbf{g}_{i_1} \in V_g$, and $\mathbf{g}_{i_2}, \mathbf{g}_{i_3}, \mathbf{g}_{i_4} \in E_g$.

Subcase 1: $\mathbf{g}_{i_1} = \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = 1 - \frac{1}{2^{t_2}}$ and $c_{RHS} = \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}} \neq c_{LHS}$ always. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

Subcase 2: $\mathbf{g}_{i_1} \neq \mathbf{g}_1$.

Using (3.3) for $l = 1$, $c_{LHS} = \frac{-1}{2^{t_2}}$ and $c_{RHS} = \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}}$. Thus, only possibility is $t_2 = t_3$ and $t_4 = t_2 - 1$. We know that $t_4 > 1$ as otherwise $\mathbf{g}_{i_4}$ will become a vertex. This implies that $t_2 = t_3 \geq 2$. Suppose now, that $\mathbf{g}_{i_1} = \mathbf{g}_z$ for some $z \in [n]$. Using (3.3) for $l = z$, $c_{LHS} \in \{1, 1 - \frac{1}{2^{t_2}}\}$ and $c_{RHS} \in \{\frac{1}{2^{t_2}} - \frac{1}{2^{t_2-1}}, \frac{1}{2^{t_2}}, \frac{-1}{2^{t_2-1}}\} = \{\frac{-1}{2^{t_2}}, \frac{1}{2^{t_2}}, \frac{-1}{2^{t_2-1}}\} \neq c_{LHS}$ at any time as $t_2 \geq 2$. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

The case for $\mathbf{g}_{i_2}$ or $\mathbf{g}_{i_3}$ or $\mathbf{g}_{i_4}$ as vertices is analogus.

Case 5: None of $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \mathbf{g}_{i_3}, \mathbf{g}_{i_4}$ are vertices in g_H.

Using (3.3) for $l = 1$, $c_{LHS} = \frac{1}{2^{t_1}} - \frac{1}{2^{t_2}}$ and $c_{RHS} = \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}}$. Hence

$$c_{LHS} = c_{RHS} \Rightarrow \frac{1}{2^{t_1}} + \frac{1}{2^{t_4}} = \frac{1}{2^{t_2}} + \frac{1}{2^{t_3}} \tag{3.6}$$

Subcase 1: $t_1 = t_4$. Putting in (3.6), we get $\frac{1}{2^{t_1-1}} = \frac{1}{2^{t_2}} + \frac{1}{2^{t_3}} \Rightarrow t_2 = t_3$ and $t_1 - 1 = t_2 - 1 \Rightarrow t_1 = t_2 = t_3 = t_4 = t$.

Recall, that S_k represent the subspace from which the internal point g_{i_k} is generated.

Claim 1: $S_1 \setminus S_2 = S_3 \setminus S_4$

Proof: Let $\mathbf{x} \in S_1 \setminus S_2$ and $z \in [n]$ be the index corresponding to $\mathbf{x}$. Using (3.3) and (3.5), for $l = z$, $c_{LHS} = \frac{1}{2^t}$. Thus, to get $c_{RHS} = \frac{1}{2^t} = c_{LHS}$, we need $a_{z_3} = \frac{1}{2^t}$ and $a_{z_4} = 0 \Rightarrow \mathbf{x} \in S_3 \setminus S_4 \Rightarrow S_1 \setminus S_2 \subseteq S_3 \setminus S_4$. Similarly, $S_3 \setminus S_4 \subseteq S_1 \setminus S_2$. Thus, we have $S_1 \setminus S_2 = S_3 \setminus S_4$.

Claim 2: $S_2 \setminus S_1 = S_4 \setminus S_3$

Proof: Similar to Claim 1.

Claim 3: $S_1 = S_3$ and $S_2 = S_4$

Proof: Let $S_1 = \{\mathbf{x}_1, .., \mathbf{x}_p, \mathbf{y}_1, ..., \mathbf{y}_q\}$ and $S_2 = \{\mathbf{z}_1, .., \mathbf{z}_p, \mathbf{y}_1, ..., \mathbf{y}_q\}$. We know that $|S_1| = |S_2| = |S_3| = |S_4|$ as they are $t-dim$ subspaces. Also, by **Claim 1** and **Claim 2**, $S_3 = \{\mathbf{x}_1, .., \mathbf{x}_p, \mathbf{w}_1, ..., \mathbf{w}_q\}$, $S_4 = \{\mathbf{z}_1, .., \mathbf{z}_p, \mathbf{u}_1, ..., \mathbf{u}_q\}$.

Consider $\mathbf{e} = \mathbf{y}_1 + \mathbf{x}_1$. If $\mathbf{e} \in S_2 \Rightarrow \mathbf{x}_1 \in S_2 - \mathbf{y}_1 \Rightarrow \mathbf{x}_1 \in S_2 \Rightarrow\Leftarrow$. Hence $\mathbf{e} \notin S_2$. This means that $\mathbf{e} \in S_1 \setminus S_2 \Rightarrow \mathbf{e} \in S_3 \setminus S_4 \Rightarrow \mathbf{e} \in S_3 \Rightarrow \mathbf{y}_1 \in S_3 - \mathbf{x}_1 \Rightarrow \mathbf{y}_1 \in S_3$. Since $\mathbf{y}_1$ is distinct from $\mathbf{x}_j$ for all j, it means that $\mathbf{y}_1 = \mathbf{w}_c$ for some $c \in [q]$. Similarly, taking $\mathbf{f} = \mathbf{y}_1 + \mathbf{z}_1$, we can show that $\mathbf{y}_1 \in S_4$. We can extend the same argument to all $\mathbf{y}_b$ $\forall b \in [q]$, to get $S_1 = S_3$ and $S_2 = S_4$.

Thus, by **Claim 3**, we get $\mathbf{g}_{i_1} = \mathbf{g}_{i_3}$ and $\mathbf{g}_{i_2} = \mathbf{g}_{i_4}$. Hence, there is a unique pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} \in supp(g_H)$ such that $\mathbf{g}_{i_1} + \mathbf{h}_1 = \mathbf{g}_{i_2} + \mathbf{h}_2$.

Subcase 2: $t_1 < t_4$. Putting in (3.6), we get that LHS is a binary representation, hence must be unique. Thus, either ($t_1 = t_2$ and $t_3 = t_4$) or ($t_1 = t_3$ and $t_2 = t_4$).

Subsubcase 1: $t_1 = t_2 < t_3 = t_4$.

