# A Super-Quadratic Lower Bound for Depth Four Arithmetic Circuits

A PROJECT REPORT
SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
**Master of Technology**
IN
**Faculty of Engineering**

BY

**Thankey Bhargav Deepakkumar**


Computer Science and Automation
Indian Institute of Science
Bangalore – 560 012 (INDIA)

July, 2020

# Declaration of Originality

I, **Thankey Bhargav Deepakkumar**, with SR No. **04-04-00-10-42-18-1-16040** hereby declare that the material presented in the thesis titled

**A Super-Quadratic Lower Bound for Depth Four Arithmetic Circuits**

represents original work carried out by me in the **Department of Computer Science and Automation** at **Indian Institute of Science** during the years **2018-20**.
With my signature, I certify that:

- I have not manipulated any of the data or results.
- I have not committed any plagiarism of intellectual property. I have clearly indicated and referenced the contributions of others.
- I have explicitly acknowledged all collaborative research and discussions.
- I have understood that any false claim will result in severe disciplinary action.
- I have understood that the work may be screened for any form of academic misconduct.

Date: 13 July 2020 Student Signature

In my capacity as supervisor of the above-mentioned work, I certify that the above statements are true to the best of my knowledge, and I have carried out due diligence to ensure the originality of the report.

Advisor Name: Advisor Signature

DEDICATED TO

*To my parents*

*for their unconditional love.*

# Acknowledgements

First and foremost, I would like to thank my advisor Prof. Chandan Saha. I am extremely fortunate to have him as my advisor and can not thank him enough for guiding and mentoring me throughout my two years as a Masters student in IISc. Not only did he provide me with invaluable advise for this work, but also helped me make an informed decision for future endeavors. It was his enthusiasm and optimism that kept me going when I was stuck on this project. This work would not have been possible without his constant support, both moral and technical.

I would also like to thank all the amazing professors in the Department of Computer Science and Automation. I would especially like to thank Chandan for teaching me Complexity Theory and Representation Theory, Prof. Siddharth Barman for teaching me Algorithms, Game Theory and Data Science, Prof. Anand Louis and Prof. Arindam Khan for teaching me Approximation Algorithms, Prof. M. Narasimha Murty for teaching me Linear Algebra and Prof. Shalabh Bhatnagar for teaching me Probability Theory; what I learned in these courses has been immensely useful in this work. I am thankful to the organisers of the Workshop on Algebraic Complexity Theory 2019; this workshop was an excellent exposure to the area of Algebraic Complexity Theory. I also thank the organizers of the Theory Lunch series in CSA; the talks in this series have helped me get a glimpse of the larger picture of Theoritical Computer Science.

I would like to thank my collaborators Nikhil Gupta and Chandan Saha, their invaluable contributions have made this work possible. I am also grateful to Prof. Arindam Khan for being the reader for mid-term and final evaluations of this project, providing valuable feedback and asking interesting questions. I thank Ankit Garg and Neeraj Kayal for sitting through presentations of this work. The insightful questions that they asked have helped me improve my understanding of some crucial aspects of this work.

**Acknowledgements**

In my two years at IISc I have made some wonderful friends. I have been lucky to have lab mates like Vineet Nair, Nikhil Gupta and Janaky Murthy. I have learned a lot from them not only about Complexity Theory but also about graduate life in general. I am especially thankful to Nikhil for all I learned from him in our many (and long) academic and a lot of non-academic discussions. I thank Swati Allabadi for our (often late night) chats and discussions, her helpful musings about life and career and being a good friend despite all my annoying idiosyncrasies. I thank Mahak Pancholi and Pooja Gupta for being amazing and forever cheerful and optimistic friends and co-organizers of extra curricular activities at IISc. I thank B Pratheek and Stanly Samuel for being amazing seniors and providing valuable advice about IISc. I would also like to thank Prasanna Patil, Raj Rajveer, Deepak Poonia, Dhiraj Shanbhag, Aakash Panda, Hemanta Makawna, Proteek Paul and all my other batch mates and juniors for making my stay at IISc enjoyable.

Last but certainly not the least, I would like to thank my family - my parents, my (late) grandfather, my grandmother, and my sister - for loving me no matter what and always being there for me. I do not have enough words to express how fortunate I am to have them in my life and how much I owe them.

# Abstract

A depth four arithmetic circuit contains a sum gate ($+$-gate) at the top followed by a layer of product gates ($\times$-gates), a layer of sum gates and again a layer of product gates at the bottom of the circuit; it can compute any multivariate polynomial. We prove a $\widetilde{\Omega}(n^{2.5})$ lower bound on the size, i.e. the number of edges (or wires) and a lower bound of $\widetilde{\Omega}(n^{1.5})$ on the number of gates of general depth four arithmetic circuits (here $n$ is the number of variables in the polynomial computed by the circuit). To the best of our knowledge this is the first super-quadratic lower bound for this model.

Recently, the problems of proving lower bounds for depth three and depth four arithmetic circuits have received a significant amount of attention, and with good reason. Due to a line of work in depth reduction for arithmetic circuits starting with [VSBR83, AV08, Koi12] and culminating in [GKKS16, Tav15], it is known that proving moderately strong exponential lower bounds for special classes of depth three and depth four circuits would solve the long standing open problem of proving a separation between VP and VNP - the algebraic analogues of the classes P and NP.

Because of a long line of recent work in lower bounds for constant depth circuits [KS16a, KS17, KLSS17, FLMS15, KSS14, GKKS14, Kay12], we now know of such strong lower bounds for special classes of depth three and depth four circuits, that improving them a little would prove VP $\neq$ VNP. While exponential lower bounds are now known for special classes of depth three and depth four circuits, not a lot is known for general depth three and four circuits. For a long time, the best known lower bound for general depth three circuits was a lower bound of $\Omega(n^2)$ [SW01], which was recently improved to $\widetilde{\Omega}(n^3)$ [KST16, BLS16, Yau16]. We show a similar improvement of a factor of $n$ for general depth four circuits. Prior to our work, the best known lower bound for depth four circuits was a lower bound of $\widetilde{\Omega}(n^{1.5})$ [Sha17] which was a slight improvement on the lower bound of roughly $\Omega(n^{1.33})$ obtained from [SS97, Raz10].

# Publications based on this Thesis

1. A Super-Quadratic Lower Bound for Depth Four Arithmetic Circuits,
   Joint work with Nikhil Gupta and Chandan Saha,
   To appear in the proceedings of the 35-th IEEE Conference on Computational Complexity (CCC), 2020.

# Contents

# List of Figures

# Chapter 1

# Introduction

An arithmetic circuit is a directed acyclic graph with input gates (nodes with in-degree zero), output gates (nodes with out-degree zero), sum and product gates. The size of an arithmetic circuit is the number of edges in the circuit, while the depth of a circuit is the length of the longest directed path in the circuit. Such a circuit takes variables and field elements as inputs and computes a multivariate polynomial over some field $\mathbb{F}$. Arithmetic circuits are algebraic analog of boolean circuits. Just as the complexity classes $\mathsf{P}$ and $\mathsf{NP}$ play a crucial role in boolean complexity theory, the classes $\mathsf{VP}$ and $\mathsf{VNP}$ introduced by Valinat in [Val79] are crucial in algebraic complexity theory. $\mathsf{VP}$ is the class of all polynomial families $\{f_n\}_{n\geq 1}$ such that $f_n$ is a polynomial in $n^{O(1)}$ many variables, of $n^{O(1)}$ degree and can be computed by a polynomial size circuit. On the other hand, $\mathsf{VNP}$ is a class of all polynomial families $\{f_n\}_{n\geq 1}$ such that there exists a polynomial family $\{g_n(\mathbf{x},\mathbf{y})\}_{n\geq 1}$ in $\mathsf{VP}$ satisfying $f_n(\mathbf{x}) = \sum_{\mathbf{y}\in\{0,1\}^{|\mathbf{y}|}} g_n(\mathbf{x},\mathbf{y})$. It is easy to see that $\mathsf{VP} \subseteq \mathsf{VNP}$. It is believed that this containment is strict; in fact it is known that if $\mathsf{VP} = \mathsf{VNP}$, then the non-uniform versions of $\mathsf{P}$ and $\mathsf{NP}$ are the same i.e $\mathsf{P/poly} = \mathsf{NP/poly}$ and the polynomial hierarchy collapses to the second level [Bür00]. However, proving that this containment is strict is a major open problem in algebraic complexity theory. Showing that there exists a polynomial family in $\mathsf{VNP}$ that can not be computed by polynomial size circuits will resolve this problem. This lower bound problem is at the heart of algebraic complexity.

Despite decades of work in algebraic complexity, only modest lower bounds are known for general circuits, formulas (circuits whose underlying graph is a tree) and algebraic branching programs (algebraic analogues of branching programs). Baur and Strassen [BS83, Str73] showed than any arithmetic circuit computing the polynomial $\sum_{i\in[n]} x_i^d$ must be of size

$\Omega(n \log d)$. A lower bound of $\Omega(n^2)$ for arithmetic circuits computing the polynomial $\sum_{i,j\in[n]} x_i^j y_j$ was shown in [Kal85]. Recently, [CKSV19] showed a lower bound of $\Omega(n^2)$ on the size of any algebraic branching program computing the polynomial $\sum_{i\in[n]} x_i^n$. Because of the apparent difficulty of proving lower bounds for general circuits, formulas and algebraic branching programs, lower bounds for special classes of these models like non-commutative circuits, monotone circuits, multilinear circuits, constant depth circuits etc... have been studied extensively.

## 1.1 Motivation

Constant depth circuits form a natural and powerful special class of arithmetic circuits. A long line of work in depth reduction starting with [VSBR83, AV08, Koi12] and culminating in [Tav15, GKKS16] has shown that proving strong enough lower bounds on the size of depth three and homogeneous depth four circuits would imply super-polynomial lower bounds on the size of general arithmetic circuits - in particular, proving a $n^{\omega(\sqrt{d})}$ lower bound on the size of homogeneous depth four circuits, with bottom fan-in bounded by $\sqrt{d}$ and computing a degree $d$ polynomial in VNP would separate VP and VNP. These results provide compelling reasons for studying constant depth circuit lower bounds.

Using the technique of shifted partials introduced by Kayal in [Kay12], a series of lower bound results for homogeneous depth four circuits [GKKS14, KSS14, FLMS15, KLSS17, KS17] followed. In particular, [KLSS17] showed a lower bound of $n^{\Omega(\sqrt{d})}$ for a $n$-variate degree $d$ polynomial in VNP and [KS17] showed the same lower bound for a polynomial in VP; improving the former to $n^{\omega(\sqrt{d})}$ would separate VP and VNP. Moreover, [KS16a] showed a lower bound of $n^{\Omega(\frac{d}{\tau})}$ for a depth three circuit with bottom fan-in bounded by $\tau$ computing a $n$-variate degree $d$ polynomial in VNP. Again improving this lower bound to $n^{\omega(\frac{d}{\tau})}$ would separate VP and VNP.

Since the machinery of shifted partials works extremely well for homogeneous depth four circuits and depth three circuits with small bottom fan-in, it is natural to wonder whether it can be used to improve upon the known lower bounds for general depth three and depth four circuits. Recently, [KST16, BLS16, Yau16] showed a lower bound of $\widetilde{\Omega}(n^3)$ for general depth three circuits which is an improvement upon the previous best lower bound of $\Omega(n^2)$ from [SW01]. Similarly, we improve upon known super linear lower bounds for general depth four circuits and prove, what is to the best of our knowledge (and mentioned in the

survey [SY10]), the first super quadratic lower bound for depth four arithmetic circuits.

## 1.2 Our contribution

We now state our result. By a depth four arithmetic circuit, we mean a $\Sigma\Pi\Sigma\Pi$ circuit i.e. a circuit with a sum gate at the top, followed by product gates in the second level, sum gates in the third level and again product gates in the forth and the bottom most level.

**Theorem 1.1** (Lower bound for depth four circuits)**.** *Over any field of characteristic zero*[1]*, there exists a family of mulitilinear polynomials* $\{f_n\}_{n\geq 1}$ *in* VNP*, where* $f_n$ *is a polynomial in* $\Theta(n)$ *variables and of degree* $\Theta(n)$ *such that any depth four circuit computing* $f_n$ *has* $\Omega\left(\frac{n^{2.5}}{(\log n)^6}\right)$ *many wires/edges and* $\Omega\left(\frac{n^{1.5}}{(\log n)^4}\right)$ *many gates.*

**A word about the polynomial family.** The polynomial family $\{f_n\}_{n\geq 1}$ is a variant of the Nisan-Wigderson design polynomial family. This family of polynomials was introduced in [KSS14] and has been used to prove several lower bounds for depth three and depth four circuits [Raz10, KLSS17, KS17, KS16a, KS16b, KST16, Sha17]. The $n$-th member of the family, $f_n$ is a polynomial in $\mathbf{x} = \{x_1, ..., x_{3m}\}$ and $\mathbf{y} = \{y_1, ..., y_{3m}\}$ variables, where $m$ is an integer in $\left[\frac{n}{2}, 2n\right]$. The degree of the polynomial in $\mathbf{y}$ variables is $\deg_{\mathbf{y}}(f_n) = m$, while its degree in $\mathbf{x}$ variables is $\deg_{\mathbf{x}}(f_n) = d_{\mathbf{x}} = \Theta(\frac{\sqrt{m}}{\ln m})$. Informally, $f_n$ contains multiple 'copies' of the design polynomial in different subsets of the $\mathbf{x}$ variables, while the $\mathbf{y}$ variables are used as 'prefixes' to uniquely identify each such copy. Note that because of the way we have defined $m$ and $d_{\mathbf{x}}$, proving that any depth four circuit computing $f_n$ has $\Omega\left(\frac{m^2 d_{\mathbf{x}}}{(\ln m)^5}\right)$ many edges and $\Omega\left(\frac{m d_{\mathbf{x}}}{(\ln m)^3}\right)$ many gates would establish the theorem. The exact description of $f_n$ is given in Chapter 4.

Much like a lot of recent work in lower bounds for constant depth circuits, we also use a complexity measure to prove our result. The measure we use is a variant of the projected shifted partials measure used in [KLSS17, KS14, KS16a, Sha17]. Much like those works, our proof is divided into two main parts: proving that the measure of any depth four circuit is 'small' (which we do in Chapter 3) and constructing an explicit polynomial with 'large' measure (done in Chapter 4).

[1] The lower bound holds even if the characteristic is sufficiently large (see Chapter 4).

