Algebraic geometric methods for algorithms in satisfiability, irreducibility of varieties, and identity testing

by

Abhibhav Garg

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2025

© Abhibhav Garg 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Shubhangi Saraf, Professor Department Computer Science University of Toronto
Supervisor:	Rafael Oliveira, Assistant Professor Cheriton School of Computer Science University of Waterloo
Internal Member:	Eric Blais, Associate Professor Cheriton School of Computer Science University of Waterloo
Internal Member:	Éric Schost, Professor Cheriton School of Computer Science University of Waterloo
Internal-External Member:	Jonathan Leake, Assistant Professor Department of Combinatorics and Optimization University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

All the results of this thesis are based on papers that I have coauthored, as well as works in preparation. The details are as follows:

- The content of Sections 1 to 9, Chapter 3 is based on Garg, Oliveira, and Sengupta [GOS22], Garg, Oliveira, Peleg, and Sengupta [GOPS23], Garg, Oliveira, and Sengupta [GOS25c], and Garg, Oliveira, and Sengupta [GOS25b].
- The content of Section 10, Chapter 3 is based on joint work with Rafael Oliveira, Akash Kumar Sengupta, Nir Shalmon, and Amir Shpilka. This work has been submitted for peer review.
- The content of Section 4.1 is based on Garg, Oliveira, and Saxena [GOS25a].
- The content of Section 4.2 is based on joint work with Robert Andrews and Éric Schost. This work is in preparation to be submitted for peer review.

Abstract

In this thesis we study three problems that lie in the intersection of abstract algebra and theoretical computer science. The first of these is the polynomial identity testing problem, which is the task of determining if an algebraic circuit computes the identically zero polynomial. We give the first polynomial time deterministic algorithm for the special case of depth four algebraic circuits, with top fan-in three, and constant bottom fan-in. We also give the first such algorithm for circuits with bottom fan-in two, and constant top fan-in. Our methods involve studying higher degree generalisations of classical incidence configurations known as Sylvester–Gallai configurations. The second of these is the problem of checking if a system of equations is satisfiable. In the regime when the number of variables in the system is a constant, we show that satisfiability can be checked in constant depth by algebraic circuits. In particular, we show that the multivariate resultant has a constant depth circuit in this regime, independent of the degrees. The previous best known constructions of the resultant required depth that was logarithmic in the degrees. The final problem we consider is the problem of deciding if an ideal theoretically defined variety is irreducible in characteristic 0. We show that this task can be solved in the polynomial hierarchy assuming the generalized Riemann hypothesis. This improves the previous best known bound of PSPACE.

Acknowledgements

I start by thanking my advisor, Rafael Oliveira, for all his support. He has instilled in me a taste for good problems, and has taught me how to solve them effectively. He was exceedingly nurturing at the start of this journey, and has constantly pushed me to improve as a researcher.

I would like to thank Eric Blais, Éric Schost, Jonathan Leake, and Shubhangi Saraf for taking the time to serve on my thesis committee. Their feedback has been very valuable, and has greatly improved this thesis.

I have been fortunate to have many wonderful collaborators over the last few years. I am grateful to Akash Kumar Sengupta, Shir Peleg, Amir Shpilka, Nitin Saxena, Éric Schost, and Robert Andrews for many insightful discussions.

I would like to thank Mrinal Kumar, Ramprasad Saptharishi, and Prahladh Harsha for hosting me at TIFR during the Monsoon 2024 term. Their love for the game is infectious. The four months I spent in Mumbai were a turning point in my academic life.

I would like to thank all the people in the algorithms and complexity group that have made it such a wonderful place. I am especially grateful to Alex Tung, Robert Wang, Yueheng Zhang, Abhiroop Sanyal, Vahid Asadi, Dante Tjowasi, Vishwas Bhargava, Argyris Mouzakis, Janani Sunderesan, Parth Mittal, Sriram Gopalakrishnan, Mohammad Ebtehaj, Vihan Shah, Remy El Sebah, Matt Regehr, Cameron Seth, Dinis Vitorino, Raymond Liu, Lap Chi Lau, and Sepehr Assadi.

Finally, I would like to thank my partner Qinglan Cao, my parents, and Yash Singh. They have stood by me through thick and thin, this thesis would have been impossible without their unconditional love and support.

Table of Contents

Examining Committee ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgements vi

1 Introduction 1
- 1.1 Algebraic complexity and identity testing . . . 1
- 1.2 Properties of algebraic varieties . . . 4
- 1.3 Organisation of this thesis . . . 6

2 Preliminaries 7
- 2.1 Notation . . . 7
- 2.2 Algebraic Circuits . . . 7
  - 2.2.1 Piecewise circuits . . . 9
- 2.3 Algebraic Geometry . . . 10
- 2.4 Complexity Theory . . . 12

3 Identity testing and Sylvester–Gallai Configurations 14
- 3.1 Linear Sylvester–Gallai configurations . . . 15
- 3.2 Polynomial identity testing and Sylvester–Gallai configurations . . . 18
- 3.3 High degree Sylvester–Gallai configurations . . . 20
- 3.4 Some high level ideas . . . 22
  - 3.4.1 Dimension bounds for quadratic radical Sylvester–Gallai configurations . . . 23
  - 3.4.2 Ideals generated by a few low degree forms . . . 24
  - 3.4.3 Strong algebras of quadratics and a structure lemma for radicals . . . 25
- 3.5 High dimension Quadratic Sylvester–Gallai configurations . . . 27
  - 3.5.1 Preliminaries . . . 27
  - 3.5.2 Integral Sequences . . . 32
  - 3.5.3 Strong Sequences . . . 34
  - 3.5.4 Proof of the Sylvester–Gallai theorem . . . 37
- 3.6 Going beyond quadratics: Strong algebras and graded quotients . . . 41
  - 3.6.1 Strong and lifted strong algebras . . . 41
  - 3.6.2 Graded Quotients . . . 46

3.7 Prime bounds . . . 49
3.7.1 Irreducibility of specialisations . . . 50
3.7.2 Bounds on forms generating non prime ideals via a Bertini theorem . . . 51
3.7.3 Absolute irreducibility with respect to strong algebras . . . 56
3.8 Dimension bounds for Product Sylvester–Gallai configurations . . . 59
3.8.1 Controlling factor sets . . . 62
3.8.2 Proof of main theorem . . . 64
3.9 Dimension bounds for high degree EK configurations . . . 68
3.9.1 Compatible Quotients . . . 71
3.9.2 A potential function . . . 74
3.9.3 Controlling highest degree forms . . . 76
3.9.4 Putting it all together . . . 82
3.10 Rank bounds for constant top fan-in depth four circuits with bottom fan-in 2 . . . 84
3.10.1 Essential variables in quotient rings . . . 85
3.10.2 Reduction to the low strength case . . . 92
3.10.3 The low strength case . . . 98
3.10.4 Rank bounds for identities . . . 108

4 Algorithms for properties of varieties **111**
4.1 Deciding irreducibility of varieties . . . 111
4.1.1 An interactive proof for system satisfiability . . . 113
4.1.2 High level ideas . . . 116
4.1.3 Heights of algebraic numbers and varieties . . . 118
4.1.4 Properties preserved by base change . . . 126
4.1.5 An interactive proof for reducibility of ideal theoretically defined varieties . . . 130
4.2 Low depth circuits for multivariate resultants . . . 131
4.2.1 Computational complexity of determinants and resultants . . . 134
4.2.2 Low depth circuits for multivariate resultants . . . 136
4.2.3 Resultant systems . . . 141

Bibliography **146**

Chapter 1

Introduction

Algebraic methods are ubiquitous in theoretical computer science. They are used in designing algorithms, analysing algorithms, proving lower bounds, proving barriers towards lower bounds, and constructing cryptographic protocols, just to name a few applications [GG13], [AB09, Chapters 16-18]. In this thesis, we use algebraic methods (in particular, methods from algebraic geometry and commutative algebra) to design algorithms for important special cases of some central problems in theoretical computer science.

The first of these problems is *polynomial identity testing*. Suppose we are given a short computation for a polynomial $P \in \mathbb{Z}[x_1, \ldots, x_n]$. This might mean we are given a sequence of additions and multiplications that start with variables and constants, and results in P. It might mean that P is described in terms of some other basic polynomials, for example P is the determinant of a matrix with entries that are easy to describe. The goal is simple: we want to test if $P \equiv 0$, that is, if P is the identically zero polynomial. This is known as the polynomial identity testing problem, often abbreviated to PIT. There are efficient randomised algorithms for this task, and we are interested in solving this task deterministically and efficiently.

The next two problems we consider are problems in computational algebraic geometry. Suppose we are given polynomials $P_1, \ldots, P_m \in \mathbb{Z}[x_1, \ldots, x_n]$. Let $V \subset \mathbb{C}^n$ be the set of common zeroes of $P_1, \ldots, P_m$, the set V is what is known as an *algebraic variety*. We consider the two most basic questions one can ask about V. The first is whether V is empty or nonempty. This is also known as the *satisfiability problem*. The second is whether V is irreducible or reducible.

In this thesis, we design algorithms for each of the three problems listed above in certain ranges of parameters and in certain models of computation. In the next few sections, we will elaborate on these problems, discuss them in context, and state the contributions of this thesis. Given that the techniques used are fairly different, we will defer discussion of proof techniques to later chapters.

1.1 Algebraic complexity and identity testing

A polynomial P of degree d in n variables has $\binom{n+d}{n}$ coefficients. When manipulating such polynomials, representing and storing them as a list of coefficients is therefore intractable, unless either the degree or the number of variables is a constant. Some saving can be made if we only list the nonzero coefficients of P. Indeed if P is a *sparse* polynomial, that is, if P only has a few nonzero coefficients, then we can represent P by a list of pairs of nonzero coefficients and monomials. This is called the *sparse representation of* P. There are two downsides to the sparse representation. The first is that many interesting polynomials are not sparse. For example, the determinant of an $n \times n$ symbolic matrix, which is arguably the single most important polynomial in linear algebra and in computer

science, is a polynomial of degree n in n^2 variables, and it has $n!$ nonzero coefficients. The second downside is that the notion of sparsity is slightly brittle. For example, if we start with a sparse polynomial P and consider a translate $P_a(x_1, \ldots, x_n) := P(x_1 + a_1, \ldots, x_n + a_n)$ then usually P_a will no longer be sparse.

An alternate data structure for representing a multivariate polynomial is an *algebraic circuit*. Informally, an algebraic circuit is a directed acyclic graph, whose leaves are labelled with variables or constants, and whose internal nodes are labelled with $+$ and $\times$. The computation at each internal node is the natural one: nodes marked $+$ compute sums of their inputs while nodes marked $\times$ compute products of their inputs. Algebraic circuits can be viewed both as data structures that succinctly represent polynomials, and also as programs in a non-uniform algebraic model of computation for polynomials. Historically, these objects were defined with slightly different syntax and were called *straight line programs* or *slps*. The notions of algebraic circuits and slps are computationally equivalent, in this thesis we will use algebraic circuit.

The usefulness of circuits as data structures was demonstrated many times over, by results that showed that natural operations on polynomials with small circuits result in polynomials with small circuits. For example, in a long line of work, Kaltofen [Kal85; Kal87; Kal89] and Kaltofen and Trager [KT90] showed that factors of polynomials of small degree with small algebraic circuits have small algebraic circuits. As another example, given a system of equations computed by small circuits, solutions to the system can be described using polynomials that have small circuits. The line of work that led into this is too long for us to review here, we refer the interested reader to Giusti, Heintz, and Sabia [GHS93] and Giusti, Heintz, Hägele, Morais, Pardo, and Montana [GHH+97] and the many references within for a historical account.

The importance of algebraic circuits as programs in a non-uniform model of computation was first observed by Valiant [Val79]. In computer science, we often identify a set of problems as being computationally equivalent (that is, these problems reduce to each other) and as being the most difficult problems in a given class of problems (that is, every problem in the class reduces to a problem in our set). Valiant observed that for collections of *algebraic problems*, we can often define the notions of reductions and completeness in a purely algebraic manner. This led him to define a number of algebraic complexity classes. For example, he defined the class VP (which stands for Valiant's P) as the class of polynomials of polynomial degree, computable by polynomial sized algebraic circuits. Similarly, he defined an algebraic version of NP, denoted VNP, and asked if $\mathsf{VP} = \mathsf{VNP}$. Given that the definitions of the model of computation and the notion of reductions are algebraic in nature, Valiant observed that the VP versus VNP problem can be phrased purely as a quantitative question about certain complete polynomials, namely the permanent and the determinant. One hopes that the VP versus VNP problem and other related problems regarding algebraic complexity class are easier than their boolean counterpart, since the tools of algebraic geometry can be deployed more easily against them. In fact it is known (see [Bür00a; Bür00b]) that a resolution to the P versus NP problem implies a resolution to the VP versus VNP problem, which shows that the latter is easier than the former.

These two avatars of algebraic circuits, as data structures and as programs, are of course very interconnected. For a very detailed exposition of all things related to algebraic circuits and algebraic complexity theory, we refer the reader to the book by Bürgisser, Clausen, and Shokrollahi [BCS97]. A more introductory survey that focuses on algebraic complexity is [SY10].

When studying algebraic circuits in either of their two roles, the following natural question seems to be of fundamental importance: given an algebraic circuit C, can we efficiently decide if C computes the zero polynomial. If we can decide whether C computes the zero polynomial, we can decide if two circuits compute the same polynomial, by performing the zero test on the difference. There is a simple randomised algorithm that solves this decision problem, based on the following

lemma (which we will call the *degree mantra*).

Lemma 1.1.1 ([Sch80; Zip79; DL78; Ore21])**.** *Let k be a field. If $P \in k[x_1, \ldots, x_n]$ is a polynomial of total degree d, and if $S \subset k$ is any subset of k, then P has at most $d|S|^{n-1}$ roots in S^n.*

Suppose C is known to compute a polynomial of degree d. We can pick S to be any subset of the field of size say $2d$, and evaluate C at a random point in S^n. If C were nonzero, then this evaluation would be nonzero with constant probability. Ignoring bit complexity issues (which can be dealt with using modular arithmetic), this is an efficient randomised algorithm for identity testing. In fact this algorithm does not even care that the polynomial has a small circuit, all it requires is the ability to evaluate the circuit at a point, and a bound on the degree of the polynomial the circuit computes.

A number of algorithms that use algebraic circuits as data structures perform identity testing as a subroutine. In fact there are also a number of algorithms that seemingly have nothing to do with manipulating polynomials that also use identity testing as a subroutine. Examples include fast parallel algorithms for graph matching problems [Lov79; KUW86; MVV87], and algorithms for primality testing of integers [AKS04]. Often identity testing is the only source of randomness, and therefore deterministic algorithms for identity testing would imply deterministic algorithms for a large number of problems (primality of integers is a case where the specific instances of identity testing that show up admit a deterministic algorithm).

Connections between algebraic complexity and identity testing were established by Heintz and Schnorr [HS80] and Agrawal [Agr05], who showed that deterministic polynomial time *blackbox* identity testing algorithms would imply the existence of explicit polynomials that do not lie in the complexity class VP. A blackbox identity testing algorithm is one where the algorithm is not allowed to look at the circuit, all it can do is evaluate the circuit at a set of points, and it has to decide if the circuit computes zero based on these evaluations. It is easy to show that for an algorithm to be able to do this correctly, it must essentially produce a small list of points such that if the circuit computes a nonzero polynomial, it must evaluate to something nonzero at at least one point. Such a set of points is called a hitting set, and the existence of polynomial sized hitting sets for families in VP was also shown by Heintz and Schnorr [HS80]. Subsequently, Kabanets and Impagliazzo [KI04] showed that the existence of any deterministic identity testing algorithm for circuits would imply either $\mathsf{NEXP} \not\subset \mathsf{P/poly}$ or that the permanent does not have small circuits. Both of these are far stronger lower bounds than anything we know today. Kabanets and Impagliazzo [KI04] also established a connection in the opposite direction: non trivial algebraic circuit lower bounds can be used to give deterministic algorithms for identity testing, using algebraic analogues of Nisan–Wigderson generators [NW94]. The set of results that relate lower bounds for algebraic circuits with deterministic algorithms for identity testing are together referred to as *hardness randomness tradeoffs*. The excellent survey by Kumar and Saptharishi [KS19] discusses recent interesting work in this direction.

The above discussion should have convinced the reader of two things. The first is that the polynomial identity testing problem is one of fundamental importance. The second is that it is probably extremely difficult to solve it in full generality. Given this state of affairs, researchers have focused on studying identity testing for syntactically restricted circuit classes. The type of restriction that we focus on in this thesis are restrictions on the *depth* of the circuits computing the identities we want to test.

The depth of a circuit C is the length of the longest path from any output node to any input node. General algebraic circuits admit efficient depth reduction: Valiant, Skyum, Berkowitz, and Rackoff [VSBR83] showed that given any circuit C computing a polynomial P of degree d, we can construct a circuit C' that computes P with depth $O(\log^2 d)$ and size polynomial in the number of variables, the degree d, and the size of C. More recently, Agrawal and Vinay [AV08], Koiran

[Koi12], Gupta, Kamath, Kayal, and Saptharishi [GKKS13], and Tavenas [Tav15] showed that any polynomial that can be computed by a circuit of size s can be computed by depth four and depth three circuits of size $s^{O(\sqrt{d})}$. These results have lead to a concentrated research effort on the identity testing problem for depth three and four circuits.

These results also show that identity testing for depth three and four circuits is almost as hard as general identity testing. In particular, polynomial time deterministic identity testing algorithms for depth four and depth three circuits imply quasipolynomial time deterministic identity testing algorithms for general circuits. Therefore the focus has been on circuits with additional restrictions on the fan-in of certain gates. One such class is the class of depth three circuits, with constant top fan-in, we denote this class by $\Sigma^k\Pi\Sigma$. Even for this very restricted class, deterministic polynomial time identity testing was the culmination of a long list of results ([DS07; KS11; SS11; KS09; SS13; SS12]). We will discuss these results in detail in Section 3.2. Towards designing algorithms, these results established a connection between polynomial identity testing of $\Sigma^k\Pi\Sigma$ and a family of results in incidence geometry known as *Sylvester–Gallai type theorems*.

A next natural class of circuits is depth four circuits with constant top and bottom fan-in. We denote such circuits by $\Sigma^k\Pi\Sigma\Pi^d$, they compute sums of products of low degree polynomials. These are perhaps the simplest families of circuits for which we do not have polynomial time identity testing algorithms. In an attempt to extend the above methods to depth four circuits, Gupta [Gup14] studied non linear generalisations of Sylvester–Gallai theorems. He defined a number of different high degree analogues of classical Sylvester–Gallai configurations, and gave a polynomial time identity testing algorithm, under the assumption that the configurations he defined all have bounded dimension. One of the main contributions of the results presented in this thesis is the resolution of some of the conjectures of Gupta [Gup14], and therefore an algorithm for identity testing of $\Sigma^3\Pi\Sigma\Pi^d$ circuits and $\Sigma^k\Pi\Sigma\Pi^2$ circuits. While this is a small step in the grand scheme of designing algorithms for identity testing, we hope that the methods we develop can be extended and used for identity testing for larger classes of circuits. We defer a detailed discussion of our results, implications for deterministic identity testing, and related results on dimension bounds for higher degree Sylvester–Gallai configurations to Chapter 3.

1.2 Properties of algebraic varieties

Most problems in computational algebraic geometry and computational commutative algebra tend to be very hard. For example, ideal membership, the most basic problem in computational algebra, is complete for EXPSPACE [MM82]. Computer scientists and mathematicians have spent many years trying to understand exactly what makes these problems difficult (see for example the survey by Bayer and Mumford [BM93]). This question is what serves as the motivation for the two problems we discuss in this section. We establish some notation to make it easy to define the problems. Let $P_1, \ldots, P_m \in \mathbb{Z}[x_1, \ldots, x_n]$ be a set of polynomials of degree at most d, and let $V \subset \mathbb{C}^n$ be the set of common zeroes of $P_1, \ldots, P_m$.

Consider the problem of satisfiability, that is, the problem of deciding if $P_1, \ldots, P_m$ have a common solution in $\mathbb{C}^n$, or equivalently the problem of deciding if V is nonempty. It might not even be clear that this problem is decidable, recall that the related task of deciding if there exist solutions in $\mathbb{Z}^n$ is undecidable. A famous result of Hilbert, now known as Hilbert's Nullstellensatz, states that V is empty if and only if there exist polynomials $Q_1, \ldots, Q_m$ such that $1 = P_1Q_1 + \cdots + P_mQ_m$. This reduces the satisfiability problem to an instance of ideal membership. The classical result of [Her26] shows that if V is empty, then we can find such $Q_1, \ldots, Q_m$ of degrees at most d^{2^n}. This turns the above problem into a problem in linear algebra, although one involving matrices of double

exponential size. Using the fast parallel algorithms for linear algebra by [Csa76] satisfiability can be decided in EXPSPACE.

The above type of *double exponential bounds* are generally unavoidable in commutative algebra. However for the special instances of ideal membership that show up when testing satisfiability (where we only want to check if 1 is in an ideal), *single exponential bounds* hold [Kol88; Bro87; Jel05]. Concretely, in the setting of the previous paragraph, if V is empty, then we can find $Q_1, \ldots, Q_m$ of degrees at most d^n, instead of d^{2^n}. In particular, if we reduce the problem to linear algebra and use the methods of Csanky [Csa76], then we can decide satisfiability in PSPACE instead of EXPSPACE. Currently, this is the best known unconditional bound on the complexity of satisfiability.

The only known hardness result for the problem of satisfiability is that it is NP-hard. This is easy to show, for example by taking instances of SAT and arithmetising them. This leaves a large gap between known upper and lower bounds for the complexity of this problem. Conditionally, this state of affairs was improved by Koiran [Koi96], who showed that assuming a version of the Riemann hypothesis, satisfiability of systems in characteristic zero lies in AM, which one can think of as NP with a verifier that can use randomness. While conditional, this almost fully characterises the complexity of satisfiability in the characteristic zero case. In fact, under reasonable hardness assumptions, AM = NP holds [MV05], therefore conditionally we have a exact characterisation.

In our work, we try to get a similar characterisation of the complexity of satisfiability, albeit at a different scale. We focus on the case when n, the number of variables, is a constant. We will also work in the algebraic model discussed in the previous section, rather than Turing model, although our results can be transferred to the Turing model easily. Recall that all the bounds we have encountered, even the double exponential ones, are still only polynomial in the degree d when n is fixed. In this setting therefore it is easy to show that satisfiability can be solved in polynomial time. Our goal however is to further situate the complexity within polynomial time. The way we stratify polynomial time is using the notion of parallel complexity, or equivalently the *depth* of the algebraic circuits that decide our problem. Our question is therefore the following: how parallelisable is the satisfiability problem, when n is a constant?

The above reductions of satisfiability to linear algebra sets a baseline. Linear algebraic tasks for systems of size $N \times N$ can be solved by circuits of depth $O(\log^2 N)$. There are good reasons to believe that this cannot be improved: circuits of depth $O(\log N)$ for the determinant would imply NL = L. This shows satisfiability can be solved in by circuits of depth $O(n^2 \log^2 d)$. In a recent breakthrough, Andrews and Wigderson [AW24] showed that for the special case of $n = 1$, satisfiability can be decided by constant depth circuits, a significant improvement over what was known before.[1] In our work, we will extend this result to all constant n.

More formally we show that the *multivariate resultant* of n homogeneous forms of degrees at most d can be computed by algebraic circuits of depth $O(n)$ and size $d^{O(n^2)}$. We also show the existence of *resultant systems* consisting of polynomials computable in the same depth. Our methods involve using the Poisson formula for the multivariate resultant, and using hidden resultants and a rational shape lemma to compute the product of evaluations of a polynomial on a zero dimension complete intersection.

The second problem we consider is that of the irreducibility of V. Any algebraic variety V is a finite union of irreducible varieties. The exact definition of irreducible varieties is slightly technical, and will be given in the preliminaries. Essentially an irreducible variety is one such that if a product of functions fg is zero on V, then one of the functions itself must be zero on V. It cannot be that f vanishes on part of V and g vanishes on another part resulting in the product being zero everywhere.

[1] In general when we measure depth we assume that the circuit has fan-in two, but when we say constant depth we allow arbitrary fan-in. Constant depth circuits with arbitrary fan-in can be converted to log depth circuits with constant fan-in, but it is believed that the latter is a strictly more powerful model than the former.

From a mathematical point of view, this irreducible decomposition is one of the first theorems one encounters in algebraic geometry, and the rest of the discipline just focuses on irreducible varieties. Computationally however, performing the irreducible decomposition is a notoriously difficult task, and the general consensus is that this task must be avoided as far as possible (see [BM93, Section 4]).

Similar to satisfiability, deciding irreducibility can also be reduced to a purely algebraic task, that of computing primary decompositions, however this task is in EXPSPACE. Also similar to satisfiability, the special case of primary decompositions that arise from studying the irreducibility of varieties can be solved much more efficiently than EXPSPACE. In particular, Scheiblechner [Sch07] and Bürgisser and Scheiblechner [BS09] show that the number of irreducible components of V can be counted in PSPACE. We give a detailed account of the history of this problem in Chapter 4. Our main result here is that assuming the generalised Riemann hypothesis, and assuming that V is defined *ideal theoretically* [2] , testing irreducibility lies in the third level of the polynomial hierarchy. This brings us one step closer to exactly classifying this problem in terms of its complexity.

We remark that in our works so far, we only consider decision versions of the above problems. A natural next step is to study the computational tasks of actually computing common roots of a set of equations, and of computing defining equations of irreducible components of a variety. In general, the bit complexity of common roots and these defining equations is exponential in the size of the input, which in some sense puts a lower bound on what we can expect to say about these problems. Currently the search problems are known to be solvable in exponential time, and an interesting open problem is to solve these problems in PSPACE. Current approaches tend to use Gröbner basis methods for these tasks, which usually require exponential space to store.

1.3 Organisation of this thesis

This thesis is divided into four chapters, the first of which is this introduction. In Chapter 2 we will establish some global preliminaries that will be used throughout the thesis. The main technical content is in Chapter 3 and Chapter 4.

In Chapter 3 we define high degree Sylvester–Gallai configurations, and the connection with identity testing. We then prove dimension bounds on some of these configurations, and as a consequence design an algorithm for deterministic identity testing of $\Sigma^3\Pi\Sigma\Pi^d$ circuits. The content of this chapter is based on joint work with Rafael Oliveira, Shir Peleg, and Akash Kumar Sengupta. Parts of this chapter have been published [GOPS23; GOS25c; GOS25b].

In Chapter 4 we study the problems of irreducibility testing and satisfiability. The work on irreducibility testing is joint with Rafael Oliveira and Nitin Saxena, and has been published [GOS25a]. The work on satisfiability is joint with Robert Andrews and Éric Schost, and is under preparation.

[2] We say that V is defined ideal theoretically if we are given a generating set for $I(V)$. We say that V is defined *set theoretically* if we are given an ideal I such that $V(I) = V$.

Chapter 2

Preliminaries

In this chapter, we introduce some notation and preliminary material that will be used throughout this thesis. Most facts are stated without proof, and references to material that discuss each topic in detail are provided. Some preliminary material that is chapter specific is provided in the corresponding chapter itself.

2.1 Notation

We denote by S the polynomial ring $\mathbb{F}[x_1, \ldots, x_n]$, with the standard grading. The underlying field will be made clear from context, and will usually be $\mathbb{C}$. We denote polynomials by lower case letters, for example f, g, h. By a *homogeneous form* or simply just *form* we mean a homogeneous polynomial in S (or any other graded ring under consideration). We usually denote forms by upper case letters, for example F, G, H.

We use boldface to denote vectors, for example $\boldsymbol{\alpha}, \mathbf{f}$. We use vector notation to index monomials as follows. If $\boldsymbol{\alpha} \in \mathbb{N}^n$ then $\mathbf{x}^{\boldsymbol{\alpha}}$ denotes the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

Given polynomials $f_1, \ldots, f_m \in S$, the ideal generated by $f_1, \ldots, f_m$ is denoted by $(f_1, \ldots, f_m)$. The radical of the above ideal is denoted by $\operatorname{rad}(f_1, \ldots, f_m)$. We say two forms F, G are non-associate if $F \notin (G)$ and $G \notin (F)$.

We call a vector space $V \subset S$ *graded* if there exists a basis of V consisting of homogeneous forms. If V is graded, then we denote by V_i the degree i piece of V, that is, the vector space generated by all elements of V that have degree i. We also denote by S_i the degree i piece of S, and we have $V_i = V \cap S_i$ and $V = \oplus V_i$.

When $\mathbb{F}$ is algebraically closed (and clear from context), we denote by $\mathbb{A}^n$ and $\mathbb{P}^n$ the affine and projective spaces of dimension n over $\mathbb{F}$. In this setting, given any ideal $I \subset S$, we denote by $V(I) \subset \mathbb{A}^n$ the set of common zeroes of the elements in I. Given any variety V we denote by $I(V)$ the ideal of polynomials that vanish on V.

2.2 Algebraic Circuits

In this section we state some basic properties of algebraic circuits. In literature these are also called arithmetic circuits, we use the two terms interchangeably. For a much more detailed exposition of these circuits we refer the reader to [SY10; BCS97].

We start by defining algebraic circuits.

Definition 2.2.1 (Algebraic Circuits)**.** Let $\mathbb{F}$ be a field and let $\mathbb{F}(\mathbf{x})$ be the field of rational functions in the variables $x_1, \ldots, x_n$. An *algebraic circuit over* $\mathbb{F}$ is a directed acyclic graph. The vertices of this graph of in-degree zero are called *input gates* and are labelled by either a variable x_i or a field element $\alpha \in \mathbb{F}$. Vertices of positive in-degree are called *internal gates* and are labelled by an operation from $\{+, \times, \div\}$. Vertices of out-degree zero are called *output gates*. Each gate of the circuit naturally computes an element of $\mathbb{F}(\mathbf{x})$, assuming no division by zero takes place in the circuit, which we require. If $\{f_1, \ldots, f_m\}$ are the functions computed by the output gates of the circuit, we say that the circuit *computes* the functions $f_1, \ldots, f_m$. The *size* of the circuit is the number of wires in the circuit. The *depth* of the circuit is the length of the longest path from an input to an output of the circuit. $\lozenge$

If an algebraic circuit with divisions computes a polynomial function, then there exists a slightly larger algebraic circuit that computes the same function without divisions. As a result, for the rest of this article, whenever we say algebraic circuit we will mean division-free algebraic circuit, unless explicitly stated otherwise.

Theorem 2.2.2 (Division Elimination [Str73])**.** *Let $\mathbb{F}$ be an infinite field and let $f \in \mathbb{F}[\mathbf{x}]$ be a polynomial. Suppose there is an algebraic circuit of size s and depth Δ that computes f. Then there is a division-free algebraic circuit of size* $\mathrm{poly}(s, \deg(f))$ *and depth $\Delta + O(1)$ that computes f.*

We will often study algebraic circuits that have additional syntactic restrictions placed on them. The most common such restriction will be a restriction on the depth Δ of the algebraic circuit. We defer the description of these restrictions to the corresponding chapters.

Algebraic circuits are fairly robust models of computation, and are closed under certain natural operations. For example, derivatives of polynomials with small circuits have small circuits, as do certain coefficients of polynomials with small circuits. We capture some of these results in the following lemmas.

Lemma 2.2.3 (Interpolation)**.** *Let $\mathbb{F}$ be an infinite field let $f \in \mathbb{F}[\mathbf{x}, y]$ be a polynomial of degree d. Let $f_0, \ldots, f_d \in \mathbb{F}[\mathbf{x}]$ be polynomials such that*

$$f(\mathbf{x}, y) = \sum_{i=0}^{d} f_i(\mathbf{x}) y^i.$$

Suppose that f can be computed by an algebraic circuit of size s and depth Δ. Then $\{f_0, \ldots, f_d\}$ can be computed by an algebraic circuit of size $O(sd)$ and depth $\Delta + O(1)$.

Lemma 2.2.4 (Derivaties [BS83])**.** *Let $\mathbb{F}$ be an infinite field let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of degree d. Suppose that f can be computed by an algebraic circuit of size s and depth Δ where the gates have fan-in at most two. Then $\{\partial_1 f, \partial_2 f, \cdots, \partial_n f\}$ can be computed by an algebraic circuit of size $O(s)$ and depth $O(\Delta)$.*

The restriction on the fan-in in the above lemma will not be relevant to us, since we will only apply this lemma in the setting where we do not care about ensuring the depth of the circuit remains small.

For some parts of this thesis, we will use algebraic circuits as a way of providing a polynomial as an input to a Turing machine succinctly. This is a generalisation of the notion of the sparse representations of polynomials, where polynomials with few nonzero coefficients are given as a list of coefficient monomial pairs consisting of only the nonzero coefficients. When using circuits this way, we have to ensure that the field constants appearing in the circuit are integers. In this setting, we

will include the bit complexity of these integers in the size of the circuit. Such a circuit computes polynomials with integer coefficients. We call these circuits *constant free circuits*. We can then deduce the following bound on the bit complexity of a polynomial computed by a constant free circuit. Recall that for a polynomial $f \in \mathbb{Z}[x_1, \ldots, x_n]$ the height of f, denoted $h(f)$, is the maximum of the bit complexities of its coefficients.

Lemma 2.2.5 ([Bür00a, Lemma 4.16]). *Let $f \in \mathbb{Z}[x_1, \ldots, x_n]$ can be computed by a constant free algebraic circuit of size s. Then $h(f) \leq 2^{2s}$.*

2.2.1 Piecewise circuits

In Section 4.2, our intermediate arguments will use a model of computation that is slightly more powerful than algebraic circuits, which we now define. *Computational trees* are a model of computation obtained by giving arithmetic circuits the ability to branch based whether an internal node computes zero on a given input. While algebraic circuits can only compute rational functions, computational trees are able to compute certain piecewise rational functions. Examples of such piecewise rational functions include the greatest common divisor of univariate polynomials (assuming that the input is given by coefficients), and squarefree decompositions of univariate polynomials.

The full definition of computation trees allows fairly complicated branching logic, and we direct the interested reader to [BCS97, Chapter 4] for details. We do not require the full power of computation trees for our results. All we will require is the ability to look at a list of k gates in a fixed order, and select the first gate in the order that computes a non zero element. Following [AW24, Section 2.2], we introduce a select gate which implements the select function, defined as follows:

$$\text{select}(y_1, \ldots, y_k) := \begin{cases} y_1 & \text{if } y_1 \neq 0 \\ y_2 & \text{if } y_1 = 0 \text{ and } y_2 \neq 0 \\ y_3 & \text{if } y_1 = y_2 = 0 \text{ and } y_3 \neq 0 \\ & \vdots \\ y_k & \text{if } y_1 = \cdots = y_{k-1} = 0 \text{ and } y_k \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Also following [AW24], arithmetic circuits with select gates will be called *piecewise arithmetic circuits*. Piecewise arithmetic circuits are a special case of computation trees.

Definition 2.2.6. Let $\mathbb{F}$ be a field and let $\mathbb{F}(\mathbf{x})$ be the field of rational functions in the variables $x_1, \ldots, x_n$. A *piecewise arithmetic circuit over* $\mathbb{F}$ is a directed acyclic graph. The vertices of this graph of in-degree zero are called *input gates* and are labeled by either a variable x_i or a field element $\alpha \in \mathbb{F}$. Vertices of positive in-degree are called *internal gates* and are labeled by an operation from $\{+, \times, \div, \text{select}\}$. Vertices of out-degree zero are called *output gates*. Each gate of the circuit naturally computes a piecewise rational function $\mathbb{F}^n \to \mathbb{F}$, assuming no division by zero takes place in the circuit, which we require. If $\{f_1, \ldots, f_m\}$ are the functions computed by the output gates of the circuit, we say that the circuit *computes* the functions $f_1, \ldots, f_m$ piecewise. The *size* of the circuit is the number of wires in the circuit. The *depth* of the circuit is the length of the longest path from an input to an output of the circuit. ◊

It follows that piecewise arithmetic circuits (and computational trees in general) can only compute piecewise function where the domain is divided into constructible sets, that is, sets that can be defined using a system of polynomial equalities and inequalities.

In Section 4.2, our goal is to construct circuits for a polynomial function, namely the multivariate resultant. When computing polynomial functions, neither divisions nor branching give us extra computational power, as the next two lemmas will show. In other words, if we can construct a circuit for a polynomial that uses divisions and select gates, then we can modify this circuit to construct another circuit of almost the same size and depth that computes the same polynomial function, without divisions or select gates. In fact something stronger is true: if we can correctly compute a polynomial function on a Zariski open set by an arithmetic circuit with divisions and branching, then we can construct a circuit without divisions or branching that computes the polynomial function everywhere. This sort of branching elimination is possible in any computation tree, a proof can be found in [BCS97, Proposition 4.26]. Branching elimination is a consequence of the fact that computation trees can only divide their domain into constructible sets, a consequence of which is that one of the pieces must essentially compute the target function almost everywhere. For the special case of piecewise arithmetic circuits, a proof of branching elimination can be found in [AW24]. We remark that while [AW24, Lemma 2.5] is only stated for piecewise arithmetic circuits computing f everywhere, their proof works when everything is restricted to any open set.

Proposition 2.2.7 ([AW24, Lemma 2.5])**.** *Let $\mathbb{F}$ be an infinite field and let $f \in \mathbb{F}[\mathbf{x}]$ be a polynomial. Suppose there is a piecewise arithmetic circuit that correctly computes f on a Zariski open set U. Then there is an arithmetic circuit of the same size, depth, and fan-in that computes f everywhere.*

2.3 Algebraic Geometry

In Chapter 4 we use some basic results from algebraic geometry, roughly at the level of [CLO15] or [SR94]. We do not repeat the most basic algebraic geometry facts here, and instead focus on computational and effective aspects of the theory.

We start by defining the notion of the degree of a variety.

Definition 2.3.1. Given an irreducible variety V of dimension r, the degree of V denoted $\deg V$ is the number of points in $V \cap H$, where H is a general linear subspace of codimension r. Given a reducible variety V, the degree of V is the sum of the degrees of its irreducible components. $\lozenge$

A proof of the fact that the degree is well defined can be found in [Hei83]. For reducible varieties, the above degree differs from the classical notion of degree in algebraic geometry. The classical degree of a reducible variety is the sum of the degrees of its highest degree components, here we sum up the degrees of all components. This notion is more useful in computer science. The following statement is [Hei83, Theorem 1], and is referred to as Bézout's inequality.

Theorem 2.3.2. *Let $\mathbb{K}$ be an algebraically closed field and $X, Y \subseteq \mathbb{A}^n_{\mathbb{K}}$ be locally closed subsets. Then, we have*

$$\deg(X \cap Y) \leq \deg X \cdot \deg Y.$$

A careful inductive application of the Bézout's inequality implies the following inequality, which we also refer to as Bézout's inequality. The statement given below is a special case of [HS80, Proposition 2.3].

Theorem 2.3.3 (Degree bounds for closed sets)**.** *Let $\mathbb{K}$ be an algebraically closed field and $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ be polynomials of degree at most d. Then, we have*

$$\deg V(f_1, \ldots, f_m) \leq d^{\min(m,n)}.$$

The following lemma gives effective bounds on hyperplane sections that reduce the dimension of projective varieties, a proof can be found in [GSS19, Lemma 19]

Lemma 2.3.4. *Suppose $V \subset \mathbb{P}^n$ is a projective zero set of dimension r and degree D. If ℓ is a linear form where each coefficient is chosen uniformly and independently from a set $B \subset \mathbb{N}$, then $\dim V \cap Z(\ell) = r - 1$ with probability at least $1 - D/|B|$. If $\ell_1, \ldots, \ell_{r+1}$ are linear forms chosen the same way then $V \cap Z(\ell_1, \ldots, \ell_{r+1}) = \emptyset$ with probability at least $1 - (r+1)D/|B|$.*

Suppose $W \subset \mathbb{A}^n$ is an affine zero set of dimension r and degree D. If ℓ is a linear polynomial where each coefficient is chosen uniformly and independently from a set $B \subset \mathbb{N}$, then $\dim W \cap Z(\ell) = r - 1$ with probability at least $1 - 2D/|B|$. If $\ell_1, \ldots, \ell_{r+1}$ are linear polynomials chosen the same way then $W \cap Z(\ell_1, \ldots, \ell_{r+1}) = \emptyset$ with probability at least $1 - 2(r+1)D/|B|$.

The following is a sufficient condition for a linear map to be a Noether normalising map. This statement, and its proof can be found in [SR94, Chapter 1, Section 5, Theorem 1.15].

Lemma 2.3.5. *Suppose $V \subset \mathbb{P}^n$ is a projective variety disjoint from a k-dimensional linear subspace $E \subset \mathbb{P}^n$. Then the projection $\pi : X \to \mathbb{P}^{n-k-1}$ with centre E defines a finite map $X \to \pi(X)$.*

Combining the two lemmas above gives us the following effective Noether Normalisation theorem.

Lemma 2.3.6 (Effective Noether Normalisation)**.** *Suppose $V \subset \mathbb{P}^n$ is a projective zero set of dimension r and degree D. If $\ell_1, \ldots, \ell_{r+1}$ are linear forms where each coefficient is chosen uniformly and independently from a set $B \subset \mathbb{N}$, then the map $\pi : V \to \mathbb{P}^r$ with coordinate functions ℓ_i is a well defined finite map with probability at least $1 - (r+1)D/|B|$.*

Suppose $W \subset \mathbb{A}^n$ is an affine zero set of dimension r and degree D. If $\ell_1, \ldots, \ell_r$ are linear polynomials where each coefficient is chosen uniformly and independently from a set $B \subset \mathbb{N}$, then the map $\pi : W \to \mathbb{A}^r$ with coordinate functions ℓ_i is a well defined finite map with probability at least $1 - 2(r+1)D/|B|$.

Proof. For the projective case, by Lemma 2.3.4, the zero set $Z(\ell_1, \ldots, \ell_{r+1})$ is a linear space disjoint from V with probability at least $1 - (r+1)D/|B|$. If this holds, then π is a finite and well defined map by Lemma 2.3.5. The affine case follows by applying the projective case to the closure of W_p, and picking $\ell_{r+1} = x_0$. □

We now discuss the notion of a tangent space We state some standard facts about tangent spaces, which can be found in [SR94, Chapter 2].

Let I be a radical ideal generated by $f_1, \ldots, f_m$. Let $\mathcal{J}$ be the Jacobian of I, which is defined to be the matrix with $\mathcal{J}_{ij} = \partial f_i / \partial x_j$. At any point $x \in V$, the tangent space $T_x(V)$ is isomorphic to the kernel of $\mathcal{J}(x)$, where $\mathcal{J}(x)$ is the Jacobian entry wise evaluated at the point x [SR94, Chapter 2, Section 1, Theorem 2.1]. At every nonsingular point x, the dimension of the tangent space is exactly equal to dimension of the component of V passing through x [SR94, Chapter 2, Section 1, Theorem 2.3]. At every singular point x, the dimension of the tangent space is greater than the dimension of the components of V passing through x. The singular locus of V is a proper algebraic subset of V, and does not contain any irreducible component of V.

We deduce that there is a point $x_0 \in V$ such that $\dim T_{x_0}(V) < r$ if and only if V has an irreducible component of dimension less than r. The following lemma is an equivalent formulation of the above statement.

Lemma 2.3.7. *If V is a algebraic set of dimension r, and if $\mathcal{J}$ is the Jacobian of a radical ideal I such that $V = Z(I)$, then $\operatorname{rank} \mathcal{J}(x_0) > n - r$ for some $x_0 \in V$ if and only if V has an irreducible component of dimension less than r.*

One of the most fundamental results in algebraic geometry is Hilbert's Nullstellensatz. This theorem states that in any algebraically closed field $\mathbb{K}$, any set of polynomials $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ are unsatisfiable if and only if they generate the trivial ideal. Equivalently, $V(f_1, \ldots, f_m) = \emptyset$ if and only if there exists $g_1, \ldots, g_m$ such that $1 = \sum_i f_i g_i$. The following is an effective version of this theorem from [Jel05] that controls the degrees of $g_1, \ldots, g_m$.

Theorem 2.3.8 (Effective Nullstellensatz)**.** *Let $\mathbb{K}$ be an algebraically closed field, and $X \subset \mathbb{K}^m$ be an affine n-dimensional variety of degree D. Let $f_1, \ldots, f_m \in \mathbb{K}[X]$ be non-constant polynomials without common zeros, where $d_i := \deg f_i$, and $d_1 \geq \cdots \geq d_m$. Lastly, let*

$$N(d_1, \ldots, d_m; n) := \begin{cases} \prod_{i=1}^{m} d_i & \text{if } n > 1 \text{ and } n \geq m \\ d_m \cdot \prod_{i=1}^{n-1} d_i & \text{if } m > n > 1 \\ d_1 & \text{if } n = 1 \end{cases}$$

There exist polynomials $g_i \in \mathbb{K}[X]$ with

$$\deg(f_i g_i) \leq \begin{cases} D \cdot N(d_1, \ldots, d_m; n), & \text{if } m \leq n \\ 2D \cdot N(d_1, \ldots, d_m, n) - 1, & \text{if } m > n \end{cases}$$

such that $1 = \sum_{i=1}^{m} f_i g_i$.

2.4 Complexity Theory

We assume familiarity with the usual complexity classes such as $\mathsf{P}, \mathsf{NP}, \mathsf{PSPACE}, \mathsf{PH}, \mathsf{AM}$. In Chapter 4, we will make repeated use of the protocol of [GS86] to certify a lower bound on the size of a set with efficient membership tests. The following statement is from [AB09, Section 8.4.1].

Lemma 2.4.1. *Suppose $S \subset \{0,1\}^n$ is a set such that the problem of membership in S is in NP. Suppose further that a number K is known, and S is guaranteed to either satisfy $|S| \geq 2K$ or $|S| < K$. Then $|S| \geq 2K$ can be certified in AM.*

We sketch the protocol here, since it will make it easier to explain the generalisation that we require. However, we omit proofs. We also assume for convenience that K is a power of 2.

The input is a boolean formula $\phi_S(x, y)$ such that $x \in S$ if and only if there exists y satisfying $\phi_S(x, y)$, and an integer K. Let $k := \log(2K)$. Arthur picks a random hash function $h : \{0,1\}^n \to \{0,1\}^{k+1}$ from a pairwise independent hash function collection, and picks $u \in \{0,1\}^{k+1}$ uniformly at random. He sends h, u to Merlin. Merlin finds x, y such that $h(x) = u$ and $\phi_S(x, y)$ is true. He sends x, y back to Arthur. Arthur accepts if and only if $\phi_S(x, y)$ is true and $h(x) = u$.

We note that constant 2 in the above can be replaced by a slightly smaller constant, that is, if S is guaranteed to satisfy either $|S| \geq 1.9K$ or $|S| < K$, then the first condition can be certified in AM using essentially the same protocol, with slightly modified parameters. We now state a slight generalisation of this protocol, to the case when membership in S itself can be verified by an AM protocol.

Corollary 2.4.2. *Suppose for each $x \in \{0,1\}^n$ there exists a subset $S_x \subset \{0,1\}^{m(x)}$, and an integer $K(x)$, such that $m(x), K(x)$ are polynomially bounded functions of x. Suppose further that there is a uniform algorithm that runs in polynomial time that given input x, returns the number $K(x)$ and also returns a boolean formula ϕ_x, such that $z \in S_x$ if and only if there exists y such that $\phi_x(z, y)$ is true. Suppose further than an integer K is known.*

If $S \subset \{0,1\}^n$ is the set of elements x such that $|S_x| \geq 2K(x)$, and if S is promised to either satisfy $|S| \geq 2K$ or $|S| \leq K$, then $|S| \geq 2K$ can be certified in AM.

There is a natural four round protocol for the above: First Arthur picks a hash function and value, and Merlin sends a x that hashes to the given value. In the previous protocol, Merlin also sent a certificate for membership of the element x, since this was decidable in NP. Here, instead of just sending this certificate, Arthur and Merlin construct the formula ϕ_x and the threshold $K(x)$, and then perform an AM protocol to certify $|S_x| \geq 2K(x)$.

Chapter 3

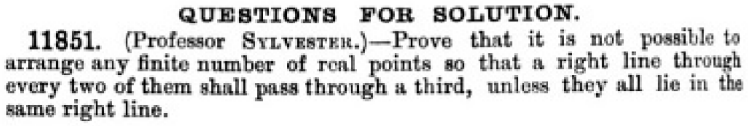
Identity testing and Sylvester–Gallai Configurations

In this chapter, we study high degree analogues of classical Sylvester–Gallai configurations. Classical Sylvester–Gallai configurations are configurations of points in affine or projective space that satisfy certain linear constraints. They have found numerous applications in computer science, some of which we list in the upcoming sections. One of these uses is in the identity testing of depth three circuits with constant top fan-in. Inspired by this, [Gup14] defined higher degree analogues of such configurations. These are essentially configurations of hypersurfaces in projective space that satisfy certain containment properties. Analogous to how classical Sylvester–Gallai configurations are used to design identity testing algorithms for depth three circuits, these higher degree Sylvester–Gallai configurations will help us design identity testing algorithms for certain depth four circuits.

This chapter is structured as follows. In Section 3.1 we discuss linear Sylvester–Gallai configurations, and numerous variations that have been studied. We also mention some applications to theoretical computer science. We state bounds on the dimensions of these linear configurations. When we define high degree configurations, these linear configurations will often act as base cases in our proofs by induction. In Section 3.2 we discuss the connection between identity testing and these Sylvester–Gallai configurations. This will motivate the definition of the higher degree Sylvester–Gallai configurations, which we will define in Section 3.3. In Section 3.4 we will sketch some high level ideas that go into the proof of dimension bounds for higher degree Sylvester–Gallai configurations. We will then make these concrete and show dimension bounds for some quadratic Sylvester–Gallai configurations in Section 3.5. Going from quadratics to higher degree forms requires a number of new tools, which we introduce in Section 3.6. In Section 3.7 we show some structural results about prime ideals, this is one of the main technical contributions of this chapter. In Section 3.8 and Section 3.9 we will show dimension bounds for some of the higher degree Sylvester–Gallai configurations that we introduce. As a consequence, we will obtain an algorithm for identity testing for depth four circuits with top fan-in three, and bottom fan-in constant. Finally, in Section 3.10 we use the methods developed to design an algorithm for identity testing for depth four circuits with constant top fan-in, and bottom fan-in two.

3.1 Linear Sylvester–Gallai configurations

The study of Sylvester–Gallai configurations started with the following question posed by Sylvester in [Syl93].



QUESTIONS FOR SOLUTION.

11851. (Professor SYLVESTER.)—Prove that it is not possible to arrange any finite number of real points so that a right line through every two of them shall pass through a third, unless they all lie in the same right line.

A proof of this impossibility was given by Gallai [Gal44], and independently by [Mel40], and this result is now called the Sylvester–Gallai theorem. Given a configuration of points, a line that passes through exactly two points is called an *ordinary* line. A line that is not ordinary is called a *special* line. An alternate way of stating the Sylvester–Gallai theorem is that any finite configuration of points in $\mathbb{R}^2$ without an ordinary line must be colinear. Equivalently, every configuration of points that is not colinear must contain an ordinary line. The theorem also holds if we consider points in $\mathbb{R}^n$ instead of $\mathbb{R}^2$.

If we instead consider points in the complex plane $\mathbb{C}^n$, then the above statement is no longer true. Indeed any cubic curve in $\mathbb{C}^2$ has nine inflection points, and a line joining any two of them passes through a third [Dic14]. However, these nine points are not colinear. Hirzebruch's work on line arrangements [Hir83] directly implies that every configuration of points in $\mathbb{C}^n$ without an ordinary line must be coplanar. This also answered a question of Serre [Ser66].

A number of generalisations of Sylvester–Gallai configurations have also been studied. These generalisations are usually obtained by relaxing the constraints on the points in the configuration. For example, Hansen [Han65] studied configurations of points in $\mathbb{R}^n$ with the constraint that every flat defined by k points contains at least $k+1$ points from the configuration. We call these high dimension Sylvester–Gallai configurations. Edelstein and Kelly [EK66] studied configurations where the points are divided into three sets, and any line through two points in two distinct sets passes through a third. We call these EK configurations. Szemerédi and Trotter studied configurations where the fraction of lines through each point that are ordinary is very small (although not 0). Equivalently, for each point p, for most other points q, the line through p and q is special. These are called robust Sylvester–Gallai configurations. In each case, it was shown that any configuration that satisfies such conditions must lie in a low dimensional linear subspace of the ambient space. We call theorems of this nature Sylvester–Gallai type theorems, or just Sylvester–Gallai theorems. For a more detailed discussion of these classical results, and a history of such theorems, we refer the reader to the survey [BM90]. We also remark that all the extensions we have discussed here are obtained by weakening the constraints that the points in the configuration satisfy. There exist other generalisations of the Sylvester–Gallai theorem that are also of interest, for example lower bounding the number of ordinary lines. The above survey also covers these variations.

In TCS, the above theorems have found applications in identity testing algorithms for circuits of depth three [KS09; SS13], and in the study of locally correctable codes [BDYW11; BDWY13]. The application to identity testing will be discussed in great detail in Section 3.2. These theorems have also found applications in circuit reconstruction [Sin16; SS25].

Most of the above discussion is about configurations of points in space. To easily define our high degree generalisations of Sylvester–Gallai configurations, and for our applications, we will use projective duality and instead consider configurations of hyperplanes in space. In this setting, lines through a pair of points correspond to intersections of hyperplanes. We can further define a hyperplane by a linear form in a polynomial ring, or by its normal vector. The condition that a

point lies on the line between two other points translates to a linear form being in the linear span of two other linear forms.

With this duality in mind, we will now formally define the types of linear Sylvester–Gallai configurations that are of interest to us, and we state appropriate bounds on the dimension of these configurations. We start with the definition of (r, δ)-linear-SG configurations.

Definition 3.1.1 (Robust linear Sylvester-Gallai configurations)**.** Let $r \in \mathbb{N}$, $0 < \delta \leq 1$ and V be a $\mathbb{K}$-vector space. Let $\mathcal{F} := \{v_1, \ldots, v_m\} \subset V$ be a finite set of *pairwise linearly independent* vectors. We say that $\mathcal{F}$ is a (r, δ)-linear-SG configuration over $\mathbb{K}$ if there exists a $\mathbb{K}$-vector subspace $U \subset V$ of dimension at most r such that the following condition holds:

- for any $v_i \in \mathcal{F} \setminus U$, there exist at least $\delta(m-1)$ indices $j \in [m] \setminus \{i\}$ such that $v_j \notin U$ and

$$|\mathrm{span}_{\mathbb{K}}\{v_i, v_j\} \cap \mathcal{F}| \geq 3 \quad \text{or} \quad \mathrm{span}_{\mathbb{K}}\{v_i, v_j\} \cap U \neq (0).$$

We will say that $\mathcal{F}$ is a (r, δ)-linear-SG configuration over the vector space U. ◊

A $(0, \delta)$-linear-SG configuration is just a robust Sylvester–Gallai configuration, and dimension bounds on such configurations (for all values of δ) were proved in [BDYW11], and then improved in [DSW14; DGOS18]. In a (r, δ)-linear-SG configuration, we allow some points to not satisfy the Sylvester–Gallai condition, although we require all such error points to lie in a small subspace. In typical applications, we will only have constantly many such error points. Such configurations with one error were studied in [Shp20], and dimension bounds for the general case were given in [OS24]. We state this bound for later use.

Proposition 3.1.2 ([OS24, Proposition 3.2])**.** *If $\mathcal{F}$ is a (r, δ)-linear-SG then* $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} \leq r + 12/\delta$.

In the special case when $r = 0$, a sharper bound is known.

Proposition 3.1.3 ([**dvir2018rank**])**.** *If $\mathcal{F}$ is a $(0, \delta)$-linear-SG then* $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} \leq 4/\delta$.

We now define higher dimensional linear Sylvester–Gallai configurations. Let $k \in \mathbb{N}$, $0 < \delta \leq 1$ and V be a $\mathbb{K}$-vector space. Let $\mathcal{F} := \{v_1, \ldots, v_m\} \subset V$ be a finite set of vectors that are pairwise linearly independent. We begin by defining ordinary and elementary spaces, as was done in [Han65; BDYW11].

Definition 3.1.4 (Ordinary spaces)**.** Let $u_1, \ldots, u_k \in \mathcal{F}$, and let $W = \mathrm{span}_{\mathbb{K}}\{u_1, \ldots, u_k\}$. The space W is called ordinary with respect to $\mathcal{F}$ if there exists an ordering $i_1, \ldots, i_k$ of $[k]$ such that $W \cap \mathcal{F} \subset \mathrm{span}_{\mathbb{K}}\{u_{i_1}, \ldots, u_{i_{k-1}}\} \cup \{u_{i_k}\}$. ◊

Definition 3.1.5 (Elementary spaces)**.** Let $u_1, \ldots, u_k \in \mathcal{F}$, and let $W = \mathrm{span}_{\mathbb{K}}\{u_1, \ldots, u_k\}$. The space W is called elementary with respect to $\mathcal{F}$ if $|W \cap \mathcal{F}| = k$. ◊

The above notions coincide when $k = 2$ and the set of vectors is pairwise linearly independent.

Definition 3.1.6. The set $\mathcal{F}$ is a $\delta-\mathrm{SG}_k^*$ configuration if for every linearly independent $v_1, \ldots, v_k \in \mathcal{F}$, there are $\delta \cdot |\mathcal{F}|$ forms u in $\mathcal{F}$ such that either

1. $u \in \mathrm{span}_{\mathbb{K}}\{v_1, \ldots, v_k\}$,
2. or the linear space $\mathrm{span}_{\mathbb{K}}\{v_1, \ldots, v_k, u\}$ contains a form in $\mathcal{F} \setminus (\mathrm{span}_{\mathbb{K}}\{v_1, \ldots, v_k\} \cup \{u\})$. ◊

Definition 3.1.7. The set $\mathcal{F}$ is a $\delta-\mathrm{SG}_k$ configuration if for every linearly independent $v_1, \ldots, v_k \in \mathcal{F}$, there are $\delta \cdot |\mathcal{F}|$ forms u in $\mathcal{F}$ such that either

1. $u \in \mathrm{span}_{\mathbb{K}}\{v_1, \ldots, v_k\}$,
2. or the linear space $\mathrm{span}_{\mathbb{K}}\{v_1, \ldots, v_k, u\}$ is not elementary. ◊

Observe that $\delta - \mathrm{SG}_k^*$ configurations are more constrained than $\delta - \mathrm{SG}_k$ configurations. The following bounds on the above configurations were proved in [DSW14], improving bounds proved in [BDYW11].

Theorem 3.1.8. *If $\mathcal{F}$ is a $\delta - \mathrm{SG}_k^*$ configuration then $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} = O(k/\delta)$. If $\mathcal{F}$ is a $\delta - \mathrm{SG}_k$ configuration then $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} = O(C^k/\delta)$ where C is a universal constant independent of k.*

The final linear Sylvester–Gallai type configuration we consider are partial linear Edelstein–Kelly (EK) configurations. These generalise the classical notion of EK configurations in three ways. The first is that some vectors are allowed to not satisfy any relationship, like the error vectors in (r, δ)-linear-SG configurations. The second is that a vector in a given set is only required to have span dependencies with vectors in the bigger of the remaining two sets. The third is that we only require dependencies with a constant fraction of this bigger set.

Definition 3.1.9 (Partial linear Edelstein–Kelly configurations)**.** Let $c \in \mathbb{N}$, $0 < \delta \leq 1$ and V be a $\mathbb{K}$-vector space. Let $\mathcal{F} := \{v_1, \ldots, v_m\} \subset V$ be a finite set of pairwise linearly independent vectors. We say that $\mathcal{F}$ is a (r, δ)-linear-EK configuration over $\mathbb{K}$ if there exists a subset $\mathcal{G} \subset \mathcal{F}$ of size at most r, and a partition $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ of the remaining forms $\mathcal{F} \setminus \mathcal{G}$ (with $|\mathcal{F}_1| \geq |\mathcal{F}_2| \geq |\mathcal{F}_3|$) such that the following condition holds:

- for any $v \in \mathcal{F}_1$ there are at least $\delta|\mathcal{F}_2|$ forms $u \in \mathcal{F}_2$ such that $|\mathrm{span}_{\mathbb{K}}\{v, u\} \cap (\mathcal{F}_3 \cup \mathcal{G})| \geq 1$.
- for any $v \in \mathcal{F}_2$ there are at least $\delta|\mathcal{F}_1|$ forms $u \in \mathcal{F}_1$ such that $|\mathrm{span}_{\mathbb{K}}\{v, u\} \cap (\mathcal{F}_3 \cup \mathcal{G})| \geq 1$.
- for any $v \in \mathcal{F}_3$ there are at least $\delta|\mathcal{F}_1|$ forms $u \in \mathcal{F}_1$ such that $|\mathrm{span}_{\mathbb{K}}\{v, u\} \cap (\mathcal{F}_2 \cup \mathcal{G})| \geq 1$.

◊

Such configurations were introduced by [Shp20] and further studied by [PS21]. The following is an improved bound on the dimensions of such configurations, and follows by a more careful analysis of the proof of [Shp20], and the above improved bound on the dimension of (r, δ)-linear-SG configurations.

Proposition 3.1.10. *If $\mathcal{F}$ is a* (r, δ)-linear-EK *configuration then*

$$\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} \leq r + c_{\mathrm{ek}} \cdot \delta^{-1} \log\left(\delta^{-1}\right)$$

for a universal constant c_{ek}.

Proof. Let $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{G}$ be a partition of $\mathcal{F}$ that satisfies the properties of the definition. Let $\delta' := \delta/2$. The main step is to find a subset $\mathcal{F}_1' \subset \mathcal{F}_1$ such that $|\mathcal{F}_2| \leq |\mathcal{F}_1'| \leq 2|\mathcal{F}_2|$ and such that $\mathcal{F}_1' \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{G}$ is itself a $(r, \delta'/2)$-linear-EK configuration, with partition $\mathcal{F}_1', \mathcal{F}_2, \mathcal{F}_3$. If such a set $\mathcal{F}_1'$ exists, then the above configuration is actually an $(r, \delta'/8)$-linear-SG, since each form satisfies as span relation with at least $\delta'/4$ fraction of the total forms. Therefore by Proposition 3.1.2 we can deduce that $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}_1' \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{G}\} \leq r + 96 \cdot \delta'^{-1}$. Every form $F \in \mathcal{F}_1 \setminus \mathcal{F}_1'$ is contained in $\mathrm{span}_{\mathbb{K}}\{\mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{G}\}$, since this form satisfies a linear condition with at least one form in $\mathcal{F}_2$. Therefore we can also deduce $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} \leq r + 96 \cdot \delta'^{-1} = r + 192 \cdot \delta^{-1}$, which will complete the proof.

It suffices therefore to show that such a set $\mathcal{F}_1'$ exists. We show that it exists as long as $|\mathcal{F}_2|$ is bigger than $c \cdot \delta^{-1} \log\left(\delta^{-1}\right)$ for some fixed constant c. If $|\mathcal{F}_2|$ is smaller than $c \cdot \delta^{-1} \log\left(\delta^{-1}\right)$ then so is $|\mathcal{F}_3|$. In this case we have $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{G}\} \leq r + 2c \cdot \delta^{-1} \log\left(\delta^{-1}\right)$, and therefore

$\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{F}\} \leq 2 + c \cdot \delta^{-1} \log\left(\delta^{-1}\right)$, which will also complete the proof. We show $\mathcal{F}_1'$ exists by a standard probabilistic method argument.

Let $m_i := |\mathcal{F}_i|$ and let $p := m_2/m_1$. If $m_3 \leq 2m_2$ then $\mathcal{F}_1' = \mathcal{F}_1$ satisfies the required properties and we are done, so we can assume $p \leq 1/2$. Let $\mathcal{F}_1' \subset \mathcal{F}_1$ be a subset constructed by picking each element of $\mathcal{F}_1$ independently with probability p.

For each $v \in \mathcal{F}_2$, let $\mathcal{F}_{\text{span}}(v) \subset \mathcal{F}_1$ be the set of forms u such that $|\operatorname{span}_{\mathbb{K}}\{v, u\} \cap (\mathcal{F}_3 \cup \mathcal{G})| \geq 1$. We have $|\mathcal{F}_{\text{span}}(v)| \geq \delta m_1$. The size of the subset $\mathcal{F}_1' \cap \mathcal{F}_{\text{span}}(v)$ is a random variable with expectation $|\mathcal{F}_{\text{span}}(v)| m_2/m_1 \geq \delta m_2$, and by a Chernoff bound, the probability that it has size less than $\delta' m_2$ is at most $2^{-a\delta m_2}$ for a fixed constant a. Similarly, for $v \in \mathcal{F}_3$ we can also define $\mathcal{F}_{\text{span}}(v)$, and the probability that $\mathcal{F}_{\text{span}}(v)$ has size less than $\delta' m_2$ is bounded by $2^{-a\delta m_2}$. Finally, we note that $|\mathcal{F}_1'|$ is itself a random variable with expectation m_2, and the probability that $|\mathcal{F}_1'| \geq 2m_2$ is bounded by 2^{-bm_2}. By a union bound, as long as $(m_2 + m_3)2^{-a\delta m_2} + 2^{-bm_2} < 1$, there is a set $\mathcal{F}_1'$ such that $|\mathcal{F}_1'| \leq 2m_2$ and such that $|\mathcal{F}_{\text{span}}(v) \cap \mathcal{F}_1'| \geq \delta' m_2$ for every $v \in \mathcal{F}_2 \cup \mathcal{F}_3$. Since $m_3 \leq m_2$, the above condition holds as long as $m_2 \geq c \cdot \delta^{-1} \log\left(\delta^{-1}\right)$ for a big enough constant c (this constant only depends on a, b).

Fix a set $\mathcal{F}_1'$ that satisfies the above conditions. If $|\mathcal{F}_1'| < m_2$, then we arbitrarily add elements to $\mathcal{F}_1'$ to ensure that $|\mathcal{F}_1'| = m_2$, this does not break any of the conditions. These conditions exactly guarantee that $\mathcal{F}_1' \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{G}$ is a $(r, \delta'/2)$-linear-EK configuration. This completes the proof. □

3.2 Polynomial identity testing and Sylvester–Gallai configurations

In this section, we discuss the connections between polynomial identity testing and the above configurations. We start by defining $\Sigma^k\Pi\Sigma$ circuits, and the notion of rank, and simple and minimal circuits.

For a fixed constant k, a $\Sigma^k\Pi\Sigma$ circuit is an algebraic circuit consisting of three layers of computation gates. The output gate is a Σ gate with fan-in k, the next layer consists of Π gates, and the lowest computational layer consists of Σ gates. These circuits compute polynomials with the following structure:

$$\prod_{i=1}^{d_1} \ell_{1,i} + \prod_{i=1}^{d_2} \ell_{2,i} + \cdots + \prod_{i=1}^{d_k} \ell_{k,i},$$

where each $\ell_{i,j}$ is a linear polynomial. We will usually work with homogeneous circuits, therefore we assume $\ell_{i,j}$ are linear forms and $d_1 = \cdots = d_k = d$. We write

$$C = T_1 + \cdots + T_k,$$

where $T_j = \prod_{i=1}^{d} \ell_{j,i}$. We say that a form ℓ *belongs to* T_i if $\ell \in (\ell_{i,j})$ for some j, and we say ℓ belongs to C if it belongs to some T_i. The *rank* of the circuit is the dimension of the linear space spanned by the forms $\ell_{i,j}$, equivalently the dimension of the forms that belong to C. The rank is the main parameter that will control the running time of algorithms for identity testing. The circuit is called *simple* if $\gcd(T_1, \ldots, T_k) = 1$, and is called *minimal* if no strict subset of $T_1, \ldots, T_k$ sums to 0.

Identity testing for this class of circuits was first studied by Dvir and Shpilka [DS07], who showed that these circuits have connections to locally decodable codes. Using this connection, they showed that the rank of any $\Sigma^k\Pi\Sigma$ circuit computing zero was bounded by polynomial function in $\log d$ and 2^{k^2}. They conjectured that the correct rank bound should be independent of d, and should only depend on k. They also noticed that in the case when $k = 3$, the set of forms $\ell_{i,j}$ that belong to a simple minimal $\Sigma^3\Pi\Sigma$ circuit computing zero formed an EK configuration. Indeed if we pick any ℓ_{1,j_1} and ℓ_{2,j_2}, then it must be that $T_3 \in (\ell_{1,j_1}, \ell_{2,j_2})$ whence $\ell_{3,j_3} \in (\ell_{1,j_1}, \ell_{2,j_2})$ for some j_3. Similar

relations with the gates swapped also hold. Crucially, since the top fan-in is three, no linear form can belong to more than one gate, since if a linear form belongs to two gates, it must belong to the third, contradicting simplicity.

Saxena and Seshadhri [SS11] improved this rank bound to polynomial in $k, \log d$, and Kayal and Saxena [KS07] proved that this cannot be improved in the case of finite fields. Subsequently, Karnin and Shpilka [KS11] proved that rank bounds will directly imply polynomial time black box identity testing for such circuits, using the notion of rank condensers introduced by Gabizon and Raz [GR08].

The first rank bounds for $\Sigma^k\Pi\Sigma$ circuits that are independent of d were shown by Kayal and Saraf [KS09], who showed that for simple minimal circuits over $\mathbb{R}$ computing zero, the rank is bounded by $2^{O(k \log k)}$. While not a polynomial function in k, this rank bound is independent of d, and therefore gives the first blackbox identity testing algorithm for this class of circuits. Their rank bounds rely on the high dimension linear Sylvester–Gallai theorem. At the time the theorem was only known over $\mathbb{R}$, subsequently the theorem was proved over $\mathbb{C}$ (Theorem 3.1.8), and therefore the technique of Kayal and Saraf also gives rank bounds for circuits over $\mathbb{C}$. Unlike the case of $k = 3$, the proof technique of Kayal and Saraf does not directly show that the linear forms belonging to C are a linear Sylvester–Gallai configuration. Instead they use the dimension bounds at intermediate points in their argument to control the dimension of certain sets of linear forms.

Subsequently, Saxena and Seshadhri [SS13] showed a stronger connection between linear Sylvester–Gallai configurations and the $\Sigma^k\Pi\Sigma$ circuits. They essentially proved that the set of forms belonging to T_1 are a $(2k^2, 0)$-linear Sylvester–Gallai configuration, which gives a bound on the dimension of $\ell_{1,j}$. The same holds for all gates, which gives a rank bound that is polynomial in k. In a follow up paper, [SS12] showed that an identity testing algorithm can be derived from an intermediate step in [SS13] itself, without having to show rank bounds. They were thus able to give an identity testing algorithm for $\Sigma^k\Pi\Sigma$ circuits for all fields, including finite fields where rank bounds do not hold.

The main difficulty in constructing linear Sylvester–Gallai configurations from circuits comes from the fact that when $k \geq 4$, even simple minimal circuits that compute zero can have linear forms that belong to more than one gate. The set of all forms that belong to C is therefore a multi set. While the multiplicities of the forms do not change the rank bounds, they do affect what relations can be derived between the linear forms, which makes these configurations hard to reason about.

We now consider depth four circuits with bounded top and bottom fan-in. We begin with $\Sigma^3\Pi\Sigma\Pi^d$ circuits, that is, circuits with top fan-in 3. Such circuits compute polynomials of the form

$$\prod_i P_i + \prod_i Q_i + \prod_i R_i,$$

where P_i, Q_i, and R_i are polynomials of degree at most d. Similar to the linear case we write $C = T_1 + T_2 + T_3$. The notion of simplicity and minimality extends to this case, as does the notion of a form belonging to a gate. We define the rank of the circuit to be the dimension of the linear span of A_i, B_i, C_i.

Gupta [Gup14] observed that similar to the linear case, $\Sigma^3\Pi\Sigma\Pi^d$ circuits computing zero satisfy a Sylvester–Gallai like condition. In particular, if we pick forms A_i and B_j, then we have $\prod_k C_k \in (A_i, B_j)$. Unlike the linear case, we cannot deduce that some fixed $C_{k'}$ lies in the ideal (A_i, B_j), since it might not be prime. He conjectured that any set of forms that satisfy this condition must have bounded dimension.

More generally, Gupta attempted to generalise the arguments of Kayal and Saraf to the case of $\Sigma^k\Pi\Sigma\Pi^d$ circuits. His approach was to try and separate the Sylvester–Gallai aspect of the proof from the other aspects of the proof. He proposed a number of different higher degree Sylvester–Gallai configurations, and conjectured that all of them have bounded dimension. If his conjectures hold, then the results of [BMS13] will give a polynomial time identity testing for $\Sigma^k\Pi\Sigma\Pi^d$ circuits. In fact

[BMS13] actually show the slightly stronger fact that rank bounds for simple and minimal $\Sigma^k\Pi\Sigma\Pi^d$ identities directly imply algorithms for identity testing.

Theorem 3.2.1 ([BMS13, Theorem 2])**.** *If r is a bound on the rank of simple and minimal $\Sigma^k\Pi\Sigma\Pi^d$ identities, then there is a deterministic algorithm for testing if a $\Sigma^k\Pi\Sigma\Pi^d$ circuit of size s in n variables computes zero that runs in time* $\mathrm{poly}(drsn)^{d^2kr}$.

In the next section, we will define some of these configurations, and in subsequent sections we will show that some of them have bounded dimension. One consequence is that we give the first polynomial time algorithm for the identity testing of $\Sigma^3\Pi\Sigma\Pi^d$ circuits. We note that while Gupta's conjectures for configurations arising from circuits with top fan-in greater than 3 are interesting in their own right, it seems that they are strictly harder to prove than rank bounds for $\Sigma^k\Pi\Sigma\Pi^d$ circuits. This is because separating out the Sylvester–Gallai aspect of the proof of Kayal and Saraf from the rest of the proof causes some loss in flexibility in what methods and techniques can be applied. Instead, it might be better to study the circuits the way that Kayal and Saraf studied depth three circuits, and use tools developed while studying high degree Sylvester–Gallai configurations at appropriate places in the argument to control various sets that turn up. By doing exactly this, we are able to obtain rank bounds and an identity testing algorithm for $\Sigma^k\Pi\Sigma\Pi$ circuits.

Before we end this section, we mention other related work on identity testing of $\Sigma^k\Pi\Sigma\Pi^d$ circuits. Dutta, Dwivedi, and Saxena [DDS21] gave a quasipolynomial time algorithm for $\Sigma^k\Pi\Sigma\Pi^d$ circuits, leveraging the Jacobian method. Their key insight involves applying the logarithmic derivative and its associated power series expansion to transform the top summation gate of the circuit into a powering gate. While this transformation technically violates the bounded top fan-in constraint, circuits with powering gates are well studied, and efficient PIT algorithms are known for such models. This reduction enables the use of existing tools for PIT in powering-gate circuits, ultimately leading to their quasipolynomial time result.

The breakthrough work of Limaye, Srinivasan, and Tavenas [LST21] on proving lower bounds for bounded-depth arithmetic circuits offers another route toward identity testing algorithms for this model. As discussed briefly above, the hardness-versus-randomness paradigm has been a powerful tool for establishing connections between circuit lower bounds and identity testing. Notably, the results of Chou, Kumar, and Solomon [CKS19] extended these trade offs to the bounded depth regime, showing that analogous principles continue to hold even in this restricted setting. By leveraging these developments, one obtains a subexponential time identity testing algorithm for depth 4 circuits. Building further on the lower bound techniques of Limaye, Srinivasan, and Tavenas [LST21], Andrews and Forbes [AF22] constructed another hitting set generator tailored to bounded depth circuits, yielding an alternative subexponential time identity testing algorithm with improved parameters. It is worth emphasizing, however, that neither of these approaches appears capable of yielding a *fully polynomial-time* algorithms for identity testing for $\Sigma^k\Pi\Sigma\Pi^d$ circuits.

3.3 High degree Sylvester–Gallai configurations

Recall that by projective duality, a linear Sylvester–Gallai configuration can be defined as a set of linear hyperplanes such that the intersection of any two hyperplanes in the set is contained in a third. The most basic versions of higher degree Sylvester–Gallai configurations directly generalise this definition. These configurations are collections of irreducible hypersurfaces of degree at most d, such that the intersection of any two hypersurfaces is contained in a third. Since $\mathbb{K}$ is algebraically closed, the Nullstellensatz allows us to state the above more conveniently as follows. Recall also that forms F, G are called non-associate if $F \notin (G)$ and $G \notin (F)$.

Definition 3.3.1 (Radical Sylvester–Gallai configuration)**.** Let $\mathcal{F} \subset \mathbb{K}[x_1, \ldots, x_N]$ be a finite set of pairwise non-associate irreducible forms of degree at most d. We say that $\mathcal{F}$ is a d-radical Sylvester-Gallai configuration if for every $F_i, F_j \in \mathcal{F}$, we have $F_k \in \mathrm{rad}(F_i, F_j)$ for some $F_k \in \mathcal{F} \setminus \{F_i, F_j\}$. ◊

The intersection of two hyperplanes is a linear subspace, which is irreducible. On the other hand, the intersection of two degree d hypersurfaces is a union of up to d^2 hypersurfaces. The above definition requires that F_k vanishes on every single irreducible component of $V(F_i, F_j)$. We can generalise this by allowing each irreducible component of $V(F_i, F_j)$ to lie in the zero set of different forms in our configuration. Since the number of irreducible components only depends on d, this condition still only depends on constantly many forms in the configuration. Such configurations are called product Sylvester–Gallai configurations, and we can succinctly represent the above condition as follows.

Definition 3.3.2 (Product Sylvester–Gallai configuration)**.** Let $\mathcal{F} \subset \mathbb{K}[x_1, \ldots, x_N]$ be a finite set of pairwise non-associate irreducible forms of degree at most d. We say that $\mathcal{F}$ is a d-product Sylvester-Gallai configuration if for every $F_i, F_j \in \mathcal{F}$, we have $\prod_{F \in \mathcal{F} \setminus \{F_i, F_j\}} F \in \mathrm{rad}(F_i, F_j)$. ◊

As we have seen in the previous section, the above types of configurations naturally show up when working with identity testing. We now define high dimensional analogues of high degree configurations. The following definition is a generalisation of linear independence of a set of vectors.

Definition 3.3.3 (Relevant sets)**.** Let $\mathcal{P} = \{P_1, \ldots, P_t\}$ be a set of forms in $\mathbb{K}[x_1, \ldots, x_N]_d$. We say that $\mathcal{P}$ is relevant if for every $1 \leq i \leq t$, $P_i \notin \mathrm{rad}(\mathcal{P} \setminus P_i)$. A relevant set of forms of size k is called a *k-relevant* set. ◊

Geometrically, a set $\mathcal{P}$ is relevant if no subset of $\mathcal{P}$ define the same variety as $\mathcal{P}$. We can now extend Definition 3.1.6 and Definition 3.1.7 to configurations with forms of higher degree.

Definition 3.3.4 (k-ordinary set)**.** Let $\mathcal{F} \subset \mathbb{K}[x_1, \ldots, x_N]$ be a finite set of pairwise non-associate irreducible forms of degree at most d. Let $\mathcal{P} \subset \mathcal{F}$ be a k-relevant set. We say that $\mathcal{P}$ is k-ordinary with respect to $\mathcal{F}$ if there are forms $F_1, \ldots, F_k \in \mathcal{F}$ such that

$$\mathrm{rad}(\mathcal{P}) \cap \mathcal{F} \subset \mathrm{rad}(F_1, \ldots, F_{k-1}) \cup \{F_k\}.$$

◊

Definition 3.3.5 (k-elementary set)**.** Let $\mathcal{F} \subset \mathbb{K}[x_1, \ldots, x_N]$ be a finite set of pairwise non-associate irreducible forms of degree at most d. Let $\mathcal{P} \subset \mathcal{F}$ be a k-relevant set. We say that $\mathcal{P}$ is k-elementary with respect to $\mathcal{F}$ if $\mathrm{rad}(\mathcal{P}) \cap \mathcal{F} = \mathcal{P}$. ◊

Definition 3.3.6 (Radical Sylvester Gallai condition for tuples)**.** Let $\mathcal{F} \subset \mathbb{K}[x_1, \ldots, x_N]$ be a finite set of pairwise non-associate irreducible forms of degree at most d. We say that $\mathcal{F}$ is a $\mathrm{SG}^*_k(d)$ configuration if for every k-relevant subset $\mathcal{P} \subset \mathcal{F}$, and every form $F \in \mathcal{F} \setminus \mathcal{P}$, either

- $F \in \mathrm{rad}(\mathcal{P})$ or
- $\mathrm{rad}(F, \mathcal{P}) \cap \mathcal{F}$ contains a form R not in $\mathrm{rad}(\mathcal{P}) \cup \{F\}$. ◊

The final configurations we study are generalisations of EK configurations.