We know that $\mathbf{g}_{i_1} \neq \mathbf{g}_{i_2}$ implying the existence of $\mathbf{g}_z$ such that $\mathbf{g}_z \in g\_i_1$ & $\mathbf{g}_z \notin \mathbf{g}_{i_2}$. Thus, using (3.3) for $l = z$, $c_{LHS} = \frac{1}{2^{t_1}}$ and $c_{RHS} \in \{0, \frac{1}{2^{t_3}}, -\frac{1}{2^{t_4}}\} \neq c_{LHS}$ always as $t_1 < t_3$. Thus, $\Rightarrow\Leftarrow$. Hence, there exists at most one pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ such that $\mathbf{h}_1 + \mathbf{g}_{i_1} = \mathbf{h}_2 + \mathbf{g}_{i_2}$.

Subsubcase 2: $t_1 = t_3 < t_2 = t_4$.

Consider some vertex $\mathbf{g}_z$ in $S_1 \setminus S_2$. Then, $c_{LHS} = \frac{1}{2^{t_1}}$ and $c_{RHS} \in \{0, \frac{1}{2^{t_3}}, \frac{-1}{2^{t_4}}, \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}}\}$. The only possibility is $g_z \in S_3 \setminus S_4$. Thus, $S_1 \setminus S_2 \subseteq S_3 \setminus S_4$. Similarly, $S_3 \setminus S_4 \subseteq S_1 \setminus S_2$. Thus, $S_1 \setminus S_2 = S_3 \setminus S_4$. Using a similar argument, $S_2 \setminus S_1 = S_4 \setminus S_3$. Consider $\mathbf{g}_z \in S_1 \cap S_2 \Rightarrow c_{LHS} = \frac{1}{2^{t_1}} - \frac{1}{2^{t_2}}$ and $c_{RHS} \in \{0, \frac{1}{2^{t_3}}, \frac{-1}{2^{t_4}}, \frac{1}{2^{t_3}} - \frac{1}{2^{t_4}}\}$. The only possibility is $\mathbf{g}_z \in S_3 \cap S_4$. Hence, $S_1 \cap S_2 \subseteq S_3 \cap S_4$. Similarly, $S_3 \cap S_4 \subseteq S_1 \cap S_2$. Finally, we have $S_1 \cap S_2 = S_3 \cap S_4$, along with $S_1 \setminus S_2 = S_3 \setminus S_4$ and $S_2 \setminus S_1 = S_4 \setminus S_3$, which gives us $S_1 = S_3$ and $S_2 = S_4$. Thus, we get $\mathbf{g}_{i_1} = \mathbf{g}_{i_3}$ and $\mathbf{g}_{i_2} = \mathbf{g}_{i_4}$. Hence, there is a unique pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} \in supp(g_H)$ such that $\mathbf{g}_{i_1} + \mathbf{h}_1 = \mathbf{g}_{i_2} + \mathbf{h}_2$.

The case of $t_1 > t_4$ is analogus

Thus, we see that in most cases there is atmost one unique pair $\mathbf{g}_{i_1}, \mathbf{g}_{i_2} \in supp(g_H)$ such that $\mathbf{g}_{i_1} + \mathbf{h}_1 = \mathbf{g}_{i_2} + \mathbf{h}_2$, and only in the very specific case of $\mathbf{g}_{i_1} = \mathbf{g}_1$, $\mathbf{g}_{i_4} = \mathbf{g}_z$, $\mathbf{g}_{i_2} = \mathbf{g}_{i_3} = \frac{\mathbf{g}_1 + \mathbf{g}_z}{2}$ (and it's analogus cases), do we get two pairs $\mathbf{g}_{i_1}, \mathbf{g}_{i_2}$ and $\mathbf{g}_{i_3}, \mathbf{g}_{i_4} \in supp(g_H)$ such that $\mathbf{h}_1 - \mathbf{h}_2 = \mathbf{g}_{i_2} - \mathbf{g}_{i_1} = \mathbf{g}_{i_4} - \mathbf{g}_{i_3}$. □

Corollary 3.6 (Sparsity of f where $f = g_H \cdot h$)**.** *Let $f = g_H \cdot h$ where g_H is the Hadamard polynomial, and h is some polynomial in $\mathbb{F}[\mathbf{x}]$ such that $\|h\|$ is not too "large" with respect to $\|g_H\|$. Then, f is denser than g_H.*

$$\|h\| \leq \frac{1}{2}\|g_H\| \Rightarrow \|f\| \geq \|g_H\|$$

Proof. By Lemma 3.5, there are at most **two** pairs of monomials in $supp(g_H)$ that have the same vector difference, for any two monomials $h_1, h_2 \in supp(h) \Rightarrow$ there are atmost **four** cancellations in $g_H \cdot (h_1 + h_2)$. Now, number of ways to choose two monomials in h is $\binom{\|h\|}{2}$. Thus

$$\|f\| = \|g_H \cdot h\| \geq \|g_H\| \cdot \|h\| - 4 \cdot \binom{\|h\|}{2}$$

Also,

$$\|g_H\| \cdot \|h\| - 4 \cdot \binom{\|h\|}{2} \geq \|g_H\| \Rightarrow 2 \cdot \|h\|^2 - (\|g_H\| + 2) \cdot \|h\| + \|g_H\| \leq 0$$

The above holds for for all h, such that $\|h\| \leq \frac{1}{2}\|g_H\|$. Thus, for any such h, we have $\|f\| \geq \|g_H\|$ where $f = g_H \cdot h$. □

Remark 3.7. An approach like this won't work for $\|h\| > \frac{1}{2}\|g_H\|$, as we might potentially have many cancellations. One way to mitigate this could be to look at, say $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3 \in supp(h)$ and say that if $\mathbf{h}_1, \mathbf{h}_2$ have some cancellations, and $\mathbf{h}_2, \mathbf{h}_3$ have some cancellations, then there can be no cancellations between $\mathbf{h}_1, \mathbf{h}_3$. However, even this wouldn't help us in bounding the maximum number of cancellations. Consider elements of $supp(h)$ as nodes of a graph G such that there is an edge between two nodes $\mathbf{h}_a, \mathbf{h}_b$ if there is some cancellation between $\mathbf{h}_a, \mathbf{h}_b$ in the product $f = g_H \cdot h$. Then, the condition mentioned above means we are looking at the edges in a triangle free graph. By Theorem A.9, the maximum number of edges in any such graph is $\mathcal{O}(|supp(h)|^2)$. Hence even on limiting the number of cancellation per edge to two by Lemma 3.5, we can still have many cancellations in f.