## 1.3 Related work

We now mention known lower bounds for general depth four circuits prior to our work. For arithmetic circuits of depth $\Delta$, a lower bound of $n \cdot \lambda_\Delta(n)$ was shown in [Val75, DDPW83, Pud94, RS03], where $\lambda_\Delta(n)$ is a slowly growing function; for $\Delta = 4$, $\lambda_\Delta(n) = \log^* n$. Hence, for depth four circuits, this yields a lower bound which is barely super linear. Shoup and Smolensky [SS97] showed a lower bound of $\Omega(\Delta \cdot n^{1+\frac{1}{\Delta}})$ for a depth $\Delta$ circuit with multiple outputs computing the polynomials $\{\sum_{j\in[n]} x_1^j y_j, \ldots, \sum_{j\in[n]} x_n^j y_j\}$. For the same model, Raz [Raz10] showed a lower bound of $\Omega(n^{1+\frac{1}{2\cdot\Delta}})$. While the lower bound in [SS97] is for polynomials of degree $\Theta(n)$, the lower bound in [Raz10] is for polynomials whose degree is $\Theta(\Delta)$. In fact, the polynomials used in [Raz10] bear a striking resemblance to the family of Nisan-Wigderson design polynomials that we use to prove our result. For $\Delta = 4$, the lower bounds in [SS97, Raz10] are roughly $\Omega(n^{1.33})$. These lower bounds were improved by Sharma in [Sha17] where they proved a lower bound of $\widetilde{\Omega}(n^{1.5})$ for a depth four circuit with single output.

## 1.4 Organisation

In Chapter 2, we establish some preliminaries about Algebraic Complexity Theory, define our complexity measure and state some well known results that will come in handy later in the thesis. In Chapter 3, we analyse the measure of a depth four circuit, while in Chapter 4, we construct an explicit polynomial with a large measure. Finally, in Chapter 5, we conclude and mention some interesting open problems for future investigations.

# Chapter 2

# Preliminaries

In this chapter we define some notations, introduce a few terms that we will use in this thesis and mention some well known numerical estimates that will later help us in the analysis.

## 2.1 Some notations

For $n \in \mathbb{N}$, we will denote by $[n]$ the set $\{1, ..., n\}$. For $n, r \in \mathbb{N}$, by $\binom{[n]}{r}$ we mean the set of all subsets of $[n]$ of size $r$.

## 2.2 Basic definitions

**Definition 2.1** (Arithmetic circuit)**.** *An arithmetic circuit* $\mathsf{C}$ *over a field* $\mathbb{F}$ *and a set of variables* $\mathbf{x} = (x_1, ..., x_n)$ *is a directed acyclic graph. The vertices of* $\mathsf{C}$ *are called gates. Each gate with in-degree 0 is called an input gate and is labelled by either a variable or a field element. Every other gate is either labelled by a* $\times$ *(called a product gate) or a* $+$ *(called a sum gate). Every edge is labelled by a field element and every gate with out-degree 0 is called an output gate. An arithmetic circuit computes a polynomial in the natural way: an input gate computes the field element or variable it is labelled with. A sum gate computes the sum of polynomials computed by its inputs, each input scaled by the field element on the corresponding edge. Similarly, a product gate computes the product of polynomials computed by its inputs, each input scaled by the field constant on the corresponding edge.*

The size of an arithmetic circuit is equal to the number of edges in it and the depth of an arithmetic circuit is equal to the length of the longest directed path in it. The fan-in of a gate is equal to the number of edges entering the gate and the fan-out of a gate is the number of

edges leaving the gate. An example of an arithmetic circuit is shown in the Figure 2.1.

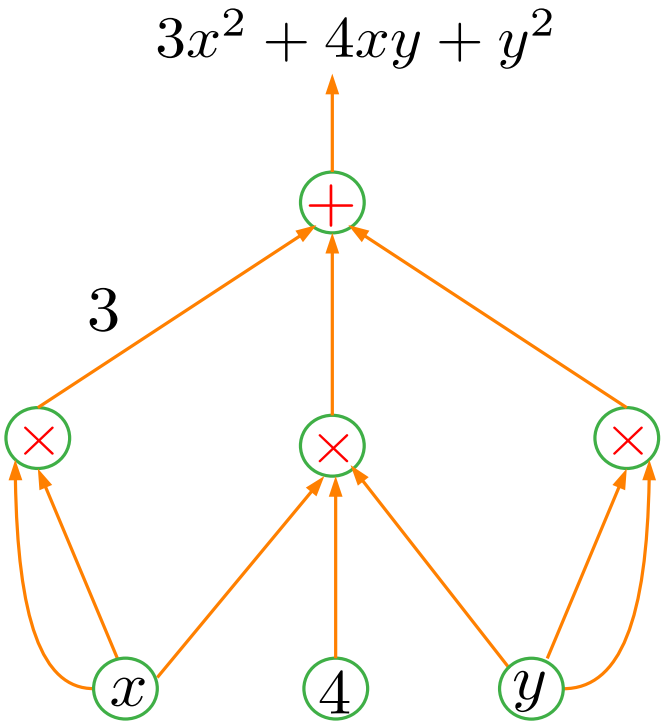


Figure 2.1: An arithmetic circuit.

A depth four circuit is a circuit with four alternating levels of gates. In this thesis, we deal with a ΣΠΣΠ circuit - such a circuit has one sum gate in the top most (first) level followed by a second level of product gates, a third level of sum gates and a bottom (fourth) level of product gates. An illustration if ΣΠΣΠ is shown in Figure 2.2.

The classes VP and VNP defined by Valiant in [Val79] are the algebraic analogs of P and NP respectively.

**Definition 2.2** (Class VP). $\mathsf{VP}_{\mathbb{F}}$ *is a class of all polynomial families* $\{f_n\}_{n\geq 1}$ *over a field* $\mathbb{F}$ *such that there exists a polynomial function* $t : \mathbb{N} \to \mathbb{N}$*, such that for all* $n \geq 1$*,* $f_n$ *is a polynomial in at most* $t(n)$ *variables, of degree at most* $t(n)$ *and computed by an arithmetic circuit of size at most* $t(n)$*.*

**Definition 2.3** (Class VNP). $\mathsf{VNP}_{\mathbb{F}}$ *is a class of all polynomial families* $\{f_n\}_{n\geq 1}$ *over a field* $\mathbb{F}$ *such that there exist polynomial functions* $k, t : \mathbb{N} \to \mathbb{N}$ *and a family of polynomials* $\{g_n\}_{n\geq 1}$ *such that*

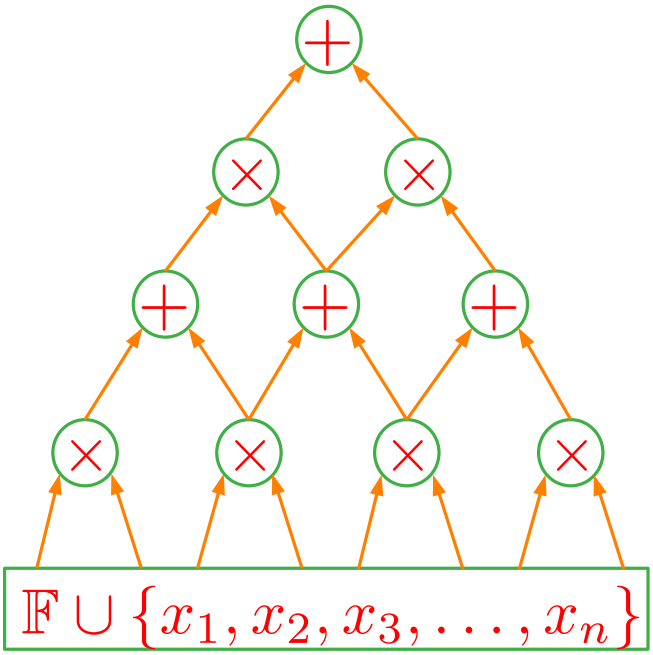


Figure 2.2: A depth four arithmetic circuit

*for all* $n \geq 1$,

$$f_n(x_1, ..., x_{k(n)}) = \sum_{w \in \{0,1\}^{t(n)}} g_{t(n)}(x_1, ..., x_{k(n)}, w_1, ..., w_{t(n)}).$$

## 2.3 The complexity measure

Now we define the complexity measure that we use to prove Theorem 1.1. Throughout this section, we will assume that $m \in \mathbb{N}$ is as stated in the paragraph following Theorem 1.1, $M \subseteq [3m]$, $|M| = m$, $f \in \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$ and $S \subseteq \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$. Note that the set $M$ is not fixed and will depend on the circuit under analysis. First let us define the support and degree of a monomial.

**Support and degree of a monomial.** The support of a monomial $\eta$, denoted $\mathrm{Supp}(\eta)$, is the set of variables appearing in it. Also, for any $\mathbf{z} \subseteq \mathbf{x} \cup \mathbf{y}$ we will use $\deg_{\mathbf{z}}(\eta)$ to denote its degree in $\mathbf{z}$ variables. We will say that $\eta$ is $\mathbf{z}$-multilinear if the degree of every $\mathbf{z}$ variable in $\eta$ is at most one.

Before defining the measure, let us define the operations that make up the measure.

1. **Partial derivatives.** Let $\eta = x_1 \cdots x_k$ be a monomial in $\mathbf{x}$ variables. Then, we define the partial derivative of $f$ with respect to $\eta$ as

$$\frac{\partial f}{\partial \eta} := \frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_2}\left(\cdots\left(\frac{\partial f}{\partial x_k}\right)\right)\right).$$

   If the degree of $\eta$ is $k$, then $\frac{\partial f}{\partial \eta}$ is said to be a $k$-th order partial derivative of $f$. We denote by $\partial^k_{\mathbf{x}} f$ the set of all $k$-th order partial derivatives of $f$ taken with respect to multilinear monomials in $\mathbf{x}$ variables.

2. **The shift operation.** Let $\eta$ be a degree $\ell$ multilinear monomial in $\mathbf{x}_M$ variables. We say that the polynomial $\eta \cdot f$ is obtained by *shifting* $f$ by $\eta$. We denote by $\mathbf{x}^\ell_M f$ the set of polynomials obtained by shifting $f$ by all degree $\ell$ multilinear monomials in $\mathbf{x}_M$ variables and $\mathbf{x}^\ell_M S := \{\mathbf{x}^\ell_M f : f \in S\}$.

3. **Multilinear projection.** We define a map $\pi_{\mathbf{x}} : \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M] \to \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$ with $\pi_{\mathbf{x}}(f)$ being the polynomial made up of exactly the $\mathbf{x}$-multilinear monomials of $f$. Formally, for a monomial $\eta$, $\pi_{\mathbf{x}}(\eta) = \eta$ if $\eta$ is $\mathbf{x}$-multilinear and 0 otherwise. The map is then linearly extended for arbitrary polynomials and $\pi_{\mathbf{x}}(S) := \{\pi_{\mathbf{x}}(f) : f \in S\}$.

4. **A degree based projection.** For $i \in \mathbb{N}$ and $f \in \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$, we define $[f]_i$ to be the polynomial made up of only those monomials of $f$ whose $\mathbf{y}$-*degree* is exactly $i$. Formally, for a monomial $\eta$, $[\eta]_i = \eta$ if $\deg_{\mathbf{y}}(\eta) = i$ and 0 otherwise. It is then linearly extended for arbitrary polynomials and $[S]_m := \{[f]_m : f \in S\}$.

5. **An evaluation map.** For $\alpha \in \mathbb{F}$ and $\mathbf{z} \subseteq \mathbf{x}_M \cup \mathbf{y}_M$, we define a map $\sigma_{\mathbf{z}=\alpha} : \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M] \to \mathbb{F}[\mathbf{x}_M \setminus \mathbf{z},\ \mathbf{y}_M \setminus \mathbf{z}]$ with $\sigma_{\mathbf{z}=\alpha}(f)$ being obtained from $f$ by setting every variable in $\mathbf{z}$ to $\alpha$ and $\sigma_{\mathbf{z}=\alpha}(S) := \{\sigma_{\mathbf{z}=\alpha}(f) : f \in S\}$.

The operations given in 1, 2 and 3 constitute the projected shifted partials measure [KLSS17]. In this work, we define and use the measure $\mathsf{PSP}_{M,k,\ell}$, which is obtained by augmenting the projected shifted partials measure with the operations in 4 and 5 as follows.

**Definition 2.4** (The measure). *For $m, k, \ell \in \mathbb{N}$, $M \subseteq [3m]$, $|M| = m$ and $f \in \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$,*

$$\mathsf{PSP}_{M,k,\ell}(f) := \dim\left\langle \sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(\mathbf{x}^\ell_M\, \partial^k_{\mathbf{x}} f\right)\right]_m\right)\right\rangle.$$

**Observation 2.1** (Sub-additivity of the measure). *For any two polynomials $f, g \in \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$,*

$$\mathsf{PSP}_{M,k,\ell}(f+g) \leq \mathsf{PSP}_{M,k,\ell}(f) + \mathsf{PSP}_{M,k,\ell}(g).$$

The above observation is easy to prove and we omit its proof here.

## 2.4 Some numerical estimates

**Proposition 2.1** (Estimating Binomial Coefficients). *For any $n, k \in \mathbb{N}$, $k \leq n$, $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} < \left(\frac{en}{k}\right)^k$.*

**Proposition 2.2** ([GKKS14, KLSS17]). *Let $a(n), f(n), g(n) : \mathbb{Z}_{>0} \to \mathbb{Z}$ be integer values functions such that $(|f| + |g|) = o(a)$. Then, $\ln \frac{(a+f)!}{(a-g)!} = (f+g)\ln(a) \pm O\left(\frac{f^2+g^2}{a}\right)$.*

# Chapter 3

# Upper Bounding the Measure for a Depth Four Circuit

In this chapter, we derive a "small" upper bound on the $\mathsf{PSP}_{M,k,\ell}(\cdot)$ measure for a depth four circuit. This forms the first part of the proof of Theorem 1.1. The contents of this chapter are from our work [GST20].

Let $\mathtt{C}$ be a depth four circuit computing the polynomial $f = f_n$ (recall that $f_n$ is a polynomial in $\mathbf{x} = \{x_1, \ldots, x_{3m}\}$ and $\mathbf{y} = \{y_1, \ldots, y_{3m}\}$ variables, and its degree in $\mathbf{x}$ variables $d_{\mathbf{x}} = \Theta\left(\frac{\sqrt{m}}{\ln m}\right)$). Then, if the top fan-in of $\mathtt{C}$ is $s$, we can express it as $\mathtt{C} = \sum_{i=1}^{s} T_i$ where $T_i = \prod_{j=1}^{a_i} Q_{ij}^{e_{ij}}$ and $Q_{ij}$ are distinct sparse polynomials computed by the sum gates in the third level of $\mathtt{C}$ while $e_{ij} \in \mathbb{N}$. Throughout this section we will refer to the polynomials computed by the second, third and forth levels of any depth four circuit as product terms, sparse polynomials and monomials respectively. Throughout this section, we will assume that the underlying field $\mathbb{F}$ is algebraically closed. We will justify this assumption in Section 4.1. The proof of the upper bound is divided into the following three steps:

**Step 1: Restricting bottom support of $\mathtt{C}$.** In this step, we remove all monomials with 'large' support. We observe that this is actually an instance of the Set Cover problem. The universe is the set of all monomials with 'large' support and there is a set corresponding to each variable $x_i$ (resp. $y_i$); this is a set of all monomials in the universe containing $x_i$ (resp. $y_i$). Since setting $x_i$ (resp. $y_i$) to 0 removes all monomials containing $x_i$ (resp. $y_i$), our goal is to simply find a small collection of variables which 'covers' all monomials in the universe. This observation leads us to use a simple greedy approximation algorithm for the Set Cover problem to restrict the bottom support (for more details, see Section 3.1.1). This restriction

helps us in removing 'heavy' gates.