Definition 3.3.7 (Non-linear Edelstein-Kelly configurations)**.** Let $\mathbb{K}$ be a field, $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ be finite sets of irreducible forms in $\mathbb{K}[x_1, \ldots, x_N]$ of degree at most d. We say that $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ forms a d-Edelstein-Kelly configuration over $\mathbb{K}$ if the following conditions hold:

1. any two forms are pairwise non-associate

2. for any $A \in \mathcal{A}, B \in \mathcal{B}$, we have that:

$$\prod_{C \in \mathcal{C}} C \in \mathrm{rad}(A, B).$$

Moreover, such relation works for any permutation of the sets $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$. ◊

The general high level question regarding these configurations is whether or not they have bounded dimension. The first such bound was proved by Shpilka [Shp20], who showed that 2-radical Sylvester–Gallai configurations have bounded dimension. In follow up work, Peleg and Shpilka showed that 2-product Sylvester–Gallai configurations and EK configurations of quadratics also have bounded dimension [PS20; PS21]. Pushing the same methods, Peleg and Shpilka and independently Garg, Oliveira, and Sengupta showed that robust versions of 2-radical Sylvester Gallai configurations also have bounded dimension [PS22; GOS22]. We will discuss these methods briefly in the next section, although we note that the above results seem to be the limit of what these methods can proof.

A major technical breakthrough was made by Oliveira and Sengupta [OS22], who showed that 3-radical Sylvester–Gallai configurations have bounded dimension. They introduced the notion of strong vector spaces[1] of linear and quadratic forms. These were inspired by the seminal work of Ananyan and Hochster on the resolution of a conjecture of Stillman [AH20a]. We will discuss these vector spaces in quite some detail in the upcoming sections, the exact properties and usefulness of these spaces are better explained once we have discussed the previous approach and its shortcomings.

Following this, in [GOPS23], we used these strong vector spaces to show that $\mathrm{SG}_k^*(2)$ configurations have bounded dimension. We give a full proof of this fact in Section 3.5. Simultaneously, Oliveira and Sengupta extended their results to all constant degrees [OS24]. They defined strong vector spaces of forms of all constant degrees, and showed that these spaces satisfy a list of nice properties. These results allowed them to reduce the degree of forms in d-radical Sylvester–Gallai configurations while controlling the dimension, and therefore they were able to show that d-radical Sylvester–Gallai configurations have bounded dimension for all d.

In [GOS25c], we showed that d-product Sylvester–Gallai configurations have bounded dimension. A full proof is given in Section 3.8. In [GOS25b] we showed that high degree EK configurations also have bounded dimension, and thus obtained the first polynomial time identity testing algorithm for depth four circuits with bounded bottom fan-in and top fan-in three. We give a full proof in Section 3.9 Finally, in Section 3.10 we show rank bounds and obtain the first identity testing algorithm for depth four circuits with bottom fan-in two and bounded top fan-in. This is joint work with Rafael Oliveira, Akash Kumar Sengupta, Nir Shalmon, and Amir Shpilka. This work has been submitted for peer review. At the point of writing this thesis, these are all the bounds we know for these configurations. In the next section, we start discussing some high level ideas of the proofs of the above dimension bounds.

3.4 Some high level ideas

In this section, we give some high level ideas of the proofs of the above dimension bounds. The goal is to try and motivate the definition of the strong vector spaces, and also to motivate some technical mathematical statements that are used in the proofs. We will roughly go over the above dimension bound results in the order that they appeared, but in some places we will forgo this for the sake of better exposition.

[1] They called them wide vector spaces at that point, and the name change took place in a later article.

I want to remark that this section mostly has my personal perspective on the ideas and techniques. I also take some liberty in rephrasing results, by making them fit into frameworks that were developed after results were shown. In my opinion, this benefits the exposition.

We will start by focusing on radical Sylvester–Gallai configurations. Let $\mathcal{F} := \{F_1, \ldots, F_m\} \subset \mathbb{K}[x_1, \ldots, x_N]$ be a set of forms of degree at most d. The following fact is a cornerstone for the dimension bounds for high degree Sylvester–Gallai configurations.

If every ideal (F_i, F_j) is radical, then the highest degree forms in $\mathcal{F}$ form a linear Sylvester–Gallai configuration.

Indeed if (F_i, F_j) is radical, then the condition that $F_k \in \mathrm{rad}(F_i, F_j)$ when F_i, F_j are of degree d reduces to $F_k \in (F_i, F_j)$, and since the ideal is generated by two homogeneous forms of degree d, and since F_k is also degree d, this reduces to $F_k \in \mathrm{span}_{\mathbb{K}}\{F_i, F_j\}$. This happens for every pair, therefore the degree d forms in the configuration are really a linear Sylvester–Gallai configuration in the coefficient vectors of the elements.

In fact we can make this statement weaker: if for each F_i, for most F_j the ideal (F_i, F_j) is radical, then we have a $(0, \delta)$-linear-SG configuration, and we have dimension bounds on the degree d forms in the configuration. Even if a few F_i fail this condition, we still have a (c, δ)-linear-SG configuration, and therefore dimension bounds.

Therefore, the hard case is when many of the ideals (F_i, F_j) are NOT radical. Here is a second cornerstone of the above dimension bounds.

If F and G are generic forms of degree d, then (F, G) is radical.

In other words, the condition that (F, G) is not radical is a special condition, and therefore we expect that if it happens a lot, we can deduce something about the forms. As an extreme example, if F, G are irreducible forms in distinct variables, then (F, G) is radical, therefore if the ideal is not radical, we can at least deduce that F, G share a variable. Of course just sharing a variable is not of much use, F and G could depend on all N variables and therefore always share variables, but we will see that we can deduce much more from the ideal being not radical.

Having established these two principles, we now go over the proof of the dimension bound for 2-radical Sylvester–Gallai configurations from [Shp20]. This will illustrate the above principles, and will also allow us to introduce a few more ideas that will be useful.

3.4.1 Dimension bounds for quadratic radical Sylvester–Gallai configurations

The early proofs of dimension bounds for quadratic forms, namely those in [Shp20; PS20; PS21; GOS22; PS22], relied on very strong forms of the second principle stated above. For example, the proof of [Shp20] relied on the following lemma regarding quadratic forms.

Lemma 3.4.1. *If $F, G \in \mathbb{K}[x_1, \ldots, x_N]$ are quadratic forms, then either (F, G) is radical, or $\ell_1\ell_2 \in (F, G)$ for linear forms ℓ_1, ℓ_2, or $(F, G) \subset (\ell_1, \ell_2)$ for linear forms ℓ_1, ℓ_2.*

When (F, G) is not radical, the above says much more than just that F, G share a variable, it says either $F = G + \ell_1\ell_2$, or that $F = \ell_1 v_1 + \ell_2 v_2$ and $G = \ell_1 u_1 + \ell_2 u_2$ for some linear forms. In [PS20; PS21; GOS22; PS22] the above statement was further refined.

We can use the notion of essential variables to rephrase the above lemma. If $F = \ell_1 v_1 + \ell_2 v_2$, then F has at most four essential variables, and if further F is irreducible, then it has at least three. If on the other hand, $F = G + \ell_1\ell_2$, then F and G almost have the same set of essential variables, there are only two variables in the symmetric difference. Therefore, the above statement says that if

(F, G) is not radical, either both of them have only three or four essential variables and two of them are common, or F, G only differ in two essential variables. This is a much stronger statement than the one we made earlier about sharing variables,

The rest of the proof of [Shp20] studies what happens if many pairs of forms satisfy the later two conditions. The latter conditions imply some relations between the essential variables of the forms involved, and enough such relations allow these forms to be controlled. Another technique that is introduced in [Shp20] that will be important throughout is the notion of a general projection, which we now explain. Suppose there exists a set of independent linear forms $v_1, \dots, v_r$ with r constant. By a change of basis (which does not affect the dimension of the configuration, or the Sylvester–Gallai condition), we can assume that these linear forms are variables $x_1, \dots, x_r$. If we now randomly specialise the variables $x_1, \dots, x_r$, then with high probability the remaining forms remain coprime, and still satisfy the Sylvester–Gallai condition. Further, if we can bound the dimension of $\mathcal{F}$ for a random setting of $x_1, \dots, x_r$, then the dimension of $\mathcal{F}$ itself must be bounded, essentially because the rank of the matrix of coefficients of the elements of $\mathcal{F}$ is unchanged by a random projection. To maintain homogeneity, instead of specialising x_i to α_i, we instead set it to $z\alpha_i$ for a new variable z, this is essentially the same as specialising and homogenising.

In [Shp20], the above projection is combined with the following observation: if v_1, v_2 are independent linear forms, then the set of forms $\mathcal{F} \cap (v_1, v_2)$ is itself a Sylvester–Gallai configuration, since for any two $F_1, F_2 \in (v_1, v_2)$ we have $\operatorname{rad}(F_1, F_2) \subset (v_1, v_2)$ since this latter ideal is prime. In this case, if a general projection is applied to the configuration $\mathcal{F} \cap (v_1, v_2)$, with respect to v_1, v_2, what we end up with (after factoring out z) is exactly a linear Sylvester–Gallai configuration. This implies that for any given form $F \in \mathcal{F}$, the set of all forms that satisfy the third condition of the above proposition with F have bounded dimension, and is a crucial step in the proof of [Shp20]. In subsequent work [PS20; PS21; PS22; GOS22], essentially the same ideas are used, with a stronger version of the above structural proposition.

Consider again Lemma 3.4.1. The result essentially states that if F, G both have many essential variables, say more then a large constant, then the only way they can fail to generate a radical ideal is if $F = G + u^2$ for a linear form u. This motivates one of the main ideas in [OS22]: if F is a form that depends on many essential variables, more than some large constant, then we should just treat F itself as a variable. This idea is also supported by the results of [AH20a], which we now briefly explain.

3.4.2 Ideals generated by a few low degree forms

This section follows the excellent survey by Erman, Sam, and Snowden [ESS19]. Given an ideal I in $S := \mathbb{K}[x_1, \dots, x_N]$, the projective dimension of the module S/I can be used to bound some aspects of the complexity of I, for example the degree and number of generators of a Gröbner basis for I (a proof of this and related results in the context of the current discussion can be found in [MS13]). Stillman conjectured that the projective dimension of an ideal generated by constantly many polynomials of bounded degree should only depend on the number of forms and the degrees of the forms, and not the number of variables in the polynomial ring that they lie in. For example, the projective dimension of four quadratics should be same if they are quadratics in ten variables or a million. There is a long line of work on this conjecture, the survey above and references within have a detailed account of its history and importance.

This conjecture was fully resolved by Ananyan and Hochster in [AH20a]. In fact they showed the following much stronger statement, before stating which we recall the definition of the strength of forms (the below definition differs from that in [AH20a] by 1, this does not change any of the results).

Definition 3.4.2. The strength of a homogeneous form $G \in \mathbb{K}[x_1, \ldots, x_N]$ is the smallest s such that $G = \sum_{i \leq s} P_i Q_i$ where P_i, Q_i have degree strictly smaller than G. We denote this by $s(G)$. If $G_1, \ldots, G_r$ are a collection of forms of the same degree, we denote by $s(G_1, \ldots, G_r)$ the smallest strength of any form non trivially spanned by $G_1, \ldots, G_s$. If V is a vector space of homogeneous forms of the same degree, we denote by $s(V)$ the minimum strength obtained by a non-zero form in V. ◊

Ananyan and Hochster proved the following: Suppose $G_1, \ldots, G_r$ are forms of degree d such every linear combination of $G_1, \ldots, G_r$ has large strength (where by large we mean bigger than a fixed function of r and d, but independent of the number of variables). Then $G_1, \ldots, G_r$ behave like a set of variables. The larger the strength of $G_1, \ldots, G_r$, the more like variables $G_1, \ldots, G_r$ behave. This phenomenon is now called the Ananyan–Hochster principle.

What it means for the forms to behave like variables can be made precise in a number of ways, and depends on the application at hand. In the context of Sylvester–Gallai configurations, what is required is that $G_1, \ldots, G_r$ form a prime sequence, and that the quotient is a UFD. This means that each quotient $S/(G_1, \ldots, G_i)$ is an integral domain and an UFD. Observe that if $G_1, \ldots, G_r$ were actually variables, then each quotient would be isomorphic to a polynomial ring, and would clearly be integral domains and UFDs.

In general, suppose we have forms $F_1, \ldots, F_r$ whose strength is smaller than the above function. After a linear transformation we can assume F_1 has small strength, say $F_1 = \sum P_i Q_i$. We can replace F_1 by the set of forms P_i, Q_i. We repeat this process of finding low strength forms in the span and replacing them with a set of forms of smaller degree. After finitely many steps, we can construct a set of forms $F'_1, \ldots, F'_s$ of high strength, with the property that $F_1, \ldots, F_r \subseteq \mathbb{K}[F'_1, \ldots, F'_s]$. The number of forms s is still independent of N, the number of variables in the polynomial ring. This essentially states that every set of constantly many forms of bounded degree lies in a subalgebra generated by constantly many forms that behave like variables (this is what the small subalgebra in the title of [AH20a] refers to). This is the main theorem of [AH20a].

Ananyan and Hochster also show the following technical lemma that is useful when studying Sylvester–Gallai configurations. Suppose $G_1, \ldots, G_s$ is a prime sequence, and let $R := \mathbb{K}[G_1, \ldots, G_s]$ be the subalgebra they generate. The algebra containment $R \subset S$ is *intersection flat*, and any prime ideal in R lifts to a prime ideal in S. This has many consequences. For example, if we have elements $F_1, F_2 \in R$, then we can compute the minimal primes of (F_1, F_2) in R, and the extension the minimal primes via the inclusion $R \to S$ gives the minimal primes in S.

3.4.3 Strong algebras of quadratics and a structure lemma for radicals

In [OS22], motivated by the above properties of forms with large strength, Oliveira and Sengupta defined the notion of wide Ananyan–Hochster vector spaces.

Definition 3.4.3 ([OS22])**.** Let $w, t \in \mathbb{N}$ be fixed. A graded vector space $V = V_1 + V_2 \subset \mathbb{K}[x_1, \ldots, x_N]$ is called (w, t)-wide AH space if $s(V_2) \geq t(\dim V + w)$. ◊

In other words, these are vector spaces where every quadratic form has high strength compared to the dimension of the vector space itself. In the case of quadratics, as long $t \geq 1$ and $w \geq 5$, any basis of V will satisfy the conditions discussed in Section 3.4.2, that is, any basis will be a prime sequence, and the quotient by any such basis will be a UFD. Further, [OS22] also showed that such wide spaces can be constructed starting from any graded vector space, that is, given any $U = U_1 + U_2$, there exists a (w, t)-wide AH space $V = V_1 + V_2$ such that $\mathbb{K}[U] \subset \mathbb{K}[V]$. The construction is based on the idea we discussed above: if V has a form of low strength then we just remove it from V and add the corresponding linear forms to V.

As an example, suppose we start with a single quadratic form Q. The space spanned by Q is a (w,t)-wide AH space if and only if $s(Q) \geq t(w+1)$. On the other hand, if $s(Q) < t(w+1)$, then the span of the essential variables of Q has dimension at most $2t(w+1)$. The space spanned by these variables is a (w,t)-wide AH space. In both cases, there exists a (w,t)-wide AH space of constant dimension such that the algebra generated by the space contains Q. The number of essential variables of Q itself might not be bounded, indeed it can be N. The Ananyan–Hochster principle states that we can instead use a basis of the above wide vector space as a set of essential variables for Q. This is the key idea in [OS22], and is fundamental to every following dimension bound for high degree Sylvester–Gallai configurations. As we will see in upcoming sections, this idea will generalise beyond quadratics.

We see another illustrative example. Before we discuss this example, we introduce the notion of relative essential variables, a small extension to the notion of essential variables. Recall that for a form Q, a set of essential variables is a set of linear forms $\ell_1, \ldots, \ell_r$ such that $Q \in \mathbb{K}[\ell_1, \ldots, \ell_r]$ and such that r is minimal with this property. This says that after a change of basis, Q only depends on r variables. Given a space of linear forms $U \subset \mathbb{K}[x_1, \ldots, x_N]$, we define the essential variables of Q over U to be a set of minimal linear forms $\ell_1, \ldots, \ell_r$ such that $Q \in \mathbb{K}[U, \ell_1, \ldots, \ell_r]$. This measures how many variables you need to write down Q, assuming that a given set of variables (those spanning U) can be used for free.

Suppose now that V is a $(w+1,t)$-wide AH space of constant dimension, and Q is a form not in $\mathbb{K}[V]$. Let W be the space obtained by adding Q to V. The space W might not be (w,t)-wide, there might be a quadratic form in W_2 whose strength is not large enough. Suppose $P \in W_2$ is such that $s(P) < t(w + \dim W)$. Every form in V_2 has strength at least $t(w+1+\dim V) = t(w+\dim W)$. Therefore, the above can only happen if $P \in W \setminus V_2$, that is, P is of the form $Q + P'$ for some $P' \in V_2$. Suppose $P = \sum u_i v_i$ is a minimal decomposition of P. In this case, we can create a new wide space by removing Q from W, and adding the linear forms u_i, v_i to W, call this space Y. The resulting space Y is also not (w,t)-wide, even though $Y_2 = V_2$, since the requirement on the strength of the elements of Y_2 is much higher than is guaranteed by V_2. Indeed for Y to be a (w,t)-wide space, we need every quadratic to have strength $t(w + \dim Y)$, which might be as high as $t(w + 2t(w + \dim V + 1) + \dim V + 1)$. However, Y is still (w',t')-wide for a smaller set of parameters w', t'. In this way, at the cost of some parameters, we have obtained a wide AH space Y of constant dimension with the following two properties: $V \subset Y$ and $Q \in \mathbb{K}[Y]$. Since $V \subset Y$, we can pick a basis of Y that extends a basis of V. Suppose this basis is $a_1, \ldots, a_p, b_1, \ldots, b_q$ where a_i are a basis for V. If w', t' are still big enough, then the Ananyan–Hochster principle applies, and these basis elements behave as variables. We can (and we will) think of $b_1, \ldots, b_q$ as a set of essential variables of Q over the vector space V, extending the previous analogy between wide AH spaces and essential variables. Things are more complicated when dealing with forms of higher degree, we will discuss these complications in a later section.

Oliveira and Sengupta also proved the following structural lemma in [OS22], which is a strong instantiation of the second cornerstone idea mentioned at the start of this section.

Theorem 3.4.4. *Suppose $S := \mathbb{K}[x_1, \ldots, x_N, y_1, \ldots, y_M]$ is a polynomial ring in two sets of variables, and suppose $P \in S$ is an irreducible form of degree d such that $P \notin (x_1, \ldots, x_N)$. Then there are at most $d^2(2d-1)$ pairwise non-associate irreducible forms $Q_i \in \mathbb{K}[x_1, \ldots, x_N]$ such that (P, Q_i) is not radical.*

The key part in the above lemma is that the number of forms Q_i is independent of the number of variables N, M and only depends on d. The way this lemma will be used is the following. Suppose we have constructed a (w,t)-wide AH space V, and we have a form Q that is not in the algebra generated by V. We can construct a (w',t')-wide AH space Y such that $Q \in \mathbb{K}[Y]$ and $V \subset Y$

as discussed above. If we construct a basis of Y extending a basis for V, then we can use the Ananyan–Hochster principle to apply the above structural lemma in the algebra generated by Y. As a consequence, if there exists more than $d^2(2d-1)$ forms in V that form a non-radical ideal with Q, we can deduce that $Q \in (V)$. In the coming sections, we will see how this is useful.

The rest of the argument of [OS22] for bounding the dimension of 3-radical Sylvester–Gallai configurations relies on the structure of ideals generated by two cubics, and does not generalise. Therefore, we do not discuss it further here. In the next section, we use the ideas above to show that $\mathrm{SG}^*_k(2)$ configurations have bounded dimension. As we will see, when dealing with quadratic forms, we can make a lot of the intuitive ideas above very concrete, much more than when we deal with higher degree forms.

3.5 High dimension Quadratic Sylvester–Gallai configurations

In this section we show that $\mathrm{SG}^*_k(2)$ configurations have bounded dimension. This is the main theorem of [GOPS23]. For the rest of this section, let $S := \mathbb{K}[x_1, \ldots, x_N]$ be a polynomial ring in N variables.

Section 3.5.1 will contain some preliminaries. This will include the definition of the linear space corresponding to a quadratic form, and relative versions of this space. These spaces will be a formal version of the discussion in the previous section, that is, spaces that we will use as proxies for essential variables. We prove some lemmas regarding these spaces. It will also include a formal definition and some properties of wide algebras and projection maps. Section 3.5.2 will introduce the notion of integral sequence and prove some basic properties. This will include the main structure theorem on radical ideals that we will use. Section 3.5.3 will introduce the notion of strong sequences and prove some properties. These will be used to create a vector space to which we can apply the results on integral sequences. Finally, in Section 3.5.4 we show the main theorem of this section, that $\mathrm{SG}^*_k(2)$ configurations have bounded dimension.

3.5.1 Preliminaries

Linear spaces corresponding to quadratic forms

Recall that the strength of a quadratic form $Q \in S$ is the smallest s such that we can write $Q = \sum_{i=1}^{s} a_i b_i$. We denote the strength of Q by $s(Q)$. A decomposition of the above nature is called a minimal decomposition. While a_i, b_i are not uniquely defined, the span $\mathrm{span}_{\mathbb{K}}\{a_i, b_i\}$ is uniquely defined, and is the same as the span of the first order partial derivatives of Q.

We now define $\mathbb{L}(Q)$, the linear space corresponding to quadratics (also also linear forms).

Definition 3.5.1 (Vector space of a quadratic form)**.** Let Q be a quadratic form of strength s, so that $Q = \sum_{i=1}^{s} a_i b_i$. Define the vector space $\mathrm{Lin}(Q) := \mathrm{span}_{\mathbb{K}}\{a_1, \ldots, a_s, b_1, \ldots, b_s\}$. Define $\mathbb{L}(Q)$ as:

$$\mathbb{L}(Q) = \begin{cases} \mathrm{span}_{\mathbb{K}}\{Q\}, & \text{if } s \geq 5 \\ \mathrm{Lin}(Q), & \text{otherwise.} \end{cases}$$

$\Diamond$

Definition 3.5.2. For a linear form $\ell \in S_1$ define $\mathbb{L}(\ell) := \mathrm{span}_{\mathbb{K}}\{\ell\}$. $\Diamond$

By the discussion in the previous section, $\mathbb{L}(Q)$ is a constant dimensional wide vector space whose algebra contains Q. The following claims easily follow from the fact that $\mathrm{Lin}(Q)$ is the span of the first order partial derivatives of Q. We omit the proofs here, they can be found in [GOPS23, Section 2.1].

Lemma 3.5.3 ([PS20, Claim 2.7]). *Suppose $Q \in \mathbb{K}[x_1, \dots, x_m]$ is a quadratic form of strength r. If y, z are new variables then $s(Q + yz) = r + 1$. In particular,* $\mathrm{Lin}(Q + yz) = \mathrm{Lin}(Q) + \mathrm{span}_{\mathbb{K}}\{y, z\}$.

Lemma 3.5.4. *Suppose $P \in \mathbb{K}[x_1, \dots, x_m]$ and $Q \in \mathbb{K}[y_1, \dots, y_n]$ are two quadratics in distinct variables. Then* $\mathrm{Lin}(P + Q) = \mathrm{Lin}(P) + \mathrm{Lin}(Q)$.

Lemma 3.5.5. *Let $W \subseteq S_1$ be a vector space. Suppose $Q \in S_2$ is such that $s(Q) = r$ in S and $s(\overline{Q}) = r' < r$ where $\overline{Q}$ is the image of Q in $S/(W)$. Then $W \cap \mathrm{Lin}(Q) \neq \{0\}$. In particular if $Q \in (W)$ then $W \cap \mathrm{Lin}(Q) \neq 0$.*

Wide vector spaces

Recall the notion of a (w, t)-wide AH space (Definition 3.4.3). In this section, we only require a special case, namely with $t = 1$. We refer to a $(r, 1)$-wide AH space as an r-wide space for convenience. We define the notion of a form being close to a vector space.

We list some properties that such spaces satisfy. The first is that such spaces can be constructed starting from any space,

Proposition 3.5.6 ([OS22], Proposition 4.11). *Suppose $U = U_1 + U_2$ is a vector space in S and suppose $r \in \mathbb{N}$. There exists an r-wide vector space $V = V_1 + V_2$ with $\mathbb{K}[U] \subseteq \mathbb{K}[V]$ such that* $\dim V_2 \leq \dim U_2$ *and* $\dim V_1 \leq 3^{\dim U_2 + 1} \cdot (r + \dim U)$.

The next result is a concrete instantiation of the Ananyan–Hochster principle.

Theorem 3.5.7 ([AH20b], Theorem 1.10). *Let $V \subset S_2$ be a vector space of dimension d such that $s(Q) \geq d - 1 + \lceil \eta/2 \rceil$. Then every sequence of linearly independent elements of V is an R_η sequence.*

The following lemma controls how much the parameters degrade when elements are added to wide spaces.

Remark 3.5.8. Suppose $V = V_1 + V_2$ is a r-wide vector space, and suppose $U \subset S_1$ is a vector space of dimension k. We have $\dim V + U \leq \dim V + k$. Further, we have $(V + U)_2 = V_2$. For every $Q \in (V + U)_2$ we therefore have $s(Q) \geq (r - k) + \dim(V + U)$. Therefore $V + U$ is a $r - k$-wide vector space. ◊

The following lemmas show that radical membership among linear forms and certain elements in the ideal (V) imply relationships between the "low strength" and "high strength" parts individually. These statements essentially follow from the definitions, we omit the proofs here and they can be found in [GOPS23, Section 5.1].

Lemma 3.5.9. *Let $F_1, \dots, F_k \in S_{\leq 2}$ be irreducible forms. Let $V = V_1 + V_2$ be r-wide with $r \geq k + 2$ and let $z \in V_1$. Suppose each F_i is either of the form $F_i = x_i$ with $x_i \in S_1$ or of the form $F_i = Q_i + zx_i$ with $Q_i \in V_2$ and $x_i \in S_1$. If*

$$F_k \in \mathrm{rad}(F_1, \dots, F_{k-1})$$

then $zx_k \in (x_1, \dots, x_k)$ and $Q_k \in \mathrm{span}_{\mathbb{K}}\{Q_1, \dots, Q_{k-1}\}$ where $Q_i = 0$ if $F_i \in S_1$.

Lemma 3.5.10. *Let $F_1, \dots, F_k \in S_{\leq 2}$ be irreducible forms. Let $V = V_1 + V_2$ be r-wide with $r \geq k + 2$ and let $z \in V_1$. Suppose each F_i is either of the form $F_i = x_i$ with $x_i \in S_1$ or of the form $F_i = Q_i + zx_i$ with $Q_i \in V_2$ and $x_i \in S_1$. Suppose further that $z, x_1, \dots, x_{k-1}$ are linearly independent. If*

$$F_k \in \mathrm{rad}(F_1, \dots, F_{k-1}),$$

and if $x_k \in (x_1)$ in $S/(z)$, then $F_1 = F_k$.

Relative linear spaces of quadratic forms

We can use the above definition of r-wide spaces to formally define relative notions $\mathbb{L}(Q)$.

Definition 3.5.11 (Forms close to a vector space)**.** Given a vector space $V = V_1 + V_2$ where $V_i \subseteq S_i$, we say that a quadratic form P is *s-close* to V if there is a form $Q \in \mathbb{K}[V]$ such that $s(P - Q) \leq s$. If a form P is not r-close to V, for any $r \leq s$, we say that P is *s-far* from V.

Given a linear form ℓ, we say ℓ is 1-close to V if $\ell \notin V_1$. ◊

The advantage of r-wide spaces in this context is that it implies the following uniqueness property, whose proof can be found in [GOPS23, Section 5.2].

Proposition 3.5.12 (Quadratics close to wide vector spaces)**.** *Let $V = V_1 + V_2$ be an r-wide vector space and $s < r/2$. If P is s-close to V, then for any $Q, Q' \in \mathbb{K}[V]$ such that $s(P-Q) = s(P-Q') = s$, we have that*

$$\mathrm{Lin}(P - Q) + V_1 = \mathrm{Lin}\big(P - Q'\big) + V_1.$$

In other words, $(\mathrm{Lin}(P - Q) + V_1)/V_1 = (\mathrm{Lin}(P - Q') + V_1)/V_1$ *for any two decompositions.*

The above uniqueness makes the following well defined.

Definition 3.5.13 (Relative space of linear forms)**.** Let r, B be integers such that $r > 2B + 1$. If V is an r-wide vector space and P is s-close to V for $s < r/2$ we can define

$$\mathbb{L}_V(P) := \begin{cases} \mathbb{L}(P) + V_1, \text{ if } P \in S_1 \\ \mathrm{Lin}(P - Q) + V_1, \text{ if } P \in S_2, s \leq B \\ \mathrm{span}_{\mathbb{K}}\{P\}, \text{ otherwise} \end{cases}$$

where $Q \in \mathbb{K}[V]$ is a form such that $s(P - Q) = s$. We also define the quotient space

$$\overline{\mathbb{L}}_V(P) := \begin{cases} \mathbb{L}_V(P)/V_1, \text{ if } s \leq B \\ 0, \text{ otherwise} \end{cases}$$

Further, we define P_V^H to be the unique polynomial in V_2 such that $P - P_V^H$ is s-close to V_1. Finally we define $P_V^L = P - P_V^H$. Note that $\mathbb{L}_V(P) = \mathbb{L}_V\big(P_V^L\big)$. The superscript H indicates that P_V^H is the high strength part of P with respect to V and the superscript L indicates that P_V^L is the low strength part of P with respect to V. ◊

We compare the above definition with the discussion on relative essential variables in the previous section. A basis for $\overline{\mathbb{L}}_V(P)$ forms a set of relative essential variables for P over V, if P is close enough to V. If not, then $V + P$ itself will be wide, and P itself can be assumed to be a variable by the Ananyan–Hochster principle.

Note that while the definition of $\mathbb{L}_V(P)$ depends on the parameter B, we suppress this from the notation for brevity. It will be clear from context the value of the parameter B whenever we use $\mathbb{L}_V(P)$.

Here are some useful results about relative linear spaces, and how they change when V is modified. Lemma 3.5.16 characterises exactly when $\dim \overline{\mathbb{L}}_V(F)$ is unchanged when $\mathbb{L}_V(G)$ is added to V_1. As the lemma shows, this happens when F and G do not share any common variables other than those that occur in V.

Proposition 3.5.14. *Suppose V is a r-wide space and P is s-close to V for $2s < r$. If $P \in (V)$ then $P_V^H \in V_2$ and $P_V^L \in (V_1)$.*

Proof. Since P is s-close to V we can write $P = P_V^H + P_V^L$. Since $P_V^H \in V_2$, we have $P_V^L \in (V)$ by assumption. We can write $P_V^L = P_2 + P_1$ with $P_1 \in (V_1)$ and $P_2 \in V_2$. In $S/(V_1)$ we have $P_2 = 0$. Since V is r-wide, this implies $P_2 = 0$ in S. Therefore $P_V^L \in (V_1)$. □

Proposition 3.5.15. *Suppose $V = V_1 + V_2$ is a r-wide vector space with $r > 2B + 1$, and suppose $P \in S_2$ is B-close to V. Then $Y := \mathbb{L}_V(P) + V_2$ is a $r - 2B$ wide vector space. If further $r > 4B + 1$ then for any other polynomial Q that is also B close to V we have $Q_V^H = Q_Y^H$.*

Proof. The first statement follows since Y is obtained by adding at most $2B$ linear forms to a basis of V. We now have $Q = Q_V^H + Q_V^L = Q_Y^H + Q_Y^L$ whence $Q_V^H - Q_Y^H = Q_Y^L - Q_V^L$. Here, we use the fact that $B < 4r + 1$ to ensure that Q_Y^H, Q_Y^L are well defined. Since both Q_Y^L, Q_V^L have strength at most B in $S/(Y)$ we obtain that $Q_V^H = Q_Y^H$. □

Lemma 3.5.16. *Suppose $V = V_1 + V_2$ is a r-wide vector space with $r > 4B + 1$, and suppose $F, G \in S_2$ are both B close to V. Let $Y := \mathbb{L}_V(G) + V_2$. Then the following hold.*

1. $\mathbb{L}_Y(F) = \mathbb{L}_V(G) + \mathbb{L}_V(F)$.
2. $\dim \overline{\mathbb{L}}_V(F) = \dim \overline{\mathbb{L}}_Y(F)$ *if and only if* $\overline{\mathbb{L}}_V(F) \cap \overline{\mathbb{L}}_V(G) = \{0\}$.
3. *If* $F \notin (V)$ *and* $\dim \overline{\mathbb{L}}_V(F) = \dim \overline{\mathbb{L}}_Y(F)$ *then* $F \notin (Y)$.

Proof. By Proposition 3.5.15 we have $H := F_V^H = F_Y^H$. Let P, R be such that $F - H - P = R$ with $P \in \mathbb{K}[V_1]$ and $\mathbb{L}_V(F) = \mathrm{Lin}(R) + V_1$. Let P', R' be such that $F - H - P' = R'$ with $P' \in \mathbb{K}[Y_1]$ and $\mathbb{L}_Y(F) = \mathrm{Lin}(R') + Y_1$. We have the equation $R' + P' = R + P$, which implies that $\mathrm{Lin}(R') + Y_1 = \mathrm{Lin}(R) + Y_1$. Since $V_1 \subset Y_1$, we have

$$\mathrm{Lin}\big(R'\big) + Y_1 = \mathrm{Lin}(R) + V_1 + Y_1. \tag{3.1}$$

Substituting $\mathbb{L}_V(F), \mathbb{L}_Y(F)$ in Eq. (3.1) and using the fact that $Y_1 = \mathbb{L}_V(G)$ we get $\mathbb{L}_Y(F) = \mathbb{L}_V(G) + \mathbb{L}_V(F)$.

Eq. (3.1) also implies

$$\overline{\mathbb{L}}_Y(F) = \frac{\mathrm{Lin}(R') + Y_1}{Y_1} = \frac{\mathbb{L}_V(F) + Y_1}{Y_1} = \frac{\mathbb{L}_V(F)}{\mathbb{L}_V(F) \cap Y_1}$$

therefore

$$\begin{aligned} \dim \overline{\mathbb{L}}_V(F) = \dim \overline{\mathbb{L}}_Y(F) &\iff \dim V_1 = \dim(\mathbb{L}_V(F) \cap Y_1) \\ &\iff V_1 = \mathbb{L}_V(F) \cap Y_1 \qquad \text{(since } V_1 \subseteq \mathbb{L}_V(F) \cap Y_1\text{)} \\ &\iff \{0\} = \overline{\mathbb{L}}_V(F) \cap \overline{\mathbb{L}}_V(G), \end{aligned}$$

proving the second item.

Assume now that $F \in (Y)$. By Proposition 3.5.14 we have $F - H \in (Y_1)$. Further, by assumption we have $F - H \notin (V_1)$. In $S/(V_1)$ we have $0 \neq \overline{F - H} = \overline{R} \in (Y_1)$ which in turn implies that $\mathrm{Lin}\big(\overline{R}\big) \cap Y_1 \neq \{0\}$ by Lemma 3.5.5. We have $R = \sum a_i b_i$ for linear forms $a_1, \ldots, a_t, b_1, \ldots, b_t$ where $\overline{a_i}, \overline{b_j}$ span $\overline{\mathbb{L}}_V(F)$. Therefore $\mathrm{Lin}\big(\overline{R}\big) = \sum \overline{a_i}\overline{b_i}$, whence $\mathrm{Lin}\big(\overline{R}\big) \subseteq \overline{\mathbb{L}}_V(F)$. This shows that $\overline{\mathbb{L}}_V(F) \cap \overline{\mathbb{L}}_V(G) \neq 0$, which by item 2 implies $\dim \overline{\mathbb{L}}_V(F) \neq \dim \overline{\mathbb{L}}_Y(F)$, contradicting the assumption. □

Note that the condition $\overline{\mathbb{L}}_V(F) \cap \overline{\mathbb{L}}_V(G) = \{0\}$ is symmetric in F and G. Therefore, we have that $\dim \overline{\mathbb{L}}_V(F) = \dim \overline{\mathbb{L}}_{\mathbb{L}_V(G)+V_2}(F)$ if and only if $\dim \overline{\mathbb{L}}_V(G) = \dim \overline{\mathbb{L}}_{\mathbb{L}_V(F)+V_2}(G)$. Further, in this case we have $F \notin (\mathbb{L}_V(G), V_2)$ and also $G \notin (\mathbb{L}_V(F), V_2)$ if $F, G \notin (V)$.

The above statement essentially allows us to compare the relative essential variables of two forms F, G as long as they are close enough to V. As we will see later, the above lemma is what makes quadratic forms significantly easier to deal with in this framework compared to higher degree forms. We will comment on this more as we proceed with the proof.

General projections

We now recall the definition and properties of projection maps from [Shp20; PS20; OS22]. Note that the following definition is a special case of a more general definition of such quotients defined by [OS24], which we will define later.

Definition 3.5.17 (Projection maps)**.** Let $S = \mathbb{K}[x_1, \cdots, x_N]$ be a polynomial ring. Let $W \subset S_1$ be a subspace of linear forms and $y_1, \ldots, y_t$ be a basis of W. Let $y_1, \ldots, y_n$ be a basis of S_1 that extends the basis $y_1, \ldots, y_t$ of W. Let z be a formal variable not in $\{y_1, \ldots, y_n\}$. For $\alpha = (\alpha_1, \ldots, \alpha_t) \in \mathbb{K}^t$, we define the projection map $\varphi_{\alpha,W}$ as the $\mathbb{K}$-algebra homomorphism $\varphi_{\alpha,W} : S \to \mathbb{K}[z, y_{t+1}, \ldots, y_n] = S[z]/(W)$ defined by

$$y_i \mapsto \begin{cases} \alpha_i z, & \text{if } 0 \leq i \leq t \\ y_i, & \text{otherwise} \end{cases}$$

$\Diamond$

For simplicity we will often drop the subscripts W or α, and write φ_α or φ for a projection map when there is no ambiguity about the vector space W or the vector α.

General projections. Fix a vector space $W \subset S_1$ as in Definition 3.5.17. We will say that a property holds for a *general projection* φ_α, if there exists a non-empty open subset $U \subset \mathbb{K}^t$ such that the property holds for all φ_α with $\alpha \in U$. Here $U \subset \mathbb{K}^t$ is open with respect to the Zariski topology, hence U is the complement of the zero set of finitely many polynomial functions on $\mathbb{K}^t$. The general choice of the element α defining a general projection φ_α allows us to say that such projection maps will avoid any finite set of polynomial constraints. As shown in [Shp20; PS20], general projection maps preserve several important properties of polynomials.

Proposition 3.5.18. *[OS22, Proposition 2.6] Let $F \in S$ be a polynomial and $W \subset S_1$ be a vector space of linear forms.*

(a) If $F \notin \mathbb{K}[W]$, then $\varphi(F) \notin \mathbb{K}[z]$ for a general projection $\varphi : S \to S[z]/(W)$.

(b) If $F \neq 0$, then $\varphi(F) \neq 0$ for a general projection.

(c) Suppose F is a form which does not have any multiple factors and $F \in (W)$. If $\varphi(F) = z^k G$ where $G \notin (z)$, then G does not have any multiple factors.

The next proposition is from [PS20, Claim 2.23].

Proposition 3.5.19. *Let $F, G \in S$ be two forms which have no common factor and $W \subset S_1$ a subspace of linear forms. For a general projection $\varphi : S \to S[z]/(W)$, we have* $\gcd(\varphi(F), \varphi(G)) \in \mathbb{K}[z]$. *In particular, if F, G are homogeneous then* $\gcd(\varphi(F), \varphi(G)) = z^k$ *for some $k \in \mathbb{N}$.*

The following result shows that general projections preserve linear independence for forms outside the algebra generated by W .

Corollary 3.5.20. *[OS22, Corollary 2.8] Let $F, G \in S$ be linearly independent irreducible forms and $W \subset S_1$ be a vector space of linear forms. If $F, G \notin \mathbb{K}[W]$ then $\varphi(F), \varphi(G)$ are linearly independent, for a general projection $\varphi : S \to S[z]/(W)$.*

The next proposition follows from [PS20, Claim 2.26].

Proposition 3.5.21. *Let $W \subset S_1$ be a vector space of linear forms. Let $\mathcal{F} \subset S_2$ be a finite set of quadratic forms. Suppose there is an integer $D > 0$ such that $\dim \operatorname{span}_{\mathbb{K}}\{\bigcup_{F\in\mathcal{F}} \mathbb{L}(\varphi(F))\} \leq D$ for a general projection $\varphi : S \to S[z]/(W)$. Then $\dim \operatorname{span}_{\mathbb{K}}\{\bigcup_{F\in\mathcal{F}} \mathbb{L}(F)\} \leq (D+1)\cdot \dim W$.*

The proposition above can be sharpened if we have extra information about the linear forms in $\mathcal{F}$. We state this sharpening in the next proposition

Proposition 3.5.22. *Let $W \subset S_1$ be a vector space of linear forms and $\mathcal{F} \subset S_2$ be a finite set of quadratic forms such that $\mathcal{F} \cap (W)$ and $s(F) < s$ for each $F \in \mathcal{F}$. Suppose there is an integer $D > 0$ such that $\dim \operatorname{span}_{\mathbb{K}}\{\bigcup_{F\in\mathcal{F}} \mathbb{L}(\varphi(F))\} \leq D$ for a general projection $\varphi : S \to S[z]/(W)$. Then we have $\dim \operatorname{span}_{\mathbb{K}}\{\bigcup_{F\in\mathcal{F}} \mathbb{L}(F)\} \leq (D+1)\cdot s$.*

Remark 3.5.23. Suppose $V = V_1 + V_2$ is a r-wide vector space and $\varphi := \varphi_{\boldsymbol{\alpha},V_1}$ is a projection mapping as defined in Definition 3.5.17. If $Q \in V_2$ is such that $s(\varphi(Q)) = a$ in $S[z]/(V_1)$ then $a-1 \leq s(Q) \leq a$ in $S/(V_1)$. Since V is r-wide, this proves that $a \geq r + \dim V_2$. Since $\dim \varphi(V_1) = 1$, and since $\dim \varphi(V_2) \leq \dim V_2$, we get $a \geq r - 1 + (\dim \varphi(V_1) + \dim \varphi(V_2))$. This shows that $\phi(V)$ is at least $r-1$ wide. ◊

3.5.2 Integral Sequences

In this subsection, we introduce the notion of integral sequences. These will be sequences of forms close to a vector space whose relative essential variables are disjointed.

Definition 3.5.24 (Integral Sequences)**.** Let r, B, t be integers with $r > 4tB+1$. Suppose $V = V_1+V_2$ is a r-wide vector space. Let $F_1, \ldots, F_t \in \mathcal{F}$ be a sequence of quadratic forms that are B-close to V. Let $U_0 := V$ and let $U_i := \mathbb{L}_{U_{i-1}}(F_i) + V_2$. The sequence $F_1, \ldots, F_t$ is called an integral sequence over V if for each i we have

- $\dim \overline{\mathbb{L}}_V(F_i) = \dim \overline{\mathbb{L}}_{U_{i-1}}(F_i)$, and
- $F_i \notin (V)$

When V is clear from context we just call $F_1, \ldots, F_t$ an integral sequence. ◊

In the rest of this section, we will assume that $r > 4tB + 1$. The following lemma and the two corollary that follows makes the statement regarding disjointed variables formal.

Proposition 3.5.25. *Suppose V is a r-wide vector space. Suppose $F_1, \ldots, F_t$ are a sequence of forms, and suppose $U_i := \mathbb{L}_{U_{i-1}}(F_i) + V_2$ with $U_0 := V$. Then*

1. *$U_t = \sum_{j=1}^{t} \mathbb{L}_V(F_j) + V_2$.*
2. *$\dim \overline{\mathbb{L}}_V(F_i) = \dim \overline{\mathbb{L}}_{U_{i-1}}(F_i)$ for every $2 \leq i \leq t$ if and only if for every $2 \leq i \leq t$ we have*

$$\overline{\mathbb{L}}_V(F_i) \cap \left(\sum_{j=1}^{i-1} \overline{\mathbb{L}}_V(F_j) \right) = \{0\}.$$

3. If additionally $F_i \notin (V)$ for every $1 \leq i \leq t$, then $F_i \notin (U_{i-1})$ for $2 \leq i \leq t$.

Proof. We prove the statements by induction on t. We will prove the additional statement that $\mathbb{L}_{U_{t-1}}(F_t) = \sum_{i=1}^{t} \mathbb{L}_V(F_t)$. Each of the three items are true by definition when $t = 1$. Suppose the statements are true for $t-1$.

Now the space U_{t-2} is $4B+1$ wide by Remark 3.5.8. Applying Lemma 3.5.16 to U_{t-2}, F_t, and F_{t-1} we can deduce that $\mathbb{L}_{U_{t-1}}(F_t) = \mathbb{L}_{U_{t-2}}(F_t) + \mathbb{L}_{U_{t-2}}(F_{t-1})$. The space U_{t-3} is also $4B+1$ wide, therefore applying Lemma 3.5.16 to U_{t-3}, F_t, and F_{t-2} we can deduce that $\mathbb{L}_{U_{t-2}}(F_t) = \mathbb{L}_{U_{t-3}}(F_t) + \mathbb{L}_{U_{t-3}}(F_{t-2})$. Repeating this and substituting, we deduce that $\mathbb{L}_{U_{t-1}}(F_t) = \mathbb{L}_V(F_t) + \sum_{i=2}^{t} \mathbb{L}_{U_{t-i}}(F_{t-i+1})$. By the induction hypothesis, we get that $\mathbb{L}_{U_{t-i}}(F_{t-i+1}) = \sum_{j=1}^{t-i} \mathbb{L}_V(F_j)$. Therefore we get $\mathbb{L}_{U_{t-1}}(F_t) = \sum_{i=1}^{t} \mathbb{L}_V(F_i)$. The first item now follows by adding V_2 to both sides.

Suppose now that $\dim \overline{\mathbb{L}}_V(F_i) = \dim \overline{\mathbb{L}}_{U_{i-1}}(F_i)$ for every $2 \leq i \leq t-1$. Suppose $\dim \overline{\mathbb{L}}_V(F_t) = \dim \overline{\mathbb{L}}_{U_{t-1}}(F_t)$. This implies $\dim \overline{\mathbb{L}}_{U_{t-1}}(F_t) = \dim \overline{\mathbb{L}}_{U_{t-2}}(F_t)$, since $V \subset U_{t-2} \subset U_{t-1}$. By item 2 of Lemma 3.5.16 applied to U_{t-2}, F_{t-1}, F_t we can deduce that $\overline{\mathbb{L}}_{U_{t-2}}(F_t) \cap \overline{\mathbb{L}}_{U_{t-2}}(F_{t-1}) = \{0\}$. Using the fact that $\mathbb{L}_{U_{t-1}}(F_t) = \sum_{i=1}^{t} \mathbb{L}_V(F_i)$, this is equivalent to $\overline{\mathbb{L}}_V(F_i) \cap \left(\sum_{j=1}^{i-1} \overline{\mathbb{L}}_V(F_j)\right) = \{0\}$. Conversely, starting with this assumption we can deduce that $\dim \overline{\mathbb{L}}_{U_{t-1}}(F_t) = \dim \overline{\mathbb{L}}_{U_{t-2}}(F_t)$. Note that $F_1, \ldots, F_{t-2}, F_t$ also satisfy the conditions of item 2. Therefore, by induction we can deduce that $\dim \overline{\mathbb{L}}_{U_{t-2}}(F_t) = \dim \overline{\mathbb{L}}_V(F_t)$. This completes the proof of item 2.

Applying the induction hypothesis to $F_1, \ldots, F_{t-2}, F_t$, we can deduce that $F_t \notin (U_{t-2})$. We can now apply Lemma 3.5.16 to U_{t-2}, F_t and F_{t-1} to deduce that $F_t \notin (U_{t-1})$, proving item 3. □

Corollary 3.5.26. *If $F_1, \ldots, F_t$ is a integral sequence, then so is any permutation of $F_1, \ldots, F_t$.*

Proof. The second condition for integral sequences holds irrespective of the order of the forms. By Proposition 3.5.25, the first condition for integral sequences is equivalent to

$$\overline{\mathbb{L}}_V(F_i) \cap \left(\sum_{j=1}^{i-1} \overline{\mathbb{L}}_V(F_j) \right) = \{0\}$$

for every $2 \leq i \leq t$. This in turn is equivalent to $\dim \sum_{j=1}^{t} \overline{\mathbb{L}}_V(F_j) = \sum_{j=1}^{t} \dim \overline{\mathbb{L}}_V(F_j)$, which is independent of the order of the forms. □

Corollary 3.5.27. *Let $F_1, \ldots, F_t$ be an integral sequence and $A := \mathbb{K}\left[V_2, \sum_{i=1}^{t} \mathbb{L}_V(F_i)\right]$. There exist vector spaces of linear forms $Y_1, \ldots, Y_t \subset A$ such that $Y_i \cap (V + Y_1 + \cdots + Y_{i-1}) = (0)$ for all i and $F_i \in \mathbb{K}[V, Y_i]$. Furthermore, $F_i \notin (V, Y_1, \cdots, Y_{i-1}, Y_{i+1}, \cdots, Y_t)$.*

Proof. By Proposition 3.5.25 we can take $Y_j := \overline{\mathbb{L}}_V(F_j)$. By Corollary 3.5.26, we may switch F_i and F_t. Then by Proposition 3.5.25 part (3), we see that $F_i \notin (V, Y_1, \cdots, Y_{i-1}, Y_{i+1}, \cdots, Y_t)$. □

We now show that integral sequences generate radical ideals, and that they are regular. An application of the above lemma along with the Ananyan–Hochster principle will reduce the statement to one about forms in almost disjoint variables, where the result is straightforward. We first show the statement for almost disjoint variables, and then show the reduction.

Lemma 3.5.28. *Suppose $F_1, \ldots, F_t$ are squarefree quadratic polynomials in $\mathbb{K}[x_1, \ldots, x_N, y_1, \ldots, y_t]$ such that $F_i \in \mathbb{K}[x_1, \ldots, x_N, y_i]$ and $F_i \notin (x_1, \ldots, x_N)$ for all i. Suppose also that $G \in \mathbb{K}[x_1, \ldots, x_N]$ is any polynomial. Then $G, F_1, \ldots, F_t$ are a regular sequence, and $(F_1, \ldots, F_t)$ is radical.*

Proof. If R is any ring, then in the polynomial ring $R[x]$, monic polynomials are always nonzerodivisors. The fact that $G, F_1, \dots, F_t$ is a regular sequence follows from the above fact by induction on t.

We now show the second fact, that the ideal is radical. Since $F_1, \dots, F_t$ is a regular sequence, every minimal prime of the ideal $(F_1, \dots, F_t)$ has codimension t. Further, the map $\mathbb{A}^{N+t} \to \mathbb{A}^N$ where we project to the x variables has finite fibers when restricted to the zero set of $(F_1, \dots, F_t)$, since for every specialisation of the x variables, there are at most two values for each y_i that make F_i vanish. Therefore the map is surjective when restricted to any irreducible component of this zero set, equivalently no element of $\mathbb{K}[x_1, \dots, x_N]$ is in any minimal prime of this ideal. We can thus check that the ideal is radical after passing to the field of fractions of $\mathbb{K}[x_1, \dots, x_N]$, and then since we are in characteristic 0, we can pass to the algebraic closure. In this algebraic closure, we have an ideal generated by squarefree forms in disjoint variables, and therefore the ideal is radical. □

Now that the statement has been established for polynomials in almost disjoint variables, we show how to reduce to the above case.

Lemma 3.5.29. *Suppose V is a r-wide vector space and $F_1, \dots, F_t$ is an integral sequence with respect to V. Suppose $F_0 \in \mathbb{K}[V] \setminus \{0\}$. Then $F_0, F_1, \dots, F_t$ is a regular sequence in S. Further $(F_1, \dots, F_t)$ is radical, and $F_1, \dots, F_t$ form a t-relevant set.*

Proof. By Corollary 3.5.27 we may assume that there exist vector spaces of linear forms $Y_1, \dots, Y_t$ in $A := \mathbb{K}\big[V_2, \sum_{i=1}^t \mathbb{L}_V(F_i)\big]$ such that $Y_i \cap (V + Y_1 + \cdots + Y_{i-1}) = (0)$ and $F_i \in \mathbb{K}[V, Y_i]$. Therefore, A is isomorphic to a polynomial ring $\mathbb{K}[V, Y_1, \dots, Y_t]$ where now $V, Y_1, \dots, Y_t$ are disjoint sets of variables, and $F_i \in \mathbb{K}[V, Y_i]$. If we further apply a random linear transformation to the variables in say Y_1, then F_1 will become monic with respect to the first variable in the set Y_1, which we call y_1. This follows since $F_1 \notin (V)$. We can apply this random transformation and then move all the remaining variables of Y_1 to V. This relabelling of the variables allows us to reduce to the case that each Y_i consists of a single variable y_i, and therefore we can invoke Lemma 3.5.28 to complete the proof. □

We note that the original proof of the above lemma in [GOPS23] was a more complicated argument using the discriminant lemma of [OS22].

3.5.3 Strong Sequences

The definition of integral sequences require that the forms are close to a vector space. This is required in Proposition 3.5.25 to give a sufficient condition for the number of relative essential variables to reduce when the vector space V is modified. In general, when we start with a Sylvester–Gallai configuration, we cannot guarantee that the forms are close to a vector space. In this section, we introduce the notion of strong sequences to handle this. We will show that either we can find a constant dimension vector space that every form in our configuration is close to, or we can directly bound the dimension of the configuration.

Definition 3.5.30. Let $k, t \in \mathbb{N}$ be such that $t \leq k + 1$. Given forms $Q_1, \dots, Q_t \in S_2$ we say that $Q_1, \dots, Q_t$ is a k-strong sequence if $Q_1, \dots, Q_t$ are linearly independent and $s(Q_1, \dots, Q_t) \geq k+5$. ◊

Remark 3.5.31. By Theorem 3.5.7, if $Q_1, \dots, Q_t$ is k-strong then $Q_1, \dots, Q_t$ is a R_3 sequence. By the discussion in [AH20b, Discussion 1.3], the ideal $(Q_1, \dots, Q_t)$ is prime and the ring $S/(Q_1, \dots, Q_t)$ is a UFD. ◊

Lemma 3.5.32. *Suppose $\mathcal{F}_2 \subset S_2$, and suppose $Q_1, \dots, Q_t$ is a maximal* k-strong *sequence in $\mathcal{F}_2$ with $t \leq k$. For any $r \geq 2(k+5)$ there exists a r-wide vector space W with $\dim W_1 \leq 7 \cdot r \cdot 3^t, \dim W_2 \leq t$ such that every $Q \in \mathcal{F}_2$ is $k+4$-close to W.*

Proof. Let $U := \mathrm{span}_{\mathbb{K}}\{Q_1, \dots, Q_t\}$. By Proposition 3.5.6, there exists r-wide vector space W such that $U \subset \mathbb{K}[W]$, $\dim W_1 \leq 3^{t+1} \cdot (r+t)$ and $\dim W_2 \leq t$. Let $Q \in \mathcal{F}_2$ be a form. Consider the sequence $Q_1, \dots, Q_t, Q$, which has length at most $k+1$. By assumption, $Q_1, \dots, Q_t, Q$ is not a k-strong sequence. Therefore, we have either $s(Q_1, \dots, Q_t, Q) \leq k+4$ or $Q \in \mathrm{span}_{\mathbb{K}}\{Q_1, \dots, Q_t\}$.

Suppose $P = \beta Q + \sum \alpha_i Q_i$ is such that $s(P) = s(Q_1, \dots, Q_t, Q) \leq k+4$. Since $Q_1, \dots, Q_t$ is k-strong we have $\beta \neq 0$. Therefore after scalar multiplication we have $Q = \sum \alpha_i Q_i + P$, and Q is $k+4$-close to W. If $Q \in \mathrm{span}_{\mathbb{K}}\{Q_1, \dots, Q_t\}$ then $Q \in W$ and therefore Q is $k+4$-close to W. □

We now define the notion of strong Sylvester-Gallai configurations. We show that a constant fraction of the forms in any such configuration is close to a vector space of constant dimension.

Definition 3.5.33. Let $\mathcal{F}_2 \subset S_2$ be a finite set of forms. Let $0 < \varepsilon \leq 1$ and $k, t \in \mathbb{N}$ with $t \leq k$. We say that $\mathcal{F}_2$ is a $\mathrm{strong}(\varepsilon, k) - \mathrm{SG}^*_t(2)$ configuration if for every k-strong sequence $Q_1, \dots, Q_t$ with $Q_i \in \mathcal{F}_2$, there are $\varepsilon(|\mathcal{F}_2| - 1)$ forms $Q_{t+1} \in \mathcal{F}_2$ such that either:

1. $Q_1, \dots, Q_t, Q_{t+1}$ is not a k-strong sequence, or
2. there is a form $R \in \mathcal{F}_2$ such that $R \in (Q_1, \dots, Q_{t+1}) \setminus (Q_1, \dots, Q_t) \cup (Q_{t+1})$.

◊

Lemma 3.5.34. *Let $\mathcal{F}_2 \subset S_2$ finite, with $m := |\mathcal{F}_2|$. Let $0 < \varepsilon \leq 1$ and $k, t \in \mathbb{N}$ with $2 \leq t \leq k$. If $\mathcal{F}_2$ is a $\mathrm{strong}(\varepsilon, k) - \mathrm{SG}^*_t(2)$ configuration then either*

1. *$\mathcal{F}_2$ is a $\mathrm{strong}(\varepsilon/4, k) - \mathrm{SG}^*_{t-1}(2)$ configuration, or*
2. *there exist a vector space W with $\dim W_1 \leq 7 \cdot r \cdot 3^{t+1+16/\varepsilon}, \dim W_2 \leq t + 1 + 16/\varepsilon$ such that at least $\varepsilon m/4$ forms in $\mathcal{F}_2$ are $k+4$ close to W.*

Proof. Let $\varepsilon' := \varepsilon/4$. Suppose $\mathcal{F}_2$ is not a $\mathrm{strong}(\varepsilon', k) - \mathrm{SG}^*_{t-1}(2)$ configuration. If there exist no k-strong sequences of length $t-1$, then there exists some maximal k-strong sequence of length at most $t-2$, and the required space W exists by Lemma 3.5.32. We can therefore assume that there exists a k-strong sequence $Q_1, \dots, Q_{t-1}$, and a set $\mathcal{B} \subset \mathcal{F}_2$ of size at least $(1-\varepsilon')m$ such that for every $Q \in \mathcal{B}$ we have that $Q_1, \dots, Q_{t-1}, Q$ is a k-strong sequence, and

$$\mathcal{F}_2 \cap (Q_1, \dots, Q_{t-1}, Q) \setminus (Q_1, \dots, Q_{t-1}) = \{Q\}. \tag{3.2}$$

Let $V := \mathrm{span}_{\mathbb{K}}\{Q_1, \dots, Q_{t-1}\}$. Forms $P_1, P_2 \in \mathcal{B}$ are pairwise independent over S_2/V, since if $(P_1) = (P_2)$ in S_2/V, then $P_2 \in (Q_1, \dots, Q_{t-1}, P_1) \setminus (Q_1, \dots, Q_{t-1}) \cup (P_1)$, contradicting $P_1 \in \mathcal{B}$.

Let $P \in \mathcal{B}$. The sequence $Q_1, \dots, Q_{t-1}, P$ is k-strong by definition of $\mathcal{B}$. Since $\mathcal{F}_2$ is a $\mathrm{strong}(\varepsilon, k) - \mathrm{SG}^*_t(2)$ configuration, there are $P_1, \dots, P_s \in \mathcal{F}_2$ with $s \geq \varepsilon m$ such that either $Q_1, \dots, Q_{t-1}, P, P_i$ is not k-strong, or there is $R_i \in \mathcal{F}_2$ such that $R_i \in (Q_1, \dots, Q_{t-1}, P, P_i) \setminus (Q_1, \dots, Q_{t-1}, P) \cup (P_i)$.

Let $\mathcal{G} := \{P_i \mid Q_1, \dots, Q_{t-1}, P, P_i \text{ is not a k-strong sequence}\}$. Let W be the r-wide vector space obtained by applying Proposition 3.5.6 to $V + \mathrm{span}_{\mathbb{K}}\{P\}$, we have $\dim W_1 \leq 7 \cdot r \cdot 3^t, \dim W_2 \leq t$. Every form in $\mathcal{G}$ is $k+4$-close to W. Hence, if $|\mathcal{G}| \geq \varepsilon' m$ then we are done.

We are left with the case that $|\mathcal{G}| \leq \varepsilon' m$. After relabelling, let $P_1, \dots, P_{s'}$ be the forms that are in $\mathcal{B} \setminus \mathcal{G}$. Since $|\mathcal{B}| \geq (1-\varepsilon')m$ and $|\mathcal{G}| \leq \varepsilon' m$ we have $s' \geq (\varepsilon - 2\varepsilon')m$.

Now for each $i \leq s'$, there is a form $R_i \in \mathcal{F}_2 \cap (Q_1, \ldots, Q_{t-1}, P, P_i) \setminus (Q_1, \ldots, Q_{t-1}, P) \cup (P_i)$ say $R_i = \sum \alpha_j Q_j + \beta P + P_i$. Since $P_i \in \mathcal{B}$ we have $\beta \neq 0$. Suppose $P_1, \ldots, P_{s''}$ are such that

$$\mathcal{B} \cap ((Q_1, \ldots, Q_{t-1}, P, P_i) \setminus (Q_1, \ldots, Q_{t-1}, P)) = \{P_i\}. \tag{3.3}$$

If $R_i = \alpha R_j$ with $\alpha \neq 0$ for $i, j \leq s''$, then we have $\alpha P_j = \sum \alpha'_i Q_i + P_i + \beta' P$, contradicting Eq. (3.3) for P_i. Therefore we have $s'' \leq |\mathcal{F}_2 \setminus \mathcal{B}| \leq \varepsilon' m$. Hence, there are at least $\varepsilon' m$ forms P_i such that $|\mathrm{span}_{\mathbb{K}}\{P, P_i\} \cap \mathcal{B}| \geq 3$ in S_2/V. Since this holds for every $P \in \mathcal{B}$, the set $\mathcal{B}$ is a $(\varepsilon', 2)$-linear-SG configuration in S_2/V. By Proposition 3.1.3 we have that $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{B}\} \leq 4/\varepsilon'$ in S_2/V and that $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{B}\} + V \leq t + 1 + 4/\varepsilon'$. Applying Proposition 3.5.6 to $\mathrm{span}_{\mathbb{K}}\{\mathcal{B}\} + V$ gives us a r-wide vector space W with $\dim W_1 \leq 7 \cdot r \cdot 3^{t+1+4/\varepsilon'}, \dim W_2 \leq t + 1 + 4/\varepsilon'$ and $\mathcal{B} \subset W$. □

Lemma 3.5.35. *Let $\mathcal{F}_2 \subset S_2$ finite, with $m := |\mathcal{F}_2|$. Suppose $\mathcal{F}_2$ is a $\mathrm{strong}(\varepsilon, k) - \mathrm{SG}^*_1(2)$ configuration. Then there is a r-wide vector space W with $\dim W_1 \leq 7 \cdot r \cdot 3^{2+16/\varepsilon}, \dim W_2 \leq 2 + 16/\varepsilon$ such that at least $\varepsilon m/4$ forms in $\mathcal{F}_2$ are $k + 5$ close to W.*

Proof. Let $\varepsilon' := \varepsilon/4$. Let $\mathcal{B}$ be the set of forms in $\mathcal{F}_2$ of strength at least $k + 5$. If $|\mathcal{B}| \leq (1 - \varepsilon')m$, then there are at least $\varepsilon' m$ forms that are $k + 5$ close to the zero vector space and we are done with $W = 0$. We are left with the case when $|\mathcal{B}| \geq (1 - \varepsilon')m$.

Let $P \in \mathcal{B}$. Let $\mathcal{G} := \{P_i \mid P, P_i \text{ is not a k-strong sequence}\}$. Let W be the r-wide vector space obtained by applying Proposition 3.5.6 to $\mathrm{span}_{\mathbb{K}}\{P\}$, we have $\dim W_1 \leq 21 \cdot r, \dim W_2 \leq 1$. Every form in $\mathcal{G}$ is $k + 4$ close to W. If therefore $|\mathcal{G}| \geq \varepsilon' m$ then we are done. We are left with the case that $|\mathcal{G}| \leq \varepsilon' m$.

Suppose $P_1, \ldots, P_{r'}$ are the forms in $\mathcal{B} \setminus \mathcal{G}$ such that P, P_i is a k-strong sequence and there exist $R_i \in (P, P_i) \setminus (P) \cup (P_i)$. We have $r' \geq 2\varepsilon' m$. Suppose $P_1, \ldots, P_{r''}$ are such that $(P, P_i) \cap \mathcal{B} = \{P, P_i\}$. If $R_i = \beta R_j$ for $i, j \leq r''$ then $P_j \in \mathrm{span}_{\mathbb{K}}\{P, P_i\}$, contradicting choice of P_i. Therefore there are at least $\varepsilon' m$ many forms P_i such that $|(P, P_i) \cap \mathcal{B}| \geq 3$. Since this holds for every P, we have that $\mathcal{B}$ is a $(\varepsilon', 2)$-linear-SG, and by Proposition 3.1.3 we have that $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{B}\} \leq 4/\varepsilon'$. If W is the r-wide space obtained by applying Proposition 3.5.6 to $\mathcal{B}$ then $\dim W_1 \leq 7 \cdot r \cdot 3^{4/\varepsilon'}, \dim W_2 \leq 2 + 4/\varepsilon'$ and $\mathcal{B} \subset W$, completing the proof. □

Corollary 3.5.36. *Suppose $\mathcal{F} = \mathcal{F}_1 \sqcup \mathcal{F}_2$ be a $\mathrm{SG}^*_k(2)$ configuration with $|\mathcal{F}_2| = m_2$. Then there exist a r-wide vector space W with $\dim W_1 \leq 7 \cdot r \cdot 3^{k+1+16 \cdot 4^{k-1}}, \dim W_2 \leq k + 1 + 16 \cdot 4^{k-1}$ such that at least $m_2/4^k$ forms in $\mathcal{F}_2$ are $k + 5$ close to W.*

Proof. We first show that $\mathcal{F}_2$ is a $\mathrm{strong}(1, k) - \mathrm{SG}^*_k(2)$ configuration. Suppose $Q_1, \ldots, Q_k$ is a k-strong sequence. Every subset of $Q_1, \ldots, Q_k$ is also a k-strong sequence, and hence generates a prime ideal by Remark 3.5.31. By definition $Q_1, \ldots, Q_k$ are linearly independent, therefore $Q_1, \ldots, Q_k$ form a k-relevant set. For every $Q_{k+1} \in \mathcal{F}_2$, if $Q_1, \ldots, Q_{k+1}$ is a k-strong sequence, then $(Q_1, \ldots, Q_{k+1})$ is prime and $Q_{k+1} \notin \mathrm{rad}(Q_1, \ldots, Q_k)$. Therefore there exists $R \in \mathcal{F}$ such that $R \in (Q_1, \ldots, Q_{k+1}) \setminus (Q_1, \ldots, Q_k) \cup (Q_{k+1})$. Since $Q_i \in \mathcal{F}_2$ it must be that $R \in \mathcal{F}_2$. This shows that $\mathcal{F}_2$ is a $\mathrm{strong}(1, k) - \mathrm{SG}^*_k(2)$ configuration.

Now let $t \geq 1$ be the smallest number such that $\mathcal{F}_2$ is a $\mathrm{strong}\left(4^{k-t}, k\right) - \mathrm{SG}^*_t(2)$ configuration. By the previous paragraph, we have $t \leq k$. If $t = 1$, the required vector space exists by Lemma 3.5.35. If $t > 1$, we apply Lemma 3.5.34. Since $\mathcal{F}_2$ is not a $\mathrm{strong}\left(4^{k-t-1}, k\right) - \mathrm{SG}^*_{t-1}(2)$ configuration, case 1 of Lemma 3.5.34 does not hold, Therefore there exists a vector space W with $\dim W_1 \leq 7 \cdot r \cdot 3^{t+1+16 \cdot 4^{k-t}}, \dim W_2 \leq t + 1 + 16 \cdot 4^{k-t}$ such that at least $4^{k-t-1} \cdot m_2$ forms in $\mathcal{F}_2$ are $k + 4$ close to W. □

Finally, in the next section we will put the above together to show the main theorem of this section, that $\mathrm{SG}^*_k(2)$ configurations have bounded dimension.

3.5.4 Proof of the Sylvester–Gallai theorem

In this section, we prove our main theorem: $\mathrm{SG}_k^*(2)$ configurations have constant vector space dimension. Throughout this section we denote our $\mathrm{SG}_k^*(2)$ configuration by $\mathcal{F} = \mathcal{F}_1 \sqcup \mathcal{F}_2$ where $\mathcal{F}_d$ is the set of forms of degree d in our configuration. Additionally, we define $m := |\mathcal{F}|$, $m_1 := |\mathcal{F}_1|$ and $m_2 := |\mathcal{F}_2|$.

Our proof has three main steps. In Section 3.5.4 we show that given $\mathcal{F}$, we can find a constant dimensional wide vector space W such that $\mathcal{F}$ is close to W. We call any such $\mathbb{K}$-algebra $\mathbb{K}[W]$ a *core algebra* of our configuration $\mathcal{F}$. This step uses the notion of strong sequences. In Section 3.5.4 we show that given such a vector space W, we can extend it to obtain a constant dimensional wide vector space $W \subset V$ such that $\mathcal{F}_2 \subset (V)$. This step uses the notion of integral sequences. In Section 3.5.4 we show that our main theorem follows given such a vector space V. This step uses general projections and the bound for linear Sylvester–Gallai configurations from [BDYW11; DSW14].

Define functions $\lambda_2(r,k) := k + 1 + 16 \cdot 4^{k-1}$, $\lambda_1(r,k) := 7 \cdot r \cdot 3^{\lambda_2(r,k)}$ and $B(k) := 3k + 15$. For the rest of this section, we set the parameter B in the definition of $\mathbb{L}_V(P)$ to $B(k)$. Note that while this parameter depends on k, it is independent of $|\mathcal{F}|$.

Constructing core algebras

We begin by showing that, to put all forms close to a wide algebra, it is enough to construct a small wide algebra which contains a constant fraction of the quadratics. More precisely, the next lemma allows us to increase the fraction of forms close to a given vector space without increasing the size of the vector space too much, so long as we start with a wide vector space which contains a constant fraction of the quadratics.[2]

Before we state and prove the lemma, the following notation will be very useful in this subsection: if $\gamma \in \mathbb{N}$, $\mathcal{G}$ is a set of forms and W is a graded vector space, we let

$$\mathcal{G}(\gamma, W) := \{P \in \mathcal{G} \mid P \text{ is } \gamma\text{-close to } W\}.$$

Lemma 3.5.37 (Increasing algebra intersection)**.** *Let $0 < \delta \leq 1$, $r, \gamma, k \in \mathbb{N}$ be such that $r > 2\gamma \geq k + 5$ and W be a r-wide vector space. If $|\mathcal{F}(\gamma, W)| \geq \delta m$ then there is a r-wide vector space Y with $\dim Y_1 \leq 3^k \cdot (\dim W + r)$, $\dim Y_2 \leq \dim W_2 + k$ such that either $|\mathcal{F}(\gamma, Y)| \geq 3\delta m/2$, or $\mathcal{F} = \mathcal{F}(3\gamma, Y)$.*

Proof. Note that $\mathcal{F}_1 \subseteq \mathcal{F}(\gamma, W)$. Let $\mathcal{H} := \mathcal{F}_2 \setminus \mathcal{F}(3\gamma, W)$. In other words, $\mathcal{H}$ is the set of forms that are 3γ-far from W. Let $H_1, \ldots, H_t \in \mathcal{H}$ be the longest sequence of linearly independent forms such that

1. $s(H_1, \ldots, H_t) \geq k + 5$, and
2. No nonzero form in $\mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$ is 2γ-close to W.

Suppose $t < k$. Let Y be the r-wide vector space obtained by applying Proposition 3.5.6 to $W + \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$. Since $H_1, \ldots, H_t$ is the longest linearly independent sequence that satisfies the above conditions, for every other $H \in \mathcal{H}$, it must be that

- either $H \in \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$, or
- $s(H, H_1, \ldots, H_t) \leq k + 4$, or

[2] This is similar in spirit to [BDYW11, Proposition 7.11] and [GOS22, Lemma 5.15].

- there exists $R \in \mathrm{span}_{\mathbb{K}}\{H, H_1, \ldots, H_t\} \setminus \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$ such that R is 2γ-close to W.

In each of these three cases, it follows that $H \in \mathcal{F}(3\gamma, Y)$. Therefore in this case, Y is the required vector space.

We are now in the case where $t \geq k$. Consider the k elements $H_1, \ldots, H_k$. Note that $H_1, \ldots, H_k$ are linearly independent, and also satisfy $s(H_1, \ldots, H_k) \geq k+5$. Therefore, $H_1, \ldots, H_k$ is a k-strong sequence. By Remark 3.5.31, the ideal $(H_1, \ldots, H_k)$ is prime and k-relevant, and $S/(H_1, \ldots, H_k)$ is a UFD. Let Y be the r-wide vector space obtained by applying Proposition 3.5.6 to $W + \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_k\}$, so $\dim Y_1 \leq 3^k \cdot (\dim W + r), \dim Y_2 \leq \dim W_2 + k$.

Now for each $G_i \in \mathcal{F}(\gamma, W)$ we have $G_i \notin (H_1, \ldots, H_k)$. In the graded UFD $S/(H_1, \ldots, H_k)$, the image of G_i must be irreducible: if not then $G_i = ab + \sum \alpha_j H_j$ in S, with $a, b \in S_1$, contradicting the fact that $\mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_k\}$ does not contain forms 2γ-close to W. The ideal $(H_1, \ldots, H_k, G_i)$ is therefore prime, and we have $R_i \in (H_1, \ldots, H_k, G_i) \setminus (H_1, \ldots, H_k)$ since $\mathcal{F}$ is a $\mathrm{SG}^*_k(2)$ configuration. We have $R_i \in \mathcal{F}(\gamma, Y)$.

If $R_i \in \mathcal{F}_1$ then we must have $R_i \in (G_i)$, contradicting the pairwise independence of elements of $\mathcal{F}_2$, therefore $R_i \in \mathcal{F}_2$. After scaling we have either $R_i - G_i \in \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_k\}$ (if $G_i \in \mathcal{F}_2$) or $R_i - aG_i \in \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_k\}$ (if $G_i \in \mathcal{F}_1$). Therefore $R_i \notin \mathcal{F}(\gamma, W)$ since otherwise $\mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_k\}$ contains a form 2γ-close to W. If G_j is another form such that $R_i = R_j$, then $R_i - G_j$ or $R_i - bG_j$ is in $\mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$, and it must be that $G_i, G_j \in \mathcal{F}_1$ and $aG_i = bG_j$ so $G_j \in (a), G_i \in (b)$. This shows that $|\{R_i\}_i| \geq \delta m/2$. Since $\mathcal{F}(\gamma, W) \cup \{R_i\}_i \subseteq \mathcal{F}(\gamma, Y)$, we are done. □

We are now ready to prove the main lemma of this subsection.

Lemma 3.5.38 (Constructing core algebras)**.** *Suppose $\mathcal{F}$ is a $\mathrm{SG}^*_k(2)$ configuration. For any r there exists a r-wide vector space W with $\dim W_1 \leq 2 \cdot 3^{k^2} \cdot \lambda_1(r, k)$ and $\dim W_2 \leq 4k^2 + \lambda_2(r, k)$ such that $\mathcal{F} = \mathcal{F}(B(k), W)$.*

Proof of Lemma 3.5.38. We build a sequence of vector spaces $W^{(i)}$ such that one of the following two statements hold: either $\mathcal{F} = \mathcal{F}(B(k), W^{(i)})$ or $\left|\mathcal{F}(k+5, W^{(i)})\right| \geq (3/2)^i \cdot m/4^k$.

Set $W^{(0)}$ to be the r-wide vector space obtained by Corollary 3.5.36. By Corollary 3.5.36, at least $m_2/4^k$ forms in $\mathcal{F}_2$ are $k+5$ close to $W^{(0)}$. Further, every form in $\mathcal{F}_1$ is 1-close to $W^{(0)}$. Since $m_1 + m_2/4^k \geq m/4^k$, we have $\left|\mathcal{F}(k+5, W^{(0)}\right| \geq m/4^k$. Therefore, $W^{(0)}$ satisfies the above property. We have $\dim W^{(0)}_i \leq \lambda_i(r, k)$.

Given $W^{(i)}$, if $\mathcal{F} = \mathcal{F}(B(k), W^{(i)})$ then terminate. If not, then apply Lemma 3.5.37 to $W^{(i)}$ with $\gamma = k+5$ and $\delta = (3/2)^i \cdot 1/4^k$ to obtain $W^{(i+1)}$. By Lemma 3.5.37, either $\mathcal{F} = \mathcal{F}(B(k), W^{(i+1)} = \mathcal{F}$ or $\left|\mathcal{F}(k+5, W^{(i+1)})\right| \geq (3/2)^{i+1} \cdot m/4^k$. Therefore $W^{(i+1)}$ also has the required property.

The above process must terminate when $(3/2)^i \cdot 1/4^k \geq 1$, which holds when $i > 4k$. Further, by induction we have $\dim W^{(4k)}_1 \leq 3^{k^2}\lambda_1(r, k) + 3^{k^2} \cdot 2^k \cdot r \leq 2 \cdot 3^{k^2} \cdot \lambda_1(r, k)$ and $\dim W_2 \leq 4k^2 + \lambda_2(r, k)$. Therefore $W^{(4k)}$ is the required space. □

Finding small ideal containing the quadratic forms

In this section we show that all quadratics in any $\mathrm{SG}^*_k(2)$ must be contained in an ideal generated by a small number of forms. The main idea is that given any wide vector space, there exist short maximal integral sequences with respect to the vector space. Recall that the we set the parameter B in the definition of relative linear spaces to $B(k) := 3k + 15$.

Lemma 3.5.39. *Suppose $r \geq 4(k+2)B(k)+1$. Suppose $\mathcal{F}$ is a $\mathrm{SG}_k^*(2)$ configuration, and suppose W is a r-wide vector space such that every $F \in \mathcal{F}$ is $B(k)$-close to W. Then there exists a maximal integral sequence with respect to inclusion of length at most k with respect to W.*

Proof. For each $F \in \mathcal{F}$ let $\overline{F_W^L}$ be the image of F_W^L in $S/(W_1)$. Define potential function Φ on integral sequences as

$$\Phi(G_1, \dots, G_c) := \sum_{i=1}^{c} \dim \mathrm{Lin}\Big(\overline{(G_i)_W^L}\Big).$$

If $\mathcal{F} \subset (W)$ then there are no integral sequences with respect to W, and the statement holds vacuously, therefore we can assume that $\mathcal{F} \setminus (W) \neq \emptyset$. Combined with the fact that W is r-wide, and that every form in $\mathcal{F}$ is $B(k)$-close to W, there exists nonempty integral sequences with respect to W. Among all integral sequences of length at most $k+1$, pick $F_1, \dots, F_c$ such that the above potential function is maximised. If $c \leq k$, then $F_1, \dots, F_c$ is maximal: if not, and if $F_1, \dots, F_{c+1}$ is an integral sequence that extends $F_1, \dots, F_c$ then $\Phi(F_1, \dots, F_{c+1}) > \Phi(F_1, \dots, F_c)$, contradicting maximality.

We are left with the case where $c = k+1$. We will find an integral sequence of length at most k with the same potential function value, and therefore the new integral sequence will be maximal. The sequence $F_1, \dots, F_k$ is an integral sequence, therefore by Lemma 3.5.29 we have that $(F_1, \dots, F_k)$ is a radical ideal. Further, by Lemma 3.5.29 we have that $F_1, \dots, F_k$ is a k-relevant set. Similarly, $F_1, \dots, F_{k+1}$ is a $k+1$-relevant set, therefore $F_{k+1} \notin (F_1, \dots, F_k)$.

Since $\mathcal{F}$ is a $\mathrm{SG}_2^*(2)$, we have $R \in (F_1, \dots, F_{k+1}) \setminus (F_1, \dots, F_k)$, that is, $R = \sum_{j=1}^{k+1} \alpha_j F_j$ with $\alpha_{k+1} \neq 0$. Without loss of generality, suppose $\alpha_j = 0$ for $j = 1, \dots, b$ and $\alpha_j \neq 0$ for $j = b+1, \dots, k+1$. Since the polynomials in $\mathcal{F}$ are pairwise linearly independent we have $b < k$.