Thus, we can see that the polynomial corresponding to the Hadamard polytope doesn't prove to be an impediment towards proving Conjecture 1.1, atleast when considering it's product with "sparse" polynomials.

3.2 Few results on sparsity bounds

We try to generalise the proof techniques used in proving Lemma 3.5 to arrive at sparsity bounds for factors of certain classes of polynomials.

Theorem 3.8. *Let $f, g, h \in \mathbb{F}[\mathbf{x}]$ s.t. $V_g = V(CS(supp(g)))$, $V_h = V(CS(supp(h)))$. Suppose, $V_g \cup V_h$ is a linearly independent set of vectors. Then, $\|f\| = \|g\|\|h\|$.*

Proof. Let $\mathbf{g}_1, \mathbf{g}_2 \in supp(g)$ and $\mathbf{h}_1, \mathbf{h}_2 \in supp(h)$ such that

$$\mathbf{g}_1 + \mathbf{h}_1 = \mathbf{g}_2 + \mathbf{h}_2 \Rightarrow \mathbf{g}_1 - \mathbf{g}_2 + \mathbf{h}_1 - \mathbf{h}_2 = 0$$

But $\mathbf{g}_1 = \sum_{v \in V_g} \lambda_v \mathbf{g}_v$, $\mathbf{g}_2 = \sum_{v \in V_g} \beta_v \mathbf{g}_v$.
Similarly, $\mathbf{h}_1 = \sum_{u \in V_h} \gamma_u \mathbf{h}_u$, $\mathbf{h}_2 = \sum_{u \in V_h} \alpha_u \mathbf{h}_u$.
Rearranging, we get

$$\sum_{v \in V_g} (\lambda_v - \beta_v)\mathbf{g}_v + \sum_{u \in V_g} (\gamma_u - \alpha_u)\mathbf{h}_u = 0$$

Using linear independence of $V_g \cup V_h$, we get

$$\lambda_v = \beta_v \; \forall v \in V_g \Rightarrow \mathbf{g}_1 = \mathbf{g}_2$$

$$\gamma_u = \alpha_u \; \forall u \in V_h \Rightarrow \mathbf{h}_1 = \mathbf{h}_2$$

Thus, product of every momomial $m_g \in g$ ad $m_h \in h$ is unique. Hence

$$\|f\| = \|g \cdot h\| = \|g\|\|h\|$$

□

However, the combined linear independence of the vertices of both the polynomials puts restrictions on both factors, which might not always be possible. Therefore, we will now look at the case of square polynomials (*i.e.*, $f = g^2$) where we only need restrictions on one polynomial g.

Lemma 3.9. *Let $f = {g_H}^2$ where g_H is the Hadamard polynomial, as per Definition 3.1. Then, $\|f\| = \Omega(\|g_H\|)$*

Proof. Let $supp(g_H) = \{\mathbf{g}_1.\mathbf{g}_2, ..., \mathbf{g}_t\}$ and $P_g = CS(supp(g_H))$. Let $V_g = V(P_g) = \{\mathbf{g}_1, ...\mathbf{g}_n\}$. For all $i \in [t]$, consider the monomial $\mathbf{g}_i^2 \in f = {g_H}^2$ which corresponds to the vector $2 \cdot \mathbf{g}_i$ in the Minkowski sum $P_f = P_g + P_g$. We claim that $2 \cdot \mathbf{g}_i$ is unique for "almost" all points $\mathbf{g}_i$ in g_H.
Suppose $\exists \mathbf{g}_j \neq \mathbf{g}_k \in P_g$ such that $2 \cdot \mathbf{g}_i = \mathbf{g}_j + \mathbf{g}_k$. If $\mathbf{g}_j = \mathbf{g}_k \Rightarrow \mathbf{g}_j = \mathbf{g}_k = \mathbf{g}_i$ and we get same points.

Case 1: $\mathbf{g}_i \in V_g$.

We know that any point in P_g can be written as convex combination of points in V_g. Thus, $\mathbf{g}_j = \sum_{v=1}^{n} a_v \mathbf{g}_v$, $\mathbf{g}_k = \sum_{v=1}^{n} b_v \mathbf{g}_v$ where $a_v, b_v \geq 0$ and $\sum_{v=1}^{n} a_v = \sum_{v=1}^{n} b_v = 1$.

$$2 \cdot \mathbf{g}_i = \mathbf{g}_j + \mathbf{g}_k \Rightarrow \sum_{\substack{v=1 \\ v \neq i}}^{n} (a_v + b_v)\mathbf{g}_v + (a_i + b_i - 2)\mathbf{g}_i = 0$$

This implies that $\mathbf{g}_i$ for $i \in [n]$ are linearly dependent which contradicts Corollary 3.4. Thus, $2 \cdot \mathbf{g}_i$ is unique for all vertices of g_H.

Case 2: $\mathbf{g}_i \in supp(g_H) \backslash V_g$. This means that $\mathbf{g}_i$ is generated by some $t-dimensional$ subspace of $\mathbb{F}_2^m$. Let $\mathbf{g}_1$ denote the column vector indexed by the $\mathbf{0} \in \mathbb{F}_2^m$.

Subcase 1: $\mathbf{g}_j, \mathbf{g}_k \in V_g$.

Using (3.3), (3.5) from the proof of Lemma 3.5 for $l = 1$, we get $c_{LHS} = 2 \cdot \frac{1}{2^t}$ and $c_{RHS} \in \{0, 1\}$. Thus, the only possibility is that one $c_{RHS} = c_{LHS} = 1 \Rightarrow t = 1$. Also, this means one of $\mathbf{g}_j, \mathbf{g}_k$ equals $\mathbf{g}_1$. Without loss of generality, let it be $\mathbf{g}_k$. Using Equation 3.3 for $l = j$, $c_{RHS} = 1$, and $c_{LHS} \in \{0, 2 \cdot \frac{1}{2}\}$. Thus, we get equality when $\mathbf{g}_i = \frac{\mathbf{g}_1 + \mathbf{g}_j}{2}$. Thus, when

$$\mathbf{g}_i = \frac{\mathbf{g}_1 + \mathbf{g}_j}{2}; \;\; \mathbf{g}_k = \mathbf{g}_1 \in V_g; \;\; \mathbf{g}_j \in V_g$$

then $2 \cdot \mathbf{g}_i$ is not unique, but this happens for only $n - 1$ choices of $\mathbf{g}_j \in V_g \setminus \{\mathbf{g}_1\}$.