**Step 2: Removing heavy gates from $\mathtt{C}$.** We say that a product term in $\mathtt{C}$ is *heavy* if it is connected to more than $\widetilde{\Omega}(md_{\mathbf{x}})$ distinct sum gates from the third level. In other words, a heavy gate has more than $\widetilde{\Omega}(md_{\mathbf{x}})$ distinct sparse factors. It is not clear to us how to obtain a 'small' lower bound on $\mathtt{C}$ in presence of heavy gates. So, much like was done in the depth three circuit lower bounds [SW01, KST16], we too remove heavy gets from the circuit. While they remove heavy gates by going modulo the affine factors of heavy gates, we do not know how to generalise this technique for depth four circuits. In the following paragraph we explain how we remove heavy gates from $\mathtt{C}$.

We assume that the underlying field is algebraically closed. Then, we get rid of heavy gates by sequentially evaluating one sparse factor of each heavy gate to 0. While there exists a heavy gate in $\mathtt{C}$, we pick a sparse factor for which the ratio of the number of heavy gates connected to it to the number of monomials it contains is maximum and evaluate it to 0. As we have restricted the bottom support, we are able to argue that this greedy procedure removes $\Theta(m)$ heavy gates from $\mathtt{C}$ at the cost of setting only a few variables to field constants (if $\mathtt{C}$ contains more than $\Theta(m)$ heavy gates, its size is already $\widetilde{\Omega}(m^{1.5}d_{\mathbf{x}})$). Since the problem of removing heavy gates can be reformulated as an instance of the Weighted Set Cover problem, this procedure is similar to a greedy approximation algorithm for the Weighted Set Cover problem [Vaz01] (Section 2.1, page-16), however its analysis - detailed in Section 3.1.2 - differs.

**Step 3: Analysing the measure of $\mathtt{C}$.** After we have pruned $\mathtt{C}$, we obtain a 'small' upper bound on its measure using the analysis in [KST16]. However, instead of the shifted partials measure used in that work, we make use of a variant of projected shifted partials measure as this measure, in conjunction with steps 1 and 2 - roughly speaking - helps us control the formal degree of the product terms of $\mathtt{C}$.

## 3.1 Pruning a depth four circuit

We define a pruned depth four circuit as follows:

**Definition 3.1** (Pruned depth four circuit)**.** *We say that a depth four circuit $\mathtt{D}$ is a* pruned circuit *if the support of all monomials in $\mathtt{D}$ is at most $\tau = \lfloor 20 \ln m \rfloor$, and it does not contain any heavy*

*gate; i.e. the number of distinct sparse polynomials feeding into any product term in* $\mathtt{D}$ *is less than* $w = \left\lfloor \frac{md_{\mathbf{x}}}{\lambda_0 \cdot (\ln m)^3} \right\rfloor$.

We prune the circuit $\mathtt{C}$ in two steps. In step 1, we restrict the bottom support of $\mathtt{C}$ (Section 3.1.1) and in step 2, we remove all heavy gates from $\mathtt{C}$ (Section 3.1.2).

### 3.1.1 Restricting the bottom support of $\mathtt{C}$

If the number of monomials in $\mathtt{C}$ is more than $\left\lfloor \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \right\rfloor$, there is nothing to prove. Otherwise, we show that we can get rid of all monomials with support more than $\tau = \lfloor 20 \ln m \rfloor$ by setting $m$ $\mathbf{x}$ and $m$ $\mathbf{y}$ variables to 0.

**Lemma 3.1.** *Let the number of monomials in* $\mathtt{C}$ *be at most* $\left\lfloor \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \right\rfloor$. *Then, for sufficiently large* $m$, *there exists* $M_1 \subseteq [3m], |M_1| = m$ *such that all monomials in* $\mathtt{C}_1$ *obtained from* $\mathtt{C}$ *by setting variables* $\mathbf{x}_{M_1}$ *and* $\mathbf{y}_{M_1}$ *to* 0 *have support at most* $\tau$.

*Proof.* We first present a greedy procedure to remove all monomials with support more than $\tau$ and then argue that it sets $m$ variables each from $\mathbf{x}$ and $\mathbf{y}$ to 0.

**Procedure 1** Restriction procedure

1. $M_1 \leftarrow \emptyset, \mathtt{C}_1 \leftarrow \mathtt{C}, H :=$ set of all monomials of $\mathtt{C}_1$ with support more than $\tau$.
2. For $j \in [3m]$, $e(j) :=$ number of monomials in $H$ containing $x_j$ or $y_j$.
3. **while** $H \neq \emptyset$ **do**
4. Pick $j' \in [3m] \backslash M_1$ such that $e(j') \geq e(j)$ for all $j \in [3m]$. Set $x_{j'} = 0$ and $y_{j'} = 0$. Update $M_1 \leftarrow M_1 \cup \{j'\}$, $\mathtt{C}_1 \leftarrow$ circuit obtained from $\mathtt{C}_1$ by setting $x_{j'}$ and $y_{j'}$ to 0, $H \leftarrow$ set of all monomials of $\mathtt{C}_1$ with support more than $\tau$, and $e(j) \leftarrow$ number of monomials in $H$ containing $x_j$ or $y_j$.
5. **end while**

It is clear that the bottom support of $\mathtt{C}_1$ obtained after the termination of the procedure is at most $\tau$. Also, since we are only setting variables to 0, it trivially follows that the procedure does not increase the number of gates nor does it increase the fan-in of any gate in the circuit. Claim 3.1 (proved below) implies that the procedure terminates in at most $m$ iterations. If it terminates before $m$ iterations, we arbitrarily add an appropriate number of $j \in [3m]$ to $M_1$ so that $|M_1| = m$ and set $x_j$ and $y_j$ to 0 for all such $j$. □

**Claim 3.1.** *Procedure 1 terminates in at most* $m$ *iterations.*

*Proof.* Let $H_i$ be the set $H$ after the $i$-th iteration of the procedure. Since each monomial in $H_i$ has support more than $\tau$, for any such monomial there are at least $\frac{\tau}{2}$ distinct $j \in [3m] \backslash M_1$ such that at least one of $x_j$ and $y_j$ appears in it. Counting the number of times at least one of $x_j$ and $y_j$ appears in a monomial in $H_i$ and summing up these counts for all $j \in [3m] \backslash M_1$, we get that

$$\sum_{j \in [3m] \backslash M_1} e(j) \geq \frac{\tau \cdot |H_i|}{2};$$

so from an averaging argument there exists a $j$ such that

$$e(j) \geq \frac{\tau \cdot |H_i|}{6m}.$$

Hence, the size of $H_{i+1}$ is upper bounded as

$$|H_{i+1}| \leq |H_i| \cdot \left(1 - \frac{\tau}{6m}\right).$$

So after $i$ iterations of the procedure we get,

$$\begin{aligned}
|H_i| &\leq |H_0| \cdot \left(1 - \frac{\tau}{6m}\right)^i \\
&\leq \left\lfloor \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \right\rfloor \cdot \left(1 - \frac{\lfloor 20 \ln m \rfloor}{6m}\right)^i \\
&\leq \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \cdot \left(1 - \frac{(20 \ln m - 1)}{6m}\right)^i \\
&\leq \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \cdot e^{-\frac{3i \cdot \ln m}{m}} && \text{(for sufficiently large } m\text{)} \\
&= \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \cdot m^{-\frac{3i}{m}}.
\end{aligned}$$

For $i = m$, $|H_i| < 1$ (for sufficiently large $m$), i.e., the procedure terminates in at most $m$ iterations. □

### 3.1.2 Removing heavy gates from the circuit $\mathtt{C}_1$

Recall that a product term is called heavy if the number of distinct sparse polynomials feeding into it is more than $w = \left\lfloor \frac{md_{\mathbf{x}}}{\lambda_0 \cdot (\ln m)^3} \right\rfloor$. We say that a sparse polynomial in $\mathtt{C}_1$ is *light* if its fan-in is at most $\frac{m}{(\ln m)^2}$ and it is connected to a heavy gate. Notice that if the number of heavy gates in $\mathtt{C}_1$ is more than $m$ or there is a heavy product term connected to fewer than $\frac{m \cdot d_{\mathbf{x}}}{2 \cdot \lambda_0 \cdot (\ln m)^3}$ light sparse polynomials, then there is nothing to prove (where $\lambda_0$ is a large enough constant which will be fixed later in the analysis). Otherwise, we prove the following lemma.

**Lemma 3.2.** *Let $\mathtt{C}_1$ be the circuit obtained from $\mathtt{C}$ after applying Lemma 3.1. If the number of heavy gates in $\mathtt{C}_1$ is at most $m$, every heavy gate is connected to at least $\frac{m \cdot d_{\mathbf{x}}}{2 \cdot \lambda_0 \cdot (\ln m)^3}$ many light sparse polynomials and the sum of fan-ins of all the light sparse polynomials is at most $\frac{m^2 \cdot d_{\mathbf{x}}}{160 \cdot \lambda_0 \cdot (\ln m)^5}$, then there exist $M_2 \subseteq [3m] \setminus M_1$, $|M_2| = m$ and $\alpha_u, \beta_u \in \mathbb{F}$ for $u \in M_2$, such that the circuit $\mathtt{D}$ obtained by setting $x_u = \alpha_u, y_u = \beta_u$ for all $u \in M_2$ is a pruned circuit.*

*Proof.* For a light sparse polynomial $Q_j$ in $\mathtt{C}_1$, let $b_j$ be its fan-in and $c_j$ denote the number of distinct heavy gates connected to it. Consider the following procedure which greedily picks the light sparse polynomial $Q_j$ for which the ratio $\frac{c_j}{b_j}$ is maximum and evaluates it to 0. This is possible since we have assumed the field $\mathbb{F}$ to be algebraically closed.

**Procedure 2** Removing Heavy Gates

1. $M_2 \leftarrow \emptyset, \mathtt{D} \leftarrow \mathtt{C}_1, H :=$ set of all heavy gates in $\mathtt{C}_1$, $i \leftarrow 1$.
2. **while** $H \neq \emptyset$ **do**
3. Let $Q_i$ be the light sparse polynomial such that the ratio $\frac{c_i}{b_i}$ is maximum among all the light sparse polynomials present in $\mathtt{D}$. Evaluate $Q_i$ to 0.
4. Add the indices of variables appearing in $Q_i$ to $M_2$, $\mathtt{D} \leftarrow$ circuit obtained form $\mathtt{D}$ by evaluating $Q_i$ to 0 and $H \leftarrow$ set of all heavy gates in $\mathtt{D}$. Increment $i$.
5. **end while**

It is clear that the circuit $\mathtt{D}$ obtained after the termination of the above procedure has no heavy gates. Moreover, at no point during the execution of the procedure does the support of any monomial increase and hence the bottom support of $\mathtt{D}$ is at most $\tau$. Hence, $\mathtt{D}$ is a pruned circuit. Claim 3.2 (proved below) implies that the procedure sets at most $m$ variables to field constants i.e. $|M_2| \leq m$. If $|M_2| < m$, then we arbitrarily add an appropriate number of $i \in [3m] \setminus M_1$ to $M_2$ so that $|M_2| = m$ and for each $i \in M_2$, set $x_i = 0$ (or $y_i = 0$) if $x_i$ (or $y_i$) has not already been set to a field constant. □

**Claim 3.2.** *Let $\overline{M}_1 = [3m]\setminus M_1$. Procedure 2 sets at most $m$ many variables in $\mathbf{x}_{\overline{M}_1} \cup \mathbf{y}_{\overline{M}_1}$ to field constants.*

*Proof.* Suppose that the procedure terminates after $t$ iterations. For $1 \leq i \leq t+1$, let $H_i$ denote the set $H$ at the beginning of the $i$-th iteration. Then, $H_{t+1} = \emptyset$, $H_t \neq \emptyset$ and for any $1 \leq i \leq t$,

$$|H_{i+1}| \leq |H_i| - c_i. \tag{3.1}$$

Let the sparse polynomials during the $i-$th iteration be $Q_{i,1}, .., Q_{i,r_i}$, and $j$ be such that $\frac{c_{i,j}}{b_{i,j}} = \max_{1\leq u\leq r_i} \frac{c_{i,u}}{b_{i,u}}$. Then, $c_i = c_{i,j}$ and $b_i = b_{i,j}$. Since every heavy gate has at least $\frac{m\cdot d_{\mathbf{x}}}{2\cdot\lambda_0\cdot(\ln m)^3}$ many light sparse polynomials connected to it,

$$\begin{aligned}
|H_i| \cdot \frac{m\cdot d_{\mathbf{x}}}{2\cdot\lambda_0\cdot(\ln m)^3} &\leq c_{i,1} + \cdots c_{i,r_i} \\
&= b_{i,1}\cdot\frac{c_{i,1}}{b_{i,1}} + \cdots + b_{i,r_i}\cdot\frac{c_{i,r_i}}{b_{i,r_i}} \\
&\leq \frac{c_i}{b_i}\cdot\left(b_{i,1} + \cdots + b_{i,r_i}\right)
\end{aligned}$$

As the sum of fan-ins of all the light sparse polynomials is at most $\frac{m^2\cdot d_{\mathbf{x}}}{160\cdot\lambda_0\cdot(\ln m)^5}$ at the beginning of the procedure and at no point in time during the execution of the procedure does the fan-in of any gate of $\mathsf{D}$ increase, $\left(b_{i,1} + \cdots + b_{i,r_i}\right) \leq \frac{m^2\cdot d_{\mathbf{x}}}{160\cdot\lambda_0\cdot(\ln m)^5}$. Hence,

$$\begin{aligned}
|H_i| \cdot \frac{m\cdot d_{\mathbf{x}}}{2\cdot\lambda_0\cdot(\ln m)^3} &\leq \frac{c_i}{b_i}\cdot\left(\frac{m^2\cdot d_{\mathbf{x}}}{160\cdot\lambda_0\cdot(\ln m)^5}\right) \\
\implies |H_i| \cdot \frac{80\cdot(\ln m)^2\cdot b_i}{m} &\leq c_i
\end{aligned} \tag{3.2}$$

From (3.1) and (3.2), we get,

$$|H_{i+1}| \leq |H_i| \cdot \left(1 - \frac{80\cdot(\ln m)^2\cdot b_i}{m}\right)$$

and hence,

$$
\begin{aligned}
|H_t| &\leq |H_1| \cdot \prod_{i=1}^{t-1} \left(1 - \frac{80 \cdot (\ln m)^2 \cdot b_i}{m}\right) \\
&\leq m \cdot \prod_{i=1}^{t-1} e^{-\frac{80 \cdot (\ln m)^2 \cdot b_i}{m}} \\
&= m \cdot e^{-\frac{80 \cdot (\ln m)^2}{m} \cdot (b_1 + \cdots + b_{t-1})}
\end{aligned}
$$

where the last inequality follows from $|H_1| \leq m$ and $1 + x \leq e^x$ which is true for all $x \in \mathbb{R}$. As $|H_t| \geq 1$, $(b_1 + \cdots b_{t-1}) \leq \frac{m}{80 \cdot \ln m}$. Since $Q_t$ is a light sparse polynomial, its fan-in $b_t \leq \frac{m}{(\ln m)^2}$ and thus $b_1 + \cdots + b_t \leq \frac{m}{40 \cdot \ln m}$. Then, as the support of each monomial in $\mathsf{D}$ is upper bounded by $\tau$, the number of variables in $\mathbf{x}_{\overline{M}_1} \cup \mathbf{y}_{\overline{M}_1}$ set to field constants is at most $\tau \cdot (b_1 + \cdots b_t) \leq \lfloor 20 \cdot \ln m \rfloor \cdot \frac{m}{40 \cdot \ln m} \leq m$. □

**Remark.** Procedure 2 resembles an approximation algorithm for the Weighted Set Cover problem [Vaz01] (Section 2.1, page-16). This is no coincidence as the problem of removing heavy gates can be formulated as an instance of Weighted Set Cover with the universe being all heavy gates and with a set corresponding to every sparse polynomial $Q$. The set corresponding to $Q$ contains all heavy gates connected to $Q$ and has a cost equal to the number of monomials feeding into $Q$.