Now $R = R_W^H + R_W^L = \sum_{j>b} \alpha_i F_i = \sum_{j>b} \alpha_i ((F_i)_W^H + (F_i)_W^L)$. Since the space W is r-wide, we have $R_W^H = \sum_{j>b} \alpha_i (F_i)_W^H$ and $R_W^L = \sum_{j>b} \alpha_i (F_i)_W^L$. By Corollary 3.5.27, after a change of basis we can assume that there are disjoint sets of variables $Y, Y_1, \dots, Y_k$ such that W_1 is spanned by Y and $(F_i)_W^L \in \mathbb{K}[Y, Y_i]$. We have $\mathrm{Lin}\big(R_W^L\big) \subseteq \mathbb{K}[Y, Y_{b+1}, \dots, Y_{k+1}]$, whence $F_1, \dots, F_b, R$ is an integral sequence by Proposition 3.5.25. Further $\overline{R_W^L} = \sum_{j>b} \alpha_j \overline{(F_j)_W^L}$, and since $\overline{(F_i)_W^L} \in \mathbb{K}[Y_i]$, by Lemma 3.5.4 we can deduce that $\dim \mathrm{Lin}\Big(\overline{R_W^L}\Big) = \sum_{j>b} \dim \mathrm{Lin}\Big(\overline{(F_j)_W^L}\Big)$. Therefore $F_1, \dots, F_b, R$ is an integral sequence of length at most k with $\Phi(F_1, \dots, F_b, R) = \Phi(F_1, \dots, F_{k+1})$. This proves that $F_1, \dots, F_b, R$ is a maximal integral sequence. □

Lemma 3.5.40. *Suppose $\mathcal{F}$ is a $\mathrm{SG}_k^*(2)$ configuration. Suppose $r \geq 8(k+2)B(k)^2+1$. There exists a $\big(r - 4kB(k)^2\big)$-wide vector space W with $\dim W_1 \leq 3 \cdot 3^{k^2} \cdot \lambda_1(r,k)$ and $\dim W_2 \leq 4k^2 + \lambda_2(r,k)$ such that $\mathcal{F}_2 \subset (W)$.*

Proof. For any r-wide vector space U such that every polynomial $F \in \mathcal{F}$ is $B(k)$ close to U, define potential function Ψ as

$$\Psi(U) = \max_{F \in \mathcal{F}_2 \setminus (U)} \dim \overline{\mathbb{L}}_U(F).$$

If $\Psi(U) = 0$ for some such U, then $\mathcal{F}_2 \subset (U)$.

We now construct W iteratively. Let $W^{(0)}$ be the r-wide vector space whose existence is guaranteed by Lemma 3.5.38. Since every $F \in \mathcal{F}$ is $B(k)$-close to $W^{(0)}$ we have $\Psi(W^{(0)}) \leq 2B(k)$. The vector space $W^{(0)}$ is r-wide, and $r \geq 4(k+2)B(k)+1$, therefore by Lemma 3.5.39 we can find a maximal integral sequence $F_1, \dots, F_c$ with respect to $W^{(0)}$ with $c \leq k$. Set $W^{(1)} = \sum_{i=1}^{c} \mathbb{L}_{W^{(0)}}(F_i) + W_2^{(0)}$. That $F_1, \dots, F_c$ is maximal implies the following: for every $G \in \mathcal{F}_2 \setminus \big(W^{(0)}\big)$ we have $\dim \overline{\mathbb{L}}_{W^{(0)}}(G) > \dim \overline{\mathbb{L}}_{W^{(1)}}(G)$. Therefore $\Psi(W^{(1)}) < \Psi(W^{(0)})$. By Lemma 3.5.16, the

vector space $W^{(1)}$ is $r - 2kB(k)$-wide since $W_1^{(1)}$ is obtained by adding at most $2kB(k)$ linear forms to $W_1^{(0)}$. In general, given $W^{(i)}$ we use Lemma 3.5.39 to find a maximal integral sequence $F_{i1}, \dots, F_{ic}$, and set $W^{(i+1)} := \sum_{j=1}^{c} \mathbb{L}_{W^{(i)}}(F_{ij}) + W_2^{(i)}$. Maximality of the sequence implies $\Psi(W^{(i+1)}) < \Psi(W^{(i)})$. By the bound on r, at every step $W^{(i)}$ is at least $4(k+2)B(k)+1$-wide. After at most $2B(k)$ steps we find a t such that $\Psi(W^{(t)}) = 0$.

By Lemma 3.5.38 we have $\dim W_1^{(0)} \leq 2 \cdot 3^{k^2} \lambda_1(r,k)$. Since $W^{(i+1)}$ is obtained by adding $2B(k)k$ linear forms to $W^{(i)}$ we get $\dim W_1^{(t)} \leq 2 \cdot 3^{k^2} \lambda_1(r,k) + 4B(k)^2 k \leq 3 \cdot 3^{k^2} \lambda_1(r,k)$. Further we have $\dim W_2^{(i)} = \dim W_2^{(i-1)}$ for all i, therefore $\dim W_2^{(t)} = \dim W_0^{(t)} \leq 4k^2 + \lambda_2(r,k)$. This completes the proof. □

Basic configuration

In this section we prove the dimension bound for the special case where all the quadratics are in the ideal generated by an r-wide algebra.

Lemma 3.5.41. *Suppose $\mathcal{F}$ is a $\mathrm{SG}_k^*(2)$ configuration. Suppose there is an r-wide linear subspace W with $r \geq k+5$ such that $\mathcal{F}_2 \subset (W)$. Then there is linear subspace W_1' with $\dim(W_1') = \left(C'^k\right) \cdot \dim W_1$, such that $\mathcal{F} \subseteq W_2 + \mathbb{K}[W_1']$.*

Proof. Let $\varphi := \varphi_{\boldsymbol{\alpha}, W_1}$ be a projection mapping as defined in Definition 3.5.17. By Remark 3.5.23, the space $\varphi(W)$ is a $r-1$-wide vector space. Let $\Delta := \dim W_1$. As $\mathcal{F}_2 \subseteq (W)$, every $F \in \mathcal{F}_2$ satisfies $\varphi(F) = \alpha \varphi(F_W^H) + z \cdot \ell$ for some linear form $\ell \in S[z]_1$.

Let $\mathcal{L}$ be the union of all the linear forms that occur in the above way, and all the linear forms in $\mathcal{F}$. Formally, $\mathcal{L} := \left\{ \ell \mid \varphi(F) = \alpha \varphi(F_W^H) + z \cdot \ell, F \in \mathcal{F}_2 \right\} \cup \varphi(\mathcal{F}_1)$. Let $\mathcal{L}/(z)$ denote the image of $\mathcal{L}$ in the vector space $(S[z]/(z))_1$, that is, the linear forms modulo z. We show that $\mathcal{L}/(z)$ is a $\mathrm{SG}_k(1)$ configuration.

Let $\overline{\ell_1}, \dots, \overline{\ell_k} \in \mathcal{L}/(z)$ be independent. Let $\overline{\ell_{k+1}} \in \mathcal{L}/(z)$. We need to show that one of the following cases holds:

1. $\bar{\ell}_{k+1} \in \mathrm{span}_{\mathbb{K}} \left\{ \overline{\ell_1}, \dots, \overline{\ell_k} \right\}$.
2. there is $\bar{g} \in \mathrm{span}_{\mathbb{K}} \left\{ \overline{\ell_1}, \dots, \overline{\ell_{k+1}} \right\} \setminus \left\{ \overline{\ell_1}, \dots, \overline{\ell_k} \right\}$ with $\bar{g} \in \mathcal{L}/(z)$.

Consider the corresponding $F_1, \dots, F_{k+1} \in \mathcal{F}$ such that $\varphi(F_i) = \alpha_i \pi(F_i{}_W^H) + z \cdot \ell_i$, with $\ell_i/(z) = \bar{\ell}_i$, or, if $F_i \in \mathcal{F}_1$ then $F_i = \ell_i$.

The first step is to show that $F_1, \dots, F_k$ form a k-relevant set. Without loss of generality, assume that $F_1 \in \mathrm{rad}(F_2, \dots, F_k)$. We have $\varphi(F_1) \in \mathrm{rad}(\varphi(F_2), \dots, \varphi(F_k))$, and by Lemma 3.5.9 we have $z\ell_1 \in (\ell_2, \dots, \ell_k)$. Since the ideal $\ell_2, \dots, \ell_k$ is prime, and since $\overline{\ell_2}, \dots, \overline{\ell_k}$ are independent, we get $\ell_1 \in \mathrm{span}_{\mathbb{K}}\{\ell_2, \dots, \ell_k\}$ contradicting choice of $\ell_1, \dots, \ell_k$. Therefore $F_1, \dots, F_k$ is k-relevant.

By the same argument, if $F_{k+1} \in \mathrm{rad}(F_1, \dots, F_k)$ then $\overline{\ell_{k+1}} \in \mathrm{span}_{\mathbb{K}} \left\{ \overline{\ell_1}, \dots, \overline{\ell_k} \right\}$. We are left with the case when $F_{k+1} \notin \mathrm{rad}(F_1, \dots, F_k)$. Since $\mathcal{F}$ is a $\mathrm{SG}_2^*(2)$ configuration, there exists $R \in \mathrm{rad}(F_1, \dots, F_{k+1}) \setminus \mathrm{rad}(F_1, \dots, F_k)$. Let $\overline{g}$ be such that $\varphi(R) = \alpha_i \varphi(R_W^H) + z \cdot g$ if $R \in \mathcal{F}_2$, and $R = g$ otherwise. We have $\varphi(R) \in \mathrm{rad}(\varphi(F_1), \dots, \varphi(F_{k+1}))$. By Lemma 3.5.9, we have $z\overline{g} \in \mathrm{span}_{\mathbb{K}} \left\{ \overline{\ell_1}, \dots, \overline{\ell_{k+1}} \right\}$ which implies $\overline{g} \in \mathrm{span}_{\mathbb{K}} \left\{ \overline{\ell_1}, \dots, \overline{\ell_{k+1}} \right\}$. Finally, by Lemma 3.5.10, we have that $\overline{g} \notin \left(\overline{\ell_i}\right)$ for any i. This completes the proof that $\mathcal{L}/(z)$ is a $SG_k(1)$ configuration.

By Theorem 3.1.8 we have

$$\dim(\mathbb{L}_W(\varphi(\mathcal{F})) = \dim(\mathcal{L}/(z)) + 1 \leq C'^k,$$

for some universal constant C'. Applying Proposition 3.5.21 it follows that $\dim(\mathbb{L}_W(\mathcal{F})) \leq C'^k \cdot \Delta$. In particular, it follows that there is a linear space of linear forms W_1', with $\dim(W_1') \leq C'^k \cdot \Delta$, satisfying $\mathcal{F} \subseteq W_2 + \mathbb{K}[W_1']$, completing the proof. □

Proof of main theorem

Theorem 3.5.42. *Let $\mathcal{F}$ be a $\mathrm{SG}_2^*(2)$ configuration. Then $\dim(\mathcal{F}) \leq 3^{c \cdot 4^k}$ where c is a universal constant independent of k.*

Proof. Let $r := 8(k+2)B(k)^2 + k + 6$. By Lemma 3.5.40, there exists a $k+5$-wide vector space W with $\dim W_1 \leq 3 \cdot 3^{k^2} \cdot \lambda_1(r,k)$ and $\dim W_2 \leq 4k^2 + \lambda_2(r,k)$ such that $\mathcal{F}_2 \subseteq (W)$. Applying Lemma 3.5.41 with this W, we obtain a vector space $W_1' \subseteq S_1$ with $\dim W_1' \leq 3 \cdot 3^{k^2} \cdot \lambda_1(r,k) \cdot C'^k$ such that $\mathcal{F} \subseteq W_2 + \mathbb{K}[W_1']$. If $Y \subseteq S_2$ is the space spanned by pairwise products of forms in W_1', then $\mathcal{F} \subseteq W_2 + Y$ and $\dim Y \leq 9 \cdot 9^{k^2} \cdot \lambda_1^2(r,k) \cdot C'^{2k}$. Substituting for λ_1, λ_2 gives us the required result. □

Remark 3.5.43. Suppose the set $\mathcal{F}$ does not have any k-relevant sequences. In this case, $\mathcal{F}$ is vacuously an $\mathrm{SG}_k^*(2)$ configuration. There are no k-strong sequences in $\mathcal{F}$ of length k, since any such sequence is a k-relevant set. Therefore every form in such a configuration is $k+5$-close to a r-wide vector space W of dimension $7 \cdot r \cdot 3^k$ by Lemma 3.5.32. Further, such a configuration has no integral sequences of length $k+1$. Therefore, by the arguments in Lemma 3.5.39 and Lemma 3.5.40, by adding $4kB(k)^2$ linear forms to W, we get a wide vector space Y such that $\mathcal{F} \subseteq Y$. If we project to Y_1 and pick out the linear forms corresponding to each element of $\mathcal{F}$ as in Lemma 3.5.41, then there are no set of $k+1$ linearly independent forms by Lemma 3.5.9. Therefore, we can deduce by the properties of the projection map that $\dim \mathrm{span}_{\mathbb{K}}\{\mathcal{F}\} = 2^{O(k)}$ in this case. ◊

3.6 Going beyond quadratics: Strong algebras and graded quotients

In the previous section, we saw an instantiation of the ideas discussed in Section 3.4 to prove rank bounds. In this section, we will discuss the barriers to extending the method in the previous section to forms of higher degree, and we will discuss the tools introduced by Oliveira and Sengupta to overcome these barriers.

The proofs of the dimension bounds for quadratic Sylvester–Gallai configurations ([Shp20; PS20; PS21; PS22; GOS22; GOPS23]) all essentially reduced to the case of linear Sylvester–Gallai configurations. Similarly, the proof of the dimension bound for 3-radical Sylvester–Gallai configurations ([OS22]) reduced the problem to a generalisation of 2-radical Sylvester–Gallai configurations. It seems natural to try and prove bounds for d-radical Sylvester–Gallai configurations by reducing them to $d-1$-Sylvester–Gallai configurations, and Oliveira and Sengupta developed a framework to do exactly this in [OS24].

The first tool in this framework is the notion of strong algebras.

3.6.1 Strong and lifted strong algebras

Strong vector spaces and algebras generalises the notion of wide AH spaces Definition 3.4.3.

Definition 3.6.1 (Strong Ananyan-Hochster vector spaces)**.** Let $R = \oplus_{d \geq 0} R_d$ be a finitely generated graded $\mathbb{K}$-algebra, generated by R_1. For any function $B = (B_1, \cdots, B_d) : \mathbb{N}^d \to \mathbb{N}^d$, we say that a non-zero graded vector subspace $V = \oplus_{i=1}^d V_i \subset R$, with dimension sequence δ, is a B-strong AH vector space if V_i is $B_i(\delta)$-strong for all i, i.e. $s(V_i) \geq B_i(\delta)$. The subalgebra $\mathbb{K}[V] \subset R$ generated by a B-strong AH vector space V is called a B-strong AH algebra. ◊

A (w,t)-wide AH space is just a strong algebra in the above sense with function $B_2(\delta_1,\delta_2) = t(w+\delta_1+\delta_2)$. The function B_1 can be arbitrary since we assume linear forms have infinite strength.

The Ananyan–Hochster principle implies that as long as B is large enough, any basis of a B-strong AH vector space behaves like variable. In the following corollary, the function $A(\eta,d):\mathbb{N}^2\to\mathbb{N}$ is the function defined in [AH20b, Theorem A]. A proof can be found in [OS24, Corollary 5.9].

Corollary 3.6.2. *Let $V=\oplus_{i=1}^d V_i\subset S$ be a B-strong AH vector space for some $B:\mathbb{N}^d\to\mathbb{N}^d$. Suppose $B_i(\delta)\geq A(\eta,i)+3(\sum_i\delta_i-1)$ for some $\eta\in\mathbb{N}$. Then any sequence of $\mathbb{K}$-linearly independent forms in V is an $\mathcal{R}_\eta$-sequence. If $\eta\geq 3$, then $S/(V)$ is a Cohen-Macaulay, unique factorization domain.*

As in the case of wide algebras, given any graded vector space U and any function B, there exists an iterative process that starts with U and gives rise to a B-strong AH vector space. For any $\mu\in\mathbb{N}^d$, we define the translation function $t_\mu:\mathbb{N}^d\to\mathbb{N}^d$ as $t_\mu=(t_{\mu,1},\cdots,t_{\mu,d})$ where the i-th component is defined by $t_{\mu,i}(\delta)=\delta_i+\mu_i$. In other words, for all $i\in[d]$ we add μ_i to the i-th component of δ. For any $n\in\mathbb{N}$, we let $t_n:=t_{(n,\cdots,n)}$.

The following lemma is proved in [OS24, Lemma 5.15].

Lemma 3.6.3 (Strengthening of Algebras)**.** *For any $d\in\mathbb{N}$ and a function $B:\mathbb{N}^d\to N^d$, there exist functions $C_B:\mathbb{N}^d\to\mathbb{N}^d$ and $h_B:\mathbb{N}^d\to\mathbb{N}^d$, depending on B, such that the following holds:*

Given a graded vector space $U=\oplus_{i=1}^d U_i\subset S$ with dimension sequence $\delta\in\mathbb{N}^d$, there exists a B-strong AH vector space $V=\oplus_{i=1}^d V_i$ such that

1. *$\mathbb{K}[U]\subset\mathbb{K}[V]$,*
2. *for all $i\in[d]$, we have $\dim(V_i)\leq C_{B,i}(\delta)$, where $C_{B,i}$ denotes the i-th component of $C_B=(C_{B,1},\cdots,C_{B,d}):\mathbb{N}^d\to\mathbb{N}^d$.*

Furthermore, suppose $H=\oplus_{i=1}^d H_i\subset U$ is a graded subspace such that $s(H_i)\geq h_{B,i}(\delta)$ for all $i\in[d]$. Then there exists a B-strong AH vector space V satisfying (1) and (2) above such that $H\subset V$.

The following corollary is from [OS24, Corollary 5.16].

Corollary 3.6.4 (Robustness of strong algebras)**.** *Let $B,G:\mathbb{N}^d\to\mathbb{N}^d$ and $\mu\in\mathbb{N}^d$. Suppose that $B_i(\delta)\geq h_{G,i}(\delta+\mu)$ for all $\delta\in\mathbb{N}^d$ and $i\in[d]$, where $h_G:\mathbb{N}^d\to\mathbb{N}^d$ is the function defined in Lemma 3.6.3. Let $U\subset S$ be a B-strong AH vector space and $W\subset S$ is a graded vector space with dimension sequences δ and μ respectively. Then there exists a G-strong AH vector space V such that*

1. *$\mathbb{K}[U+W]\subset\mathbb{K}[V]$,*
2. *$U\subset V$,*
3. *for all $i\in[d]$, $\dim(V_i)\leq C_{G,i}(\delta+\mu)$, where $C_G:\mathbb{N}^d\to\mathbb{N}^d$ is the function defined in Lemma 3.6.3.*

We recall the discussion right after Definition 3.4.3. In a typical application of the above corollary, we will start with a vector space U that is already strong, and add a small vector space to W. The vector space V is obtained by applying some rounds of cleaning up, that is, looking at elements that have low strength, decomposing them and adding new forms to the vector space. This will result in a strong space. If we ensure that U was much stronger than the target strength we want from V, then the iterative process will never decompose any element of U, which will guarantee that $U\subset V$.

Therefore, in the function h essentially controls how much the strength parameter degrades when we add a small vector space to U.

As before, we can think of strong vector spaces as a proxy for essential variables. If P is a degree d form, and V is the B-strong AH space obtained by applying Corollary 3.6.4 with $U = 0$ and $W = \text{span}_{\mathbb{K}}\{P\}$, then a basis of the resulting space acts as essential variables for V. If Y is any fixed strong AH space, then a basis for the space obtained by applying Corollary 3.6.4 with $U = Y$ and $W = \text{span}_{\mathbb{K}}\{P\}$ acts as relative essential variables of P over Y. In this setting, the fact that U is unchanged is useful: it means that our set of variables with respect to which we were considering essential variables remains unchanged.

The Ananyan–Hochster principle holds in polynomial rings, but not in quotients of polynomial rings. The following example is from [OS24]. Let $F = x_1^2 + \cdots + x_5^2$, and $R = S/(F)$. The ring R is factorial since F has high strength. In this ring, the linear forms x_1, x_2 generate a strong vector space since they are linearly independent, but the quotient by these two linear forms is not a UFD. This happens because in the ring S, the vector space generated by F, x_1, x_2 is not strong enough for the Ananyan–Hochster principle to apply. To remedy this, [OS24] introduced the notion of lifted strength.

Definition 3.6.5 (Lifted strength)**.** Let $U \subset S$ be a graded vector space and $R = S/(U)$. Let $F \in R_d$ be a non-zero form. We define the lifted strength of F with respect to U as

$$\widetilde{s}(U, F) := \min\{s\Big(U_d + \text{span}_{\mathbb{K}}\Big\{\widetilde{F}\Big\}\Big)\}$$

where $\widetilde{F}$ varies over all forms in S_d such that the image of $\widetilde{F}$ in R is F. Given a set of forms $F_1, \cdots, F_m \in R_d$, we define

$$\widetilde{s}(U, F_1, \cdots, F_m) = \min\{s\Big(U_d + \text{span}_{\mathbb{K}}\Big\{\widetilde{F}_1, \cdots, \widetilde{F}_m\Big\}\Big)\},$$

where $\widetilde{F}_i$ varies over all forms in S_d such that the image of $\widetilde{F}_i$ in R is F_i. Given a non-zero vector space $V \subset R_d$, we define

$$\widetilde{s}(U, V) = \min\{\widetilde{s}(U_d, F_1, \cdots, F_m)\},$$

where $F_1, \cdots, F_m$ vary over all possible bases of V. We say that $V \subset R_d$ is k-lifted strong with respect to U if $\widetilde{s}(U, V) \geq k$. For simplicity, we omit U from the notation and write $\widetilde{s}V$ when U is clear from the context.

Let $V = \oplus_{i=1}^{d} V_i \subset R$ be a graded vector space. For any function, $B : \mathbb{N}^d \to \mathbb{N}^d$ we will say that V is B-lifted strong with respect to U, if V_i is $B_i(\dim(U_i) + \dim(V_i))$-lifted strong, i.e. $\widetilde{s}(U, V_i) \geq B_i(\dim(U_i) + \dim(V_i))$ for all $i \in [d]$. In other words, V is B-lifted strong with respect to U, if the vector space $U + \text{span}_{\mathbb{K}}\Big\{\widetilde{F}_1, \cdots, \widetilde{F}_m\Big\}$ is B-strong in S, for any homogeneous basis $F_1, \cdots, F_m \in R$ of V and any set of homogeneous lifts $\widetilde{F}_1, \cdots, \widetilde{F}_m \in S$. $\Diamond$

Essentially the idea is to ensure that preimage of V under the quotient map $S \to S/U$ is a strong algebra, so the Ananyan–Hochster principle can be applied in S. In particular, Corollary 3.6.2 can be applied in S/U, if we replace the assumption that V is B-strong with the assumption that V is B-lifted strong.

Corollary 3.6.4 can be extended to the setting of lifted strong vector spaces. As discussed above, the typical application will be to start with a lifted strong vector space and extend this by adding constantly many forms. [OS24, Corollary 5.17] captures this, and we state it below.

Corollary 3.6.6. *Let $B : \mathbb{N}^d \to \mathbb{N}^d$. Let $U \subset S$ be a graded vector space with dimension sequence $\delta_U \in \mathbb{N}^d$ and let $R = S/(U)$. Let $V \subset R$ is a graded vector space with dimension sequence $\delta_V \in \mathbb{N}^d$. Suppose V is $h_{2B} \circ t_k$-lifted strong with respect to U. Let $P_1, \cdots, P_k \in R_{\leq d}$ be homogeneous elements. Then there exists a graded vector space $V' \subset R_{\leq d}$ such that:*

1. *V' is B-lifted strong with respect to U.*
2. *$P_1, \cdots, P_k \in \mathbb{K}[V']$.*
3. *$V \subset V'$.*
4. *for all $i \in [d]$, we have $\dim(V'_i) \leq C_{2B,i}(t_k(\delta_U + \delta_V)) - \delta_{U,i}$. In particular, $\dim(V') \leq C(B, \delta_U, \delta_V, k) := \sum(C_{2B,i}(t_k(\delta_U + \delta_V)) - \delta_{U,i})$.*

The following definition allows us to conveniently refer to the above spaces.

Definition 3.6.7 (Ananyan-Hochster spaces)**.** In the situation of Corollary 3.6.4, we define $\mathrm{AH}(U, W)$ to be any graded vector space V provided by Corollary 3.6.4. Similarly, in the situation of Corollary 3.6.6, we define $\mathrm{AH}_R(V, P_1, \cdots, P_k)$ to be any vector space V' provided by Corollary 3.6.6. ◊

Remark 3.6.8. Note that, for given $V, P_1, \cdots, P_k$, the vector space $\mathrm{AH}_R(V, P_1, \cdots, P_k)$ is not necessarily unique. As stated in our definition, we use the notation $\mathrm{AH}_R(V, P_1, \cdots, P_k)$ to denote any vector space satisfies the properties in Corollary 3.6.6, whose existence is guaranteed. We will only use these properties of these spaces and in all our arguments we work with any fixed choice of such a vector space $\mathrm{AH}_R(V, P_1, \cdots, P_k)$. ◊

Similar to the quadratic case, our proofs will have some iterative steps, where we start with some strong vector space V, add some forms to V and the apply the above lemmas to obtain a strong space that contains V and whose algebra contains the forms we added. Each such step will increase the size of the vector space, and degrade the strength of the vector space. The following definition gives us a convenient way to keep track of these parameters.

Definition 3.6.9 ((k,t)-process)**.** Let $U \subseteq S$ be a graded finitely generated vector space such that $R := S/(U)$ is a UFD, and $k, t \geq 1$ be integers. Let $V \subseteq R$ be a vector space and let $\mathcal{F} \subseteq R$. A (k,t) process on V is defined to be a process that starts with the vector space V, and performs at most t rounds. In each round, at most k forms $F_1, \ldots, F_k$ of degree at most d are picked from the set $\mathcal{F}$, and V is updated to $AH_R(V, F_1, \ldots, F_k)$. In each iteration, the picked forms are allowed to be chosen based on the new V. Note that we are dropping the dependence on d from the notation for convenience, as d will always be clear from context. ◊

We introduce some auxiliary functions that will provide uniform bounds for (k,t)-processes (see Lemma 3.6.10).

Strength bound function. Given a function $B : \mathbb{N}^d \to \mathbb{N}^d$, and positive integers k, t, we define a function $H(B, k, t) : \mathbb{N}^d \to \mathbb{N}^d$ recursively as follows. Let $H(B, k, 0) := B$, and

$$H(B, k, t) := h_{2H(B,k,t-1)} \circ t_{k+1},$$

where h is as defined in Lemma 3.6.3. This function $H(B, k, t)$ captures the strength needed for a vector space V, so that after applying a (k,t)-process to V, the resulting vector space V' is still B-lifted strong and this (k,t)-process preserves V, i.e. $V \subseteq V'$.

Dimension bound function. Fix $\delta \in \mathbb{N}^d$, and B, k, t as above. We define a function $D(B,k,t,\delta) : \mathbb{N}^d \to \mathbb{N}^d$ recursively as follows. Let $D(B,k,0,\delta)$ be the identity function. To define $D(B,k,t,\delta)(\mu)$, assuming $D(B,k,t-1,\delta)$ is already defined, we first define

$$\eta := \max_{\substack{x \in \mathbb{N}^d \\ \|x\|_1 = \|\mu\|_1}} \|D(B,k,t-1,\delta)(x)\|_1,$$

where $\|x\|_1 = x_1 + \cdots + x_d$ for $x = (x_1, \cdots, x_d) \in \mathbb{N}^d$. We then define the i^{th} coordinate function $D_i(B,k,t,\delta)(\mu)$ as

$$D_i(B,k,t,\delta)(\mu) = \max_{x, \|x\|_1 = \eta} C_{2H(B,k,t),i}(t_{k+1}(x+\delta)),$$

where again C is the function defined in Lemma 3.6.3. This function $D(B,k,t,\delta)$ captures the dimension upper bound for the result of a (k,t)-process applied on a vector space V.

We note these bounds and basic properties of (k,t) processes below.

Lemma 3.6.10. *Let $U \subseteq S$ be a graded finitely generated vector space such that $R := S/(U)$ is a UFD, and $k, t, k', t' \geq 1$ be integers.*

1. *The composition of a (k,t) process and a (k',t') process is a $(\max(k,k'), r+r')$ process.*
2. *If V is $H(B,k,t)$-lifted strong, and if V' is the result of a (k,t) process on V, then $V \subset V'$ and V' is B-lifted strong. Moreover, each intermediate vector space that appears as part of the process is $h_{2B} \circ t_1$-lifted strong.*
3. *If V is $H(B,k,t)$-lifted strong, and if V' is the result of a (k,t) process on V, then the dimension of V' is bounded by $D(B,k,t,\delta_U)(\delta_V)$, where δ_U, δ_V are the dimension vectors of V, U respectively.*

Proof. The first property follows by definition. Properties (2) and (3) follow by induction. □

Before we end this subsection on strong algebras, we compare the general case with the quadratic case. In the quadratic case, given a vector space V, the relative essential variables of a form Q (which with our current definitions is a set of forms extending basis of V to a basis for the space $\mathrm{AH}(V,Q)$) is either Q itself, or a collection of linear forms. If Q is close to V then the relative essential variables are linear forms, and if W is any strong vector space such that $\mathbb{K}[V] \subset \mathbb{K}[W]$, then Q will also be close to W. Further, we can easily compare the relative essential variables with respect to V and W, since these are just two vector space, one containing the other. The argument in the quadratic case was therefore to first find a vector space such that every form is close to it, and then use this number of relative essential variables as a potential function to measure progress towards the condition that $\mathcal{F} \subset (V)$ as we updated V by adding linear forms to it.

In the high degree case, the relative essential variables are graded vector spaces of degree at most d. In general, it is unclear how the dimension of $\mathrm{AH}(V,Q)$ and $\mathrm{AH}(W,Q)$ compare even when $V \subset W$, due to the fact that these vector spaces are not unique (in the quadratic case we had some notion of uniqueness). In general, the dimension of $\mathrm{AH}(W,Q)$ can be bigger than $\mathrm{AH}(V,Q)$: for example consider the case where $\mathrm{AH}(V,Q)$ contains a quadratic form (which must have high strength), but when constructing $\mathrm{AH}(W,Q)$ this quadratic form is decomposed because W contains a quadratic form close to this one. In this case, the quadratic form will be replaced by a set of linear forms. If we order elements of $\mathbb{N}^d$ lexicographically, then we can in fact show that $\mathrm{AH}(V,Q) \geq \mathrm{AH}(W,Q)$. For example, when constructing $\mathrm{AH}(W,Q)$, we can start by making the same choices as we would in constructing $\mathrm{AH}(V,Q)$, before potentially doing more steps in the process. However, unless we

bound how much a coordinate increases when a larger coordinate decreases, we cannot use this order to define a potential function. This increase is related to the dimension of W, and the target strength, and therefore is not easy to control.

In [OS24], the authors use a different potential function to bound the number of iterations, defined as follows. Given a form F and a graded vector space V, let k be the largest integer such that F is contained in the ideal generated by $\mathbb{K}[V]_k$, the degree k piece of the algebra $\mathbb{K}[V]$. For example, if $k \geq 1$ then F is in the ideal (V), and if $k = \deg F$ (which is the highest value k can take) then $F \in \mathbb{K}[V]$. If V consisted of only linear forms, then k would be the largest power of the ideal (V) that F lies in. The proxy for the distance between V and F is then simply $\deg F - k$. Clearly this distance is between 0 and d for all forms F, and if V, W are such that $\mathbb{K}[V] \subset \mathbb{K}[W]$, then the distance between F and W measured this way is smaller than that between F and V. Therefore, it is feasible to use this to define a potential. The difficulty in using this potential lies in the fact that since it is so crude, we have to work slightly harder to ensure that a certain iterative process drops the potential.

In [GOS25c; GOS25b], we work with a modified version of the above distance, which will be easier to describe once we introduce graded quotients, which are the second tool introduced by Oliveira and Sengupta.

3.6.2 Graded Quotients

Before we begin this subsection, we note that these were originally called general quotients, but in [GOS25b] we decided to call them graded quotients instead, which is a more accurate name.

In [Shp20], Shpilka introduced the notion of a general projection, which is simply a map that sends a set of variables $x_1, \ldots, x_r$ to random multiples of a new variable z. This is essentially equivalent to specialising the variables $x_1, \ldots, x_r$ randomly and then homogenising. If a form F belongs to the ideal generated by $x_1, \ldots, x_r$, then such a map sends F to $z \cdot \ell$, where ℓ is a linear form. These maps can therefore reduce the degree of forms, and were used to reduce quadratic Sylvester–Gallai configurations to linear ones. The images of general projections lie in the ring $S[z]/(x_1 - \alpha_1 z, \ldots, x_r - \alpha_r z)$, which is isomorphic to a polynomial ring. These general projections were also used in [PS20; PS21; PS22; GOS22; GOPS23; OS22].

In general, we do not expect our forms to lie in ideals generated by a few linear forms, therefore general projections do not always reduce the degree of forms in our configuration. The main observation of Oliveira and Sengupta is the following: since a collection of forms or large strength behave like variables (The Ananyan–Hochster principle), projections can be defined with respect to not just variables, but also collections of forms of large strength, and these projections will behave like projections of variables. This ensures that such projections can be used to reduce degree, while simultaneously ensuring that the rings we encounter, while no longer polynomial rings, still behave like polynomial rings.

Definition 3.6.11 (Graded Quotients)**.** Let $U = \oplus_{i=1}^{d} U_i \subset S$ be a graded vector space of dimension sequence δ in S and $R := S/(U)$ be the quotient ring. Let $V = \oplus_{i=1}^{d} V_i \subset R$ be a graded subspace of dimension sequence μ.

Let $F_1, \ldots, F_n$ be a homogeneous basis for V and z be a new variable. For $\alpha \in \mathbb{K}^n$, let $I_\alpha \subseteq R[z]$ be the homogeneous ideal generated by the forms $\{F_1 - \alpha_1 z^{\deg(F_1)}, \ldots, F_n - \alpha_n z^{\deg(F_n)}\}$. We define the graded quotient map $\varphi_{V,\alpha}$ as the quotient homomorphism of finitely generated graded $\mathbb{K}$-algebras given by

$$\varphi_{V,\alpha} : R[z] \to R[z]/I_\alpha.$$

$\Diamond$

Remark 3.6.12. The definition above depends on the choice of the basis $F_1, \cdots, F_n$ for V. We omit this from the notation for simplicity. We will apply this graded quotient construction for sufficiently strong vector spaces V. Therefore [AH20b] shows that any basis of V_α forms an $\mathcal{R}_\eta$ sequence. Similarly, several of our statements regarding these graded quotients will apply independent of the choice of the basis. $\Diamond$

The next proposition and lemma correspond to [OS24, Proposition 6.3] and [OS24, Lemma 6.4].

Proposition 3.6.13. *Suppose $V \subset R$ is B-lifted strong with respect to U. Then $R[z]$ and $R[z]/I_\alpha$ are quotients of $S[z]$ by $\mathcal{R}_\eta$-sequences, for any choice of $\alpha \in \mathbb{K}^n$. In particular, they are Cohen-Macaulay UFDs.*

General points. We say that a property $\mathcal{P}$ holds for a *general* $\alpha \in \mathbb{K}^m$, if there exists a non-empty open subset $\mathcal{U} \subseteq \mathbb{K}^m$ such that the property $\mathcal{P}$ holds for all $\alpha \in \mathcal{U}$. Here $\mathcal{U} \subseteq \mathbb{K}^m$ is open with respect to the Zariski topology. Hence $\mathcal{U}$ is the complement of the zero set of finitely many polynomial functions on $\mathbb{K}^m$. Note that, equivalently a property $\mathcal{P}$ holds for a general $\alpha \in \mathbb{K}^m$, if there is a closed subset $\mathcal{Z} \subseteq \mathbb{K}^m$ such that the $\mathcal{P}$ holds for all $\alpha \notin \mathcal{Z}$.

Lemma 3.6.14. *Let $S = \mathbb{K}[x_1, \cdots, x_N]$ and z be a new variable. Fix positive integers $d_1, \cdots, d_n \in \mathbb{N}$. For $\alpha \in \mathbb{K}^n$, let $I_\alpha = (x_1 - \alpha_1 z^{d_1}, \cdots, x_n - \alpha_n z^{d_n})$. Let $\varphi_\alpha : S[z] \to S[z]/I_\alpha$ be the quotient ring homomorphism.*

1. *The ideal I_α is prime in $S[z]$, and the composition morphism $\mathbb{K}[z] \hookrightarrow S[z] \to S[z]/I_\alpha$ is injective.*
2. *If $F \in \mathbb{K}[x_1, \cdots, x_n]$ is a non-zero polynomial, then $\varphi_\alpha(F) \neq 0$ in $S[z]/I_\alpha$ for a general $\alpha \in \mathbb{K}^n$.*
3. *Let $F \in S \setminus \mathbb{K}[x_1, \cdots, x_n]$, then $\varphi_\alpha(F) \notin \mathbb{K}[z]$ in $S[z]/I_\alpha$, for a general $\alpha \in \mathbb{K}^n$.*
4. *If $F \in S$ is a non-zero polynomial, then $\varphi_\alpha(F) \neq 0$ in $S[z]/I_\alpha$ for a general $\alpha \in \mathbb{K}^n$.*
5. *If $F, G \in S$ have no common factor, then $\gcd(\varphi_\alpha(F), \varphi_\alpha(G)) \in \mathbb{K}[z]$ for a general $\alpha \in \mathbb{K}^n$.*
6. *If $F \in S$ is square-free. Then, for a general $\alpha \in \mathbb{K}^n$, the multiple factors of $\varphi_\alpha(F)$ must be in $\mathbb{K}[z]$.*

The next proposition corresponds to [OS24, Proposition 6.5].

Proposition 3.6.15. *Let $V \subset R$ be a B-lifted strong vector space and $\varphi_\alpha : R[z] \to R[z]/I_\alpha$ be a graded quotient as defined in Definition 3.6.11.*

1. *The ideal I_α is a prime ideal in $R[z]$ and the composition $\mathbb{K}[z] \hookrightarrow R[z] \to R[z]/I_\alpha$ is injective.*
2. *If $F \in R \setminus \{0\}$, then $\varphi_\alpha(F) \neq 0$ for a general $\alpha \in \mathbb{K}^n$.*
3. *If $F \notin \mathbb{K}[V] \subset R$, then $\varphi_\alpha(F) \notin \mathbb{K}[z]$ for a general $\alpha \in \mathbb{K}^n$.*
4. *If $F \notin (V)$ then $\varphi_\alpha(F) \notin (z)$ in $R[z]/I_\alpha$.*

The next proposition corresponds to [OS24, Proposition 6.6].

Proposition 3.6.16. *Let $G : \mathbb{N}^d \to \mathbb{N}^d$ be a function such that $G_i(\delta) \geq h_{B,i} \circ t_2(\delta)$ for all $\delta \in \mathbb{N}^d$. Let $V \subset R_{\leq d}$ be a G-lifted strong vector space and $\varphi_\alpha : R[z] \to R[z]/I_\alpha$ be a graded quotient as defined in Definition 3.6.11. Let $F, G \in R_{\leq d}$ be such that they have no common factor. There exists a non-empty open subset $\mathcal{U} \subseteq \mathbb{K}^{\dim(V)}$ such that for all $\alpha \in \mathcal{U}$ we have:*

1. $\gcd(\varphi_\alpha(F), \varphi_\alpha(G)) \in \mathbb{K}[z]$.

2. *If F, G are homogeneous, then* $\gcd(\varphi_\alpha(F), \varphi_\alpha(G)) = z^k$ *for some* $k \in \mathbb{N}$. *In particular, we have* $\gcd(\varphi_\alpha(zF), \varphi_\alpha(zG)) = z^{k+1}$ *for some* $k \in \mathbb{N}$. *Furthermore, if* $F, G \notin K[V] \subset R$ *then* $\varphi_\alpha(F)$, $\varphi_\alpha(G)$ *are linearly independent.*

3. *If $F \in R$ is a square-free form, then $\varphi_\alpha(F)$ does not have multiple factors other than z^k for some $k \in \mathbb{N}$.*

The next proposition corresponding to [OS24, Proposition 6.9], tells us that if a finite set of forms resulting from a graded quotient has small vector space dimension, then it must be the case that the original set of forms must have small vector space dimension. We refer to this result as *lifting from graded quotients*, as we are lifting our upper bounds for the quotient configurations to the original configurations.

Proposition 3.6.17 (Lifting from graded quotients)**.** *Let $d, e \in \mathbb{N}$ such that $1 \leq d \leq e$. Let $U \subset S_{\leq e}$ be a graded vector space generated by forms $H_1, \cdots, H_t$. Let $R = S/(U)$. Let $V \subset R_{\leq e}$ be a B-lifted strong vector space with basis $F_1, \cdots, F_n \in R$. Let $\varphi_\alpha : R[z] \to R[z]/I_\alpha$ be a graded quotient as defined in Definition 3.6.11. Let $\mathcal{F} \subset R_{\leq d}$ be a finite set of homogeneous elements. Suppose that there exists $D \in \mathbb{N}$ and a dense set $\mathcal{U} \subset \mathbb{K}^n$ such that $\dim \operatorname{span}_{\mathbb{K}}\{\varphi_\alpha(\mathcal{F})\} \leq D$ for every $\alpha \in \mathcal{U}$. Then*

$$\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{F}\} \leq d^2(1+d)^{2n+2} D \cdot \Pi_{i=1}^{t} \deg(H_i) \cdot \Pi_{j=1}^{n} \deg(F_j).$$

While the statement of [OS24, Proposition 6.9] requires that the set of α for which the rank bound holds is an open set, the proof only requires the weaker condition that the set of such α is dense. The fact that this weaker condition suffices will be crucial for us.

For our dimension bounds, especially for those in [GOS25b], we need a more refined version of the above graded quotients, essentially to allow us to compare changes across a series of graded quotients conveniently. We introduce this idea in Section 3.9.1.

We end this section by a quick discussion on how the two ingredients above combine to give the proof of the dimension bound in [OS24]. First, the definition of Sylvester–Gallai configurations is extended to configurations of forms in quotients of polynomial rings, that is, rings of the type $S/(U)$. The definitions of Section 3.3 are treated as special cases when $U = (0)$. The inductive statement is that such generalised configurations have bounded dimension, as long as U is strong enough. The bound on the dimension is only a function of the dimension of U.

Now given a Sylvester–Gallai configuration consisting of forms in ring $R := S/(U)$, consider the forms of highest degree, say d. If this is a (c, δ)-linear-SG, then there exists a small vector space whose span contains these degree d forms, and applying a graded quotient with respect to a strong vector space containing the above small vector space results in a $d-1$ radical Sylvester–Gallai configuration in a ring R' where induction can be applied.

Suppose the degree d forms are not a (c, δ)-linear-SG. Recall the radical structure result (Theorem 3.4.4). By picking c larger than $2d^2(d-1)$ and δ a small enough constant, we can find a set of constantly many degree d forms in $\mathcal{F}$ such that all except a very small fraction of the remaining degree d forms are outside the strong algebra generated by these forms. Now a graded quotient is applied with respect to this space, and this process is repeated. Oliveira and Sengupta show that each such step decreases the potential function defined above by a large amount, and therefore after constantly many steps, every degree d form drops degree, and induction can be applied.

In the next section, we introduce two new tools that we introduced in order to control the dimensions of product Sylvester–Gallai configurations and high degree EK configurations.

3.7 Prime bounds

In this section, we discuss prime bounds, which are an instantiation of the second cornerstone principle discussed in Section 3.4, and one of the main technical contributions of our work. We fix $A := \mathbb{K}[x_1, \ldots, x_n]$ and $S := \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ to be polynomial rings. We fix $\mathfrak{m}$ to be the maximal ideal generated by $x_1, \ldots, x_n$ (we will use $\mathfrak{m}$ to denote both this ideal in A and its extension under the inclusion $A \to S$).

Recall the structural result for radical ideals of [OS22], that for any irreducible polynomial $P \in S \setminus \mathfrak{m}$ of degree d, there are at most $2d^2(d-1)$ pairwise non associate irreducible forms $Q \in A$ such that (P, Q_i) is not radical. The contrapositive of this statement is that if we can find enough Q_i such that (P, Q_i) is not radical, then we must have $P \in \mathfrak{m}$. This is a core part of the proof of the dimension bounds of [OS24].

When dealing with product Sylvester–Gallai configurations, the Sylvester–Gallai relation is guaranteed to be linear only if the ideal generated by the two forms in question is prime, not just radical. Therefore, we are interested in studying how many forms Q_i can exist, such that (P, Q_i) is not prime. Unfortunately, this number is not bounded. For example, if $P = y_1^4 - y_2^2 H$ for some quadratic form $H \in A$, then every Q of the form $H - \ell^2$ is such that (P, Q) is not prime, where ℓ is an arbitrary linear form in A. Indeed we have $y_1^4 - \ell^2 y_2^2 = \left(y_1^2 - \ell y_2\right)\left(y_1^2 + \ell y_2\right) \in (P, Q)$, but neither of these forms are in the ideal. If $s(H) > 2$ then every such Q_i is irreducible, and these Q_i are clearly non associate for distinct ℓ.

Let $K := \overline{K(A)}$ denote the algebraic closure of the fraction field of A. Notice that while P is irreducible as a form in S, it factors in the ring $K[y_1, \ldots, y_m]$, as $P = \left(y_1^2 - \sqrt{H} y_2\right)\left(y_1^2 + \sqrt{H} y_2\right)$. Such a polynomial is said to be absolutely reducible over $K(A)[y_1, \ldots, y_m]$. As we will see, this will be the only barrier to the number of forms Q_i being finite. We show that as long as P is absolutely irreducible, that is, as long as it is irreducible as a polynomial in $K[y_1, \ldots, y_m]$, the number of such Q_i will be finite. In fact we will show that not only are the number of such Q_i finite, but they can be bounded by a function of only d, independent of m and n.

Intuitively, the reason this holds is that being absolutely reducible is a closed condition on the coefficients of a polynomial. This means that if we consider the space of coefficient vectors of polynomials of degree d in $K[y_1, \ldots, y_m]$ up to scalar multiplication with elements of K (which is $\mathbb{P}_K^D$, where $D = \binom{d+m}{m}$) then the set of coefficient vectors of absolutely reducible polynomials is a proper variety in this space. This can be seen as follows. Fix an integer $e < d$. Consider the spaces of polynomials of degrees e and $d - e$, these will be projective spaces of dimension $\binom{e+m}{m}$ and $\binom{d-e+m}{m}$ respectively, call these spaces V_e, V_{d-e}. There is a natural map from the space $V_e \times V_{d-e}$ to the space of polynomials of degree d, which we call V_d. This map is just the multiplication map: given the coefficient vectors of a polynomial of degree e and one of degree $d - e$, the map sends this to the coefficient vector of the product of the polynomials. Observe further that each coordinate function of this map is defined using just the coefficients $0, 1$, that is, they do not use any element of K other than $0, 1$ and in fact the maps do not actually depend on the field K. Since V_e, V_{d-e} are projective, so is the product $V_e \times V_{d-e}$, and therefore the image is closed. This image contains the coefficient vectors of all polynomials that have a factor of degree e. If we repeat this for all $1 \leq e \leq d-1$, and take a union of all the images, we get a set that is closed, and is exactly the set of all absolutely irreducible polynomials. Over the course of the rest of this section, we will make the above formal, and see why it helps.

We divide the rest of this section into three parts. In Section 3.7.1, we show some results above specialisations of absolutely irreducible polynomials. In particular we study when a given specialisation of the variables $x_1, \ldots, x_n$ results in an absolutely irreducible form to get mapped to

an absolutely reducible form. In Section 3.7.2 , we relate prime ideals of the form (P, Q_i) to the above specialisations, and therefore get bounds on the number of such Q_i. In Section 3.7.3, we show how absolutely irreducible and reducible forms behave with respect to general quotients.

3.7.1 Irreducibility of specialisations

In this part, we answer the following question: given a polynomial $P \in S$ that is irreducible as a polynomial in $K[y_1, \dots, y_m]$, which specialisations of the variables α of the variables $x_1, \dots, x_n$ result in P being mapped to a reducible polynomial in $\mathbb{K}[y_1, \dots, y_m]$. To see that such specialisations can exist, we can consider a simple example. Suppose $P = y_1^2 + \ell y_2$, where ℓ is a linear form in the x variables. This form is absolutely irreducible in $K[y_1, \dots, y_m]$, since it is linear in y_2. Any specialisation of the x variables such that $\ell = 0$ causes the image to become reducible. This example shows that the number of such specialisations is not bounded. However, the argument before the start of the section implies that the set of such specialisations is closed (we will formalise this soon). Therefore, our goal will be to control the degree of the closed set of such bad specialisations.

For an absolutely irreducible polynomial $P \in S$, and an element $\alpha \in \mathbb{K}^n$, we denote by P_α the specialisation $P(\alpha_1, \dots, \alpha_n, y_1, \dots, y_m)$. Recall that the goal is to give a condition on α that results in the reducibility of $P_\alpha \in \mathbb{K}[y_1, \dots, y_m]$.

Lemma 3.7.1. *Let $S = \mathbb{K}[x_1, \dots, x_n, y_1, \dots, y_m]$ be a polynomial ring and $A = \mathbb{K}[x_1, \dots, x_n]$. Let $P \in S$ be a polynomial of degree d. Suppose that P is absolutely irreducible. Then there exists a polynomial $H \in A$ of degree $\deg(H) \leq d^2 \cdot 2^{4(m+1)^2 m^{2d}}$ such that $P_\alpha \in \mathbb{K}[y_1, \dots, y_m]$ is irreducible for any $\alpha \in \mathbb{K}^n$ with $H(\alpha) \neq 0$.*

Proof. The proof will follow [Stacks, Tag 0557], with effective versions of the required bounds.

Since $\mathbb{K}$ has characteristic 0, after a linear transformation of the y variables with $\mathbb{Z}$ coefficients, we can assume that the leading coefficient of P with respect to y_m is an element of A, that is, we have $P = \sum_{i \leq d} y_m^i P_i$ where $P_d \in A$. By changing d we can also assume $P_d \neq 0$. The advantage of this change of basis is that now after specialisation, as long as $P_d(\alpha) \neq 0$, the polynomial P_α is monic in y_m, therefore any factors of P_α have to have smaller y_m degree than d.

Fix an integer $1 \leq e \leq d-1$. Consider the universal polynomials Q_1, Q_2 of y_m degrees $e, d-e$ respectively, and total degree less than d. In other words, these are polynomials in the variables $y_1, \dots, y_m$ whose coefficients are distinct variables. We use $z_1, \dots, z_a$ to denote the coefficients, we have $a \leq 2(m+1)m^d$ as a crude upper bound.

Consider the product $Q_1 Q_2$, and write this product as $Q_1 Q_2 = \sum_m q_m(z) y^m$ where m ranges over all monomials of degree at most $2d$, and y_m degree at most d (we use y, z as short hands for sets of variables). If we consider $Q_1 Q_2 - P = \sum_m (q_m(z) - P_m) y^m$, and consider the set of all coefficients $q_m(z) - P_m$, this is a set of unsatisfiable equations in the ring $K[z_1, \dots, z_a]$. Indeed if we had a solution to this system, then we can specialise Q_1, Q_2 to this solution to get a factorisation of P, which we know does not exist.

By the Nullstellensatz therefore, we have $1 \in (q_m(z) - P_m)$, that is, there are polynomials $g_m \in K[z]$ such that $1 = \sum g_m (q_m(z) - P_m)$. Each $q_m(z) - P_m$ is a polynomial of degree 2 in the z variables, and of degree at most d in the x variables. Using the effective Nullstellensatz Theorem 2.3.8 we can deduce that the g_m in fact lie in the ring $K(A)[z_1, \dots, z_a]$, and the common denominator of all the g_m is a polynomial in A of degree at most $d \cdot 2^{4(m+1)^2 m^{2d}}$ (the common denominator is a determinant with entries having degree at most d, and the size of the matrix is as most $2^{4(m+1)^2 m^{2d}}$). Let this common denominator be denoted by H_e.

We claim that for any α such that $H_e(\alpha) \neq 0$, the polynomial P_α does not have a factor of y_m degree e. Indeed if P_α had such a factor, then the set of equations $q_m(z) - (P_\alpha)_m$ would have a

solution in $\mathbb{K}$. However, specialising the equation $1 = \sum g_m(q_m(z) - P_m)$ at α contradicts the above (the fact that $H_e(\alpha) \neq 0$ means that this specialisation is well defined, that is, we are not dividing by 0).

Repeating the above construction for all $1 \leq e \leq d-1$, and setting $H := \prod H_e$ finishes the proof, since any α such that $H(\alpha) \neq 0$ is such that P_α cannot have a valid factorisation. The degree of H is at most $d^2 \cdot 2^{4(m+1)^2 m^{2d}}$. □

We will also need the following lemma that states that absolute irreducibility is mostly unchanged by the addition of elements of A.

Lemma 3.7.2. *Let $S = \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ be a polynomial ring and $A = \mathbb{K}[x_1, \ldots, x_n]$. Let $P \in S$ be a polynomial of degree d. Suppose that P is absolutely irreducible. The number of $Q_i \in S$ such that $P + Q_i$ is not absolutely irreducible is bounded by a function $\mathcal{C}(d, m)$.*

Proof. By [Eis95, Corollary 14.3], the set of reducible forms are a closed subvariety of $\mathbb{P}_K^{\binom{m+d}{d}-1}$. The degree of this variety only depends on m and d, let $\mathcal{C}(d, m)$ be the degree. If $P + Q$ is not absolutely irreducible for some Q, then it must be that $P + Q$ lies on this variety. If this holds for more than $\mathcal{C}(m, d)$ many Q, then the whole line of polynomials of the type $P + Q$ has to lie in the variety, contradicting the fact that P is absolutely irreducible. □

The proof that the set of reducible forms is closed is similar to the proof of Lemma 3.7.1, the only missing ingredient is the main theorem of elimination theory. Therefore it is also possible to obtain an explicit bound for $\mathcal{C}$, but we will not require this.

3.7.2 Bounds on forms generating non prime ideals via a Bertini theorem

We start this section by relating the above absolute irreducibility preserving specialisations with forms $Q_i \in S$ such that (P, Q_i) is not prime. Essentially if (P, Q_i) is not prime, then it must be the case that a general $\alpha \in V(Q)$ is such that P_α is reducible. We require the following technical lemma that is a modification of [OS24, Lemma 4.21].

Lemma 3.7.3. *Let $\phi : \mathcal{A} \to \mathcal{B}$ be a homomorphism of finitely generated $\mathbb{K}$-algebras. Let $Y = \operatorname{Spec}(\mathcal{B})$, $X = \operatorname{Spec}(\mathcal{A})$ and $\pi : Y \to X$ be the corresponding morphism of affine schemes. Suppose that X is irreducible and reduced, that is, $\mathcal{A}$ is an integral domain. Furthermore, suppose that every irreducible component of Y dominates X.*

1. *If $\pi^{-1}(x)$ is irreducible for a general closed point $x \in X$, then Y is irreducible.*

2. *Suppose Y satisfies Serre's property $\mathcal{S}_1$, that is, Y does not have embedded primes. If for every irreducible component W of X, we have that $\pi^{-1}(x)$ is reduced for a dense set of closed points $x \in W$, then Y is reduced.*

Proof. We note that (1) follows directly from [OS24, Lemma 4.21]. Furthermore, in [OS24, Lemma 4.21], statement (2) was proved under the assumption that $\pi^{-1}(x)$ is reduced for a general closed point $x \in W$. However, the same argument works if we have that a dense set of fibers $\pi^{-1}(x)$ is reduced. □

We can apply the above in our setting to deduce the following.

Proposition 3.7.4. *Let $S = \mathbb{K}[x_1, \dots, x_n, y_1, \dots, y_m]$ be a graded polynomial ring with $\deg(x_i) = d_i \geq 1$ and $\deg(y_j) = e_j \geq 1$. Let $A = \mathbb{K}[x_1, \dots, x_n]$ and $K(A)$ the fraction field of A. Let $d \geq 1$ and $P \in S_d \setminus (x_1, \dots, x_n)$ such that P is absolutely irreducible over $K(A)$. Then there are at most $d^2 \cdot 2^{4(m+1)^2 m^{2d}}$ pairwise non-associate irreducible homogeneous elements $Q_i \in A$ such that (P, Q_i) is not prime.*

Proof. Let $H \in \mathbb{K}[x_1, \dots, x_n]$ be the polynomial given by Lemma 3.7.1. We show that any $Q \in A$ that does not divide H is such that (P, Q) is prime, this will complete the proof.

Since $P \notin \mathfrak{m} := (x_1, \dots, x_n)$, we deduce that P is not a zero divisor in $S/(Q)$, therefore Q, P is a regular sequence and $S/(P, Q)$ is Cohen–Macaulay. Furthermore, we also have that for any minimal prime $\mathfrak{p}$ over (P, Q) in S, the ideal $\mathfrak{p} \cap A$ is a minimal prime of (Q) in A. In particular, $\mathfrak{p} \cap A = (Q)$, since Q is irreducible. Let $\pi : \mathbb{A}^{m+n} \to \mathbb{A}^n$ be the morphism corresponding to $A \subset S$. Note that $\pi(x_1, \dots, z_n, y_1, \dots, x_m) = (x_1, \dots, x_n)$ for closed points. Consider the affine schemes given by $Y = \operatorname{Spec}(S/(P, Q))$ and $X = \operatorname{Spec}(A/(Q))$. Note that we have a homomorphism of finitely generated $\mathbb{K}$-algebras $A/(Q) \to S/(P, Q)$, as $(Q) \subset (P, Q) \cap A$. The corresponding morphism of affine schemes is given by $\pi|_Y : Y \to X$ and we have commutative diagram.

$$\begin{array}{ccc} Y & \longrightarrow & \mathbb{A}^{m+n} \\ \downarrow{\scriptstyle \pi|_Y} & & \downarrow{\scriptstyle \pi} \\ X & \longrightarrow & \mathbb{A}^n \end{array}$$

Since Q is irreducible, we have that the scheme X is irreducible and reduced. Since $\mathfrak{p} \cap A = (Q)$ for any minimal prime $\mathfrak{p}$ over (P, Q) in S, we have that any irreducible component of Y dominates X. Therefore, by Lemma 3.7.3, it is enough to show that $\pi|_Y^{-1}(x)$ is irreducible and reduced for a general closed point $x \in X$. Let $x = (c_1, \dots, c_n) \in X$ be a closed point. Let $P_x \in \mathbb{K}[y_1 \dots, y_m]$ denote the polynomial $P(c_1, \dots, c_n, y_1, \dots, y_m)$. Note that $\pi|_Y^{-1}(x) = \operatorname{Spec}(\mathbb{K}[y_1, \dots, y_m]/(P_x))$, as $x \in X$. Since $H \notin (Q)$, we have that $H(x) \neq 0$ for a general closed point $x \in X$. Therefore, by Lemma 3.7.1, we conclude that P_x is irreducible in $\mathbb{K}[y_1, \dots, y_m]$ for a general closed point $x \in X$. Therefore, $\pi|_Y^{-1}(x) = \operatorname{Spec}(\mathbb{K}[y_1, \dots, y_m]/(P_x))$ is irreducible and reduced for a general closed point $x \in X$. Hence (P, Q) is prime. □

The above bound depends on m, which is insufficient for our application. We now use a Bertini theorem to get a bound independent of m. The high level idea is simple: suppose $Q_1, \dots, Q_r$ are such that (P, Q_i) is not prime. We take general hyperplane sections of the form $\sum \alpha_i y_i = 0$, until the number of y variables is reduced to say 5. A Bertini type theorem will show that (P, Q_i) remains not prime after these hyperplane sections, and the forms Q_i clearly remain non associate. Another Bertini type theorem will show that P remains absolutely irreducible. We can now invoke the above results with $m = 5$, to get a bound on r that only depends on d, and not on m.

There are two technical challenges in carrying out the above strategy. The first is that our hyperplane sections are not complete general, since they do not involve the x variables. The second is that we want to preserve non primality, which is slightly more delicate than preserving irreducibility (which is the more familiar setting where Bertini is applied). To overcome these, we use versions of Bertini's theorem that also guarantee that hyperplane sections preserve properties such as reducedness and the property of being Cohen–Macaulay. A list of properties preserved by such hyperplane sections can be found in [FOV99, Section 3.4].

For $\alpha \in \mathbb{K}^m$, let φ_α be the quotient homomorphism

$$\varphi_\alpha : S \to R_\alpha := S/(\alpha_1 y_1 + \cdots + \alpha_m y_m)$$

Note that R_α is also a polynomial ring. In fact, if $\alpha_1 \neq 0$, then $R_\alpha \simeq \mathbb{K}[x_1, \ldots, x_n, y_2, \ldots, y_m]$. Furthermore, $\varphi_\alpha(x_i) = x_i$ for all $i \in [n]$ and hence φ_α is an A-algebra homomorphism. Therefore φ_α extends to a $\overline{K(A)}$-algebra homomorphism

$$\overline{K(A)}[y_1, \ldots, y_m] \to \overline{K(A)}[y_1, \ldots, y_m]/(\alpha_1 y_1 + \cdots + \alpha_m y_m),$$

which we continue to denote by φ_α. For the remainder of this subsection we fix the notations φ_α and R_α defined as above.

The following lemma states that absolute irreducibility is preserved by the above hyperplane sections.

Lemma 3.7.5. *Let $n, m \in \mathbb{N}$ such that $m \geq 4$. Let $S = \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ and $A = \mathbb{K}[x_1, \ldots, x_n]$. Suppose that $P \in S$ is absolutely irreducible over $K(A)$. Then there exists an open dense subset $\mathcal{T} \subset \mathbb{K}^m$ such that $\varphi_\alpha(P) \in R_\alpha$ is absolutely irreducible over $K(A)$ for all $\alpha \in \mathcal{T}$.*

Proof. Let $S' := K[y_1, \ldots, y_m]$. For $\beta \in K^m$, we let $R'_\beta = S'/(\beta_1 y_1 + \cdots + \beta_m y_m)$ and φ_β be the quotient homomorphism $S' \to R'_\beta$. Consider the affine hypersurface $X = \mathrm{Spec}(S'/(P)) \subset \mathbb{A}^m_K$. Since P is an irreducible polynomial over K, we know that X is an irreducible and reduced scheme over K. Note that $\dim(X) \geq 3$ as $m \geq 4$. For any $\beta \in K^m$, let X_β denote the hyperplane section $X \cap V(\beta_1 y_1 + \cdots + \beta_m y_m)$. We will apply the Bertini theorems [FOV99, Corollary 3.4.9, Theorem 3.4.10] to show that a general hyperplane section X_β is also irreducible and reduced.

We consider the linear system on X spanned by $y_1, \ldots, y_m$. Then reducedness of X_β follows from [FOV99, 3.4.8, Corollary 3.4.9]. For irreducibility, we verify that the conditions in [FOV99, Theorem 3.4.10] hold in our setting. We note that the base locus of the linear system $y_1, \ldots, y_m$ is, by definition, the common zero set of $P, y_1, \ldots, y_m$ in K^m. Therefore the base locus is contained in $\{(0, \ldots, 0)\}$, and hence at most 0-dimensional. Since $\dim(X) \geq 3$, we have that the codimension of the base locus of is at least 3. Moreover, consider the rational map $\phi : X \dashrightarrow \mathbb{P}^{m-1}_K$ defined by $x \to [y_1(x) : \cdots : y_m(x)]$. Since ϕ factors through the natural map $\mathbb{A}^m_K \dashrightarrow \mathbb{P}^{m-1}_K$, we have that $\dim(\phi^{-1}(p)) \leq 1$ for any $p \in \mathbb{P}^{m-1}_K$. Since $\dim(X) \geq 3$, we conclude that $\dim(\phi(X)) \geq 2$. Hence the linear system is not composed with a pencil. Therefore, by [FOV99, Theorem 3.4.10], we conclude that X_β is irreducible for a general $\beta \in K^m$.

Since $X_\beta = \mathrm{Spec}(R'_\beta/(\varphi_\beta(P))$, we conclude that $\varphi_\beta(P)$ is irreducible over K for a general $\beta \in K^m$. Therefore there exists an open subset $\mathcal{U} \subset K^m$ such that for all $\beta \in \mathcal{U}$ we have that $\varphi_\beta(P)$ is irreducible over K. Finally, by [Ras99, Proposition 3.3], we know that $\mathcal{U} \cap \mathbb{K}^m$ is an open dense subset of $\mathbb{K}^m$, and we conclude by defining $\mathcal{T} := \mathcal{U} \cap \mathbb{K}^m$. □

The following lemma shows that non prime ideals of the form (P, Q) remain non prime after hyperplane sections.

Lemma 3.7.6. *Let $n, m \in \mathbb{N}$ such that $m \geq 5$ and $S = \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Let $P, Q \in S$ be a regular sequence. Suppose that (P, Q) is not prime. Then the ideal $(\varphi_\alpha(P), \varphi_\alpha(Q))$ is not prime in R_α for a general $\alpha \in \mathbb{K}^m$.*

Proof. Let $X = \mathrm{Spec}(S/(P, Q))$ be the affine scheme defined by P, Q. Since (P, Q) is not prime, we have that X is not irreducible or not reduced. Since P, Q is a regular sequence, we know that X is Cohen-Macaulay. In particular, X does not have any embedded components and all the irreducible components of X have the same dimension. Let $X = X_1 \cup \cdots X_k$ be the irreducible decomposition of X. For $\alpha \in \mathbb{K}^m$, let $H_\alpha \subset \mathbb{A}^{n+m}$ be the hyperplane defined by $\ell_\alpha := \alpha_1 y_1 + \cdots + \alpha_m y_m$. We let X_α and $X_{i,\alpha}$ denote the corresponding hyperplane sections $X \cap H_\alpha$ and $X_i \cap H_\alpha$ respectively. Note that $X_\alpha = \mathrm{Spec}(S/(P, Q, \ell_\alpha)) = \mathrm{Spec}(R_\alpha/(\varphi_\alpha(P), \varphi_\alpha(Q)))$.

Case 1. Non-irreducible X. We will show that if X is not irreducible, then a general hyperplane section X_α is not irreducible. Suppose that X is not irreducible, i.e. $k > 1$.

First, we will apply the Bertini theorem [FOV99, Theorem 3.4.10] to the schemes X_i and show that $X_{i,\alpha}$ is irreducible for general α. Fix $i \in [k]$. Consider the linear system on X_i spanned by $y_1, \ldots, y_m$. Let $\phi : X_i \dashrightarrow \mathbb{P}^{m-1}_{\mathbb{K}}$ be the rational map defined as $x \to [y_1(x) : \cdots : y_m(x)]$. Note that $\dim(X_i) = \dim(X) = n + m - 2 \geq n + 3$. Since $\dim(\phi^{-1}(p)) \leq n + 1$ for all $p \in \mathbb{P}^{m-1}_{\mathbb{K}}$, we conclude that $\dim(\phi(X_i)) > 1$. Therefore the linear system spanned by $y_1, \ldots, y_m$ is not composed of a pencil. We note that the base locus or the common zero set of $y_1, \ldots, y_m$ on X_i is contained in the affine subspace $V(y_1, \ldots, y_m) \subset \mathbb{A}^{n+m}$. Hence the dimension of the base locus is at most n. Since $\dim(X_i) \geq n + 3$, we see that the codimension of the base locus is at least 3. By [FOV99, Theorem 3.4.10], we conclude that $X_{i,\alpha}$ is irreducible for a general α. Note that there are finitely many X_i. Hence, for a general α, we have that $X_{i,\alpha}$ are irreducible for all $i \in [k]$.

Now, we have that $X_\alpha = X_{1,\alpha} \cup \cdots \cup X_{k,\alpha}$. We will show that $X_{i,\alpha} \not\subset X_{j,\alpha}$ for $i \neq j$. Then $X_\alpha = X_{1,\alpha} \cup \cdots \cup X_{k,\alpha}$ will be the irreducible decomposition of X_α. Hence X_α will be non-irreducible, since $k > 1$.

We have $X_i \neq X_j$ for $i \neq j$ and we know that $\dim(X_i) = \dim(X_j)$. Therefore $U := X_i \setminus X_j$ is a non-empty open subset of X_i. Let $Z = X_i \cap X_j$. Note that $\dim(Z) \leq \dim(X_i) - 1 \leq n + m - 3$. We will show that for a general α, we must have $X_{i,\alpha} \cap U \neq \emptyset$. Otherwise, we have $X_{i,\alpha} \subset Z$ for a dense set of α in $\mathbb{K}^m$. Now $\dim(X_{i,\alpha}) = \dim(X_i \cap H_\alpha) \geq n + m - 3$ by [Har77, Propsition 7.1]. Therefore, we must have $\dim(Z) = n + m - 3$ and $X_{i,\alpha}$ is an irreducible component of Z for a dense set of α in $\mathbb{K}^m$. For any irreducible component W of Z, the set of hyperplanes containing W is a vector subspace of $\mathrm{span}_{\mathbb{K}}(y_1, \ldots, y_m)$. In particular, the set of $\alpha \in \mathbb{K}^m$ such that $W \subset H_\alpha$ is an linear subspace of $\mathbb{K}^m$. Since there are only finitely many possibilities for W, we must have that there is an irreducible component W of Z such that $H_\alpha \supset W$ for a dense set of $\alpha \in \mathbb{K}^m$. This is a contradiction since $\dim(V(y_1, \ldots, y_m)) \leq n$. Therefore, $X_{i,\alpha} \cap U \neq \emptyset$, and hence $X_{i,\alpha} \not\subset X_{j,\alpha}$ for a general α, if $i \neq j$.

Case 2. Non-reduced X. By the previous case, we may assume that X is irreducible. We will show that if X is not reduced, then a general hyperplane section X_α is also non-reduced.

Since P, Q is a regular is sequence, the ideal (P, Q) does not have any embedded primes. Since $X = \mathrm{Spec}(S/(P, Q))$ is irreducible, we know that (P, Q) has a unique minimal prime $\mathfrak{p}$. In particular, $\mathrm{rad}(P, Q) = \mathfrak{p}$. Since X is Cohen-Macaulay, it is reduced iff it is generically reduced. Therefore any non-empty open subset $U \subset X$ is a non-reduced scheme. Suppose that there is a dense set of $\alpha \in \mathbb{K}^m$, such that X_α is reduced. We let $U := X \setminus V(y_1, \ldots, y_m)$ and in the following we will show that U is a reduced scheme, and obtain a contradiction.

Let $t_1 \ldots, t_m$ be new variables and let $\mathbb{A}^m$ be the corresponding affine space. We define $S' = \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m, t_1, \ldots, t_m]$. We consider the following affine scheme $\mathcal{X} \subset X \times \mathbb{A}^m \subset \mathbb{A}^{2m+n}$ defined by the ideal $(P, Q, \sum t_i y_i) \subset S'$. Let $q : \mathcal{X} \to \mathbb{A}^m$ and $p : \mathcal{X} \to X$ denote the projection morphisms. For any $\alpha \in \mathbb{K}^m$, we know that the fibre $q^{-1}(\alpha)$ is the hyperplane section X_α. By [FOV99, Lemma 3.4.4], we know that the morphism $p : p^{-1}(U) \to U$ is a locally trivial fibration with fibers $\mathbb{A}^{m-1}$. Therefore, it is enough to show that $p^{-1}(U)$ is reduced.

We note that $P, Q, \sum t_i y_i$ is a regular sequence in S'. Otherwise, $\sum t_i y_i \in \mathfrak{p}\dot{S}'$, as $\mathfrak{p}\dot{S}'$ is the minimal prime of (P, Q) in S'. This is a contradiction, since we may substitute $t_i = 1$ and $t_j = 0$ for $j \neq i$, to obtain that $y_i \in \mathfrak{p}$ for all i. This is a contradiction since $\mathrm{ht}(\mathfrak{p}) = 2$ and $m \geq 5$. Therefore $\mathcal{X}$ is Cohen-Macaulay. Let $\mathcal{X}' \subset \mathcal{X}$ be the union of the irreducible components of $\mathcal{X}$ which dominate $\mathbb{A}^m$ under the morphism $q : \mathcal{X}' \to \mathbb{A}^m$. Note that $\mathcal{X}'$ does not have any embedded primes, as $\mathcal{X}$ is Cohen-Macaulay. Since X is irreducible, we know that X_α is irreducible for a general α (by the Bertini theorem as proved in Case 1.). Thus a general fibre $q^{-1}(\alpha) = X_\alpha$ is irreducible, and we conclude that $\mathcal{X}'$ is irreducible by Lemma 3.7.3. Moreover, we know that the fibers X_α are

non-reduced for a dense set of α, by assumption. Therefore, $\mathcal{X}'$ is reduced by Lemma 3.7.3.

Now $p^{-1}(U)$ is an irreducible open subset of $\mathcal{X}$, as U is irreducible. Therefore $p^{-1}(U)$ is contained in a unique irreducible component of $\mathcal{X}$. Now $p^{-1}(U)$ dominates $\mathbb{A}^m$ under the morphism q, since a general point of X is contained in a general hyperplane H_α parametrised by $\alpha \in \mathbb{A}^m$. Therefore, $p^{-1}(U)$ must be contained in $\mathcal{X}'$. Since $\mathcal{X}'$ is reduced, we obtain that $p^{-1}(U)$ is reduced, as desired. □

With this, we can get what we call a prime bound that is independent of m.

Theorem 3.7.7. *There exists a function $\Pi_1 : \mathbb{N} \to \mathbb{N}$ such that the following holds.*

Let $S = \mathbb{K}[x_1, \dots, x_n, y_1, \dots, y_m]$ be a graded polynomial ring with $\deg(x_i) = d_i \geq 1$ and $\deg(y_j) = e_j \geq 1$. Let $A = \mathbb{K}[x_1, \dots, x_n]$ and $K(A)$ the fraction field of A. Let $d \geq 1$ and $P \in S_d \setminus (x_1, \dots, x_n)$ such that P is absolutely irreducible over $K(A)$. Then there are at most $\Pi_1(d)$ pairwise non-associate irreducible homogeneous elements $Q_i \in A$ such that (P, Q_i) is not prime.

Proof. By Proposition 3.7.4, we know that there exist at most $\widetilde{C}(m, d)$ pairwise non-associate irreducible homogeneous elements $Q_i \in A$ such that (P, Q_i) is not prime. Since there are finitely many such Q_i, we may apply Lemma 3.7.6 to each pair (P, Q_i) to assume that $(\varphi_\alpha(P), \varphi_\alpha(Q_i))$ is not prime for a general $\alpha \in \mathbb{K}^m$. Note that $\varphi_\alpha(Q_i) = Q_i$. We may apply $(m-5)$ number of successive hyperplane sections corresponding to $\alpha^{(\mathbf{1})}, \dots, \alpha^{(\mathbf{m-5})} \in \mathbb{K}^m$ to obtain a quotient ring $S \to R_\alpha$. Note that R_α is isomorphic to $A[y_1, \dots, y_5]$. Let $\widetilde{P}$ be the image of P in R_α under the successive quotient morphism. By choosing $\alpha = (\alpha^{(\mathbf{1})}, \dots, \alpha^{(\mathbf{m-5})})$ generally, we may assume that $\widetilde{P}$ is absolutely irreducible over $K(A)$ by Lemma 3.7.5. Since $\alpha^{(\mathbf{1})}, \dots, \alpha^{(\mathbf{m-5})} \in \mathbb{K}^m$ are chosen generally and there are finitely many Q_i, we conclude that $(\widetilde{P}, Q_i)$ is not prime in R_α for all i. Then by applying Proposition 3.7.4 to R_α, we know that there are at most $\widetilde{C}(5, d)$ number of $Q_i \in A$ such that $(\widetilde{P}, Q_i)$ is not prime in R_α. Therefore we are done by defining $\Pi_1(d) := \widetilde{C}(5, d)$. □

We can also use the above Bertini theorems to improve the bound in Lemma 3.7.2 and make it independent of m.

Theorem 3.7.8. *There exists a function $\Pi_2 : \mathbb{N} \to \mathbb{N}$ such that the following holds.*

Let $S = \mathbb{K}[x_1, \dots, x_n, y_1, \dots, y_m]$ be a graded polynomial ring with $\deg(x_i) = d_i \geq 1$ and $\deg(y_j) = e_j \geq 1$. Let $A = \mathbb{K}[x_1, \dots, x_n]$ and $K(A)$ the fraction field of A. Let $d \geq 1$ and $P \in S_d \setminus (x_1, \dots, x_n)$ such that P is absolutely irreducible over $K(A)$. Then there are at most $\Pi_2(d)$ elements $Q_i \in A$ such that $P + Q_i$ is not absolutely irreducible.

In particular, if $Q \in A$ is a form of degree d, then there are at most $\Pi_2(d)$ non-associate forms in $\operatorname{span}_{\mathbb{K}}\{P, Q\} \setminus Q$ that are not absolutely irreducible.

Proof. The reduction is essentially the same as that in Theorem 3.7.7.

Define $\Pi_2(d) := \max_{c \leq d, 1 \leq m' \leq 5} \mathcal{C}(m', c)$, where $\mathcal{C}$ is the function defined in Lemma 3.7.2. If $m \leq 5$, then the claim is true by definition of $\mathcal{C}$. Assume $m \geq 6$, and towards a contradiction assume that there are more than $\Pi_2(d)$ forms such that $P + Q_i$ is absolutely reducible. Apply $m-5$ random hyperplane sections to P, to obtain a degree d polynomial P'. The polynomial P' is irreducible by Lemma 3.7.5. Since the hyperplane sections do not act on $\mathbb{F}$, each of the forms $P' + Q_i$ is absolutely reducible, which contradicts the definition of $\mathcal{C}(5, d)$. □

We will use $\Pi_3(d)$ to denote the function $\Pi_1(d) + \Pi_2(d)$. This completes our definition on the bounds on the number of ideals that are not prime.

3.7.3 Absolute irreducibility with respect to strong algebras

The final part of this section discusses the relation between absolute irreducibility, strong algebras and general quotients. Suppose V is a vector space that is sufficiently strong, and P is a form of degree d that is not in the ideal (V). We have discussed how we can construct a vector space of small dimension whose algebra contains P, namely $\mathrm{AH}(V, P)$. We have also discussed how the generators of this vector space acts like variables via the Ananyan–Hochster principle. We can use this to define the following useful notion of absolute reducibility with respect to a vector space.

The following definition is from [GOS25c]. In the following definition, the function h is the function whose existence was proved in Lemma 3.6.3.

Definition 3.7.9. Let $B : \mathbb{N}^d \to \mathbb{N}^d$. Suppose $B_i(\delta) \geq A(\eta, i)+3(\sum_i \delta_i - 1)$ for some $\eta \in \mathbb{N}$. Suppose $R = S/(U)$. Suppose $V \subset R$ is a graded vector space that is $h_{2B} \circ t_1$-lifted strong with respect to U. Suppose $P \in R$ is a form. Let V' be the vector space obtained by applying Corollary 3.6.6 to V and P. Let $y_1, \dots, y_a$ be a basis of homogeneous forms of V, and $y_{a+1}, \dots, y_b$ extend this to a basis of V'. We say P is absolutely reducible over V if P is absolutely reducible as a polynomial in $\mathbb{K}(y_1, \dots, y_a)[y_{a+1}, \dots, y_b]$. ◊

The first application of this definition is that it allows Theorem 3.7.7 and Theorem 3.7.8 to be applied in more general settings.

Corollary 3.7.10 (Corollary of Theorem 3.7.7, Theorem 3.7.8)**.** *Let $B : \mathbb{N}^d \to \mathbb{N}^d$. Suppose $B_i(\delta) \geq A(\eta, i) + 3(\sum_i \delta_i - 1)$ for some $\eta \in \mathbb{N}$. Suppose $R = S/(U)$. Suppose $V \subset R$ is a graded vector space that is $h_{2B} \circ t_1$-lifted strong with respect to U. Suppose $P \in R$ is absolutely irreducible with respect to V such that $\deg P = d$. There are at most $\Pi_1(d)$ irreducible non-associate forms $Q_i \in \mathbb{K}[V]$ such that (Q_i, P) is not prime. Further, there are at most $\Pi_2(d)$ pairwise non-associate degree d irreducible forms $H_i \in \mathbb{K}[V]$ such that $\mathrm{span}_{\mathbb{K}}\{P, H_i\} \setminus \{H_i\}$ contains a form that is absolutely reducible over V.*

Proof. Let V' be the vector space obtained by applying Corollary 3.6.6 to V and P. Let $y_1, \dots, y_a$ be a basis of homogeneous forms of V, and $y_{a+1}, \dots, y_b$ extend this to a basis of V'. The ring $\mathbb{K}[V']$ is isomorphic to a polynomial ring in the variables y_i. Further, P is irreducible as a form in $\mathbb{K}[V']$ by definition of absolute irreducibility with respect to V.