Subcase 2: $\mathbf{g}_j \in V_g, \; \mathbf{g}_k \in supp(g_H) \setminus V_g$.

Subsubcase 1: $\mathbf{g}_j = \mathbf{g}_1$

Using (3.3) for $l = 1$, we get $c_{LHS} = 2 \cdot \frac{1}{2^t}$ and $c_{RHS} = 1 + \frac{1}{2^{t_k}}$. Thus, $c_{LHS} \neq c_{RHS}$ giving us a $\Rightarrow\!\!\!\!\!\!\!\!\Leftarrow$. Thus, $2 \cdot \mathbf{g}_i$ is unique in this case.

Subsubcase 2: $\mathbf{g}_j \neq \mathbf{g}_1$

Using (3.3) for $l = 1$, we get $c_{LHS} = 2 \cdot \frac{1}{2^t}$ and $c_{RHS} = \frac{1}{2^{t_k}}$. Thus, $c_{LHS} = c_{RHS} \Rightarrow t_k = t - 1 \Rightarrow t \geq 2$. Using (3.3) for $l = j$, $c_{LHS} \in \{0, 2 \cdot \frac{1}{2^t}\} \leq \frac{1}{2}$ while $c_{RHS} \in \{1, 1 + \frac{1}{2^{t_k}}\} \geq 1$. Thus, $c_{LHS} \neq c_{RHS}$ giving us a $\Rightarrow\!\!\!\!\!\!\!\!\Leftarrow$. Thus, $2 \cdot \mathbf{g}_i$ is unique in this case.

Subcase 3: $\mathbf{g}_j, \mathbf{g}_k \in supp(g_H) \setminus V_g$.

Using (3.3) for $l = 1$, we get $c_{LHS} = 2 \cdot \frac{1}{2^t}$ and $c_{RHS} = \frac{1}{2^{t_j}} + \frac{1}{2^{t_k}}$. Thus, $c_{LHS} = c_{RHS} \Rightarrow t = t_j = t_k$. Since $\mathbf{g}_j \neq \mathbf{g}_k$, there exists some $\mathbf{g}_z \in V_g$ that generates $\mathbf{g}_j$ but not $\mathbf{g}_j$. Using (3.3) for $l = z$, $c_{RHS} = \frac{1}{2^{t_j}} = \frac{1}{2^t}$ and $c_{LHS} \in \{0, 2 \cdot \frac{1}{2^t}$. Thus, $c_{LHS} \neq c_{RHS} \Rightarrow$ we have a $\Rightarrow\!\!\!\!\!\!\!\!\Leftarrow$. Thus, $2 \cdot \mathbf{g}_i$ is unique in this case.

Thus, we see that $2 \cdot \mathbf{g}_i$ is unique for all $i \in [t]$ except $n-1$ cases. Thus we get that $\|f\| \geq t-(n-1)$. Using Definition 3.1, this means $\|f\| \geq n^{\Omega(log\, n)} - (n-1) = \Omega(\|g_H\|)$ □

We will now look at another class of polynomials which has "sparse" factors.

Definition 3.10. (Uniform convex polynomial) Let $g \in \mathbb{F}[\mathbf{x}]$. Let $V_g = \{\mathbf{g}_1, ..., \mathbf{g}_t\}$ denote set of points that act as vertices for the polytope pf g. Define

$$UCC_g = \left\{ \sum_{i \in S} \frac{1}{|S|} \mathbf{g}_i \mid \phi \subset S \subseteq V_g \right\}$$

Let E_g be some subset of UCC_g. Then, g is called an *uniform convex polynomial* if $supp(g) = V_g \cup E_g$.

Remark 3.11. Note that $|UCC_g| \leq 2^t - 1$ where $t = |V_g|$. Thus, such polynomials can be "dense" while having only a "sparse" vertex set.

Theorem 3.12. *Let $\mathbb{F}$ be a field of char $\neq 2$ and $g \in \mathbb{F}[\mathbf{x}]$ be a* uniform convex polynomial *with linearly independent vertices. Let $f = g^2$. Then, $\|f\| \geq \|g\|$.*

Proof. Let $supp(g) = \{\mathbf{g}_1, ..., \mathbf{g}_T\}$ with $V_g = V(CS(supp(g)) = \{\mathbf{g}_1, ..., \mathbf{g}_t\}$. For any point $\mathbf{g}_r \in supp(g)$, let $m_{\mathbf{g}_r}$ denote the monomial corresponding to it in g. We have,

$$g = \alpha_1 m_{\mathbf{g}_1} + ... + \alpha_T m_{\mathbf{g}_T} \;\; where \;\; \alpha_i \in \mathbb{F}$$

$$f = g^2 \Rightarrow f = \sum_{i=1}^{T} \alpha_i^2 m_{g_i}^2 + \sum_{i=1}^{T} \sum_{j=1}^{T} 2\alpha_i \alpha_j m_{\mathbf{g}_i} m_{\mathbf{g}_j}$$

When $\mathbb{F}$ has $char = 2$, the terms containing $2\alpha_i \alpha_j m_{\mathbf{g}_i} m_{\mathbf{g}_j}$ vanish. Hence, here we are only looking at fields with $char \neq 2$.

Consider any internal point $\mathbf{g}_p$ in g such that it is the uniform convex combination of $\mathbf{g}_{i_1}, ..., \mathbf{g}_{i_k}$ from V_g. Formally,

$$\mathbf{g}_p = \frac{1}{k}\sum_{l=1}^{k} \mathbf{g}_{i_l} \;\; ; \;\; i_l \in [t]$$

We claim that for all $l \in [k]$, the monomial $m_{\mathbf{g}_p} \cdot m_{\mathbf{g}_{i_l}}$ in $f = g^2$ can't be cancelled by any other monomial in f. We will show this using the linear independence of vertices of g.

Without loss of generality, let $l = 1$. Suppose that $m_{\mathbf{g}_p} \cdot m_{\mathbf{g}_{i_1}}$ is cancelled by some monomial. Then, $\exists \mathbf{g}_a, \mathbf{g}_b \in supp(g)$ such that

$$\mathbf{g}_a + \mathbf{g}_b = \mathbf{g}_p + \mathbf{g}_{i_l} \Rightarrow \mathbf{g}_a + \mathbf{g}_b = \frac{1}{k}\sum_{l=1}^{k} \mathbf{g}_{i_l} + \mathbf{g}_{i_1} \tag{3.7}$$

Since V_g has independent vertices, in the above equation the coefficient for each vertex must match in both sides.