## 3.2 Analysing the measure of a pruned depth four circuit

**Lemma 3.3.** *Let $\mathsf{D}$ be a pruned depth four circuit obtained from Lemma 3.2. Also, let $d_{\mathbf{x}}, \tau, w$ be as defined earlier, $t = \left\lfloor \frac{d_{\mathbf{x}}}{(\ln m)^3} \right\rfloor$, $\delta = \frac{1}{(\ln m)^2}$, $k = \left\lfloor \frac{\delta d_{\mathbf{x}}}{t} \right\rfloor$ and $\ell = \left\lfloor \frac{m}{m^{\delta/t}+1} \right\rfloor$. Then, for sufficiently large $m$,*

$$
\mathsf{PSP}_{M,k,\ell}(\mathsf{D}) \leq s \cdot m^{O(1)} \binom{m}{\ell + 2kt\tau} \binom{\lceil \frac{w}{t} \rceil + k - 1}{k}.
$$

We prove the lemma at the end of this section. As $\mathsf{D} = T_1 + \cdots + T_s$, where $T_i$ is a product term and as $\mathsf{PSP}_{M,k,\ell}$ is sub-additive, to prove the lemma it suffices to show that for all $i \in [s]$,

$$
\mathsf{PSP}_{M,k,\ell}(T_i) \leq m^{O(1)} \binom{m}{\ell + 2kt\tau} \binom{\lceil \frac{w}{t} \rceil + k - 1}{k}.
$$

Consider any such product term $T = \prod_{i\in[a]} Q_i^{e_i}$, where $Q_i \in \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$, and since $\mathtt{D}$ is a pruned depth four circuit, $a \leq w$. Write $Q_i = Q_i' + Q_i''$, where $Q_i'$ is the sum of all monomials of $Q_i$ wherein the individual degree of every $\mathbf{x}$ variable is at most two and $Q_i'' = Q_i - Q_i'$. Then,

$$T = \prod_{i\in[a]} (Q_i' + Q_i'')^{e_i} = \prod_{i\in[a]} Q_i'^{e_i} + Q'',$$

where $Q''$ is a polynomial whose every monomial has a $\mathbf{x}$ variable with degree at least three. Thus, $\mathsf{PSP}_{M,k,\ell}(Q'') = 0$ and hence from the sub-additivity of $\mathsf{PSP}_{M,k,\ell}$ we have that

$$\mathsf{PSP}_{M,k,\ell}(T) \leq \mathsf{PSP}_{M,k,\ell}\Big(\prod_{i\in[a]} Q_i'^{e_i}\Big).$$

Let $T' = \prod_{i\in[a]} Q_i'^{e_i}$. We will now upper bound $\mathsf{PSP}_{M,k,\ell}(T')$. First, we assume without loss of generality that $a = w$ since if $a < w$ then we can multiply with additional sparse polynomials all of which are 1. Next we divide the sparse polynomials into disjoint sets such that each set (except perhaps the last) has size exactly $t$. Then, we have that

$$T' = P_1 \cdots P_{\lceil \frac{w}{t} \rceil}, \text{ where } P_i = \prod_{j=(i-1)t+1}^{\min(it,w)} Q_j'^{e_j}.$$

**Claim 3.3.** *Let $P = Q_1'^{e_1} \cdots Q_t'^{e_t}$ be one of the polynomials $P_i$. For $k \geq 0$, let $P^{(k)} := \prod_{i\in[t]} Q_i'^{\max(e_i-k,0)}$. Then, $\partial_{\mathbf{x}}^k P \subseteq \mathbb{F}\text{-span}\{\mathbf{y}_M^{\leq\infty}\mathbf{x}_M^{\leq k(2t\tau-1)} P^{(k)}\}$.*

*Proof.* We prove the claim by induction on $k$. If $k = 0$, then $\partial_{\mathbf{x}_M}^0 P = \{P\} = \{P^{(0)}\}$ and hence the claim is true. Assume that the claim is true for $k$. Let $X$ be a multilinear monomial of degree $k+1$ in $\mathbf{x}$ variables. Then $X = xX'$ where $X'$ is a multilinear monomial of degree $k$ in $\mathbf{x}$ variables and $x$ one of the $\mathbf{x}$ variables. From the induction hypothesis we have that,

$$\frac{\partial P}{\partial X'} = g \cdot P^{(k)}$$

where $g$ is a polynomial in $\mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$ with $\mathbf{x}_M$ degree of $g$ being at most $k(2t\tau - 1)$ while its $\mathbf{y}_M$ degree can be arbitrarily large.

Let $J := \{j \in [t] : e_j > k\}$. We have that,

$$\begin{aligned}
\frac{\partial P}{\partial X} &= \frac{\partial}{\partial x}\left(g \cdot P^{(k)}\right) \\
&= \frac{\partial}{\partial x}\left(g \cdot \prod_{j\in J} {Q'_j}^{e_j-k}\right) \\
&= \frac{\partial g}{\partial x} \cdot \prod_{j\in J} {Q'_j}^{e_j-k} + g \cdot \sum_{j\in J}(e_j-k) \cdot {Q'_j}^{e_j-k-1} \cdot \frac{\partial Q'_j}{\partial x} \cdot \prod_{i\in J\setminus\{j\}} {Q'_i}^{e_i-k} \\
&= \left(\frac{\partial g}{\partial x} \cdot \prod_{j\in J} Q'_j + g \cdot \sum_{j\in J}(e_j-k) \cdot \frac{\partial Q'_j}{\partial x} \cdot \prod_{i\in J\setminus\{j\}} Q'_i\right) \cdot \prod_{j\in J} {Q'_j}^{e_j-k-1}
\end{aligned}$$

Observe that as $\mathtt{D}$ is a pruned depth four circuit, the support of all monomials of $Q'_j$ is upper bounded by $\tau$ and as in any monomial the individual degree of any $\mathbf{x}$ variable is at most two, $\deg_{\mathbf{x}}(Q'_j) \leq 2\tau$. Also, $|J| \leq t$ and hence

$$\deg_{\mathbf{x}}\left(\frac{\partial g}{\partial x} \cdot \prod_{j\in J} Q'_j + g \cdot \sum_{j\in J}(e_j-k) \cdot \frac{\partial Q'_j}{\partial x} \cdot \prod_{i\in J\setminus\{j\}} Q'_i\right) \leq (k+1)(2t\tau-1).$$

As $\prod_{j\in J} {Q'_j}^{e_j-k-1} = P^{(k+1)}$, the claim is true for $k+1$. □

*Proof of Lemma 3.3.* Recall that it is enough to show the following

$$\mathsf{PSP}_{M,k,\ell}(T') \leq m^{O(1)}\binom{m}{\ell+2kt\tau}\binom{\lceil \frac{w}{t} \rceil + k - 1}{k},$$

where $T' = P_1 \cdots P_{\lceil \frac{w}{t} \rceil}$. Let $v = \lceil \frac{w}{t} \rceil$. Now,

$$\begin{aligned}
\partial^k_{\mathbf{x}} T' &\subseteq \mathbb{F}\text{-span}\left\{\partial^{k_1}_{\mathbf{x}} P_1 \cdots \partial^{k_v}_{\mathbf{x}} P_v \;:\; k_1 + \cdots + k_v = k\right\} \\
&\subseteq \mathbb{F}\text{-span}\left\{\mathbf{y}_M^{\leq \infty}\mathbf{x}_M^{\leq k_1(2t\tau-1)} P_1^{(k_1)} \cdots \mathbf{y}_M^{\leq \infty}\mathbf{x}_M^{\leq k_v(2t\tau-1)} P_v^{(k_v)} \;:\; k_1 + \cdots + k_v = k\right\} \\
&\subseteq \mathbb{F}\text{-span}\left\{\mathbf{y}_M^{\leq \infty}\mathbf{x}_M^{\leq k(2t\tau-1)} P_1^{(k_1)} \cdots P_v^{(k_v)} \;:\; k_1 + \cdots + k_v = k\right\},
\end{aligned}$$

where the second to last inclusion follows from Claim 3.3. Hence,

$$\mathbf{x}^{\ell}_M \partial^k_{\mathbf{x}} T' \subseteq \mathbb{F}\text{-span}\left\{\mathbf{y}_M^{\leq \infty}\mathbf{x}_M^{\leq \ell + k(2t\tau-1)} P_1^{(k_1)} \cdots P_v^{(k_v)} \;\; k_1 + \cdots + k_v = k\right\}.$$

In other words, the space of shifted partials of $T'$ is contained in the $\mathbb{F}$-span of polynomials of the form $Y \cdot X \cdot P_1^{(k_1)} \cdots P_v^{(k_v)}$ where $Y$ is a monomial in $\mathbf{y}_M$ variables and $X$ is a monomial in $\mathbf{x}_M$ variables of degree at most $\ell + k(2t\tau - 1)$. Let us analyse the effect of the operations $\sigma_{\mathbf{y}_M=1}$, $[\cdot]_m$ and $\pi_{\mathbf{x}}$ on one such polynomial. We will assume that $\deg_{\mathbf{y}}(Y) \leq m$ and $X$ is multilinear for otherwise the polynomial will vanish after the operations are applied. Then, we have that,

$$\sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(Y \cdot X \cdot P_1^{(k_1)} \cdots P_v^{(k_v)}\right)\right]_m\right) = X \cdot \sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(\sigma_{\mathrm{Supp}(X)=0}\left(P_1^{(k_1)} \cdots P_v^{(k_v)}\right)\right)\right]_{m-\deg(Y)}\right).$$

Thus,

$$\sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(\mathbf{x}_M^{\ell}\partial_{\mathbf{x}}^{k}T'\right)\right]_m\right) \subseteq \mathbb{F}\text{-span}\Bigg\{X \cdot \sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(\sigma_{\mathrm{Supp}(X)=0}\left(P_1^{(k_1)} \cdots P_v^{(k_v)}\right)\right)\right]_i\right) : X \text{ is a multilinear monomial in } \mathbf{x}_M \text{ variables, } \deg(X) \text{ is at most } \ell + k(2t\tau - 1), 0 \leq i \leq m \text{ and } k_1 + \cdots + k_v = k\Bigg\}.$$

Once we fix $i$, $X$, and $k_1, ..., k_v$, $X \cdot \sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(\sigma_{\mathrm{Supp}(X)=0}\left(P_1^{(k_1)} \cdots P_v^{(k_v)}\right)\right)\right]_i\right)$ is fixed. So,

$$\begin{aligned}
\mathsf{PSP}_{M,k,\ell}(T') &= \dim\left\langle \sigma_{\mathbf{y}_M=1}\left(\left[\pi_{\mathbf{x}}\left(\mathbf{x}_M^{\ell}\partial_{\mathbf{x}}^{k}T'\right)\right]_m\right)\right\rangle \\
&\leq (m+1)\cdot \sum_{j=0}^{\ell+k(2t\tau-1)} \binom{m}{j}\binom{v+k-1}{k} \\
&\leq (m+1)\cdot(\ell+2kt\tau)\cdot \binom{m}{\ell+2kt\tau}\binom{v+k-1}{k} \\
&= m^{O(1)}\cdot \binom{m}{\ell+2kt\tau}\binom{\lceil \frac{w}{t}\rceil+k-1}{k},
\end{aligned}$$

where the second last inequality follows from Claim 3.4.