Therefore we can apply Theorem 3.7.7 to deduce that there are at most $\Pi_1(d)$ forms $Q_i \in \mathbb{K}[V]$ such that (P, Q_i) is not prime as an ideal of $\mathbb{K}[V']$. Further, since $y_1, \dots, y_b$ form a prime sequence, prime ideals of $\mathbb{K}[V']$ extend to prime ideals in R.

Similarly, by Theorem 3.7.8, there are at most $\Pi_2(d)$ forms $H_i \in \mathbb{K}[V]$ such that $(P, H_i)_d \setminus \{H_i\}$ contains a form that is absolutely reducible over V, when we treat (P, H_i) as an ideal in $\mathbb{K}[V']$. Since both P, H_i have degree d, the ideal (P, H_i) contains the same elements in degree d when treated as an ideal of $\mathbb{K}[V']$ and as an ideal of R. This completes the proof. □

Next we show how the above notion of absolute irreducibility with respect to V interacts with the notion of graded quotients. Essentially, if a form is absolutely reducible with respect to V then a graded quotient will map the form to something reducible. To prove this, we need a few technical lemmas. The first relates irreducibility of polynomials in different rings.

Proposition 3.7.11. *Let $\mathbb{K}[z, y_1, \dots, y_m]$ be a graded polynomial ring with $\deg(z) = 1$. Suppose $P \in \mathbb{K}[z, y_1, \dots, y_m]$ is a homogeneous polynomial. Then the following holds.*

1. *If $\gcd(P, z) = 1$ and P is reducible in $\mathbb{K}[z, y_1, \dots, y_m]$, then P is reducible in $\overline{\mathbb{K}(z)}[y_1, \dots, y_m]$.*

2. If P is reducible in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$, then P is reducible in $\mathbb{K}[z, y_1, \ldots, y_m]$.

Proof. Part 1: Consider the factorisation of P in $\mathbb{K}[z, y_1, \ldots, y_m]$. As long as there exists at least two factors that are not pure polynomials in z, the same factorisation shows that P is reducible in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$. By homogeneity, any factor of P in $\mathbb{K}[z, y_1, \ldots, y_m]$ must be homogeneous, and therefore any factor that is a polynomial in z must be of the form z^e. Since $\gcd P, z = 1$, such a factor does not exist.

Part 2: We can assume that $\gcd(P, z) = 1$, otherwise P is clearly reducible in $\mathbb{K}[z, y_1, \ldots, y_m]$. Since P is homogeneous and $\gcd(z, P) = 1$, [Ful89, Section 2.6, Corollary 5] shows that P reducible in $\mathbb{K}[z, y_1, \ldots, y_m]$ iff $P' := P(1, y_1, \ldots, y_m)$ is reducible in $\mathbb{K}[y_1, \ldots, y_m]$. Suppose P factors in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$ as $P = AB$. Let $A = \sum_\alpha p_\alpha(z) y^\alpha$ and $B = \sum_\beta p_\beta(z) y^\beta$, where y^α, y^β are monomials, and $p_\alpha(z), p_\beta(z) \in \overline{\mathbb{K}(z)} \neq 0$. Consider the universal polynomials $C = \sum_\alpha c_\alpha y^\alpha$ and $D = \sum_\beta d_\beta y^\beta$ where c_α, d_β are formal variables. We can write the product $CD = \sum_\gamma q_\gamma(c, d) y^\gamma$ where q_γ are polynomials with coefficients in $\mathbb{K}$. That $P = AB$ is equivalent to the fact that when we set $c_\alpha = p_\alpha, d_\beta = p_\beta$, each q_γ is equal to the corresponding coefficient of P. Suppose P has degree d as a polynomial in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$, and suppose the space of monomials of degree d in $y_1, \ldots, y_m$ has dimension N.

The coefficients of CD define a map to the space $\mathbb{P}^{N-1}_{\overline{\mathbb{K}[z]}}$, whose image is closed, since projective maps are closed. The image of this map is exactly those polynomials in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$ that factor into two polynomials of the same multidegree as C, D respectively.

We construct the defining equations of this image $F_1, \ldots, F_r$. Each F_i has coefficients in $\mathbb{K}$, since the coefficients of CD themselves are in $\mathbb{K}$. That P is reducible implies that each F_i vanishes on the coefficients of P. In particular, each F_i vanishes when the coefficients of P are specialised to $z = 1$.

Similar to this construction, we can construct the defining equations for the space of polynomials in $\mathbb{K}[y_1, \ldots, y_m]$ of degree d that factor into two polynomials of the same multidegree as C, D. These defining equations are also exactly $F_1, \ldots, F_r$, since the only coefficients that occur in the construction in both cases are those from $\mathbb{K}$. By the above, the coefficients of P', which are exactly the coefficients of P specialised to $z = 1$, satisfy these equations. This shows that P' is reducible, whence P is reducible. □

The next two lemmas discusses the irreducibility of images of forms under certain maps.

Lemma 3.7.12. *Let $B = \mathbb{K}[z, x_1, \ldots, x_n]$, where $\deg(z) = 1$ and $\deg(x_i) = d_i$ for some $d_i \geq 1$. For $\alpha \in \mathbb{K}^n$, let $\mathfrak{p}_\alpha$ be the prime ideal $(x_1 - \alpha_1 z^{d_1}, \ldots, x_n - \alpha_1 z^{d_n}) \subseteq \mathbb{K}[z, x_1, \ldots, x_n]$. If $\mathcal{T} \subseteq \mathbb{K}^n$ is a dense subset, then the set $\mathcal{S} := \{\mathfrak{p}_\alpha \mid \alpha \in \mathcal{T}\}$ is dense in $\operatorname{Spec}(B)$.*

Proof. Suppose that $\mathcal{S}$ is not dense in $\operatorname{Spec}(B)$. Then $\mathcal{S}$ is contained in a proper closed subset $V(I)$ of $\operatorname{Spec}(B)$, where $I \subseteq B$ is a non-zero ideal. Let $V(I) = \cup_j V(\mathfrak{q}_j)$ be the decomposition into irreducible components where $\mathfrak{q}_1, \ldots, \mathfrak{q}_r \subseteq B$ are non-zero prime ideals. Since $\mathcal{S} \subseteq \cup_j V(\mathfrak{q}_j)$, each prime $\mathfrak{p}_\alpha \in \mathcal{S}$ contains at least one of the prime ideals $\mathfrak{q}_1, \ldots, \mathfrak{q}_r$. Let $\mathcal{T}_j := \{\alpha \in \mathcal{T} \mid \mathfrak{q}_j \subseteq \mathfrak{p}_\alpha\}$. Since $\cup_{j=1}^r \mathcal{T}_j = \mathcal{T}$, at least one of the sets $\mathcal{T}_j$ must be dense in $\mathbb{K}^n$. Suppose $\mathcal{T}_k$ is dense. Then we may replace $\mathcal{T}$ with $\mathcal{T}_k$. Let $\mathfrak{q} = \mathfrak{q}_k$. Then we have that $\mathfrak{q} \subseteq \mathfrak{p}_\alpha$ for all $\alpha \in \mathcal{T}$. We will see that this leads to a contradiction.

Let $f \in \mathfrak{q} \subseteq \mathbb{K}[z, x_1, \ldots, x_n]$ be a non-zero polynomial. Let $\deg(f) = e$ and $f = f_0 + \cdots + f_e$ be the decomposition of f into its homogeneous parts with the grading of B. We may write $f_e = g_0 + g_1 z + \cdots + g_d z^d$ where $g_1, \ldots, g_d \in \mathbb{K}[x_1, \ldots, x_n]$ are homogeneous with $\deg(g_i) = e - i$. Consider the map $\varphi_\alpha : B \to B/\mathfrak{p}_\alpha = \mathbb{K}[z]$. Then $\varphi_\alpha(f_e) = (\sum_{i=0}^d g_i(\alpha)) z^e$. As $\mathfrak{p}_\alpha \subseteq B$ is a homogeneous ideal and $f \in \mathfrak{p}_\alpha$, we must have $f_e \in \mathfrak{p}_\alpha$. Hence $\varphi_\alpha(f_e) = 0$ in $\mathbb{K}[z]$. Therefore,

$(\sum_i g_i)(\alpha) = 0$ for all $\alpha \in \mathcal{U}$. Hence $\sum_i g_i = 0$ in $\mathbb{K}[x_1, \ldots, x_n]$, as $\mathcal{U} \subseteq \mathbb{K}^n$ is a non-empty open subset. Now $g_i \in \mathbb{K}[x_1, \ldots, x_n]$ is homogeneous with $\deg(g_i) = e - i$. Hence we must have $g_i = 0$ for all i. Thus $f_e = 0$, which is a contradiction. □

Proposition 3.7.13. *Let $S := \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m]$ be a graded polynomial ring with $\deg(x_i) = d_i \geq 1$ and $\deg(y_j) = e_j \geq 1$. Let $A = \mathbb{K}[x_1, \ldots, x_n]$ and $K(A)$ the fraction field of A. Let $d \geq 1$ and $P \in S_d \setminus (x_1, \ldots, x_n)$. Suppose $\varphi_\alpha : S \to \mathbb{K}[z, y_1, \ldots, y_m]$ is a map defined by $\varphi_\alpha(x_i) = \alpha_i z^{d_i}$ with $\alpha \in \mathbb{K}^n$ and $\varphi_\alpha(y_j) = y_j$.*

1. *If P is absolutely irreducible over $K(A)$, then $\varphi_\alpha(P)$ is absolutely irreducible over $\mathbb{K}(z)$ for a general choice of α.*

2. *Suppose there exists a dense subset $\mathcal{T} \subseteq \mathbb{K}^n$, such that $\varphi_\alpha(P)$ is absolutely irreducible over $\mathbb{K}(z)$ for all $\alpha \in \mathcal{T}$. Then P is absolutely irreducible over $K(A)$.*

Proof. (1) Suppose that P is absolutely irreducible over $K(A)$. Then, by [Stacks, Tag 0557], there exists $f \in A$ such that $\varphi(P)$ is irreducible over $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$ for any ring homomorphism $\varphi : A_f \to \overline{\mathbb{K}(z)}$. Let $\varphi_\alpha : A \to \mathbb{K}[z] \hookrightarrow \mathbb{K}(z) \hookrightarrow \overline{\mathbb{K}(z)}$ be defined by $x_i \mapsto \alpha_i z^{d_i}$. Note that $\varphi_\alpha(f) \in \mathbb{K}[z]$. Let $\deg(f) = e$ and $f = f_0 + \cdots + f_e$ be the decomposition of f into its homogeneous parts. Then $\varphi_\alpha(f) = \sum_{i=0}^{e} f_i(\alpha) z^i$. Let $\mathcal{U} := \mathbb{K}^n \setminus V(f_e)$. Then $\mathcal{U} \subseteq \mathbb{K}^n$ is a non-empty open subset as f_e is non-zero. Moreover, $\varphi_\alpha(f) \neq 0$ for all $\alpha \in \mathcal{U}$. Therefore, for all $\alpha \in \mathcal{U}$, the map φ_α descends to a ring homomorphism $A_f \to \overline{\mathbb{K}(z)}$ by the universal property of localization. Hence $\varphi_\alpha(P)$ is irreducible in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$for a general choice of α, that is, for all $\alpha \in \mathcal{U}$.

(2) Let $\mathcal{T} \subseteq \mathbb{K}^n$ be a dense subset such that $\varphi_\alpha(P)$ is irreducible for all $\alpha \in \mathcal{T}$. Suppose that P is reducible in $\overline{K(A)}[y_1, \ldots, y_m]$. Then it is also reducible in $\overline{K(B)}[y_1, \ldots, y_m]$, where $B := A[z] = \mathbb{K}[z, x_1, \ldots, x_n]$. We will show that this leads to a contradiction.

Consider the affine schemes $X := \operatorname{Spec}(B[y_1, \ldots, y_m]/(P))$ and $Y := \operatorname{Spec}(B)$. The ring homomorphism $B \to B[y_1, \ldots, y_m]/(P)$ gives rise to a morphism of finite type $f : X \to Y$. Let $\eta \in Y$ be the generic point of Y, that is, η is the prime ideal (0). By applying [Stacks, Tag 055A], we know that there exists a non-empty open subset $\mathcal{V} \subseteq Y$ such that the number of irreducible components of the geometric fibre $X_{\overline{\mathfrak{p}}}$ is constant for all $\mathfrak{p} \in \mathcal{V}$.

Note that by definition $X_{\overline{\mathfrak{p}}} = \operatorname{Spec}(\frac{B[y_1,\ldots,y_m]}{(P)} \otimes_B \overline{k(\mathfrak{p})})$, where $k(\mathfrak{p}) = B_{\mathfrak{p}}/\mathfrak{p}B_{\mathfrak{p}}$ is the residue field at $\mathfrak{p}$. Since $k(\mathfrak{p}) = K(B/\mathfrak{p})$, we know that $X_{\overline{\mathfrak{p}}} = \operatorname{Spec}(\overline{K(B/\mathfrak{p})}[y_1, \ldots, y_m]/(P))$. Therefore the number of irreducible components of $X_{\overline{\mathfrak{p}}}$ is the number of irreducible factors of the image of P in $\overline{K(B/\mathfrak{p})}[y_1, \ldots, y_m]$. Note that the geometric generic fibre $X_{\overline{\eta}}$ has at least two irreducible components as P is reducible in $\overline{K(B)}[y_1, \ldots, y_m]$ and $\eta = (0)$. Therefore, $X_{\overline{\mathfrak{p}}}$ is reducible for all $\mathfrak{p} \in \mathcal{V}$. Hence the image of P is reducible in $\overline{K(B/\mathfrak{p})}[y_1, \ldots, y_m]$ for all $\mathfrak{p} \in \mathcal{V}$.

For $\alpha \in \mathbb{K}^n$, let $\mathfrak{p}_\alpha$ be the prime ideal $(x_1 - \alpha_1 z^{d_1}, \ldots, x_n - \alpha_1 z^{d_n}) \subseteq B = \mathbb{K}[z, x_1, \ldots, x_n]$. Note that $B/\mathfrak{p}_\alpha = \mathbb{K}[z]$ and $K(B/\mathfrak{p}_\alpha) = \mathbb{K}(z)$. By assumption we know that the image of P, namely $\varphi_\alpha(P)$, is irreducible in $\overline{\mathbb{K}(z)}[y_1, \ldots, y_m]$ for all $\alpha \in \mathcal{T}$. By Lemma 3.7.12, the set of primes $\mathcal{S} := \{\mathfrak{p}_\alpha \mid \alpha \in \mathcal{T}\}$ is dense in Y. Hence $\mathcal{S} \cap \mathcal{V} \neq \emptyset$. This is a contradiction. □

With all of this, we can prove the following extension of Proposition 3.6.15. For ease of reference, we copy the whole statement here.

Proposition 3.7.14. *Let $V \subset R$ be a B-lifted strong vector space and $\varphi_\alpha : R[z] \to R[z]/I_\alpha$ be a graded quotient as defined in Definition 3.6.11.*

1. *The ideal I_α is a prime ideal in $R[z]$ and the composition $\mathbb{K}[z] \hookrightarrow R[z] \to R[z]/I_\alpha$ is injective.*

2. *If $F \in R \setminus \{0\}$, then $\varphi_\alpha(F) \neq 0$ for a general $\alpha \in \mathbb{K}^n$.*

3. *If $F \notin \mathbb{K}[V] \subset R$, then $\varphi_\alpha(F) \notin \mathbb{K}[z]$ for a general $\alpha \in \mathbb{K}^n$.*

4. *If $F \notin (V)$ then $\varphi_\alpha(F) \notin (z)$ in $R[z]/I_\alpha$.*

5. *If F is absolutely reducible with respect to V then $\varphi_\alpha(F)$ is a reducible form in $R[z]/I_\alpha$ for general α.*

6. *If F is absolutely irreducible with respect to V and $F \notin (V)$, then $\varphi_\alpha(F)$ is irreducible for general α.*

Proof. The proof of item (4) is similar to that of item (3). We now prove item (5). Let V' be the vector space from the definition of F being absolutely reducible with respect to V: in particular, V' is obtained by applying Corollary 3.6.6 to V, F. Suppose $x_1, \ldots, x_a$ is a basis for V, and $y_1, \ldots, y_b$ extends this basis to a basis of V'. By definition, F is absolutely reducible in the ring $A := \mathbb{K}(x_1, \ldots, x_a)[y_1, \ldots, y_b]$.

Let $J_\alpha := I_\alpha \cap A[z]$. The proof of [OS24, Proposition 6.5] shows the following facts: The ideal J_α is generated by $x_i - \alpha_i z^{\deg x_i}$, and that $A[z]/J_\alpha \to R[z]/I_\alpha$ is an injective map. Further, $A[z]/J_\alpha$ is identified by a subalgebra of $R[z]/I_\alpha$ that is generated by a prime sequence.

Since $J_\alpha = \left(x_i - \alpha_i z^{\deg x_i}\right)$, we can apply Proposition 3.7.13 to deduce that F is absolutely reducible in $A(z)/J_\alpha \cong A(z)[y_1, \ldots, y_b]$. Therefore, we can apply Proposition 3.7.11 to deduce that F is reducible in $A[z]/J_\alpha$. Since this a subalgebra of $R[z]/I_\alpha$ generated by a prime sequence, the image of F under this inclusion, which is the same as $\varphi_\alpha(F)$, is also reducible in $R[z]/I_\alpha$. □

This ends our section on prime bounds.

3.8 Dimension bounds for Product Sylvester–Gallai configurations

In this section, we will show dimension bounds for product Sylvester–Gallai configurations. We present the proof of this fact from [GOS25c]. The high level idea of this proof is similar to the proof of dimension bounds for radical Sylvester–Gallai configurations from [OS24].

The proof strategy of Oliveira and Sengupta [OS24] was to induct on d, the highest degree of the forms in a Sylvester–Gallai configuration. By a combinatorial argument, a small vector space V is constructed such that a large fraction of the forms of degree d in the configuration are in the ideal generated by V. Now a graded quotient is applied, which maps each of these forms to a form of lower degree. Of course the codomain of a graded quotient is no longer a polynomial ring, but it is the quotient of a polynomial ring by a set of forms of large strength, therefore by an instantiation of the Ananyan–Hochster principle, it behaves like the quotient of a polynomial ring by a set of variables (which is just a polynomial ring). This requires defining a generalisation of radical Sylvester–Gallai configurations. Finally, this process is repeated till every degree d form drops degree, and a radical Sylvester–Gallai configuration of lower degree is obtained, where induction can be applied. This last step is also delicate, since the number of iterations has to be bounded by a constant, to ensure dimension bounds. Therefore, it is not sufficient for the set of degree d forms to say half at each step. The lower degree forms are used to show that the iteration terminates in constantly many steps.

We start with the definition of a generalisation of product Sylvester–Gallai configurations.

Definition 3.8.1 (Product Sylvester–Gallai configurations)**.** Let $U \subset S$ be a graded finitely generated vector space such that $R := S/(U)$ is a UFD, and let $z \in R_1$. Let $\mathcal{F} := \{z, F_1, \ldots, F_m\} \subset R$ be a finite set of square-free forms of degree at most d. We say that $\mathcal{F}$ is a (d, c, z, R)-product Sylvester–Gallai configuration if the following conditions hold:

1. $\gcd(F_i, F_j) = 1$ for all $i \neq j$, and $\gcd(z, F_i) = 1$ for all $i \in [m]$ (non-associate forms)
2. every F_i is a product of irreducible forms of degree at most c. (factor bound)
3. for every $i \neq j$ (Sylvester–Gallai condition)

$$z \cdot \prod_{k \neq i,j} F_k \in \operatorname{rad}(F_i, F_j).$$

$\Diamond$

The above definition generalises Definition 3.3.2 in a number of ways. The underlying ring is allowed to be any finitely generated UFD over $\mathbb{K}$ as opposed to just a polynomial ring. The forms are allowed to be reducible as long as they are square free and relatively prime. Further, the definition keeps track of the highest degree of any factor of any form in the configuration.

Suppose F_i, F_j are two forms in the configuration. The radical ideal $\operatorname{rad}(F_i, F_j)$ is the intersection of all minimal primes of (F_i, F_j). The Sylvester–Gallai condition above implies that each such minimal prime contains either z, or some form from the configuration other than F_i and F_j. Now suppose F_i factors as $G_i H_i$. Every minimal prime of (F_i, F_j) contains either G_i or H_i. Further, if F_i, F_j are non associate, then so are H_i, F_j and G_i, F_j. This implies that every such minimal prime is a minimal prime of either (G_i, F_j) or (H_i, F_j). The converse is also true. Finally note that if F_k is in some minimal prime of (F_i, F_j) and if F_k factors, one of the factors is in this minimal prime. The above discussion implies essentially that if the forms in our configuration are reducible, we might has well factor them and include all factors in the configuration, and this configuration will satisfy essentially the same Sylvester–Gallai like properties. This leads us to the following definition.

Definition 3.8.2 (Factor set)**.** Let $\mathcal{F}$ be a (d, c, z, R)-product Sylvester–Gallai configuration. A set $\mathcal{I} \subset R$ consisting of every irreducible factor of every form $F \in \mathcal{F}$ is called a factor set of $\mathcal{F}$. [3] $\Diamond$

We write $\mathcal{I}(\mathcal{F})$ if the underlying configuration $\mathcal{F}$ is not clear from context. Since factors of forms are forms, we can write $\mathcal{I} = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_c$, where $\mathcal{I}_j \subset R_j$. The degree of $\mathcal{I}$ is the largest e such that $\mathcal{I}_e \neq \emptyset$. If $\mathcal{F}$ is a (d, c, z, R)-configuration then the factor set $\mathcal{I}$ has degree at most c.

The following lemma captures a Sylvester–Gallai like structural result that $\mathcal{I}$ satisfies.

Lemma 3.8.3. *Suppose $U \subset S$ is a graded finitely generated vector space such that $R := S/(U)$ is a UFD. Suppose $\mathcal{F}$ is a (d, c, z, R)-product Sylvester–Gallai configuration, and suppose $\mathcal{I}$ is a factor set of $\mathcal{F}$. For every $G_1, G_2 \in \mathcal{I} \setminus (z)$, at least one of the following is true.*

- *There is some $F \in \mathcal{F}$ such that $G_1 | F$ and $G_2 | F$.*
- *(G_1, G_2) is not prime.*
- *$|(G_1, G_2) \cap \mathcal{I}| \geq 3$.*

Proof. Let $G_1, G_2 \in \mathcal{I} \setminus (z)$. Suppose (G_1, G_2) is prime, and suppose there is no $F \in \mathcal{F}$ such that $G_1 | F$ and $G_2 | F$. We must show that the third item holds. If $z \in (G_1, G_2)$ then we are done, since $z \in \mathcal{I}$. We are therefore left with the case when $z \notin (G_1, G_2)$.

We have $\operatorname{rad}(F_1, F_2) \subset (G_1, G_2)$. By the Sylvester–Gallai condition on $\mathcal{F}$, and the fact that (G_1, G_2) is prime, we have $F_3 \in (G_1, G_2)$ for some form $F_3 \in \mathcal{F}$ such that $F_3 \neq F_1, F_2$. This implies that $G_3 \in (G_1, G_2)$ for some irreducible factor G_3 of F_3. Since F_1, F_3 are relatively prime and similarly F_2, F_3 are relatively prime we have $G_3 \notin (G_1), (G_2)$. $\square$

[3] The factor set is only defined up to units in R, which are just the nonzero scalars. The properties we are interested in, such as the dimension of the span of the factor set, and ideals generated by pairs of polynomials in $\mathcal{I}$ are independent of such scaling, therefore this ambiguity in the definition does not matter to us.

The following result shows that in order to bound the dimension of a product SG-configuration $\mathcal{F}$, it is enough to bound the dimension of a factor set $\mathcal{I}$.

Proposition 3.8.4. *Suppose $U \subset S$ is a graded finitely generated vector space such that $R := S/(U)$ is a UFD. Suppose $\mathcal{F}$ is a (d, c, z, R)-product Sylvester–Gallai configuration, and suppose $\mathcal{I}$ is a factor set of $\mathcal{F}$. If $\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{I}\} = s$ then $\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{F}\} \leq \binom{s+d}{d}$.*

Proof. Let $P_1, \ldots, P_s \in \mathcal{I}$ be such that every form in $\mathcal{I}$ is a linear combination of forms $P_1, \ldots, P_s$. Consider the subalgebra $\mathbb{K}[\mathcal{I}] \subset R$ generated by $\mathcal{I}$. We know that $\mathcal{F} \subset \mathbb{K}[\mathcal{I}]_{\leq d}$. Now, $\mathbb{K}[\mathcal{I}] = \mathbb{K}[P_1, \cdots, P_s]$. Thus, every form in $\mathcal{F}$ is a linear combination of monomials in $P_1, \ldots, P_s$ of degree at most d. Hence $\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{F}\} \leq \binom{s+d}{d}$. □

Therefore, our goal will be to study these factor sets. Before we do so, we show that the images of product Sylvester–Gallai configurations under graded quotients are also product Sylvester–Gallai configurations.

Definition 3.8.5. Suppose $U \subset S$ is a graded finitely generated vector space such that $R := S/(U)$ is a Cohen-Macaulay UFD. Let $\mathcal{F}$ be a (d, c, z, R)-product Sylvester–Gallai configuration. Let $V \subset R_{\leq d}$ be a $h_B \circ t_2$-lifted strong vector space such that $z \in V$, and $\varphi_\alpha : R[y] \to R[y]/I_\alpha$ be a graded quotient, where $I_\alpha := (V_\alpha)$.

We define a configuration $\varphi_\alpha(\mathcal{F})$ as

$$\varphi_\alpha(\mathcal{F}) := \left\{F' \mid F \in \mathcal{F}, \varphi_\alpha(F) = y^r F', F' \neq 1, (F', y) = 1\right\} \cup \{y\}.$$

◊

Informally, we apply the map φ_α to every element of $\mathcal{F}$, and then factor out all powers of y from the images, and add the result to our set as long as it is different from 1. Further, we also add y to the set.

The next proposition shows that this new set is also a product Sylvester–Gallai configuration.

Proposition 3.8.6. *Suppose $U \subset S$ is a graded finitely generated vector space such that $R := S/(U)$ is CM and a UFD. Suppose $\mathcal{F}$ is a (d, c, z, R)-product Sylvester–Gallai configuration. Suppose $V \subset R_{\leq d}$ is a $h_B \circ t_2$-lifted strong vector space such that $z \in V$, and $\varphi_\alpha : R[y] \to R[y]/I_\alpha$ is a graded quotient. Then $\varphi_\alpha(\mathcal{F})$ is a $(d, c, y, R[y]/I_\alpha)$-product Sylvester–Gallai configuration.*

Proof. We have $y \in \varphi_\alpha(\mathcal{F})$ by construction. By item 3 of Proposition 3.6.16, for any F and F' such that $\varphi_\alpha(F) = y^r F'$, we have that F' is squarefree.

Suppose $F', G' \in \varphi_\alpha(\mathcal{F}) \setminus \{y\}$. Suppose further that $F, G \in \mathcal{F}$ are such that $\varphi_\alpha(F) = y^r F'$ and $\varphi_\alpha(G) = y^s G'$. By item 2 of Proposition 3.6.16 we have $(\varphi_\alpha(F), \varphi_\alpha(G)) = y^t$ for some t. Since $(F', G')|(\varphi_\alpha(F), \varphi_\alpha(G))$, and since $(y, F') = (y, G') = 1$, we deduce that $(F', G') = 1$. Finally, since the map φ_α is degree preserving, the set $\varphi_\alpha(\mathcal{F})$ is a collection of squarefree and pairwise relatively prime forms of degree at most d, each of which factor into forms of degree at most c.

By the definition of $\mathcal{F}$, for any $F, G \in \mathcal{F}$ different from z we have

$$z \cdot \prod_{H \in \mathcal{F} \setminus \{F, G, z\}} H \in \operatorname{rad}(F, G).$$

Applying the map φ_α, we obtain

$$y \cdot \prod_{H \in \mathcal{F} \setminus \{F, G, z\}} y^{t_H} \varphi_\alpha(H) \in \operatorname{rad}\left(y^r F', y^s G'\right) \subset \operatorname{rad}\left(F', G'\right).$$

For each such H, we either have $\varphi_\alpha(H) = y^{t_H}$ for some t_H, or $\varphi_\alpha(H) = y^{t_H} H'$ for some $H' \in \varphi_\alpha(\mathcal{F})$. Therefore we have

$$y \cdot \prod_{H' \in \varphi_\alpha(\mathcal{F}) \setminus \{F', G', y\}} H' \in \operatorname{rad}(F', G').$$

This completes the proof. □

3.8.1 Controlling factor sets

This subsection has a few lemmas that allow us to control factor sets in certain ways. The first lemma states that forms of maximum degree in a factor set cannot generate a prime ideal with many forms of lower degree in the factor set. In our iterative process, this lemma will be crucial in measuring how much progress we are making, and ensuring that our iteration terminates after constantly many steps.

Lemma 3.8.7. *Let U be a graded finitely generated vector space such that $R := S/(U)$ is a CM UFD. Let $1 < c \leq d$ be integers and $\mathcal{F}$ be a (d, c, z, R)-product Sylvester–Gallai configuration with factor set $\mathcal{I}$ of degree c. For any $G \in \mathcal{I}_c$, there are at most $d + c^2 \cdot |\mathcal{I}_c|$ forms $H \in \mathcal{I}_{<c}$ such that (G, H) is prime.*

Proof. Suppose $H_1, \ldots, H_t \in \mathcal{I}_{<c}$ are such that (G, H_i) is prime for all i, and $t > d + c^2 \cdot |\mathcal{I}_c|$. Among these forms, there are at most d forms H_i such that $G | F$ and $H_i | F$ for some $F \in \mathcal{F}$. For any other H_i, the product SG-condition implies $|\mathcal{I} \cap (G, H_i)| \geq 3$, and since $G \in \mathcal{I}_c$ and $H_i \in \mathcal{I}_{<c}$ we have $|\mathcal{I}_c \cap (G, H_i)| \geq 3$.

If we assume $t > d + c^2 \cdot |\mathcal{I}_c|$, by the pigeonhole principle, we have without loss of generality that there is some $G' \in \mathcal{I}_c \setminus \{G\}$ such that $G' \in (G, H_i)$ for $1 \leq i \leq c^2 + 1$. We have $\operatorname{codim}((G, G')) = 2$, and $\deg(G, G') \leq c^2$. Since $(G, G') \subset (G, H_i)$ for each $1 \leq i \leq c^2 + 1$, each such (G, H_i) is a minimal prime of (G, G'). By the associativity formula, (G, G') has at most c^2 distinct minimal primes, therefore without loss of generality we have $(G, H_1) = (G, H_2)$. Since $\deg H_1, \deg H_2 < c$, this implies that $H_1 \in (H_2)$, which contradicts the definition of $\mathcal{I}$. This shows that $t \leq d + c^2 \cdot |\mathcal{I}_c|$. □

The next lemma we show states that if the algebra generated by a strong vector space contains a large fraction of a factor set, then every form of highest degree in the factor set is either in the ideal generated by the vector space, or absolutely reducible with respect to it. If we apply a graded quotient with respect to such an algebra, then the resulting product Sylvester–Gallai configuration will have factor sets of smaller maximum degree, and the dimension of such configurations can be bounded by induction.

Lemma 3.8.8. *Let U be a graded finitely generated vector space such that $R := S/(U)$ is a CM UFD. Let $\mathcal{F}$ be a (d, c, z, R)-product Sylvester–Gallai configuration with factor set $\mathcal{I}$ where $|\mathcal{I}_c| \geq 12(c+1)\Pi_1(c) + d$. Suppose $W \subset S_{\leq d}$ is a $h_{2B} \circ t_1$-lifted strong vector space with respect to U with $z \in W$, such that W satisfies one of the following conditions:*

$$|\mathcal{I}_c \cap \mathbb{K}[W]| \geq 3|\mathcal{I}_c|/4 > 0 \qquad \text{or} \qquad |\mathcal{I} \cap \mathbb{K}[W]| \geq 2(c+1)|\mathcal{I}_c|$$

If $G \in \mathcal{I}_c$ is absolutely irreducible with respect to W, then $G \in (W)$.

Proof. Let $\{z, H_1, \ldots, H_t\} = \mathcal{I} \cap \mathbb{K}[W]$. Let $G \in \mathcal{I}_c$ be absolutely irreducible with respect to W such that $G \notin (W)$. We will show that the existence of such a G results in a contradiction.

Let V be the B-lifted strong vector space $\mathrm{AH}(W, G)$, that is, obtained from W and G by applying Corollary 3.6.6. By Corollary 3.7.10 there are at most $\Pi_1(c)$ many H_i such that (G, H_i) is not prime.

There are at most d forms H_i such that G, H_i divide the same form in $\mathcal{F}$. By Lemma 3.8.3 there are at least $s := t - \Pi_1(c) - d$ forms H_i such that (G, H_i) is prime. Without loss of generality we assume that these are $H_1, \ldots, H_s$. We now consider the two cases, and derive a contradiction from each of them.

Suppose $|\mathcal{I}_c \cap \mathbb{K}[W]| \geq 3|\mathcal{I}_c|/4$. By the assumption on the size of $|\mathcal{I}_c|$ we have $|\mathcal{I}_c \cap \mathbb{K}[W]| \geq 2|\mathcal{I}_c|/3 + \Pi_1(c) + d$. We can assume $H_1, \ldots, H_r \in \mathcal{I}_c$ for $r = 2|\mathcal{I}_c|/3$ and $r \leq s$. For each $1 \leq i \leq r$ we have (G, H_i) is prime, therefore by Lemma 3.8.3 we have $G_i \in (G, H_i) \cap \mathcal{I}_c$. If $G_i \in (W)$ for any i, then $G \in (W)$, which is contrary to assumption. Therefore $G_i \notin (W)$ for all i, and in particular we have $G_i \neq H_j$ for any i, j. Since $r = 2|\mathcal{I}_c|/3$, by the pigeonhole principle we must have $G_j = G_{j'}$ for some j, j', which in turn implies $G \in \left(H_j, H_{j'}\right) \subset (W)$, contradicting our assumption.

Suppose now that the second condition holds, that is $|\mathcal{I} \cap \mathbb{K}[W]| \geq 2(c+1)|\mathcal{I}_c|$. By the assumption on the size of $\mathcal{I}_c$ we have $|\mathcal{I} \cap \mathbb{K}[W]| \geq (c+1)|\mathcal{I}_c| + \Pi_1(c) + d$. Let $r = (c+1)|\mathcal{I}_c|$ so $r \leq s$. Since (G, H_i) is prime, by Lemma 3.8.3 we have $G_i \in (G, H_i)$ for each $1 \leq i \leq r$. Further, we must have $G_i \in \mathcal{I}_c$, as $(G, H_i) \cap R_{<c} = (H_i) \cap R_{<c}$. By the pigeonhole principle, we can assume without loss of generality that $G_1 = \cdots = G_{c+1}$. Therefore for each H_j with $1 \leq j \leq c+1$ we have $G_1 = \alpha_j G + R_j H_j$. If $\alpha_j \neq \alpha_{j'}$ for some $1 \leq j, j' \leq c+1$ then $G \in \left(H_j, H_{j'}\right) \subset (W)$, contradicting assumption. Therefore we can assume that $\alpha_j = \alpha_1$ for all j, and $G_1 - \alpha_1 G = R_j H_j$. Since the forms H_i are non associate, we have $c+1$ distinct factors of $G - G_1$, which contradicts the fact that this form has degree c. □

The final lemma in this section gives us a way of constructing small vector spaces whose algebras contain a large fraction of the forms. It is an instantiation of the first cornerstone principle discussed in Section 3.4. If the set of highest degree forms of our factor set do not form a (r, δ)-linear-SG configuration, then we can essentially find a set of r forms H_i such that for all except a $r\delta$ fraction of the remaining forms F, the ideals (F, H_i) are not prime. If r is bigger than $\Pi_1(c)$, we can deduce that F is not absolutely irreducible over $\mathrm{AH}(H_1, \ldots, H_r)$.

Lemma 3.8.9. *Let $\mu \in (0,1)$, $c \leq d \in \mathbb{N}$, $r := 2 \cdot \Pi_1(c) + d$ and $k \geq 20 \cdot dr/\mu$. Let U be a finitely generated graded vector space with dimension sequence δ_U such that $R := S/(U)$ is a CM UFD. Moreover, suppose $\widetilde{B} \geq h_{2B} \circ t_1$ and U is $h_{2\widetilde{B}} \circ t_k$ strong in S. Let $\mathcal{F}$ be a (d, c, z, R)-product Sylvester–Gallai configuration, with factor set $\mathcal{I}$ satisfying $|\mathcal{I}_c| \geq r$. Then there exists a $\widetilde{B}$-lifted strong vector space W such that*

1. $\dim W \leq C_{2\widetilde{B},i}(t_k(\delta_U)) - \delta_U$.
2. $z \in W$.
3. $|\mathcal{I}_c \cap \mathbb{K}[W]| \geq 2 \cdot \Pi_1(c) + d$.
4. *There are at most $\mu|\mathcal{I}_c|$ forms $G \in \mathcal{I}_c$ s.t. $G \notin (W)$ and G is absolutely irreducible relative to W.*
5. *If further $|\mathcal{I}_c| < \varepsilon|\mathcal{I}|$ for some $0 < \varepsilon < \left(4 \cdot r \cdot c^2\right)^{-1}$ then there are at most $|\mathcal{I}|/2$ forms $G' \in \mathcal{I}_{<c}$ such that $G' \notin (W)$ and G' is absolutely irreducible with respect to W.*

Proof. Set $\nu := \mu/2r$. For any $H \in \mathcal{I}_c$, define

$$\mathcal{F}_{\mathrm{span}}(H) := \{G \mid G \in \mathcal{I}_c, |(G, H) \cap \mathcal{I}_c| \geq 3\}.$$

If $\mathcal{I}_c$ is a (r, ν)-linear SG configuration, then $\mathcal{I}_c$ has a basis of size at most $r + 1 + 8/\nu < k$ by Proposition 3.1.2. Let W be the vector space obtained by applying Corollary 3.6.6 to (0), and the set consisting of z and a basis for $\mathcal{I}_c$. This W has the first four required properties.

Suppose $\mathcal{I}_c$ is not a (r,ν)-linear SG configuration. Let $H_1,\dots,H_{r+1}$ be a set of witnesses to this fact, that is, each H_i is such that $|\mathcal{F}_{\mathrm{span}}(H_i)| < 2\nu|\mathcal{I}_c|$. For each $i \in [r+1]$, let $\{H_{ij}\}_j$ be the set of forms such that $H_i|F$ and $H_{ij}|F$ for some $F \in \mathcal{F}$. For each i, there are at most d such forms H_{ij}. Let W be the vector space obtained by applying Corollary 3.6.6 to (0), and the set consisting of z, H_{ij}, H_i for all $i \in [r+1], j \in [d]$. This set has size at most k, therefore W is $\widetilde{B}$-lifted strong. So far we have shown that W satisfies the first three properties.

Let $\mathcal{B} := \mathcal{I}_c \setminus \bigcup_{i\le r+1} \mathcal{F}_{\mathrm{span}}(H_i)$. We have $|\mathcal{B}| \ge (1 - r\cdot 2\nu)|\mathcal{I}_c| = (1-\mu)|\mathcal{I}_c|$. Let $G \in \mathcal{B}$ be such that G is absolutely irreducible with respect to W, we will show that $G \in (W)$. If $G = H_{ij}$ for some i, j then $G \in \mathbb{K}[W]$ by construction, therefore we can assume that this does not hold. If (G, H_i) is prime for some i, then $|\mathcal{I} \cap (G, H_i)| \ge 3$. Since H_i, G both have degree c, it must be that $|\mathcal{I}_c \cap (G, H_i)| \ge 3$, contradicting the definition of $\mathcal{B}$. Therefore we deduce that (G, H_i) is not prime for every $1 \le i \le r$. Let V be the B-lifted strong vector space $\mathrm{AH}(W, G)$. If $G \notin (W)$ then there at most $\Pi_1(c)$ forms $H \in \mathbb{K}[W]$ such that (G, H) is not prime. Since $r > \Pi_1(c)$, and since (G, H_i) is not prime for all i, we reach a contradiction to Corollary 3.7.10. Therefore $G \in (W)$. This shows that W has the first four properties.

We now show the fifth property for both constructions of W above. We have $|\mathcal{I}_c \cap \mathbb{K}[W]| \ge r$. Let $H_1,\dots,H_r \in \mathcal{I}_c \cap \mathbb{K}[W]$. Let $\mathcal{B} \subset \mathcal{I}_{<c}$ be the set of forms G such that (G, H_i) is prime for some $1 \le i \le r$. By Lemma 3.8.7 we have $|\mathcal{B}| \le r \cdot \left(c^2 \cdot |\mathcal{I}_c| + d\right) \le 2 \cdot r \cdot c^2 |\mathcal{I}_c| \le 2 \cdot r \cdot c^2 \cdot \varepsilon \cdot |\mathcal{I}|$. Suppose $G' \notin \mathcal{B}$ is such that G' is absolutely irreducible with respect to W. Let V be the B-lifted strong vector space $\mathrm{AH}(W, G')$. If $G' \notin (W)$ then there at most $\Pi_1(c)$ forms $H \in \mathbb{K}[W]$ such that (G', H) is not prime. Since $r > \Pi_1(c)$, and since (G', H_i) is not prime for all $1 \le i \le r$, we reach a contradiction. Therefore $G' \in (W)$. The desired bound now holds by the assumption on ε. □

The final statement of the above lemma states that if $\mathcal{I}_c$ is small to start with, then when we construct W, not only is a large fraction of $\mathcal{I}_c$ absolutely reducible with respect to W or in the ideal (W), but so are a large fraction of the forms in $\mathcal{I}_{<c}$. This will be essential in controlling the number of iterations in our iterative process. With this we can show the main theorem of this section.

3.8.2 Proof of main theorem

We now show the main theorem. We will start with a (d, c, z, R)-product Sylvester–Gallai configuration, and repeatedly apply Lemma 3.8.9, and apply a graded quotient with the vector space. The size of the set $\mathcal{I}_c$ will drop by a multiplicative factor in each iteration, however this alone does not guarantee that the iterative process will terminate in constantly many steps. To ensure this, we also use the final statement of Lemma 3.8.9. This statement ensures that at least half of the lower degree forms drop degree under the graded quotient. Again this fact alone is not sufficient to control the number of iterations, since the sizes of the product Sylvester–Gallai configuration can change drastically as we are applying graded quotients. This is where Lemma 3.8.8 will apply: if the size of the configuration drops drastically, it must be because a large fraction of the configuration is in the algebra we are currently working with, and the algebra satisfies the hypothesis of Lemma 3.8.8, which will complete the induction step.

We start by setting up parameters that control the dimensions and strengths of all the spaces we construct, in an effort to make our bounds explicit. We note that the exact parameters are irrelevant, all that matters is that they are independent of the number of variables or the size of the configuration. Let $d \in \mathbb{N}$ be a fixed constant. Let $B : \mathbb{N}^d \to \mathbb{N}^d$ be the ascending function given by $B_i(\delta) := A(\eta, i) + 3(\sum_i \delta_i - 1)$ for all $i \in [d]$, where $A(\eta, i)$ is the function provided by [AH20b, Theorem A].

Let $c \in \mathbb{N}$ be such that $1 \le c \le d$. Define the following parameters:

$$r_c := 12(c+1)(\Pi_1(c)+d), \quad \varepsilon_c := \left(12 \cdot r_c \cdot c^2\right)^{-1}, \quad \mu_c := \varepsilon_c/3, \quad k_c := (40 \cdot r_c \cdot d)/\mu_c, \quad g_c = 9c^2.$$

We inductively define functions $\Lambda_{i,j}$ for each $1 \le i \le d$ and $0 \le j \le g_i$. Let $\Lambda_{1,0} := 2B \circ t_2$. Given $\Lambda_{i,j}$ with $j < g_i$, define $\Lambda_{i,j+1} = h_{2\Lambda_{i,j}} \circ t_{k_i}$. Given Λ_{i,g_i}, define $\Lambda_{i+1,0} = h_{2\Lambda_{i,g_i}} \circ t_{k_{i+1}}$.

For each $1 \le i \le d$ and $0 \le j \le g_i$, let $\Gamma_{i,j} : \mathbb{N} \to \mathbb{N}$, be the following function:

$$\Gamma_{i,j}(n) := \max_{\delta \in \mathbb{N}^d, \|\delta\|_1 \le n} \sum_{\ell=1}^{d} C_{2\Lambda_{i,j},\ell}(t_{k_i}(\delta)),$$

where $\|\delta\|_1$ is the ℓ_1-norm of the integer vector δ.

Define $\eta_c : \mathbb{N} \to \mathbb{N}$ as

$$\eta_c(n) = (d+3)^{4g_c^2+4n+\sum_{i=0}^{g_c} 4\Gamma_{c,i}(n)}.$$

Finally, define $\xi_c : \mathbb{N} \to \mathbb{N}$ as $\xi_1(n) = (51d)^d$, and $\xi_c(n) = \eta_c(n) \cdot \xi_{c-1}(\Gamma_{c,0}(n))$.

Lemma 3.8.10. *Let $U \subset S_{\le d}$ be a Λ_{c,g_c}-strong vector space and $R := S/(U)$ (thus R is a CM UFD). Let $\mathcal{F}$ be a (d,c,z,R)-product Sylvester-Gallai configuration.*

For each $0 \le i < g_c$, there exist vector spaces $V^{(i)} \subset S[z_1,\dots,z_i]$ and graded quotient maps

$$\varphi_{\alpha_i}^{(i)} : (S[z_1,\dots,z_i]/V^{(i)})[z_{i+1}] \to S[z_1,\dots,z_{i+1}]/V^{(i+1)}$$

such that the following hold.

1. *Each $V^{(i)}$ is Λ_{c,g_c-i}-strong.*
2. $\dim V^{(i)} \le \Gamma_{c,g_c-i}(\dim U)$.
3. *$\varphi_{\alpha_{g_c}}^{(g_c)} \circ \cdots \circ \varphi_{\alpha_0}^{(0)}(\mathcal{F})$ is a $(d, c-1, z_{g_c}, S[z_1,\dots,z_{g_c}]/V^{(g_c)})$-product Sylvester-Gallai configuration.*
4. $\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{F}\} \le \eta_c(\dim U) \cdot \dim \operatorname{span}_{\mathbb{K}}\left\{\varphi_{\alpha_{g_c}}^{(g_c)} \circ \cdots \circ \varphi_{\alpha_0}^{(0)}(\mathcal{F})\right\}$.

Proof. We construct $V^{(i)}$ via the following iterative process.

Set $R^{(0)} \leftarrow R$, $z_0 = z$, $V^{(0)} \leftarrow U$, $\mathcal{F}^{(0)} \leftarrow \mathcal{F}$, and let $\mathcal{I}^{(0)}$ be the factor set of $\mathcal{F}^{(0)}$
for $0 \le i \le g_c - 1$ **do**
  if $\left|\mathcal{I}_c^{(i)}\right| \ge r_c$ **then**
    Apply Lemma 3.8.9 to $\mathcal{F}^{(i)}$ with $\mu = \mu_c$, and $k = 20 \cdot d \cdot r_c/\mu_c$ to obtain $W^{(i)}$
  else
    Let $W^{(i)}$ be the vector space obtained by applying Corollary 3.6.6 to $\mathcal{I}_c$
  end if
  Let z_{i+1} be a new variable, and let $\varphi_{\alpha_i}^{(i)} : R^{(i)} \to R^{(i)}[z]/W_{\alpha_i}^{(i)}$ be a graded quotient, where $\alpha_i \in \mathbb{K}^{\dim W^{(i)}}$.
  Set $R^{(i+1)} \leftarrow R^{(i)}[z_{i+1}]/W_{\alpha_i}^{(i)}$, $V^{(i+1)} \leftarrow \ker\left(S[z_1,\dots,z_{i+1}] \to R^{(i+1)}\right)$.
  Set $\mathcal{F}^{(i+1)} \leftarrow \varphi_{\alpha_i}^{(i)}(\mathcal{F}^{(i)})$ according to Definition 3.8.5, and $\mathcal{I}^{(i+1)}$ to be the factor set of $\mathcal{F}^{(i+1)}$.
end for

For $i \in [g_c]$, let $m_i := \left|\mathcal{I}_c^{(i)}\right|$ and $M_i := \left|\mathcal{I}^{(i)}\right|$. We establish the following facts by induction on i.

Fact 1: The space $V^{(i)}$ is Λ_{c,g_c-i}-strong for every i, and satisfies $\dim V^{(i)} \le \Gamma_{c,g_c-i}(\dim U)$.

Proof of fact 1. The base case holds by assumption for $V^{(0)} = U$. Now, if $V^{(i)}$ is Λ_{c,g_c-i}-strong, then $W^{(i)}$ is Λ_{c,g_c-i-1}-lifted strong in both cases of the construction: by Lemma 3.8.9 in the if branch, and by Corollary 3.6.6 in the else branch. Therefore $V^{(i+1)}$ is also Λ_{c,g_c-i-1}-lifted strong. Further, by induction and Lemma 3.8.9, Corollary 3.6.6 the vector space $W^{(i)}$ has dimension sequence bounded by $C_{2\Lambda_{c,g_c-i-1}}(t_k(\delta_{V^{(i-1)}})) - \delta_{V^{(i-1)}}$. Since $V^{(i)} = V^{(i-1)} + W^{(i)}$, we deduce that $V^{(i)}$ has dimension sequence bounded by $C_{2\Lambda_{c,g_c-i-1}}(t_k(\delta_{V^{(i-1)}}))$. By definition of Γ_{c,g_c-i}, we have $\dim V^{(i)} \le \Gamma_{c,g_c-i}(\dim U)$. □

Fact 2: Suppose $m_i \neq 0$ for every $1 \le i \le g_c$. Then $m_i \le \mu_c^i m_0$ and

$$M_i \ge M_1 - 2\mu_c(c+1)m_0/(1-\mu_c) \ge (1/4-\mu_c)m_0 - 2\mu_c(c+1)m_0/(1-\mu_c).$$

Proof of fact 2. Consider the space $W^{(i-1)}$ for $i \ge 1$. If $W^{(i-1)}$ is constructed by following the else branch above, then $\mathcal{I}_c^{(i)} \subset \mathbb{C}\big[W^{(i-1)}\big]$. If $W^{(i-1)}$ satisfies either one of the conditions in Lemma 3.8.8, then every form in $\mathcal{I}_c^{(i)}$ is either in the ideal $\big(W^{(i-1)}\big)$, or absolutely reducible with respect to $W^{(i-1)}$, and hence the images of every such under the map $\varphi_{\alpha_{i-1}}^{(i-1)}$ has degree less than c by item 4 of Proposition 3.7.14. In all of these cases, we get $m_i = 0$. Therefore, we can assume that $W^{(i-1)}$ is constructed by applying Lemma 3.8.9.

Every form in $\mathcal{I}_c^{(i)}$ that is either in the ideal $\big(W^{(i-1)}\big)$, or absolutely reducible with respect to $W^{(i-1)}$ is mapped to a form of degree less than c by the map $\varphi_{\alpha_{i-1}}^{(i-1)}$. By Lemma 3.8.9 there are at most $\mu_c \cdot m_{i-1}$ forms that do not satisfy either of these properties, therefore we get that $m_i \le \mu_c m_{i-1} \le \mu_c^i m_0$ by induction. We further have $\big|\mathcal{I}^{(i-1)} \cap \mathbb{C}\big[W^{(i-1)}\big]\big| < 2(c+1)m_{i-1}$. By item 2 of Proposition 3.7.14, every element in $\mathcal{I}^{(i-1)} \setminus \mathbb{C}\big[W^{(i-1)}\big]$ is mapped to a unique element in $\mathcal{F}^{(i)}$ by the map $\varphi_{\alpha_{i-1}}^{(i-1)}$, therefore we have $M_i \ge M_{i-1} - 2(c+1)m_{i-1}$.

By Lemma 3.8.9, we have $\Big|\mathcal{I}_c^{(0)} \cap \mathbb{C}\big[W^{(0)}\big]\Big| < 3m_0/4$. Further, there are at most $\mu_c m_0$ forms in $\mathcal{I}_c^{(0)} \setminus \big(W^{(0)}\big)$ which are absolutely irreducible with respect to $W^{(0)}$. Thus, at least $(1/4-\mu_c)m_0$ forms are either absolutely reducible with respect to $\big(W^{(0)}\big)$, or in $\big(W^{(0)}\big) \setminus \mathbb{C}\big[W^{(0)}\big]$. By item 2 of Proposition 3.7.14, each of these forms is mapped to a unique element in $\mathcal{F}^{(1)}$ under $\varphi_{\alpha_0}^{(0)}$, therefore $M_1 \ge (1/4-\mu_c)m_0$. Combined with the above derived equation $M_i \ge M_{i-1} - 2(c+1)m_{i-1}$, and the fact that $m_i \le \mu_c^i m_0$ we get the claimed bound on M_i. □

Note that Fact 1 establishes items 1 and 2 of our lemma.

Establishing item 3: Note that, if $m_i = 0$ for some $1 \le i \le g_c$, then $V^{(i)}, \varphi_{\alpha_i}^{(i)}$ satisfies item 3. Thus, all we need to show is that the iterative process above will reach $m_i = 0$ for some $i \le g_c$.

Assume, for the sake of contradiction, that $m_i \neq 0$ for $1 \le i \le g_c$. Facts 1 and 2, combined with $\mu_c \le (12(c+1))^{-1}$, and the fact that $m_i \neq 0$ for $1 \le i \le g_c$ imply the following inequalities:

- $m_i/M_i \le \varepsilon_c$ for every $i \ge 2$.
- $M_i \ge M_1/2$ for every $1 \le i \le g_c$.

Now, define potentials Φ_i for each $1 \le i \le g_c$ by

$$\Phi_i := \sum_{G \in \mathcal{I}^{(i)}} (\deg G)^2.$$

We have $\Phi_1 \leq c^2 M_1$ by definition. Consider a form $G \in \mathcal{I}^{(i)}$ such that $\deg G = e$. This form contributes e^2 to Φ_i. We now consider the contribution of the image of $\varphi^{(i)}_{\alpha_i}(G)$ in $\mathcal{F}^{(i+1)}$. There are 3 cases to analyze:

Case 1: $G \in \left(W^{(i)}\right)$ and G is absolutely irreducible with respect to $W^{(i)}$.

In this case, $\varphi^{(i)}_{\alpha_i}(G) = z^\beta \cdot G'$, for some $\beta \in \mathbb{N}^*$ and G' irreducible. Therefore, $\varphi^{(i)}_{\alpha_i}(G)$ only contributes G' to the factor set $\mathcal{I}^{(i+1)}$. Since $\deg(G') \leq e-1$, its contribution to Φ_{i+1} is at most $(e-1)^2$, and we have $(e-1)^2 < e^2$. So the potential decreases in this case.

Case 2: G is absolutely reducible with respect to $W^{(i)}$.

In this case, $\varphi^{(i)}_{\alpha_i}(G) = z_{i+1}^{e'} \prod G_j$, with $\sum \deg G_j \leq e$. Therefore t factors G_i of G contribute $\sum (\deg G_j)^2$ to Φ_{i+1}. Since $\sum(\deg G_j) < e$ we have $\sum(\deg G_j)^2 < e^2$. So the potential also decreases in this case.

Case 3: $G \notin (W^{(i)})$ and it remains absolutely irreducible with respect to $W^{(i)}$.

In this case, $\varphi^{(i)}_{\alpha_i}(G) \in \mathcal{I}^{(i+1)}$ and $\deg(\varphi^{(i)}_{\alpha_i}(G)) = e$, so the contribution of G to the potential doesn't change.

By Lemma 3.8.9, and the fact that $m_i/M_i \leq \varepsilon_c$ for $i \geq 2$ proved above, at least $(M_i - m_i)/2$ forms in $\mathcal{I}^{(i)}$ fall under cases 1 and 2. The bounds on M_i, m_i imply that $(M_i - m_i)/2 \geq M_i/4$. Therefore, combined with the above case analysis, we have $\Phi_{i+1} \leq \Phi_i - M_i/4 \leq \Phi_i - M_1/8$. After $8c^2$ steps, we have $\Phi_i = 0$, which implies that the factor set is empty. Therefore, in iteration $8c^2 + 1 < 9c^2$, we must have $m_i = 0$, which contradicts our assumption.

Proving item 4 (dimension bounds): Since $\dim V^{(i)} \leq \Gamma_{c,g_c-i}(\dim U)$, and since $V^{(i)}$ is generated in degree at most d, we can bound the product of the degrees of a basis of $V^{(i)}$ by $e_i := (d+3)^{\Gamma_{c,g(c)-i}(\dim U)}$. Let $e_0 := (d+3)^{\dim U}$ be a bound on the product of the degrees of a basis of U. By Proposition 3.6.17 we have

$$\operatorname{span}_{\mathbb{K}}\left\{\varphi^{(g_i)}_{\alpha_{g_i}} \circ \cdots \circ \varphi^{(0)}_{\alpha_0}(\mathcal{F})\right\} \leq (d+3)^{4+\dim V^{(i-1)}} e_i e_{i-1} \operatorname{span}_{\mathbb{K}}\left\{\varphi^{(g_{i-1})}_{\alpha_{g_{i-1}}} \circ \cdots \circ \varphi^{(0)}_{\alpha_0}(\mathcal{F})\right\}.$$

By induction, we can deduce that that

$$\dim \operatorname{span}_{\mathbb{K}}\{\mathcal{F}\} \leq (d+3)^{4g_c^2 + 4(\dim U) + \sum_{i=0}^{g_c} 4\Gamma_{c,i}(\dim U)} \cdot \dim \operatorname{span}_{\mathbb{K}}\left\{\varphi^{(g_c)}_{\alpha_{g_c}} \circ \cdots \circ \varphi^{(0)}_{\alpha_0}(\mathcal{F})\right\}.$$

Substituting the definition of $\eta_c(n)$ proves item 4.

□

We can now prove our main theorem, with Lemma 3.8.10 acting as our inductive step.

Theorem 3.8.11. *Let $U \subset S_{\leq d}$ be a Λ_{d,g_d}-strong vector space and $R = S/U$ (hence R is a CM UFD). If $\mathcal{F}$ is a (d,c,z,R)-product Sylvester-Gallai configuration then* $\dim \operatorname{span}_{\mathbb{C}}\{\mathcal{F}\} \leq \xi_c(\dim U)$. *In particular, if $\mathcal{F}$ is a $(d,c,z,S[z])$-product Sylvester-Gallai configuration then* $\dim \operatorname{span}_{\mathbb{C}}\{\mathcal{F}\} \leq \xi_c(0)$.

Proof. The second statement follows from the first since the zero vector space is Λ_{d,g_d}-strong. We prove the theorem by induction on c. Let $\mathcal{I}$ be the factor set of $\mathcal{F}$.

Base case: $c = 1$. That is, when every $G \in \mathcal{I}$ is a linear form.

By Lemma 3.8.3, for every $G_1, G_2 \in \mathcal{I} \setminus \{z\}$, either $G_1 | F$ and $G_2 | F$ for some form $F \in \mathcal{F}$, or $|(G_1, G_2) \cap \mathcal{I}| \geq 3$. Consequently, as long as $|\mathcal{I}| \geq 3d$, the set $\mathcal{I}$ is a $(1, 1/2)$-linear SG configuration, and $\dim \operatorname{span}_{\mathbb{C}}\{\mathcal{I}\} \leq 51 \leq 51d$. If $|\mathcal{I}| \leq 3d$ then we trivially have $|\mathcal{I}| \leq 51d$, and Proposition 3.8.4 implies $\dim \operatorname{span}_{\mathbb{C}}\{\mathcal{F}\} \leq (51d)^d = \xi_1(\dim U)$.

Inductive step: suppose now that $c > 1$.

In this case, U satisfies the hypothesis of Lemma 3.8.10. Thus, there is a vector space V and a sequence of graded quotient maps $\varphi_{\alpha_i}^{(i)}$ such that $\varphi(\mathcal{F})$ is a $(d, c-1, z_{g_c}, S[z_1, \ldots, z_{g_c}]/V)$-product Sylvester-Gallai configuration, where $\varphi(\mathcal{F})$ is the composition of the graded quotients $\varphi_{\alpha_i}^{(i)}$. The lemma also guarantees that $\dim V \leq \Gamma_{c,0}(\dim U)$, and that V is $\Lambda_{c,0}$-strong, and hence $\Lambda_{c-1,g_{c-1}}$-strong. Therefore we can apply the inductive hypothesis, and we have $\dim \operatorname{span}_{\mathbb{C}}\{\varphi(\mathcal{F})\} \leq \xi_{c-1}(\dim V) \leq \xi_{c-1}(\Gamma_{c,0}(\dim U))$. By item 4 of Lemma 3.8.10 we have

$$\dim \operatorname{span}_{\mathbb{C}}\{\mathcal{F}\} \leq \eta_c(\dim U) \cdot \xi_{c-1}(\Gamma_{c,0}(\dim U)) \leq \xi_c(\dim U). \qquad \square$$

3.9 Dimension bounds for high degree EK configurations

In this section, we show dimension bounds for high degree EK configurations. As in the previous section, our proof is inductive and therefore we first define a generalisation of Definition 3.3.7. We note a slight change in notion in this section. In the previous sections, we typically used B to denote the strength of a vector space. In this section, the letter B is reserved for elements of our configuration, and therefore we use μ to denote the strength.

Definition 3.9.1 (High degree EK configurations)**.** Let $U \subset S$ be a graded finitely generated vector space such that $R := S/(U)$ is a UFD, and let $z \in R_1$. Let $\mathcal{A}, \mathcal{B}, \mathcal{C} \subset R$ be pairwise disjoint finite sets of irreducible forms of degree at most d. We say that $\mathcal{A}, \mathcal{B}, \mathcal{C}$ is a (d, z, R)-EK configuration if the following conditions hold:

1. $z \notin \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$, and the union $\{z\} \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ consists of pairwise non-associate forms.

2. For every $A \in \mathcal{A}, B \in \mathcal{B}$ we have

$$z \cdot \prod_{C \in \mathcal{C}} C \in \operatorname{rad}(A, B).$$

3. For every $A \in \mathcal{A}, C \in \mathcal{C}$ we have

$$z \cdot \prod_{B \in \mathcal{B}} B \in \operatorname{rad}(A, C).$$

4. For every $B \in \mathcal{B}, C \in \mathcal{C}$ we have

$$z \cdot \prod_{A \in \mathcal{A}} A \in \operatorname{rad}(B, C).$$

$\Diamond$

In the case when the ideals occurring are prime, for example if (A, B) is prime for some $A \in \mathcal{A}, B \in \mathcal{B}$, then the condition above implies that either $z \in (A, B)$ or $C \in (A, B)$ for some $C \in \mathcal{C}$. We then say that C is the EK-image of A, B. The image C might not be unique, we arbitrarily fix a choice for every pair.

Similar to product Sylvester–Gallai configurations, we can define images of EK configurations under graded quotients.

Definition 3.9.2. Let $U \subset S$ be a graded finitely generated vector space such that $R := S/(U)$ is a CM UFD. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be a (d, z, R)-EK configuration. Let $V \subset R_{\leq d}$ be a $h_B \circ t_2$-lifted strong vector space such that $z \in V$, and $\varphi_\alpha : R[y] \to R[y]/I_\alpha$ be a graded quotient.

We define a configuration $\varphi_\alpha(\mathcal{A}), \varphi_\alpha(\mathcal{B}), \varphi_\alpha(\mathcal{C})$ as

$$\varphi_\alpha(\mathcal{A}) := \big\{A' \,|\, A \in \mathcal{A}, A'|\varphi_\alpha(A), A' \neq 1, (A', y) = 1, A' \text{ irreducible}\big\},$$
$$\varphi_\alpha(\mathcal{B}) := \big\{B' \,|\, B \in \mathcal{B}, B'|\varphi_\alpha(B), B' \neq 1, (B', y) = 1, B' \text{ irreducible}\big\},$$
$$\varphi_\alpha(\mathcal{C}) := \big\{C' \,|\, C \in \mathcal{C}, C'|\varphi_\alpha(C), C' \neq 1, (C', y) = 1, B' \text{ irreducible}\big\}.$$

$\Diamond$

Proposition 3.9.3. *Let $U \subset S$ be a graded finitely generated vector space such that $R := S/(U)$ is a CM UFD. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be a (d, z, R)-EK configuration. Let $V \subset R_{\leq d}$ be a $h_B \circ t_2$-lifted strong vector space such that $z \in V$, and $\varphi_\alpha : R[y] \to R[y]/I_\alpha$ be a graded quotient. Let $\varphi_\alpha(\mathcal{A}), \varphi_\alpha(\mathcal{B}), \varphi_\alpha(\mathcal{C})$ be the image of the EK configuration under the graded quotient as defined in Definition 3.9.2. There is an open subset $\mathcal{U} \subset \mathbb{K}^{\dim V}$ such that for every $\alpha \in \mathcal{U}$, the sets $\varphi_\alpha(\mathcal{A}), \varphi_\alpha(\mathcal{B}), \varphi_\alpha(\mathcal{C})$ are an $(d, y, R[y]/I_\alpha)$-EK configuration.*

Proof. Let $\mathcal{F} := \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ and $\varphi_\alpha(\mathcal{F}) := \varphi_\alpha(\mathcal{A}) \cup \varphi_\alpha(\mathcal{B}) \cup \varphi_\alpha(\mathcal{C})$. We start by verifying the first condition for the configuration to be an EK configuration, which is equivalent to $\varphi_\alpha(\mathcal{F}) \cup \{y\}$ consisting of pairwise non associate irreducible forms.

By construction, the forms in $\varphi_\alpha(\mathcal{F}) \cup \{y\}$ are irreducible. By item 3 of Proposition 3.6.16, for any F and F' such that $\varphi_\alpha(F) = y^r F'$, we have that F' is squarefree. Suppose $F', G' \in \varphi_\alpha(\mathcal{F}) \setminus \{y\}$ are irreducible factors of $\varphi_\alpha(F), \varphi_\alpha(G)$ respectively, for $F, G \in \mathcal{F}$. By item 2 of Proposition 3.6.16 we have $(\varphi_\alpha(F), \varphi_\alpha(G)) = y^t$ for some t. Since $(F', G')|(\varphi_\alpha(F), \varphi_\alpha(G))$, and since $(y, F') = (y, G') = 1$, we deduce that $(F', G') = 1$. Therefore the forms in $\varphi_\alpha(\mathcal{F}) \cup \{y\}$ are pairwise non associate.

Suppose $A' \in \varphi_\alpha(\mathcal{A}), B' \in \varphi_\alpha(\mathcal{B})$ are irreducible factors of $\varphi_\alpha(A), \varphi_\alpha(B)$ respectively, for $A \in \mathcal{A}, B \in \mathcal{B}$. We have

$$z \cdot \prod_{C \in \mathcal{C}} C \in \operatorname{rad}(A, B).$$

Applying the map φ_α, we obtain

$$y \cdot \prod_{C \in \mathcal{C}} y^{t_C} \varphi_\alpha(C) \in \operatorname{rad}(\varphi_\alpha(A), \varphi_\alpha(B)) \subset \operatorname{rad}\big(A', B'\big).$$

For each such C, we either have $\varphi_\alpha(C) = y^{t_C}$ for some t_C, or $C'|\varphi_\alpha(C)$ for some $C' \in \varphi_\alpha(\mathcal{C})$. Therefore we have

$$y \cdot \prod_{C' \in \varphi_\alpha(\mathcal{C})} C' \in \operatorname{rad}\big(A', B'\big).$$

This shows the second condition for EK configurations, and a symmetric argument shows the third and fourth conditions. $\square$

Before we discuss any more specific results about EK configurations, we show some general statements that will be useful. The first of these is a structural statement that says that we can control any set of forms where the strength of large sets is one.

Lemma 3.9.4. *Let $\mu : \mathbb{N}^e \to \mathbb{N}^e$ be an ascending function such that $\mu_i(\delta) \geq A(\eta, i) + 3\|\delta\|_1$, where A is the function defined in [AH20b, Theorem A]. Let $U \subseteq S_{\leq e}$ be a graded finitely generated vector space such that $R := S/(U)$ is a UFD, and $z \in R_1$. Let $\mathcal{G}$ be a finite set non-associate irreducible homogeneous forms in R_d. Suppose that for any subset $\mathcal{S} \subseteq \mathcal{G}$ with $|\mathcal{S}| > b$ we have $s(\mathcal{S}) \leq 1$. Then there exists a subset $T \subseteq \mathcal{G}$ such that the following holds.*

1. $|T| \leq (\Pi_2(d)+1)\cdot b$.

2. *Let $Y \subseteq R$ be any $h_{2\mu}\circ t_1$-lifted strong vector space with $T \subseteq \mathbb{K}[Y]$. Then every form in $\mathcal{G}$ is absolutely reducible over Y or in the ideal (Y).*

Proof. For convenience, in this proof, any subset $\mathcal{S}$ with $s(\mathcal{S}) > 1$ is called a high strength set. Consider all high strength subsets of $\mathcal{G}$ partially ordered by inclusion. Let $\mathcal{S}_1 \subseteq \mathcal{G}$ be a maximal high strength subset. We will define subsets $\mathcal{H}_i \subseteq \mathcal{G}$ iteratively.

We set $\mathcal{H}_1 := \mathcal{S}_1$. For every form $F \in \mathcal{G} \setminus \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_1\}$ we have $s(F, H_1) = 1$ for some $H_1 \in \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_1\}$.

Now let $\mathcal{S}_2 \subseteq \mathcal{G} \setminus \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_1\}$ be a maximal high strength set. If $\{F\} \cup \mathcal{S}_2$ is high strength for some $F \in \mathcal{G}$, then we must have $F \in \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_1\}$. Therefore if we define $\mathcal{H}_2 := \mathcal{H}_1 \cup \mathcal{S}_2$, then for every form $F \notin \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_2\}$ there is a $H_2 \in \mathcal{S}_2$ such that $s(F, H_2) = 1$.

We repeat this process iteratively and set $\mathcal{H}_{i+1} := \mathcal{H}_i \cup \mathcal{S}_{i+1}$ where $\mathcal{S}_{i+1}$ is a maximal strong subset of $\mathcal{G} \setminus \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_i\}$. Note that $|\mathcal{S}_i| \leq b$ for all i by assumption. Therefore, in each iteration, we add at most b forms. We do this $\Pi_2(d)+1$ times. Then we have $\mathcal{H}_{\mathcal{C}(d)+1} \leq (\Pi_2(d)+1)b$. Let $T := \mathcal{H}_{\Pi_2(d)+1}$. We will show that this set T satisfies the desired property. Let $Y \subseteq R$ be a $h_{2B}\circ t_1$-lifted strong vector space such that $T \subseteq \mathbb{K}[Y]$.

Let $F \in \mathcal{G}$ be a form that is not in the ideal (Y). Hence $F \notin \operatorname{span}_{\mathbb{K}}\{\mathcal{S}_i\}$ for all i and also $F \notin \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_i\}$ for all i. The set $\mathcal{S}_i$ is a maximal high strength subset of $\mathcal{G} \setminus \operatorname{span}_{\mathbb{K}}\{\mathcal{H}_{i-1}\}$. Therefore, for every $1 \leq i \leq \Pi_2(d)+1$, there is a $H_i \in \operatorname{span}_{\mathbb{K}}\{\mathcal{S}_i\}$ such that $s(F, H_i) = 1$. Let $P_i = F + \beta_i H_i$ with $s(P_i) = 1$. In particular, P_i is reducible in R. Note that $P_i \notin (Y)$, as $F \notin (Y)$. Hence no irreducible factor of P_i can lie in $\mathbb{K}[Y]$. Hence, the forms P_i are absolutely reducible over Y. If $P_i = \gamma P_j$, then $F \in (H_i, H_j)$ by the assumption that the forms in $\mathcal{G}$ are non-associate, therefore $F \in \mathbb{K}[Y]$. If the forms P_i, P_j are all non-associate and not in (Y), then by Corollary 3.7.10 the form F must be absolutely reducible with respect to Y. This completes the proof. □

The above lemma can be compared to the fact that any set of quadratic forms close to a vector space W where every maximal integral sequence has bounded length can be controlled, which was a key fact in establishing dimension bounds on high dimensional quadratic Sylvester–Gallai configurations.

When studying EK configurations, we will find ourselves in the following situation: in the notion from the statement of Lemma 3.9.4, we will find a strong vector space V such that set of forms $\mathcal{G}$ are all absolutely irreducible over V, and the images of $\mathcal{G}$ under a graded quotient φ_α satisfy the hypothesis of Lemma 3.9.4. Of course in this case, we can apply a graded quotient and apply Lemma 3.9.4, and deduce the existence of a strong vector space Y_α in a quotient ring $S[z]/U'$ such that every form in $\varphi_\alpha(\mathcal{G})$ is either in the in the ideal (Y_α) or absolutely reducible with respect to Y_α. If we continue with the proof, and eventually deduce a dimension bound on $\varphi_\alpha(\mathcal{G})$ (the dimension bound itself might be obtained by applying more graded quotients), then we can use Proposition 3.6.17 to deduce a dimension bound on $\mathcal{G}$. Indeed this is the proof strategy in [OS24; OS24; GOS25c].

However, what we cannot deduce directly from the discussion so far is existence of a vector space Y of constant dimension in the ring R/U itself such that every form in $\mathcal{G}$ is absolutely reducible over Y, or in the ideal (Y). In other words, we can lift dimension bounds from graded quotients, but not any weaker structural result. An obvious attempt at doing so would be to take lifts of the vector spaces Y_α obtained by applying Lemma 3.9.4. However, it is not clear how to combine various lifts to get a single vector space of constant dimension. As we will see, in this particular setting, the fix will be to take preimages of the sets $T_\alpha \subset \varphi_\alpha(\mathcal{G})$, instead of preimages of the vector spaces Y_α. Since there are only finitely many possible choices of subsets, there will be one particular subset T such

that for all α in some open set, we have $\varphi_\alpha(T) = T_\alpha$. We will be able to show that $\mathrm{AH}(V,T)$ for this choice of T will have the required property. A small price we will have to pay is that we will have to start keeping track of which α are good. However this will be a fairly mechanical process, and does not add any technical difficulty to the proofs. It will turn out the set of good α will no longer be an open set, but just a dense a set. This is still sufficient for us to lift both dimension bounds, and the other structural results. We now introduce the notion of compatible graded quotients, which will allow us to carry out the above.

3.9.1 Compatible Quotients

Compatible quotients are a formal way of writing the composition of two graded quotients as a single graded quotient.

Definition 3.9.5 (Compatible graded quotients)**.** Let $R = S/(U)$ and $V, Y \subseteq R$ be finite dimensional $\mathbb{K}$-vector spaces such that $V \subseteq Y$. Let $\mathcal{F}_V, \mathcal{F}_Y$ be $\mathbb{K}$-linear bases of V, Y respectively. We say that $\mathcal{F}_V, \mathcal{F}_Y$ are compatible bases if $\mathcal{F}_V \subseteq \mathcal{F}_Y$.

Consider the compatible bases given by $\mathcal{F}_V = \{F_1, \ldots, F_m\}$ and $\mathcal{F}_Y = \{F_1, \ldots, F_m, G_1, \ldots, G_n\}$. Let $\alpha \in \mathbb{K}^m$,$\beta \in \mathbb{K}^n$ and let y_1, y_2 be new variables. Let $\varphi_{V,\alpha} : R[y_1] \to R[y_1]/I_{V,\alpha}$ be the graded quotient given by $F_i \mapsto \alpha_i y_1^{\deg(F_i)}$. Let $\varphi_{Y,(\alpha,\beta)} : R[y_1] \to R[y_1]/I_{Y,(\alpha,\beta)}$ be the graded quotient that sends $F_i \mapsto \alpha_i y_1^{\deg(F_i)}$ and $G_j \mapsto \beta_j y_1^{\deg(G_j)}$. Then we will say that $\varphi_{V,\alpha}$ and $\varphi_{Y,(\alpha,\beta)}$ are compatible graded quotients. ◊

Given two compatible quotients $\varphi_{V,\alpha}$ and $\varphi_{Y,(\alpha,\beta)}$, we would like to express $\varphi_{Y,(\alpha,\beta)}$ as a composition of $\varphi_{V,\alpha}$ and another graded quotient of the form $\varphi_{W,\beta}$, where $W \subseteq R[y_1]/I_{V,\alpha}$ is some vector space. However, in order to make such a composition work, we need to include the variable y_1 in the vector space W, and we will need to adjoin a new variable y_2 for the second quotient map. This leads to an incompatibility where it is not not enough to use a quotient of the form $\varphi_{W,\beta}$, since we also need to send $y_1 \mapsto \gamma y_2$, for some scalar $\gamma \in \mathbb{K}$, under the second quotient. This naturally leads to the notion of a twisted quotient defined below. Moreover, in Proposition 3.9.8 we note that we can indeed decompose $\varphi_{Y,(\alpha,\beta)}$ into a composition of $\varphi_{V,\alpha}$ and a twisted quotient.

Definition 3.9.6 (Twists)**.** Let $\gamma \in \mathbb{K}^*$ and $d_1, \ldots, d_n \in \mathbb{N}$. For $\beta \in \mathbb{K}^n$, we define the twist of β by γ as the natural action of the torus given by $\gamma \cdot \beta = (\gamma^{d_1}\beta_1, \ldots, \gamma^{d_n}\beta_n)$. ◊

Definition 3.9.7 (Twisted quotient)**.** Let $V \subseteq R$ with a basis $\{F_1, \ldots, F_n\}$ and $\beta \in \mathbb{K}^n$. For $\gamma \in \mathbb{K}^*$, we will say that the graded quotient $\varphi_{V,\gamma\cdot\beta}$ is the γ-twist of the graded quotient $\varphi_{V,\beta}$. ◊

We will often consider a tuples (α, β, γ) where β varies in an open subset depending on α. Furthermore, γ varies in an open subset depending on β. We note the following elementary result that will help us formalize this notion of relative generality.

Proposition 3.9.8. *Fix integers $d_1, \ldots, d_n \geq 1$. Let $\mathcal{V} \subseteq \mathbb{K}^{n+1}$ be a non-empty open subset. Let $f : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^{n+1}$ be the map $f(\gamma, \beta) = (\gamma, \gamma^{d_1}\beta_1, \ldots, \gamma^{d_n}\beta_n)$. The we have the following.*

1. *The inverse image $f^{-1}(\mathcal{V})$ is an open subset of $\mathbb{K}^{n+1}$.*
2. *The image $p_2(f^{-1}(\mathcal{V})) \subseteq \mathbb{K}^n$ is a non-empty open subset, where p_2 is the projection map $\mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$.*
3. *For each $\beta \in \mathcal{U}_2 := p_2(f^{-1}(\mathcal{V}))$, there exists a non-empty open subset $\mathcal{T}_\beta \subseteq \mathbb{K}$, such that $f(\gamma, \beta) \in \mathcal{V}$ for all $\gamma \in \mathcal{T}_\beta$.*

Proof. We know that f is a continuous map in Zariski topology. Therefore $f^{-1}(\mathcal{V})$ is an open subset. We have the second statement since projection maps are open maps with respect to Zariski topology. Let $\mathcal{U}_2 := p_2(f^{-1}(\mathcal{V}))$. For each $\beta \in \mathcal{U}_2$, we define $\mathcal{T}_\beta := f^{-1}(\mathcal{V}) \cap p_2^{-1}(\beta)$. Note that the fibre $p_2^{-1}(\beta)$ is isomorphic to the affine space $\mathbb{K}$. Since $\beta \in p_2(f^{-1}(\mathcal{V}))$, we have that $\mathcal{T}_\beta$ is non-empty. Furthermore, $\mathcal{T}_\beta$ is open by the openness of $f^{-1}(\mathcal{V})$. □

Proposition 3.9.9. *Let $R = S/(U)$ and $V, Y \subseteq R$ be finite dimensional $\mathbb{K}$-vector spaces such that $V \subseteq Y$. Let $\alpha \in \mathbb{K}^{\dim(V)}$ and $\beta \in \mathbb{K}^{\dim(Y)-\dim(V)}$. Let $\varphi_{V,\alpha}$ and $\varphi_{Y,(\alpha,\beta)}$ be compatible graded quotients with compatible bases $\mathcal{F}_V \subseteq \mathcal{F}_Y$. Let $Y' = \operatorname{span}_{\mathbb{K}}\{\mathcal{F}_Y \setminus \mathcal{F}_V\}$. We denote $Y'_\alpha = \varphi_{V,\alpha}(Y') \subseteq R'_\alpha$, where $R'_\alpha := R[y]/I_{V,\alpha}$. Let $Y_\alpha = \operatorname{span}_{\mathbb{K}}\{y_1, Y'_\alpha\}$. Then the following holds.*

1. *Let $B : \mathbb{N}^d \to \mathbb{N}^d$ be such that $B_k(\delta + \mu) \geq B_k(\delta) + \|\mu\|_1$ for all $\delta, \mu \in \mathbb{N}^d$. If Y is B-lifted strong, then Y'_α is also B-lifted strong.*

2. *Let $\gamma \in \mathbb{K}^*$. For a general α, we have a commutative diagram of $\mathbb{K}$-algebra homomorphisms.*

$$\begin{array}{ccccc} R[y_1] & \xrightarrow{\varphi_{V,\alpha}} & R'_\alpha & \xrightarrow{\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}} & R'_\alpha[y_2]/I_{Y_\alpha,\gamma\cdot(1,\beta)} \\ \downarrow{\scriptstyle id} & & & & \downarrow{\scriptstyle \simeq} \\ R[y_1] & & \xrightarrow{\qquad\qquad\varphi_{Y,(\alpha,\beta)}\qquad\qquad} & & R[y_1]/I_{Y,(\alpha,\beta)} \end{array}$$

where the map $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}$ sends $y_1 \mapsto \gamma y_2$ and on Y'_α it is same as the twisted quotient $\varphi_{Y'_\alpha,\gamma\cdot\beta}$. Moreover, the vertical isomorphism sends $y_2 \mapsto \frac{1}{\gamma} y_1$.