Case 1: $\mathbf{g}_a, \mathbf{g}_b \in V_g$

Suppose $\mathbf{g}_a = \mathbf{g}_{i_1} \Rightarrow \mathbf{g}_b = \mathbf{g}_p$ which is a $\Rightarrow\!\!\!\Leftarrow$ as $\mathbf{g}_p$ is an internal point and $\mathbf{g}_b$ a vertex. Hence $\mathbf{g}_a \neq \mathbf{g}_{i_1}$. Similarly, $\mathbf{g}_b \neq \mathbf{g}_{i_1}$. Thus, looking at the coefficients for the vertex $\mathbf{g}_{i_1}$ in (3.7), we get that $c_{LHS} = 0$ and $c_{RHS} = 1 + \frac{1}{k} \neq c_{LHS}$. Hence, there don't exist any such $\mathbf{g}_a, \mathbf{g}_b$.

Case 2 : $\mathbf{g}_a \in V_g, \mathbf{g}_b \in supp(g) \setminus V_g$

Suppose $\mathbf{g}_a = \mathbf{g}_{i_1} \Rightarrow \mathbf{g}_b = \mathbf{g}_p$. But this would give a term $2\alpha_a\alpha_b m_{\mathbf{g}_a} m_{\mathbf{g}_b}$ and won't lead to a cancellation. When $\mathbf{g}_a \neq \mathbf{g}_{i_1}$, looking at coefficient of $\mathbf{g}_{i_1}$ in (3.7), $c_{RHS} = 1 + \frac{1}{k}$ and $c_{LHS} \in \{0, \frac{1}{t_b}\}$, where t_b is the number of vertices required to generate $\mathbf{g}_b$. Since, $\mathbf{g}_b$ is an internal point, $t_b \geq 2$ implying that $c_{LHS} \leq \frac{1}{2} \neq c_{RHS}$. Hence, there don't exist any such $\mathbf{g}_a, \mathbf{g}_b$.

Case 3 : $\mathbf{g}_a, \mathbf{g}_b \in supp(g) \setminus V_g$

Looking at coefficient of $\mathbf{g}_{i_1}$ in (3.7), $c_{RHS} = 1 + \frac{1}{k}$ and $c_{LHS} \in \{0, \frac{1}{t_a}, \frac{1}{t_b}, \frac{1}{t_a} +$

$\frac{1}{t_b}\} \leq \frac{1}{t_a} + \frac{1}{t_b} \leq \frac{1}{2} + \frac{1}{2} = 1 < c_{RHS}$. Thus, $c_{LHS} \neq c_{RHS}$. Hence, there don't exist any such $\mathbf{g}_a, \mathbf{g}_b$.

Hence, we see that product of each internal point in g with atleast one vertex survives in f. Also, by item (2) of Theorem 2.15, we know that each vertex of g survives in f. Thus, we have $\|f\| \geq \|g\|$. □

3.3 Some limitations and counterexamples

We now look at some limitations of our proof techniques and explore why they can't be generalised further, by listing some counterexamples.

1. A natural extension to consider for the approach used in Theorem 3.12 would be to extend it to polynomials where the internal points can be *non-uniform convex combinations* of vertices. We now shown an example of a polynomial for which this fails.

   Example 3.1. *Let g be a polynomial in $\mathbb{F}[\mathbf{x}]$ such that*

   $$supp(g) = \left\{\mathbf{v}_1, \mathbf{v}_2, \frac{3\mathbf{v}_1 + \mathbf{v}_2}{4}, \frac{7\mathbf{v}_1 + \mathbf{v}_2}{8}\right\}$$

   Clearly, $\mathbf{v}_1, \mathbf{v}_2$ constitute the vertex set, and the others are internal points. Then,

   $$\frac{3\mathbf{v}_1 + \mathbf{v}_2}{4} + \mathbf{v}_1 = 2 \cdot \frac{7\mathbf{v}_1 + \mathbf{v}_2}{8}$$

   refuting the claim that we make in the proof of Theorem 3.12.

2. Taking inspiration from item (2) of Theorem 2.15, a natural extension to investigate is the following:

   Conjecture 3.13 (Uniquely generated set)**.** *Let $f = g \cdot h$ with $P_f = P_g + P_h$. Then, there exists a general set $S \supsetneq V_g$ such that for every point $\mathbf{u} \in S$, there*

exists a point $\mathbf{v} \in P_h$ *such that* $\mathbf{u}+\mathbf{v}$ *is uniquely generated in* $P_g + P_h$*, that is* $\mathbf{u}+\mathbf{v} \neq \mathbf{u}'+\mathbf{v}'$ *for any other* $\mathbf{u}' \in P_g$ *and* $\mathbf{v}' \in P_h$*.*

We will now provide a counterexample to the above conjecture.

Example 3.2. *Consider the case of* $f = g^2$*. Then for all* $\mathbf{u} \neq \mathbf{v}$ *in* $supp(g)$*, we know that* $\mathbf{u}+\mathbf{v}$ *is trivially also generated as* $\mathbf{v}+\mathbf{u}$*. Therefore, at most only the* $2\mathbf{u} \in P_g + P_h$ *can be uniquely generated.*
Let $Z = \{\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n\}$ *where* $n = 2^m$ *for some* $m \in \mathbb{Z}$*. Let* Z_k *denotes the set formed by taking the uniform convex combination of any* k *points from* Z*, i.e.,*

$$Z_k = \left\{ \frac{\mathbf{z}_{i_1} + \mathbf{z}_{i_2} + ... + \mathbf{z}_{i_k}}{k} \mid 0 < i_1 < i_2 < ... < i_k \leq n \right\}$$

Thus, $|Z_k| = \binom{n}{k}$*.*
Define $supp(g) = Z_1 \cup Z_2 \cup Z_4 \cup Z_8 \cup ... \cup Z_{\frac{n}{2}}$*. Thus,*

$$|supp(g)| = \sum_{t=1}^{m-1} \binom{n}{2^t} \geq \binom{n}{2^{m-1}} = \binom{n}{\frac{n}{2}} \approx exp(n) \ \ (using \ \ Lemma \ \ A.8)$$

Also,

$$V_g = V(CS(supp(g))) = Z_1 \Rightarrow |V_g| = n$$

Thus, this is a "dense" polynomial with a "sparse" vertex set.
Notice that for $k > 1$*, for any* $\mathbf{g}_1 \in Z_k \subseteq supp(g)$*, there exist* $\mathbf{g}_2, \mathbf{g}_3 \in Z_{\frac{k}{2}} \subseteq supp(g)$ *such that* $2 \cdot \mathbf{g}_1 = \mathbf{g}_2 + \mathbf{g}_3$*.*
Thus, in $f = g^2$*, the only uniquely generated monomials* $2\mathbf{u}$ *correspond to the vectors* $\mathbf{u}$ *in* Z_1 *inside* $supp(g)$*. Hence, we have only* n *uniquely generated points in* $f = g^2$*. Thus, it is possible that we might have an* f *with* $\|f\| = n$ *and its factor* g *having* $\|g\| = exp(n)$*.*

Thus, a purely combinatorial approach to the factor sparsity problem which tries to get a bound by analysing the *uniquely generated points* in $f = g \cdot h$ can give us a "sparse" polynomial with "dense" factors. Hence, this approach can't be used to prove Conjecture 1.1.