□

**Claim 3.4.** *Let $\ell, k, t$ and $\tau$ be as defined earlier. Then, $\ell + 2kt\tau < \frac{m}{2}$.*

*Proof.* We will show that the ratio $\frac{\frac{m}{2}-2kt\tau}{\ell} > 1$. Putting the values of $k$ and $\ell$,

$$\frac{\frac{m}{2} - 2kt\tau}{\ell} = \frac{\frac{m}{2} - 2\left\lfloor \frac{\delta d_{\mathbf{x}}}{t} \right\rfloor t\tau}{\left\lfloor \frac{m}{m^{\delta/t}+1} \right\rfloor}$$
$$\geq \left(\frac{1}{2} - \frac{2\delta d_{\mathbf{x}}\tau}{m}\right)(m^{\delta/t}+1).$$

So, we need to show that

$$\frac{1}{\frac{1}{2} - \frac{2\delta d_{\mathbf{x}}\tau}{m}} < m^{\delta/t} + 1 \iff \frac{1}{\frac{1}{2} - \frac{2\delta d_{\mathbf{x}}\tau}{m}} - 1 < m^{\delta/t}$$
$$\iff \frac{1 + \frac{4\delta d_{\mathbf{x}}\tau}{m}}{1 - \frac{4\delta d_{\mathbf{x}}\tau}{m}} < m^{\delta/t}.$$

For large enough $m$, $\frac{4\delta d_{\mathbf{x}}\tau}{m} \leq \frac{1}{2}$. Using $1 + x \leq e^x$, which holds for all $x \in \mathbb{R}$, and $\frac{1}{1-x} \leq e^{2x}$, which holds for $0 \leq x \leq \frac{1}{2}$ we get:

$$\frac{1 + \frac{4\delta d_{\mathbf{x}}\tau}{m}}{1 - \frac{4\delta d_{\mathbf{x}}\tau}{m}} \leq e^{\frac{12\delta d_{\mathbf{x}}\tau}{m}}.$$

So showing that $e^{\frac{12\delta d_{\mathbf{x}}\tau}{m}} < m^{\delta/t}$ would suffice. Now,

$$e^{\frac{12\delta d_{\mathbf{x}}\tau}{m}} < m^{\delta/t} \iff e^{\frac{12 d_{\mathbf{x}} t\tau}{m}} < m.$$

Putting the values of $d_{\mathbf{x}}, t$ an $\tau$, we get that $\frac{12 d_{\mathbf{x}} t\tau}{m} = \frac{12 d_{\mathbf{x}} \left\lfloor \frac{d_{\mathbf{x}}}{(\ln m)^3} \right\rfloor \lfloor 20 \ln m \rfloor}{m} \leq \frac{12 d_{\mathbf{x}}^2 \cdot 20 \ln m}{m(\ln m)^3} = \Theta\left(\frac{m}{m(\ln m)^2(\ln m)^2}\right) = \Theta\left(\frac{1}{(\ln m)^4}\right) = o(1)$ as $d_{\mathbf{x}} = \Theta\left(\frac{\sqrt{m}}{\ln m}\right)$. Thus $e^{\frac{12 d_{\mathbf{x}} t\tau}{m}} < m$. □

# Chapter 4

# An Explicit Polynomial Family with High Measure

In this chapter we construct a polynomial family $\{f_n\}_{n\geq 1}$ with a "large" $\mathsf{PSP}_{M,k,\ell}(\cdot)$ measure - this forms the second part of the proof of Theorem 1.1. We also prove Theorem 1.1 in this chapter. The contents of this chapter up to and including Section 4.1 are from our work [GST20]. Section 4.2 contains a proof Lemma 4.1, which was proved in [KLSS17, KS16a, Sha17]; we provide its proof for the sake of completeness.

We now describe the family $\{f_n\}_{n\geq 1}$, whose $n$-th member $f_n$ is a polynomial in variables $\mathbf{x} = \{x_1, ..., x_{3m}\}$ and $\mathbf{y} = \{y_1, ..., y_{3m}\}$, where $m \in \left[\frac{n}{2}, 2n\right]$ will be fixed later.

$$f_n := \sum_{S\subseteq[3m],|S|=m} \left(\prod_{i\in S} y_i\right) \cdot \mathsf{NW}_r(\mathbf{x}_S),$$

where $\mathsf{NW}_r$ is a variant of the Nisan-Wigderson design polynomial (introduced in [KSS14]), the construction of which is described later and $r$ is a parameter fixed in this construction. Note that $\{f_n\}_{n\geq 1}$ is in VNP. Given a monomial, in order to find its coefficient in $f_n$, we first check if the monomial is multilinear and of degree $m$ in $\mathbf{y}$ variables. If it is so and $S$ is the set of the indices of the $m$ many $\mathbf{y}$ variables in the monomial then simply return the coefficient of the part of the monomial in $\mathbf{x}$ variables in $\mathsf{NW}_r(\mathbf{x}_S)$ – this can be done as the Nisan-Wigderson polynomial family is in VNP.

Let $M_1$ and $M_2$ be as in Chapter 3 and $M = [3m] \setminus (M_1 \cup M_2)$. Let $f_1$ be the polynomial computed by the pruned circuit $\mathsf{D}$, which is obtained from $f = f_n$ by setting the variables $\mathbf{x}_{\overline{M}}$ and $\mathbf{y}_{\overline{M}}$ to field constants as in Section 3.1. Let us now see how $\mathsf{PSP}_{M,k,\ell}(f_1)$ is related to

$\dim \left\langle \pi_{\mathbf{x}} \left( \mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} \mathsf{NW}_r \right) \right\rangle$.

**Lemma 4.1.** *Let $f_1$ be as defined above. Then,* $\mathsf{PSP}_{M,k,\ell}(f_1) = \dim \left\langle \pi_{\mathbf{x}} \left( \mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} \mathsf{NW}_r(\mathbf{x}_M) \right) \right\rangle$.

*Proof.* The proof follows easily from the following two observations:

1. The two operations in $\mathbf{y}$ variables and the three operations in $\mathbf{x}$ variables (in the definition of $\mathsf{PSP}_{M,k,\ell}$) commute. That is, we have

$$\sigma_{\mathbf{y}_M=1} \left( \left[ \pi_{\mathbf{x}} (\mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} f_1) \right]_m \right) = \pi_{\mathbf{x}} \left( \mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} \left( \sigma_{\mathbf{y}_M=1} \left( [f_1]_m \right) \right) \right).$$

2. $f_1 = \left( \prod_{i \in M} y_i \right) \cdot \mathsf{NW}_r(\mathbf{x}_M) + f'$, where $f' \in \mathbb{F}[\mathbf{x}_M, \mathbf{y}_M]$ and $\deg_{\mathbf{y}}(f') < m$.

From these observations we have that

$$\begin{aligned}
&\mathsf{PSP}_{M,k,\ell}(f_1) \\
&= \dim \left\langle \sigma_{\mathbf{y}_M=1} \left( \left[ \pi_{\mathbf{x}} (\mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} f_1) \right]_m \right) \right\rangle \\
&= \dim \left\langle \pi_{\mathbf{x}} \left( \mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} \left( \sigma_{\mathbf{y}_M=1} \left( \left[ \left( \prod_{i \in M} y_i \right) \cdot \mathsf{NW}_r(\mathbf{x}_M) + f' \right]_m \right) \right) \right) \right\rangle \\
&= \dim \left\langle \pi_{\mathbf{x}} \left( \mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} \mathsf{NW}_r(\mathbf{x}_M) \right) \right\rangle.
\end{aligned}$$

The last equality follows from the fact that $\left( \sigma_{\mathbf{y}_M=1} \left( [(\prod_{i \in M} y_i) \cdot \mathsf{NW}_r(\mathbf{x}_M) + f']_m \right) \right) = \mathsf{NW}_r(\mathbf{x}_M)$. □

**Construction of $\mathsf{NW}_r$.** Let $d_{\mathbf{x}} = \left\lfloor \frac{\sqrt{n}}{\ln n} \right\rfloor$. Pick an $\alpha$ such that $d_{\mathbf{x}} \left\lceil d_{\mathbf{x}}^{1+\alpha} \right\rceil \leq n \leq 2 d_{\mathbf{x}} \left\lceil d_{\mathbf{x}}^{1+\alpha} \right\rceil$; this forces $\alpha$ to be $\Theta(\frac{\ln \ln n}{\ln n})$. Let $q$ be a prime number between $\left\lceil d_{\mathbf{x}}^{1+\alpha} \right\rceil$ and $2 \left\lceil d_{\mathbf{x}}^{1+\alpha} \right\rceil$ – such a prime exists [Erd32] – and let $m = d_{\mathbf{x}} q$. Thus, $d_{\mathbf{x}} \left\lceil d_{\mathbf{x}}^{1+\alpha} \right\rceil \leq m \leq 2 d_{\mathbf{x}} \left\lceil d_{\mathbf{x}}^{1+\alpha} \right\rceil$ and hence $\frac{n}{2} \leq m \leq 2n$; moreover, it can be easily verified that $d_{\mathbf{x}} \in \left[ \frac{\sqrt{m}}{2\sqrt{2} \cdot \ln m}, \frac{2\sqrt{2} \cdot \sqrt{m}}{\ln m} \right]$; both being as required in Section 1.2. Also notice that this means $q = \Theta(\sqrt{n} \ln n)$. Let $\beta = \frac{1}{\ln m}$ and $r = \left\lfloor \frac{\alpha+\beta}{2(1+\alpha)} d_{\mathbf{x}} \right\rfloor - 1$, $\mathbf{u} = (u_{1,1}, \ldots, u_{1,q}, \ldots, u_{d_{\mathbf{x}},1}, \ldots, u_{d_{\mathbf{x}},q})$ and define

$$\mathsf{NW}_r(\mathbf{u}) := \sum_{h(z) \in \mathbb{F}_q[z],\, \deg(h) \leq r} u_{1,h(1)} \cdots u_{d_{\mathbf{x}},h(d_{\mathbf{x}})}.$$

A lower bound on $\dim \left\langle \pi_{\mathbf{x}} \left( \mathbf{x}_M^{\ell} \partial_{\mathbf{x}}^{k} \mathsf{NW}_r \right) \right\rangle$ was proved in [KS16a, Sha17]. Their analysis continues to hold for our choice of parameters – which only slightly differ from the parameters in [Sha17]. Moreover, while they prove this lower bound over fields of characteristic

zero, the same proof also works if the characteristic is greater than $q^{(r+1)\cdot\min\left\{\binom{m}{k}\binom{m}{\ell},\binom{m}{\ell-d_{\mathbf{x}}-k}\right\}}$. For the sake of completeness, we provide a proof of the following lemma in Section .

**Lemma 4.2** (Lemma 5.2 of [KS16a], Lemma 4.1 of [Sha17])**.**

$$\dim\left\langle \pi_{\mathbf{x}}\left(\mathbf{x}_M^{\ell}\partial_{\mathbf{x}}^{k}\,\mathsf{NW}_r(\mathbf{x}_M)\right)\right\rangle \geq \frac{1}{m^{O(1)}}\min\left\{\frac{1}{4^k}\cdot\binom{m}{\ell}\binom{m}{k},\binom{m}{\ell+d_{\mathbf{x}}-k}\right\}.$$

Hence, from Lemmas 4.1 and 4.2 we get

**Lemma 4.3.**

$$\mathsf{PSP}_{M,k,\ell}(f_1) \geq \frac{1}{m^{O(1)}}\min\left\{\frac{1}{4^k}\cdot\binom{m}{\ell}\binom{m}{k},\binom{m}{\ell+d_{\mathbf{x}}-k}\right\}.$$

## 4.1 Proof of Theorem 1.1

Before proving the theorem, let us first justify the assumption that $\mathbb{F}$ is an algebraically closed field that we made in Chapter 3. Suppose not. Then, let $\overline{\mathbb{F}}$ be its algebraic closure. Since $\mathsf{C}$ is also a circuit over $\overline{\mathbb{F}}$ and $f_n$ a polynomial over $\overline{\mathbb{F}}$, we can make all arguments assuming the underlying field to be $\overline{\mathbb{F}}$. Since the size of a circuit does not depend on the underlying field, the lower bound so obtained will continue to hold when we treat $\mathsf{C}$ as a circuit over $\mathbb{F}$.

First we will prove a lower bound on the number of wires of $\mathsf{C}$. If the number of monomials in $\mathsf{C}$ is $\left\lfloor\frac{m^2 d_{\mathbf{x}}}{(\ln m)^5}\right\rfloor$ then there is nothing to prove. Otherwise from Lemma 3.1, we can obtain a circuit $\mathsf{C}_1$ such that the support of all the monomials of $\mathsf{C}_1$ is at most $\tau = \lfloor 20\ln m\rfloor$, the number of gates in $\mathsf{C}_1$ is at most the number of gates in $\mathsf{C}$ and the fan-in of each gate in $\mathsf{C}_1$ is upper bounded by the fan-in of the corresponding gate in $\mathsf{C}$. Then, if $\mathsf{C}_1$ does not satisfy the hypothesis of Lemma 3.2, the size of $\mathsf{C}_1$ and hence the size of $\mathsf{C}$ is at least $\Omega\left(\frac{m^2 d_{\mathbf{x}}}{(\ln m)^5}\right)$. Otherwise, we can obtain a pruned circuit $\mathsf{D}$ such that the top fan-in and the bottom support of $\mathsf{D}$ are upper bounded by the top fan-in and bottom support of $\mathsf{C}_1$ and so proving a lower bound on the top fan-in of $\mathsf{D}$ would suffice.

As $\mathsf{D}$ computes $f_1$, $\mathsf{PSP}_{M,k,\ell}(\mathsf{D}) = \mathsf{PSP}_{M,k,\ell}(f_1)$. Lemma 3.3 and 4.3 imply

$$
\begin{aligned}
s &\geq \frac{\frac{1}{m^{O(1)}} \min\left\{\frac{1}{4^k}\cdot\binom{m}{\ell}\binom{m}{k}, \binom{m}{\ell+d_{\mathbf{x}}-k}\right\}}{m^{O(1)}\cdot\binom{m}{\ell+2kt\tau}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \\
&\geq \frac{\frac{1}{m^{O(1)}} \min\left\{\frac{1}{4^k}\cdot\binom{m}{\ell}\binom{m}{k}, \binom{m}{\ell+d_{\mathbf{x}}-k}\right\}}{m^{O(1)}\cdot\binom{m}{\ell+2kt\tau}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \\
&\geq \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot\frac{\binom{m}{\ell+1}}{\binom{m}{\ell+2kt\tau+1}}, \frac{\binom{m}{\ell+d_{\mathbf{x}}-k}}{\binom{m}{\ell+2kt\tau}}\right\} \\
&= \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot\frac{(m-\ell-2kt\tau-1)!}{(m-\ell-1)!}\cdot\frac{(\ell+2kt\tau+1)!}{(\ell+1)!},\right. \\
&\qquad\qquad \left.\frac{(m-\ell-2kt\tau)!}{(m-\ell-d_{\mathbf{x}}+k)!}\cdot\frac{(\ell+2kt\tau)!}{(\ell+d_{\mathbf{x}}-k)!}\right\} \\
&= \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot e^{(-2kt\tau)\ln\frac{m-\ell-1}{\ell+1}\pm o(1)}, e^{(d_{\mathbf{x}}-2kt\tau-k)\ln\frac{m-\ell}{\ell}\pm o(1)}\right\}
\end{aligned}
$$

(Using Proposition 2.2.)

$$
\begin{aligned}
&\geq \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot\left(\frac{m}{\ell+1}-1\right)^{-2kt\tau}, \left(\frac{m}{\ell}-1\right)^{(d_{\mathbf{x}}-2kt\tau-k)}\right\} \\
&= \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot\left(\frac{m}{\left\lfloor\frac{m}{m^{\delta/t}+1}\right\rfloor+1}-1\right)^{-2kt\tau}, \left(\frac{m}{\left\lfloor\frac{m}{m^{\delta/t}+1}\right\rfloor}-1\right)^{(d_{\mathbf{x}}-2kt\tau-k)}\right\} \\
&\geq \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot\left(\frac{m}{\frac{m}{m^{\delta/t}+1}}-1\right)^{-2kt\tau}, \left(\frac{m}{\frac{m}{m^{\delta/t}+1}}-1\right)^{(d_{\mathbf{x}}-2kt\tau-k)}\right\} \\
&\geq \frac{1}{m^{O(1)}\binom{\lceil\frac{w}{t}\rceil+k-1}{k}} \min\left\{\frac{\binom{m}{k}}{4^k}\cdot m^{-2k\delta\tau}, m^{(1-2\delta\tau-\frac{\delta}{t})k}\right\}
\end{aligned}
$$