Proof. (1) If $Y = \oplus_{k=1}^d Y_k$ is B-lifted strong in R, then we have $\widetilde{s}(U, Y_k) \geq B_k(\delta_U + \delta_Y)$ for any graded piece Y_k. Let $Q_1, \ldots, Q_r \in S$ lifts of a basis of Y_k. Now $\delta_Y = \delta_V + \delta_{Y'}$ and hence $\widetilde{s}(U, Y_k) \geq B_k(\delta_U + \delta_{Y'}) + \|\delta_V\|_1$. Let $H_1, \ldots, H_t \in S$ be lifts of a basis of a graded piece $(Y'_\alpha)_k \subseteq R'_\alpha$. Then, for any $F \in U_k + \operatorname{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$, we may write $F = P + Q + T$, where $P \in U_k$, $Q \in \operatorname{span}_{\mathbb{K}}\{Q_1, \ldots, Q_r\}$ and $T \in I_{V,\alpha}$. Therefore,

$$s(F) \geq s(P + Q) - s(T) \geq \widetilde{s}(U, Y_k) - \|\delta_V\|_1 \geq B_k(\delta_U + \delta_{Y'}) \geq B_k(\delta_U + \delta_{Y'_\alpha})$$

as $s(T) \leq \|\delta_V\|_1$ and $\delta_Y \geq \delta_{Y'_k}$. Since $F \in U_k + \operatorname{span}_{\mathbb{K}}\{H_1, \ldots, H_t\}$ was arbitrary, we get that $\widetilde{s}(U, (Y'_\alpha)_k) \geq B_k(\delta_U + \delta_{Y'_\alpha})$, and hence Y'_α is B-lifted strong.

(2) Let $\mathcal{F}_V = \{F_1, \ldots, F_m\}$ and $\mathcal{F}_Y = \{F_1, \ldots, F_m, G_1, \ldots, G_n\}$. For a general α, we have that the images of $G_1, \ldots, G_n$ in R'_α are linearly independent, and hence they are a basis of Y'_α. Then composition of the first two horizontal maps sends $F_i \mapsto \alpha_i(\gamma y_2)^{\deg(F_i)}$ and $G_j \mapsto \beta_j(\gamma y_2)^{\deg(G_j)}$. Therefore, we have an isomorphism $R'_\alpha[y_2]/I_{Y_\alpha,\gamma\cdot(1,\beta)} \simeq R[y_1, y_2]/I$, where I is generated by the forms $y_1 - \gamma y_2, F_i - \alpha_i y_1^{\deg(F_i)}, G_j - \beta_j(\gamma y_2)^{\deg(G_j)}$ with all $i \in [m], j \in [n]$. Now we have an isomorphism $R[y_1, y_2]/(y_1 - \gamma y_2) \xrightarrow{\sim} R[y_1]$ given by $y_2 \mapsto \frac{1}{\gamma} y_1$. This isomorphism descends to an isomorphism $R[y_1, y_2]/I \simeq R[y_1]/I_{Y,(\alpha,\beta)}$. By composing these two isomorphisms, we get the desired vertical isomorphism $R'_\alpha[y_2]/I_{Y_\alpha,\gamma\cdot(1,\beta)} \xrightarrow{\sim} R[y_1]/I_{Y,(\alpha,\beta)}$. □

We can use these compatible quotients to show the following extension of Lemma 3.9.4 that was discussed in the previous section.

Lemma 3.9.10. *Let $\mu : \mathbb{N}^e \to \mathbb{N}^e$ be an ascending function such that $\mu_i(\delta) \geq A(\eta, i) + 3\|\delta\|_1$, where A is the function defined in [AH20b, Theorem A]. Let $U \subseteq S_{\leq e}$ be a graded finitely generated vector*

space such that $R := S/(U)$ is a UFD, and $z \in R_1$. Let V be a $H(\mu, 1, (\Pi_2(d)+1)\cdot r)$-lifted strong vector space, where H is the function defined in Section 3.6.1 Let $\mathcal{G}$ be a finite set non-associate irreducible homogeneous forms in R_d such that every $G \in \mathcal{G}$ is absolutely irreducible with respect to V.

Suppose that there exists a dense subset $\mathcal{Z} \subseteq \mathbb{K}^{\dim(V)}$ such that for all $\alpha \in \mathcal{Z}$, the following holds: for any subset $\mathcal{S} \subset \mathcal{G}$ with $|\mathcal{S}| > b$ we have $s(\varphi_{V,\alpha}(\mathcal{S})) \leq 1$. Then there exists a μ-lifted strong vector space W obtained by applying a $(1, (\Pi_2(d)+1)\cdot b)$-process to V such that every form in $\mathcal{G}$ is absolutely reducible over W or in (W).

Proof. By Proposition 3.6.16, we can further replace $\mathcal{Z}$ by a possibly a possible smaller dense subset, and assume that $\varphi_{V,\alpha}(\mathcal{G})$ consists of pairwise non-associate forms for all $\alpha \in \mathcal{Z}$. For each $\alpha \in \mathcal{Z}$, we can apply Lemma 3.9.4 to the set $\varphi_{V,\alpha}(\mathcal{G}) \subseteq R[y_1]/I_{V,\alpha}$ and obtain a subset $T_\alpha \subseteq \varphi_{V,\alpha}(\mathcal{G})$ with properties (1) and (2) from Lemma 3.9.4. For each $\alpha \in \mathcal{Z}$ let us choose $T_\alpha \subseteq \mathcal{G}$ which a lift of T_α. Note that there are finitely many choices for T_α. Therefore we may replace $\mathcal{Z}$ by a possibly smaller dense subset and assume that there exists a fixed set $T \subseteq \mathcal{G}$ such that $T = T_\alpha$ for all $\alpha \in \mathcal{Z}$. In particular, the following holds:

1. $|T| \leq (\Pi_2(d)+1)\cdot r$, and
2. For all $\alpha \in \mathcal{Z}$ and for any $Y_\alpha \subseteq R[y_1]/I_{V,\alpha}$ which is $h_{2\mu} \circ t_1$-lifted strong vector space with $\varphi_{V,\alpha}(T) \subseteq \mathbb{K}[Y]$, every form in $\varphi_{V,\alpha}(\mathcal{G})$ is absolutely reducible over Y or in the ideal (Y_α).

We let $W := \mathrm{AH}(V, T)$. We now show that every form in $\mathcal{G}$ is either in the ideal W or absolutely reducible with respect to W. Note that $V \subseteq W$ by Corollary 3.6.6. Let $W = V + W'$, where W' is obtained by extending a basis of V to a basis of W. Then $\varphi_{V,\alpha}(W')$ is a $h_{2\mu} \circ t_2$-lifted strong vector space in $R[y]/I_\alpha$ by part (1) of Proposition 3.9.9. Let $W'_\alpha = \varphi_{V,\alpha}(W')$. We also have $\varphi_{V,\alpha}(T) \subseteq \mathbb{K}[y_1, W'_\alpha]$ for all $\alpha \in \mathcal{Z}$. In particular, by taking $W_\alpha := \mathrm{span}_{\mathbb{K}}\{y_1, W'_\alpha\}$, we know that every form in $\varphi_{V,\alpha}(\mathcal{G})$ is absolutely reducible over W_α or in (W_α). By Proposition 3.6.16, we can replace $\mathcal{Z}$ with a possibly smaller dense subset, and assume that $\dim(W') = \dim(W'_\alpha)$ for all $\alpha \in \mathcal{Z}$.

Let $F \in \mathcal{G}$. For $\beta \in \mathbb{K}^{\dim(W')}$ and $\gamma \in \mathbb{K}$, consider the map $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}$ as in Proposition 3.9.9. Let us denote $R'_\alpha = R[y]/I_{V,\alpha}$. Let $G := \varphi_{V,\alpha}(F)$. We consider two situations below.

Case 1. Suppose that $G \in (W_\alpha)$ in $R[y_1]/I_{V,\alpha}$. Then we may compose with map $\varphi_{W_\alpha,\gamma\cdot(1,\beta)}$ as in Proposition 3.9.8. Under this composition $\varphi_{W_\alpha,\gamma\cdot(1,\beta)} \circ \varphi_{V,\alpha}$, we see that F mapped to a form in (y_2). By the isomorphism in Proposition 3.9.9, we conclude that $\varphi_{W,(\alpha,\beta)}$ maps F to a form in (y_1). Since α is from a dense set, and β can also be chosen to be general, we have $F \in (W)$ in R by part (4) of Proposition 3.7.14.

Case 2. Now suppose that we have $G \notin (W_\alpha)$ and G is absolutely reducible over W_α for all $\alpha \in \mathcal{Z}$. By Proposition 3.7.14 there exists a non-empty open subset $\mathcal{V} \subseteq \mathbb{K}^{n+1}$ such that $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(G)$ is reducible for all $\gamma \cdot (1, \beta) \in \mathcal{V}$. We let $\mathcal{U}_\alpha$ be the image of $f^{-1}(\mathcal{V})$ under the projection map $p_2 : \mathbb{K} \times \mathbb{K}^{\dim(W')} \to \mathbb{K}^{\dim(W')}$, where $f(\gamma, \beta) = \gamma \cdot (1, \beta)$ as in Proposition 3.9.8.

Note that $\varphi_{W,(\alpha,\beta)} = \varphi_{Y_\alpha,\gamma\cdot(1,\beta)} \circ \varphi_{V,\alpha}$ by Proposition 3.9.9. Let $\mathcal{Z}' := \{(\alpha, \beta) \mid \alpha \in \mathcal{Z}, \beta \in \mathcal{U}_\alpha\}$. Then $\mathcal{Z}' \subseteq \mathbb{K}^{\dim(W)}$ is a dense set. Now for all $(\alpha, \beta) \in \mathcal{Z}'$, we have $\varphi_{W,(\alpha,\beta)}(F)$ is reducible in $R[y_1]/I_{W,(\alpha,\beta)}$, by the isomorphism in Proposition 3.9.9. As $F \notin (W)$, we conclude that F is absolutely reducible over W by Proposition 3.7.14. Since $F \in \mathcal{G}$ was arbitrary, we conclude that every form in $\mathcal{G}$ is either in the ideal (W) or absolutely reducible over it. The claimed bound on the strength of W follows from Lemma 3.6.10. □

These compatible quotients allow us to better argue about the types of potential functions introduced when bounding the dimension of product Sylvester–Gallai configurations. We show this in the next section.

3.9.2 A potential function

In this section, we define a potential function that will be useful in controlling EK configurations. Informally, the potential function measures how close a form is to a given vector space. It will be convenient to define the potential so that the closer the form is to the space, the higher the potential is. Our main arguments will therefore involve steps that increase the potential by a sufficient amount.

Let $U \subseteq S$ be a graded finitely generated vector space such that $R := S/(U)$ is a CM UFD. Let $V \subseteq R$ be a $h_{2B} \circ t_1$-lifted strong vector space and $\alpha \in \mathbb{K}^{\dim(V)}$. The quotient map $\varphi_{V,\alpha} : R[y] \to R[y]/I_{V,\alpha}$ is defined in Section 3.6.2.

Definition 3.9.11. For a form $F \in R$, let $\varphi_{V,\alpha}(F) = y^r \prod_{i=1}^{s} F_i^{r_i}$ be an irreducible factorization. We define the potential of F with respect to V and α as $\psi_{V,\alpha}(F) := 2r + s$. ◊

We remark that this potential measures the same thing as the potential defined in Lemma 3.8.10: it is an increasing function of the number of factors and the power of z under the graded quotient, with the power of z having a larger weight. We could have used the potential from Lemma 3.8.10 here instead and vice versa. The advantage of this potential is that it is linear, which means we can easily write down the changes in potential.

Proposition 3.9.12. *For any two forms $F, G \in R$. Suppose that $\varphi_{V,\alpha}(F), \varphi_{V,\alpha}(G)$ have c number of common irreducible factors other than y. Then*

$$\Psi_{V,\alpha}(FG) = \Psi_{V,\alpha}(F) + \Psi_{V,\alpha}(G) - c.$$

In particular,

$$\Psi_{V,\alpha}(FG) \leq \Psi_{V,\alpha}(F) + \Psi_{V,\alpha}(G).$$

Moreover, equality occurs iff $\gcd(F, G) = y^t$ for some $t \geq 0$.

Proof. Suppose we have the irreducible factorisations $\varphi_{V,\alpha}(F) = y^a \prod_{i=1}^{s} F_i^{d_i}$ and $\varphi_{V,\alpha}(G) = y^b \prod_{i=1}^{t} G_i^{e_i}$. Since $\varphi_{V,\alpha}(FG) = \varphi_{V,\alpha}(F)\varphi_{V,\alpha}(G)$, we have

$$\Psi_{V,\alpha}(FG) = 2(a+b) + s + t - c \leq 2a + s + 2b + t = \Psi_{V,\alpha}(F) + \Psi_{V,\alpha}(G).$$

Moreover, we have equality iff $c = 0$. □

Group potential. For a finite set of forms $\mathcal{G} \subseteq R$, we define the potential of $\mathcal{G}$ with respect to the pair (V, α) as

$$\Phi_{V,\alpha}(\mathcal{G}) = \sum_{G \in \mathcal{G}} \psi_{V,\alpha}(G).$$

Remark 3.9.13. Note that if $\mathcal{G} \subseteq R_{\leq d}$, then $\Phi_{V,\alpha}(\mathcal{B}) \leq 2d|\mathcal{B}|$ for any (V, α). ◊

Potential increasing forms. Let $V \subseteq Y$ be both $h_{2B} \circ t_1$-lifted strong vector spaces in R. Let $\alpha \in \mathbb{K}^{\dim(V)}$ and $\beta \in \mathbb{K}^{\dim(Y)-\dim(V)}$. We say that $F \in \mathcal{G}$ is a *potential increasing* form with respect to the pairs (V, α) and $(Y, (\alpha, \beta))$ if

$$\Psi_{Y,(\alpha,\beta)}(F) > \Psi_{V,\alpha}(F).$$

The notion of potential increasing captures the property that a form factors more when quotiented with respect to Y than V. First we note that for general quotients potential does not decrease.

Proposition 3.9.14 (Non-decreasing property)**.** *Let $V \subseteq Y$ be both $h_{2B} \circ t_1$-lifted strong vector spaces in R of dimensions m and $m+n$ respectively. Let $\varphi_{V,\alpha}$ and $\varphi_{Y,(\alpha,\beta)}$ be compatible graded quotients with respect to bases $\mathcal{F}_V \subseteq \mathcal{F}_Y$. Let $Y' = \operatorname{span}_{\mathbb{K}}\{\mathcal{F}_Y \setminus \mathcal{F}_V\}$, $R'_\alpha = R[y_1]/I_{V,\alpha}$, $Y'_\alpha = \varphi_{V,\alpha}(Y')$ and $Y_\alpha = \operatorname{span}_{\mathbb{K}}\{y_1, Y'_\alpha\}$.*

Fix a square-free form $F \in R$. For a general $\alpha \in \mathbb{K}^m$, there exists a non-empty open subset $\mathcal{U} \subseteq \mathbb{K}^n$ (depending on α) such that for all $\beta \in \mathcal{U}$ the following holds.

1. *We have $\Psi_{Y,(\alpha,\beta)}(F) \geq \Psi_{V,\alpha}(F)$.*
2. *Moreover, $\Psi_{Y,(\alpha,\beta)}(F) = \Psi_{V,\alpha}(F)$ iff one of the following holds*
   (a) *$F \in \mathbb{K}[V]$.*
   (b) *Let F_i be an irreducible factor of $\varphi_{V,\alpha}(F)$ such that $F_i \notin (y_1)$. Then its image $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_i)$ is irreducible and $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_i) \notin (y_2)$ for general $\gamma \in \mathbb{K}$.*

Proof. Note that the quotient map $\varphi_{Y,(\alpha,\beta)} : R \to R[y_2]/I_{Y,(\alpha,\beta)}$ factors, up to isomorphism, as $\varphi_{Y,(\alpha,\beta)} = \varphi_{Y_\alpha,\gamma\cdot(1,\beta)} \circ \varphi_{V,\alpha}$ by Proposition 3.9.9. Suppose we have the irreducible factorization

$$\varphi_{V,\alpha}(F) = y_1^r \prod_{i=1}^{s} F_i^{r_i}.$$

Then we have

$$\varphi_{Y,(\alpha,\beta)}(F) = y_1^r \prod_{i=1}^{s} \varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_i)^{r_i}.$$

If $s = 0$, then $\varphi_{V,\alpha}(F)$ is a power of y_1 up to scalar multiples. By Proposition 3.7.14, we have $V \in \mathbb{K}[V]$. In particular, $\Psi_{Y,(\alpha,\beta)}(F) = \Psi_{V,\alpha}(F) = 2\deg(F)$.

Suppose that $s > 0$. By Proposition 3.6.16, there exists a non-empty open subset $\mathcal{V} \subseteq \mathbb{K}^{n+1}$ such that for all $\gamma \cdot (1,\beta) \in \mathcal{V}$, the forms $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_i)$ and $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_j)$ do not have any common factors other than y_2 for $i \neq j$. We let $\mathcal{U}$ be the image of $f^{-1}(\mathcal{V})$ under the projection $\mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$, where $f(\gamma,\beta) = \gamma \cdot (1,\beta)$ as in Proposition 3.9.8. Let $G_i = \varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_i)$. We have an irreducible factorization $G_i = y_2^{a_i} \prod_{i=1}^{s_i} H_i$. Note that since G_i is irreducible, the forms H_i appear with multiplicity 1 by Proposition 3.6.16. Now, using the isomorphism in Proposition 3.9.8, we have

$$\Psi_{Y,(\alpha,\beta)}(F) = 2r + \sum_{i=1}^{s} 2r_i a_i + \sum_{i=1}^{s} s_i \geq 2r + s = \Psi_{V,\alpha}(F).$$

Moreover, equality occurs iff $a_i = 0$ and $s_i = 1$ for all $i \in [s]$. In other words G_i is irreducible and G_i not a scalar multiple of y_2. □

We note the following criterion for when the potential strictly increases.

Lemma 3.9.15 (Potential increasing criterion)**.** *Let $V \subseteq Y$ be both $h_{2B} \circ t_1$-lifted strong vector spaces in R of dimensions m and $m+n$ respectively. Let $\varphi_{V,\alpha}$ and $\varphi_{Y,(\alpha,\beta)}$ be compatible graded quotients with respect to bases $\mathcal{F}_V \subseteq \mathcal{F}_Y$. Let $Y' = \operatorname{span}_{\mathbb{K}}\{\mathcal{F}_Y \setminus \mathcal{F}_V\}$, $R'_\alpha = R[y_1]/I_{V,\alpha}$, $Y'_\alpha = \varphi_{V,\alpha}(Y')$ and $Y_\alpha = \operatorname{span}_{\mathbb{K}}\{y_1, Y'_\alpha\}$.*

Let $F \in R$ be a square-free form. Let $\alpha \in \mathbb{K}^m$ be a general element. Suppose that there exists a factor F_k of $\varphi_{V,\alpha}(F)$ in R'_α, such that

1. *F_k is not a scalar multiple of y_1, and*

2. F_k is absolutely reducible with respect to Y_α or F_k is in the ideal (Y_α) in R'_α.

Then there exists a non-empty open subset $\mathcal{U} \subseteq \mathbb{K}^n$ (depending on α) such that for all $\beta \in \mathcal{U}$, the form F is potential increasing with respect to the pairs (V, α) and $(Y, (\alpha, \beta))$.

Proof. Suppose that we have factorization $\varphi_{V,\alpha}(F) = y_1^r \prod_{i=1}^s F_i^{r_i}$ in R'_α. Note that the quotient map $\varphi_{Y,(\alpha,\beta)} : R \to R[y_2]/I_{Y,(\alpha,\beta)}$ factors as $\varphi_{Y,(\alpha,\beta)} = \varphi_{Y_\alpha,\gamma\cdot(1,\beta)} \circ \varphi_{V,\alpha}$ by Proposition 3.9.9. Let $\mathcal{U} \subseteq \mathbb{K}^n$ be the non-empty open subset provided by *Proposition* 3.9.14.

Let F_k be a factor of $\varphi_{V,\alpha}(F)$ with properties (1) and (2) in the hypothesis. In particular, F_k is not a scalar multiple of y_1. Therefore $F \notin \mathbb{K}[V]$ by Proposition 3.7.14.

Case 1. Suppose that $F_k \in (Y_\alpha)$. Then for any γ, β, the form $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_k)$ is divisible by y_2. Therefore $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_k)$ is reducible if $\deg(\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_k)) > 1$. If $\deg(\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_k)) = 1$, then it is a scalar multiple of y_2.

Case 2. Now suppose that F_k is absolutely reducible with respect to Y_α. Then by Proposition 3.7.14, there exists a non-empty open subset $\mathcal{V} \subseteq \mathbb{K}^{n+1}$ such that $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_k)$ is a reducible form for $\gamma \cdot (1, \beta) \in \mathbb{K}^{n+1}$. We let $\mathcal{U}_2$ to be the image of $f^{-1}(\mathcal{V})$ under the projection $\mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$, where $f(\gamma, \beta) = \gamma \cdot (1, \beta)$ as in Proposition 3.9.8.

Therefore, by combining both cases above, there exists a non-empty open subset $\mathcal{U}' := \mathcal{U}_2 \cap \mathcal{U}$ such that for all $\beta \in \mathcal{U}'$, the form $\varphi_{Y_\alpha,\gamma\cdot(1,\beta)}(F_k)$ is either reducible or it is a scalar multiple of y_2, for a general γ depending on β. Then Proposition 3.9.14 implies that we must have a strict inequality $\Psi_{Y,(\alpha,\beta)}(F) > \Psi_{V,\alpha}(F)$ for all $\beta \in \mathcal{U}'$. □

With all these tools at hand, we can prove our dimension bounds.

3.9.3 Controlling highest degree forms

As in the case of product Sylvester–Gallai configurations, we will start by controlling the highest degree forms in our configuration. For the rest of this section, fix positive integer $\eta \geq 3$ and set $\varepsilon := 1/16$. We will use the auxiliary functions H, D defined in Section 3.6.1

We begin by establishing some notation specific to this subsection. Fix an integer e. Let $\mu : \mathbb{N}^e \to \mathbb{N}^e$ be an ascending function such that $\mu_i(\delta) \geq A(\eta, i) + 3\|\delta\|_1$, where A is the function defined in [AH20b, Theorem A]. Let $U \subseteq S_{\leq e}$ be a graded finitely generated vector space such that $R := S/(U)$ is a UFD, and $z \in R_1$. Let $\mathcal{A}, \mathcal{B}, \mathcal{C} \subset R_{\leq d}$ be sets that form a (d, z, R)-EK configuration (Definition 3.9.1) for $2 \leq d \leq e$. Let $\mathcal{F} := \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$.

Suppose $V \subseteq R_{\leq d}$ is a $h_{2\mu} \circ t_1$-lifted strong vector space. Let $\mathcal{A}_d^V$ denote the set of forms in $\mathcal{A}_d$ which are not in (V) and are absolutely irreducible with respect to V.

$$\mathcal{A}_d^V := \{F \in \mathcal{A}_d \mid F \notin (V) \text{ and } F \text{ absolutely irreducible with respect to } V\}.$$

We denote $m_a^V := |\mathcal{A}_d^V|$. We use M_a^V to denote $|\mathcal{A}| - m_a^V$, this is the number of forms in $\mathcal{A}$ that either have degree less than d, or are absolutely reducible over V, or are in the ideal generated by V.

In particular, if we take $V = (0)$, then $m_a^{(0)}, M_a^{(0)}$ are the sizes of the degree d and degree less than d pieces of $\mathcal{A}$ respectively. We similarly define m_b, M_b, m_c, M_c. We will assume that $|\mathcal{A}| \geq |\mathcal{B}| \geq |\mathcal{C}|$.

Our goal in this section is to construct a vector space W such that all the degree d forms in $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are absolutely reducible over W or in the ideal (W), i.e. $m_a^W = m_b^W = m_c^W = 0$. First, we show that if the algebra $\mathbb{K}[V]$ contains a sufficient fraction of forms, then we may find W such that at least one of m_a^W, m_b^W, m_c^W is 0.

Lemma 3.9.16. *Suppose $|\mathcal{C}| \geq 2\Pi_2(d)\varepsilon^{-1}$. Further suppose V is a $H(\mu, 1, 2 \cdot d \cdot \varepsilon^{-1})$-lifted strong vector space such that one of the following holds.*

- $|\mathbb{K}[V] \cap \mathcal{A}| \geq \varepsilon|\mathcal{A}|$ *or*
- $|\mathbb{K}[V] \cap \mathcal{B}| \geq \varepsilon|\mathcal{B}|$ *or*
- $|\mathcal{C}| \geq |\mathcal{B}|/2$ *and* $|\mathbb{K}[V] \cap \mathcal{C}| \geq \varepsilon|\mathcal{C}|$.

Then there is a μ-lifted strong vector space W obtained by applying a $(1, 4 \cdot d \cdot \varepsilon^{-1})$ process to V such that $m_b^W \cdot m_b^W \cdot m_c^W = 0$, that is, at least one of the sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ have no degree d forms absolutely irreducible over W.

Proof. We first assume $|\mathbb{K}[V] \cap \mathcal{A}| \geq \varepsilon|\mathcal{A}|$. Let $A_1, \ldots, A_r$ be the elements in $|\mathbb{K}[V] \cap \mathcal{A}|$. In this case a $(1, 2 \cdot d \cdot \varepsilon^{-1})$ process will suffice.

We now describe one iteration of our $(1, 2 \cdot d \cdot \varepsilon^{-1})$ process. Let $B \in \mathcal{B}_d$ be a form that is absolutely irreducible over V. If no such B exists, then every form in $\mathcal{B}_d$ is absolutely reducible over V, which implies $m_b^V = 0$, and we terminate the process. Since B is absolutely irreducible over V, there are at most $\Pi_1(d)$ elements among $A_1, \ldots, A_r$ such that (B, A_i) is not prime. Reordering, we can assume that (A_i, B) is prime for $1 \leq i \leq s$ with $s = r - \Pi_1(d)$. Since $r \geq \varepsilon|\mathcal{A}|$ and $|\mathcal{A}| \geq |\mathcal{C}| \geq 2\Pi_1(d)\varepsilon^{-1}$, we have $s \geq \varepsilon|\mathcal{A}|/2$.

Note that $z \notin (A_i, B)$ as $\deg(B) = d > 1$ and z, A_i are non-associate forms. Hence, by the EK condition, there exists $C_i \in (A_i, B) \cap \mathcal{C}$ for $1 \leq i \leq s$, as (A_i, B) is prime. Since $\deg B = d$, we have $\deg C_i = d$, and we can write $C_i = \alpha_i B + H_i A_i$. If $C_i \in (V)$ then $B \in (V)$, contradicting assumption. If $C_i = C_j$, then $\alpha_j B + H_j A_j = \alpha_i B + H_i A_i$. If we also have $\alpha_i \neq \alpha_j$ then again $B \in (V)$, a contradiction. Therefore, if $C_i = C_j$ then $\alpha_i = \alpha_j$, and further A_i, A_j are distinct factors of $C_i - \alpha_i B = C_j - \alpha_j B$. Since $\deg C_j - \alpha_j B = d$, and since these factors are distinct, there are at most d indices such that $C_{i_1} = \cdots = C_{i_a}$. In conclusion, the above paragraph shows that the set of forms $C_1, \ldots, C_s$ consists of at least s/d distinct forms, none of which are in the ideal (V).

We now update V to $AH_R(V, B)$. After this update, m_c^V reduces by at least s/d, which is at least $\varepsilon|\mathcal{A}|/2d$. Note that we still have $|\mathbb{K}[V] \cap A| \geq \varepsilon|A|$. Therefore if $m_b^V \neq 0$, we may apply the argument above again and reduce m_c^V by s/d. Since m_c^V is at most $|\mathcal{A}|$ to start with, after $2 \cdot d \cdot \varepsilon^{-1}$ complete iterations of the above, we have a vector space W such that $m_b^W = 0$ or $m_c^W = 0$. The claimed bound on the lifted strength of W follows from Lemma 3.6.10.

In the case when $|\mathbb{K}[V] \cap \mathcal{B}| \geq \varepsilon|\mathcal{B}|$ the exact same argument applies with every occurrence of $\mathcal{A}$ and $\mathcal{B}$ swapped. In the last case, we have $|\mathbb{K}[V] \cap \mathcal{C}| \geq \varepsilon|\mathcal{C}|$ and $|\mathcal{C}| \geq |\mathcal{B}|/2$. If we just swap the roles of $\mathcal{A}$ and $\mathcal{C}$ in the above argument, then essentially in each iteration we are reducing m_a^V by $\varepsilon|\mathcal{C}|/2$. However, this does not suffice, because m_a^V might be significantly bigger than $\mathcal{C}$, and therefore we cannot guarantee that the iteration terminates after constantly many steps. Therefore, instead of just swapping the roles of $\mathcal{A}$ and $\mathcal{C}$, we replace the role of $\mathcal{A}$ with $\mathcal{C}$ and the role of $\mathcal{B}$ with $\mathcal{A}$. Now in each iteration we are reducing m_b^V by at least $\varepsilon|\mathcal{C}|/2$. The further assumption that $|\mathcal{C}| \geq |\mathcal{B}|/2$ implies that we are reducing m_b^V by at least $\varepsilon|\mathcal{B}|/4$. Therefore the above iteration terminates after $4 \cdot d \cdot \varepsilon^{-1}$ steps. □

We now define the notion of unbreakable pencils.

Definition 3.9.17. Let $F_1, \cdots, F_a \in R$ be homogeneous of degree d. We will say that $F_1, \cdots, F_a$ *generate unbreakable pencils* if they are linearly independent and $s(F_i, F_j) > 1$ for all $i, j \in [a]$. Alternatively, given a set of degree d forms $\mathcal{S} \subseteq R_d$, we will say that the set $\mathcal{S}$ is *unbreakable pencil generating* if $\mathcal{S}$ is a linearly independent set and $s(F, G) > 1$ for any $F, G \in \mathcal{S}$. ◊

Note that if $F_1, \cdots, F_a$ generate unbreakable pencils, then F_i is irreducible for all $i \in [a]$. Recall that for $\alpha \in \mathbb{K}^{\dim(V)}$, we denote by $\varphi_{V,\alpha}$ the graded quotient homomorphism $R[y] \to R[y]/I_\alpha$ as

defined in Section 3.6.2. We have the following corollary of Lemma 3.9.10, which follows just by the fact that any set $\mathcal{S}$ that satisfies $s(\mathcal{S}) \geq 1$ is also an unbreakable pencil.

Lemma 3.9.18. *Fix an integer $r > 0$. Suppose V is $H(\mu, 1, (\Pi_2(d)+1)\cdot r)$-lifted strong. Suppose that there exists a dense subset $\mathcal{Z} \subseteq \mathbb{K}^{\dim(V)}$ such that for all $\alpha \in \mathcal{Z}$ we have*

$$r > \max\{|\mathcal{S}| \mid \mathcal{S} \subseteq \mathcal{A}_d^V \text{ and } \varphi_{V,\alpha}(\mathcal{S}) \text{ is an unbreakable pencil generating set}\}.$$

Then there exists a μ-lifted strong vector space W obtained by applying a $(1, (\Pi_2(d)+1)\cdot r)$-process to V such that every form in $\mathcal{A}_d^V$ is absolutely reducible over W or in (W), that is, we have $m_a^W = 0$.

Lemma 3.9.19. *Suppose $|\mathcal{C}| \geq 2\Pi_1(d)\varepsilon^{-1}$. Let $k := 10 \cdot d \cdot (\Pi_2(d) + \Pi_1(d)) \cdot (\Pi_2(d)+1)$. Let V be a $H(\mu, k, 8\cdot d\cdot \varepsilon^{-1})$-lifted strong vector space such that one of the following holds.*

- *$M_a^V \geq 6\varepsilon|\mathcal{A}|$ or*
- *$M_b^V \geq 6\varepsilon|\mathcal{B}|$ or*
- *$M_c^V \geq 6\varepsilon|\mathcal{C}|$ and $|\mathcal{C}| \geq |\mathcal{B}|/2$.*

Then there is a μ-lifted strong vector space W obtained by applying a $(k, 8\cdot d\cdot\varepsilon^{-1})$ process to V such that $m_a^W \cdot m_b^W \cdot m_c^W = 0$.

Proof. We first assume $M_b^V \geq 6\varepsilon|\mathcal{B}|$. In this case a $(k, 4\cdot d\cdot\varepsilon^{-1})$ process will suffice. Recall that for any vector space W, the set $\mathcal{A}_d^W$ is the set of forms in $\mathcal{A}_d$ which are not in (W) and are absolutely irreducible with respect to W. Also, we have $m_a^W := |\mathcal{A}_d^W|$ and $M_a^W := |\mathcal{A}| - m_a^W$.

Let $\Pi_3(d) := \Pi_2(d) + \Pi_1(d)$. Let $a := 10 \cdot d \cdot \Pi_3(d)$. We will describe one iteration of the desired $(k, 4\cdot d\cdot\varepsilon^{-1})$ process below and we will show that at most $4\cdot d\cdot\varepsilon^{-1}$ iterations suffice.

First let us note two special cases which yield a vector space W with $m_a^W \cdot m_b^W \cdot m_c^W = 0$. If we reach any of these two *breaking conditions* during the iterative process, we will terminate the process. The first condition will involve properties of $\mathcal{B}$ and the second will involve $\mathcal{A}$.

Breaking condition 1. Suppose that $|\mathcal{B} \cap \mathbb{K}[V]| \geq \varepsilon|\mathcal{B}|$. Then we apply Lemma 3.9.16 to V, and call the resulting space W. We terminate the iterative process.

Breaking condition 2. Suppose that there exists a dense subset $\mathcal{Z} \subseteq \mathbb{K}^{\dim(V)}$ such that for all $\alpha \in \mathcal{Z}$ we have

$$a > \max\{|\mathcal{S}| \mid \mathcal{S} \subseteq \mathcal{A}_d^V \text{ and } \varphi_{V,\alpha}(\mathcal{S}) \text{ is an unbreakable pencil generating set}\}.$$

Here $\varphi_{V,\alpha} : R[y_1] \to R[y_1]/I_{V,\alpha}$ is a graded quotient. Then we apply Lemma 3.9.18 and let W be the resulting vector space. We terminate the iterative process.

Iterative step. Suppose that none of the breaking conditions occur. Then we have $|\mathcal{B} \cap \mathbb{K}[V]| < \varepsilon|\mathcal{B}|$. Moreover, there exists a non-empty open subset $\mathcal{U} \subseteq \mathbb{K}^{\dim(V)}$, such that for all $\alpha \in \mathcal{U}$ there exists a set $\mathcal{S}_\alpha \subseteq \mathcal{A}_d^V$ where $|\mathcal{S}_\alpha| = a$ and $\varphi_{V,\alpha}(\mathcal{S}_\alpha)$ is an unbreakable pencil generating set. Since there are finitely many possibilities for such $\mathcal{S}_\alpha$, we may replace $\mathcal{U}$ by a dense subset $\mathcal{Z} \subseteq \mathbb{K}^{\dim(V)}$ and assume that all the $\mathcal{S}_\alpha$ are the same set for all $\alpha \in \mathcal{Z}$. Let $\mathcal{S} = \mathcal{S}_\alpha$ for all α. Note that $\mathcal{S} \not\subseteq \mathbb{K}[V]$, as $\mathcal{S} \subseteq \mathcal{A}_d^V$. Therefore, we may apply Proposition 3.6.16 and assume that $|\varphi_{V,\alpha}(\mathcal{S})| = |\mathcal{S}|$, after possibly replacing $\mathcal{Z}$ with a smaller dense set. Moreover, we may also assume that all elements of $\varphi_{V,\alpha}(\mathcal{S})$ are irreducible of degree d. We update V to $\mathrm{AH}(V, \mathcal{S})$ and go to the next iteration.

Now we will show that after at most $2\cdot d\cdot\varepsilon^{-1}$ iterations of the above process, one of the two break conditions must be met. This implies that the above process is a $(k, 4\cdot d\cdot\varepsilon^{-1})$ process as required, and that the required conclusion holds.

Analysis of one iteration. Suppose that $|\mathcal{B} \cap \mathbb{K}[V]| < \varepsilon|\mathcal{B}|$. Moreover, there exists a set $\mathcal{S} \subseteq \mathcal{A}_d^V$ with $|\mathcal{S}| = a$ and a dense set $\mathcal{Z} \subseteq \mathbb{K}^{\dim(V)}$, such that for all $\alpha \in \mathcal{Z}$ the set $\varphi_{V,\alpha}(\mathcal{S})$ is an unbreakable pencil generating in $R[y_1]/I_{V,\alpha}$. Let $R'_\alpha := R[y_1]/I_{V,\alpha}$. We define $Y := AH_R(V, \mathcal{S})$. Note that Y is a $h_{2\mu} \circ t_1$-lifted strong vector space in R'_α by Lemma 3.6.10, since it is a vector space that is intermediate in a $(k, 4 \cdot d \cdot \varepsilon^{-1})$ process and since V is strong enough. Moreover we have $V \subseteq Y$. By choosing compatible bases of $V \subseteq Y$, we let $Y = V + Y'$ and $Y'_\alpha := \varphi_{V,\alpha}(Y) \subseteq R'_\alpha$, as in Lemma 3.9.15. Let $Y_\alpha = \mathrm{span}_{\mathbb{K}}\{y_1, Y'_\alpha\}$.

Let $B_1, \ldots, B_r \in \mathcal{B}$ be the forms that either have degree less than d, or that are absolutely reducible with respect to V or in the ideal (V). By definition, we have $r = M_b^V$. Among these forms, let $B_1, \ldots, B_s$ be the set of forms that are not in $\mathbb{K}[V]$. Since $|\mathcal{B} \cap \mathbb{K}[V]| < \varepsilon|\mathcal{B}|$, we have $s \geq M_b^V - \varepsilon|\mathcal{B}|$.

For $\alpha \in \mathcal{Z}$, let $\mathcal{G}_\alpha \subseteq \{B_1, \cdots, B_r\}$ be the (possibly empty) subset of forms such that some factor of $\varphi_{V,\alpha}(B_i)$, other than y_1, is either absolutely reducible over Y_α, or in the ideal (Y_α). Now there are finitely many possibilities for $\mathcal{G}_\alpha$ as $B_1, \cdots, B_r$ are fixed. Therefore, by replacing $\mathcal{Z}$ with a possibly smaller dense subset, we may assume that there exists a set $\mathcal{G}$ such that $\mathcal{G}_\alpha = \mathcal{G}$ for all $\alpha \in \mathcal{Z}$. In particular, if $B_i \in \mathcal{G}$ then some factor of $\varphi_{V,\alpha}(B_i)$, other than y_1, is either absolutely reducible over Y_α, or in the ideal (Y_α), for all $\alpha \in \mathcal{Z}$. Therefore, if $B_i \in \mathcal{G}$ then $\Psi_{Y,(\alpha,\beta)}(B_i) > \Psi_{V,\alpha}(B_i)$, for all $\alpha \in \mathcal{Z}$ and for general $\beta \in \mathbb{K}^{\dim(Y')}$ (depending on α) by Lemma 3.9.15. We will show below that $|\mathcal{G}| \geq \varepsilon|\mathcal{B}|$.

First, let us show that this condition implies that there can be at most $2 \cdot d \cdot \varepsilon^{-1}$ number of iterations before a breaking condition is reached.

Bounding total number of iterations. Suppose we know that $|\mathcal{G}| \geq \varepsilon|\mathcal{B}|$ at each iteration. Now each form in $\mathcal{G}$ is potential increasing with respect to (V, α) and $(Y, (\alpha, \beta))$, for all $\alpha \in \mathcal{Z}$ and general choice of β (depending on each α). Then, in each iteration, when we increase V to Y, the group potential of $\mathcal{B}$ increases by at least $\varepsilon|\mathcal{B}|$, i.e.

$$\Phi_{Y,(\alpha,\beta)}(\mathcal{B}) > \Phi_{V,\alpha}(\mathcal{B})$$

for a dense set of $\alpha \in \mathcal{Z}'$ and general β (depending on α). In other words, the group potential jumps by at least $\varepsilon|\mathcal{B}|$ for a dense set of (α, β), when we replace V by Y. Note that the group potential of $\mathcal{B}$ is upper bounded by $2d|\mathcal{B}|$ by Remark 3.9.13, irrespective of the choice of α, β. Therefore there can be at most $2 \cdot d \cdot \varepsilon^{-1}$ iterations that do not reach a break condition.

Showing $|\mathcal{G}| \geq \varepsilon|\mathcal{B}|$ for each iteration. Suppose that we have fixed V and a dense set of $\mathcal{Z} \subseteq \mathbb{K}^{\dim(V)}$ as above. Suppose that $|\mathcal{G}| < \varepsilon|\mathcal{B}|$. We can further reorder $B_1, \ldots, B_s$ so that the elements of $\mathcal{G}$ are at the end of the ordering. Suppose $B_1, \ldots, B_t \notin \mathcal{G}$, we have $t \geq M_b^V - 2\varepsilon|\mathcal{B}| \geq 4\varepsilon|\mathcal{B}|$.

For each B_i with $i \leq t$, fix an irreducible factor B'_i of $\varphi_{V,\alpha}(B_i)$, such that B'_i is not a scalar multiple of y_1. Note that $\deg(B'_j) < d$, by the choice of B_j and Proposition 3.7.14. Since $B_i \notin \mathcal{G}$, we know that B'_i is absolutely irreducible over Y_α and $B'_i \notin (Y_\alpha)$. Recall that $\mathcal{S} \subseteq \mathbb{K}[Y]$, and hence $\varphi_{V,\alpha}(\mathcal{S}) \subseteq \mathbb{K}[Y_\alpha]$. Moreover, Y_α is sufficiently strong by Proposition 3.9.9, so that we may apply Corollary 3.7.10. Hence each B'_i is not prime with at most $\Pi_1(d)$ elements in $\varphi_{V,\alpha}(\mathcal{S})$. Moreover each B'_i spans something reducible with at most $\Pi_2(d)$ many elements in $\varphi_{V,\alpha}(\mathcal{S})$.

Recall that $|\varphi_{V,\alpha}(\mathcal{S})| = |\mathcal{S}| = a$. Let $\varphi_{V,\alpha}(\mathcal{S}) = \{F_1, \cdots, F_a\}$. We denoted $\Pi_3(d) = \Pi_1(d) + \Pi_2(d)$. Then by a double counting, at most $2\Pi_3(d)$ of the forms among $F_1, \ldots, F_a$ are not prime with more than $t/2$ of the B'_i. Let $b := a - 2\Pi_3(d)$. We may reorder and assume that for any fixed $i \in [b]$, there are at least $t/2$ forms among the B'_j, such that (F_i, B'_j) is prime.

By Proposition 3.9.3 we know that the sets $\varphi_{V,\alpha}(\mathcal{A}), \varphi_{V,\alpha}(\mathcal{B}), \varphi_{V,\alpha}(\mathcal{C})$ form a (d, z, R)-EK configuration. Therefore, for each F_i, B'_j which form a prime ideal, there exists a form $C \in \mathcal{C}$ such that $\varphi_{V,\alpha}(C) \in \left(F_i, B'_j\right)$. For $i \in [b]$, we let $\mathcal{P}_{i,\alpha}$ be the set of forms $C \in \mathcal{C}_d$ such that $\varphi_{V,\alpha}(C) \in \left(F_i, B'_j\right)$

for some j such that $\left(F_i, B'_j\right)$ is prime. Now there are finitely many possibilities for the sets $\mathcal{P}_{i,\alpha} \subseteq \mathcal{C}_d$. Therefore, by replacing $\mathcal{Z}$ with a possibly smaller dense subset, we may assume that there exists a set $\mathcal{P}_i$ such that $\mathcal{P}_i = \mathcal{P}_{i,\alpha}$ for all $\alpha \in \mathcal{Z}$. Recall that F_i is of degree d. Moreover B'_j does not divide $\varphi_{V,\alpha}(C)$ by Proposition 3.6.16, since $B'_j \notin (y_1)$ and α varies in a dense set. Therefore $\varphi_{V,\alpha}(C)$ is irreducible and has degree d, as $\varphi_{V,\alpha}(C) \in (F_i, B'_j)$. Since this holds for all α in a dense set, we see that C must be absolutely irreducible over V, by Proposition 3.7.14.

We first prove that each $|\mathcal{P}_i| \geq t/2(d+1) \geq t/4d$. Fix $i = 1$ without loss of generality. Suppose F_1 is prime with $B'_1, \ldots, B'_{t/2}$. Suppose $|\mathcal{P}_1| < t/2(d+1)$. Then there exists $C \in \mathcal{P}_i$ such that $\varphi_{V,\alpha}(C) \in (F_1, B'_j)$ for all $j \in [d+1]$, after a possible reordering of $B'_1, \ldots, B'_{t/2}$. Therefore we have

$$c_1 F_1 + B'_1 H_1 = \cdots = c_{d+1} F_1 + B'_{d+1} H_{d+1}$$

for some scalars $c_i \in \mathbb{K}$ and forms $H_i \in R'_\alpha$. Recall that and $\deg(B'_i) < d$ for all i and $\deg(F_1) = d$. If $c_i \neq c_j$ for some $i \neq j$, then we have $(F_1, B'_i) \subseteq \left(B'_i, B'_j\right)$. This is a contradiction since (F_1, B'_i) is prime. Therefore we have $c_1 = \cdots = c_{d+1}$. Now each of $B'_2, \ldots, B'_{d+1}$ are factors of H_1, which is a contradiction since H_1 has at most d factors.

Since $t \geq 4\varepsilon|\mathcal{B}|$, we have $|\mathcal{P}_i| \geq \varepsilon|\mathcal{B}|/d \geq \varepsilon m_c^V/d$. The number of sets $\mathcal{P}_i$ is $b := a - 2\Pi_3(d)$. Hence $b > d(\Pi_3(d)+2)\varepsilon^{-1}$. By double counting, we can assume that there is a common element in $\mathcal{P}_1, \ldots, \mathcal{P}_{\Pi_3(d)+2}$. We use this to derive the contradiction, completing the proof.

Suppose C lies in $\mathcal{P}_1, \ldots, \mathcal{P}_{\Pi_1(d)+2}$ with $C' = \varphi_{V,\alpha}(C)$. Therefore, for each F_i there exists B'_i such that $C' = c_i F_i + H_i B'_i$. We therefore have $(c_1 F_1 - c_i F_i, B'_1) \subseteq (B'_1, B'_i)$ for all $2 \leq i \leq \Pi_3(d)+1$. Recall that $\varphi_{V,\alpha}(\mathcal{S})$ is an unbreakable pencil generating set. Hence the forms $c_1 F_1 - c_i F_i$ are nonzero and pairwise non associate, by the linear independence of $F_1, \ldots, F_a$. Further, each element is irreducible as $s(F_i, F_j) \geq 1$. Recall that B'_1 is absolutely irreducible over Y_α and $B'_1 \notin (Y_\alpha)$. Therefore, by Corollary 3.7.10, there is a j such that $(c_1 F_1 - c_j F_j, B'_1)$ is prime. Since B'_1, B'_j are irreducible in R'_α, which is a UFD, the ideal has height 2. Therefore we must have an equality $(c_1 F_1 - c_j F_j, B'_1) = \left(B'_1, B'_j\right)$. This is a contradiction.

In the case when $M_a^V \geq 6\varepsilon|\mathcal{A}|$ the exact argument holds with all occurrences of $\mathcal{A}, \mathcal{B}$ swapped. Now suppose we are in the case when $M_c^V \geq 6\varepsilon|\mathcal{C}|$ and $|\mathcal{C}| \geq |\mathcal{B}|/2$. We swap the roles of $\mathcal{C}$ and $\mathcal{B}$. Note that in this case, $|\mathcal{C} \cap \mathbb{K}[V]| \geq \varepsilon|\mathcal{C}|$ is also a valid breaking condition, since Lemma 3.9.16 can be applied with $\mathcal{C}$ by assumption. The only other place in the arguments where the relative sizes of $\mathcal{A}, \mathcal{B}, \mathcal{C}$ play a role is to ensure that the sets $\mathcal{P}_i$ have a common intersection. Each set in this case will have size $\varepsilon|\mathcal{C}|/d$, which is at least $\varepsilon m_b^V/2d$ by assumption. The rest of the argument remains the same. □

Lemma 3.9.20. *Suppose V is a $H(\mu, \Pi_1(d) + \Pi_2(d) + 1, 2)$-lifted strong space such that one of m_a^V, m_b^V, m_c^V is zero. Then there is a μ-lifted strong space W obtained by applying a $(\Pi_1(d) + \Pi_2(d) + 1, 2)$ process to V such that $m_a^W = m_b^W = m_c^W = 0$.*

Proof. In this proof, the roles of $\mathcal{A}, \mathcal{B}, \mathcal{C}$ can be freely permuted. Therefore without loss of generality assume $m_a^V = 0$. Let $\Pi_3(d) := \Pi_1(d) + \Pi_2(d)$.

Let $B_1, \ldots, B_r$ be the forms in $\mathcal{B}_d$ that are absolutely irreducible over V. If the span of $B_1, \ldots, B_r$ is less than $\Pi_3(d) + 1$, then we add a basis of this set to V, and call this resulting space Y. We have $m_b^Y = m_a^Y = 0$, and we move to the next step. Suppose the span is larger than $\Pi_3(d) + 1$. Without loss of generality, assume $B_1, \ldots, B_{\Pi_3(d)+1}$ are linearly independent. Let Y be the space $\mathrm{AH}(V, B_1, \ldots, B_{\Pi_3(d)+1})$. Now let C be a form in $\mathcal{C}_d$ that is absolutely irreducible over Y, if such a form exists. There is at least one form among $B_1, \ldots, B_{\Pi_3(d)+1}$ say B_i such that (C, B_i) is prime, and does not span anything absolutely reducible. However, the EK image of B_i, C has to be a form

in $\mathcal{A}_d$ that is absolutely irreducible over V, but no such form exists. Therefore, we have $m_c^Y = 0$. This completes the first step of our process.

After rearranging again, we can assume now that $m_a^Y = m_b^Y = 0$. Among the forms in $\mathcal{A}_d, \mathcal{B}_d$, pick a set of forms $F_1, \dots, F_{\Pi_3(d)+1}$ that are linearly independent. If such forms exist, set $W = \mathrm{AH}(Y, F_1, \dots, F_{\Pi_3(d)+1})$. Each form $C \in \mathcal{C}_d$ that is absolutely irreducible over W has at least one F_i such that (F_i, C) is prime and does not span anything absolutely reducible, but the EK image of this ideal is in $\mathcal{A}_d \cup \mathcal{B}_d$ which only consists of absolutely reducible forms.

Suppose $F_1, \dots, F_{\Pi_3(d)+1}$ cannot be picked. Then we instead define $F_1, \dots, F_{\Pi_3(d)+1}$ to be a set of linearly independent forms in $\mathcal{A} \cup \mathcal{B}$ that consist of a basis of $\mathcal{A}_d \cup \mathcal{B}_d$ (the forms are now allowed to be of lower degree), if possible. Again let $W = \mathrm{AH}(Y, F_1, \dots, F_{\Pi_3(d)+1})$. Again if $C \in \mathcal{C}_d$ is absolutely irreducible over W then we have (C, F_i) is prime for some i, and therefore $G = \alpha C + F_i H_i$ for some degree d form $G \in \mathcal{A}_d \cup \mathcal{B}_d$. This implies that C is in the ideal generated by W, since by assumption every form in $\mathcal{A}_d \cup \mathcal{B}_d$ is in the span of W. This is again a contradiction.

The only remaining case is when $\mathcal{A} \cup \mathcal{B}$ has span at most $\Pi_3(d) + 1$. In this case, we just pick $F_1, \dots, F_a$ to be a basis of $\mathcal{A} \cup \mathcal{B}$, and set $W = \mathrm{AH}(V, F_1, \dots, F_a)$. Every element of $\mathcal{A} \cup \mathcal{B}$ is in the algebra $\mathbb{K}[W]$. Now for any $C \in \mathcal{C}_d$ that is not in the ideal, and for any element say $A \in \mathcal{A}$, the radical ideal $\mathrm{rad}(A, C)$ has to contain $\prod_{B \in \mathcal{B}} B$. The latter element is in $\mathbb{K}[W]$, and by the elimination theorem [OS24, Lemma 4.26], the only element in $\mathrm{rad}(A, C) \cap \mathbb{K}[W]$ is A, which is a contradiction.

Therefore in all cases, we have $m_a^W = m_b^W = m_c^W = 0$. The claimed bound on the strength of W follows from Lemma 3.6.10. □

The following lemma combines the above results and allows us to fully control the highest degree forms in an EK-configuration.

Lemma 3.9.21. *Let $c := 3(\Pi_1(d) + \Pi_2(d) + 1)$ and $\delta := 8 \cdot \varepsilon \cdot (\Pi_1(d) + \Pi_2(d) + 1)^{-1}$. Let $k := 6 \cdot d \cdot (\Pi_1(d) + \Pi_2(d) + 1)^3 \cdot \varepsilon^{-1} + c + c_{\mathrm{ek}} \cdot \delta^{-1}$ and $t := 8 \cdot d \cdot \varepsilon^{-1} + 3$. Let μ be a function, and suppose U is $H(\mu, k, t)$-strong. There is a (k, t) process starting with 0 such that the resulting space $W \subset R$ is μ-lifted strong and satisfies $m_a^W = 0$ and $m_b^W = 0$ and $m_c^W = 0$. Further, $\dim W_i \leq D_i(\mu, k, t, \delta_U)$, where δ_U is the dimension vector of U.*

Proof. The first step is to find a space V such that one of m_a^V, m_b^V, m_c^V is zero.

If $|\mathcal{C}| < 2\Pi_1(d)\varepsilon^{-1}$ then we pick $V = \mathrm{AH}(0, \mathcal{C})$ and we are done. If $M_b^{(0)} \geq 6\varepsilon|\mathcal{B}|$ or $M_a^{(0)} \geq 6\varepsilon|\mathcal{A}|$ then we apply Lemma 3.9.19 starting with 0 and we are done.

Now consider the sets $\mathcal{A}_d, \mathcal{B}_d, \mathcal{C}_d$. Even though we have $|\mathcal{A}| \geq |\mathcal{B}| \geq |\mathcal{C}|$, we cannot assume $|\mathcal{A}_d| \geq |\mathcal{B}_d| \geq |\mathcal{C}_d|$. For each form $F \in \mathcal{F}_d$ we define a subset $\mathcal{F}_{\mathrm{span}}(F)$ as follows. If $F \in \mathcal{A}_d$ and $|\mathcal{B}_d| \geq |\mathcal{C}_d|$ then $\mathcal{F}_{\mathrm{span}}(F) \subset \mathcal{B}_d$ is the set of forms $B \in \mathcal{B}_d$ such that $|\mathrm{span}_{\mathbb{K}}\{F, B\} \cap \mathcal{C}_d| \geq 1$. If instead $|\mathcal{C}_d| \geq |\mathcal{B}_d|$ then $\mathcal{F}_{\mathrm{span}}(F) \subset \mathcal{C}_d$ is the set of forms $C \in \mathcal{C}_d$ such that $|\mathrm{span}_{\mathbb{K}}\{F, C\} \cap \mathcal{B}_d| \geq 1$. We can similarly define $\mathcal{F}_{\mathrm{span}}(F)$ for forms $F \in \mathcal{B}_d, \mathcal{C}_d$. In each case, $\mathcal{F}_{\mathrm{span}}(F)$ will be a subset of the bigger of the two degree d parts of the remaining sets. The fractional size of $\mathcal{F}_{\mathrm{span}}(F)$ is defined to be the ratio of $|\mathcal{F}_{\mathrm{span}}(F)|$ and the size of whichever among $\mathcal{A}_d, \mathcal{B}_d, \mathcal{C}_d$ that $\mathcal{F}_{\mathrm{span}}(F)$ is a subset of.

For a form $F \in \mathcal{A}_d \cup \mathcal{B}_d$, we have $\mathcal{F}_{\mathrm{span}}(F) \subset \mathcal{C}_d$ only if $|\mathcal{C}_d| \geq |\mathcal{B}_d|$ or $|\mathcal{C}_d| \geq |\mathcal{A}_d|$. We are in the case where $M_b^{(0)} < 6\varepsilon|\mathcal{B}| < |\mathcal{B}|/2$ and $M_a^{(0)} < 6\varepsilon|\mathcal{A}| < |\mathcal{A}|/2$, equivalently in the case where $|\mathcal{B}_d| \geq |\mathcal{B}|/2$ and $|\mathcal{A}_d| \geq |\mathcal{A}|/2$. Therefore if $\mathcal{F}_{\mathrm{span}}(F) \subset \mathcal{C}_d$ then we have $|\mathcal{C}| \geq |\mathcal{B}|/2$. This observation will be useful later in the proof.

If the set $\mathcal{A}_d \cup \mathcal{B}_d \cup \mathcal{C}_d$ is a (c, δ)-linear EK configuration (as defined in Definition 3.1.9), then by Proposition 3.1.10 we can find a basis of $\mathcal{A}_d \cup \mathcal{B}_d \cup \mathcal{C}_d$ of size at most $c + c_{\mathrm{ek}} \cdot \delta^{-1}$, and we are done. Suppose they do not form such a configuration. In particular, there are at least c forms

in $\mathcal{F}_d$ with fractional $\mathcal{F}_{\text{span}}(F)$ smaller than δ. By the pigeonhole principle, there are at least $c/3 = \Pi_1(d) + \Pi_2(d) + 1$ such forms within the same set. We now do a case analysis.

Case $|\mathcal{C}| \leq |\mathcal{B}|/2$: The key observation in this case is that none of the $\mathcal{F}_{\text{span}}(F)$ are subsets of $\mathcal{C}_d$. We analyse the subcase when there are $c/3$ forms in $\mathcal{C}_d$ with $\mathcal{F}_{\text{span}}(C_i) \subset \mathcal{A}_d$ having fractional size less than δ. The other subcases have the exact same proof (with the roles of $\mathcal{A}, \mathcal{B}$ swapped if $|\mathcal{A}_d| \leq |\mathcal{B}_d|$). Suppose forms $C_1, \ldots, C_{c/3} \in \mathcal{C}_d$ have fractional size less than δ, and suppose $\mathcal{F}_{\text{span}}(C_i) \subset \mathcal{A}_d$. Let $Y := \text{AH}(0, C_1, \ldots, C_{c/3})$. If there is some $A \in \mathcal{A}_d$ such that (C_i, A) is prime, then $A \in \mathcal{F}_{\text{span}}(C_i)$. Therefore each C_i is prime with at most $\delta|\mathcal{A}_d|$ elements of $\mathcal{A}_d$. By a union bound, there are are at most $(c/3)\delta|\mathcal{A}_d| \leq 4\varepsilon|\mathcal{A}|$ elements in $\mathcal{A}_d$ that are prime with some element among $C_1, \ldots, C_{c/3}$. The remaining $(1-10\varepsilon)|\mathcal{A}|$ elements of $\mathcal{A}$ are either absolutely reducible over Y or in the algebra. Therefore, we have $M_a^Y \geq (1-10\varepsilon)|\mathcal{A}| \geq 6\varepsilon|\mathcal{A}|$. Now we can apply Lemma 3.9.19 starting with Y and the resulting space V is such that one of m_a^V, m_b^V, m_c^V is zero.

Case $|\mathcal{C}| \geq |\mathcal{B}|/2$: In this case, it is possible that some forms $F \in \mathcal{F}_d$ are such that $\mathcal{F}_{\text{span}}(F) \subset \mathcal{C}_d$. If we are in this case and we have $M_c^{(0)} \geq 6\varepsilon|\mathcal{C}|$ then we can Lemma 3.9.19 starting with 0 and we are done. Therefore we have to deal with the subcase when $M_c^{(0)} \leq 6\varepsilon|\mathcal{C}|$.

If there are $c/3$ forms $F_1, \ldots, F_{c/3}$ all in $\mathcal{A}_d$ or $\mathcal{B}_d$ with $\mathcal{F}_{\text{span}}(F) \subset \mathcal{C}_d$ and fractional size less than δ, then we can set $Y := \text{AH}(0, F_1, \ldots, F_{c/3})$, and we have $M_c^Y \geq (1-10\varepsilon)|\mathcal{C}| \geq 6\varepsilon|\mathcal{C}|$. Therefore we can invoke Lemma 3.9.19. In the other cases, for example if the $c/3$ forms are in $\mathcal{C}_d$, then the analysis from the previous case applies. Therefore, we have V such that one of m_a^V, m_b^V, m_c^V is zero.

Now that we have the space V, we apply the process of Lemma 3.9.20 to obtain W. The fact that this entire process is a (k, t) process for the claimed k, t follows simply by adding up the number of iterations in the intermediate processes, and taking a max for the number of forms added at each step of the intermediate processes (for incomparable terms we use the sum to bound the max). The claimed bounds on the strength and dimension of W follow from Lemma 3.6.10 and the observation that since U is $H(\mu, k, t)$-strong, the 0 vector space in R is $H(\mu, k, t)$-lifted strong. □

3.9.4 Putting it all together

We can now prove our main theorem for EK-configurations.

Theorem 3.9.22. *For any two positive integers d, e such that $e \geq 2$ and $d \leq e$, there exist ascending functions $\Lambda_{d,e} : \mathbb{N}^e \to \mathbb{N}^e$ and $\lambda_{d,e} : \mathbb{N} \to \mathbb{N}$, both independent of $\mathbb{K}$ and N, such that the following holds.*

Let $\mathbb{K}$ be an algebraically closed field of characteristic 0 and $S := \mathbb{K}[x_1, \cdots, x_N]$. Let $U \subset S_{\leq e}$ be a $\Lambda_{d,e}$-strong graded vector space and $R = S/(U)$. If $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ is a (d, z, R)-EK configuration in R for some $z \in S_1$, then we have

$$\dim(\text{span}_{\mathbb{K}}\{\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}\}) \leq \lambda_{d,e}(\dim(U)).$$

That is, the dimension of the $\mathbb{K}$-linear span of $\mathcal{A}, \mathcal{B}, \mathcal{C}$ is upper bounded by a function of d, e, δ and $\dim(U)$, which is independent of the field $\mathbb{K}$, the number of variables N and the cardinality of $\mathcal{A}, \mathcal{B}, \mathcal{C}$.

Proof. For each fixed e, we prove the result by induction on d. Therefore, for the rest of this proof we fix e. Fix $\mu : \mathbb{N}^e \to \mathbb{N}^e$ such that $\mu_i(\delta) = A(\eta, i) + 3\|\delta\|_1$, where A is the function defined in [AH20b, Theorem A]. We will define the functions $\Lambda_{d,e}$ and $\lambda_{d,e}$ inductively, and simultaneously prove that they have the claimed properties.

The base case is when $d = 1$. Define $\Lambda_{1,e} := h_{2\mu} \circ t_2$. If U is $\Lambda_{1,e}$-strong for this choice, then any two linear forms in R that are non associate form a prime ideal by [OS24, Proposition 5.10]. In particular, for any $A \in \mathcal{A}, B \in \mathcal{B}$ the ideal (A, B) is prime, and $\operatorname{rad}(A, B) = (A, B)$. The second condition for $(1, z, R)$-EK configurations reduces to the condition that the linear span of A, B contains some element in $\mathcal{C} \cup \{z\}$. The third and fourth conditions reduce to similar symmetric statements. This shows that the set $\mathcal{F}$ is a $(1, 1)$-partial linear EK configuration, where the error subset $\mathcal{G}$ is the singleton $\{z\}$, and the partitions $\mathcal{A}, \mathcal{B}, \mathcal{C}$ witness the required partition in the definition of $(1, 1)$-partial linear EK configurations. Therefore we can bound the dimension of $\mathcal{F}$ by $\lambda_{1,e} := 1 + c_{\mathrm{ek}}$. This completes the base case.

Inductively, suppose we have proved the theorem for $d-1$, and in particular suppose the functions $\Lambda_{d-1,e}, \lambda_{d-1,e}$ are defined. Let $c := 3(\Pi_1(d) + \Pi_2(d) + 1)$ and $\delta := 8 \cdot \varepsilon \cdot (\Pi_1(d) + \Pi_2(d) + 1)^{-1}$. Let $k := 6 \cdot d \cdot (\Pi_1(d) + \Pi_2(d) + 1)^3 \cdot \varepsilon^{-1} + c + c_{\mathrm{ek}} \cdot \delta^{-1} \cdot \log \delta^{-1}$ and $t := 8 \cdot d \cdot \varepsilon^{-1} + 3$. Set $\Lambda_{d,e} := H(\Lambda_{d-1,e}, k, t)$.

Suppose U is $\Lambda_{d,e}$-strong, and $\mathcal{A}, \mathcal{B}, \mathcal{C}$ is an (d, R, z)-EK configuration. By Lemma 3.9.21, there exists a vector space $W \subset R$ that is $\Lambda_{d-1,e}$-lifted strong such that every degree d form in $\mathcal{F}$ is either in the ideal generated by W, or absolutely reducible with respect to W. For a general choice of α, consider the graded quotient $R[y] \to R[y]/I_\alpha$ where I_α is the ideal is the ideal corresponding to W and α in $R[y]$. By Proposition 3.9.3, the image of the configuration $\mathcal{A}, \mathcal{B}, \mathcal{C}$ is itself an $(d, y, R[y]/I_\alpha)$-EK configuration. Further, since every degree d form in $\mathcal{F}$ is either in the ideal generated by W, or absolutely reducible with respect to W, we can deduce by Proposition 3.7.14 that the image is in fact a $(d-1, y, R[y]/I_\alpha)$-EK configuration. Note that $R[y]/I_\alpha = S[y]/U + I_\alpha$. Further, since W is $\Lambda_{d-1,e}$-lifted strong, we can deduce that $U + I_\alpha$ is generated by a $\Lambda_{d-1,e}$-strong vector space of dimension at most $\dim U + \dim W$. We can apply the inductive hypothesis to deduce that the image of the configuration under the graded quotient has dimension at most $\lambda_{d-1,e}(\dim W + \dim U)$. By Proposition 3.6.17 we can deduce that the dimension of the original configuration is bounded by $\lambda_{d-1,e}(\dim W + \dim U) \cdot (1 + d)^{3 \dim W + 4} \cdot e^{\dim U}$. We also have $\dim W \leq \sum_{i=1}^{d} D_i(\Lambda_{d-1,e}, k, t, \delta_U)$. Therefore, if we set

$$\lambda_{d,e}(u) := \max_{\substack{\delta \in \mathbb{N}^e \\ \|\delta\|_1 = u}} \lambda_{d-1,e}\left(\sum_{i=1}^{d} D_i(\Lambda_{d-1,e}, k, t, \delta) + u\right) \cdot (1 + d)^{3 \sum_{i=1}^{d} D_i(\Lambda_{d-1,e}, k, t, \delta) + 4} \cdot e^u,$$

then we are done. □

A corollary is the claimed algorithm for identity testing for $\Sigma^3\Pi\Sigma\Pi^d$ circuits. To show this, we first establish a connection between these circuits and EK configurations.

Definition 3.9.23 (Configuration corresponding to $\Sigma^3\Pi\Sigma\Pi^d$ circuits.)**.** Given a $\Sigma^3\Pi\Sigma\Pi^d$ circuit $T = \sum_{i=1}^{3} \prod_j P_{ij}$, we define a configuration of forms as follows. Let $\mathcal{A}'$ be the set of irreducible factors of the polynomials $\{P_{1j}\}_j$. Define $\mathcal{A}$ to be the set of forms obtained by homogenising the forms in $\mathcal{A}'$, where we homogenise using a new variable z. If $\mathcal{A}$ contains a subset of forms that are pairwise associate, then we discard all except one form from this subset from $\mathcal{A}$, so the final set $\mathcal{A}$ does not have any pair of forms that are associate to each other. Similarly, the sets $\mathcal{B}, \mathcal{C}$ are defined using the polynomials $\{P_{2j}\}_j$ and $\{P_{3j}\}_j$ respectively.

The sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ consist of irreducible homogeneous forms of degree at most d from the ring $S[z]$. Further, neither of the three sets contains the form z. ◊

Lemma 3.9.24. *Suppose $T = \sum_{i=1}^{3} \prod_j P_{ij} = \sum_{i=1}^{3} T_i$ is a simple minimal circuit that computes 0. The sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ as defined in Definition 3.9.23 are a $(d, z, S[z])$-EK configuration. Further, $\operatorname{rank} T \leq \lambda^d$ where λ is the rank of the EK configuration.*

Proof. Note that T continues to compute zero, and continues to be simple and minimal when we consider it as a polynomial in $S[z]$. For a circuit computing zero with top fan-in three, minimality is equivalent to the pairwise linear independence of T_1, T_2, T_3. In particular, none of $T_1 + T_2, T_2 + T_3, T_1 + T_3$ are zero. This will be used implicitly throughout the rest of the argument.

By construction, the sets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ each consist of non-associate forms. Suppose a form $A \in \mathcal{A}$ and a form $B \in \mathcal{B}$ are associate. Suppose A, B are obtained by homogenising A', B' respectively. Since A, B are associate, the polynomials A', B' are scalar multiples of each other. Since A', B' are factors of T_1, T_2 respectively, and since $T = 0$, it must be that A' is also a factor of T_3. This implies that $A' | \gcd(T_1, T_2, T_3)$, contradicting simplicity. Therefore $\mathcal{A}$ and $\mathcal{B}$ are disjoint, and the union $\mathcal{A} \cup \mathcal{B}$ consists of non-associate forms. A symmetric argument with the other pairs of sets shows that $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are pairwise disjoint, and that the union consists of non-associate forms. Since none of these sets contain z, the union $\{z\} \cup \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ consists of pairwise non-associate forms. This shows the first condition.

Now suppose $A \in \mathcal{A}, B \in \mathcal{B}$ are forms that are obtained by homogenising A', B' respectively. We have $T_3 \in (A', B') \subset \mathrm{rad}(A', B')$. Let $\mathcal{C}'$ be the set of irreducible factors of $\{P_{3j}\}_j$, we have $\prod_{C' \in \mathcal{C}'} C' \in \mathrm{rad}(T_3) \subset \mathrm{rad}(A', B')$. Upon homogenisation, we obtain $\prod_{C \in \mathcal{C}} C \in \mathrm{rad}(zA, zB) \subset \mathrm{rad}(A, B)$. Therefore $z \cdot \prod_{C \in \mathcal{C}} C \in \mathrm{rad}(A, B)$. This shows that the second condition for the configuration to be an EK configuration holds, and by symmetric arguments so do the third and fourth condition.

We now show the final statement. The homogenisation of P_{ij} is a polynomial of degree at most d in a basis for the EK configuration. Therefore, these homogenisations span a vector space of dimension at most λ^d. Picking a basis and setting $z = 1$ gives a basis for the forms P_{ij}. □

Corollary 3.9.25. *There is a deterministic polynomial-time algorithm for identity testing of* $\Sigma^3\Pi\Sigma\Pi^d$ *circuits.*

Proof. By Definition 3.9.23 we can define an EK configuration from the circuit Φ. By Theorem 3.9.22, the rank of the configuration is bounded by $\lambda_{d,d}(0)$ for the function $\lambda_{d,d}$ defined in the theorem. By Lemma 3.9.24, the rank of the circuit is bounded by $\rho(d) := \lambda_{d,d}(0)^d$. Therefore by Theorem 3.2.1, there is an algorithm for identity testing of such circuits that runs in time $(ds\rho(d)n)^{O(d^2\rho(d))}$, where s is the size of the circuit. □

3.10 Rank bounds for constant top fan-in depth four circuits with bottom fan-in 2

In this section, we show that simple and minimal $\Sigma^k\Pi\Sigma\Pi^2$ circuits computing identity have rank that is bounded by only a function of k. As a consequence, when k is a constant, we obtain a polynomial time algorithm for identity testing of these circuits. We obtain our rank bounds by adapting the proof of Kayal and Saraf [KS09], who proved rank bounds for $\Sigma^k\Pi\Sigma$ circuits. As discussed in Section 3.2, Kayal and Saraf do not show that the set of all linear forms in a given simple minimal $\Sigma^k\Pi\Sigma$ identity form a linear Sylvester–Gallai configuration. Instead, they identify certain subsets of these forms (those that fail to satisfy some nice properties that we will describe later), and only show that these subsets satisfy Sylvester–Gallai type properties, which allows them to control the dimension of these bad subsets. This extra flexibility will be crucial in our arguments also.

The overall structure of the proof will be similar to the proof of rank bounds for high dimensional quadratic Sylvester–Gallai configurations (Section 3.5). The rank bounds will proceed in two steps. In the first step, in Section 3.10.2, we will show that all forms in a $\Sigma^k\Pi\Sigma\Pi^2$ are close to a vector

space V. This will be similar to Section 3.5.3. As an intermediate step, we will define and prove rank bounds for a generalisation of strong$(\varepsilon, k) - \mathrm{SG}_t^*(2)$ configurations (Definition 3.5.33. The exact definition of what it means to be close to a vector space will be a generalisation of the notion introduced in Section 3.5.1. We will formally define this notion, and show relevant properties in Section 3.10.1. Once we have this vector space V, we will use an argument similar to that in Section 3.5.4 to study the circuit. This is done in Section 3.10.3. In both of these steps, we will use a Kayal and Saraf style argument to control certain bad subsets. Finally, in Section 3.10.4 we combine all the structural results and prove our rank bounds.

3.10.1 Essential variables in quotient rings

Recall the definitions of linear spaces corresponding to quadratic forms (Section 3.5.1) and the notion of relative linear spaces of quadratic forms (Section 3.5.1). These notions were defined in the ring S. We extend these notions to quadratics in a ring of the form $R := S/(U)$ where U is a strong enough vector space.

Fix $\eta \geq 3$. Throughout this section we will assume that all our ascending functions, such as $\Lambda, \Gamma, \mu : \mathbb{N}^2 \to \mathbb{N}^2$, satisfy the following property: the i-th component is larger than the function

$$B_i(\delta) := A(\eta, i) + 3\left(\sum_i \delta_i\right)$$

for all $\delta \in \mathbb{N}^2$. Here the functions $A(\eta, i)$ are given by [AH20b, Theorem A].

Definition 3.10.1. Fix $a \in \mathbb{N}$. For $G \in R_2$, we define $\mathcal{W}_a(G)$ to be the collection of all subspaces $W \subseteq R_1$ such that $G \in \mathbb{K}[W]$ and $\dim(W) \leq a$. ◊

Proposition 3.10.2. *Suppose that $U \subseteq S$ is Λ-strong and $R = S/(U)$. Let $a \in \mathbb{N}$ be such that $\Lambda_i(\delta_U) \geq B \circ t_{(2a,0)}(\delta_U)$. Let $G \in R_2$. For $r \in \mathbb{N}$, if G is not r-strong, then the following holds.*

1. *We have $\mathcal{W}_{2r}(G) \neq \emptyset$.*
2. *If $a \geq 2r$, then there is an unique vector space $W \in \mathcal{W}_a(G)$ of minimum dimension. Moreover, for any $W' \in \mathcal{W}_a(G)$, we have $W \subseteq W'$.*

Proof. (1) Since G is not r-strong, there exist linear forms $\ell_1, \cdots, \ell_{2r} \in R_1$ such that $G \in \mathbb{K}[\ell_1, \cdots, \ell_{2r}] \subseteq R$. We let $W = \mathrm{span}_{\mathbb{K}}\{\ell_1, \cdots, \ell_{2r}\}$.

(2) Let $W_1, W_2 \subseteq R_1$ be two vector spaces of minimum dimension such that $G \in \mathbb{K}[W_i]$ and $\dim(W_i) \leq a$. Let $\ell_1, \cdots, \ell_m$ be a basis of $W_1 + W_2$ for some $m \leq 2a$. Note that we must have $\dim(W_i) \leq a$, by part (1). Now $\ell_1, \cdots, \ell_m$ is a prime sequence in R by Theorem 3.5.7. Let $A = \mathbb{K}[\ell_1, \cdots, \ell_m]$, which is isomorphic to a polynomial ring in m many variables. Therefore, we may apply [Shp20, Claim 14] in the polynomial ring A, and conclude that $W_1 = W_2$.

Let $W \in \mathcal{W}_a(G)$ be the vector space of minimum dimension. Now let $W' \in \mathcal{W}_a(G)$ be another subspace. We may again find a subalgebra $A \subseteq R$, which is isomorphic to a polynomial ring in at most $2a$ variables that contains basis of $W + W'$. By applying [PS20, Fact 2.15] in the polynomial ring A, we obtain $W \subseteq W'$. □

Corollary 3.10.3. *Let $\mu, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions with $\Lambda \geq B \circ t_{(4\mu_2,0)}$. Suppose that $U \subseteq S$ is $\Lambda \circ t_{(0,1)}$-strong and $R = S/(U)$. Let $G \in R_2$ be a form that is not μ-lifted strong. Then there exists an unique vector space $W \subseteq R_1$ of minimum dimension such that $G \in \mathbb{K}[W]$. Moreover, we have $\dim(W) \leq 2\mu_2(\delta_U + (0,1))$.*

Proof. The strength of G in R is upper bounded by its lifted strength, that is,

$$s(G) \leq s(U, G).$$

Since G is not μ-lifted strong, we know that there is a lift $\widetilde{G}$ such that $U' := \operatorname{span}_{\mathbb{K}}\left\{U + \widetilde{G}\right\}$ is not μ-strong. Now the degree 1 graded piece $U_1' = \operatorname{span}_{\mathbb{K}}\left\{U + \widetilde{G}\right\} \cap R_1$ is μ-strong, hence we must have that $U_2' = \operatorname{span}_{\mathbb{K}}\left\{U + \widetilde{G}\right\} \cap R_2$ is not $\mu_2(\delta_U + (0,1))$-strong. In particular, $s(U, G) \leq \mu_2(\delta_U + (0,1))$, and hence $s(G) \leq \mu_2(\delta_U + (1,0))$. Let $r = \mu_2(\delta_U + (0,1))$ and $\Lambda' = \Lambda \circ t_{(0,1)}$. Then we are done by applying Proposition 3.10.2, as U is Λ'-strong and $\Lambda' \geq B \circ t_{(4r,0)}$. □

Proposition 3.10.4. *Let $\mu, \Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions such that $\Gamma \geq 2\mu$. Suppose that $U \subseteq S$ is Λ-strong and $R = S/(U)$. Let $V \subseteq R_{\leq 2}$ be a Γ-lifted strong vector space. Let $F \in R_2$. Up to scalar multiples, there exists at most one quadratic form $G \in \operatorname{span}_{\mathbb{K}}\{V, F\}$ such that $\operatorname{span}_{\mathbb{K}}\{G\}$ is not μ-lifted strong in R.*

Proof. If $V_2 = (0)$, then the only quadratic forms in $\operatorname{span}_{\mathbb{K}}\{V, F\}$ are scalar multiples of F. Hence we may assume that $V_2 \neq (0)$ and $\delta_V \geq (0,1)$. Let $G_1, G_2 \in \operatorname{span}_{\mathbb{K}}\{V, F\}$ be two non-associate forms which are not μ-lifted strong in R. Then we have $\widetilde{s}_{\min}(U, G_i) < \mu_2(\delta_U + (0,1))$. Let $G_1 = \alpha_1 H_1 + \beta_1 F$ and $G_2 = \alpha_2 H_2 + \beta_2 F$ for $H_1, H_2 \in V$. Since V is Γ-lifted strong and $\Gamma \geq \mu$, we must have $\beta_1, \beta_2 \neq 0$. Let $H := \beta_2 G_1 - \beta_1 G_2 \neq 0$. The lifted strength of H satisfies

$$\widetilde{s}_{\min}(U, H) \leq \widetilde{s}_{\min}(U, G_1) + \widetilde{s}_{\min}(U, G_1) < 2\mu_2(\delta_U + (0,1))$$

Now $\Gamma \geq 2\mu$, hence $2\mu_2(\delta_U + (0,1)) \leq \Gamma_2(\delta_U + \delta_V)$. Therefore, we have a contradiction as V is Γ-lifted strong and $H \in V$ is a non-zero form. □

Definition 3.10.5. Let $R = S/(U)$. For a vector space $V \subseteq R_{\leq 2}$ and a form $F \in R_2$, we define the pseudo-distance of F from the algebra $\mathbb{K}[V]$ as

$$\text{ps-dist}(F, V) = \begin{cases} 1 + \min_Q\{s(F - Q) \mid Q \in \mathbb{K}[V] \cap R_2\} & \text{if } F \notin \mathbb{K}[V] \\ 0 & \text{if } F \in \mathbb{K}[V] \end{cases}$$

We say that F is *s-close* to the algebra $\mathbb{K}[V]$ if $\text{ps-dist}(F, V) < s$. ◊

We note a few properties of the pseudo-distance.