Chapter 4

Conclusions and future work

In conclusion, we looked at upper bounds for sparsity of factors of $s - sparse$ polynomials with bounded individual degree with the view of resolution of the Conjecture 1.1. We analysed the best known bound ([BSV20]) currently by looking at it's limiting case, and showed that the polynomial representation of it in-fact doesn't refute polynomial sparsity conjecture. We also gave improved factor sparsity bounds for certain special classes of polynomials. Additionally, we showed the limitations of our proof techniques and some approaches that won't help in solving the problem. However, there still remains a lot of scope for work in this problem. The most important question still remains establishing existence/non-existence of polynomial sparsity of factors of polynomials of bounded individual degree. The best lower bound known to us is s^d over general fields, as shown in Example 1.1. However, the polynomial considered in that example is a *symmetric polynomial*, for which an $s^{\mathcal{O}(d^2 \log d)}$ upper bound on the factor-sparsity of an s-sparse, symmetric polynomials with individual degree $\leq d$ has already been established in [BS22]. For non-symmetric polynomials, the best known lower bound is even "smaller" than s^d while the best known upper bound result is $s^{\mathcal{O}(d^2 \log n)}$ by [BSV20]. Thus, there remains a considerable gap between the upper bound and lower bounds, and is a very intriguing question to settle.

Appendix A

Additional theorems and proofs

Theorem A.1. *(Carathéodory theorem) Let $M \in \mathbb{R}^n$ be a finite set. Let $\mu \in CS(M)$. Then, μ can be written as a convex combination of atmost $n+1$ points in M.*

Proof. If $|V(M)| \leq n+1$, then the claim always holds, as any point inside the convex hull of M can be expressed as a convex combination of the vertices.
Otherwise, let $V(M) = \{\mathbf{x}_1, ..., \mathbf{x}_r\}$ where $r > n+1$. Then, $\mathbf{x}_i - \mathbf{x}_r \ \forall i \in [r-1]$ are linearly dependent, as there are atleast $n+1$ such vectors in $\mathbb{R}^n$. Thus, $\exists \beta_i$ s.t. not all are 0 and $\sum_{i=1}^{r-1} \beta_i(\mathbf{x}_i - \mathbf{x}_r) = 0$. Define $\beta_r = -\sum_{i=1}^{r-1} \beta_i$. Then,

$$\sum_{i=1}^{r-1} \beta_i(\mathbf{x}_i - \mathbf{x}_r) = \sum_{i=1}^{r-1} \beta_i \mathbf{x}_i - (\sum_{i=1}^{r-1} \beta_i)\mathbf{x}_r = \sum_{i=1}^{r-1} \beta_i \mathbf{x}_i + \beta_r \mathbf{x}_r = \sum_{i=1}^{r} \beta_i \mathbf{x}_i = 0 \quad \text{(A.1)}$$

$$\sum_{i=1}^{r} \beta_i = \sum_{i=1}^{r-1} \beta_i + \beta_r = 0 \quad \text{(A.2)}$$

Now, for any point $\mu \in CS(M)$,
$\mu = \sum_{i=1}^{r} \lambda_i \mathbf{x}_i$ where $\lambda_i \geq 0 \ \forall i \in [r]$, and $\sum_{i=1}^{r} \lambda_i = 1$
Using A.1 and A.2, $\mu = \sum_{i=1}^{r} \lambda_i \mathbf{x}_i - \alpha(\sum_{i=1}^{r} \beta_i \mathbf{x}_i) = \sum_{i=1}^{r} (\lambda_i - \alpha\beta_i)\mathbf{x}_i$ where $\sum_{i=1}^{r}(\lambda_i - \alpha\beta_i) = \sum_{i=1}^{r} \lambda_i - \alpha(\sum_{i=1}^{r} \beta_i) = 1$. Thus if we can show that $\lambda_i - \alpha\beta_i \geq 0$

for all $i \in [r]$, this would also represent a convex combination.
Now, $\forall i \in [r]$ s.t. $\beta_i \leq 0$, we have $\lambda_i - \alpha\beta_i \geq 0$. Let $T = \{j \mid \beta_j \geq 0\}$. Set

$$\alpha = min_{j \in T}\{\frac{\lambda_j}{\beta_j}\} \quad , \quad k = argmin_{j \in T}\{\frac{\lambda_j}{\beta_j}\} \tag{A.3}$$

Now $\forall\, i \in T$, we have $\lambda_i - \alpha\beta_i = \lambda_i - \frac{\lambda_k}{\beta_k}.\beta_i \geq \lambda_i - \frac{\lambda_i}{\beta_i}\beta_i \geq 0$. Also, $\lambda_k - \alpha\beta_k = 0$.
With α, k as per A.3, we get $\mu = \sum_{i=1}^{k-1}(\lambda_i - \alpha\beta_i)\mathbf{x}_i + \sum_{i=k+1}^{r}(\lambda_i - \alpha\beta_i)\mathbf{x}_i$ and hence a convex combination of only $r-1$ points. Repeat the process until you get μ as a convex combination of $n+1$ points. We can always do this as for all values of $r > n+1$, as we will have linear dependence between the chosen vectors. □

Theorem A.2. *(Chernoff-Hoeffding's inequality) Let $\theta_1, ..., \theta_m$ be identical independently distributed random variables such that $\mathbb{E}(\theta_i) = \mu$ and $\mathbb{P}(a \leq \theta_i \leq b) = 1$. Then,*

$$\forall\, \epsilon > 0, \;\; \mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m}\theta_i - \mu\right| > \epsilon\right) \leq 2\, exp\left(\frac{-2m\epsilon^2}{(b-a)^2}\right)$$