Since $\frac{\binom{m}{k}}{4^k}\cdot m^{-2k\delta\tau} = \frac{\binom{m}{k}}{4^k\cdot m^{(1-\frac{\delta}{t})k}}\cdot m^{(1-2\delta\tau-\frac{\delta}{t})k} \leq (\frac{em}{k})^k\cdot\frac{m^{\frac{\delta k}{t}}}{4^k m^k}\cdot m^{(1-2\delta\tau-\frac{\delta}{t})k}$. For our choice of parameters $\delta, k$ and $t$, $m^{\frac{\delta k}{t}} = O(1)$. Hence, $\frac{\binom{m}{k}}{4^k}\cdot m^{-2k\delta\tau} \leq m^{(1-2\delta\tau-\frac{\delta}{t})k}$ and thus,

$$
s \geq \frac{1}{m^{O(1)}}\cdot\frac{\binom{m}{k}\cdot m^{-2k\delta\tau}}{4^k\cdot\binom{\lceil\frac{w}{t}\rceil+k-1}{k}}
$$

$$
\begin{aligned}
&\geq \frac{1}{m^{O(1)}} \cdot \left( \frac{m \cdot k}{4e \cdot k \cdot m^{2\delta\tau} \cdot (\frac{w}{t} + k)} \right)^k && \text{(Using Proposition 2.1.)} \\
&\geq \frac{1}{m^{O(1)}} \cdot \left( \frac{m \cdot t}{8e \cdot m^{2\delta\tau} \cdot w} \right)^k && \text{(Since } kt \leq w = \left\lfloor \tfrac{md_{\mathbf{x}}}{\lambda_0 \cdot (\ln m)^3} \right\rfloor .) \\
&= \frac{1}{m^{O(1)}} \cdot \left( \frac{m \cdot \left\lfloor \frac{d_{\mathbf{x}}}{(\ln m)^3} \right\rfloor}{8e \cdot m^{2\delta\tau} \cdot \left\lfloor \frac{md_{\mathbf{x}}}{\lambda_0 \cdot (\ln m)^3} \right\rfloor} \right)^k \\
&\geq \frac{1}{m^{O(1)}} \cdot \left( \frac{m \cdot \frac{d_{\mathbf{x}}}{(\ln m)^3}}{16e \cdot m^{2\delta\tau} \cdot \frac{md_{\mathbf{x}}}{\lambda_0 \cdot (\ln m)^3}} \right)^k \\
&\geq \frac{1}{m^{O(1)}} \cdot \left( \frac{\lambda_0}{16e \cdot m^{2 \cdot \frac{1}{(\ln m)^2} \cdot \lfloor 20 \ln m \rfloor}} \right)^{\ln m} && \text{(Since } k \geq \lfloor \ln m \rfloor .) \\
&= \frac{1}{m^{O(1)}} \cdot \left( \frac{\lambda_0}{16e \cdot e^{O(1)}} \right)^{\ln m} \\
&= \omega \left( \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \right),
\end{aligned}
$$

if we choose $\lambda_0$ to be a large enough constant.

Now let us prove the lower bound on the number of gates. Notice that if the circuit $\mathtt{C}$ computing $f$ has a heavy gate as defined in Chapter 3 then we are done. So assume that it does not have any heavy gates. Now, if the number of monomials in $\mathtt{C}$ is $\left\lfloor \frac{m^2 d_{\mathbf{x}}}{(\ln m)^5} \right\rfloor$ then there is nothing to prove. Otherwise from Lemma 3.1, we can obtain a circuit $\mathtt{C}_1$ such that the support of all the monomials of $\mathtt{C}_1$ is at most $\tau = \lfloor 20 \ln m \rfloor$, the number of gates in $\mathtt{C}_1$ is at most the number of gates in $\mathtt{C}$ and the fan-in of each gate in $\mathtt{C}_1$ is upper bounded by the fan-in of the corresponding gate in $\mathtt{C}$. Obtain a circuit $\mathtt{D}$ from $\mathtt{C}_1$ by picking a set $M_2 \subseteq [3m] \setminus M_1$ (where $M_1$ is as in Lemma 3.1), $|M_2| = m$ arbitrarily and setting variables in $\mathbf{x}_{M_2}$ and $\mathbf{y}_{M_2}$ to 0 (notice that the top fan-in and bottom support of $\mathtt{D}$ are upper bounded by the top fan-in and bottom support of $\mathtt{C}_1$ ). Now $\mathtt{D}$ computes $f_1$ and just as it was done in the preceding paragraph, we can show that the top fan-in of $\mathtt{D}$ is $\omega\left(\frac{m^2 d_{\mathbf{x}}}{(\ln m)^5}\right)$. However, we only get an $\Omega\left(\frac{md_{\mathbf{x}}}{(\ln m)^3}\right)$ lower bound on the number of gates since the definition of a heavy gate is the bottleneck.

## 4.2 Proof of Lemma 4.1

For the sake of completeness, we now give a proof of Lemma 4.2 by replicating the analysis in [KS16a] and [Sha17]. We first construct a matrix $N$ with $0,1$ entries whose rank is a lower bound on $\dim\left\langle \pi_{\mathbf{x}}\left(\mathbf{x}_M^{\ell}\partial_{\mathbf{x}}^{k}\,\mathsf{NW}_r(\mathbf{x}_M)\right)\right\rangle$ and then derive a lower bound on the rank $N$ using a lemma in real matrix analysis. Before we describe the construction of $N$, let us establish some conventions.

By simply re-indexing the variable set $\mathbf{x}_M$, we can assume that $M=[m]$. For the sake of making the notation a little simple, we will drop the subscript $M$ from $\mathbf{x}_M$. Notice that there is a $1-1$ correspondence between the indices of variables $\mathbf{x}$ and the set $[q]\times[d_{\mathbf{x}}]\equiv[m]$. These is also a $1-1$ correspondence between the monomials in $NW_r$ and the set $\binom{[q]\times[d_{\mathbf{x}}]}{d_{\mathbf{x}}}\equiv\binom{[m]}{d_{\mathbf{x}}}$ as $NW_r$ is a homogeneous and multilinear polynomial of degree $d_{\mathbf{x}}$. Moreover, every monomial in $NW_r$ corresponds to a unique polynomial of degree at most $r$ in $\mathbb{F}_q[z]$. Because of these reasons, going forward, we will represent a monomial of $NW_r$ using an either an element of the set $\binom{[m]}{d_{\mathbf{x}}}$ or a polynomial of degree at most $r$ in $\mathbb{F}_q[z]$.

**Definition 4.1** (Support of $NW_r$)**.** *We define the support of $NW_r$ - denoted be* $\mathrm{Supp}(NW_r)$ *- as follows:*

$$\mathrm{Supp}(NW_r):=\left\{D\in\binom{[m]}{d_{\mathbf{x}}}:\prod_{i\in[m]}x_i \text{ is a monomial in } NW_r\right\}$$

**Construction of the matrix** $N$**.** The rows of $N$ are indexed by ordered pairs of the form $(A,C)\in\binom{[m]}{\ell}\times\binom{[m]}{k}$ such that $A\cap C=\emptyset$ and its columns are indexed by sets $S\in\binom{[m]}{\ell+d_{\mathbf{x}}-k}$. The row indexed by $(A,C)$ corresponds to the polynomial

$$g_{A,C}=\left(\prod_{i\in A}x_i\right)\cdot\sigma_{\mathbf{x}_A=0}\left(\frac{\partial}{\partial(\prod_{j\in C}x_j)}NW_r\right).$$

The $S$-th entry of the row $(A,C)$ is the coefficient of the monomial $\prod_{i\in S}x_i$ in $g_{A,C}$. Note that as monomials in $NW_r$ have $0,1$ coefficients, every entry of $N$ is either 0 or 1. Now as

$$\dim\left\langle\pi_{\mathbf{x}}\left(\mathbf{x}^{\ell}\partial_{\mathbf{x}}^{k}\,\mathsf{NW}_r(\mathbf{x})\right)\right\rangle=\dim\left\langle\left\{g_{A,C}:A\in\binom{[m]}{\ell},C\in\binom{[m]}{k}\right\}\right\rangle$$

we have the following proposition.

**Proposition 4.1.** $rank(N)\leq\dim\left\langle\pi_{\mathbf{x}}\left(\mathbf{x}^{\ell}\partial_{\mathbf{x}}^{k}\,\mathsf{NW}_r(\mathbf{x})\right)\right\rangle$.

For sets $A, B$ we define the operations $A \setminus\setminus B$ and $A \uplus B$ as follows

1.
$$A \setminus\setminus B := \begin{cases} A \setminus B & B \subseteq A \\ \texttt{InvalidSet} & \text{otherwise} \end{cases}$$

2.
$$A \uplus B := \begin{cases} A \cup B & A \cap B = \emptyset \\ \texttt{InvalidSet} & \text{otherwise} \end{cases}$$

We label the $((A, C), S)$-th entry of $N$ by the set $D = (S \setminus\setminus A) \uplus C$. It is not too hard to see that $D \in \text{Supp}(NW_r)$ if and only if $N_{((A,C),S)} = 1$.

Note that while $N$ is a matrix over some characteristic 0 field $\mathbb{F}$, since it has $0, 1$ entries, we can treat it as a matrix over the field $\mathbb{R}$ of real numbers. As the determinant of any real $0, 1$ matrix is an integer and as any characteristic 0 field contains the field of rational numbers $\mathbb{Q}$ as a sub-field, the determinant of any sub-matrix of $N$ will be the same over $\mathbb{R}$ and $\mathbb{F}$. Thus the rank of $N$ over $\mathbb{F}$ is the same as its rank over $\mathbb{R}$. Because of this reason, from now on we will treat $N$ as a real matrix. We will now focus our attention on deriving a lower bound on $rank(N)$ using the notion of surrogate rank.

**Deriving a lower bound on** $rank(N)$**.** Let $B = N^T N$. Then $B$ is a real positive semi-definite matrix and it is easy to show that,

**Proposition 4.2.** *Over any field* $\mathbb{F}$, $rank(B) \leq rank(N)$. *Moreover, over the field of real numbers* $\mathbb{R}$, $rank(B) = rank(N)$.

So in order to lower bound $rank(N)$, we can simply lower bound $rank(B)$. Let us now define the surrogate rank of $B$.

**Definition 4.2.** *The surrogate rank of* $B$ *- denoted* $\text{SurRank}(B)$ *- is the ratio* $\frac{\text{Tr}(B)^2}{\text{Tr}(B^2)}$.

The following lemma gives us a way of obtaining a lower bound on $rank(B)$ by lower bounding $\text{SurRank}(B)$.

**Lemma 4.4** ([Alo09])**.** $\text{SurRank}(B) \leq rank(B)$.

The above lemma can be proved by using the Cauchy-Schwarz inequality the vector of non-zero eigenvalues of $B$.

In what follows, we will show that $\text{SurRank}(B) \geq \frac{1}{m^{O(1)}} \min\{\frac{1}{4^k} \cdot \binom{m}{\ell}\binom{m}{k}, \binom{m}{\ell + d_{\mathbf{x}} - k}\}$ by first deriving a lower bound on $\text{Tr}(B)$ and then an upper bound on $\text{Tr}(B^2)$.

### 4.2.1 Lower bound on $\mathrm{Tr}(B)$

**Claim 4.1.** $\mathrm{Tr}(B) = q^{r+1} \cdot \binom{d_{\mathbf{x}}}{k} \cdot \binom{m-d_{\mathbf{x}}}{\ell}$.

*Proof.* Since all entries of $N$ are either 0 or 1, we have that

$$\mathrm{Tr}(B) = \mathrm{Tr}(N^T N) = \text{number of non zero entries in } N.$$

The number of non-zero entries in $N$ is just the sum over all $D \in \mathrm{Supp}(NW_r)$ of the number of cells of $N$ labelled by $D$. The $((A, C), S)$-th entry of $N$ is labelled by $D$ if and only if $D = (S \setminus\setminus A) \uplus C$ i.e. $S = (D \setminus\setminus C) \uplus A$. Then, as $A \cap C = \emptyset$, the number of cells labelled by $D$ is $\binom{d_{\mathbf{x}}}{k} \cdot \binom{m-d_{\mathbf{x}}}{\ell}$. Since $|\mathrm{Supp}(NW_r)| = q^{r+1}$, the claim follows. □

### 4.2.2 Upper bound on $\mathrm{Tr}(B^2)$

The following proposition easily follows from the definition of $B$.

**Proposition 4.3.**

$$\mathrm{Tr}(B^2) = \sum N_{(A_1,C_1),S_1} \cdot N_{(A_1,C_1),S_2} \cdot N_{(A_2,C_2),S_1} \cdot N_{(A_2,C_2),S_2}.$$

*where the sum is over all tuples* $((A_1, C_1), (A_2, C_2), S_1, S_2)$ *such that* $(A_1, C_1), (A_2, C_2) \in \binom{[m]}{\ell} \times \binom{[m]}{k}$, $A_1 \cap C_1 = \emptyset$, $A_2 \cap C_2 = \emptyset$ *and* $S_1, S_2 \in \binom{[m]}{\ell+d_{\mathbf{x}}-k}$.

Let us define the notion of a box which shall use in the calculations. For any pair of row indices $(A_1, C_1), (A_2, C_2) \in \binom{[m]}{\ell} \times \binom{[m]}{k}$ and any pair of columns indices $S_1, S_2 \in \binom{[m]}{\ell+d_{\mathbf{x}}-k}$, we define the box

$$\mathbf{b} = \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

to be the ordered tuple

$$(((A_1, C_1), S_1), ((A_1, C_1), S_2), ((A_2, C_2), S_1), ((A_2, C_2), S_2))$$

of the cells of $N$. As every entry of $N$ is either 0 or 1,

$$\mathrm{Tr}(B^2) = \text{number of boxes } \mathbf{b} \text{ whose all four entries are 1.}$$

Now, all the four entries of

$$\mathbf{b} = \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2))$$

can be non-zero only if all four entries of $\mathbf{b}$ are labelled by sets in $\mathrm{Supp}(NW_r)$ i.e. only if

$$(S_1 \setminus\setminus A_1) \uplus C_1, (S_2 \setminus\setminus A_1) \uplus C_1, (S_1 \setminus\setminus A_2) \uplus C_2, (S_2 \setminus\setminus A_2) \uplus C_2 \in \mathrm{Supp}(NW_r).$$

For a box

$$\mathbf{b} = \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2),$$

$\mathrm{labels}(\mathbf{b})$ is the tuple of labels of entries in $\mathbf{b}$,

$$\mathrm{labels}(\mathbf{b}) = ((S_1 \setminus\setminus A_1) \uplus C_1, (S_2 \setminus\setminus A_1) \uplus C_1, (S_1 \setminus\setminus A_2) \uplus C_2, (S_2 \setminus\setminus A_2) \uplus C_2).$$

**Proposition 4.4.** $\mathrm{Tr}(B^2)$ *is the number of boxes*

$$\mathbf{b} = \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

*such that all four labels in* $\mathrm{labels}(\mathbf{b})$ *are valid sets in* $\mathrm{Supp}(NW_r)$.