Proposition 3.10.6. *Let $R = S/(U)$. Let $V \subseteq R_{\leq 2}$. We have the following.*

1. *Let $F \in R_2$. If $\text{ps-dist}(F, V) \leq s$, then there exist linear forms $\ell_1, \cdots, \ell_{2s} \in R_1$ such that $F \in \mathbb{K}[V, \ell_1, \cdots, \ell_{2s}]$.*
2. *For any $F_1, F_2 \in R_2$, we have*

$$\text{ps-dist}(F_1 + F_2, V) \leq \text{ps-dist}(F_1, V) + \text{ps-dist}(F_2, V).$$

3. *If $V' \subseteq R_{\leq 2}$ is such that $V \subseteq \mathbb{K}[V']$, then*

$$\text{ps-dist}(F, V') \leq \text{ps-dist}(F, V)$$

for all $F \in R_2$.

Proof. (1) There exists $Q \in \mathbb{K}[V]$ such that $s(F-Q) < s$. Hence there exist $\ell_1, \cdots, \ell_s \in R_1$ such that $F - Q \in (\ell_1, \cdots, \ell_s)$. Since $F - Q \in R_2$, there exist $\ell_{s+1}, \cdots, \ell_{2s}$ such that $F - Q \in \mathbb{K}[\ell_1, \cdots, \ell_{2s}]$. Hence $F = Q + (F - Q) \in \mathbb{K}[V, \ell_1, \cdots, \ell_{2s}]$.

(2) If $F_1 - Q_1 \in (\ell_1, \cdots, \ell_{s_1})$ and $F_2 - Q_2 \in (\ell'_1, \cdots, \ell'_{s_2})$ for some $Q_1, Q_2 \in \mathbb{K}[V]$, then $F_1 + F_2 - (Q_1 + Q_2) \in (\ell_1, \cdots, \ell_{s_1}, \ell'_1, \cdots, \ell'_{s_2})$. Hence $s(F_1 + F_2 - (Q_1 + Q_2)) \leq s_1 + s_2 - 1$. We are done by taking $s_1 = \text{ps-dist}(F_1, V)$ and $s_2 = \text{ps-dist}(F_2, V)$.

(3) Let $Q \in \mathbb{K}[V]$ such that $s(F - Q) = \text{ps-dist}(F, V) - 1$. Then $Q \in \mathbb{K}[V']$, as $V \subseteq \mathbb{K}[V']$. Hence $\text{ps-dist}(F, V') \leq s(F - Q) + 1 = \text{ps-dist}(F, V)$. □

Proposition 3.10.7. *Let $\mu, \Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Let $F \in R_2$.*

1. *Suppose $\Gamma \geq \mu$. If $\text{span}_{\mathbb{K}}\{V, F\}$ is not μ-lifted strong, then*

$$\text{ps-dist}(F, V) \leq \mu_2(\delta_U + \delta_V + (0, 1)).$$

2. *If we have*

$$\text{ps-dist}(F, V) + \dim(V_1) + \dim(U_1) \leq \mu_2(\delta_U + \delta_V + (0, 1)),$$

then $\text{span}_{\mathbb{K}}\{V, F\}$ is not μ-lifted strong.

Proof. (1) Note that $F \notin V$ as $\Gamma \geq \mu$. Since V is $\Gamma \circ t_{(0,1)}$-lifted strong, we know that $s(U, G) \geq \mu_2(\delta_U + \delta_V + (0, 1))$ for all $G \in V_2$. Therefore, there exist lifts of a basis of V and F, given by $\widetilde{F_1}, \cdots, \widetilde{F_m}, \widetilde{F}$ and $\widetilde{Q} \in \text{span}_{\mathbb{K}}\left\{\widetilde{F_1}, \cdots, \widetilde{F_m}\right\}$ such that $s\left(U, \widetilde{F} - \widetilde{Q}\right) < \mu_2(\delta_U + \delta_V + (0, 1))$. Let Q be the image of $\widetilde{Q}$ in the quotient ring $R = S/(U)$. Then we have $Q \in \mathbb{K}[V]$ and $s(F - Q) < \mu_2(\delta_U + \delta_V + (1, 0))$.

(2) Let $Q \in \mathbb{K}[V]$ such that $s(F - Q) = \text{ps-dist}(F, V) - 1$. Let $P = F - Q$ and $r = \text{ps-dist}(F, V)$. Let us first choose appropriate lifts of all the forms involved as follows. Note that there exist $\ell_1, \cdots, \ell_{2r} \in R_1$ such that $P \in \mathbb{K}[\ell_1, \cdots, \ell_{2r}]$. Let $\widetilde{\ell_1}, \cdots, \widetilde{\ell_{2r}}$ be lifts in S. Then there exists $\widetilde{P} \in \mathbb{K}[\widetilde{\ell_1}, \cdots, \widetilde{\ell_{2r}}]$ such that $\widetilde{P}$ is a lift of P. Let $Q = G + H$, where $G \in \mathbb{K}[V_1]$ and $H \in V_2$. Now there exists a lift $\widetilde{G} \in S$ such that $s(\widetilde{G}) < \dim(V_1)$. Let $\widetilde{F}, \widetilde{H}$ be lifts of F, H respectively. Now we have $(\widetilde{F} - \widetilde{G} - \widetilde{H}) - \widetilde{P} \in (U)$. Therefore, there exists a form $H' \in U_2$ such that $(\widetilde{F} - \widetilde{G} - \widetilde{H}) - \widetilde{P} - H' \in (U_1)$. Hence, we have

$$s(\widetilde{F} - \widetilde{H} - H') < r + \dim(V_1) + \dim(U_1),$$

as $s(\widetilde{P}) < r$ and $s(\widetilde{G}) < \dim(V_1)$. Recall that $\widetilde{H}, \widetilde{F}$ are lifts of forms in $\text{span}_{\mathbb{K}}\{V, F\}$. Since $\mu_2(\delta_U + \delta_V + (0, 1)) \geq r + \dim(V_1) + \dim(U_1)$, we have that $\text{span}_{\mathbb{K}}\{V, F\}$ is not μ-lifted strong. □

In Proposition 3.10.2, we saw that in a strong enough quotient ring, given a weak form, there is a unique vector space of linear forms of minimal dimension that contains it in it's algebra. Given a form F is not weak, but is close (of low pseudo-distance) to some $V \subseteq R_{\leq 2}$, we can apply that proposition to the weak form $F - Q$ which shows the low pseudo-distance, resulting in a low-dimension addition to V_1 whose algebra contains F. The following proposition shows that and proves a similar uniqueness guarantee.

Proposition 3.10.8. *Let $\Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Let $F \in R_2$. Let $a, r \in \mathbb{N}$ such that $\text{ps-dist}(F, V) \leq r$ and $a \geq 2r + \dim(V_1)$. Suppose that $\Lambda_i(\delta_U) \geq B_i \circ t_{(2a,0)}(\delta_U)$ and $\Gamma_2(\delta_U + \delta_V) \geq a + \dim(U_1)$. Then there exists a vector space $W \subseteq R_1$ such that the following holds.*

1. *There exists $Q \in \mathbb{K}[V]$ such that $s(F - Q) = \text{ps-dist}(V, F) - 1$ and $F - Q \in \mathbb{K}[W]$.*

2. *We have* $V_1 \subseteq W$.

3. *We have* $\dim(W) \leq 2\,\text{ps-dist}(F, V) + \dim(V_1)$.

4. *Let* $W' \subseteq R_1$ *be a vector space such that* $V_1 \subseteq W'$ *and* $\dim(W') \leq a$. *Suppose that* $F - Q' \in \mathbb{K}[W']$ *for some* $Q' \in \mathbb{K}[V]$ *with* $s(F - Q) < r$. *Then we have* $W \subseteq W'$.

5. *A vector space* W *of minimum dimension that satisfies properties* (1) *and* (2) *above is unique. Moreover, if* W *is the unique vector space of minimum dimension satisfying* (1) *and* (2), *then for all* $Q' \in \mathbb{K}[V]$ *with* $s(F - Q') = \text{ps-dist}(F, V) - 1$, *we have* $F - Q' \in \mathbb{K}[W]$.

Proof. Let $s = \text{ps-dist}(F, V) \leq r$. So we have $a \geq 2s + \dim(V_1)$. Let $Q_1 \in \mathbb{K}[V]$ such that $s(F - Q_1) = \text{ps-dist}(V, F) - 1$. Since $\Lambda_i(\delta_U) \geq B \circ t_{(2a,0)}(\delta_U)$, we may apply Proposition 3.10.2 to the form $F - Q_1 \in R$ which is not s-strong. Let $W_1 \in \mathcal{W}_a(F - Q_1)$ be the unique vector space of minimal dimension provided by Proposition 3.10.2. We define $W := W_1 + V_1$. We will show that W satisfies the properties stated above.

Property (1) is satisfied by construction. We also have $V_1 \subseteq W$ by construction, so (2) is satisfied. Moreover, by Proposition 3.10.2, we have $\dim(W_1) \leq 2s$. Hence we have (3).

Now we will show property (4) for W as defined above. Let $W' \subseteq R_1$ be a vector space of $\dim(W') \leq a$ such that $V_1 \subseteq W'$. Let $Q_2 \in \mathbb{K}[V] \cap R_2$ be a form such that $s(F - Q_2) < r$ and $F - Q_2 \in \mathbb{K}[W']$. Let $W_2 \in \mathcal{W}_a(F - Q_2)$ be the unique vector space of minimal dimension given by applying Proposition 3.10.2 to $F - Q_2$. Then $W_2 \subseteq W'$ by Proposition 3.10.2. It is enough to show that $W_2 + V_1 = W$.

First, we will show that $Q_1 - Q_2 \in \mathbb{K}[V_1]$. Let $P := Q_1 - Q_2 \in \mathbb{K}[V]$. Note that $s(P) < s + r \leq 2r$. We may choose a lift $\widetilde{P} \in S$ such that $s(\widetilde{P}) \leq s(P) < 2r$. Let $P = G + H$ where $G \in \mathbb{K}[V_1]$ and $H \in V_2$. Now, let us show that $H = 0$. Suppose that $H \neq 0$. Then for any lift $\widetilde{H}$ in S, there exists $H' \in U_2$ such that $\widetilde{H} + \widetilde{G} - \widetilde{P} - H' \in (U_1)$. Hence we have $s(\widetilde{H} - H') < 2r + \dim(V_1) + \dim(U_1)$. This is a contradiction, since V is Γ-lifted strong and $\Gamma_2(\delta_U + \delta_V) \geq 2r + \dim(V_1) + \dim(U_1)$. Therefore we must have $Q_1 - Q_2 \in \mathbb{K}[V_1]$.

Now we have $F - Q_2 = (F - Q_1) + (Q_1 - Q_2) \in \mathbb{K}[W_1 + V_1] = \mathbb{K}[W]$. Since $\dim(W) \leq a$, we have that $W \in \mathcal{W}_q(F - Q_2)$. Therefore by Proposition 3.10.2, we have $W_2 \subseteq W$. Similarly, we have $F - Q_1 = (F - Q_2) + (Q_2 - Q_1) \in \mathbb{K}[W_2 + V_1]$. Now $\dim(W_2 + V_1) \leq a$ and hence $W_2 + V_1 \in \mathcal{W}_a(F - Q_1)$. Again, by minimality of W_1, we have that $W = W_1 + V_1 \subseteq W_2 + V_1$. Hence $W = W_2 + V_1 \subseteq W'$ as desired.

If W' is of minimal dimension satisfying property (1) and (2), then $\dim(W') \leq \dim(W)$. Hence we must have $W' = W$, as $W \subseteq W'$ by (4) above. Moreover, let $Q_2 \in \mathbb{K}[V] \cap R_2$ be a form such that $s(F - Q_2) = \text{ps-dist}(F, V) - 1 = s - 1$ and $F - Q_2 \in \mathbb{K}[W']$. Let $W_2 \in \mathcal{W}_a(F - Q_2)$ be the unique vector space of minimal dimension given by applying Proposition 3.10.2 to $F - Q_2$. Then the argument above shows that $W_2 + V_1 = W$. Hence $F - Q_2 \in \mathbb{K}[W]$. □

We now define the notion of relative space of essential variables in quotient rings.

Definition 3.10.9 (Relative space of essential variables)**.** Let $\Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Suppose that $\Lambda_i(\delta_U) \geq B_i \circ t_{(2a,0)}(\delta_U)$ and $\Gamma_2(\delta_U + \delta_V) \geq a + \dim(U_1)$ for some $a \in \mathbb{N}$.

Quadratics. For $F \in R_2$ such that ps-dist(V, F) satisfies

$$2 \cdot \text{ps-dist}(F, V) + \dim(V_1) \leq a,$$

we will define the space of essential variables of F relative to V in R and we denote it by $\mathbb{L}_{R,V}(a, F)$.

Let $W \subseteq R_1$ be the unique vector space of minimal dimension provided by applying Proposition 3.10.8. We will denote

$$\mathbb{L}_R(a, V, F) := W.$$

Recall that $V_1 \subseteq W$. We define the relative space of essential variables $\mathbb{L}_{R,V}(a, F)$ as the vector space

$$\mathbb{L}_{R,V}(a, F) = \mathbb{L}_R(a, V, F)/V_1.$$

Linear forms. For $F \in R_1$, we define the relative spaces as follows:

$$\mathbb{L}_R(a, V, F) = V_1 + \operatorname{span}_{\mathbb{K}}\{F\},$$

and

$$\mathbb{L}_{R,V}(a, F) = \mathbb{L}_R(a, V, F)/V_1.$$

We will omit R from the notation and write $\mathbb{L}(a, V, F)$ and $\mathbb{L}_V(a, F)$, when the ambient ring is evident from the context. ◊

We note the following properties of the relative space of essential variables.

Proposition 3.10.10. *Let $\Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Let $F \in R_2$. Let $a \in \mathbb{N}$ such that $a \geq 2 \cdot \text{ps-dist}(F, V) + \dim(V_1)$. Suppose that $\Lambda_i(\delta_U) \geq B_i \circ t_{(2a,0)}(\delta_U)$ and $\Gamma_2(\delta_U + \delta_V) \geq a + \dim(U_1)$. The following holds.*

1. *We have $\operatorname{span}_{\mathbb{K}}\{V, F\} \subseteq \mathbb{K}[V_2, \mathbb{L}_R(a, V, F)]$ and*

$$\dim(\mathbb{L}_{R,V}(a, F)) \leq \dim(\mathbb{L}_R(a, V, F)) \leq a.$$

2. *Let V' be a $\Gamma \circ t_{(0,1)}$-lifted strong vector space such that $V \subseteq \mathbb{K}[V']$. Suppose that $a \geq 2 \cdot \text{ps-dist}(F, V') + \dim(V_1')$. Then we have*

$$\dim(\mathbb{L}_{R,V'}(a, F)) \leq \dim(\mathbb{L}_{R,V}(a, F)).$$

Proof. (1) We have $V_1 \subseteq \mathbb{L}_R(a, V, F)$ by Proposition 3.10.8. Therefore we can deduce

$$\mathbb{K}[V, \mathbb{L}_R(a, V, F)] = \mathbb{K}[V_2, \mathbb{L}_R(a, V, F)].$$

Moreover, there exists $Q \in \mathbb{K}[V]$ such that $F - Q \in \mathbb{K}[\mathbb{L}_r(a, V, F)]$. Hence $F \in \mathbb{K}[V_2, \mathbb{L}_R(a, V, F)]$. Moreover, we have the dimension upper bound by part (3) of Proposition 3.10.8.

(2) First let us note that $\mathbb{L}_R(a, V', F)$ is well-defined. Since Γ is an ascending function, we have $\Gamma_2(\delta_U + \delta_{V'}) \geq \Gamma_2(\delta_U + \delta_V) \geq a + \dim(U_1)$. Hence the conditions in Definition 3.10.9 are satisfied. Let $v_1, \cdots, v_m$ be a basis of V_1 and $v_1, \cdots, v_m, w_1, \cdots, w_n$ be an extension to a basis of $\mathbb{L}_R(a, V, F)$. We have $n = \dim(\mathbb{L}_{R,V}(a, F))$. Now there exists $Q \in \mathbb{K}[V]$ such that $F - Q \in \mathbb{K}[\mathbb{L}_R(a, V, F)]$ and $s(F - Q) = \text{ps-dist}(F, V) - 1$. Consider $W' := V_1' + \operatorname{span}_{\mathbb{K}}\{w_1, \cdots, w_n\}$. We have $\dim(W') \leq a$ and $V_1' \subseteq W'$. Moreover, we have $F - Q \in \mathbb{K}[W']$ and $s(F - Q) = \text{ps-dist}(F, V) - 1$. Set $r := \text{ps-dist}(F, V)$. Then we have $\text{ps-dist}(F, V') \leq r$ by Proposition 3.10.6. Hence we can apply Proposition 3.10.8 to the vector space V' and the form F to obtain that $\mathbb{L}_R(a, V', F) \subseteq W'$. Hence $\mathbb{L}_{R,V'}(a, F) \subseteq W'/V_1'$ and

$$\dim(\mathbb{L}_{R,V'}(a, F)) \leq \dim(W'/V_1') \leq n = \dim(\mathbb{L}_{R,V}(a, F)).$$ □

Definition 3.10.11 (Lin-separated forms)**.** Let $\Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Let $F \in R_2$. Let $a \in \mathbb{N}$. Suppose that $\Lambda_i(\delta_U) \geq B \circ t_{(2a,0)}(\delta_U)$ and $\Gamma_2(\delta_U + \delta_V) \geq a + \dim(U_1)$.

Let $\mathcal{F} \subseteq R_{\leq 2}$ be a finite set of forms such that for any $F \in \mathcal{F} \cap R_2$ we have $2 \cdot \text{ps-dist}(F, V) + \dim(V_1) \leq a$. We will say that $\mathcal{F}$ is *lin-separated* modulo V if $\mathcal{F} \cap \mathbb{K}[V] = \emptyset$ and

$$\dim(\sum_{F \in \mathcal{F}} \mathbb{L}_{R,V}(a, F)) = \sum_{F \in \mathcal{F}} \dim(\mathbb{L}_{R,V}(a, F)).$$

$\Diamond$

Proposition 3.10.12. *Let $\Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Let $F \in R_2$. Let $a \in \mathbb{N}$. Suppose that $\Lambda_i(\delta_U) \geq B_i \circ t_{(2a,0)}(\delta_U)$ and $\Gamma_2(\delta_U + \delta_V) \geq a + \dim(U_1)$. The following holds.*

1. *Let $\mathcal{F} \subseteq R_{\leq 2}$ be a finite set of forms that is lin-separated modulo V. Then any subset $\mathcal{G} \subseteq \mathcal{F}$ is also lin-separated modulo V.*

2. *Let $F, G \in R_{\leq 2} \setminus \mathbb{K}[V]$ be such $\mathbb{L}_R(a, V, F)$ and $\mathbb{L}_R(a, V, G)$ are well-defined. Then F, G are lin-separated modulo V iff $\mathbb{L}_R(a, V, F) \cap \mathbb{L}_R(a, V, G) = V_1$.*

Proof. (1) Let $\mathcal{F} = \{F_1, \cdots, F_m\}$. Without loss of generality, we may assume that $\mathcal{G} = \{F_1, \cdots, F_n\}$ for some $n \leq m$. Suppose that $\mathcal{G}$ is not lin-separated modulo V. Then we have

$$\dim(\sum_{i=1}^{n} \mathbb{L}_{R,V}(a, F_i)) < \sum_{i=1}^{n} \dim(\mathbb{L}_{R,V}(a, F_i)).$$

Therefore, we have

$$\dim(\sum_{i=1}^{m} \mathbb{L}_{R,V}(a, F_i)) \leq \dim(\sum_{i=1}^{n} \mathbb{L}_{R,V}(a, F_i)) + \dim(\sum_{i=n+1}^{m} \mathbb{L}_{R,V}(a, F_i)) < \sum_{i=1}^{m} \dim(\mathbb{L}_{R,V}(a, F_i)).$$

This is a contradiction as $\mathcal{F}$ is lin-separated modulo V.

(2) Let $W_1 = \mathbb{L}_{R,V}(a, F)$ and $W_2 = \mathbb{L}_{R,V}(a, G)$. We have $\dim(W_1) + \dim(W_2) = \dim(W_1 + W_2)$ iff $W_1 \cap W_2 = (0)$ in R_1/V_1. By lifting back to R_1, we obtain that F, G are lin-separated modulo V iff $\mathbb{L}_R(a, V, F) \cap \mathbb{L}_R(a, V, G) = V_1$.

$\square$

We note the following property of lin-separated forms under quotients.

Lemma 3.10.13. *Let $\Gamma, \Lambda : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong. Let $a \in \mathbb{N}$. Suppose that $\Lambda_i(\delta_U) \geq B_i \circ t_{(2a,0)}(\delta_U)$ and $\Gamma_i(\delta_U + \delta_V) \geq B_i \circ t_{(2a,0)}(\delta_U + \delta_V)$.*

Let $F_1, F_2 \in R_2$ be such that $F_1, F_2 \notin (V)$. Suppose that $2\,\text{ps-dist}(F_i, V) + \dim(V_1) \leq a$, i.e. $\mathbb{L}_{R,V}(a, F_i)$ is well-defined. Consider the quotient ring $R' = R/(V)$ and let F_1', F_2' be the images of F_1, F_2 in R'. Then the following holds.

1. *The relative spaces of essential variables $\mathbb{L}_{R',(0)}(a, F_i')$ are well-defined.*

2. *In the vector space $R_1' \simeq R_1/V_1$, we have $\mathbb{L}_{R',(0)}(a, F_i') \subseteq \mathbb{L}_{R,V}(a, F_i')$.*

3. *Suppose that F_1, F_2 are lin-separated modulo V in R, then F_1', F_2' are lin-separated modulo (0) in R'.*

4. *Suppose that F_1, F_2 are lin-separated modulo V. Let $P = F_1 + F_2 + H \in R$ where $H \in \mathbb{K}[V]$ such that $\mathbb{L}_{R,V}(a, P)$ is well-defined. Let $P' \in R'$ be the image of P. Then we have $\mathbb{L}_{R',(0)}(a, P') = \mathbb{L}_{R',(0)}(a, F_1') + \mathbb{L}_{R',(0)}(a, F_2')$. Moreover, we have*

$$\dim(\mathbb{L}_{R',(0)}(a, P')) = \dim(\mathbb{L}_{R',(0)}(a, F_1')) + \dim(\mathbb{L}_{R,V}(a, F_2')).$$

Proof. (1) Let us check that the forms F_i', the vector space $V' = (0)$ and the ring R' satisfy the conditions in Definition 3.10.9. Let $\widetilde{V} \subseteq S$ be the vector space generated by lifts of a basis of V. Then $R' = S/(U + \widetilde{V})$. Now note that $U + \widetilde{V}$ is $\Gamma \circ t_{(0,1)}$-strong, as V is $\Gamma \circ t_{(0,1)}$-lifted strong. By assumption, we have $\Gamma_i(\delta_U + \delta_V) \geq B \circ t_{(2a,0)}(\delta_U + \delta_V)$. Now, $B \circ t_{(2a,0)}(\delta_U + \delta_V) \geq a + \dim(U_1) + \dim(V_1)$. Moreover, we have ps-dist$(F_i', (0)) \leq$ ps-dist(F, V). Hence in the quotient ring R' we have ps-dist$(F_i', (0)) \leq a$. Hence the spaces $\mathbb{L}_{R',(0)}(a, F_i')$ are well-defined.

(2) Suppose $Q_i \in \mathbb{K}[V]$ is such that $s(F_i - Q_i) =$ ps-dist$(F_i, V) - 1$. Let $W_i \in \mathcal{W}_a(F_i - Q_i)$ be the unique vector space of minimal dimension such that $F_i - Q_i \in \mathbb{K}[W_i]$. Recall that $\mathbb{L}_R(a, V, F_i) = W_i + V_1$ and $\mathbb{L}_{R,V}(a, F_i) = (W_i + V_1)/V_1$. Let $\phi : R_1 \to R_1' = R_1/V_1$ be the quotient map. Let $G_i = F_i - Q_i$ and $G_i' \in R'$ be the image. Then $G_i' = F_i'$ in R'. Moreover, $\phi(\mathbb{L}_R(a, V, F_i)) = \mathbb{L}_{R,V}(a, F_i) \subseteq R_1/V_1 = R_1'$ is a vector space of linear forms with dimension at most a and such that $F_i' \in \mathbb{K}[\phi(\mathbb{L}_{R,V}(a, F_i))]$. Therefore, by minimality, we have $\mathbb{L}_{R',(0)}(a, F_i') \subseteq \mathbb{L}_{R,V}(a, F_i)$.

(3) Follows from part (2). Indeed, we have $\mathbb{L}_{R',(0)}(a, F_1) \cap \mathbb{L}_{R',(0)}(F_2) \subseteq \mathbb{L}_{R,V}(a, F_1) \cap \mathbb{L}_{R,V}(F_2) = (0)$.

(4) Note that $F_1', F_2' \neq 0$ as $F_1, F_2 \notin (V)$. Let P' be the image of P in R'. Note that $P' = F_1' + F_2'$. Recall that $\dim(\mathbb{L}_{R'}(a, F_i')) \leq a$ and $\mathbb{L}_{R'}(a, F_1') \cap \mathbb{L}_{R'}(a, F_2') = (0)$. Let $x_1, \cdots, x_m \in R_1'$ be a basis of $\mathbb{L}_{R'}(a, F_1')$. Similarly, let $y_1, \cdots, y_n \in R_1'$ be a basis of $\mathbb{L}_{R'}(a, F_2')$. Since $\Gamma_i(\delta_U + \delta_V) \geq B_i \circ t_{(2a,0)}(\delta_U + \delta_V)$, we have that $x_1, \cdots, x_m, y_1, \cdots, y_n \in R'$ is a prime sequence. Therefore the subalgebra $A = \mathbb{K}[x_1, \cdots, x_m, y_1, \cdots, y_n] \subseteq R'$ is isomorphic to a polynomial ring in $m + n$ variables. Now $P' = F - 1' + F_2'$ and $F_1' \in \mathbb{K}[x_1, \cdots, x_m]$ and $F_2' \in \mathbb{K}[y_1, \cdots, y_n]$. Therefore, by applying [PS20, Fact 2.17] repeatedly, we obtain that $\mathbb{L}_{R'}(a, P') = \mathbb{L}_{R'}(a, F_1') + \mathbb{L}_{R'}(a, F_2')$. Moreover, the dimensions add up as $\mathbb{L}_{R'}(a, F_1') \cap \mathbb{L}_{R'}(a, F_2') = (0)$

□

The following result shows that, under suitable conditions, the relative space of essential variables $\mathbb{L}_{R,V}(a, V, F)$ is well-defined if $\text{span}_{\mathbb{K}}\{V, F\}$ is not μ-lifted strong.

Corollary 3.10.14. *Let $\mu, \Gamma, \Lambda, f : \mathbb{N}^2 \to \mathbb{N}^2$ be ascending functions. Suppose that the functions μ, Γ, Λ, f satisfy the following conditions:*

1. *$\Gamma_i(\delta) \geq 2\mu_i \circ t_{(0,1)}(\delta) + \delta_1$ for all $\delta \in \mathbb{N}$, and $i = 1, 2$.*

2. *$\Lambda \geq B \circ t_{(2\mu_2 + f_1, 0)} \circ t_f \circ t_{(0,1)}$.*

Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong such that $\dim(V_i) < f_i(\delta_U)$. Let $F \in R_2$ be such that $\text{span}_{\mathbb{K}}\{V, F\}$ is not μ-lifted strong. Then the spaces $\mathbb{L}_R(a, V, F)$ and $\mathbb{L}_{R,V}(a, F)$ are well-defined for $a = 2\mu_2(\delta_U + \delta_V + (0, 1)) + f_1(\delta_U)$.

Proof. We have $\Gamma \geq \mu$. Hence by Proposition 3.10.7, we have

$$\text{ps-dist}(V, F) \leq 2\mu_2(\delta_U + \delta_V + (0, 1)).$$

Therefore $a = 2\mu_2(\delta_U + \delta_V + (0,1)) + f_1(\delta_U) \geq 2\,\text{ps-dist}(V,F) + \dim(V_1)$. Now we have

$$\Lambda_i(\delta_U) \geq B_i \circ t_{(2\mu_2+f_1,0)} \circ t_f \circ t_{(0,1)}(\delta_U) \geq B \circ t_{(2a,0)}(\delta_U).$$

Moreover, we have $\Gamma_2(\delta_U + \delta_V) \geq 2\mu_2(\delta_U + \delta_V + (0,1)) + \dim(U_1) + \dim(V_1) \geq a + \dim(U_1)$. Hence Proposition 3.10.8 and Definition 3.10.9 apply and the spaces $\mathbb{L}_R(a,V,F)$ and $\mathbb{L}_{R,V}(a,F)$ are well-defined. □

3.10.2 Reduction to the low strength case

In this subsection, we prove the following technical lemma.

Lemma 3.10.15. *Suppose $\mu_{k-1} : \mathbb{N}^2 \to \mathbb{N}^2$ and $\xi_{k-1} : \mathbb{N}^2 \to \mathbb{N}$ are fixed functions. Define the function $\xi_k^1 : \mathbb{N}^2 \to \mathbb{N}$ as $\xi_k^1(\delta) := \xi_{k-1}\big(\big(C_{\mu_{k-1}} \circ t_1\big)(\delta)\big) + 1$. Let $\mu_k^{far} := h_{\mu_{k-1}} \circ t_1$. Here, $C_{\mu_{k-1}}, h_{\mu_{k-1}}$ are the functions defined in Lemma 3.6.3. Define the function $M(\delta) : \mathbb{N}^2 \to \mathbb{N}$ as*

$$M(\delta) := \max_{\delta' \in \mathbb{N}^2, \|\delta'\| = \binom{k}{2}\xi_k^1(\delta)} \mu_{k-1}(\delta + \delta' + (0, k^2\xi_k^1(\delta))).$$

For every vector space U that is μ_k^{far}-strong, and every $\Sigma^k\Pi\Sigma\Pi^2$ simple minimal identity $\mathcal{C} = \sum_{i=1}^k T_i$ in $R := S/(U)$, one of the following holds:

1. *for some i, j, the symmetric difference $T_i \Delta T_j$ satisfies $\dim \text{span}_{\mathbb{K}}\{T_i \Delta T_j\} \leq k^2 \xi_k^1(\delta_U)$.*
2. *there is a form $H \in \mathcal{C}$ that is μ_{k-1}-lifted strong, such that the image of $\mathcal{C}$ in $R/(H)$ has pairwise rank at least $\xi_{k-1}(\delta_U + (0,1))$.*
3. *let $b := k^2\big(\xi_k^1(\delta_U) + 1\big)$, there exists a graded vector space V generated by forms in $\mathcal{C}$ with $\dim V \leq 4^{b+1} + b^3 + b + 48$ such that every quadratic form $Q \in \mathcal{C}$ is $3^b M$-close to V.*

The rest of this subsection is devoted to the proof of this lemma. Fix a dimension sequence $\delta \in \mathbb{N}^2$. Let U be a vector space with dimension sequence $\delta_U := \delta$ that is μ_k^{far}-strong. Let $\mathcal{C}$ be any simple and minimal $\Sigma^k\Pi\Sigma\Pi^2$ identity in $R := S/(U)$

We consider all vector spaces U' obtained by adding a single form to U, and making the result vector space μ_{k-1}-strong using Lemma 3.6.3 while preserving U. Such spaces exist since $\mu_k^{\text{far}} \geq h_{\mu_{k-1}} \circ t_1$. Given δ_U, we know by Lemma 3.6.3 that the dimension sequence of any U' obtained above is bounded coordinate wise by $(C_{\mu_{k-1}} \circ t_1)(\delta_U)$. By definition, $\xi_k^1(\delta_U)$ is greater than $\xi_{k-1}(\delta_{U'})$ for any such U'.

Constructing a core: Start by picking a core set of forms as follows. Go over the $\binom{k}{2}$ many symmetric differences $T_i \Delta T_j$, and pick $\xi_k^1(\delta_U)$ many forms from each symmetric difference, such that the set of all forms picked is linearly independent. For each symmetric difference, we first attempt to pick quadratic forms, and only pick linear forms if we already have a set that spans the quadratics in $T_i \Delta T_j$. If we cannot pick such a core, then some symmetric difference must satisfy $\dim \text{span}_{\mathbb{K}}\{T_i \Delta T_j\} \leq k^2 \xi_k^1(\delta_U)$, which gives us Case 1 of Lemma 3.10.15. Hence, we can assume that the above process succeeds, and we have a core consisting of $\binom{k}{2}\xi_k^1(\delta_U)$ linearly independent forms. Let V^{core} be the span of the forms that are in the core. We have $\dim V^{\text{core}} = \binom{k}{2}\xi_k^1(\delta_U)$. Let $M := M(\delta_U)$ for brevity, where $M(\cdot)$ is the function from the statement of Lemma 3.10.15.

Strong forms outside V^{core}: Let $\mathcal{B} \subset \mathcal{C}$ be the set of forms $Q \in \mathcal{C}$ such that $\text{span}_{\mathbb{K}}\{Q\}$ is μ_{k-1}-lifted strong with respect to U, and that do not belong to V^{core}. Suppose $P \in \mathcal{C} \setminus \mathcal{B}$. Then either $P \in V^{\text{core}}$, or $\text{span}_{\mathbb{K}}\{P\}$ is not μ_{k-1}-lifted strong. The latter implies that $s(P) \leq \mu_{k-1}(\delta_U + (0,1))$,

by [OS24, Proposition 5.14.1]. In particular, $s(P) \leq M$, and therefore P is M-close to the zero vector space, therefore also V^{core}. The forms in $\mathcal{B}$ are exactly those that we have to now control.

Define $\mathcal{B}' \subset W := \mathbb{P}(R_2/V_2^{\mathrm{core}})$ to be the image of $\mathcal{B} \subset R_2$ under the vector space quotient $R_2 \to R_2/V_2^{\mathrm{core}}$ followed by projectivisation. By definition, no element of $\mathcal{B}$ is mapped to 0 under the quotient map, however it is possible for different elements of $\mathcal{B}$ to get mapped to linearly dependent elements of the vector space W. We will control the set $\mathcal{B}'$, which will imply control over the set $\mathcal{B}$.

Definition 3.10.16 (Strong span configurations)**.** A set $\mathcal{G}' \subset W := \mathbb{P}(R_2/V_2^{\mathrm{core}})$ is called a (M', b')-strong span configuration if the following holds. For every $h_1, \ldots, h_{b'} \in \mathcal{G}'$ which are linearly independent, either $|\mathrm{span}_{\mathbb{K}}\{h_1, \ldots, h_{b'}\} \cap \mathcal{G}'| > b'$, or there exist preimages $H_1, \ldots, H_{b'} \in R_2$ of $h_1, \ldots, h_{b'}$, and a form $H \in \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_{b'}\}$ such that H is M'-close to V^{core}. $\lozenge$

Note that in the above definition, the choice of preimages does not matter, since the pseudo-distance to V^{core} is unchanged if different preimages are picked. Returning to the set $\mathcal{B}'$, which is the image of $\mathcal{B}$ in W, we consider two cases. Recall that in Case 3 we defined

$$b := k^2\big(\xi_k^1(\delta_U) + 1\big).$$

Either $\mathcal{B}'$ is an (M, b)-strong span configuration, or $\mathcal{B}'$ fails to be an (M, b)-strong configuration. If $\mathcal{B}'$ fails to be an (M, b)-strong configuration, then we can deduce that there exist $H_1, \ldots, H_b \in \mathcal{B}$ with the following properties: The forms $H_1, \ldots, H_b$ are linearly independent, and no form in their span is M-close to V_2^{core}. In particular this implies that no non zero form in the span lies in V_2^{core}. Finally, every form in $\mathcal{B} \cap \mathrm{span}_{\mathbb{K}}\{H_1, \ldots, H_b, V_2^{\mathrm{core}}\}$, actually lies in $\mathrm{span}_{\mathbb{K}}\{H_i, V_2^{\mathrm{core}}\}$ for some i, since the images $h_1, \ldots, h_b$ satisfy $|\mathrm{span}_{\mathbb{K}}\{h_1, \ldots, h_b\} \cap \mathcal{B}'| = b$.

In the rest of this section, we handle these two cases. We will show that given any (M', b')-strong span configuration, we can find a small set of forms $\mathcal{H}$ such that every form in the configuration is close to $\mathrm{span}_{\mathbb{K}}\{\mathcal{H}\} + V^{\mathrm{core}}$, as in Case 3 of Lemma 3.10.15. The exact parameters will be stated below. We also show that if $\mathcal{B}'$ fails to be an (M, b)-strong configuration, that is, $H_1, \ldots, H_b$ with the above properties exist, then one of the forms H_i satisfies Case 2 of Lemma 3.10.15. This will use a Kayal–Saraf type argument. The rest of this section is split into two parts, to handle these two cases. We start with the second case.

A fan-in reduction lemma when $\mathcal{B}'$ fails to be an (M, b)-strong configuration

Let $\mathcal{H} := \{H_1, \ldots, H_b\}$ be the witness to the fact that $\mathcal{B}'$ is not an (M, b)-strong span configuration. In this section, we will show the existence of a form $H \in \mathcal{H}$, such that the image of $\mathcal{C}$ in the ring $R/(H)$ has rank at least $\xi_k^1(\delta_U)$. Each of the elements in $\mathcal{H}$ has lifted strength at least μ_{k-1}. For any such element H, and any lift $\widetilde{H}$ of H in S, the vector space $U + \widetilde{H}$ is μ_{k-1}-strong. This ring is the same as $R/(H)$, and is a Cohen–Macaulay UFD by Corollary 3.6.2. The image of $\mathcal{C}$ in these rings is well defined, as is the notion of pairwise rank for these images. The dimension sequence of the vector space $U + \widetilde{H}$ is $\delta_U + (0, 1)$, and by construction we have $\xi_k^1(\delta_U) > \xi_{k-1}(\delta_U + (0, 1))$. Therefore, the existence of such a H shows that Case 2 of Lemma 3.10.15 is satisfied, as we also show that pairwaise rank remains high (Lemmas 3.10.18 and 3.10.19).

Recall that V^{core} was constructed by picking $\xi_k^1(\delta_U)$ many linearly independent elements from each symmetric difference $T_i \Delta T_j$. Suppose $G_1, \ldots, G_a$ are the forms we picked from $T_1 \Delta T_2$, with $a = \xi_k^1(\delta_U)$. For a form $H \in \mathcal{H}$, we define the notion of H being *core preserving* and *core collapsing* for $T_1 \Delta T_2$ as follows.

Definition 3.10.17 (Core preserving/collapsing)**.** Let $H \in \mathcal{H} \setminus (T_1 \cup T_2)$. We say that H is core preserving for $T_1 \Delta T_2$ if $|\mathrm{span}_{\mathbb{K}}\{G_i, H\} \cap T_1 \Delta T_2| = 1$ for all $i \in [a]$. The form G_i is always in this

span, therefore the condition states that no other form is in the span. We say H is core collapsing for $T_1 \Delta T_2$ if there is $i \in [a]$ such that $|\mathrm{span}_{\mathbb{K}}\{G_i, H\} \cap T_1 \Delta T_2| \geq 2$. $\Diamond$

The notion extends naturally to every symmetric difference. The following lemma justifies the name core preserving.

Lemma 3.10.18. *Suppose $H \in \mathcal{H} \setminus (T_1 \cup T_2)$ is core preserving for $T_1 \Delta T_2$. The image of T_1, T_2 in $R' := R/(H)$ has pairwise rank at least $\xi_k^1(\delta_U)$.*

Proof. We claim that $G_1, \ldots, G_a$ are witnesses to the pairwise strength of T_1, T_2. Condition $H \notin T_1 \cup T_2$ ensures these gates survive in R'. The forms $G_1, \ldots, G_a$ remain linearly independent in R', otherwise $H \in V^{\mathrm{core}}$.

Fix G_1, and suppose without loss of generality that $G_1 \in T_1 \setminus T_2$. The form G_1 is either a linear form, or a quadratic form. If G_1 is a quadratic form, then the image of G_1 in R' remains irreducible: indeed if it was reducible, say $G_1 = \ell_1 \ell_2$ in R', then $H = \alpha G_1 + \ell_1 \ell_2$ in R, contradicting the fact that H is not M-close to V^{core}.

Suppose that G_1 is no longer in the symmetric difference in R'. It must be that for some form $F \in T_2 \setminus T_1$, the image of F in R' has a non-trivial gcd with the image of G_1. We consider two cases. First suppose G_1 is a quadratic form (i.e. $G_1 \in R_2$). Since the image of G_1 is irreducible, it must be that $F \in (G_1, H)$ in the ring R. Since F, G_1, H are quadratics, this implies $F \in \mathrm{span}_{\mathbb{K}}\{G_1, H\}$, contradicting the fact that H is core preserving.

Suppose now that $G_1 \in R_1$. If F were also linear, then G_1, F are associate in R, contradicting the choice of G_1, F (namely that they lie in the symmetric differences). Therefore F must be a quadratic, and we must have $F = H + G_1 \ell$ in R. It must also hold that $F \notin V_2^{\mathrm{core}}$, otherwise H will be 1-close to V^{core}. However this contradicts the construction of the core: we always first pick as many quadratics as possible before picking linear forms, but here we show that we picked G_1 as part of the core but did not pick F. $\square$

We now consider the core collapsing case.

Lemma 3.10.19. *Suppose $H \in \mathcal{H} \setminus (T_1 \cup T_2)$. Assume that there are c other elements in $\mathcal{H} \setminus (T_1 \cup T_2)$ that are core collapsing for $T_1 \Delta T_2$. Then the image of T_1, T_2 in $R' := R/(H)$ has pairwise rank at least c.*

Proof. Without loss of generality, we assume that the c elements that are core collapsing other than H are $H_1, \ldots, H_c$. The form H_1 is core collapsing for $T_1 \Delta T_2$. By definition, this means that there is a form G_{i_1} among $G_1, \ldots, G_a$ such that $|\mathrm{span}_{\mathbb{K}}\{G_{i_1}, H_1\} \cap T_1 \Delta T_2| > 1$. If G_{i_1} is a linear form, then the only homogeneous forms in $\mathrm{span}_{\mathbb{K}}\{G_{i_1}, H_1\}$ are multiples of G_{i_1} and multiples of H_1, neither of which are in $T_1 \Delta T_2$. Therefore, in the above setting, it must be that G_{i_1} is a quadratic form. Let F_1 be any form in the above intersection other than G_{i_1}, so $F_1 \in T_1 \Delta T_2$ and $F_1 \in \mathrm{span}_{\mathbb{K}}\{G_{i_1}, H_1\}$. Similarly we define G_{i_j} and F_j for all $j = 2, \ldots, c$.

We claim that $F_1, \ldots, F_c$ are witnesses to the pairwise rank of T_1, T_2 in R'. First observe that F_1 is irreducible in R': if not, and if $F_1 = \ell_1 \ell_2$ in R', then $F_1 = \ell_1 \ell_2 + \alpha H$ in R with $\alpha \neq 0$. Further we have $F_1 \in \mathrm{span}_{\mathbb{K}}\{G_{i_1}, H_1\}$, which would imply that an element in the span of $H, H_1, V^{\mathrm{core}}$ has strength 1, contradicting the fact that no form in the span of $H_1, \ldots, H_b$ is M-close to V^{core}.

Since F_1 is a quadratic form in R', the only way it is no longer in the symmetric difference of $T_1 \Delta T_2$ is if there is another form $F' \in T_1 \Delta T_2$ such that F', F_1 are associate in R'. This implies $F' = F_1 + \alpha H$ in R, with $\alpha \neq 0$. Further we have $F_1 \in \mathrm{span}_{\mathbb{K}}\{G_{i_1}, H_1\}$, say $F_1 = \alpha' G_{i_1} + \alpha'' H_1$ with $\alpha', \alpha'' \neq 0$. We have $F' = \alpha' G_{i_1} + \alpha'' H_1 + \alpha H$. Now if $F' \notin \mathcal{B}$, then $\alpha'' H_1 + \alpha H$ is M-close to V^{core}, since $s(F') < \mu_{k-1}(\delta_U + (0,1)) \leq M$. This cannot hold, because of our choice of $H_1, \ldots, H_b$

(as witnesses to the fact that $\mathcal{B}'$ is not an (M,b)-strong span configuration). On the other hand, the only forms in $\mathrm{span}_{\mathbb{K}}\{H_1,\dots,H_b,V^{\mathrm{core}}\}\cap\mathcal{B}$ are those that lie in $\mathrm{span}_{\mathbb{K}}\{H_i,V^{\mathrm{core}}\}$ for some i (again, as they witness to the fact that $\mathcal{B}'$ is not an (M,b)-strong span configuration). However if F' had this structure, it would contradict the fact that no form in $\mathrm{span}_{\mathbb{K}}\{H_1,\dots,H_b\}$ lies in V^{core}. In either case we have a contradiction, therefore such an F' does not exist.

Finally, by the same token, we observe that $F_1,\dots,F_c$ are linearly independent in R'. Any linear dependence between $F_1,\dots,F_c$ would lift to a linear dependence between $H_1,\dots,H_b$ and an element of V^{core}, which contradicts the choice of $H_1,\dots,H_b$. Therefore the pairwise rank of T_1,T_2 in R' is at least c. □

While Lemma 3.10.18 and Lemma 3.10.19 are stated for $T_1\Delta T_2$, the same results hold for other symmetric differences $T_i\Delta T_j$. Recall that our objective was to find a form H such that the pairwise rank of $\mathcal{C}$ in $R/(H)$ is at least ξ^1_k. Given the above two lemmas, a sufficient condition for H to satisfy this property is that for every $i\neq j\in[k]$, either H is core preserving for $T_i\Delta T_j$, or there are $\xi^1_k(\delta_U)$ many forms other than H that are core collapsing for $T_i\Delta T_j$, or H belongs to one of T_i,T_j.

The following counting argument shows that such a $H\in\mathcal{H}$ exists. For each $T_i\Delta T_j$ let $\mathcal{H}^{i,j}\subset\mathcal{H}$ be the set of elements H such that H is core collapsing for $T_i\Delta T_j$. If $\left|\mathcal{H}^{i,j}\right|\geq\xi^1_k(\delta_U)+1$, then for any $H\in\mathcal{H}$, either one of T_i,T_j vanishes in $R/(H)$ (which happens if H belongs to one of T_i,T_j), or their pairwise rank is at least $\xi^1_k(\delta_U)$, by either Lemma 3.10.18 or Lemma 3.10.19.

Let $\mathcal{H}'$ be the union of $\mathcal{H}^{i,j}$ over those i,j such that $\left|\mathcal{H}^{i,j}\right|\leq\xi^1_k(\delta_U)$. The size of $\mathcal{H}'$ is at most $\binom{k}{2}\xi^1_k(\delta_U)$, and by the choice of b, we can pick $H\in\mathcal{H}\setminus\mathcal{H}'$. We claim that this H has the required property. The pairs i,j with $\mathcal{H}^{i,j}\geq\xi^1_k(\delta_U)+1$ have already been handled. If i,j is such that $\mathcal{H}^{i,j}\leq\xi^1_k(\delta_U)$, then by construction either H belongs to one of T_i,T_j or H is core preserving for $T_i\Delta T_j$.

Controlling $\mathcal{B}'$ when it is an (M,b)-strong configuration

In this subsection, we show that strong span configurations, defined in Definition 3.10.16, can be controlled. This is the last step in completing the proof of Lemma 3.10.15. Specifically, we show the following result.

Theorem 3.10.20. *Suppose $\mathcal{G}'\subset W:=\mathbb{P}(R_2/V_2^{core})$ is an (M',b')-strong span configuration. There exists a subset $\mathcal{H}'\subset\mathcal{G}'$ of size at most $4^{b'+1}+b'^3+48$ such that every form in $\mathcal{G}'$ is $3^{b'}M'$-close to $V^{core}+\mathrm{span}_{\mathbb{K}}\{\mathcal{H}'\}$.*

In our setting, we can apply Theorem 3.10.20 to $\mathcal{B}'$ with $b'=b$ and $M'=M$. If we then add the elements of $\mathcal{H}'$ obtained by the above application to V^{core}, we deduce Case 3 of Lemma 3.10.15. To prove Theorem 3.10.20, we define a generalization of (M',b')-strong span configurations. To make the definition and the arguments convenient we establish some notation. First, we use small letters to denote elements of W, and corresponding capital letters to denote preimages in R_2. For example, if $h,h_i\in W$ then H,H_i will denote a choice of preimage of h,h_i. Our arguments and results will not depend on which preimage is chosen. Given a set of elements $h_1,\dots,h_c\in W$, we say that $h_1,\dots,h_c$ are M'-far from V^{core} if $h_1,\dots,h_c$ are linearly independent, and there is no nonzero form $H\in\mathrm{span}_{\mathbb{K}}\{H_1,\dots,H_c\}$ such that H is M'-close to V^{core}. Therefore, if we say $h_1,\dots,h_c$ are not M'-far from V^{core}, we mean either that they are linearly dependent, or that some nonzero element in the span of $H_1,\dots,H_c$ is M'-close to V^{core}. [4] Given a set $\mathcal{H}\subset W$, with

[4] We could have chosen to say $h_1,\dots,h_c$ are M'-close to V^{core} in this situation, as opposed to saying not M'-far. However the former makes it seem as though all the forms are close to V^{core}, while what we want to capture is that some form in the span is close to V^{core}.

$\mathcal{H} := \{h_1, \ldots, h_c\}$, we say that g is M'-close to $\mathcal{H}, V^{\text{core}}$ (or g is M'-close to $h_1, \ldots, h_c, V^{\text{core}}$) if G is M'-close to $\text{span}_{\mathbb{K}}\{H_1, \ldots, H_c\} + V^{\text{core}}$.

Definition 3.10.21 (Fractional strong span configuration)**.** Let $0 < \nu \leq 1$. A set $\mathcal{G}' \subset W := \mathbb{P}(R_2/V_2^{\text{core}})$ is called a (ν, M', b')-strong span configuration if the following holds. For every $h_1, \ldots, h_{b'-1} \in \mathcal{G}'$ which are M'-far from V^{core}, for ν fraction of the remaining elements $h \in \mathcal{G}'$, either $|\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{b'-1}, h\} \cap \mathcal{G}'| > b'$, or h is M-close to $h_1, \ldots, h_{b'-1}, V^{\text{core}}$. $\lozenge$

An (M', b')-strong span configuration is the same as a $(1, M', b')$-strong span configuration. As a warm up, we show how $(\nu, M', 2)$-span configurations can be controlled. This will also act as the base case in our induction arguments.

Lemma 3.10.22. *Suppose $\mathcal{G}'$ is a $(\nu, 2, M')$-strong span configuration such that no $0 \neq g \in \mathcal{G}$ is M'-close to V^{core}. Then there is a set $\mathcal{H} \subset W$ of size at most $8/\nu$, and a set $\mathcal{G}_c$ with $|\mathcal{G}_c| \geq (\nu/2)|\mathcal{G}'|$ such that for all $g \in \mathcal{G}_c$, the element g is M'-close to $\mathcal{H}, V^{core}$.*

Proof. For $g \in \mathcal{G}'$, let $\mathcal{F}_{\text{weak}}(g)$ be the set of forms $g' \in \mathcal{G}'$ that are M'-close to g, V^{core}. If for some $g \in \mathcal{G}'$ we have $|\mathcal{F}_{\text{weak}}(g)| \geq \nu/2|\mathcal{G}'|$ then we pick $\mathcal{H} = \{g\}$ and we are done with $\mathcal{G}_c = \mathcal{F}_{\text{weak}}(g)$. Therefore we assume that for all $g \in \mathcal{G}'$, we have $|\mathcal{F}_{\text{weak}}(g)| < \nu/2|\mathcal{G}'|$

For $g \in \mathcal{G}'$, let $\mathcal{F}_{\text{span}}(g)$ be the set of forms $h \in \mathcal{G}' \setminus \mathcal{F}_{\text{weak}}(g)$ such that $|\text{span}_{\mathbb{K}}\{g, h\} \cap \mathcal{G}'| \geq 3$. We have $|\mathcal{F}_{\text{span}}(g)| \geq \nu/2|\mathcal{G}'|$ for all $g \in \mathcal{G}'$, since $|\mathcal{F}_{\text{weak}}(g)| \leq \nu/2|\mathcal{G}'|$. This shows that the set $\mathcal{G}'$ is a $\nu/2$-linear SG configuration. By [DGOS18, Theorem 1.6] $\dim \text{span}_{\mathbb{K}}\{\mathcal{G}'\} \leq 8/\nu$ and thus it has a basis consisting of at most $8/\nu$ forms. Picking $\mathcal{H}$ to be such a basis, we are done with $\mathcal{G}_c = \mathcal{G}'$. $\square$

Note that if a form g is M'-close to V^{core} in a $(\nu, 2, M')$-strong span configuration, then every other form h automatically satisfies a relation with g, since the second condition in the definition of $(\nu, 2, M')$-strong span configurations holds. This is why when discussing fractional configurations, we focus on the forms that are not M'-close.

Continuing our warm up, we now show how $(1, 2, M')$-strong span configurations can be controlled. This will also act as a base case in our main proof.

Lemma 3.10.23. *Suppose $\mathcal{G}'$ is a $(1, 2, M')$-strong span configuration. There is a set $\mathcal{H} \subset \mathcal{G}'$ of size at most 8 such that every $g \in \mathcal{G}'$ is $2M'$-close to $\mathcal{H}, V^{core}$.*

Proof. Let $\mathcal{G}_w$ be the set of forms in $\mathcal{G}'$ that are M'-close to V^{core}. Let $\mathcal{G}_s := \mathcal{G}' \setminus \mathcal{G}_w$. For brevity let $m := |\mathcal{G}_s|$. Suppose $\mathcal{G}_s$ is a 1/2-linear SG configuration. In this case, by [DGOS18, Theorem 1.6] $\dim \text{span}_{\mathbb{K}}\{\mathcal{G}_s\} \leq 8$. Hence, picking $\mathcal{H}$ to be a basis of $\mathcal{G}_s$ of size 8, we are done.

Therefore we can assume that $\mathcal{G}_s$ is not a 1/2-linear SG configuration. Let $h_1 \in \mathcal{G}_s$ be a witness to this fact and let $\mathcal{H} := \{h_1\}$. Let $\mathcal{G}_1$ be those elements that are M'-close to $\mathcal{H}, V^{\text{core}}$, we have $\mathcal{G}_w \subset \mathcal{G}_s$. Since $\mathcal{G}'$ is a $(1, 2, M')$-strong span configuration, this means that for at least $m/2$ elements $g \in \mathcal{G}_s$, namely those with which h_1 does not span anything in $\mathcal{G}_s$, either g is M'-close to h_1, V^{core}, or $|\text{span}_{\mathbb{K}}\{h_1, g\} \cap \mathcal{G}_w| \geq 1$. Observe that if $|\text{span}_{\mathbb{K}}\{h_1, g\} \cap \mathcal{G}_w| \geq 1$ then this also implies that g is M'-close to h_1, V^{core}. Therefore $|\mathcal{G}_1 \cap \mathcal{G}_s| \geq m/2$.

We now have to show that every element $g \in \mathcal{G}' \setminus \mathcal{G}_1$ is $2M'$-close to $\mathcal{H}, V^{\text{core}}$. It suffices to show that every element $g \in \mathcal{G}' \setminus \mathcal{G}_1$ is either in the span of two elements of $\mathcal{G}_1$ or is M'-close to g', V^{core} for some $g' \in \mathcal{G}_1$.

Let g be any element that does not satisfy the second property. Then for every $g' \in \mathcal{G}_1 \cap \mathcal{G}_s$ we have $|\text{span}_{\mathbb{K}}\{g, g'\} \cap \mathcal{G}'| \geq 3$. If for some g', one of the elements in the span (other than g') lies in $\mathcal{G}_1$, then g is in the span of two elements of $\mathcal{G}_1$. If for $g', g'' \in \mathcal{G}_1 \cap \mathcal{G}_s$, the intersection $\text{span}_{\mathbb{K}}\{g, g'\} \cap \text{span}_{\mathbb{K}}\{g, g''\}$ contains an element other than g, then it must be that $g \in \text{span}_{\mathbb{K}}\{g', g''\}$. By the pigeonhole principle, one of the conditions is forced to hold since $|\mathcal{G}_1 \cap \mathcal{G}_s| > m/2$. $\square$

We now turn to the case $b' > 2$. Our argument is inspired by the arguments in [BDYW11], where rank bounds for higher dimensional linear SG configurations are given. In particular, we induct on b'. If the given (ν, b', M')-strong span configuration is a $(\nu', b'-1, M')$-strong configuration for some smaller ν' that we choose, then have bounds by induction. If not then we use a witness to the failure to identify a fractional linear SG configuration, where we can apply the bounds from [DGOS18, Theorem 1.6].

Theorem 3.10.24. *Suppose $\mathcal{G}'$ is a (ν, b', M')-strong span configuration such that no $g \in \mathcal{G}'$ is M'-close to V^{core}. Then there is a set $\mathcal{H}$ of at most $b' + 4^{b'}/\nu$ elements, and a set $\mathcal{G}_c$ with $|\mathcal{G}_c| \geq \left(\nu/4^{b'-1}\right)|\mathcal{G}'|$ such that for all $g \in \mathcal{G}_c$, the element g is M'-close to $\mathcal{H}, V^{core}$.*

Proof. Define $\nu_i := \nu/4^{b'-i}$ for all $i \leq b'$. Let r be the smallest positive integer such that $\mathcal{G}'$ is a not a (ν_r, r, M')-strong span configuration, but $\mathcal{G}'$ is a $(\nu_{r+1}, r+1, M')$-strong span configuration. Note that $\nu_{b'} = \nu$, therefore $1 \leq r \leq b'-1$. Suppose $r = 1$, so $\mathcal{G}'$ is a $(\nu_2, 2, M')$-strong span configuration. Applying Lemma 3.10.22 then gives us the required sets $\mathcal{H}, \mathcal{G}_c$.

Now suppose $r > 1$. In particular, $\mathcal{G}'$ is not a (ν_r, r, M')-strong span configuration. Let $h_1, \ldots, h_{r-1}$ be a witness to this fact. This implies that $h_1, \ldots, h_{r-1}$ are M'-far from V^{core}. This also implies that for at least $1 - \nu_r$ fraction of $g \in \mathcal{G}'$, the element g is M'-far from $h_1, \ldots, h_r, V^{\text{core}}$, and further that $|\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g\} \cap \mathcal{G}'| = r$. In particular, we must have $|\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}\} \cap \mathcal{G}'| = r-1$ (if not, then the condition on the previous line cannot hold for many g). Let $\mathcal{G}_1$ be this set.

For any $g \in \mathcal{G}_1$, define the set $\mathcal{F}_{\text{weak}}(g) \subset \mathcal{G}_1$ to be the set of elements g' such that g' is M'-close to $h_1, \ldots, h_{r-1}, g, V^{\text{core}}$. If for any $g \in \mathcal{G}_1$, we have $|\mathcal{F}_{\text{weak}}(g)| \geq \nu_r|\mathcal{G}'|$, then the required statement holds with $\mathcal{H} := \{h_1, \ldots, h_{r-1}, g\}$ and $\mathcal{G}_c := \mathcal{F}_{\text{weak}}(g)$. Therefore we can assume that $|\mathcal{F}_{\text{weak}}(g)| \leq \nu_r|\mathcal{G}'|$ for all $g \in \mathcal{G}_1$.

For any $g \in \mathcal{G}_1$ define $\mathcal{F}_{\text{span}}(g) \subset \mathcal{G}'$ to be those g' such that $\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g, g'\} > r+1$. Since $\mathcal{G}'$ is a $(\nu_{r+1}, r+1, M')$-strong span configuration, for every $g \in \mathcal{G}_1$ we have

$$\left|\mathcal{F}_{\text{span}}(g) \cup \mathcal{F}_{\text{weak}}(g) \cup \left(\mathcal{G}' \setminus \mathcal{G}_1\right)\right| \geq 4\nu_r|\mathcal{G}|.$$

In particular, under the above assumptions, for every $g \in \mathcal{G}_1$ we have $|\mathcal{F}_{\text{span}}(g)| \geq 2\nu_r|\mathcal{G}'|$.

Let $Y := \text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}\}$. We now show that under the assumptions above, the image of $\mathcal{G}_1$ in $\mathbb{P}(W/Y)$ is a ν_r-linear SG configuration. Since $\mathcal{G}_1$ consists of more than half the elements of $\mathcal{G}$, setting $\mathcal{H}$ to be a basis of $\mathcal{G}_1$ will complete the proof, with $\mathcal{G}_c = \mathcal{G}_1$. Observe first that the above map is injective: indeed if g, g' get mapped to the same element, then $g' \in \text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g\}$, contradicting the definition of $\mathcal{G}_1$. Now for each $g \in \mathcal{G}_1$ we have $|\mathcal{F}_{\text{span}}(g)| \geq 2\nu_r|\mathcal{G}'|$. If we can ensure that for each g, for at least ν_r elements $g' \in \mathcal{F}_{\text{span}}(g)$ we have $|\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g, g'\} \cap \mathcal{G}_1| > r+1$ then we will be done. For this, we simply observe that if $\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g, g'\} \cap \text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g, g''\}$ has an element in $\mathcal{G}' \setminus \mathcal{G}_1$ then $g'' \in \text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}, g, g'\}$. Here we crucially use the fact that $\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{r-1}\} \cap \mathcal{G}'$ only has $r-1$ elements. Therefore, there can be at most $\nu_r|\mathcal{G}'|$ elements in $\mathcal{F}_{\text{span}}(g)$ that only span elements of $\mathcal{G}' \setminus \mathcal{G}_1$ with $h_1, \ldots, h_{r-1}, g$. The remaining elements, of which there are at least $\nu_r|\mathcal{G}'|$ many, are witness to necessary SG condition. □

With these lemmas, we can complete the proof of Theorem 3.10.20, and control $(1, b', M')$-strong span configurations.

Proof of Theorem 3.10.20. Let $\mathcal{G}'_w$ be the elements in $\mathcal{G}'$ that are M'-close to V^{core}. Let $\mathcal{G}'_s := \mathcal{G}' \setminus \mathcal{G}'_w$. Let $m := |\mathcal{G}'_s|$. The set $\mathcal{G}'_s$ is itself a $(1, b', M')$-strong span configuration. Indeed for some choice of $h_1, \ldots, h_{b'-1} \in \mathcal{G}'_s$ that is M'-far from V^{core}, and some $h \in \mathcal{G}'_s$, if $\text{span}_{\mathbb{K}}\{h_1, \ldots, h_{b'-1}, h\}$ contains

an element of $\mathcal{G}'_w$, then this implies that h is M'-close to $h_1, \dots, h_{b'-1}, V^{\text{core}}$. Therefore we focus on the set $\mathcal{G}'_s$.

We start by applying Theorem 3.10.24 to the set $\mathcal{G}'_s$ with $\nu = 1$. Let $\mathcal{H}'_1$ be the set of elements guaranteed by Theorem 3.10.24, we have $|\mathcal{H}'_1| \leq b' + 12 \cdot 4^{b'-1}$. Let $\mathcal{G}'_1 \subset \mathcal{G}'_s$ be the set of elements that are M'-close to $\mathcal{H}_1, V^{\text{core}}$. We have $|\mathcal{G}_1| \geq m/4^{b'-1}$.

Define $\mathcal{J}'_1 := \mathcal{G}'_s \backslash \mathcal{G}'_1$, these are the images of the forms that we still have to control. We pick a set of forms $f_1, \dots, f_{b'-1} \in \mathcal{J}'_1$ with the following properties: The set $f_1, \dots, f_{b'-1}$ is $3M'$ far from $\mathcal{H}'_1, V^{\text{core}}$, and $|\text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}\} \cap \mathcal{J}'_1| = b' - 1$. The case where such a set of forms does not exist will be handled later in the proof. For each $g \in \mathcal{G}_1$, the element g is not M'-close to $f_1, \dots, f_{b'-1}$, since it if was, then $f_1, \dots, f_{b'-1}$ would be not be $3M'$-far from $\mathcal{H}'_1, V^{\text{core}}$. Here we use the fact that g is itself not M'-close to V^{core}. Therefore, we must have $|\text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}, g\} \cap \mathcal{J}'_1| > b'$. If an element in this span other than g lies in either $\mathcal{G}'_w$ or $\mathcal{G}'_1$, then this would imply that $f_1, \dots, f_{b'-1}$ is $3M'$-close to $\mathcal{H}'_1, V^{\text{core}}$, contradicting assumption. Therefore, for each g, there is at least one element in this span other than $f_1, \dots, f_{b'-1}$ that lies in $\mathcal{J}'_1$. Note that since $|\text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}\} \cap \mathcal{J}'_1| = b' - 1$, the element in the $\text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}, g\}$ non-trivially depends on g. Further note that any element in this span is M'-close to $\mathcal{H}'_1, \{f_1, \dots, f_{b'-1}, V^{\text{core}}\}$.

Suppose now for $g, g' \in \mathcal{G}'_1$, there is an element in

$$\text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}, g'\} \cap \text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}, g''\}$$

that is not one of the f_i. Since elements in these two spans non-trivially depend on g, g', this would imply that $\text{span}_{\mathbb{K}}\{g, g'\} \cap \text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}\} \neq \{0\}$. This again shows that $f_1, \dots, f_{b'-1}$ is not $3M'$-far from $\mathcal{H}'_1, V^{\text{core}}$, contradicting assumption. Define $\mathcal{H}'_2 := \{f_1, \dots, f_{b'-1}\}$, and define $\mathcal{G}'_2 \subset \mathcal{G}'_s$ to be the set of elements that are M'-close to $\mathcal{H}'_1, \mathcal{H}'_2, V^{\text{core}}$. The above argument shows $|\mathcal{H}'_2| \geq 2m/4^{b'-1}$. Define $\mathcal{J}'_2 := \mathcal{G}'_s \setminus \mathcal{B}'_2$.

We now repeat this, and look for a set $\mathcal{H}'_3 \subset \mathcal{J}_2$ of size $b' - 1$ such that $\mathcal{H}'_3$ is $3M'$-far from $\mathcal{H}'_1, \mathcal{H}'_2, V^{\text{core}}$ and such that $|\text{span}_{\mathbb{K}}\{\mathcal{H}'_3\} \cap \mathcal{J}'_2| = b' - 1$. If we can find such a set, then we can repeat this argument, and define $\mathcal{G}'_3$ to be the set of forms that are M'-close to $\mathcal{H}'_1, \mathcal{H}'_2, \mathcal{H}'_3, V^{\text{core}}$ and we will have $|\mathcal{H}'_3| \geq 4m/4^{b'-1}$.

We do this process as many times as possible. Suppose we managed to iterate this argument t times, and construct $\mathcal{H}'_1, \dots, \mathcal{H}'_t$. Note that the case where we could not pick $f_1, \dots, f_{b'-1}$ is simply the case when $t = 0$. Since the size of the set that is M'-close doubles each time, we must have $t \leq 2(b' - 2)$. At the end, one of two cases holds. Either every element of $\mathcal{G}'_s$ is M'-close to $\mathcal{H}'_1, \cdots, \mathcal{H}'_t, V^{\text{core}}$, in which case the lemma is proved. If not, then every set of $b' - 1$ forms that is $3M'$-far from $\mathcal{H}'_1, \cdots, \mathcal{H}'_t, V^{\text{core}}$ is such that $|\text{span}_{\mathbb{K}}\{f_1, \dots, f_{b'-1}\} \cap \mathcal{J}'_t| > b' - 1$.

In this latter case, the set $\mathcal{J}'_t$ is essentially a $(1, b' - 1, 3M')$-strong span sequence. The only difference is that we replace the condition $g_1, \dots, g_{b'}$ is M'-close to V^{core} with the condition that $g_1, \dots, g_{b'-1}$ is $3M'$-close to $\mathcal{H}'_1, \cdots, \mathcal{H}'_t, V^{\text{core}}$. Note that the span relations still depend only on V^{core}. All our arguments apply verbatim to this slightly more general setting. Therefore, we can apply induction, with Lemma 3.10.23 acting as a base case. Letting $\mathcal{H}'$ denote the union of all the sets constructed (namely the $\mathcal{H}'_i$ and the sets obtained by induction) we deduce that every element of $\mathcal{G}'_s$ is $3^{b'}M'$-close to $\mathcal{H}', V^{\text{core}}$. The size bound on $\mathcal{H}'$ also follows by induction. This completes the proof. □

3.10.3 The low strength case

In this subsection we consider identities $\mathcal{C}$ where all the quadratic forms are close to an algebra $\mathbb{K}[V]$ for some vector space $V \subseteq R_{\leq 2}$. We start by defining useful potential functions the capture the proximity to the algebra $\mathbb{K}[V]$ and absolute irreducibility over V.

Potential functions and absolutely irreducible forms

Let $\Gamma, \Lambda, f : \mathbb{N}^2 \to \mathbb{N}^2$ and $g : \mathbb{N}^2 \to \mathbb{N}$ be ascending functions. Let $R = S/(U)$, where $U \subseteq S_{\leq 2}$ is $\Lambda \circ t_{(0,1)}$-strong. Set $a := 2g(\delta_U) + f_1(\delta_U)$. Suppose that U satisfies the following conditions:

1. $\Lambda_i(\delta_U) \geq B \circ t_{(2a,0)}(\delta_U)$.
2. $\Gamma \geq h_{2B} \circ t_1$.
3. $\Gamma_2(\delta_U + f_1(\delta_U) + f_2(\delta_U)) \geq a + \dim(U_1)$.

Relative linear spaces. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong such that $\dim(V_i) \leq f_i(\delta_U)$. Note that for $F \in R_2$, such that ps-dist$(F, V) < g(\delta_U)$, the relative spaces $\mathbb{L}_R(a, V, F) \subseteq R_1$ and $\mathbb{L}_{R,V}(a, F) \subseteq R_1/V_1$ are well-defined by Definition 3.10.9.

Individual potential. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong and $\dim(V_i) \leq f_i(\delta_U)$. Let $F \in R_{\leq 2}$ be such that ps-dist$(F, V) < g(\delta_U)$ if $\deg(F) = 2$. We define the individual potential of F with respect to V (and the bounding functions f, g) as follows

$$\Phi(f, g, V, F) = \begin{cases} 0 & F \in V_1 \\ 1 & \text{if } F \in R_1 \setminus V_1 \\ \dim(\mathbb{L}_{R,V}(a, F)) & \text{if } F \in R_2 \end{cases}$$

Absolutely irreducible forms. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong and $\dim(V_i) \leq f_i(\delta_U)$. Let $\mathcal{C} = \sum_{i=1}^{k} T_i$ be a $\Sigma^k\Pi\Sigma\Pi^2$ simple and minimal identity in R. For $i, j \in [k]$ we denote the set of absolutely irreducible forms in $T_i \Delta T_j$ as

$$\mathcal{B}_{ij}(V, \mathcal{C}) = \{F \in T_i \Delta T_j \mid F \text{ is absolutely irreducible over } V \text{ and } F \notin (V)\}.$$

We denote the set of absolutely irreducible forms in $\mathcal{C}$ as

$$\mathcal{B}(V, \mathcal{C}) = \{F \in \mathcal{C} \mid F \text{ is absolutely irreducible over and } F \notin (V)\}.$$

Proposition 3.10.25. *For all irreducible forms $F \in \mathcal{C}$, there exists $i \neq j$ such that $F \in T_i \Delta T_j$. Moreover, we have $\mathcal{B}(V, \mathcal{C}) = \cup_{i \neq j} \mathcal{B}_{ij}(V, \mathcal{C})$.*

Proof. Since $\mathcal{C}$ is a simple and minimal identity, we know that for every irreducible form $F \in \mathcal{C}$, there exists T_i, T_j such that $T_i \in (F)$ and $T_j \notin (F)$. In particular, $F \in T_i \Delta T_j$. Therefore, every form $F \in \mathcal{B}(V, \mathcal{C})$ belongs to some $\mathcal{B}_{ij}(V, \mathcal{C})$. Hence $\mathcal{B}(V, \mathcal{C}) = \cup_{i \neq j} \mathcal{B}_{ij}(V, \mathcal{C})$. □

Potential of a circuit. Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong and $\dim(V_i) \leq f_i(\delta_U)$. Let $\mathcal{C} = \sum_{i=1}^{k} T_i$ be a $\Sigma^k\Pi\Sigma\Pi^2$ simple and minimal identity in R such that ps-dist$(F, V) < g(\delta_U)$ for all $F \in \mathcal{C} \cap R_2$. We define the potential of the circuit $\mathcal{C}$ as follows:

$$\Phi(f, g, V, \mathcal{C}) = \sum_{i \neq j} \max_{F \in T_i \Delta T_j} \Phi(f, g, V, F).$$

We define the *irreducibility potential* of the circuit $\mathcal{C}$ as follows:

$$\Phi_{\text{irr}}(f, g, V, \mathcal{C}) = \sum_{i \neq j} \max_{F \in \mathcal{B}_{ij}(V, \mathcal{C})} \Phi(f, g, V, F).$$

Proposition 3.10.26. *Let $V \subseteq R_{\leq 2}$ be $\Gamma \circ t_{(0,1)}$-lifted strong and $\dim(V_i) \leq f_i(\delta_U)$. Let $\mathcal{C} = \sum_{i=1}^{k} T_i$ be a $\Sigma^k\Pi\Sigma\Pi^2$ simple and minimal identity in R such that $\text{ps-dist}(F, V) < g(\delta_U)$ for all $F \in \mathcal{C} \cap R_2$. Then we have $\Phi(f, g, V, F) \leq a$ for all $F \in T_i \Delta T_j$, and*

$$\Phi_{\text{irr}}(f, g, V, \mathcal{C}) \leq \Phi(f, g, V, \mathcal{C}) \leq a\binom{k}{2}.$$

Proof. If $F \in R_1$, then $\Phi(f, g, V, F) \leq 1 \leq a$. If $F \in R_2$, then we have $\Phi(f, g, V, F) = \dim(\mathbb{L}_{R,V}(a, F)) \leq a$ by Proposition 3.10.10. By summing over all the $\binom{k}{2}$ symmetric differences, we obtain $\Phi(f, g, V, \mathcal{C}) \leq a\binom{k}{2}$. We have $\Phi_{\text{irr}}(f, g, V, \mathcal{C}) \leq \Phi(f, g, V, \mathcal{C})$ as $\mathcal{B}_{ij}(V, \mathcal{C}) \subseteq T_i \Delta T_j$. □

Identities with bounded pseudo-distance

In this subsection, we state our main result that controls identities where all forms $F \in \mathcal{C}$ have bounded pseudo-distance with respect to a vector space V.

Notation. Suppose $\mu_{k-1} : \mathbb{N}^2 \to \mathbb{N}^2$ and $\xi_{k-1} : \mathbb{N}^2 \to \mathbb{N}$ are fixed functions. We define the function $\xi_k^2 : \mathbb{N}^2 \to \mathbb{N}$ as

$$\xi_k^2(\delta) := \xi_{k-1}\big(\big(C_{2\mu_{k-1}} \circ t_1\big)(\delta)\big) + 1.$$

Let $g^{\dim} : \mathbb{N}^2 \to \mathbb{N}^2$ and $g^{ps} : \mathbb{N}^2 \to \mathbb{N}$ be given functions. These functions will upper bound the dimension of the initial vector space V and the pseudo-distances $\text{ps-dist}(F, V)$ for forms $F \in \mathcal{C}$.

Auxiliary functions. Define $a_k : \mathbb{N}^2 \to \mathbb{N}$ as $a_k(\delta) := 2g^{ps}(\delta) + \delta_1$ for $\delta \in \mathbb{N}^2$. We define $\lambda_k(\delta) := 2a_k(\delta) + \xi_k^2(\delta) + 1$ and $\beta_k(\delta) := k^2 \cdot \xi_k^2(\delta)$. Moreover, define the function $c_k : \mathbb{N}^2 \to \mathbb{N}$ as

$$c_k := a_k\binom{k}{2}\left(a_k\left(\binom{k}{2} + \beta_k\right) + 2a_k^2\lambda_k\binom{k}{2} + a_k\lambda_k\binom{k}{2}\right).$$

Strength and dimension bounds under strengthening process. We define two new functions $\Gamma, f : \mathbb{N}^2 \to \mathbb{N}^2$ (depending on the functions above). Let $\Gamma = h_{2B} \circ t_{(2c_k,1)}$ and $f = C_\Gamma \circ t_{g^{dim}}$.

In the situation of Lemma 3.10.27, we will have a vector space V of dimension sequence at most $g^{\dim}(\delta_U)$. The function Γ will be the strength lower bound that we will enforce when we apply a strengthening process on V. Moreover, the function f will upper bound the dimension sequence of the resulting vector space after strengthening V.

Strength lower bound function for U. Let $\mu_k^{\text{low}} := h_{2\Gamma} \circ t_{g^{\dim}}$. This function will provide the strength lower bound required on U in Lemma 3.10.27.

Constants. Given $U \subseteq S_{\leq 2}$ as in Lemma 3.10.27 below, let us set $a := 2g^{ps}(\delta_U) + f_1(\delta_U)$. We define $\lambda(k) := 2a + \xi_k^2(\delta_U) + 1$ and we have $\beta_k(\delta_U) = k^2 \cdot \xi_k^2(\delta_U)$.

Dimension upper bound. Let $\rho_k(\delta_U) := g_1^{\dim}(\delta_U) + g_2^{\dim}(\delta_U) + c_k(\delta_U + f_1(\delta_U) + f_2(\delta_U)) + C_\Gamma(\delta_U + g^{\dim}(\delta_U))$. This quantity will provide the dimension upper bound on vector spaces Y in outcomes (1), (2) of Lemma 3.10.27.

Lemma 3.10.27. *Suppose $\mu_{k-1} : \mathbb{N}^2 \to \mathbb{N}^2$ and $\xi_{k-1} : \mathbb{N}^2 \to \mathbb{N}$ are fixed functions. Let $g^{\dim} : \mathbb{N}^2 \to \mathbb{N}^2$ and $g^{ps} : \mathbb{N}^2 \to \mathbb{N}$ be functions. Let $\xi_k^2, \mu_k^{low}, \rho_k$ be the functions defined above.*

Let $U \subseteq S_{\leq 2}$ be a μ_k^{low}-strong vector space and $R := S/(U)$. Let $\mathcal{C} = \sum_{i=1}^{k} T_i$ be a $\Sigma^k\Pi\Sigma\Pi^2$ simple and minimal identity in R. Suppose there exists a vector space $V \subseteq R_{\leq 2}$ such that

- $\dim(V_i) \leq g_i^{\dim}(\delta_U)$, *and*
- *for all quadratic forms $Q \in \mathcal{C} \cap R_2$, we have $\text{ps-dist}(Q, V) < g^{ps}(\delta_U)$, i.e. all quadratic forms of $\mathcal{C}$ are $g^{ps}(\delta_U)$-close to $\mathbb{K}[V]$.*

Then one of the following holds.