Theorem A.3. *Let V_q be a $q-$dimensional vector space over a finite field $\mathbb{F}_p$. Then,*

1. *Number of sets of basis for a $k-$dimensional subspace W is equal to $\prod_{i=0}^{k-1}(p^k - p^i)$*
2. *Number of $k-$dimensional subspaces W is equal to $\frac{\prod_{i=0}^{k-1} p^q - p^i}{\prod_{i=0}^{k-1} p^k - p^i}$*

Proof. **1)** Cardinality of $W = p^k$. For finding a basis, we need k linearly independent vectors, thus we can choose the first vector $\mathbf{v}_1$ in $p^k - 1$ ways, as $\mathbf{0}$ can't be part of any basis. For second choice, we need to avoid all vectors in span $\{\mathbf{0}, \mathbf{v}_1\}$, and hence can make the choice in $p^k - p$ ways. Proceeding similarly, the r^{th} choice can be made in $p^k - p^{r-1}$ ways. Thus the total number of basis $= \prod_{i=0}^{k-1}(p^k - p^i)$.
2) To generate a $k-$dimensional subspace W, we only need k linearly independent vectors from V_q. Using a similar argument as above, we can pick them in $\prod_{i=0}^{k-1} p^q - p^i$ ways. But from **1)**, we know that each such W will have $\prod_{i=0}^{k-1}(p^k - p^i)$ sets of basis. Hence, number of <u>distinct</u> $k-$dimensional subspaces W is equal to $\frac{\prod_{i=0}^{k-1} p^q - p^i}{\prod_{i=0}^{k-1} p^k - p^i}$. □

Corollary A.4. *Total Number of subspaces of a $q-$dimensional vector space V_q over a finite field $\mathbb{F}_p$ is $p^{\Omega(q^2)}$.*

Proof. Let S_n denote number of $n-$*dimensional* subspaces of V_q. Let T denote total number of subspaces.

$$T = S_0 + \sum_{i=1}^{q} S_i = 1 + \sum_{k=1}^{q} \left(\frac{\prod_{i=0}^{k-1} p^q - p^i}{\prod_{i=0}^{k-1} p^k - p^i} \right) \geq \sum_{k=1}^{q} \left(\frac{\prod_{i=0}^{k-1} p^q - p^{k-1}}{\prod_{i=0}^{k-1} p^k - 1} \right)$$

$$\geq \sum_{k=1}^{q} \left(\frac{(p^q - p^{k-1})^k}{(p^k-1)^k} \right) \geq \sum_{k=1}^{q} \left(\frac{(p^q - p^k)^k}{p^{k^2}} \right) \geq \frac{(p^q - p^{\frac{q}{2}})^{\frac{q}{2}}}{p^{\frac{q}{4}^2}} = (p^{\frac{q}{2}} - 1)^{\frac{q}{2}} = p^{\Omega(q^2)}$$

□

Lemma A.5. *Let V be a vector space over $\mathbb{F}$. Let $S \subseteq V$ be a subspace, and $S^\perp$ be the orthogonal complement of S. Then, $S \oplus S^\perp = V$.*

Proof. $S^\perp = \{\mathbf{v} \mid \langle \mathbf{v}, \mathbf{u} \rangle \ \ \forall\ \mathbf{u} \in S\}$. This implies that $\mathbf{0} \in S^\perp$. Also, let $\mathbf{x}, \mathbf{y} \in S^\perp$. Then for any $\mathbf{u} \in S$, $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{u} \rangle = \langle a\mathbf{x}, \mathbf{u} \rangle + \langle b\mathbf{y}, \mathbf{u} \rangle = 0$. Thus $a\mathbf{x} + b\mathbf{y} \in S^\perp$ for any $a, b \in \mathbb{F}$. Hence, we can see that $S^\perp$ is a subspace.
Suppose, $\mathbf{u} \in S, S^\perp \Rightarrow < \mathbf{u}, \mathbf{u} >= 0 \Rightarrow \mathbf{u} = \mathbf{0}$. Thus, $S \cap S^\perp = \{\mathbf{0}\}$.
Also, $S + S^\perp \subseteq V$ trivially as both are subspaces of V. Consider $\mathbf{v} \in V$. Let $S = \{\mathbf{x}_1, ..., \mathbf{x}_m\}, S^\perp = \{\mathbf{y}_1, ..., \mathbf{y}_n\}$.

Case 1: $\mathbf{v} \in span(S)$, then $\mathbf{v} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i \Rightarrow \mathbf{v} \in S + S^\perp$.

Case 2: $\mathbf{v} \in span(S^\perp)$, then $\mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{y}_i \Rightarrow \mathbf{v} \in S + S^\perp$.

Case 3: $\mathbf{v} \notin span(S^\perp), span(S)$. $\mathbf{v} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i + \beta \mathbf{z}$ where $\mathbf{z} \perp \mathbf{x}_i\ \forall i$. Thus, $\mathbf{z}$ can be written as a linear combination of $\mathbf{y}_j$. Therefore, $\mathbf{v} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i + \sum_{i=1}^{n} \beta_i \mathbf{y}_i \Rightarrow \mathbf{v} \in S + S^\perp$.

Thus, $V = S + S^\perp$. Since $S \cap S^\perp = \{\mathbf{0}\} \Rightarrow S \oplus S^\perp = V$. □

Theorem A.6. *(Uniqueness of orthogonal complements) Let $V \subset \mathbb{F}^n$ be a vector space. Unique subspaces of V have unique orthogonal complements.*

Proof. Suppose $S \oplus T_1 = S \oplus T_2 = V$. Consider $t \in T_1 \subseteq V \Rightarrow t \in S \oplus T_2 \Rightarrow t = s + t'$ for some $s \in S, t' \in T_2$. Taking dot product with s, we get $\langle t, s\rangle = \langle s, s\rangle + \langle t', s\rangle \Rightarrow 0 = \langle s, s\rangle + 0 \Rightarrow s = 0 \Rightarrow t = t' \Rightarrow T_1 \subseteq T_2$. Similarly $T_2 \subseteq T_1$. Hence $T_1 = T_2$. □

Claim A.7. *Let $\mathbf{y} \neq \mathbf{0} \in \mathbb{F}_2^n$ Then, $\sum_{\mathbf{z}\in\mathbb{F}_2^n}(-1)^{\langle \mathbf{y},\mathbf{z}\rangle} = 0$*

Proof. Proof by induction on n.