Now we will compute $\mathrm{Tr}(B^2)$, by counting the number of boxes whose all four labels are sets in $\mathrm{Supp}(NW_r)$. To do this, we need to analyse the structure of such a box

$$\mathbf{b} = \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_2).$$

Let $\mathrm{labels}(\mathbf{b}) = (D_1, D_2, D_3, D_4)$. Then,

$$\begin{aligned} D_1 &= (S_1 \setminus\setminus A_1) \uplus C_1, & D_2 &= (S_2 \setminus\setminus A_1) \uplus C_1, \\ D_3 &= (S_1 \setminus\setminus A_2) \uplus C_2, & D_4 &= (S_2 \setminus\setminus A_2) \uplus C_2. \end{aligned}$$

Let us define the following sets:

$$\begin{aligned} E_1 &:= A_1 \setminus A_2 & E_2 &:= A_2 \setminus A_1 \\ E_3 &:= C_1 & E_4 &:= C_2 \\ E_5 &:= D_1 \setminus (E_2 \uplus E_3) & E_6 &:= D_2 \setminus (E_2 \uplus E_3) \\ &= D_3 \setminus (E_1 \uplus E_4) & &= D_4 \setminus (E_1 \uplus E_4) \end{aligned}$$

Notice that:

1. As $D_2$ and $D_4$ are valid sets, $A_1, A_2 \subseteq S_2$ and so $E_2 \subseteq S_2 \setminus\setminus A_1$. Also, as $D_2$ is a valid set, $S_2 \setminus\setminus A_1$ and $C_1$ are disjoint. These together imply that $E_2$ and $E_3$ are disjoint. Similarly

$E_1$ and $E_4$ are disjoint.

2. As $D_1$ and $D_3$ are valid sets, $A_1, A_2 \subseteq S_1$, so $E_2 \subseteq S_1 \setminus\setminus A_1$. Then, since $D_1 = (S_1 \setminus \setminus A_1) \uplus C_1$, $E_2 \uplus E_3 \subseteq D_1$. Similarly, $E_2 \uplus E_3 \subseteq D_2$, $E_1 \uplus E_4 \subseteq D_3$ and $E_1 \uplus E_4 \subseteq D_4$.

3. Since $D_1 = (S_1 \setminus\setminus A_1) \uplus C_1$, $D_1 \setminus C_1 = (S_1 \setminus\setminus A_1)$ and thus $(D_1 \setminus C_1) \setminus E_2 = (S_1 \setminus\setminus A_1) \setminus E_2$; i.e. $D_1 \setminus (E_2 \uplus E_3) = S_1 \setminus\setminus (A_1 \cup A_2)$ as $C_1 = E_3$ and $E_2 \cap E_3 = \emptyset$. Similarly, it can be shown that $D_3 \setminus (E_1 \uplus E_4) = S_1 \setminus\setminus (A_1 \cup A_2)$ and hence $D_1 \setminus (E_2 \uplus E_3) = D_3 \setminus (E_1 \uplus E_4)$. Similarly, $D_2 \setminus (E_2 \uplus E_3) = D_4 \setminus (E_1 \uplus E_4)$.

It is easy to see that $D_1, D_2, D_3$ and $D_4$ can be expressed as follows:

$$\begin{aligned} D_1 &= E_2 \uplus E_3 \uplus E_5, & D_2 &= E_2 \uplus E_3 \uplus E_6, \\ D_3 &= E_1 \uplus E_4 \uplus E_5, & D_4 &= E_1 \uplus E_4 \uplus E_6. \end{aligned} \tag{4.1}$$

Then, if $|A_1 \cap A_2| = v$, we have that,

$$\begin{aligned} |E_1| &= |E_2| = \ell - v, \\ |E_3| &= |E_4| = k, \\ |E_5| &= |E_6| = d_{\mathbf{x}} - (\ell - v + k). \end{aligned} \tag{4.2}$$

**Claim 4.2.** *Exactly one of the following holds for the sets $D_1, D_2, D_3$ and $D_4$:*

1. *All sets are distinct,*
2. *All four are the same i.e. $D_1 = D_2 = D_3 = D_4$,*
3. *$D_1 = D_2$, $D_3 = D_4$ and $D_1 \neq D_3$,*
4. *$D_1 = D_3$, $D_2 = D_4$ and $D_1 \neq D_2$.*

*Moreover, if $D_1, D_2$ and $D_3$ are distinct, then $\ell - v + k \leq r$ and $d_{\mathbf{x}} - (\ell - v + k) \leq r$.*

*Proof.* If all sets are distinct, then we are done. Otherwise we show that if $D_1$ is the same as at least one of $D_2, D_3$ and $D_4$, the proposition holds; the argument for other cases is similar.

If $D_1 = D_2$, then by equation (4.1), $E_5 = E_6$ and hence $D_3 = D_4$; thus either 2 or 3 holds.

If $D_1 = D_3$, then by equation (4.1), $E_2 \uplus E_3 = E_1 \uplus E_4$ and hence $D_2 = D_4$; thus either 2 or 4 holds.

If $D_1 = D_4$, then by equation (4.1), $E_6, E_1 \uplus E_4 \subseteq D_1$. Thus $D_2, D_3 \subseteq D_1$. However, as $|D_1| = |D_2| = |D_3|$, this is only possible if $D_1 = D_2 = D_3$ i.e. if 2 holds.

For the 'moreover' part of the proposition, notice that $|D_1 \cap D_2| \geq |E_2 \uplus E_3| = \ell - v + k$. So, if $\ell - v + k \geq r + 1$, then $D_1 = D_2$ because of the low intersection property of $NW_r$. Also, $|D_1 \cap D_3| \geq |E_5| = d_{\mathbf{x}} - (\ell - v + k)$. Hence, if $d_{\mathbf{x}} - (\ell - v + k) \geq r + 1$, then $D_1 = D_3$. □

The above claim gives a characterization of all the boxes

$$\mathbf{b} = \text{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

that can contribute to $\text{Tr}(B^2)$ and the following corollary follows.

**Corollary 4.1.** *For any four distinct sets $D_1, D_2, D_3, D_4 \in \binom{[m]}{d}$ define*

$$\begin{aligned}
\nu_0(D_1) &:= \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_1, D_1, D_1)\},\\
\nu_1(D_1, D_2) &:= \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_2, D_1, D_2)\},\\
\nu_2(D_1, D_2) &:= \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_1, D_2, D_2)\},\\
\nu_3(D_1, D_2, D_3, D_4) &:= \{box\ \mathbf{b} : \text{labels}(\mathbf{b}) = (D_1, D_2, D_3, D_4)\}.
\end{aligned}$$

*Also, define*

$$\begin{aligned}
T_0 &:= \sum_{D_1 \in \text{Supp}(NW_r)} |\nu_0(D_1)|,\\
T_1 &:= \sum_{D_1, D_2 \in \text{Supp}(NW_r)} |\nu_1(D_1, D_2)|,\\
T_2 &:= \sum_{D_1, D_2 \in \text{Supp}(NW_r)} |\nu_2(D_1, D_2)|,\\
T_3 &:= \sum_{D_1, D_2, D_3, D_4 \in \text{Supp}(NW_r)} |\nu_3(D_1, D_2, D_3, D_4)|.
\end{aligned} \tag{4.3}$$

*Then*

$$\text{Tr}(B^2) = T_0 + T_1 + T_2 + T_3.$$

We now upper bound $T_0, T_1, T_2$ and $T_3$.

### 4.2.3 Upper bound on $T_0$

First we observe that

**Observation 4.1.** *A $D_1 \in \binom{[m]}{d}$ can label at most one cell of a row $(A, C)$ of the matrix $N$.*

Hence for any box

$$\mathbf{b} = \text{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

that contributes to either $\nu_0(D_1)$ or $\nu_2(D_1, D_2)$, $S_1 = S_2$.

Now every box $\mathbf{b} \in \nu_0(D_1)$, $D_1 = D_3$, from equation (4.1), $E_1 \subseteq D_3 = D_1$. However, $E_1 \subseteq A_1$ and $A_1$ and $D_1$ are disjoint. Thus, $E_1 = A_1 \setminus A_2 = \emptyset$. Similarly, $E_2 = A_2 \setminus A_1 = \emptyset$. This and Equation (4.1) imply that $E_3 = E_4$ (i.e. $C_1 = C_2$). Hence we have the following claim,

**Claim 4.3.**

$$|\nu_0(D_1)| = \binom{m - d_{\mathbf{x}}}{\ell}\binom{d_{\mathbf{x}}}{k} \quad \textit{and} \quad T_0 = q^{r+1} \cdot \binom{m - d_{\mathbf{x}}}{\ell}\binom{d_{\mathbf{x}}}{k}.$$

*Proof.* Fix a $D_1 \in \binom{[m]}{d}$. As $A_1 = A_2$ and $C_1 = C_2$, the number of cells labelled by $D_1$ is $\binom{m-d_{\mathbf{x}}}{\ell} \cdot \binom{d_{\mathbf{x}}}{k}$ - the number of ways of picking $C_1$ times the number of ways of picking an $A_1$ that is disjoint from $C_1$. Summing over all $D_1 \in \text{Supp}(NW_r)$ yields the desired expression for $T_0$. □

### 4.2.4 Upper bound on $T_1$

In this section, we will need an estimate of the number of polynomials in $\mathbb{F}_q[z]$ of degree at most $r$ having exactly $w$ distinct roots in $[d_{\mathbf{x}}]$; so let us first estimate this number. We denote this number by $R(w, r)$. Since any polynomial $h(z) \in \mathbb{F}_q[z]$ of degree at most $r$ having $w$ roots in $[d_{\mathbf{x}}]$ is of the form

$$h(z) = (z - \alpha_1) \cdot (z - \alpha_2) \cdots (z - \alpha_w) \cdot \hat{h}(z)$$

where $\alpha_1, \alpha_2, ..., \alpha_w$ are in $[d_{\mathbf{x}}]$ and $\hat{h}(z) \in \mathbb{F}_q[z]$ is a polynomial of degree at most $r - w$, we have that

$$R(w, r) \leq q^{r-w+1} \cdot \binom{d_{\mathbf{x}}}{w} \leq q^{r+1} \cdot \left(\frac{d_{\mathbf{x}}}{q}\right)^w \cdot \frac{1}{w!}. \tag{4.4}$$

Let $D_1$ and $D_2$ be distinct sets in $\text{Supp}(NW_r)$. Consider the box

$$\mathbf{b} = \text{box}((A_1, C_1), (A_2, C_2), S_1, S_2)$$

in the set $\nu_1(D_1, D_2)$. Just as in the previous section, even here $D_1 = D_3$; thus $A_1 = A_2$ and $C_1 = C_2$. We then have the following claim,

**Claim 4.4.** *If $|D_1 \cap D_2| = w$, then*

$$|\nu_1(D_1, D_2)| = \binom{m - 2d_{\mathbf{x}} + w}{\ell}\binom{w}{k} \quad \text{and} \quad T_1 \leq d_{\mathbf{x}} \cdot \frac{q^{2(r+1)}}{d_{\mathbf{x}}^{\alpha k} \cdot k!} \cdot \binom{m - 2d_{\mathbf{x}} + k}{\ell}.$$

*Proof.* Let us fix $D_1$ and $D_2$ and then count the number of rows $(A, C)$ where in $D_1$ and $D_2$ both can occur as labels. $C \subseteq D_1 \cap D_2$ can be picked in $\binom{w}{k}$ ways. Then, since $A$ must be disjoint from $D_1 \cup D_2$ and $|D_1 \cup D_2| = 2d_{\mathbf{x}} - w$, $A$ can be picked in $\binom{m-2d_{\mathbf{x}}+w}{\ell}$ ways. Then, from equation (4.3), we get,

$$\begin{aligned}
T_1 &= \sum_{D_1 \in \text{Supp}(NW_r)} \sum_{k \leq w < d_{\mathbf{x}}} \sum_{\substack{D_2 \in \text{Supp}(NW_r), \\ |D_1 \cap D_2| = w}} \binom{m - 2d_{\mathbf{x}} + w}{\ell} \cdot \binom{w}{k} \\
&= \sum_{D_1 \in \text{Supp}(NW_r)} \sum_{k \leq w < d_{\mathbf{x}}} R_d(w, r) \cdot \binom{m - 2d_{\mathbf{x}} + w}{\ell} \cdot \binom{w}{k} \\
&\leq \sum_{D_1 \in \text{Supp}(NW_r)} \sum_{k \leq w < d_{\mathbf{x}}} q^{r+1} \cdot \left(\frac{d}{q}\right)^w \cdot \frac{1}{w!} \cdot \binom{m - 2d_{\mathbf{x}} + w}{\ell} \cdot \binom{w}{k} \\
&\leq q^{r+1} \sum_{D_1 \in \text{Supp}(NW_r)} \sum_{k \leq w < d_{\mathbf{x}}} \left(\frac{1}{d^\alpha}\right)^w \cdot \frac{1}{w!} \cdot \binom{m - 2d_{\mathbf{x}} + w}{\ell} \cdot \binom{w}{k}
\end{aligned}$$

The maxima of $\left(\frac{1}{d_{\mathbf{x}}^\alpha}\right)^w \frac{1}{w!}\binom{m-2d_{\mathbf{x}}+w}{\ell}\binom{w}{k}$ is attained at $w = k$. Hence,

$$T_1 \leq d_{\mathbf{x}} \cdot \frac{q^{2(r+1)}}{d_{\mathbf{x}}^{\alpha k} \cdot k!} \cdot \binom{m - 2d_{\mathbf{x}} + k}{\ell}.$$