1. *There exists a vector space $Y \subseteq R_{\leq 2}$ with $\dim(Y) \leq \rho_k(\delta_U)$, and a pair $i, j \in [k]$, such that*
$$T_i \Delta T_j \subseteq \mathbb{K}[Y].$$
2. *There exists a vector space $Y \subseteq R_{\leq 2}$ with $\dim(Y) \leq \rho_k(\delta_U)$, such that for all $Q \in \mathcal{C} \cap R_2$ we have that Q is not absolutely irreducible over Y.*
3. *There exists $H \in \mathcal{C}$ with the following property. Let $W := AH_R(\mu_{k-1}, H)$ be the μ_{k-1}-lifted strong space obtained by applying Corollary 3.6.6 to $\mathrm{span}_{\mathbb{K}}\{H\}$ in R. Let $\varphi_{W,\alpha} : R[y] \to R' := R[y]/I_{W,\alpha}$ be the graded quotient corresponding to $\alpha \in \mathbb{K}^{\dim(W)}$. For a general $\alpha \in \mathcal{Z}(H) \subseteq \mathbb{K}^{\dim(W)}$, we have that the circuit $\varphi_{W,\alpha}(\mathcal{C}) \subseteq R'$ has pairwise rank at least $\xi_{k-1}(\delta_W)$.*

The rest of this section will be devoted to the proof of Lemma 3.10.27. We will construct a vector space Y via an iterative process, and show that if Y does not satisfy condition 1, then there exists a form $H \in \mathcal{C}$ that enables us to carry out fan-in reduction.

Constructing cores and disjoint sequences

We will now iteratively construct a set of forms that will make up the core. We will then use another iterative process to pick a sequence of forms that we call a good disjoint sequence. This is analogous to what we did in Section 3.10.2. Throughout these iterative processes, we will use our potential functions to measure progress.

Construction of a core: *Set-up.* Suppose we are in the setting of Lemma 3.10.27. Let Γ, f be the functions defined in the notation before Lemma 3.10.27. In the following, we will write $g := g^{ps}$, where g^{ps} is given in the statement of Lemma 3.10.27. Recall that $a := 2g(\delta_U) + f_1(\delta_U)$. We also have $\lambda(k) := 2a + \xi_k^2(\delta_U) + 1$. For simplicity, we will write β_k in place of $\beta_k(\delta_U) = k^2 \cdot \xi_k^2(\delta_U)$.

Strengthening. First, let us replace V by the vector space $AH_R(\Gamma, V)$ which is $\Gamma \circ t_{(0,1)}$-lifted strong. Note that, after updating V, we have $\dim(V_i) \leq f_i(\delta_U)$ and for all quadratic forms $Q \in \mathcal{C}$, we have $\text{ps-dist}(Q, V) \leq g(\delta_U)$ by Proposition 3.10.6.

Construction of core. We will pick a set of core forms using an iterative process and we will denote the set of these forms as $\mathcal{A}_{ij}^{\mathrm{core}}(V, \mathcal{C}) \subseteq T_i \Delta T_j$. We define
$$\mathcal{A}^{\mathrm{core}}(V, \mathcal{C}) := \cup_{i \neq j} \mathcal{A}_{ij}^{\mathrm{core}}(V, \mathcal{C})$$
and
$$W^{\mathrm{core}}(V, \mathcal{C}) := \sum_{G \in \mathcal{A}^{\mathrm{core}}(V, \mathcal{C})} \mathbb{L}_R(a, V, G) \subseteq R.$$

Iterative process for core. We describe the iterative process below.
Start by setting $Y := V$, let $\mathcal{A}^{\mathrm{core}}(Y, \mathcal{C}) = \emptyset$ and $\mathcal{A}_{ij}^{\mathrm{core}}(Y, \mathcal{C}) = \emptyset$ for $i, j \in [k]$.
If $T_i \Delta T_j \subseteq \mathbb{K}[Y]$ for some $i \neq j \in [k]$, then we stop.
Otherwise, for $i \in [k]$ and $j > i$, we do the following:

- If $\mathcal{B}_{ij}(Y, \mathcal{C}) \neq \emptyset$, i.e. there are quadratics in $T_i \Delta T_j$ that are absolutely irreducible over Y, do the following:
  - If $\mathcal{B}_{ij}(Y, \mathcal{C})$ contains a sequence of quadratic forms $F_1, \cdots, F_{\lambda(k)}$ such that the set
  $$\mathcal{A}^{\mathrm{core}}(Y, \mathcal{C}) \cup \{F_1, \cdots, F_{\lambda(k)}\}$$
  is lin-separated modulo Y (recall Definition 3.10.11), then we let $\mathcal{A}_{ij}^{\mathrm{core}}(Y, \mathcal{C}) := \{F_1, \cdots, F_{\lambda(k)}\}$.

– Otherwise, we have that the maximum length of a lin-separated sequence in $\mathcal{B}_{ij}(Y,\mathcal{C})$ is less than $\lambda(k)$. Let $F_1,\cdots,F_m$ be such a maximal sequence where $m<\lambda(k)$. We replace Y with $\bigcup_{n=1}^{m}\mathbb{L}_R(a,Y,F_n)$, and restart the iterative process with the new Y and $i=1$.[5]

- If $\mathcal{B}_{ij}(Y,\mathcal{C})=\emptyset$, we do the following:
  – If there exist (not necessarily quadratic) forms $F_1,\cdots,F_{\lambda(k)}$ in $T_i\Delta T_j$ such that the set
  $$\mathcal{A}^{\mathrm{core}}(Y,\mathcal{C})\cup\{F_1,\cdots,F_{\lambda(k)}\}$$
  is lin-separated modulo Y, then we let $\mathcal{A}_{ij}^{\mathrm{core}}(Y,\mathcal{C}):=\{F_1,\cdots,F_{\lambda(k)}\}$.
  – Otherwise, we have that the maximum length of a lin-separated sequence in $T_i\Delta T_j$ is less than $\lambda(k)$. Let $F_1,\cdots,F_m$ be such a maximal sequence where $m<\lambda(k)$. We replace Y with $\bigcup_{n=1}^{m}\mathbb{L}_R(a,Y,F_n)$, and restart the iterative process with the new Y and $i=1$.

Now we prove termination and properties of the outcome of this iterative process.

Proposition 3.10.28. *We have the following.*

1. *The iterative process restarts at most $2a\binom{k}{2}$ times. In particular, the iterative process terminates.*
2. *At the end of the iterative process, we have a vector space Y, containing V, that is $\Gamma\circ t_{(0,1)}$-lifted strong and $\dim(Y)\leq\dim(V)+2a^2\lambda(k)\binom{k}{2}$. Furthermore, we have $\dim(W^{core})\leq a\lambda(k)\binom{k}{2}$ and the vector space $Y+W^{core}$ is $\Gamma\circ t_{(0,1)}$-lifted strong.*
3. *Moreover, one of the following holds.*
   (a) *We have $T_i\Delta T_j\subseteq\mathbb{K}[Y]$.*
   (b) *We have $\mathcal{B}_{ij}(Y,\mathcal{C})=\emptyset$ for all $i\neq j$. In particular, $\mathcal{B}(Y,\mathcal{C})=\emptyset$.*
   (c) *There exists $i\neq j$ such that $\mathcal{B}_{ij}(Y,\mathcal{C})\neq\emptyset$. In this case, for all i,j, we have one of the following.*
      i. *If $\mathcal{B}_{ij}(Y,\mathcal{C})\neq\emptyset$, then $\mathcal{A}_{ij}^{core}(Y,\mathcal{C})\subseteq\mathcal{B}_{ij}(Y,\mathcal{C})$ and the core $\mathcal{A}_{ij}^{core}(Y,\mathcal{C})$ consists of $\lambda(k)$ number of absolutely irreducible quadratic forms that are lin-separated modulo Y.*
      ii. *If $\mathcal{B}_{ij}(Y,\mathcal{C})=\emptyset$, then the core $\mathcal{A}_{ij}^{core}(Y,\mathcal{C})$ consists of $\lambda(k)$ number of forms that are lin-separated modulo Y. In particular, for each quadratic form F in $\mathcal{A}_{ij}^{core}(Y,\mathcal{C})$, we have $F\in(Y)$ or F is absolutely reducible.*

Proof. (1) It is enough to show that the iterative process restarts at most $2a\binom{k}{2}$ times. Suppose that for certain i,j, we have $\mathcal{B}_{ij}(Y,\mathcal{C})\neq\emptyset$ and the iterative process restarts because there is no $\lambda(k)$-length sequence which is lin-separated modulo Y. Let $Y'=\bigcup_{n=1}^{m}\mathbb{L}_R(a,Y,F_n)$, which is the new Y when we restart the iterative process. Note that $\mathcal{B}_{ij}(Y',\mathcal{C})\subseteq\mathcal{B}_{ij}(Y,\mathcal{C})$ as $Y\subseteq Y'$. Moreover, for all $F\in\mathcal{B}_{ij}(Y,\mathcal{C})\setminus\{F_1,\cdots,F_m\}$, we know that the dimension of the relative linear space drops, i.e. $\dim(\mathbb{L}_{R,Y'}(a,F))<\dim(\mathbb{L}_{R,Y}(a,F))$, as $F_1,\cdots,F_m,F$ is not lin-separated modulo Y. Also, $\mathbb{L}_{R,Y'}(a,F)=(0)$, whereas we have $\mathbb{L}_{R,Y}(a,F)\neq(0)$. Therefore we have

$$\Phi(f,g,Y',F)<\Phi(f,g,Y,F),$$

[5] Proposition 3.10.28 shows that Y remains $\Gamma\circ t_{(0,1)}$-lifted strong throughout the process, and therefore it satisfies the conditions of Definition 3.10.11.

for all $F \in \mathcal{B}_{ij}(Y, \mathcal{C})$. Hence the irreducibility potential of $\mathcal{C}$ must decrease, i.e. we have

$$\Phi_{\text{irr}}(f, g, Y', \mathcal{C}) < \Phi_{\text{irr}}(f, g, Y, \mathcal{C}).$$

By Proposition 3.10.26, we have an upper bound $\Phi_{\text{irr}}(f, g, Y, \mathcal{C}) \leq a\binom{k}{2}$. Therefore, the first kind of restart occurs at most $a\binom{k}{2}$ times. Similarly, we see that every time the second kind of restart works, the circuit potential $\Phi(f, g, Y, \mathcal{C})$ must decrease. Hence the second kind of restart also occurs at most $a\binom{k}{2}$ times.

(2) At the beginning of the iterative process we have $Y = V$. Now the vector space Y is updated only when a restart step occurs in the process. Recall that $\dim(\mathbb{L}_R(a, Y, F) \leq a$ for all F_i. Hence, whenever a restart step occurs, the dimension of Y increases by at most $a\lambda(k)$. Since there are at most $2a\binom{k}{2}$ restart steps, we have that the final Y satisfies $\dim(Y) \leq \dim(V) + 2a^2\lambda(k)\binom{k}{2}$. Let $c := 2a^2\lambda(k)\binom{k}{2}$. Since V was $\Gamma \circ t_{(c,1)}$-lifted strong, we have that Y is $\Gamma \circ t_{(0,1)}$-lifted strong. Note that $\mathcal{A}^{\text{core}}_{ij}(Y, \mathcal{C})$ has $\lambda(k)$ forms. For each such form $F \in \mathcal{A}^{\text{core}}_{ij}(Y, \mathcal{C})$ we have $\dim(\mathbb{L}_{R,Y}(a, F)) \leq a$. Therefore, $\dim(W^{\text{core}}) \leq a\lambda(k)\binom{k}{2}$. Again, let $c' = a\lambda(k)\binom{k}{2} + 2a^2\lambda(k)\binom{k}{2}$. Since V is $\Gamma \circ t_{(c',1)}$-lifted strong, we have that $Y + W^{\text{core}}$ is $\Gamma \circ t_{(0,1)}$-lifted strong.

(3) If $\mathcal{B}_{ij}(V, \mathcal{C}) = \emptyset$ for all $i \neq j$, then $\mathcal{B}(V, \mathcal{C}) = \emptyset$ by Proposition 3.10.25. At the last iteration of the process (when no more restart conditions occur), we end up with $\mathcal{A}^{\text{core}}(Y, \mathcal{C})$ with the stated properties due to the construction. □

Definition 3.10.29. After the iterative process ends, we will say that $T_i \Delta T_j$ is an *irreducible symmetric difference* if $\mathcal{B}_{ij}(Y, \mathcal{C}) \neq \emptyset$ (corresponding to case 3.2.1 in Proposition 3.10.28). Otherwise, we say that $T_i \Delta T_j$ is a *non-irreducible symmetric difference* (corresponding to case 3.2.2 in Proposition 3.10.28). ◊

Disjoint sequences: We continue with the set up above and assume that we have constructed a core using Proposition 3.10.28. We will now define the notion of *disjoint* sequences, which will be the analogue of the witness set $\mathcal{H}$ in Section 3.10.2.

Notation. Recall that we defined $\beta_k := k^2 \cdot \xi_k^2(\delta_U)$. Let Y and $W^{\text{core}}(Y, \mathcal{C})$ be the vector spaces given by the core construction above. Let $R' = R/(Y + W^{\text{core}}(Y, \mathcal{C}))$. For $F \in R$ let F' denote the image of F in R'.

Definition 3.10.30. We say that $H_1, \cdots, H_{\beta_k} \in \mathcal{C} \cap R_2$ is a disjoint sequence with respect to (Y, W^{core}) if the following holds.

1. We have $H_i \in \mathcal{B}(Y + W^{\text{core}}, \mathcal{C})$, i.e. H_i is absolutely irreducible over $Y + W^{\text{core}}$ and $H_i \notin (Y + W^{\text{core}})$.

2. The forms $H_1, \cdots, H_{\beta_k}$ are lin-separated modulo $Y + W^{\text{core}}(Y, \mathcal{C})$.

3. Then for any other $G_1, \cdots, G_{\beta_k}$ with properties (1) and (2) above, we have

$$\sum_{i=1}^{\beta_k} \dim(\mathbb{L}_{R'}(a, G_i')) \leq \sum_{i=1}^{\beta_k} \dim(\mathbb{L}_{R'}(a, H_i')).$$

In other words $H_1, \cdots, H_{\beta_k}$ maximizes the sum of dimensions of essential variables in R' (among sequences of size β_k with properties (1) and (2) above).

◊

Definition 3.10.31 (Good disjoint sequences.)**.** Let $H_1, \cdots, H_{\beta_k}$ be a disjoint sequence with respect to (Y, W^{core}). We say that $H_1, \cdots, H_{\beta_k}$ is a *good* disjoint sequence if there is no form $P \in \mathcal{B}(Y + W^{\text{core}})$ such that $P = \alpha H_i + \beta H_j + Q$ for some $Q \in \mathbb{K}[Y + W^{\text{core}}]$ and $\alpha, \beta \neq 0$. $\Diamond$

Lemma 3.10.32. *Let $H_1, \cdots, H_{\beta_k}$ be a disjoint sequence with respect to (Y, W^{core}) which is not a good disjoint sequence. Let $P \in \mathcal{B}(Y + W^{core}, \mathcal{C})$ such that $P = \alpha H_i + \beta H_j + Q$ for some $Q \in \mathbb{K}[Y + W^{core}]$ and $\alpha, \beta \neq 0$. Then the following holds.*

1. *The sequence*
$$H_1 \cdots, H_{i-1}, H_{i+1}, \cdots, H_{j-1}, H_{j+1}, \cdots, H_{\beta_k}, P$$
is lin-separated modulo $Y + W^{core}$.

2. *For any $F \in \mathcal{B}(Y + W^{core}, \mathcal{C})$, the sequence*
$$H_1 \cdots, H_{i-1}, H_{i+1}, \cdots, H_{j-1}, H_{j+1}, \cdots, H_{\beta_k}, P, F$$
is not lin-separated modulo $Y + W^{core}$.

Proof. (1) Let $Y' = Y + W^{\text{core}}$. We have that $\mathbb{L}_{R,Y'}(a, P) \subseteq \mathbb{L}_{R,Y'}(a, H_i) + \mathbb{L}_{R,Y'}(a, H_j)$. Since $H_1, \cdots, H_{\beta_k}$ are lin-separated, we have
$$(\mathbb{L}_{R,Y'}(a, H_i) + \mathbb{L}_{R,Y'}(a, H_j)) \cap \sum_{\ell \neq i,j} \mathbb{L}_{R,Y'}(a, H_\ell) = (0).$$
Hence we are done.

(2) By Lemma 3.10.13, we have that $\dim(\mathbb{L}_{R'}(a, P')) = \dim(\mathbb{L}_{R'}(a, H_i') + \dim(\mathbb{L}_{R'}(a, H_i'))$. Therefore, the sequence obtain by removing H_i, H_j and adding P has the same sum of dimensions of relative spaces in R'. However, it is of length $\beta_k - 1$. If we have a lin-separated sequence after adding F, then this would contradict condition (3) in the definition of disjoint sequences. $\square$

Note that it is not necessary that a disjoint sequence exists, even without the additional good property above. However, if a good disjoint sequence does not exist then we can control the forms in $\mathcal{C}$. In particular, after increasing the vector space Y we will be able to obtain $\mathcal{B}(Y, \mathcal{C}) = \emptyset$. We will use the following iterative process for constructing a good disjoint sequence.

Iterative process for constructing good disjoint sequences. Start with Y and W^{core} as above.

1. If $\mathcal{B}(Y + W^{\text{core}}, \mathcal{C}) \neq \emptyset$, we do the following.

   (a) If there exists a lin-separated sequence of $H_1, \cdots, H_{\beta_k} \in \mathcal{B}(Y + W^{\text{core}}, \mathcal{C})$: we choose one with maximum dimensional sum and obtain a disjoint sequence. Now we do the following:

      i. If the disjoint sequence is good, we stop.

      ii. If the disjoint sequence is not good, i.e. there exists $\alpha H_i + \beta H_j + Q \in \mathcal{B}(Y + W^{\text{core}}, \mathcal{C})$. Then we consider
      $$H_1 \cdots, H_{i-1}, H_{i+1}, \cdots, H_{j-1}, H_{j+1}, \cdots, H_{\beta_k}, P.$$
      We replace Y with $Y + W^{\text{core}} + \sum_{\ell \neq i,j}^{\beta_k} \mathbb{L}_{R,Y+W^{\text{core}}}(a, H_\ell) + \mathbb{L}_{R,Y+W^{\text{core}}}(a, P)$. We run the core constructing iterative process with this new Y.

      A. If the core construction ends in $\mathcal{B}(Y, \mathcal{C}) = \emptyset$, we stop.

B. Otherwise, we restart our iterative process with the output (Y, W^{core}) of the core construction process.

(b) Otherwise, we pick a maximal length lin-separated sequence $H_1, \cdots, H_m \in \mathcal{B}(Y+W^{\mathrm{core}}, \mathcal{C})$ and replace Y with $Y + W^{\mathrm{core}} + \sum_{i=1}^{m} \mathbb{L}_{R,Y+W^{\mathrm{core}}}(a, H_i)$. We run the core constructing iterative process with this new Y.

i. If the core construction ends in $\mathcal{B}(Y, \mathcal{C}) = \emptyset$, we stop.

ii. Otherwise, we restart our iterative process with the output (Y, W^{core}) of the core construction process.

2. If $\mathcal{B}(Y + W^{\mathrm{core}}, \mathcal{C}) = \emptyset$, we stop.

Lemma 3.10.33. *The iterative process for good disjoint sequences terminates. At the end of the process, we have one of the following:*

1. *We have an output of the core construction (Y', W^{core}) where Y' is $\Gamma \circ t_{(0,1)}$-lifted strong. Moreover we have a good disjoint sequence $H_1, \cdots, H_{\beta_k}$ with respect to (Y', W^{core}).*

2. *We have a $\Gamma \circ t_{(0,1)}$-lifted strong vector space Y' such that $\mathcal{B}(Y', \mathcal{C}) = \emptyset$.*

Moreover, in both the cases above we have

$$\dim(Y') \leq \dim(Y) + a\binom{k}{2}\left(a(\binom{k}{2} + \beta_k) + 2a^2\lambda(k)\binom{k}{2} + a\lambda(k)\binom{k}{2}\right).$$

Proof. Termination. Suppose that for a certain (Y, W^{core}) the iterative process restarts. In each of the two restart steps, i.e. 1.1.2.2 and 1.2.2, we have a maximal lin-separated sequence $F_1, \cdots, F_m \in \mathcal{B}(Y + W^{\mathrm{core}}, \mathcal{C})$ with $m < \beta_k$. Let $Y' = Y + W^{\mathrm{core}} + \sum_{i=1}^{m} \mathbb{L}_{R,Y+W^{\mathrm{core}}}(a, F_i)$. Now any form $F \in \mathcal{B}(Y + W^{\mathrm{core}}, \mathcal{C})$, we have that $\Phi(f, g, Y', F) < \Phi(f, g, Y, F)$, as $F_1, \cdots, F_m, F$ is not lin-separated modulo Y. Therefore, every time a restart occurs the irreducibility potential $\Phi_{\mathrm{irr}}(f, g, Y, \mathcal{C})$ decreases. Since we have an upper bound $\Phi_{\mathrm{irr}}(f, g, Y, \mathcal{C}) \leq a\binom{k}{2}$, the iterative process terminates after at most $a\binom{k}{2}$ restart steps.

Dimension bound and strength. Every time we encounter step 1.1.2 or 1.2 above, we first increase the dimension of Y by at most $a(\binom{k}{2} + \beta_k)$. Then if we run the core construction with this Y, the output (Y', W^{core}) satisfies

$$\dim(Y' + W^{\mathrm{core}}) \leq \dim(Y) + a(\binom{k}{2} + \beta_k) + 2a^2\lambda(k)\binom{k}{2} + a\lambda(k)\binom{k}{2}.$$

Since there are at most $a\binom{k}{2}$ restart steps, we have an upper bound on dimension of the output Y' (in all cases):

$$\dim(Y') \leq \dim(Y) + a\binom{k}{2}\left(a(\binom{k}{2} + \beta_k) + 2a^2\lambda(k)\binom{k}{2} + a\lambda(k)\binom{k}{2}\right).$$

Since if we let c be the summand above such that $\dim(Y') = \dim(Y) + c$, then Y' is $\Gamma \circ t_{(0,1)}$-lifted strong, since Y is $\Gamma \circ t_{(c,1)}$-lifted strong. $\square$

A fan-in reduction lemma

We now use the core construction (Y, W^{core}) and a good disjoint sequence from Section 3.10.3 to show a fan-in reduction lemma. As in Section 3.10.2, we again define the notions of core preserving and core collapsing forms. The definitions will be more intricate, because of the changed setting. Let $\mathcal{H}$ denote the good disjoint sequence constructed in the previous subsection.

Definition 3.10.34. Suppose $G_1, \ldots, G_s$ are the forms from $T_1 \Delta T_2$ that were picked when constructing the core, with $s = \lambda(k)$. Suppose H is one of the forms in the good disjoint sequence that does not belong to $T_1 \cup T_2$.

We say H is core preserving for $T_1 \Delta T_2$ if the following happens: There exists at least $\xi_k^2(\delta_U)$ many forms among the G_i, say $G_1, \ldots, G_t$ such that

- The forms H and G_i are lin-separated modulo Y for $i \leq t$, and
- $|\text{span}_{\mathbb{K}}\{G_i, H\} \cap T_i \Delta T_j| = 1$ for $i \leq t$.

We say H is core collapsing for $T_1 \Delta T_2$ if it is not core preserving. $\lozenge$

In the previous subsection, in order to kill a form H, we studied the image of the circuit $\mathcal{C}$ under the quotient map $R \to R/(H)$. This was because we were focusing on those H that are lifted strong, and therefore the ring $R/(H)$ had the nice properties for induction to apply. If on the other hand, H is a form of slow strength, then $R/(H)$ might not even be a UFD. In this case, we will kill H using a different map.

Given any vector space W, recall the definition of the graded quotient $\varphi_{W,\alpha}$, for $\alpha \in \mathbb{K}^{\dim W}$. This map sends a basis element w_i of W to $\alpha_i z^{\deg w_i}$. For a general choice of α, this map has a number of desirable properties (Proposition 3.6.15). If $H \in \mathbb{K}[W]$, then we can write H as an element in $\mathbb{K}[W]$, and pick α to be a root of H. For this special choice of α we have $\varphi_{W,\alpha}(H) = 0$. This will be the type of map we use to kill H, we call these targeted graded quotients.

Proposition 3.10.35 (Targeted Quotients)**.** *Let $S = \mathbb{K}[x_1, \cdots, x_n, y_1, \cdots, y_m]$ and z be a new variable. Fix positive integers $d_1, \cdots, d_n \in \mathbb{N}$. Let $F \in S$ be an irreducible homogeneous polynomial such that $F \in \mathbb{K}[x_1, \cdots, x_n]$. For $\alpha \in \mathbb{K}^n$, let $I_\alpha = (x_1 - \alpha_1 z^{d_1}, \cdots, x_n - \alpha_n z^{d_n})$. Let $\varphi_\alpha : S[z] \to S[z]/I_\alpha$ be the quotient ring homomorphism. Then we have the following.*

1. *Let $Q \in S$ such that $Q \notin (F)$. Then $\varphi_\alpha(Q) \neq 0$ for general $\alpha \in \mathcal{Z}(F)$, where $\mathcal{Z}(F) \subseteq \mathbb{K}^n$ denotes the set of zeroes of F.*
2. *Let $Q, H \in S$ be forms of positive degree in at least one of the variables y_i. If $\varphi_\alpha(Q), \varphi_\alpha(H)$ have a common factor which is not in $\mathbb{K}[z]$, then F divides the resultant $\text{Res}_{y_i}(Q, H)$ for some y_i.*

If H is already μ_{k-1}-lifted strong in R, then we will directly analyse the image of the circuit under the quotient map. If not, then we will analyse the image of the circuit under the composition of two maps: the first is a general projection of Y, and the second is a targeted quotient of the image of H. We will show that under the composition of these maps, the image of $\mathcal{C}$ has large pairwise rank. We can alternatively write this composition as a targeted quotient of H, followed by a general projection of the remaining elements of Y. The fact that the pairwise rank is large under the composition will imply that it will be large under the first of these maps.

Lemma 3.10.36. *Suppose H is core preserving for $T_1 \Delta T_2$. Let $W := AH_R(\mu_{k-1}, H)$ be the μ_{k-1}-lifted strong space obtained applying Corollary 3.6.6 to $\mathrm{span}_{\mathbb{K}}\{H\}$ in R. Let $\varphi_{W,\alpha} : R[y] \to R' := R[y]/I_{W,\alpha}$ be the graded quotient corresponding to $\alpha \in \mathbb{K}^{\dim(W)}$. For a general $\alpha \in \mathcal{Z}(H) \subseteq \mathbb{K}^{\dim(W)}$, the image of T_1, T_2 under the map $\varphi_{W,\alpha}$ has pairwise rank at least $\xi_{k-1}(\delta_W)$.*

Proof. As in the definition Definition 3.10.34, we assume that $G_1, \ldots, G_t$ are the forms from the core that are lin-separated modulo Y with H, and such that $|\mathrm{span}_{\mathbb{K}}\{G_i, H\} \cap T_i \Delta T_j| = 1$. Here we have $t \geq \xi_k^2(\delta_U)$ by definition. We will show that under a targeted quotient $\varphi_{W,\alpha}$, the images of these G_i witness the pairwise rank of T_1, T_2. Fix G_1, and assume without loss of generality that $G_1 \in T_1 \setminus T_2$.

We perform a case analysis, based on the lifted strength of H. First, let us assume that H is already μ_{k-1}-lifted strong. In this case, the map $\varphi_{W,\alpha}$ is the same as the map $R \to R/(H)$. Suppose G_1 is a quadratic form. If G_1 becomes reducible in $R/(H)$, then $G_1 = H + \ell_1\ell_2$ in R. However, this contradicts the fact that G_1 and H are lin-separated modulo Y with H absolutely irreducible modulo Y. Therefore, if G_1 is a quadratic, and there is a $F \in T_2 \setminus T_1$ such that G_1 and F have non trivial gcd in $R/(H)$, it must be that $F \in \mathrm{span}_{\mathbb{K}}\{H, G_1\}$, in contradiction to the definition of H. If G_1 is a linear form, then any such F must be a quadratic form, and it can be written as $F = H + G_1\ell$ for a linear form ℓ. Now G_1 is in the core, therefore every element of $T_1 \Delta T_2$ is absolutely reducible over Y, and this includes F. This is in contradiction to H and G_1 being lin-separated modulo Y. This shows our result in the case when H is μ_{k-1}-lifted strong.

We now analyse the case when H is not μ_{k-1}-lifted strong, and therefore W is a vector space of linear forms. We start with a basis of $W \cap Y_1$ (call this basis M_1, and extend this to a basis of Y (call the new elements M_2) and also extend this to a basis of W (call the new elements M_3). Under the map $\varphi_{W,\alpha}$, the elements of M_1 and M_3 are mapped to multiples of a new variable. On M_2, the map acts as identity. We can now compose this map with a general projection of the image of Y, call this map ψ_Y. On the other hand, we can also write the composed map $\psi_Y \circ \varphi_{W,\alpha}$ as the composition of a general quotient of Y (call this map ψ'_Y), followed by a targeted quotient of the image of M_3, denoted φ'_W.

We will show that under the composition of the above maps, T_1, T_2 have pairwise rank at least $\xi_{k-1}(\delta_W)$. This implies the same is true under just the map $\varphi_{W,\alpha}$.

Suppose $F \in T_2 \setminus T_1$ is such that $\gcd(\psi_Y \circ \varphi_{W,\alpha}(F), \psi_Y \circ \varphi_{W,\alpha}(G_1)) \neq 1$. We now do a case analysis, based on whether $T_1 \Delta T_2$ is an irreducible symmetric difference, or a reducible one.

If $T_1 \Delta T_2$ is an irreducible symmetric difference, then $\psi'_Y(G_1)$ is irreducible. Since G_1 and H are lin-separated, this implies that $\psi_Y \circ \varphi_{W,\alpha}(G_1)$ is also irreducible. Therefore, if F exists, then it must be that $\psi_Y \circ \varphi_{W,\alpha}(F)$and $\psi_Y \circ \varphi_{W,\alpha}(G_1)$ are associate. This implies that some linear combination of F and G_1 lies in the kernel of $\psi_Y \circ \varphi_{W,\alpha} = \varphi'_W \circ \psi'_Y$, whence $H \in \mathrm{span}_{\mathbb{K}}\{F, G_1\}$. This contradicts the definition of H.

On the other hand, suppose $T_1 \Delta T_2$ is a reducible symmetric difference. In this case, either G_1 is a linear form and so is $\psi'_Y(G_1)$, or G_1 is a quadratic form and $\psi'_Y(G_1)$ is a product of linear forms. The same is true of F and $\psi_Y(F)$. The forms G_1 and F are relatively prime in R, therefore, the only common factor of $\psi'_Y(F)$ and $\psi'_Y(G_1)$ is z. Suppose $\psi'_Y(G_1) = \ell_1\ell_2$ and $\psi'_Y(F) = \ell_3\ell_4$. For $\gcd(\varphi'_W \circ \psi'_Y(F), \varphi'_W \circ \psi'_Y(G_1))$ to be different from 1, either one of ℓ_1, ℓ_2 has to become equal to ℓ_3, ℓ_4 under the second part of the map (which cannot happen, since H is a quadratic and these forms are linear), or $\ell_1\ell_2$ and $\ell_3\ell_4$ must be associate under the second part of the map, which would imply that their difference is a multiple of $\psi'_Y(H)$. This in turn again implies that $H \in \mathrm{span}_{\mathbb{K}}\{F, G_1\}$, a contradiction.

Therefore, such an F cannot exist in either case, and the pairwise rank of T_1, T_2 is at least $\xi_k^2(\delta)$. By construction, this value is at least $\xi_{k-1}(\delta_W)$. □

Lemma 3.10.37. *Suppose $H \in \mathcal{H} \setminus (T_1 \cup T_2)$. Assume that there are c other elements in $\mathcal{H} \setminus (T_1 \cup T_2)$ that are core collapsing for $T_1 \Delta T_2$. Let $W := AH_R(\mu_{k-1}, H)$ be the μ_{k-1}-lifted strong space obtained applying Corollary 3.6.6 to $\mathrm{span}_{\mathbb{K}}\{H\}$ in R. Let $\varphi_{W,\alpha} : R[y] \to R' := R[y]/I_{W,\alpha}$ be the graded quotient corresponding to $\alpha \in \mathbb{K}^{\dim(W)}$. For a general $\alpha \in \mathcal{Z}(H) \subseteq \mathbb{K}^{\dim(W)}$, the image of T_1, T_2 under the map $\varphi_{W,\alpha}$ has pairwise rank at least c.*

Proof. Without loss of generality, we assume that the c elements are $H_1, \ldots, H_c$. We focus on H_1. Recall that the core consists of $G_1, \ldots, G_s$ from $T_1 \Delta T_2$. Among these s forms, there are at most $2a$ forms such that G_i and H_1 are not lin-separated modulo Y among the remaining forms, there are at most $\xi_k^2(\delta)$ many forms such that $\mathrm{span}_{\mathbb{K}}\{G_i, H_1\} \cap T_1 \Delta T_2$ has only one element. In particular, there exists an index i_1 such that G_{i_1} and H_1 are lin-separated modulo Y, and such that $F_1 \in \mathrm{span}_{\mathbb{K}}\{H_1, G_{i_1}\} \cap T_1 \Delta T_2$. Similarly, we construct $F_2, \ldots, F_c$. Note that in this case, G_{i_j} cannot be linear: if it were, then F_i would also be linear, and this would contradict the fact that G_{i_j} and F_i are relatively prime.

We now claim that $F_1, \ldots, F_c$ are witnesses to the pairwise rank of T_1, T_2 under the map $\varphi_{W,\alpha}$. Suppose towards a contradiction that F_1 is no longer in the symmetric difference. Suppose we write $F_1 = \gamma H_1 + \gamma' G_{i_1}$. If G_{i_1} is an absolutely irreducible quadratic over Y, then so is F_1. Further, H and F_1 are lin-separated modulo Y. We are now in the same situation as in the proof of Lemma 3.10.36, and we can deduce that if F' is such that F_1 and H have non trivial gcd under the map $\varphi_{W,\alpha}$, then $H \in \mathrm{span}_{\mathbb{K}}\{F', F_1\}$. Therefore, $H \in \mathrm{span}_{\mathbb{K}}\{F', H_1, G_{i_1}\}$, which implies $F' \in \mathrm{span}_{\mathbb{K}}\{H, H_1, G_{i_1}\}$. For the same reason as above, F' must be absolutely irreducible over Y, W^{core}. This is a contradiction to the good disjoint sequence property. □

As in Section 3.10.2, while the above lemmas are stated for $T_1 \Delta T_2$, they also apply to every other symmetric difference. The rest of the argument is completed the same way. For each $T_i \Delta T_j$, we let $\mathcal{H}^{i,j} \subset \mathcal{H}$ be the set of forms which are core collapsing for $T_i \Delta T_j$. If $\left|\mathcal{H}^{i,j}\right| \geq \xi_k^2(\delta_U) + 1$, then for any $H \in \mathcal{H}$, either one of T_i, T_j vanishes under a targeted quotient of H (which happens if H belongs to one of T_i, T_j), or their pairwise rank is at least $\xi_k^2(\delta_U)$, by either Lemma 3.10.36 or Lemma 3.10.37.

Let $\mathcal{H}'$ be the union of $\mathcal{H}^{i,j}$ over those i, j such that $\left|\mathcal{H}^{i,j}\right| \leq \xi_k^2(\delta_U)$. The size of $\mathcal{H}'$ is at most $\binom{k}{2}\xi_k^1(\delta_U)$, and by the choice of β_k, we can pick $H \in \mathcal{H} \setminus \mathcal{H}'$. We claim that this H has the required property. The pairs i, j with $\mathcal{H}^{i,j} \geq \xi_k^2(\delta_U) + 1$ have already been handled. If i, j is such that $\mathcal{H}^{i,j} \leq \xi_k^2(\delta_U)$, then by construction either H belongs to one of T_i, T_j or H is core preserving for $T_i \Delta T_j$.

3.10.4 Rank bounds for identities

In this final subsection, we will combine the results from the previous section and prove our rank bounds. The proof proceeds by induction on the top fan-in k.

Theorem 3.10.38. *For each integer $k \geq 2$, there exists functions $\mu_k : \mathbb{N}^2 \to \mathbb{N}^2$ and $\xi : \mathbb{N}^2 \to \mathbb{N}$ with the following property: Let $S = \mathbb{K}[x_1, \cdots, x_N]$ be a polynomial ring. If $U \subseteq S_{\leq 2}$ is a graded vector space that is μ_k-lifted strong, and $\mathcal{C}$ is any simple and minimal $\Sigma^k \Pi \Sigma \Pi^2$ identity in $R := S/(U)$ then the rank of $\mathcal{C}$ is at most $\xi_k(\delta_U)$.*

Proof. We prove the theorem by induction on k. The base case is $k = 2$. In any ring R which is a UFD, there are no simple and minimal depth four identities with fan-in two. We define $\mu_2(d_1, d_2) := (A(3,1) + 3(d_1 + d_2), A(3,2) + 3(d_1 + d_2)$, where A is the function from Corollary 3.6.2. By the corollary, if U is any vector space that is μ_2-strong, then $S/(U)$ is a UFD, thus we can define $\xi_2(d_1, d_2) := 0$. This completes the base case.

We assume ξ_{k-1} and μ_{k-1} have been defined. Define $\xi_k^2 : \mathbb{N}^2 \to \mathbb{N}$ as $\xi_k^1(\delta) := \xi_{k-1}\big(\big(C_{\mu_{k-1}} \circ t_1\big)(\delta)\big)+1$ and $\xi_k^2 : \mathbb{N}^2 \to \mathbb{N}$ as $\xi_k^2(\delta) := \xi_{k-1}\big(\big(C_{2\mu_{k-1}} \circ t_1\big)(\delta)\big) + 1$. Define the function $M_k(\delta) : \mathbb{N}^2 \to \mathbb{N}$ as

$$M_k(\delta) := \max_{\delta' \in \mathbb{N}^2, \|\delta'\| = \binom{k}{2}\xi_k^1(\delta)} \mu_{k-1}(\delta + \delta' + (0, k^2\xi_k^2(\delta))).$$

Note that these are the same functions defined in Lemma 3.10.15, instantiated with our inductive bounds. Define $b_k(\delta) := k^2\big(\xi_k^1(\delta) + 1\big)$. Define $g_k^{dim}(\delta) := 4^{b_k(\delta)+1} + b_k(\delta)^3 + b_k(\delta) + 16$. Define $g_k^{ps}(\delta) := 3^{b_k(\delta)}M_k(\delta)$. Finally, let μ_k be the function μ_k^{low} defined in Lemma 3.10.27, instantiated with the above parameters. This completes the definition of μ_k.

We now define $\xi_k(\delta)$. This will be the max of three terms, one coming from Lemma 3.10.15, one coming from Lemma 3.10.27, and one coming from the rank bound of Kayal and Saraf [KS09]. The first of these terms is $\xi_{k-1}\big(C_{2\mu_{k-1}}(\delta + b_k(\delta))\big) \cdot 72 \cdot 2^{|\delta|_1} \cdot 6^{2 \cdot b_k(\delta)} \cdot C_{2\mu_{k-1}}(\delta + b_k(\delta))^2$. The second of these terms is $\xi_{k-1}\big(C_{2\mu_{k-1}}(\delta + \rho_k(\delta))\big) \cdot 72 \cdot 2^{|\delta|_1} \cdot 6^{2 \cdot \rho_k(\delta)} \cdot C_{2\mu_{k-1}}(\delta + \rho_k(\delta))^2$, where ρ_k is the function defined in Lemma 3.10.27 instantiated with the parameters above. The third term is $\xi_k^{\text{KS}}\big(C_{2\mu_{k-1}}(\delta + \rho_k(\delta))\big) \cdot 72 \cdot 2^{|\delta|_1} \cdot 6^{2 \cdot \rho_k(\delta)} \cdot C_{2\mu_{k-1}}(\delta + \rho_k(\delta))^2$, where ξ_k^{KS} is the rank bound for $\Sigma^k\Pi\Sigma$ identities over $\mathbb{C}$ proved in [KS09].

We now prove that the functions ξ_k and μ_k have the required property. Let U be any vector space with dimension sequence δ_U, and let $\mathcal{C}$ be any $\Sigma^k\Pi\Sigma\Pi^2$ identity in $R = S/(U)$. Suppose U is μ_k-strong. Note in particular that U is $\mu_{k-1} \circ t_1$-strong. By Lemma 3.10.15 instantiated with the above parameters, we deduce that one of three things happens (corresponding to the three cases).

- The first case is that $\dim \operatorname{span}_{\mathbb{K}}\{T_i \Delta T_j\} \leq b_k(\delta_U)$ for some i, j. In this case, if we define W to be the vector space obtained by starting with a basis of this span making the resulting vector space μ_{k-1}-lifted strong Corollary 3.6.6, then the dimension sequence of W is bounded coordinate wise by $C_{2\mu_{k-1}}(\delta + b_k(\delta))$. The image of $\mathcal{C}$ under a graded quotient of W has top fan-in $k-1$, and therefore by induction it has rank at most $\xi_{k-1}\big(C_{2\mu_{k-1}}(\delta + b_k(\delta))\big)$. We can deduce a rank bound on $\mathcal{C}$ by lifting the above rank bound using Proposition 3.6.17, and this rank bound is exactly the first of the two terms above that we maximised over to define ξ_k. Therefore, in this case the desired conclusion holds.

- The second case is the existence of a H that is μ_{k-1}-lifted strong, such that the image of $\mathcal{C}$ under the map $R \to R/(H)$ has pairwise rank at least $\xi_{k-1}^2(\delta)$, therefore also rank at least $\xi_{k-1}^2(\delta)$. However note that in this case, $R/(H)$ is isomorphic to a quotient of S by a μ_{k-1}-strong vector space of dimension sequence $\delta + (0, 1)$, and $\xi_{k-1}^2(\delta)$ is larger than $\xi_{k-1}(\delta + (0, 1))$. This contradicts the induction hypothesis, and therefore under the above parameters, this case cannot happen.

- The final case is the existence of a vector space V of dimension at most $g_k^{dim}(\delta)$ such that every quadratic in $\mathcal{C}$ is $g_k^{ps}(\delta)$-close to V. In this case, by Lemma 3.10.27, we can deduce that one of three things happens (here we use the fact that $\mu_k \geq \mu_k^{\text{low}}$ to be able to invoke this lemma): either there is a vector space Y of dimension at most $\rho_k(\delta_U)$ whose algebra contains a symmetric difference, or there is a vector space Y of dimension at most $\rho_k(\delta_U)$ such that every form in $\mathcal{C}$ is absolutely irreducible over $\mathbb{K}[Y]$, or there exists a form H such that a targeted quotient with respect to H kills a gate, and maintains pairwise rank at least $\xi_{k-1}^2(\delta)$. The first and the third case above are identical to the two cases above, and the same arguments apply. In the second case, the image of $\mathcal{C}$ under a graded quotient of Y is a $\Sigma^k\Pi\Sigma$ identity, and therefore a rank bound can be deduced by the proof of [KS09]. Lifting this rank bound gives us the desired conclusion. This completes this case. □

As a corollary, we have an identity testing algorithm.

Corollary 3.10.39. *There is a deterministic polynomial-time algorithm for identity testing of* $\Sigma^k\Pi\Sigma\Pi^2$ *circuits.*

Proof. This follows from the above rank bounds, and Theorem 3.2.1. □

Chapter 4

Algorithms for properties of varieties

From a computational point of view, a natural way of defining a variety $V \subset \mathbb{C}^n$ is by giving a set of equations $f_1, \ldots, f_m \in \mathbb{C}[x_1, \ldots, x_n]$ such that V is exactly the set of common zeroes of $f_1, \ldots, f_k$. In this chapter, we are interested in two fundamental computational problems related to such a variety V. The first is the task of checking if V is irreducible. The second is the even more basic task of checking if V is empty. This second problem is also known as the system satisfiability problem, since it is equivalent to checking if the system of equations $f_1, \ldots, f_m$ is satisfiable.

We will consider these two problems in two very different models of computation. For the first problem, we will consider the Turing model. Here, the inputs will be restricted to elements of $\mathbb{Z}[x_1, \ldots, x_n]$, and the algorithm will be allowed to depend polynomially on the bit complexity of the input. In fact we will allow the inputs to be given fairly succinctly, as algebraic circuits. It is easy to show that this problem is NP-hard, and in his thesis, Scheiblechner showed that the problem is in PSPACE [Sch07]. Our goal will be to try and situate the problem within PSPACE, in some special cases, for example when the zero set is defined ideal theoretically, and not just set theoretically. [1]

The second problem will be considered in the algebraic circuit model. We will design algebraic circuits that solve the projective version of the satisfiability problem. This model of computation allows the inputs to be arbitrary elements of any extension of $\mathbb{Q}$. We will focus on the regime where we know that this problem is solvable in polynomial time, in particular in the regime when the number of variables is a constant. We know that the problem is in VP, and like before, our goal is to situate the problem *within* VP. More concretely, the size of the inputs is $s\binom{d+n}{n}$, and it is well known that satisfiability can be solved by algebraic circuits of size $d^{O(n)}$, which is polynomial in the input size, and depth $O(n \log d)$ (see for example [Ier89] and the references within). We will focus on constructing circuits that have depth $O(n)$, and size $d^{O(n^2)}$.

4.1 Deciding irreducibility of varieties

Every non empty algebraic variety is a finite union of irreducible algebraic varieties. This is usually one of the first results proved in any treatment of algebraic varieties, and as a consequence, the study of these objects usually reduces to the case of irreducible varieties. From a computational perspective however, performing this decomposition is an expensive task, and the conventional wisdom is that this task should be avoided as much as possible ([BM93]).

[1] We say that V is *set theoretically defined* by the polynomials $f_1, \ldots, f_m$ if V is the zero set of $f_1, \ldots, f_m$. If $f_1, \ldots, f_m$ further have the property that $(f_1, \ldots, f_m)$ is radical, or equivalently by the Nullstellensatz that every polynomial g that vanishes on V lies in $(f_1, \ldots, f_m)$, then we say that V is *ideal theoretically defined* by $f_1, \ldots, f_m$.

Suppose $V \subset \mathbb{C}^n$ is an algebraic variety that is set theoretically defined by $f_1, \ldots, f_m$. Let $I := (f_1, \ldots, f_m)$. The ideal I admits a primary decomposition $I = \cap_{i=1}^{r} Q_j$ where Q_j are P_j-primary for prime ideals P_j. The irreducible components of V correspond exactly to those P_j among $P_1, \ldots, P_r$ that are minimal with respect to inclusion. Therefore, the above geometric task of deciding if an algebraic variety is irreducible reduces to the commutative algebraic task of checking if a given ideal $(f_1, \ldots, f_m)$ has one or more minimal primes. If V is defined ideal theoretically, then I has a primary decomposition of the form $I = \cap_{i=1}^{s} P_i$, where all the primes that occur are now minimal. Therefore, in this special case, the above geometric task reduces to the task of checking if I is prime.

Like most algebraic tasks, the above can be solved using Gröbner bases. Two well known algorithms using Gröbner bases are the algorithm of Gianni, Trager, and Zacharias [GTZ88] and the algorithm of Eisenbud, Huneke, and Vasconcelos [EHV92]. We also note that constructive methods for performing such decompositions were also studied in [Her26; Las05], and [Sei74; Sei78]. In the worst case, computing Gröbner bases requires exponential space [Dub90; Huỳ86]. Therefore, the above algorithms lie in EXPSPACE. This is not surprising, given that these algorithms solve a commutative algebraic problem, and even the most basic commutative algebra problem, that of ideal membership, is complete for EXPSPACE [MM82].

The original task at hand, the geometric task of deciding if V is irreducible is expected to be significantly easier than the task of primary decomposition. At a high level, this is because the primary decomposition task has some arithmetical aspects, and it seems like a large part of the complexity lies in these aspects of the problem. A more detailed discussion on this can be found in [BM93, Chapter 4]. This was made concrete in the works of [CG83] and [GH91], who gave single exponential time algorithms for computing the minimal primes P_i. If we want to actually compute these minimal primes, then exponential time might be unavoidable simply because the size of the output might be exponential in the input. In [BS07], a PSPACE algorithm was given for the task of counting the number of irreducible components. We note that all the above algorithms work for varieties that are defined set theoretically.

Our main result in this part of the thesis is to show that the task of deciding if an ideal theoretically defined variety is irreducible lies in PH, assuming the generalised Riemann hypothesis. We formally state the above results.

Theorem 4.1.1 (Constant round interactive protocols for primality)**.** *Let C be a constant free algebraic circuit of size s with integer constants that computes $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$. If $I := (f_1, \ldots, f_m)$ is radical and if the dimension of I is given, then the complexity of testing if I is prime lies in* coAM*, assuming GRH.*

Combining the above theorem with the algorithm from [Koi97, Theorem 4.1] that computes the dimension of a set theoretically defined variety, we obtain the following for radical ideals.

Theorem 4.1.2 (Primality testing in PH)**.** *Let C be a constant free algebraic circuit of size s with integer constants that computes $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$. If the ideal $I := (f_1, \ldots, f_m)$ is radical then the complexity of testing if I is prime lies in $\Sigma_3^p \cap \Pi_3^p$, assuming GRH.*

Before we proceed, we give a quick proof that the above task of deciding irreducibility is NP–hard, even for varieties defined ideal theoretically. The reduction is a slight variation of arithmetisation of SAT.

Proposition 4.1.3 (coNP-hardness of ideal primality testing)**.** *Given a 3CNF Φ on n variables, there exists a polynomial time algorithm whose output is a constant free circuit C_Φ computing polynomials $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$, along with their partial derivatives $\partial_i f_j$, such that the ideal $I := (f_1, \ldots, f_m)\mathbb{C}[x_1, \ldots, x_n] \subset \mathbb{C}[x_1, \ldots, x_n]$ has the following properties:*

- *I is a radical, equidimensional Cohen-Macaulay ideal*
- *I is prime if and only if Φ is unsatisfiable.*

Proof. Let $x_1^2 - x_1, x_2^2 - x_2, \ldots, x_n^2 - x_n, g_1, \ldots, g_r$ be the arithmetisation of Φ, where r is the number of clauses in Φ. The instance Φ is satisfiable if and only if the arithmetised system has a solution.

The reduction algorithm proceeds as follows. First check if $(0, \ldots, 0)$ is a solution to the above system. If it is, then the algorithm returns the system $f_1 = x_1(x_1 - 1)$, clearly there is a constant-sized circuit that computes $f_1, \partial_i f_1$.

Suppose $(0, \ldots, 0)$ is not a solution to the above system. Consider the set of polynomials $\{x_i^2 - x_i\}_{i \in [n]} \cup \{x_j g_i\}_{i \in [r], j \in [n]}$. Since each g_i has a small circuit, there is a small circuit that computes all the above polynomials and all their derivatives. Further, the set of zeroes of the above system is exactly those points on the boolean hypercube that correspond to a satisfying assignment of ϕ, and also the point $(0, \ldots, 0)$.

In each of the above cases, by the Lemma in [Sei74, Page 310], the ideal constructed is radical. In the first case the ideal is a complete intersection therefore CM. In the second case the ideal is zero dimensional, therefore CM. If Φ is satisfiable, then in either case the system created has at least two components and is therefore not prime. If Φ is not satisfiable, the system created has unique solution $(0, \ldots, 0)$, thus the ideal is prime. □

Our techniques are inspired by, and are a generalisation of the methods introduced by Koiran in [Koi96], where he showed that satisfiability can be decided in AM, assuming the generalised Riemann hypothesis. In the next section (Section 4.1.1), we will go over Koiran's argument in some detail. Then in Section 4.1.2 we will explain our results and techniques at a high level. To formalise our results, we will require the notion of the height of algebraic numbers and heights of varieties. In Section 4.1.3 we state these definitions and various useful properties of these heights. In Section 4.1.4 we will prove that certain properties of V are preserved when we go modulo a prime. Finally in Section 4.1.5 we will combine all of these and prove our main theorem.

4.1.1 An interactive proof for system satisfiability

Given $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$ of degree at most d, the satisfiability problem is deciding if the polynomials have a common zero in $\mathbb{C}^n$. We assume that the polynomials are given by means of a constant free algebraic circuit of size s. We denote by h the maximum of the bit complexities of the coefficients of the polynomials $f_1, \ldots, f_m$.

By the Nullstellensatz, the satisfiability problem is equivalent to deciding if the ideal $(f_1, \ldots, f_m)$ is trivial, that is, if $1 \in (f_1, \ldots, f_m)$. The decidability of this problem was established in [Her26], where Hermann proved that if $g \in (f_1, \ldots, f_m)$ then $g = \sum g_i f_i$ for some polynomials $g_i \in \mathbb{Q}[x_1, \ldots, x_n]$ whose degrees are bounded by a double exponential function. This reduces the above problem to a linear algebra problem of double exponential size, which can be solved in exponential space using the methods of Csanky [Csa76] and Berkowitz [Ber84]. For general ideal membership, we cannot do any better: Mayr and Meyer [MM82] proved that general ideal membership is complete for EXPSPACE.

For the special case of $1 \in (f_1, \ldots, f_m)$, single exponential degree bounds were shown for the polynomials $g_1, \ldots, g_m$ [Bro87; Kol88; Jel05]. This result is called the effective Nullstellensatz. Combined with the fast parallel linear algebra methods of [Csa76], the effective Nullstellensatz shows that satisfiability of systems can be checked in PSPACE. This is still the best known unconditional bound on the complexity of this problem. We note that there is a large body of work that studies the more general problem of solving systems, and finding polynomials $g_1, \ldots, g_m$ that witness $1 = \sum f_i g_i$ if the system is unsatisfiable. We refer the reader to the introduction of [KP96] for an extensive

survey of these results. In [KP96], the authors also give an algorithm for decidability that runs in time polynomial in $s, d^{O(n)}$.

One might expect the satisfiability problem to be in NP, since a common root of $f_1, \ldots, f_m$ is a certificate for yes instances for satisfiability. The issue with this is of course that the bit complexity of every common root of $f_1, \ldots, f_m$ can be exponential in the parameters h, m, d, n. Koiran's main idea was the following: if the system $f_1, \ldots, f_m$ is satisfiable, then for many primes $p \in \mathbb{N}$, the system $f_1, \ldots, f_m$ should have solutions in $\mathbb{F}_p$. Here and in the rest of this section, when we say that $f_1, \ldots, f_m$ have solutions in a field k, we mean that images of $f_1, \ldots, f_m$ under the extension of the natural map $\mathbb{Z} \to k$ have solutions in k^n. For any prime p whose bit complexity is polynomial in h, m, d, n, such a root can be certified, since the bit complexity of the root is polynomial in that of p.

Of course we need quantitative bounds on the number of primes for which this happens. Further, we will also require that the above proof system is sound, in the sense that if the system unsatisfiable, then there are no solutions in $\mathbb{F}_p^n$ for most p. This soundness is straightforward to show, using the effective Nullstellensatz (Theorem 2.3.8). If the system $f_1, \ldots, f_m$ is unsatisfiable, there exists polynomials $g_1, \ldots, g_m$ of degrees at most d^n such that $1 = \sum f_i g_i$. This is a linear system in the coefficients of g_i, and therefore if it has a solution then it has a rational solution. Further, by computing the size of the system and invoking Cramer's rule, we can deduce that there exists a set of rational polynomials $g_1, \ldots, g_m$ that further have a common denominator of bit complexity polynomial in h, d^{n^2}. Call this common denominator a, then for any prime p such that p does not divide a, the equation $1 = \sum f_i g_i$ is well defined in $\mathbb{F}_p[x_1, \ldots, x_n]$. For any such p, the system cannot even have a solution in $\overline{\mathbb{F}}_p$, let alone in $\mathbb{F}_p$. Therefore, there are only finitely many *bad primes*, that is, primes where there are no $\mathbb{F}_p$ solutions, and the above argument gives an upper bound that is polynomial in h, d^{n^2} for this set of bad primes.

Note that almost the exact same argument shows that if $f_1, \ldots, f_m$ is satisfiable, then for all except a constant number of primes p, the system $f_1, \ldots, f_m$ has solutions in $\overline{\mathbb{F}}_p$. However $\overline{\mathbb{F}}_p$ solutions are not of much use as certificates, because similar to $\mathbb{C}$ solutions, they could have exponential bit complexity, since the root might lie in an extension of $\mathbb{F}_p$ of exponential degree. For this argument, we really require roots that lie in $\mathbb{F}_p$, or at most in some extension of $\mathbb{F}_p$ of polynomial degree. This is what makes the statement non trivial.

Let us consider the most basic case, that of a single univariate polynomial. Suppose $f \in \mathbb{Z}[x]$ is a degree d univariate polynomial that is irreducible (for reducible polynomials we can just apply the following argument to its factors). We want to get a handle on the number of primes p such that f has a solution in $\mathbb{F}_p$. For each prime, the image of f in $\mathbb{F}_p[x]$ will factor into irreducible polynomials, and f has an $\mathbb{F}_p$-root if and only if one of the factors of f in $\mathbb{F}_p[x]$ is linear. The factorisation pattern of irreducible integer polynomials over primes has been studied classically. A theorem of Frobenius states that for any fixed set of positive numbers $d_1, \ldots, d_t$ such that $d_1 + \cdots + d_t = d$, the density of the set of primes p such that f factors into irreducible polynomials of degrees $d_1, \ldots, d_t$ in $\mathbb{F}_p[x]$ is exactly equal to the fraction of elements in the Galois group of the splitting field of f that have a cycle decomposition consisting of cycles of lengths $d_1, \ldots, d_t$. Specialising to what we require, the statement says that the density of primes p such that f has a solution in $\mathbb{F}_p$ is equal to the fraction of elements of the Galois group that have at least one fixed point. For a transitive subgroup of the symmetric group, one can show that this fraction is at least $1/d$. Therefore, the density fraction of primes p such that f has a root in $\mathbb{F}_p$ is at least $1/d$. While this theorem was proved by Frobenius, it is more commonly known as the Chebotarev density theorem, which is a strict generalisation of the above. We refer the reader to the survey [SL96] for more on the history of these theorems.

The one issue with the theorem of Frobenius is that it is only a statement about the density, that is, it is an asymptotic statement. From a computational perspective, we have to not only ensure that

density is high enough, but also that enough primes of polynomial bit complexity admit solutions to f. For an irreducible polynomial $f \in \mathbb{Z}[x]$, we define

$$\Pi_f(a) := \{p \leq a \,|\, f \text{ has a } \mathbb{F}_p \text{ solution}\},$$

and $\pi_f(n) := |\Pi_f(n)|$. The function $\pi(a) := \pi_0(a)$ is the usual prime counting function. The density statement says that $\pi_f(a)/\pi(a)$ converges to a number c satisfying $c > 1/d$. An error term can be derived from the above, which will imply that $\pi_f(a) = c\pi(a) + O(a^{1-\varepsilon} \log a)$. In this statement, the constants implicit in the big O statement depend on the polynomial f. To ensure that the value of $\pi_f(a)$ is large enough for some integer a of polynomial bit size, we also have to control the size of the these implicit constants. Such a bound is currently only known assuming the generalised Riemann hypothesis, and is not known unconditionally. Assuming the GRH, based on the works of [AO83; LO77], the following bound was deduced in [Koi96].

Theorem 4.1.4. *Assuming the GRH holds, there exists an absolute constant c such that*

$$\pi_f(a) \geq \frac{1}{d}\Big(\pi(a) - \log \mathrm{Disc}_x(f) - ca^{1/2} \log\Big(\mathrm{Disc}_x(f)a^d\Big)\Big).$$

In other words, the constant is proportional to the logarithm of the discriminant of f, and therefore can be bounded as a function of the bit complexity and degree of f.

With the univariate case handled, we return to the multivariate setting. The idea is to reduce to the univariate case. Suppose α is a solution to $f_1, \ldots, f_m$ that lies in $\overline{\mathbb{Q}}$ (if the system has $\mathbb{C}$-roots it must have $\overline{\mathbb{Q}}$-roots. Each coordinate α_i is an element of some algebraic extension K_i of $\mathbb{Q}$, and by the primitive element theorem we have $K_i = \mathbb{Q}(\gamma_i)$ for some algebraic number γ_i. By the primitive element theorem, there is a single element γ such that $\mathbb{Q}(\gamma) = \mathbb{Q}(\gamma_1, \ldots, \gamma_n)$. Let Q be the minimal polynomial of γ in $\mathbb{Z}[z]$, and let Q_i be polynomials such that $\alpha_i = Q_i(\gamma)$. The polynomials Q_i will have rational coefficients.

For any prime p such that Q has an $\mathbb{F}_p$-root, call this root γ_p, we can make sense of γ as an element of $\mathbb{F}_p$. Indeed γ_p will satisfy every algebraic relation that γ satisfies. Further, for such every prime p that does not divide any denominator in the coefficients of Q_i, we can make sense of the element α_i in $\mathbb{F}_p$, by setting $\alpha_{i,p} = Q_i(\gamma_p)$ in $\mathbb{F}_p$. The condition of not dividing the denominators ensures that this expression is well formed. Now again $\alpha_{i,p}$ satisfies the algebraic relations that α_i does, and it is straightforward to check that $\alpha_{1,p}, \ldots, \alpha_{n,p}$ is a $\mathbb{F}_p$-root of the original system.

All that remains is to bound the complexities of Q, Q_i. One way to do this is the following: take generic hyperplane sections to reduce to the case of dimension 0 (note that bounding the complexities is part of the analysis, and not the algorithm, so we can assume that we know what the dimension of V is). Now use multivariate resultants to project to each coordinate, which will allow us to bound the complexity of the minimal polynomial of γ_i. Finally, an effective version of the primitive element theorem will bound the complexities of Q, Q_i.

At the end of this process, the degree bound obtained on Q, Q_i is of the form d^{n^c} for some fixed constant c, and the bit complexities are bounded by $h \cdot d^{n^c}$. This exponential bound is undesirable, but examples given in [Koi96] show that it is unavoidable. Indeed if our original system was $x_1 - p_1^2, x_2 - p_2^2, \ldots, x_n - p_n^2$ where $p_1, \ldots, p_n$ are distinct primes, then the smallest extension of $\mathbb{Q}$ that contains a root of this system has degree 2^n.

Applying the univariate analysis to Q, and accounting for the denominators in Q_i, we can deduce the following: the density of primes p for which the system has $\mathbb{F}_p$ solutions is at least $1/d^{n^c}$, with effective error bounds which depend on h. This density is extremely small, and a random prime will not have have this property.

To summarise the above discussion, we have the following: if the original system $f_1, \ldots, f_m$ is unsatisfiable, then there are at most $h^c \cdot d^{n^c}$ primes p such that $f_1, \ldots, f_m$ have an $\mathbb{F}_p$ solution, for some constant c. If the system is satisfiable, then the density of primes for which the system has an $\mathbb{F}_p$-root is at least $1/d^{n^c}$, with effective error bounds. Therefore, if we pick a bound N of size polynomial in $h^c \cdot d^{n^c}$, then $\pi_I(N)$ will be less than $h^c \cdot d^{n^c}$ if the system is unsatisfiable, and more than $2h^c \cdot d^{n^c}$ if the system is satisfiable. The exact bound N depends on the error terms, and therefore getting effective error bounds is crucial. To distinguish these two cases, the Goldwasser Sipser protocol [GS86] is used. We state the final result, so we can invoke it later.

Theorem 4.1.5 ([Koi96])**.** *Let C be a constant free algebraic circuit of size s with integer constants that computes $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$. There is a* AM *protocol that certifies the existence of a common solution of $f_1, \ldots, f_m$.*

We make two remarks about Theorem 4.1.5 and its proof that will be crucial in the way we invoke them. The first of these is regarding the representation of input polynomials. The AM protocols for the above problems involve evaluating the polynomials $f_1, \ldots, f_m \pmod p$ at points in $\mathbb{F}_p$, where p is a prime of bit complexity polynomial in the input. Therefore, the polynomials $f_1, \ldots, f_m$ are allowed to be given either as white box circuits of polynomial size (here the size includes the bit complexity of the constants in the circuit), or more generally as black-boxes that allow mod p queries.

The second remark is regarding the parameters. Suppose $\deg f_i \leq d$ and $\mathrm{ht}(f_i) \leq h$. The length of the messages and the computation done by Arthur in the protocol is polynomial in $\log\left(h \cdot 2^{(n \log \sigma)^c}\right)$, for a universal constant c, where $\sigma := dm + 2$. Therefore, if we create a system of polynomials $g_1, \ldots, g_m$ with $\deg g_i \leq d^n$ and $\mathrm{ht}(g) \leq h \cdot 2^{(n \log \sigma)^c}$, then the protocol applied to $g_1, \ldots, g_m$ still run in time $\mathrm{poly} n, m, d, h$ as long as we ensure that $g_1, \ldots, g_m$ have circuits of size $\mathrm{poly} n, m, d, h$. The same applies for the results of [Koi97].

In the next section, we will give a high level overview of how we adapt this framework for our problem.

4.1.2 High level ideas

Let us begin by establishing the notation we will use throughout. We will use $f_1, \ldots, f_m$ to denote the input polynomials, these will be elements of $\mathbb{Z}[x_1, \ldots, x_n]$. We use d to denote a bound on the degrees of f_i, and h to denote the bound on the bit complexities. In other words, every coefficient of every f_i is an integer with absolute value at most 2^h. This quantity is usually called the height of the polynomials $f_1, \ldots, f_m$. We will use I to denote the ideal $(f_1, \ldots, f_m)\mathbb{Q}[x_1, \ldots, x_n]$, and V to denote the variety $V(I) \subset \overline{\mathbb{Q}}^n$.

Suppose $p \in \mathbb{N}$ is a prime. We will often overload notation and also use f_i to denote the image of f_i under the map $\mathbb{Z}[x_1, \ldots, x_n] \to \mathbb{F}_p[x_1, \ldots, x_n]$. If there is chance of ambiguity we will use $f_{i,p}$ instead. We will use I_p to denote the ideal of $\mathbb{F}_p[x_1, \ldots, x_n]$ generated by $f_{1,p}, \ldots, f_{n,p}$. We will use V_p to denote the zero set of I_p, this is a subset of $\overline{\mathbb{F}}_p^n$. We will use V_p' to denote $V_p \cap \mathbb{F}_p^n$, the set of $\mathbb{F}_p$ rational points of V_p.

We remark that I_p and V_p depend on the inputs $f_1, \ldots, f_m$, and not just the ideal I. For example, if $f_1 = 2x_1$ and $f_2 = 2x_2$, then $I = (x_1, x_2)$, but $I_2 = (0)$. We admit that this is not optimal notation, but we are willing to live with it for the sake of some brevity.

With the notation established, we move on to the high level ideas. For this section, we only deal with the case of varieties that are ideal theoretically defined. Suppose $V = V_1 \cup \cdots \cup V_s$ is the irreducible decomposition of V, and r is the dimension of V. We consider two cases. The first is that $\dim V_i = r$ for all i, and the second is that $\dim V_i < r$ for some i. The latter case will be the easy case, and we start with that.

Recall the notion of the tangent space at a point v on a variety V. If v is nonsingular, then the tangent space at v, denoted T_vV is a linear approximation of V at v. It is a linear variety passing through v whose dimension is the same as the dimension of the component of V containing v. If a point lies on two components, then the point is singular, therefore the above dimension is well defined. Since V is ideal theoretically defined, the dimension of the tangent space at v is the dimension of the kernel of the Jacobian of $f_1, \ldots, f_m$ evaluated at v. If $J_1, \ldots, J_t$ is the set of all $(n-r) \times (n-r)$ minors of the Jacobian $\mathcal{J}$, then V has a component of dimension less than r if and only if the system $f_1, \ldots, f_m, J_1, \ldots, J_t$ has a common root. As described the number of polynomials in this system is exponential in the input, however some standard methods will allow us to reduce the number down to something polynomial. We can thus reduced the task of certifying reducibility to the task of certifying that a system of equations has a solution. This case can therefore be handled by the methods of Koiran discussed above.

We are left with the hard case, when every V_i has dimension r. We show that if V is reducible, then V_p is reducible for most primes p, and similarly if V is irreducible then V_p is irreducible for most primes p. To show this, we essentially want to express these conditions in terms of conditions that are preserved when we reduce mod p.

As an illustrative example, suppose $n = 2$, and $r = 1$, so V is union of curves. Suppose further that I is irreducible. In this case we have $V = V(h)$ for some polynomial $h \in \mathbb{Q}[x_1, x_2]$ that is irreducible. We also have $h \in I$, therefore $h = \sum g_i f_i$ for some polynomials $g_i \in \mathbb{Q}[x_1, x_2]$.

If p is any prime that does not divide the denominators of g_i, then $h \in I_p$. It turns out that the property of being absolutely irreducible is also preserved when reducing mod p for most primes p. If p satisfies both these conditions, then V_p must be irreducible.

There still remains one problem, that of certifying that V_p is reducible or irreducible for a given prime p. Recall that in Koiran's proof, $\mathbb{F}_p$-roots were being certified, since these have polynomial bit size. To obtain polynomial sized certificates for the reducibility of V_p, we will use the Lang–Weil bounds. These bounds essentially relate the number of components of V_p with the sizes of V_p', which recall are the $\mathbb{F}_p$-points of V_p. This latter set is finite, and therefore we can hope to certify its size. We use the following effective version of the bounds from [CM06].

Theorem 4.1.6 (Effective Lang–Weil bound)**.** *Let $V_p \subset \overline{\mathbb{F}}_p^n$ be an absolutely irreducible affine variety defined over $\mathbb{F}_p$ of dimension r and degree D. Suppose also that $p > 2(r+1)D^2$. Then*

$$\left|\left|V_p \cap \mathbb{F}_p^n\right| - p^r\right| < (D-1)(D-2)p^{r-1/2} + 5D^{13/3}p^{r-1}.$$

As a consequence of the above bounds, we can deduce the following. If V_p is irreducible, then V_p' has size at most $p^r + E$, where E is the error term $(D-1)(D-2)p^{r-1/2} + 5D^{13/3}p^{r-1}$. If V_p is reducible, and has at least two components that are defined over $\mathbb{F}_p$, then V_p' has size at least $2p^r - E$. The gap between the sizes of these two sets is large enough in the two cases to apply the Goldwasser Sipser protocol. We note that the $\mathbb{F}_p$-definability condition in the above bounds is crucial, and the bounds do not hold without this assumption. Therefore, we will have to show something slightly stronger than what we claimed above: if V is reducible, we have to show not only that V_p is reducible, but also that at least two of its components are $\mathbb{F}_p$-definable.

The general case is harder since V might not be defined by a single equation. We will tackle the general case using projections to the bivariate case. Similarly, showing that reducedness is preserved is also harder. For this, we will essentially apply the above argument to say V_1 and V_2 instead of just to V. The hard part here is that V_1, V_2 are no longer defined over $\mathbb{Z}$, but over some finite extension of $\mathbb{Q}$. Therefore we will have to resort to using the theory of heights on algebraic numbers in order to control which primes are good and bad. We will go over this in more detail in the next subsection.

In the next section, we will introduce some basic notions from algebraic number theory that we will require to carry out the above. We will also define a notion of height of varieties.

4.1.3 Heights of algebraic numbers and varieties

Algebraic number theory preliminaries

We state some algebraic number theory facts that we will need. We do not provide any proofs since these facts are standard. An excellent exposition for much of this material is [Mil20, Chapter 2, 3].

Let $q \in \mathbb{Z}[z]$ be a monic irreducible polynomial with $\deg q = e$, and let ξ be a root of q. Let $K := \mathbb{Q}(\xi)$ be the algebraic extension of $\mathbb{Q}$ generated by ξ, we have $K \cong \mathbb{Q}[z]/(q)$. Let O_K be the ring of integers of K, that is, the set of elements $\beta \in K$ such that β satisfies a monic equation with coefficients in $\mathbb{Z}$. We have $\mathbb{Z} \subset O_K$, and O_K is integrally closed in K. The discriminant of q, denoted $\mathrm{Disc}_z(q)$ is defined to be $\mathrm{Res}_z(q, \partial q)$.

Usually the ring of integers O_K is different from $\mathbb{Z}[\xi]$. We have $\mathbb{Z}[\xi] \subset O_K$, but this inequality might be strict. However, it holds that $O_K \subset (1/\mathrm{Disc}_z(q))\mathbb{Z}[\xi]$. Therefore we can represent any algebraic integer as an element of $\mathbb{Q}[z]$, where the coefficients have common denominator $\mathrm{Disc}_z(q)$.

The ring O_K is a Dedekind domain, and every nonzero prime ideal of O_K is maximal. For any prime $\mathfrak{p} \subset O_K$, there is a unique prime $p \in \mathbb{N}$ such that $\mathfrak{p} \cap \mathbb{Z} = (p)$, in which case we say $\mathfrak{p}$ lies above p. While O_K is not factorial in general, it admits unique factorisation of ideals. Every ideal $I \subset O_K$ can be the written uniquely as a product of primes, $I = \prod \mathfrak{p}_i^{e_i}$. In particular, the ideal (p) for a prime $p \in \mathbb{N}$ itself can be factored as $(p) = \prod \mathfrak{p}_i^{e_i}$. The set of prime ideals of O_K occurring in this factorisation of (p) are exactly the prime ideals of O_K that lie above p. When $p \nmid \mathrm{Disc}_z(q)$, $e_i = 1$ for all i, if this holds we say that p is unramified.

For any unramified p, and $\mathfrak{p}$ lying above p, the quotient map $O_K \to O_K/\mathfrak{p}$ is well structured. Since $\mathfrak{p}$ is maximal, and since $p \in \mathfrak{p}$, the quotient $O_K/\mathfrak{p}$ is a field of characteristic p, and is in fact a finite extension of $\mathbb{F}_p$. The prime $\mathfrak{p}$ corresponds to an irreducible factor q_1 of $q \pmod p$, such that $\mathfrak{p} = pO_K + q_1(\xi)O_K$. It follows that $O_K/\mathfrak{p} \cong \mathbb{F}_p[z]/q_1(z)$. Composing with the inclusion map $\mathbb{Z}[\xi] \to O_K$ gives us a map $\mathbb{Z}[\xi] \to \mathbb{F}_p[z]/q_1(z)$. We usually apply this to the case when q_1 is a linear polynomial, and therefore $\mathbb{F}_p[z]/q_1(z) = \mathbb{F}_p$. This map therefore extends the usual map $\mathbb{Z} \to \mathbb{F}_p$. A proof of the facts in the last paragraph can be found in [Mar18, Theorem 27, Chapter 3] and the discussion preceding it.

Heights of algebraic numbers

Let us for a moment consider Koiran's argument again. If $f_1, \ldots, f_m$ is an unsatisfiable system, there is a number a such that for any prime p that does not divide a, the system has no $\overline{\mathbb{F}}_p$-roots. If the system is satisfiable, there is a number b (which is the common denominator of the polynomials Q_i in the notation of the previous section), and a polynomial Q such that for any prime p that does not divide b, and such that Q has a $\mathbb{F}_p$-root, the system has $\mathbb{F}_p$-roots. In each case, the number of primes p that divide a or b was bounded simply by the height of a, b, that is, $\log(a), \log(b)$. Rephrased in the language of the above paragraph, these heights give a bound on the number of primes p such that the image of a or b is zero under the map $\mathbb{Z} \to \mathbb{F}_p$.

In our applications, we will similarly want to control primes p under which a given algebraic integer $\alpha \in O_K$ is mapped to zero under the map $O_K \to O_K/\mathfrak{p}$ for some prime $\mathfrak{p}$ over p. Here is one way to do this for elements of $\mathbb{Z}[\xi]$. Suppose $\alpha = Q(\xi)$ where $Q \in \mathbb{Z}[z]$ is a polynomial. We can assume without loss of generality that Q has degree less than e. The polynomial Q is relatively prime to the minimal polynomial q of ξ. Therefore, the resultant $\mathrm{Res}_z(Q, q)$ is a nonzero integer. For any prime p that does not divide this resultant, Q and q do not have any common roots in

$\overline{\mathbb{F}}_p$. Therefore the image of α under the quotient is nonzero for any such prime. The heights of the coefficients of Q act as a measure of the height of α. We call this the *elementary height* of α in $\mathbb{Z}[\xi]$. The original proof of our result in [GOS25a] used this notion of height of algebraic numbers. While easy to define, this notion of height suffers from some drawbacks. For example it depends not just on α but also on ξ, and it is tedious to track how this notion of height changes under various operations.

In this thesis, we will instead use a slightly more sophisticated notion of height for algebraic numbers. The definition of the height function will assume some background on algebraic number theory. We refer the reader to [Wal00, Chapter 3] for the necessary background. Let M_K be the set of absolute values in the number field K that extend the absolute values in $\mathbb{Q}$, let M_K^∞ denote the Archimedean absolute values. For each $\nu \in M_K$ we denote the local degree of K at ν as the degree of the extension $[K_\nu : \mathbb{Q}_{\nu_0}]$, where K_ν is the completion of K with respect to ν and ν extends ν_0. We assume that the absolute values are picked such that the product formula holds, that is, $\prod_{\nu \in M_K} |\alpha|_\nu^{N_\nu} = 1$ for all nonzero $\alpha \in K$. A proof that we can pick such absolute values can be found in the references within [Wal00, Chapter 3].

Given a nonzero element $\alpha \in K$, the logarithmic Weil height is defined as

$$\mathrm{ht}(\alpha) := \frac{1}{[K:\mathbb{Q}]} \sum_{\nu \in M_K} N_\nu \log(\max(1, |\alpha|_\nu)).$$

This is a value that only depends on α, and does not depend on K. The height function satisfies $\mathrm{ht}(\alpha_1\alpha_2) \leq \mathrm{ht}(\alpha_1)\,\mathrm{ht}(\alpha_2)$ and $\mathrm{ht}(\alpha_1 + \alpha_2) \leq \mathrm{ht}(\alpha_1) + \mathrm{ht}(\alpha_2) + \log 2$ (for example see [Wal00, Property 3.3]).

The following lemma shows the logarithmic Weil height function can be used to control the number of primes modulo which a given algebraic number is mapped to 0, similar to how the bit complexity was used in Koiran's argument.

Lemma 4.1.7. *Suppose $\alpha \in O_K$ is a nonzero algebraic integer. There are at most $[K:\mathbb{Q}]\,\mathrm{ht}(\alpha)$ many primes $\mathfrak{p}$ such that $\alpha \in \mathfrak{p}$.*

Proof. For all non-archimedean ν we have $|\alpha|_\nu \leq 1$, therefore the non-archimedean absolute values do not contribute to the height. Recall that our absolute values were picked so the product formula holds, therefore

$$\sum_{\nu \in M_K^\infty} N_\nu \log |\alpha|_\nu = - \sum_{\nu \notin M_K^\infty} N_\nu \log |\alpha|_\nu.$$

The expression $[K:\mathbb{Q}]\,\mathrm{ht}(\alpha)$ is larger than the expression on the left (because of the max term), therefore is also an upper bound for the expression on the right of the above equation. For any prime $\mathfrak{p}$ that extends a prime p, we have $\alpha \in \mathfrak{p}$ if and only if the valuation ν corresponding to $\mathfrak{p}$ is such that $|\alpha|_\nu < 1$. From the definition of the absolute value, it follows that if $|\alpha|_\nu < 1$ then in fact $-N_\nu \log |\alpha|_\nu \geq 1$. We can therefore conclude that there are at most $[K:\mathbb{Q}]\,\mathrm{ht}(\alpha)$ many primes $\mathfrak{p}$ such that $\alpha \in \mathfrak{p}$. □

We will require a notion of simultaneous height of a set of elements. Given a set of elements $\mathcal{A} \subset K$, we define the height $\mathrm{ht}(\mathcal{A})$ as

$$\mathrm{ht}(\mathcal{A}) := \frac{1}{[K:Q]} \sum_{\nu \in M_K} N_\nu \log\left(\max_{\alpha \in \mathcal{A}}(1, |\alpha|_\nu)\right).$$

We have $\mathrm{ht}(a) \leq \mathrm{ht}(\mathcal{A})$ for all $a \in \mathcal{A}$. Finally, for a polynomial f, we define $\mathrm{ht}(f)$ to be the height of the set of coefficients of f. For a set of polynomials, the simultaneous height is just the height

of the set of all coefficients. Note that for subsets of $\mathbb{Z}$ or polynomials with coefficients in $\mathbb{Z}$, this notion of height is exactly the same as the max bit complexity of the elements of the set and the max bit complexity of the coefficients of the polynomial respectively, which matches the usual notion of height. Indeed in this case we can compute the height in the field $K = \mathbb{Q}$, therefore $[K : \mathbb{Q}] = 1$ and there is only absolute value, the usual one.

The following will be used frequently, and follows at once from [Wal00, Lemma 3.8].

Lemma 4.1.8. *Suppose M is a $D \times D$ matrix and suppose $h = \mathrm{ht}(\{M_{i,j}\})$ is the simultaneous height of its entries. Then* $\mathrm{ht}(\det M) \leq D \log(D) + Dh$.