Basis (n=1): We have $\mathbf{y} \neq \mathbf{0}$ giving us $\mathbf{y} = 1$. Thus $\sum_{\mathbf{z}\in\mathbb{F}_2}(-1)^{\langle 1,\mathbf{z}\rangle} = (-1)^{\langle 1,0\rangle} + (-1)^{\langle 1,1\rangle} = 1 + (-1) = 0$

Hypothesis: Let it be true for all positive integers $n < m$.

Induction: For $n = m$.

$$\sum_{\mathbf{z}\in\mathbb{F}_2^m}(-1)^{\langle \mathbf{y},\mathbf{z}\rangle} = \sum_{\mathbf{z}'\in\mathbb{F}_2^{m-1}}(-1)^{\langle \mathbf{y},(\mathbf{z}',0)\rangle} + (-1)^{\langle \mathbf{y},(\mathbf{z}',1)\rangle}$$

Case 1: Last bit of $\mathbf{y}$ is 1, i.e. $\mathbf{y}_m = 1$.
Then, $1 + \langle \mathbf{y}, (\mathbf{z}', 0)\rangle = \langle \mathbf{y}, (\mathbf{z}', 1)\rangle \Rightarrow (-1)^{\langle \mathbf{y},(\mathbf{z}',0)\rangle} + (-1)^{\langle \mathbf{y},(\mathbf{z}',1)\rangle} = 0$

Case 2: Last bit of $\mathbf{y}$ is 0, i.e. $\mathbf{y}_m = 0$.
Then, $\langle \mathbf{y}, (\mathbf{z}', 0)\rangle = \langle \mathbf{y}, (\mathbf{z}', 1)\rangle \Rightarrow (-1)^{\langle \mathbf{y},(\mathbf{z}',0)\rangle} + (-1)^{\langle \mathbf{y},(\mathbf{z}',1)\rangle} = 2(-1)^{\langle \mathbf{y},(\mathbf{z}',0)\rangle} = 2(-1)^{\langle (\mathbf{y}',0),(\mathbf{z}',0)\rangle} = 2(-1)^{\langle \mathbf{y}',\mathbf{z}'\rangle}$ where $\mathbf{y}', \mathbf{z}' \in \mathbb{F}_2^{m-1}$. Finally we get,
$\sum_{\mathbf{z}\in\mathbb{F}_2^m}(-1)^{\langle \mathbf{y},\mathbf{z}\rangle} = \sum_{\mathbf{z}'\in\mathbb{F}_2^{m-1}} 2(-1)^{\langle \mathbf{y}',\mathbf{z}'\rangle} = 0$ (by Hypothesis)

□

Lemma A.8.

$$\binom{2n}{n} \approx \Omega(exp(n))$$

Proof. By Stirling's approximation, $n! \approx \sqrt{2n\pi}(\frac{n}{e})^n$.

$$\binom{x+y}{x} = \frac{(x+y)!}{x!y!} \approx \frac{\sqrt{2(x+y)\pi}(\frac{x+y}{e})^{x+y}}{\sqrt{2x\pi}(\frac{x}{e})^x \sqrt{2y\pi}(\frac{y}{e})^y} = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{x+y}{xy}}\frac{(x+y)^{x+y}}{x^x\, y^y}$$

$$\Rightarrow \quad \binom{x+y}{x} \approx \sqrt{\frac{1}{2\pi}\left(\frac{1}{x}+\frac{1}{y}\right)}\left(1+\frac{y}{x}\right)^x \left(1+\frac{x}{y}\right)^y$$

Using this, we get

$$\binom{2n}{n} \approx \frac{1}{\sqrt{n\pi}}\, 4^n = \Omega(exp(n))$$

□

Theorem A.9 (Turan's theorem [Aig95])**.** *Let $G = (V, E)$ be a graph on n vertices without a $k-$clique. Then,*

$$|E| \leq \frac{(k-2)n^2}{2(k-1)}$$

Bibliography

[Aig95] Martin Aigner. Turán's graph theorem. *The American Mathematical Monthly*, 102(9):808–816, 1995.

[Bar15] Siddharth Barman. Approximating nash equilibria and dense bipartite subgraphs via an approximate version of caratheodory's theorem. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 361–369, 2015.

[BS22] Pranav Bisht and Nitin Saxena. Derandomization via symmetric polytopes: Poly-time factorization of certain sparse polynomials. 2022.

[BSV20] Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. Deterministic factorization of sparse polynomials with bounded individual degree. *Journal of the ACM (JACM)*, 67(2):1–28, 2020.

[CR88] Benny Chor and Ronald L Rivest. A knapsack-type public key cryptosystem based on arithmetic in finite fields. *IEEE Transactions on Information Theory*, 34(5):901–909, 1988.

[DdO14] Zeev Dvir and Rafael Mendes de Oliveira. Factors of sparse polynomials are sparse. *arXiv preprint arXiv:1404.4834*, 2014.

[Kal87] Erich Kaltofen. Single-factor hensel lifting and its application to the straight-line complexity of certain polynomials. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 443–452, 1987.

[Kal89] Erich Kaltofen. Factorization of polynomials given by straight-line programs. *Adv. Comput. Res.*, 5:375–412, 1989.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *computational complexity*, 13(1):1–46, 2004.

[KSS14] Swastik Kopparty, Shubhangi Saraf, and Amir Shpilka. Equivalence of polynomial identity testing and deterministic multivariate polynomial factorization. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 169–180. IEEE, 2014.

[Ost21] A Ostrowski. U on the meaning of the theory of convex polyhedra for the formal algebra. *Annual Reports German Math. Association*, 20:98–99, 1921.

[Sch00] Andrzej Schinzel. *Polynomials with special regard to reducibility*, volume 77. Cambridge University Press, 2000.

[Sud97] Madhu Sudan. Decoding of reed solomon codes beyond the error-correction bound. *Journal of complexity*, 13(1):180–193, 1997.

[SV10] Amir Shpilka and Ilya Volkovich. On the relation between polynomial identity testing and finding variable disjoint factors. In *International Colloquium on Automata, Languages, and Programming*, pages 408–419. Springer, 2010.

[Vol17] Ilya Volkovich. On Some Computations on Sparse Polynomials. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*, volume 81, pages 48:1–48:21, 2017.

[vzGK85] J von zur Gathen and E Kaltofen. Factoring sparse multivariate polynomials. *Journal of Computer and System Sciences*, 31(2):265–287, 1985.

[Zie12] Günter M Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 2012.