□

### 4.2.5 Upper bound on $T_2$

Let $D_1$ and $D_2$ be distinct sets in $\mathrm{Supp}(NW_r)$. Then, from Observation 4.1, any box $\mathbf{b} \in \nu_2(D_1, D_2)$ is of the form

$$\mathbf{b} = \mathrm{box}((A_1, C_1), (A_2, C_2), S_1, S_1).$$

Let $|C_1 \cap C_2| = u$. Then, we get

**Claim 4.5.** *If $|D_1 \cap D_2| = w$, then*

$$|\nu_2(D_1, D_2)| = \sum_{u=0}^{k} \binom{m - 2d_{\mathbf{x}} + w}{\ell - d_{\mathbf{x}} + k + w - u} \binom{d_{\mathbf{x}} - w}{k - u}^2 \binom{w}{u}$$

$$\text{and } T_2 \leq d_{\mathbf{x}} \cdot k \cdot q^{2(r+1)} \cdot \binom{m - 2d_{\mathbf{x}}}{\ell - d_{\mathbf{x}} + k} \cdot \binom{d_{\mathbf{x}}}{k}^2.$$

*Proof.* The calculation for $T_2$ is similar to that for $T_1$ and hence is omitted. In this case, the maxima of the expression is attained at $w = u = 0$. □

### 4.2.6 Upper bound on $T_3$

**Claim 4.6.** *For our choice of the parameter $r$ and large enough $m$, $T_3 = 0$.*

*Proof.* From Claim 4.2, we have that if any box

$$\mathbf{b} \in \nu_3(D_1, D_2, D_3, D_4),$$

then $\ell - v + k \leq r$ and $d_{\mathbf{x}} - (\ell - v + k) \leq r$ and hence $d_{\mathbf{x}} \leq 2r$. However, recall that

$$r = \left\lfloor \frac{\alpha + \beta}{2(1 + \alpha)} d_{\mathbf{x}} \right\rfloor - 1,$$

where $\alpha = \Theta(\frac{\ln \ln m}{\ln m})$ and $\beta = \frac{1}{\ln m}$. Thus,

$$\begin{aligned} r &\leq \frac{\frac{c \cdot \ln \ln m}{\ln m} + \frac{1}{\ln m}}{2\left(1 + \frac{c \cdot \ln \ln m}{\ln m}\right)} d_{\mathbf{x}} && \text{(where } c \text{ is a constant)} \\ &= \frac{c \cdot \ln \ln m + 1}{2(\ln m + c \cdot \ln \ln m)} d_{\mathbf{x}} \\ &\leq \frac{c \cdot \ln \ln m}{\ln m} d_{\mathbf{x}} && \text{(for large enough } m) \end{aligned}$$

$$\leq \frac{1}{3} d_{\mathbf{x}} \qquad \text{(for large enough } m\text{)}$$

Thus we have $3r \leq d_{\mathbf{x}} \leq 2r$ which can only happen when $r \leq 0$. However, for large enough $m$, $r > 0$ and hence $d_{\mathbf{x}} \leq 2r$ is not possible. □

### 4.2.7 Lower bound on SurRank$(B)$

Comparing $\binom{m-2d_{\mathbf{x}}}{\ell-d_{\mathbf{x}}+k}$ and $\binom{m-d_{\mathbf{x}}}{\ell}$ we get

$$\binom{m-2d_{\mathbf{x}}}{\ell-d_{\mathbf{x}}+k} \geq \frac{1}{3^{d_{\mathbf{x}}}} \cdot \binom{m-d_{\mathbf{x}}}{\ell}.$$

Hence, from Claims 4.3 and 4.5, the upper bound on $T_2$ dominates the upper bound on $T_0$. This in conjunction with Corollary 4.1 and Claims 4.6, 4.4 and 4.5 yields,

$$\begin{aligned}\mathrm{Tr}(B^2) \leq\ & d_{\mathbf{x}} \cdot \frac{q^{2(r+1)}}{d_{\mathbf{x}}^{\alpha k} \cdot k!} \cdot \binom{m-2d_{\mathbf{x}}+k}{\ell} \\ & + 2d_{\mathbf{x}} \cdot k \cdot q^{2(r+1)} \cdot \binom{m-2d_{\mathbf{x}}}{\ell-d_{\mathbf{x}}+k} \cdot \binom{d_{\mathbf{x}}}{k}^2.\end{aligned}$$

This along with Claim 4.1 implies that SurRank$(B) \geq \min(R_1, R_2)$, where

$$R_1 = \frac{q^{2(r+1)} \cdot \binom{d_{\mathbf{x}}}{k}^2 \cdot \binom{m-d_{\mathbf{x}}}{\ell}^2}{2d_{\mathbf{x}} \cdot \frac{q^{2(r+1)}}{d^{\alpha k} \cdot k!} \cdot \binom{m-2d_{\mathbf{x}}+k}{\ell}}$$

and

$$R_2 = \frac{q^{2(r+1)} \cdot \binom{d_{\mathbf{x}}}{k}^2 \cdot \binom{m-d_{\mathbf{x}}}{\ell}^2}{4d_{\mathbf{x}} \cdot k \cdot q^{2(r+1)} \cdot \binom{m-2d_{\mathbf{x}}}{\ell-d_{\mathbf{x}}+k} \cdot \binom{d_{\mathbf{x}}}{k}^2}.$$

Since

$$\frac{\binom{m-d_{\mathbf{x}}}{\ell}^2}{\binom{m-2d_{\mathbf{x}}+k}{\ell}} \geq \frac{1}{2^k d_{\mathbf{x}}^{O(1)}} \cdot \binom{m}{\ell} \text{ and}$$

$$d_{\mathbf{x}}^{\alpha k} \cdot k! \cdot \binom{d_{\mathbf{x}}}{k}^2 \geq \frac{1}{2^k d_{\mathbf{x}}^{O(1)}} \cdot \binom{m}{k},$$

$R_1 \geq \frac{1}{d^{O(1)}} \cdot \frac{1}{4^k} \cdot \binom{m}{k} \cdot \binom{m}{\ell}$. Similarly, as

$$\frac{\binom{m-d_{\mathbf{x}}}{\ell}^2}{\binom{m-2d_{\mathbf{x}}}{\ell-d_{\mathbf{x}}+k}} \geq \frac{1}{d_{\mathbf{x}}^{O(1)}}\binom{m}{\ell+d_{\mathbf{x}}-k},$$

$R_2 \geq \frac{1}{d^{O(1)}}\binom{m}{\ell+d_{\mathbf{x}}-k}$. Hence,

$$\begin{aligned}\mathrm{SurRank}(B) &\geq \frac{1}{d_{\mathbf{x}}^{O(1)}} \min\left(\frac{1}{4^k}\cdot\binom{m}{k}\binom{m}{\ell}, \binom{m}{\ell+d_{\mathbf{x}}-k}\right)\\ &= \frac{1}{m^{O(1)}} \min\left(\frac{1}{4^k}\cdot\binom{m}{k}\binom{m}{\ell}, \binom{m}{\ell+d_{\mathbf{x}}-k}\right).\end{aligned}$$

# Chapter 5

# Conclusion and Future Work

In this thesis, we proved, what is to the best of our knowledge, the first super-quadratic lower bound for depth four arithmetic circuits. Some interesting avenues for future work are as follows:

1. Prove a super quadratic lower bound on the number of *gates* of a depth four circuit. The almost cubic lower The almost cubic lower bound for depth three circuits in [KST16] is on the number of gates.

2. Improve the lower bound to an almost cubic lower bound.

3. Prove a super-quadratic lower bound for a polynomial in VP. An almost cubic lower bound for depth three circuits is known for polynomials computed by polynomial size depth five circuits [BLS16, Yau16].

4. Prove lower bounds for the $IMM_{2,n}$ "polynomial". It follows from ideas in [CT19] that if $IMM_{2,n}$ - the $2 \times 2$ matrix of polynomials obtained by multiplying $n$ many $2 \times 2$ symbolic matrices whose entries ae distinct variables - can be computed by a depth $\Delta$ circuit of size $n^k$, then it can also be computed by a depth $\Delta_0$ circuit of size $O(\frac{\Delta}{\Delta_0} \cdot n^{1+\exp(-\frac{\Delta}{\Delta_0 k})})$.[1] So, proving a lower bound of $\Omega(\Delta \cdot n^{1+\frac{1}{\Delta}})$ for a depth circuit computing $IMM_{2,n}$ would give a super-polynomial lower bound for constant depth circuits! In fact, even for depth five circuits, we have the following: a quadratic lower bound on the number of gates of a depth five circuit computing $IMM_{2,n}$ would yield a super-cubic lower bound on the size of depth three circuits.

[1] We thank Ankit Garg for pointing this out to us.

# Bibliography


[Alo09] Noga Alon. Perturbed Identity Matrices Have High Rank: Proof and Applications. *Combinatorics, Probability & Computing*, 18(1-2):3–15, 2009. 27

[AV08] Manindra Agrawal and V. Vinay. Arithmetic Circuits: A Chasm at Depth Four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 67–75. IEEE Computer Society, 2008. iii, 2

[BLS16] Nikhil Balaji, Nutan Limaye, and Srikanth Srinivasan. An almost cubic lower bound for ΣΠΣ circuits computing a polynomial in VP. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:143, 2016. iii, 2, 37

[BS83] Walter Baur and Volker Strassen. The Complexity of Partial Derivatives. *Theor. Comput. Sci.*, 22:317–330, 1983. 1

[Bür00] Peter Bürgisser. Cook's versus valiant's hypothesis. *Theor. Comput. Sci.*, 235(1):71–88, 2000. 1

[CKSV19] Prerona Chatterjee, Mrinal Kumar, Adrian She, and Ben Lee Volk. A Quadratic Lower Bound for Algebraic Branching Programs. *Electronic Colloquium on Computational Complexity (ECCC)*, page 170, 2019. 2

[CT19] Lijie Chen and Roei Tell. Bootstrapping results for threshold circuits "just beyond" known lower bounds. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 34–41. ACM, 2019. 37

[DDPW83] Danny Dolev, Cynthia Dwork, Nicholas Pippenger, and Avi Wigderson. Superconcentrators, Generalizers and Generalized Connectors with Limited Depth

(Preliminary Version). In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing, 25-27 April, 1983, Boston, Massachusetts, USA*, pages 42–51. ACM, 1983. 4

[Erd32] Paul Erdős. Beweis eines Satzes von Tschebyschef. *Acta Litt. Sci. Szeged*, 5:194–198, 01 1932. 22

[FLMS15] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower Bounds for Depth-4 Formulas Computing Iterated Matrix Multiplication. *SIAM J. Comput.*, 44(5):1173–1201, 2015. Conference version appeared in the proceedings of STOC 2014. iii, 2

[GKKS14] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the Chasm at Depth Four. *J. ACM*, 61(6):33:1–33:16, 2014. Conference version appeared in the proceedings of CCC 2013. iii, 2, 9

[GKKS16] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic Circuits: A Chasm at Depth 3. *SIAM J. Comput.*, 45(3):1064–1079, 2016. Conference version appeared in the proceedings of FOCS 2013. iii, 2

[GST20] Nikhil Gupta, Chandan Saha, and Bhargav Thankey. A super-quadratic lower bound for depth four arithmetic circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 27:28, 2020. 10, 21

[Kal85] K. Kalorkoti. A Lower Bound for the Formula Size of Rational Functions. *SIAM J. Comput.*, 14(3):678–687, 1985. 2

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012. iii, 2

[KLSS17] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. *SIAM J. Comput.*, 46(1):307–335, 2017. Conference version appeared in the proceedings of FOCS 2014. iii, 2, 3, 8, 9, 21

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012. iii, 2

[KS14] Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 136–145, 2014. 3

[KS16a] Neeraj Kayal and Chandan Saha. Lower Bounds for Depth-Three Arithmetic Circuits with small bottom fanin. *Computational Complexity*, 25(2):419–454, 2016. Conference version appeared in the proceedings of CCC 2015. iii, 2, 3, 21, 22, 23, 26

[KS16b] Mrinal Kumar and Shubhangi Saraf. Sums of Products of Polynomials in Few Variables: Lower Bounds and Polynomial Identity Testing. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 35:1–35:29, 2016. 3

[KS17] Mrinal Kumar and Shubhangi Saraf. On the Power of Homogeneous Depth 4 Arithmetic Circuits. *SIAM J. Comput.*, 46(1):336–387, 2017. Conference version appeared in the proceedings of FOCS 2014. iii, 2, 3

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 146–153, 2014. iii, 2, 3, 21

[KST16] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An Almost Cubic Lower Bound for Depth Three Arithmetic Circuits. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 33:1–33:15, 2016. iii, 2, 3, 11, 37

[Pud94] Pavel Pudlák. Communication in Bounded Depth Circuits. *Combinatorica*, 14(2):203–216, 1994. 4

[Raz10] Ran Raz. Elusive Functions and Lower Bounds for Arithmetic Circuits. *Theory of Computing*, 6(1):135–177, 2010. Conference version appeared in the proceedings of STOC 2008. iii, 3, 4

[RS03] Ran Raz and Amir Shpilka. Lower Bounds for Matrix Product in Bounded Depth Circuits with Arbitrary Gates. *SIAM J. Comput.*, 32(2):488–513, 2003. Conference version appeared in the proceedings of STOC 2001. 4

[Sha17] Abhijat Sharma. An Improved Lower Bound for Depth Four Arithmetic Circuits. Master's thesis, Indian Institute of Science, Bangalore, India, 2017. https://www.csa.iisc.ac.in/~chandan/thesis_reports/AbhijatSharma_MScThesis.pdf. iii, 3, 4, 21, 22, 23, 26

[SS97] Victor Shoup and Roman Smolensky. Lower Bounds for Polynomial Evaluation and Interpolation Problems. *Computational Complexity*, 6(4):301–311, 1997. Conference version appeared in the proceedings of FOCS 1991. iii, 4

[Str73] Volker Strassen. Vermeidung von divisionen. *The Journal für die Reine und Angewandte Mathematik*, 264:182–202, 1973. 1

[SW01] Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10(1):1–27, 2001. Conference version appeared in the proceedings of CCC 1999. iii, 2, 11

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic Circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010. 3

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Inf. Comput.*, 240:2–11, 2015. Conference version appeared in the proceedings of MFCS 2013. iii, 2

[Val75] Leslie G. Valiant. On Non-linear Lower Bounds in Computational Complexity. In William C. Rounds, Nancy Martin, Jack W. Carlyle, and Michael A. Harrison, editors, *Proceedings of the 7th Annual ACM Symposium on Theory of Computing, May 5-7, 1975, Albuquerque, New Mexico, USA*, pages 45–53. ACM, 1975. 4

[Val79] Leslie G. Valiant. Completeness Classes in Algebra. In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 249–261, 1979. 1, 6

[Vaz01] Vijay V. Vazirani. *Approximation algorithms*. Springer, 2001. 11, 16

[VSBR83] Leslie G. Valiant, Sven Skyum, S. Berkowitz, and Charles Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM J. Comput.*, 12(4):641–644, 1983. iii, 2

[Yau16] Morris Yau. Almost cubic bound for depth three circuits in VP. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:187, 2016. iii, 2, 37