Relationship between Weil height and the elementary height

In this section, we show how bounds on the Weil height h imply bounds on the elementary height, and how bounds on the elementary height imply bounds on the Weil height. As we will see, the Weil height has a number of advantages, but the bit complexity of actually storing an algebraic number on a computer is controlled by the elementary height. The below arguments will allow us to start with the elementary heights of a given set of polynomials, switch to the Weil heights, use various results to manipulate them, and finally switch back to the elementary height to show our complexity theoretic results.

For any algebraic integer $\alpha \in \overline{\mathbb{Q}}$, we can use the height $\mathrm{ht}(\alpha)$ to bound the coefficients of the minimal polynomial of α, as long as we also have a bound on the degree of this minimal polynomial. Suppose $z^d + a_{d-1}z^{d-1} + \cdots + a_0$ is the minimal polynomial of α. By [Wal00, Lemma 3.11], we have $\log(|a_i|) \leq d(\mathrm{ht}(\alpha) + \log 2)$.

Now suppose we have a monic polynomial $q \in \mathbb{Z}[z]$ of degree e, and ξ is a root of q. By [Wal00, Lemma 3.11] we have also $e\,\mathrm{ht}(\xi) \leq \log(\mathrm{ht}(q)) + (1/2)\log(e+1)$. Here since q is an integer polynomial, $\mathrm{ht}(q)$ is simply the logarithm of the largest coefficient of q in absolute value. In this way, we can use the coefficients of the minimal polynomial to also bound the Weil height.

Now suppose we have $\alpha \in \mathbb{Z}[\xi]$. Suppose $\alpha = \sum_{i<e} Q_i \xi^i$. Recall that our definition of the elementary height of such α was simply the height of the integer polynomial $\sum Q_i z^i$. Suppose H is the elementary height of this integer polynomial. We can use this to bound its length, that is, the sum of the absolute values of its coefficients, by $e \cdot 2^H$. Using [Wal00, Lemma 3.7], we can bound $\mathrm{ht}(\alpha)$ by $H + \log e + e\,\mathrm{ht}(\xi)$.

Finally, suppose we have $\alpha \in O_K$ and a height bound on α, we show how to derive a bound on the elementary height of α. The following method was suggested to me by Sriram Gopalakrishnan. We assume $\alpha \in \mathbb{Z}[\xi]$ since the elementary height is only defined in this case. In any case, since $\mathrm{Disc}_z(q)\alpha$ is always in $\mathbb{Z}[\xi]$, the same method can be used to obtain a bound for the representation of α as an element of $\left(\frac{1}{\mathrm{Disc}_z(q)}\right)\mathbb{Z}[\xi]$.

Lemma 4.1.9. *Suppose $q \in \mathbb{Z}[z]$ is the monic minimal polynomial of an algebraic integer ξ, and suppose* $\deg q = e$. *Suppose $\alpha = \sum Q_i \xi^i$ is the unique representation of an element $\alpha \in \mathbb{Z}[\xi]$ with degree less than e. Then the elementary height of α is bounded by $2eh + e\log e$ where $h := \max(\mathrm{ht}(\alpha), \mathrm{ht}(\xi))$.*

Proof. We have $\mathbb{Q}(\alpha) \subset \mathbb{Q}(\xi)$, therefore the minimal polynomial of α has degree less than e, call this degree d. Let $\xi_1, \ldots, \xi_e$ be the conjugates of ξ and let $\alpha_1, \ldots, \alpha_d$ be the conjugates of α. Suppose $\xi = \xi_1$ and $\alpha = \alpha_1$. Suppose $\sigma_1, \ldots, \sigma_e$ are automorphisms of the splitting field of ξ such that $\sigma_i(\xi) = \xi_i$. We have $\sigma_i(\alpha) = Q(\xi_i) = \alpha_{j_i}$ for some $j_i \in [d]$. In fact the multiset of elements $\sigma_i(\alpha)$ for all $i \in [e]$ is exactly the set $\alpha_1, \ldots, \alpha_e$, with each element occurring e/d times.

If we let M be the matrix defined by $M_{i,j} = \xi_i^{j-1}$, and we let $\boldsymbol{\alpha}$ be the vector with $\boldsymbol{\alpha}_i = \alpha_{j_i}$ then we get the linear equation $Mv = \boldsymbol{\alpha}$, and the solution v are the coefficients of Q. By Cramer's rule, each coefficient is a quotient of two determinants of size e, with entries having heights bounded by h. Using Lemma 4.1.8, we can deduce that each coefficient has height bounded by $2eh + 2e\log e$. Therefore, the elementary height of α is bounded by $2eh + e\log e$. □

Given either a bound on the height of ξ or a bound on the height of the minimal polynomial $q \in \mathbb{Z}[z]$ of ξ, we can bound the height of $\mathrm{Disc}_z(q)$ as follows.

Lemma 4.1.10. *Suppose $q \in \mathbb{Z}[z]$ is the monic minimal polynomial of an algebraic integer ξ, and suppose $\deg q = e$. Then $\mathrm{ht}(\mathrm{Disc}_z(q)) \leq O(e\,\mathrm{ht}(q)\log e)$ and $\mathrm{ht}(\mathrm{Disc}_z(q)) \leq O(e^2\,\mathrm{ht}(\alpha)\log e)$.*

Proof. Recall that $\mathrm{Disc}_z(q) = \mathrm{Res}(q, q')$. We have $\mathrm{ht}(q') \leq \mathrm{ht}(q) + \log e$. Each entry of $\mathrm{Res}\,(q, q')$ is a coefficient of either q or q' therefore has height at most $\mathrm{ht}(q) + \log e$. The size of the matrix is at most $2e$, therefore by Lemma 4.1.8 we have $\mathrm{ht}(\mathrm{Res}(q, q')) \leq O(e(\mathrm{ht}(q) + \log e))$ which we crudely bound by $O(e\,\mathrm{ht}(q)\log e)$. Finally, since $\mathrm{ht}(q) \leq e\,\mathrm{ht}(\alpha) + e\log 2$ ([Wal00, Lemma 3.11]), we obtain the second statement in the lemma. □

Recall the definition of $\Pi_f(a)$ for an irreducible polynomial $f \in \mathbb{Z}$. It is the set of primes $p \leq a$ such that f has an $\mathbb{F}_p$-root. For an algebraic integer ξ we can define $\Pi_\xi(a)$ to be $\Pi_q(a)$ where q is a minimal polynomial for ξ. Similarly we can define the counting function $\pi_\xi(a)$. Using the above bound on the discriminant along with Theorem 4.1.4, and the fact that $\pi(a) \geq a/\log a$ for a large enough, we can deduce the following.

Corollary 4.1.11. *Suppose ξ is an algebraic integer of height h and degree e. Then*

$$\pi_\xi(a) \geq \frac{1}{e}\left(\frac{a}{\log a} - c \cdot h \cdot a^{1/2}\log a \cdot e^2 \log e\right),$$

where c is a universal constant.

Absolute irreducibility of polynomials under quotients

Recall that a polynomial $f \in O_K[x_1, \ldots, x_n]$ is called absolutely irreducible if it is irreducible as a polynomial in $\mathbb{C}[x_1, \ldots, x_n]$. The height of the polynomial $\mathrm{ht}(f)$ will allow us to control how many primes p are such that an absolutely irreducible polynomial f is mapped to an absolutely reducible polynomial under the extension of the map $O_K \to O_K/\mathfrak{p}$ for some prime $\mathfrak{p}$ above p. To get this bound, we will use the Noether equations of [Kal95].

Lemma 4.1.12. *Suppose $f \in O_K[x_1, \ldots, x_n]$ is an absolutely irreducible polynomial of degree d. There exists an element $b_f \in O_K$ with $\mathrm{ht}(b_f) \leq c \cdot h \cdot d^6 \cdot n$ for some universal constant c with the following property. For any prime $\mathfrak{p}$ such that $b_f \notin \mathfrak{p}$, the image of f under the extension of $O_K \to O_K/\mathfrak{p}$ remains nonzero.*

Proof. Let b_1 be any nonzero coefficient of a degree d monomial in f. We have $\mathrm{ht}(b_1) \leq \mathrm{ht}(f)$, and for any prime $\mathfrak{p}$ such that b_1 is nonzero in $O_K/\mathfrak{p}$, the image of f under the quotient has degree d.

In [Kal95, Theorem 7], a finite set of polynomials $\{\Phi_t\}$ are constructed with integer coefficients that define the set of absolutely reducible polynomials of degree d in n variables over any field. In other words, suppose $\mathbb{F}$ is any field, and g is any degree d polynomial in $\mathbb{F}[x_1, \ldots, x_n]$. Let $\Phi_t^{\mathbb{F}}$ be the image of Φ_t under the map $\mathbb{Z} \to \mathbb{F}'$, where $\mathbb{F}'$ is the prime subfield of $\mathbb{F}$ (so if $\mathbb{F}$ has characteristic 0 then $\mathbb{F}' = \mathbb{Q}$ otherwise $\mathbb{F}' = \mathbb{F}_p$ where p is the characteristic of $\mathbb{F}$). The polynomial g is reducible in $\overline{\mathbb{F}}[x_1, \ldots, x_n]$ if and only if $\Phi_t^{\mathbb{F}}$ vanish on the coefficients of g for all t.

Now since f is absolutely irreducible, and since $\Phi_t^K = \Phi_t$, there is some polynomial in this set, say Φ_1 such that Φ_1 does not vanish on the coefficients of f. [Kal95, Theorem 7] also guarantees that $\deg \Phi_t \leq 12d^6$ for all t, and the length of Φ_t, that is, the sum of absolute values of its coefficients is bounded by $d^{c_1 d^6 n}$ for some universal constant c_1. Let b_2 be the evaluation of Φ_1 at the coefficients of f. An application of [Wal00, Lemma 3.7] using the bound on the length of the Φ_1 above along with the bound of h on the heights of the coefficients of f allows us to bound $\mathrm{ht}(b_2)$ by $c_1 d^6 n \log(d) + h \cdot 12d^6$, which is less than $c_2 h d^6 n$ for a universal constant c_2. Set $b_f := b_1 b_2$, we have $\mathrm{ht}(b_f) \leq c h d^6 n$ for a constant c.

Now let $\mathfrak{p}$ be any prime such that b_f is nonzero in $O_K/\mathfrak{p}$, and let $\mathfrak{p}$ be above $p \in \mathbb{Z}$. The image of f under the extension of the map $O_K \to O_K/\mathfrak{p}$ still has degree d since b_1 remains nonzero. By [Kal95, Theorem 7], the evaluation of $\Phi_1^{O_K/\mathfrak{p}}$ on the image of f is exactly equal to the image of b_2 under the map $O_K \to O_K/\mathfrak{p}$. Therefore, this evaluation also remains nonzero, and we can deduce that f remains absolutely irreducible. □

We note that a version of the above statement with the elementary height function appears as [Kal95, Theorem 8]. In our proofs, since we will be manipulating Weil heights, we will instead use the above version of the statement to avoid having to translate back and forth.

An arithmetic Nullstellensatz and heights of varieties

The following is an arithmetic version of the effective Nullstellensatz.

Theorem 4.1.13 ([KPS01, Theorem 3.6]). *Let K be a number field. Let $f_1, \ldots, f_m \in O_K[x_1, \ldots, x_n]$ be polynomials without common zeroes. Set $d := \max \deg f_i$ and $h := \mathrm{ht}(f_1, \ldots, f_m)$.*

Then there exists $a \in O_k$ nonzero and $g_1, \ldots, g_m \in O_K[x_1, \ldots, x_n]$ such that

- $a = \sum f_i g_i$,
- $\deg g_i \leq 4nd^n$,
- $\mathrm{ht}(a, g_1, \ldots, g_m) \leq 4n(n+1)d^n(h + \log m + (n+7)\log(n+1)d)$.

Note that the naive bound obtained by using Cramer's rule for the height of a was of the type $h \cdot d^{n^2}$, which is much worse than the above bound. The advantage of using this height machinery might still not be obvious, especially since the above improvement only changes the polynomial factors in our algorithm. The real advantage of this machinery comes from extending it to define heights on varieties. [2]

The height function on varieties we use is the height function defined in [KPS01], which itself is inspired by the works [Phi91; Phi94; Phi95; Nes85; Nes87]. We require the definition of the S_n^r Mahler measure. Suppose $f \in \mathbb{C}[X_r, \ldots, X_r]$ is a multihomogeneous polynomial in r groups of n variables each. We define the logarithmic Mahler measure as

$$m(f; S_n^r) = \int_{S_n^r} \log|f| \mu_n^r,$$

where S_n^r is the r-fold product of the unit sphere $S_n \subset \mathbb{C}^n$ and μ_n is the Haar measure on sphere. The exact definition of the Mahler measure will not be important to us. We will just require the following bound relating the Mahler measure to absolute values of the coefficients of f:

$$\log(|f|_1) \leq m(f; S_n^r) + c \cdot r \cdot d \cdot \log(nr+1), \tag{4.1}$$

[2] I must admit that the main reason for using this machinery here is that it gives me an opportunity to learn it.

where d is the degree of f in each set of variables, c is a universal constant, and $|f|_1$ is the sum of the absolute values of the coefficients of f.

We can now state the definition of the height of an equidimensional affine variety V of dimension r. Given such a variety, the Chow form of the projective closure $\overline{V}$, denoted f_V, is a polynomial in $(r+1)(n+1)$ variables. We define

$$\mathrm{ht}(V) = \frac{1}{[K:\mathbb{Q}]}\left(\sum_{\nu \in M_K^\infty} N_\nu m\big(\sigma_\nu(f_V); S_{n+1}^{r+1}\big) + \sum_{\nu \notin M_K^\infty} N_v \log|f_V|_\nu\right) + \deg V(r+1)\left(\sum_{i=1}^{n}\frac{1}{2i}\right).$$

Here, $\sigma_\nu(f)$ is the natural embedding of $K[x_1,\ldots,x_n]$ corresponding to the Archimedean absolute value ν. Again the exact form of the above expression will not matter to us. Apart from the final term (which only depends on the degree), the height of V is essentially an invariant height of a Chow form of f_V. One difference is that for Archimedean ν, instead of measuring using the usual absolute value, the Mahler measure is used instead (this has better properties with respect to multiplication). The second main difference is there is no max term for each absolute value (this should be compared to the previous definition of the height of a polynomial). As a consequence, if we scale f_V by an element of K different from 0, then the first term in the summand remains unchanged. We expect a height function on a polynomial to scale when the polynomial is scaled, but here we are measuring the height of a variety V, which is unchanged when f_V is scaled. Therefore this invariant notion of height is preferred.

We now state some properties about height functions from [KPS01] that will be useful to us. The function ht is additive on irreducible components, that is, if $V = V_1 \cup \cdots \cup V_r$ is the irreducible decomposition of V then $\mathrm{ht}(V) = \sum \mathrm{ht}(V_i)$. In this way, and in many other ways, the height function behaves like the degree of a variety. For example, the height function satisfies an arithmetic Bézout inequality [Phi95, Theorem 3], which states that

$$\mathrm{ht}(V \cap W) \leq \mathrm{ht}(V)\deg W + \mathrm{ht}(W)\deg V + 2n^2 \deg V \deg W.$$

We note that the statement above can be made more precise, see for example [KPS01, Page 14]. By repeated applications of this, and keeping track of the irreducible components, the following is obtained in [KPS01, Corollary 2.10].

Lemma 4.1.14 ([KPS01, Corollary 2.10])**.** *If $f_1,\ldots,f_s$ are polynomials of degrees $d_1,\ldots,d_s$ with $d_1 \geq \cdots \geq d_s$ and $h := \mathrm{ht}(f_1,\ldots,f_s)$ then $\mathrm{ht}(V(f_1,\ldots,f_s)) \leq n d_1^n (h + 2\log n)$.*

A second important property we will need is the behaviour of heights of varieties under regular maps.

Lemma 4.1.15 ([KPS01, Proposition 2.4])**.** *Let $V \subset \overline{\mathbb{A}}^n$ be a variety of dimension r and let $\varphi : \mathbb{A}^n \to \mathbb{A}^N$ be an affine map. Then*

$$\mathrm{ht}(\phi(V)) \leq \mathrm{ht}(V) + (r+1)(\mathrm{ht}(\varphi) + 8\log(n+N+1))\deg V.$$

Using the above machinery, we can for example show the following result about defining equations for irreducible components of varieties. We start with a lemma that will allow us to go from height bounds on hypersurfaces back to height bounds on defining equations.

Lemma 4.1.16. *Suppose $V \subset \mathbb{A}^n$ is a hypersurface that is defined over K with $\deg V = d$ and $\mathrm{ht}(V) = h$. Then there exists a polynomial $f \in O_K[x_1,\ldots,x_n]$ with*

$$\mathrm{ht}(f) \leq [K:\mathbb{Q}]\big(\mathrm{ht}(V) + c \cdot n^2 \log n \cdot d \log d\big)$$

such that $V = V(f)$, where c is a universal constant.

Proof. We start by assuming $(0,\dots,0) \notin V$. Since V is a hypersurface, a chow form of $\overline{V}$ is a polynomial in n groups of $n+1$ variables each, and is homogeneous of degree d in each group. Each group of variables corresponds to a linear form $L_i = \sum_{j=0}^{n} U_{i,j}x_j$, for $i = 1,\dots,n$. Here we assume that x_0 is the new variable with respect to which we homogenise. Let f_V be a fixed Chow form, it is defined up to multiples.

Let g be the specialisation of f_V obtained by the following specialisation: set $U_{i,i} = 1$ for all i, and $U_{i,j} = 0$ for all $i \neq j$ with $j > 0$. This is a polynomial in the variables $U_{0,1},\dots,U_{0,n}$. For each choice of $\alpha_1,\dots,\alpha_n$, we have $g(\alpha_1,\dots,\alpha_n) = 0$ if and only if V intersects $x_1+\alpha_1,\dots,x_n+\alpha_n$. This implies that $V = V(g(-x_1,\dots,-x_n))$.

Now consider the specialisation $\widehat{g}$ of f_V to $U_{i,j} = 0$ for all $i \neq j$. Note that $\widehat{g}$ must be a monomial of f_V, because it has to be homogeneous in each of its variables. We have $\widehat{g}(1,\dots,1) = g(0,\dots,0) \neq 0$, since $(0,\dots,0) \notin V$. We normalise f_V so that this monomial has coefficient 1. [3]

Now we have by definition

$$\mathrm{ht}(V) = \frac{1}{[K:\mathbb{Q}]}\left(\sum_{\nu\in M_K^\infty} N_\nu m\big(\sigma_\nu(f_V); S_{n+1}^{r+1}\big) + \sum_{\nu\notin M_K^\infty} N_v \log|f_V|_\nu\right) + d\cdot(r+1)\left(\sum_{i=1}^{n}\frac{1}{2i}\right).$$

Rearranging, and using the fact that $\mathrm{ht}(V) = h$ we obtain

$$\sum_{\nu\in M_K^\infty} N_\nu m\big(\sigma_\nu(f_V); S_{n+1}^{r+1}\big) + \sum_{\nu\notin M_K^\infty} N_v \log|f_V|_\nu = [K:\mathbb{Q}]\left(\mathrm{ht}(V) - d\cdot(r+1)\left(\sum_{i=1}^{n}\frac{1}{2i}\right)\right).$$

For each σ_ν we have $m(\sigma_\nu(f_V)) \leq m\big(\sigma_\nu(f_V); S_{n+1}^{r+1}\big) + d\cdot(r+1)\sum_{i=1}^{n}(2i)^{-1}$, by [KPS01, Equation 1.2]. Here, m denotes the usual Mahler measure. Using the fact that $\sum_{\nu\in M_K^\infty} N_\nu = [K:\mathbb{Q}]$ we can deduce

$$\frac{1}{[K:\mathbb{Q}]}\left(\sum_{\nu\in M_K^\infty} N_\nu m(\sigma_\nu(f_V)) + \sum_{\nu\notin M_K^\infty} N_v \log|f_V|_\nu\right) \leq \mathrm{ht}(V).$$

The expression on the left is exactly the invariant height of the polynomial f_V as defined in [Phi86, Definition 1.11]. Following the proof of [Phi86, Item 7, Proposition 1.12] our choice of normalisation allows us to deduce that the invariant height of f_V agrees with the height of f_V in the sense of [Phi86, Definition 1.11]. Indeed the proof shows that as long as the least monomial with respect to a certain lex ordering of the variables has coefficient 1, the invariant height and height agree. This is almost the same as the height $\mathrm{ht}(f_V)$ that we defined earlier, the only difference being the use of the Mahler measure instead of the usual absolute value for the Archimedean places. Correcting for this using Eq. (4.1) we obtain $\mathrm{ht}(f_V) \leq \mathrm{ht}(V) + c_1 \cdot n\log n \cdot d$, for an absolute constant c_1.

The process of obtaining a defining equation for V from f_V has been outlined above, we now calculate the height bounds. We start by specialising $U_{i,j} = 0$ for $i \neq j$ and $j > 0$, this does not increase the height since we just shrink the set of coefficients of f_V. Next, we specialise $U_{i,i} = 1$ for all i. The coefficients of this specialisation are just sums of coefficients of f_V. Therefore we end up with a polynomial of height at most $\mathrm{ht}(V) + c_2 \cdot n^2\log n \cdot d\log d$ for a constant c_2. Finally we replace $U_{i,0}$ with $-x_i$, this does not change the height. Therefore we can conclude that in the given situation, we have $V = V(f')$ for a polynomial $f' \in K[x_1,\dots,x_n]$ with $\mathrm{ht}(f') \leq \mathrm{ht}(V) + c_2 \cdot n^2\log n \cdot d\log d$. To complete the proof, we have to find a polynomial $f \in O_K[x_1,\dots,x_n]$ that is of the form

[3] In the notation of [KPS01], this normalised Chow form is denoted Ch_V. Since we just need this here in one proof, we do not give it a name.

$\alpha f'$ for some $\alpha \in K \setminus \{0\}$. This will ensure $V(f) = V(f')$. To do this, we invoke [KPS01, Lemma 1.3]. This result guarantees the existence of a $b \in \mathbb{Z}$ with the above property, that $bf' \in O_K[x_1, \dots, x_n]$ and that satisfies $\operatorname{ht}(bf') \leq [K : \mathbb{Q}] \operatorname{ht}(f')$. Setting $f = bf'$ we obtain $V = V(f)$ and $\operatorname{ht}(f) \leq [K : \mathbb{Q}]\big(\operatorname{ht}(V) + c_3 \cdot n^2 \log n \cdot d \log d\big)$ as claimed.

To finish the proof, we have to remove the assumption that $(0, \dots, 0) \notin V$. The idea is just to translate V to ensure this condition. By the degree mantra, there exists a point in $[d]^n$ that does not lie on V. Let φ be the translation map that sends this point to the origin, by Lemma 4.1.15 we have $\operatorname{ht}(\varphi(V)) \leq \operatorname{ht}(V) + c_4 \cdot n \log n \cdot d \log d$ for a universal constant c_4. By construction $\varphi(V)$ is definable over K, has the same degree as V and does not contain $(0, \dots, 0)$. By the above argument, we have a defining equation for this translate with height bounded by $[K : \mathbb{Q}]\big(\operatorname{ht}(V) + c_5 \cdot n^2 \log n \cdot d \log d\big)$. The equation f for V is obtained by shifting this equation for $\varphi(V)$, and by [KPS01, Lemma 1.2 (c)] we obtain that $\operatorname{ht}(f) \leq [K : \mathbb{Q}]\big(\operatorname{ht}(V) + c \cdot n^2 \log n \cdot d \log d\big)$, finishing the proof. □

As a consequence we can deduce the following about the irreducible components of V.

Lemma 4.1.17. *Suppose $K = \mathbb{Q}[\xi]$. Suppose $V \subset \mathbb{A}^n$ is a hypersurface that is defined over K with $\deg V = d$ and $\operatorname{ht}(V) = h$. If $V_1, \dots, V_r$ are the irreducible components of V, then each V is defined over an extension K_i of K with $[K_i : K] \leq d$. Further $K_i = \mathbb{Q}[\mu_i]$ for an algebraic integer of height at most $h \cdot (d \cdot e \cdot n)^c$, where c is a universal constant and $e = [K : \mathbb{Q}]$.*

Proof. We have $\operatorname{ht}(V) = \sum \operatorname{ht}(V_i)$ therefore $\operatorname{ht}(V_i) \leq \operatorname{ht}(V)$ for each i. Let $f \in O_K[x_1, \dots, x_n]$ be the polynomial defining V whose existence is guaranteed by Lemma 4.1.16. Now each V_i is defined by a factor of f as a polynomial in $\overline{\mathbb{Q}}[x_1, \dots, x_n]$. Our goal is to control the degree of extensions that contain these factors.

We now recall the factoring algorithm for factoring over the algebraic closure from [Kal95]. From the algorithm we can deduce the following: if $g \in K[x_1, \dots, x_n]$ is a polynomial that is monic in x_1 and such that $g(x_1, 0, \dots, 0)$ is squarefree, then each absolutely irreducible factor of g lies in an extension of the type $K[\eta_i]$, where η_i is a root of $g(x_1, 0, \dots, 0)$.

Returning to our setting, we can ensure that f in monic in x_1 and that $f_1(x_1, 0, \dots, 0)$ is squarefree by a random linear transformation using integer coefficients of height $c_1 \log d$ for some constant c_1. Call this new polynomial g, we have $\operatorname{ht}(g) \leq \operatorname{ht}(f) + c_2 \cdot n \log n \cdot d \log d$ for a constant c_2. Further, the degree of g is exactly d. We also have $\operatorname{ht}(g(x_1, 0, \dots, 0)) \leq \operatorname{ht}(g)$. Let $G := g(x_1, 0, \dots, 0)$. If η_i is a root of G and if $K_i := K[\eta_i]$ then $[K_i : K] \leq d$. Combined with the fact that the factors of f have coefficients in K_i, this shows the first of the required claims. Note that η_i might not be an algebraic integer, we will eventually find an algebraic integer that generates the same field.

We now control the heights of the extension. Let L be a extension field of K that contains all the conjugates of ξ. Let $\xi_1, \dots, \xi_e$ be the conjugates of ξ with $\xi_1 = \xi$ and let $\sigma_1, \dots, \sigma_e$ be automorphisms of L that map ξ to $\xi_1, \dots, \xi_e$ respectively. Further let $G_i := \sigma_i(G)$. We have $\operatorname{ht}(G_i) = \operatorname{ht}(G)$ for all i. The polynomial $\widehat{G} := \prod_{i=1}^{e} G_i$ satisfies $\deg \widehat{G} = de$ and $\operatorname{ht}(\widehat{G}) \leq e \operatorname{ht}(G) + ed \log 2$ by [KPS01, Lemma 1.2]. By construction, $\widehat{G} \in \mathbb{Z}[x_1]$.

We can factor out the gcd of the coefficients of $\widehat{G}$ without increasing its height or degree. This polynomial is not monic, since G does not necessarily have leading term 1. If α is the leading coefficient of $\widehat{G}$ then $\operatorname{ht}(\alpha) \leq \operatorname{ht}(\widehat{G})$. We can multiply by α^{de-1} and replace the variable x_1 with x_1/α to get a monic polynomial. An extension that contains the root of this new $\widehat{G}$ also contains a root of the original, and this process increases the height by at most $(de - 1) \operatorname{ht}(\widehat{G})$. We know that η_i is a root of $\widehat{G}$, therefore the minimal polynomial of η_i is a factor of $\widehat{G}$. By Gelfonds inequality [HS13, B.7.3], we can deduce that every factor of $\widehat{G}$ has height at most $\operatorname{ht}(\widehat{G}) + ed$.

Pulling all the bounds above together, we can deduce that there exists an algebraic integer μ_i with a minimal polynomial $\widehat{G}_i$ satisfying $\operatorname{ht}(\widehat{G_i}) \leq e^2 d \operatorname{ht}(f) + c_3 e \cdot n \log n \cdot d^2 \log d$ such that

$\mathbb{Q}[\eta_i] = \mathbb{Q}[\mu_i]$. By [Wal00, Lemma 3.11] we can deduce that $\mathrm{ht}(\mu_i) \leq e^2 d\,\mathrm{ht}(f) + c_3 e \cdot n \log n \cdot d^2 \log d$. To finish, we wish to find a primitive element for the field extension K_i. By the primitive element theorem, an element of the form $\mu_i + \beta\xi$ will be a primitive element for all except finitely many β. From the proof of the theorem we can deduce that the set of bad β has size at most the product of the degrees of ξ, μ_i which is $e^2 d$ in our setting. Therefore we can pick an integer $\beta_i \in [e^2 d]$ with this property, and we can deduce that $\mathrm{ht}(\mu_i + \beta_i\xi) \leq e^2 d\,\mathrm{ht}(f) + c \cdot e \cdot n \log n \cdot d^2 \log d$. Since $\mathbb{Q}[\mu_i + \beta_i\xi] = K[\mu_i] = K[\eta_i] = K_i$, we are done. □

We made no efforts to optimise the above bound. In particular, keeping track of the degrees more accurately might allow us to reduce the polynomial factors. The step of constructing $\widehat{G}$ is also lossy, but we are not sure how to avoid this.

Finally, using projections, we can deduce the following about irreducible components of equidimensional varieties.

Lemma 4.1.18. *Suppose $K = \mathbb{Q}[\xi]$. Suppose $V \subset \mathbb{A}^n$ is an equidimensional variety that is defined over K with $\dim V = r$ and $\deg V = d$ and $\mathrm{ht}(V) = h$. If $V_1, \ldots, V_s$ are the irreducible components of V, then each V is defined over an extension K_i of K with $[K_i : K] \leq d(r+1)$. Further $K_i = \mathbb{Q}[\mu_i]$ for an algebraic integer of height at most $h \cdot (d \cdot e \cdot n)^c$, where c is a universal constant and $e = [K : \mathbb{Q}]$.*

Proof. Let V_1 be a fixed component of V. The idea is the following: we will apply a random linear projection $\varphi : \mathbb{A}^n \to \mathbb{A}^{r+1}$. The image $\varphi(V)$ will be a hypersurface, whose irreducible components are in bijection with those of V. At this point, we can use Lemma 4.1.17 to deduce a defining equation $f_{1,1}$ for $\varphi(V_1)$. If we take the image of this equation under the map φ^*, we get an equation $g_{1,1}$ with the following properties: $V(g_{1,1}) = V_1$ and $\dim(V_i \cap V(g_{1,1})) < r$ for $i > 1$. This is not quite enough because $V(g_{1,1})$ still intersects the other components. To fix this, we will just apply the above argument $r+1$ times, to obtain $g_{1,1}, \ldots, g_{1,r+1}$. We will then have $V_i \cap V(g_{1,1}, \ldots, g_{1,r+1}) = \emptyset$. The field extension K_i will be the compositum of the extensions containing $g_{1,j}$ and we can again use the primitive element theorem to deduce a bound on the height of μ_i. Tracking the height bounds is a routine calculation extremely similar to the ones above, therefore we do not repeat it here. □

4.1.4 Properties preserved by base change

In this section, we collect some properties about zero sets that remain unchanged when we go modulo a prime. The set up is as follows. We have a finite extension $K := \mathbb{Q}(\xi)$ if the rational numbers, and $q \in \mathbb{Z}[z]$ is a monic minimal polynomial for ξ. We have polynomials $f_1, \ldots, f_m \in O_K[x_1, \ldots, x_n]$. The heights of $\xi, f_1, \ldots, f_m$ are all bounded by a number h. The degrees of f_i are bounded by d, and the degree of q is e. We denote by I the ideal $(f_1, \ldots, f_m)$ in $\overline{\mathbb{Q}}[x_1, \ldots, x_n]$, and by V the zero set of I.

For any prime $p \in \mathbb{N}$ that does not divide the discriminant of q, and for any prime $\mathfrak{p}$ of O_K lying above p, we denote by $f_{i,\mathfrak{p}}$ the image of f_i under the map $O_K[x_1, \ldots, x_n] \to (O_K/\mathfrak{p})[x_1, \ldots, x_n]$. We denote by $I_\mathfrak{p}$ the ideal generated by $f_{1,\mathfrak{p}}, \ldots, f_{m,\mathfrak{p}}$, and by $V_\mathfrak{p}$ the zero set of $I_\mathfrak{p}$ in $\overline{\mathbb{F}}_p[x_1, \ldots, x_n]$.

Let us start with an easy property, to demonstrate how the arguments will proceed. We show that if V is non-empty, then $V_\mathfrak{p}$ is non-empty for all except constantly many primes. We will also give an upper bound on this number of primes. Before we begin, we remark that we cannot expect $V_\mathfrak{p}$ to be non-empty for every prime. Consider the univariate example with $f_1 = 12x - 1$. This has a solution, but $f_{1,(2)}$ does not have a solution, neither does $f_{1,(3)}$. Recall the effective Nullstellensatz Theorem 2.3.8. It states that if $f_1, \ldots, f_m$ is unsatisfiable, then there are $g_1, \ldots, g_m$ with $\deg f_i g_i \leq d^n$ such that $1 = \sum f_i g_i$. In fact this is a necessary and sufficient condition for the polynomials to be unsatisfiable. The contrapositive is that as long as there are no polynomials

$g_1, \dots, g_m$ with $\deg f_i g_i \leq d^n$ such that $1 = \sum f_i g_i$, the system must be satisfiable. The converse also holds.

In our setting, we know that $f_1, \dots, f_m$ is satisfiable, so there is no $g_1, \dots, g_m$ with $\deg f_i g_i \leq d^n$ such that $1 = \sum f_i g_i$. The existence of such a $g_1, \dots, g_m$ is a linear system in the coefficients of $f_1, \dots, f_m$. The unknowns in this system are the coefficients of each monomial of each g_i. The number of such coefficients is bounded by $m \cdot \binom{d^n+n+1}{n+1}$, which we bound by $m \cdot d^{c_1 n^2}$. Each term $f_i g_i$ has degree d^n. Therefore, these is one equation for each monomial of degree d^n. For all non-constant monomials, these equations say that the coefficient of the monomial is zero. The equation corresponding to the constant term says that the monomial is 1. Therefore the total number of equations in this system are $\binom{d^n+n}{n}$, which we bound again by $d^{c_1 n^2}$.

Call the matrix of this system M, it has one column for each coefficient of each g_i, and it has one row for each monomial of degree at most d^n. When talking about this linear system, we will overload notation and use the same symbol to denote polynomials and their coefficient vectors. Therefore $[M|1]$ denotes the augmented matrix where the vector corresponding to the constant polynomial 1 is added as a column. That the linear system has no solutions means that $\operatorname{rank}[M|1] = \operatorname{rank} M + 1$. Let M' be a square submatrix of $[M|1]$ such that $\operatorname{rank} M' = \operatorname{rank}[M|1]$, this submatrix M' has to include the last column and row of $[M|1]$ (let us assume that the row corresponding to the constant term is the last row). The number of rows and columns of M' is bounded by $d^{c_1 n^2}$, and each entry of M' has height at most h. We have $\det M' \in O_K$, and further by Lemma 4.1.8, we have $\operatorname{ht}(\det M') \leq h \cdot d^{c_2 n^2}$. Note that our bounds are fairly crude, keeping track of exact bounds will be extremely tedious and will not change any of our results. By Lemma 4.1.7, there are at most $e \cdot \operatorname{ht}(\det M')$ many primes $\mathfrak{p} \subset O_K$ such that $\det M' \in \mathfrak{p}$.

Let $\mathfrak{p} \subset O_K$ be a prime that $p := \mathfrak{p} \cap \mathbb{Z}$ is not ramified in O_K and such that $\det M' \notin \mathfrak{p}$. For any such $\mathfrak{p}$ we claim that $V_{\mathfrak{p}}$ is nonempty. Let $\mathbb{F}_q := O_K/\mathfrak{p}$, the fact that $\mathfrak{p} \cap \mathbb{Z}$ is unramified ensures that this quotient is a finite extension of $\mathbb{F}_p$. By the effective Nullstellensatz, to show that $\mathfrak{p}$ is nonempty, it suffices to show that there do not exist polynomials $g_1, \dots, g_m \in \mathbb{F}_q[x_1, \dots, x_n]$ with $\deg f_{i,\mathfrak{p}} g_i \leq d^n$ with $1 = \sum f_{i,\mathfrak{p}} g_i$. As before, we can write this as a linear system in the coefficients of $f_{1,\mathfrak{p}}, \dots, f_{m,\mathfrak{p}}$. The key observation is that the matrix corresponding to this system, call it $M_{\mathfrak{p}}$, is exactly the matrix obtained by starting with the matrix M above and reducing every entry mod $\mathfrak{p}$. As a consequence, we can immediately deduce that $\operatorname{rank} M_{\mathfrak{p}} \leq \operatorname{rank} M$. The same is true for $[M_{\mathfrak{p}}|1]$ and $[M|1]$. Now we consider the submatrix of $[M_{\mathfrak{p}}|1]$ with the same rows and columns as M', call this submatrix $M'_{\mathfrak{p}}$. That $\det M' \notin \mathfrak{p}$ implies that $\det M'_{\mathfrak{p}} \neq 0$, and therefore that $\operatorname{rank}[M_{\mathfrak{p}}|1] = \operatorname{rank}[M|1]$. Combined with the inequality that $\operatorname{rank} M < \operatorname{rank}[M|1]$, we deduce that $\operatorname{rank} M_{\mathfrak{p}} < \operatorname{rank}[M_{\mathfrak{p}}|1]$. This shows that the linear system is unsatisfiable, therefore that $V_{\mathfrak{p}}$ is nonempty.

Finally we can count the number of primes $\mathfrak{p}$ that do not satisfy the above. By Lemma 4.1.10 there are at most $c_3 \cdot e^2 \log e \cdot h$ primes $p \in \mathbb{Z}$ that are ramified, and each prime has at most e primes $\mathfrak{p}$ above it. Further there are at most $e \cdot h \cdot d^{c_2 n^2}$ primes $\mathfrak{p}$ with $\det M' \in \mathfrak{p}$. Combining these, we deduce that there are at most $h \cdot e^3 \log e \cdot d^{cn^2}$ primes $\mathfrak{p}$ such that $V_{\mathfrak{p}}$ is empty, for some absolute constant c. We capture this is a lemma so we can refer to it.

Lemma 4.1.19. *There is a universal constant c that satisfies the following: if V is satisfiable, then there are at most $h \cdot e^3 \log e \cdot d^{cn^2}$ primes $\mathfrak{p}$ such that either $\mathfrak{p} \cap \mathbb{N}$ is ramified or such that $V_{\mathfrak{p}}$ is empty.*

The key step in the above argument is showing that the property of V being empty is implied by some element staying nonzero under the quotient. We will do essentially the same for every property. In fact even the fact that an absolutely irreducible polynomial remains absolutely irreducible when going modulo $\mathfrak{p}$ for almost all $\mathfrak{p}$ was proved this way.

The next property we will show remains unchanged is the dimension of V. In particular, we show the following.

Lemma 4.1.20. *There is a universal constant c that satisfies the following: there are at most $h \cdot e^3 \log e \cdot d^{c_2 n^3}$ primes $\mathfrak{p}$ such that either $\mathfrak{p} \cap \mathbb{Z}$ is unramified or such that $\dim V_{\mathfrak{p}} \neq \dim V$.*

As before we will express the fact that $\dim V = r$ as the non vanishing of some algebraic integer. We can do this in a number of ways. One way is to look at all projections onto sets of r and $r+1$ coordinates. Since V is dimension r, the dimension of any such projection will be at most r. Therefore the ideal I will contain a polynomial in every set of $r+1$ variables, and we can control its degree. Further there will be a set of r variables such that the projection is dense, therefore there is a set of r variables such that I does not contain any polynomial in these r variables. This property will also be preserved by reducing modulo $\mathfrak{p}$.

Here we will do something different, and introduce a method that will be useful later on too. The main idea is that if V has dimension r then the intersection of V with a random linear subspace H_r of dimension $n-r$ is nonempty, and with a random linear subspace H_{r+1} of dimension $n-r-1$ is empty. If this happens for a fixed H_r we say that H_r intersects V properly, similarly for H_{r+1}. For a fixed such linear subspace H_r, we can count the number of primes $\mathfrak{p}$ such that the image of $H_{r,\mathfrak{p}} \cap V_{\mathfrak{p}}$ is nonempty. However, showing that $H_{r,\mathfrak{p}} \cap V_{\mathfrak{p}} \neq \emptyset$ does say anything about the dimension of $V_{\mathfrak{p}}$. Indeed it might just be that $H_{r,\mathfrak{p}}$ is a linear subspace that intersects $V_{\mathfrak{p}}$ improperly.

One way of fixing this situation is to quantify which subspaces intersect V improperly. The set of such subspaces will be a closed subspace of the appropriate Grassmanian. We can write down the equation of this closed subspace, and show that the corresponding equation for $V_{\mathfrak{p}}$ is obtained by reducing the equation for V modulo $\mathfrak{p}$.

Here we will do something slightly more crude. We will just take a large number of subspaces H_r^i, and study those primes $\mathfrak{p}$ for which $H_{r,\mathfrak{p}}^i \cap V_{\mathfrak{p}}$ is non empty for many i. While we cannot deduce anything about the dimension of $V_{\mathfrak{p}}$ from a single such intersection, we will be able to do so from enough intersections. In particular, if we take a large number of subspaces, at least one of them will have to intersect $V_{\mathfrak{p}}$ properly.

Proof of Lemma 4.1.20. Let $N := (n+1) \cdot d^n + 1$. Let $\mathcal{H}'$ be a collection of linear subspaces picked as follows. We pick every possible set of r linear polynomials with coefficients in the set $\{0, \dots, N-1\}$, and add the linear subspace defined by this set of equations to $\mathcal{H}$, as long as the linear subspace has dimension $n-r$. Different sets of linear forms can give rise to the same linear subspace, therefore elements of $\mathcal{H}'$ can repeat. Note that $|\mathcal{H}'| \geq (1 - r/N) N^{r(n+1)}$ by the degree mantra. Let $\mathcal{H} \subset \mathcal{H}'$ be those subspaces $H \in \mathcal{H}'$ such that $\dim V \cap H \neq \emptyset$. By Lemma 2.3.4 we can deduce that $|\mathcal{H}| \geq (1 - r \cdot d^n/N)|\mathcal{H}'|$ therefore $|\mathcal{H}| \geq (1 - r \cdot d^n/N)(1 - r/N) N^{r(n+1)}$.

Let $H \in \mathcal{H}$ be a fixed linear subspace. Suppose $H = V(\ell_1, \dots, \ell_r)$. We have $\mathrm{ht}(\ell_1, \dots, \ell_r) \leq \log N$. By Lemma 4.1.19, there are at most $(h + \log N) \cdot e^3 \log e \cdot d^{c_1 n^2}$ primes $\mathfrak{p}$ such that $V_{\mathfrak{p}} \cap H_{\mathfrak{p}} = \emptyset$. Here, $H_{\mathfrak{p}}$ is the zero set of the images of $\ell_1, \dots, \ell_r$ under the quotient map $O_K \to O_K/\mathfrak{p}$. By taking a union over all $H \in \mathcal{H}$, we can deduce that there are at most $|\mathcal{H}|(h + \log N) \cdot e^3 \log e \cdot d^{c_1 n^2}$ primes $\mathfrak{p}$ such that $V_{\mathfrak{p}} \cap H_{\mathfrak{p}} = \emptyset$ for some $H \in \mathcal{H}$. In other words, for every other prime $\mathfrak{p}$ we have $V_{\mathfrak{p}} \cap H_{\mathfrak{p}} \neq \emptyset$ for all $H \in \mathcal{H}$.

Now consider any such prime $\mathfrak{p}$, that has the additional property that $\mathfrak{p}$ lies above a prime p with $p > N$. We will show that $\dim V_{\mathfrak{p}} \geq r$. Suppose towards a contradiction that $\dim V_{\mathfrak{p}} < r$. Similar to our analysis above, using the degree mantra and Lemma 2.3.4 we can deduce that the number of linear subspaces H with $V_{\mathfrak{p}} \cap H_{\mathfrak{p}} \neq \emptyset$ is bounded by $(d^n/N)(1 - r/N) N^{r(n+1)}$. Since $N > (r+1) \cdot d^n$, this contradicts the choice of $\mathfrak{p}$. Combining the bounds above and plugging in the value of N, we deduce that there are at most $h \cdot e^3 \log e \cdot d^{c_2 n^3}$ primes $\mathfrak{p}$ such that $\dim V_{\mathfrak{p}} < r$.

To control the primes $\mathfrak{p}$ for which $\dim \mathfrak{p} > r$, we do a very similar argument. We will pick $\mathcal{H}'$ to be subspaces of dimension $n-r-1$ and define $\mathcal{H}$ to be those subspaces such that $V \cap H = \emptyset$. Using

Theorem 4.1.13, in place of Lemma 4.1.19 we can deduce that for all except constantly many primes, we have $V_{\mathfrak{p}} \cap H_{\mathfrak{p}} = \emptyset$ for all $H \in \mathcal{H}$ (the number of primes will be bounded by a similar function to the one above). For any such prime, again using the same argument we can deduce that $\dim V_{\mathfrak{p}} \leq r$, and complete the proof. Since the arguments are almost exactly the same, we do not repeat them again. □

The next property we show is preserved by the above operation is the irreducibility of V.

Lemma 4.1.21. *There is a universal constant c that satisfies the following: If V is irreducible then there are at most $h \cdot \left(d^{n^c} \cdot n \cdot e\right)^c$ primes $\mathfrak{p}$ such that either $\mathfrak{p} \cap \mathbb{Z}$ is unramified or such that V is reducible.*

Proof. First suppose V is a hypersurface. We can deduce by Lemma 4.1.16 that V is defined by a polynomial $f \in O_K[x_1, \ldots, x_n]$ that satisfies $\mathrm{ht}(f) \leq e\big(\mathrm{ht}(V) + c_1 \cdot n^2 \log n \cdot d \log d\big)$. Note that f will be a factor of the polynomials f_i, therefore this bound is far from optimal but will give us the same result asymptotically. Let b_f be the element whose existence is given by Lemma 4.1.12. For any $\mathfrak{p}$ such that $b_f \notin \mathfrak{p}$, as long as $\mathfrak{p}$ is above an unramified prime, the image of f remains absolutely irreducible. Finally, consider the system $f_1, \ldots, f_m, 1 - yf$ where y is a new variable. This system is unsatisfiable, and we have $h(f_1, \ldots, f_m, 1 - yf) \leq \mathrm{ht}(1 - yf) + \sum \mathrm{ht}(f_i)$. By Theorem 4.1.13, there is an element $a \in O_K$ and $g_1, \ldots, g_{m+1}$ satisfying $a = g_{m+1}(1 - yf) + \sum f_i g_i$ with $\mathrm{ht}(a) \leq d^{c_2 n} \cdot \mathrm{ht}(f) \cdot n^3 \log n \cdot d \log d$.

Now consider any prime $\mathfrak{p}$ that lies above an unramified prime, and such that $a, b_f \notin \mathfrak{p}$. Assume further that $\dim V_{\mathfrak{p}} = \dim V$. The image of f in $O_K/\mathfrak{p}$ remains absolutely irreducible. Since $a \notin \mathfrak{p}$, the system $f_{1,\mathfrak{p}}, \ldots, f_{m,\mathfrak{p}}, 1 - yf_{\mathfrak{p}}$ remains unsatisfiable, whence $f_{\mathfrak{p}} \in \mathrm{rad}(I_{\mathfrak{p}})$. These facts imply that $V_{\mathfrak{p}}$ is irreducible. We can use Lemma 4.1.14 to bound height of $\mathrm{ht}(V)$. Putting all of this together rather crudely, we get that the above holds for all except $h \cdot \left(d^{n^{c_3}} \cdot n \cdot e\right)^{c_3}$ many primes $\mathfrak{p}$.

Now suppose V is an arbitrary irreducible variety of dimension r. Since V is irreducible, for a random linear map $\varphi : \mathbb{A}^n \to \mathbb{A}^{r+1}$, the image will be irreducible. We have bounds on $\mathrm{ht}(\varphi(V))$ just as above. As above, we can get a height bound on a defining equation for $\varphi(V)$, and a height bound on the image of this equation under φ^*, call this equation f. For all except a bounded number of primes, we have $f_{\mathfrak{p}} \in \mathrm{rad}(I_{\mathfrak{p}})$ and $\dim V_{\mathfrak{p}} = \dim V$. For any such prime, the image of $V_{\mathfrak{p}}$ under the random linear map $\varphi_{\mathfrak{p}}$ is irreducible. As in the proof of Lemma 4.1.20, the way we deal with this randomness is to take a large set of such linear maps, and pick those primes for which the above holds for all the maps. For any such prime, the image $V_{\mathfrak{p}}$ will be irreducible. Since the argument is essentially the exact same, we do not repeat it here. Putting together the height bounds, one can easily show the above holds for all except $h \cdot \left(d^{n^c} \cdot n \cdot e\right)^c$ many primes $\mathfrak{p}$ where c is some universal constant. □

Using the results above and the results from the previous section, we can deduce the following results about varieties defined over integers. Similar statements will hold for varieties over other rings.

Lemma 4.1.22. *Suppose $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$ are polynomials of degree at most d and height at most h. Suppose V is equidimensional and reducible. If $\pi_V(a)$ is the number of primes $p \leq a$ such that V_p has at least two components that are $\mathbb{F}_p$ definable, then we have*

$$\pi_V(a) \geq \frac{1}{d^{2n}} \left(\frac{a}{\log a} - h \cdot a^{1/2} \log a \cdot (d^{cn} \cdot n)^c \right).$$

Proof. Let $D := d^n$ be a bound on the degree of V, which is just the cardinality of V. By Lemma 4.1.18, each component V_i of V is defined over an extension of $\mathbb{Q}$ of degree at most D

generated by an element μ_i of height at most $h \cdot \left(d^{n^{c_1}} \cdot n\right)^{c_1}$. Suppose we fix μ_1, μ_2. The extension $\mathbb{Q}[\mu_1, \mu_2]$ has degree at most D^2, and by the primitive element theorem, is generated by an element μ_{12} of the form $\mu_{12} = \mu_1 + \alpha\mu_2$. The number α can be picked to be an integer less than D^2 in absolute value, therefore $\mathrm{ht}(\mu_{12}) \leq \mathrm{ht}(\mu_1) + \mathrm{ht}(\mu_2) + c_2 \log D$.

At least two components of V are defined over the extension $K := \mathbb{Q}[\mu_{12}]$. By Lemma 4.1.20 and Lemma 4.1.21, there are at most $h \cdot \left(d^{n^{c_3}} \cdot n\right)^{c_3}$ primes p, such that for some prime $\mathfrak{p}$ above p, the varieties $V_{1,\mathfrak{p}}$ and $V_{2,\mathfrak{p}}$ either have dimension different from $\dim V$, or are reducible. If q is the minimal polynomial of μ_{12}, then for any prime p such that q has an $\mathbb{F}_p$-root and such that p is not one of the primes above, we have that V_1, V_2 are $\mathbb{F}_p$ definable. By Corollary 4.1.11 and the above bounds, we can deduce the required conclusion. □

4.1.5 An interactive proof for reducibility of ideal theoretically defined varieties

Putting all the above ingredients together, we will show our main theorem. We restate the theorem for convenience.

Theorem 4.1.23. *Let C be an algebraic circuit of size s with integer constants that computes $f_1, \ldots, f_m \in \mathbb{Z}[x_1, \ldots, x_n]$. Suppose $I := (f_1, \ldots, f_m)$ is promised to be radical, and suppose $r = \dim I$ is given. There is a* coAM *protocol that decides the irreducibility of $V(I)$.*

Proof. Arthur and Merlin will run two protocols. The first protocol is to certify that V contains a component of dimension less than r. If Merlin is unable to certify this, then they run a protocol to certify that V has two components of dimension r. If V indeed is reducible, then Merlin can correctly convince Arthur in one of the two rounds. If not, then Merlin will be unable to convince Arthur in either round.

The first protocol proceeds as follows. Merlin constructs the Jacobian matrix $\mathcal{J}_{i,j} = \partial_j f_i$, and checks if there is a minor M of $\mathcal{J}$ of size at least $n - r + 1$ such that $M, f_1, \ldots, f_m$ is satisfiable. If such a minor exists, then he sends the rows and columns corresponding to M to Arthur. Arthur can construct a circuit that computes M, and then use the protocol from Theorem 4.1.5 to verify the satisfiability of $M, f_1, \ldots, f_m$. By Lemma 2.3.7, such a minor M exists if and only if V has a component of dimension less than r. Therefore the first protocol behaves as claimed.

If the first protocol fails to certify the reducibility, then it must be that V only has components of dimension r. If V is reducible then there must be at least two such components. By Lemma 4.1.22, the number of primes $p \leq a$ for which V_p has at least two $\mathbb{F}_p$ definable components satisfies

$$\pi_V(a) \geq \frac{1}{d^{2n}} \left(\frac{a}{\log a} - h \cdot a^{1/2} \log a \cdot (d^{c_1 n} \cdot n)^{c_1} \right),$$

where d, h are bounds on the degrees and heights of $f_1, \ldots, f_m$. Using Lemma 2.2.5, we can bound h by 2^{2s}, and similarly we can bound d by 2^s. We then get

$$\pi_V(a) \geq \frac{1}{2^{2s \cdot n}} \left(\frac{a}{\log a} - 2^{2s} \cdot a^{1/2} \log a \cdot (2^{c_1 s \cdot n} \cdot n)^{c_1} \right).$$

On the other hand, if V were irreducible, then the number of such primes would be bounded by $B := 2^{2s} \cdot \left(2^{s \cdot n^{c_2}} \cdot n\right)^{c_2}$ by Lemma 4.1.21. Let $C := 200 \cdot 2^{5s}$.

We can pick an A large enough that $\pi_V(A) - \pi_V(C) \geq 2B$ in the case when V is reducible, this only requires a to have height polynomial in s, n. It suffices for Merlin to certify that $\Pi_V(A) \setminus \Pi_V(C)$ has size at least $2B$ to show that V is reducible. For a prime $p \in \Pi_V(A) \setminus \Pi_V(C)$, by the Lang Weil bound Theorem 4.1.6, we have $\left|V_p \cap \mathbb{F}_p^n\right| \geq 2p^r - 6 \cdot 2^{5s} \cdot p^{r-1/2}$. If on the other hand p was such that V_p is irreducible, then $\left|V_p \cap \mathbb{F}_p^n\right| \leq p^r + 6 \cdot 2^{5s} \cdot p^{r-1/2}$. The ratio between these two sizes, is at least 1.9 since $p > C$. Therefore membership in $\Pi_V(A) \setminus \Pi_V(C)$ can be certified in AM. □

4.2 Low depth circuits for multivariate resultants

In this section we will focus on the homogeneous system satisfiability problem. We recall the statement of the problem. The input is a system of homogeneous forms $F_1, \ldots, F_m \in \mathbb{F}[x_0, \ldots, x_n]$ and the YES instances are those sets of forms that have a common solution other than $(0, \ldots, 0)$ in the algebraic closure $\mathbb{F}$. Our model of computation will be piecewise circuits. These are algebraic circuits that have the additional power of branching based on equality with zero. The formal definition of piecewise circuits, and the properties we use can be found in Section 2.2.1.

We usually think of algebraic circuits as computing formal polynomials, but when using them to design algorithms it is helpful to think of them more like boolean circuits, where the inputs, values on the gates, and outputs are all field elements. An algebraic circuit computes a polynomial function of the input field elements. We say that an algebraic circuit solves a decision problem if the circuit takes value 0 exactly on the YES instances of the problem. Note that the only decision problems that are solvable by algebraic circuits are those where the YES instances are a closed subset in the Zariski sense of the space of inputs. Further, this closed subset has to be defined by a single equation (that which the circuit computes) which means it must be a hypersurface.

More generally, we can consider algebraic circuits with multiple outputs, each of which computes a polynomial function of the inputs. We say a multioutput algebraic circuit solves a decision problem if all outputs take value 0 exactly on the YES instances of the problem. Equivalently, an input is a NO instance if and only if there is at least one output of the circuit that is non zero on this input. Similar to the above, the only decision problems that multioutput algebraic circuits can solve are those where the YES instances are a closed subset of the space of inputs. The difference here is that the closed subset can be defined by more than one equation.

Given that algebraic circuits can only solve decision problems with the above properties, it is not even clear if the decision problem of satisfiability can be solved by algebraic circuits. Let us consider a few motivating examples.

The first example is that of homogeneous linear systems. Consider a collection of homogeneous linear forms $L_0, \ldots, L_m \in \mathbb{F}[x_0, \ldots, x_n]$ given by $L_i(\mathbf{x}) = \sum_{j=0}^{n} \ell_{i,j} x_j$. The system of equations $L_0(\mathbf{x}) = \cdots = L_m(\mathbf{x}) = 0$ has a nonzero solution exactly when the matrix of coefficients of $L_0, \ldots, L_m$ is not full rank. In the case when $m = n$, this equivalent to the condition

$$\det \begin{pmatrix} \ell_{0,0} & \ell_{0,1} & \cdots & \ell_{0,n} \\ \ell_{1,0} & \ell_{1,1} & \cdots & \ell_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n,0} & \ell_{n,1} & \cdots & \ell_{n,n} \end{pmatrix} = 0.$$

As a polynomial in n^2 variables, the determinant has circuits of polynomial size. A direct consequence is that satisfiability of square homogeneous linear systems can be solved by algebraic circuits with a single output as described above.

In the general case, the fact the matrix is not full rank is equivalent to every maximal minor vanishing. The number of such minors is $\binom{m}{n+1}$. We can construct a multioutput algebraic circuit that computes all maximal minors, and this circuit correctly decides solvability of the system, however the size of this circuit is not polynomially bounded by n. This can be remedied using the notion of rank condensers as we will see later, for now we are only interested in the fact that decidability can be solved by circuits.

The second motivating example will be homogeneous bivariates in two variables. Before we discuss this example, we recall the definition of the Sylvester matrix and the resultant of two univariates.

Definition 4.2.1. Let $f, g \in \mathbb{F}[x]$ be polynomials of degrees d and e, respectively. The *Sylvester matrix* of f and g, denoted $\mathrm{Syl}(f,g)$, is the $(d+e) \times (d+e)$ matrix

$$\mathrm{Syl}(f,g) \coloneqq \begin{pmatrix} f_d & & & & g_e & & & & \\ f_{d-1} & f_d & & & g_{e-1} & g_e & & & \\ \vdots & \vdots & \ddots & & \vdots & \vdots & \ddots & & \\ \vdots & \vdots & & f_d & g_1 & \vdots & & \ddots & \\ \vdots & \vdots & & f_{d-1} & g_0 & \vdots & & & \ddots \\ \vdots & \vdots & & \vdots & & g_0 & & & g_e \\ f_0 & \vdots & & \vdots & & & \ddots & & \vdots \\ & f_0 & & \vdots & & & & \ddots & \vdots \\ & & \ddots & \vdots & & & & & \ddots \\ & & & f_0 & & & & & g_0 \end{pmatrix},$$

where the first e columns are formed from the coefficients of f and the remaining d columns are formed from the coefficients of g. The *resultant* of f and g is $\mathrm{Res}(f,g) \coloneqq \det \mathrm{Syl}(f,g)$. ◊

An important property that the above resultant has is that it detects when f, g have nontrivial gcd, and therefore when they have a common root.

Lemma 4.2.2 (see, e.g., [GG13, Corollary 6.17])**.** *Let $f, g \in \mathbb{F}[x]$. Then $\gcd(f,g) \neq 1$ if and only if $\mathrm{Res}(f,g) = 0$. Equivalently, the system of equations $f(x) = g(x) = 0$ has a solution in $\overline{\mathbb{F}}$ if and only if $\mathrm{Res}(f,g) = 0$.*

Now suppose $F, G \in \mathbb{F}[x,y]$ are homogeneous bivariates of degrees d, e respectively, and we are interested in the problem of deciding if they have a common solution. Let $F = \sum_{i \le d} f_i x^i y^{d-i}$ and $G = \sum_{i \le e} g_i x^i y^{e-i}$. We define $\mathrm{Res}(F,G)$ as

$$\mathrm{Res}(F,G) \coloneqq \det \begin{pmatrix} f_d & & & & g_e & & & & \\ f_{d-1} & f_d & & & g_{e-1} & g_e & & & \\ \vdots & \vdots & \ddots & & \vdots & \vdots & \ddots & & \\ \vdots & \vdots & & f_d & g_1 & \vdots & & \ddots & \\ \vdots & \vdots & & f_{d-1} & g_0 & \vdots & & & \ddots \\ \vdots & \vdots & & \vdots & & g_0 & & & g_e \\ f_0 & \vdots & & \vdots & & & \ddots & & \vdots \\ & f_0 & & \vdots & & & & \ddots & \vdots \\ & & \ddots & \vdots & & & & & \ddots \\ & & & f_0 & & & & & g_0 \end{pmatrix}.$$

This is almost the same as the resultant of the univariates $F(x,1)$ and $G(x,1)$. The difference is that if $F(x,1)$ has lower degree than $d' < d$, or if $G(x,1)$ has degree $e' < e$ then $\mathrm{Res}(F(x,1), G(x,1))$ will be the determinant of a submatrix of the above matrix.

The system F, G has a common root if and only if $\mathrm{Res}(F,G) = 0$. This can be checked by case analysis. If the system has a common zero, it either of the form $(\alpha, 0)$ for some $\alpha \in \overline{\mathbb{F}} \setminus \{0\}$ or of the

form $(\alpha, 1)$. In the first case, it must be that $f_d = g_e = 0$ whence $\mathrm{Res}(F, G) = 0$. If $f_d \neq 0$, then the system does not have a root of the first form, and any root of the system is a root of $F(x, 1)$ and $G(x, 1)$, and therefore if such a root exists, then the resultant vanishes by Lemma 4.2.2.

Macaulay generalised the above results, and proved that satisfiability of all homogeneous square systems can be decided by a single polynomial. This polynomial is called the multivariate resultant.

We fix a choice of natural numbers $d_0, \ldots, d_n \in \mathbb{N}$. For $d \in \mathbb{N}$, we write $M_{n,d}$ for the set of homogeneous degree-d monomials in $n+1$ variables. We take $\mathbf{u} = \{u_{i,\boldsymbol{\alpha}} \mid i \in \{0, \ldots, n\}, \boldsymbol{\alpha} \in M_{n,d}\}$ to be a set of variables, corresponding to the coefficients of $n+1$ homogeneous forms of degrees $d_0, \ldots, d_n$, respectively. With this notation in hand, we now define the resultant.

Definition 4.2.3 (see, e.g., [CLO05, Chapter 3, Theorem 2.3])**.** The *resultant* $\mathrm{Res}_{d_0,\ldots,d_n} \in \mathbb{Z}[\mathbf{u}]$ is the unique polynomial satisfying the following conditions.

1. If $F_0, \ldots, F_n \in \mathbb{C}[x_0, \ldots, x_n]$ are homogeneous forms of degrees $d_0, \ldots, d_n$, respectively, then $\mathrm{Res}(F_0, \ldots, F_n) = 0$ if and only if the system of equations $F_0(\mathbf{x}) = \cdots = F_n(\mathbf{x}) = 0$ has a nonzero solution in $\mathbb{C}^{n+1}$ (equivalently, a solution in $\mathbb{P}^n$).
2. $\mathrm{Res}_{d_0,\ldots,d_n}(x_0^{d_0}, \ldots, x_n^{d_n}) = 1$.
3. The polynomial $\mathrm{Res}_{d_0,\ldots,d_n}$ is irreducible in $\mathbb{C}[\mathbf{u}]$.

$\lozenge$

When we say $\mathrm{Res}(F_0, \ldots, F_n)$ we mean the evaluation of the appropriate sized resultant on the coefficients of $F_0, \ldots, F_n$. The existence of the resultant is a non trivial fact, and we do not prove it here. A proof of the existence can be found in a number of places, for example [GKZ94; Mac94; Phi86; Jou91]. The bivariate resultant $\mathrm{Res}(F, G)$ described above, and the usual determinant of square systems are both special cases of multivariate resultants.

Although Definition 4.2.3 is stated for the special case where the underlying field is the complex numbers, resultants exist over all fields. Just like the bivariate resultant and the determinant, the multivariate resultant does not meaningfully depend on the field of coefficients.

Definition 4.2.4 (see, e.g., [Jou91, Proposition 2.3])**.** Let $\mathbb{F}$ be any field. The *resultant* $\mathrm{Res}_{d_0,\ldots,d_n} \in \mathbb{F}[\mathbf{u}]$ is the unique polynomial satisfying the following conditions.

1. If $F_0, \ldots, F_n \in \mathbb{F}[x_0, \ldots, x_n]$ are homogeneous forms of degrees $d_0, \ldots, d_n$, respectively, then $\mathrm{Res}(F_0, \ldots, F_n) = 0$ if and only if the system of equations $F_0(\mathbf{x}) = \cdots = F_n(\mathbf{x}) = 0$ has a nonzero solution in $\overline{\mathbb{F}}^{n+1}$.
2. $\mathrm{Res}_{d_0,\ldots,d_n}(x_0^{d_0}, \ldots, x_n^{d_n}) = 1$.

Moreover, if we write $\mathrm{Res}^{\mathbb{Z}}_{d_0,\ldots,d_n}$ and $\mathrm{Res}^{\mathbb{F}}_{d_0,\ldots,d_n}$ for the resultants over $\mathbb{Z}$ and $\mathbb{F}$, respectively, then $\mathrm{Res}^{\mathbb{F}}_{d_0,\ldots,d_n}$ is the image of $\mathrm{Res}^{\mathbb{Z}}_{d_0,\ldots,d_n}$ under the natural ring homomorphism $\mathbb{Z}[\mathbf{u}] \to \mathbb{F}[\mathbf{u}]$. $\lozenge$

As with the resultant over the complex numbers, the existence of a polynomial satisfying Definition 4.2.4 is not obvious. We refer the reader interested in the existence of the resultant to Jouanolou [Jou91], who provides a proof of the existence of the resultant over an arbitrary field $\mathbb{F}$. The final statement in Definition 4.2.4, the fact that the resultant over any field is the image of the resultant over the integers will be crucial in our proofs. A slightly weaker version of this fact (which is sufficient for us), that the resultant is the image of the resultant over the prime subfield of $\mathbb{F}$, can be proved using an elementary Gröbner basis argument.

We make one final remark on resultants and inhomogeneous systems of equations before we move on to the computational aspects. Suppose we have a collection of linear polynomials $h_0, \ldots, h_n \in \mathbb{F}[x_1, \ldots, x_n]$ given by $h_i(\mathbf{x}) = \sum_{j=1}^{n} h_{i,j} x_j + h_{i,0}$, then this system has a solution if and only if the column vector $[h_{0,0}, \ldots, h_{n,0}]^T$ is in the span of the matrix $(h_{i,j})_{0 \leq i \leq n, 1 \leq j \leq n}$. There is no polynomial in the $h_{i,j}$ that vanishes exactly when this condition holds. Therefore in general there is no resultant like polynomial that detects when inhomogeneous polynomials have common roots.

The situation can be salvaged to some extent, as follows. Suppose we have inhomogeneous polynomials $f_0, \ldots, f_{n+1} \in \mathbb{F}[x_1, \ldots, x_n]$ of degrees $d_0, \ldots, d_n$. We can homogenise these polynomials with respect to variable x_0. Let F_i be the form obtained by homogenising f_i. We can define an inhomogeneous version of the resultant by setting $\mathrm{Res}(f_0, \ldots, f_n) := \mathrm{Res}(F_0, \ldots, F_n)$. If $f_0, \ldots, f_n$ have a common root, then so do $F_0, \ldots, F_n$, therefore the above inhomogeneous resultant vanishes if $f_0, \ldots, f_n$ have a common root. However, the converse is not true, and the inhomogeneous resultant can vanish even if $f_0, \ldots, f_n$ do not have a common root. This happens exactly when $F_0, \ldots, F_n$ only have roots at the hyperplane at infinity. Therefore care must be taken when dealing with inhomogeneous resultants (these are also sometimes called affine resultants).

If we start with homogeneous forms $F_0, \ldots, F_n$, then set $x_0 = 1$ to obtain inhomogeneous polynomials say $f_0, \ldots, f_n$, and then apply the above construction of the inhomogeneous resultant, we can potentially get something different from $\mathrm{Res}(F_0, \ldots, F_n)$. This occurs when one of the forms F_i is a multiple of x_0. In this case, homogenising f_i does not give us back F_i. As a concrete example, we can consider $F(x, y) = y(x - 2y)$ and $G(x, y) = y(x - 3y)$. The resultant of F, G is zero, since $(1, 0)$ is a non-zero common root. If we set $y = 1$, we obtain polynomials $f = x - 2$ and $g = x - 3$. Homogenising f, g gives us $x - 2y$ and $x - 3y$ respectively, and $\mathrm{Res}(x - 2y, x - 3y) \neq 0$, since these two forms have no common non-zero root. However, if $F_0, \ldots, F_n$ are generic, then none of them will be a multiple of x_0, and in this case by definition the inhomogeneous resultant of $f_0, \ldots, f_n$ and the resultant of $F_0, \ldots, F_n$ will be the same.

4.2.1 Computational complexity of determinants and resultants

In this section, we list some upper and lower bounds on the complexity of resultants and determinants.

The determinant can be computed by bounded fan-in polynomial circuits of depth $O(\log^2 n)$. This was originally proved in [Csa76] when the underlying field had large enough characteristic, and effective circuits in all characteristics were given in [Ber84].

On the other hand, we know strong lower bounds for constant depth circuits computing the determinant: [LST21] showed that any such circuit has super polynomial size when then field has large enough characteristic, and [For24] extended this result to all fields. While it is still open if the determinant has bounded fan-in polynomial sized circuits of depth $O(\log n)$, it is conjectured that this is not the case.

As we saw above, the bivariate resultant can be written as the determinant of the Sylvester matrix $\mathrm{Syl}(F, G)$. As a consequence, the bivariate resultant also has a polynomial sized circuit of depth $O(\log^2 n)$. In a recent breakthrough, Andrews and Wigderson [AW24] showed that the bivariate resultants can actually be computed in constant depth, with polynomial sized circuits, when the underlying field has large enough characteristic. Their result relies on the following formula for the bivariate resultant.

Lemma 4.2.5 (see, e.g., [CLO05, Chapter 3, Section 1]). *Let $f(x) = \sum_{i=0}^{d} f_i x^i$ and $g(x) = \sum_{i=0}^{e} g_i x^i$ be univariate polynomials. Let $\alpha_1, \ldots, \alpha_d \in \overline{\mathbb{F}}$ be the roots of f in $\overline{\mathbb{F}}$ with multiplicity, and likewise let*

$\beta_1, \ldots, \beta_e \in \overline{\mathbb{F}}$ *be the roots of* g *with multiplicity. Then the resultant* $\mathrm{Res}(f, g)$ *satisfies the identity*

$$\mathrm{Res}(f, g) = f_d^e g_e^d \prod_{i=1}^{d} \prod_{j=1}^{e} (\alpha_i - \beta_j).$$

This result was extended to all fields in [BKR+25]. Combined with the lower bound of [LST21; For24], this shows that the bivariate resultant is strictly easier than the determinant for constant depth algebraic circuits. We state the result formally since we will need it in our proofs.

Theorem 4.2.6 ([AW24; BKR+25])**.** *Let* $\mathbb{F}$ *be an infinite field. Let* $f, g \in \mathbb{F}[x]$ *be univariate polynomials given by their coefficients. Then the resultant* $\mathrm{Res}(f, g)$ *can be computed in* $\mathsf{AC}^0_{\mathbb{F}}$.

Before we discuss the complexity of the multivariate resultant, we note its degree.

Lemma 4.2.7 (see, e.g., [CLO05, Chapter 3, Theorem 3.1])**.** *The resultant* $\mathrm{Res}_{d_0,\ldots,d_n}(\mathbf{u})$ *is homogeneous of degree* $d_0 \cdots d_{i-1} d_{i+1} \cdots d_n$ *with respect to the variables* $\{u_{i,\boldsymbol{\alpha}} \mid \boldsymbol{\alpha} \in M_{n,d_i}\}$ *and is homogeneous of total degree* $\sum_{i=0}^{n} d_0 \cdots d_{i-1} d_{i+1} \cdots d_n$.

Multivariate resultants are generally not determinants of matrices of coefficients, in the way bivariate resultants are. However, Macaulay proved that they can be written as quotients of the determinants of coefficient matrices, or the gcd of the determinants of coefficient matrices as follows.

One means of doing this is by forming the *Macaulay matrix* of the forms $F_0, \ldots, F_n$, a square matrix of size $\binom{d_0+\cdots+d_n}{n}$ whose entries are the coefficients of the forms $F_0, \ldots, F_n$. Writing $M(\mathbf{u})$ for the Macaulay matrix of when the coefficients of F_i are the variables $u_{i,\boldsymbol{\alpha}}$, one can show that the resultant is a factor of the determinant of the Macaulay matrix, i.e., that

$$\det M(\mathbf{u}) = \mathrm{Res}_{d_0,\ldots,d_n}(\mathbf{u}) \cdot G(\mathbf{u}).$$

One can show that the cofactor $G(\mathbf{u})$ is a minor of the Macaulay matrix $M(\mathbf{u})$, so the resultant $\mathrm{Res}_{d_0,\ldots,d_n}(\mathbf{u})$ can be written as a quotient of two determinants of size at most $\binom{d_0+\cdots+d_n}{n}$. Alternatively, one can express the resultant as the greatest common divisor of $n+1$ determinants of size $\binom{d_0+\cdots+d_n}{n}$: the construction of the Macaulay matrix is specified by an ordering of the variables $x_0, \ldots, x_n$, and if we let $M_i(\mathbf{u})$ be any Macaulay matrix where the variable x_i comes last in this ordering, then one can show

$$\mathrm{Res}_{d_0,\ldots,d_n}(\mathbf{u}) = \gcd(\det M_0(\mathbf{u}), \ldots, \det M_n(\mathbf{u})).$$

The precise details regarding the Macaulay matrix will not be important for our work, so we refer the reader to Cox, Little, and O'Shea [CLO05, Chatper 3, Section 4] to learn more.

For the purposes of this work, the Macaulay matrix is important insofar as it provides a baseline for the complexity of the multivariate resultant. We can always compute the resultant as a quotient of two determinants of size at most $\binom{d_0+\cdots+d_n}{n}$, or as the greatest common divisor of $n+1$ such determinants. Both constructions yield an arithmetic branching program of size polynomial in $\binom{d_0+\cdots+d_n}{n}$ that computes $\mathrm{Res}_{d_0,\ldots,d_n}(\mathbf{u})$. An improved construction due to Jeronimo and Sabia [JS07] results in an arithmetic circuit whose size is polynomially bounded in the number of variables and degree of $\mathrm{Res}_{d_0,\ldots,d_n}(\mathbf{u})$. In general, the resultant is NP-hard to compute [GKP13], so one should not expect to construct polynomial-size arithmetic circuits for the resultant.

At this point, we can state the main theorem of this section.

Theorem 4.2.8. *Let* $n, d_0, \ldots, d_n \in \mathbb{N}$ *be natural numbers and let* $d := \max(2, d_0, \ldots, d_n)$. *There is an arithmetic circuit of size* $d^{O(n^2)}$ *and depth* $O(n)$ *that computes* $\mathrm{Res}_{d_0,\ldots,d_n}$.

We make some remarks on the parameters of the above circuit. When n is a constant, the size of the input is of the order d^n. The above circuit has polynomial size, and constant depth. In this regime, the circuits obtained by using the Macaulay matrix will have depth $O(n \log d)$. Combined with the lower bound of [LST21], this shows that multivariate resultants in this regime of parameters are strictly easier than determinants for constant depth circuits.

When n is growing, and when d is also growing at least polynomially in n (so $d > n^c$ for some $c > 0$), then the above circuits have quasipolynomial size. Therefore, in this regime we are able to reduce depth but only at the cost of a super polynomial increase in the size.

4.2.2 Low depth circuits for multivariate resultants

We now give our construction for the circuits computing the resultant. Our circuits rely on the following expression for the resultant, which is called the Poisson formula.

Theorem 4.2.9 (see, e.g., [CLO05, Chapter 3, Section 3, Exercise 8])**.** *Let $F_0, \dots, F_n \in \mathbb{F}[\mathbf{x}]$ be homogeneous forms of degrees $d_0, \dots, d_n$, respectively. For $i \in \{0, 1, \dots, n\}$, let*

$$\overline{F}_i(\mathbf{x}) \coloneqq F_i(0, x_1, \dots, x_n)$$
$$f_i(\mathbf{x}) \coloneqq F_i(1, x_1, \dots, x_n).$$

Suppose $\mathrm{Res}_{d_1,\dots,d_n}(\overline{F}_1, \dots, \overline{F}_n) \neq 0$. Then the resultant $\mathrm{Res}_{d_0,\dots,d_n}(F_0, \dots, F_n)$ satisfies the identity

$$\mathrm{Res}_{d_0,\dots,d_n}(F_0, \dots, F_n) = \mathrm{Res}_{d_1,\dots,d_n}(\overline{F}_1, \dots, \overline{F}_n)^{d_0} \cdot \prod_{\boldsymbol{\alpha} \in V(f_1,\dots,f_n)} f_0(\boldsymbol{\alpha})^{m(\boldsymbol{\alpha})},$$

where $V(f_1, \dots, f_n) \subseteq \overline{\mathbb{F}}^n$ is the set of common zeroes of $f_1, \dots, f_n$ in the algebraic closure $\overline{\mathbb{F}}^n$ and $m(\boldsymbol{\alpha}) \in \mathbb{N}$ is the multiplicity of $\boldsymbol{\alpha}$ in $V(f_1, \dots, f_n)$.

We note that the Poisson formula is a generalisation of Lemma 4.2.5 for bivariate resultants. The Poisson formula is false without the assumption that $\mathrm{Res}(\overline{F}_1, \dots, \overline{F}_n) \neq 0$. For example, if $F_1, \dots, F_n$ have a common root with $x_0 = 0$, but $F_0, \dots, F_n$ do not have any common roots, then $\mathrm{Res}(\overline{F}_1, \dots, \overline{F}_n) = 0$ but $\mathrm{Res}(F_0, \dots, F_n) \neq 0$. As we will see, the fact that the condition $\mathrm{Res}(\overline{F}_1, \dots, \overline{F}_n) \neq 0$ holds generically will allow us to apply the Poisson formula without having to worry about this condition.

The first term of the Poisson formula can be computed recursively. To compute the second term of the formula, we need to evaluate the polynomial f_0 at the points of $V(f_1, \dots, f_n)$ while only having access to the coefficients of $f_1, \dots, f_n$, not the points of $V(f_1, \dots, f_n)$ themselves. The $n = 1$ case of this problem is well-understood: this evaluation can be performed using the Newton identities in zero or large characteristic [AW24], or in arbitrary characteristic as a corollary of a more general result on the complexity of symmetric polynomials [BKR+25].

We briefly go over the $n = 1$ case from [AW24; BKR+25], to motivate the rest of the discussion. In this setting, we have univariates f, g and we want to evaluate $\prod_{\beta \in V(g)} f(\beta)$. While we do not have access to the roots β, we do have access to some nice functions of these roots. In particular, the sum of the roots of g is the constant term g_0 of g (up to sign). The product of the roots of g is the term g_{e-1} (assuming g is monic degree e), and in general the coefficients of g are the set of elementary symmetric functions of g. The function $\prod_{\beta \in V(g)} f(\beta)$ is itself another symmetric function in the set of roots of β, and the arguments of [AW24; BKR+25] essentially show that this symmetric function can be computed with just the elementary symmetric functions, without requiring access to the roots themselves.

Returning to the multivariate case, our first goal is to relate the coefficients of the roots with the coefficients of $f_1, \ldots, f_n$. We do this using the hidden variable method. We essentially treat the polynomials $f_1, \ldots, f_n$ as polynomials with coefficients in $\mathbb{F}(x_1)[x_2, \ldots, x_n]$. If we compute the resultant of these polynomials, we get a univariate in x_1. Generically, this univariate will exactly vanish on the first coordinates of the roots of $f_1, \ldots, f_n$.

Lemma 4.2.10 ([CLO05, Chapter 3, Proposition 5.15, Chapter 4, Equation 2.9])**.** *Let $F_1, \ldots, F_n \in \mathbb{F}[\mathbf{x}]$ be generic homogeneous forms of degrees $d_1, \ldots, d_n$, respectively. For $i \in [n]$, let*

$$\overline{F}_i(\mathbf{x}) \coloneqq F_i(0, x_1, \ldots, x_n)$$
$$f_i(\mathbf{x}) \coloneqq F_i(1, x_1, \ldots, x_n).$$

For $i \in [n]$, let $\mathrm{Res}^{x_i}_{d_1,\ldots,d_n}(f_1, \ldots, f_n)$ be the resultant of $f_1, \ldots, f_n$ when viewed as polynomials in $\mathbb{F}(x_i)[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]$. Then there exists a constant $\gamma \in \mathbb{F}$ such that

$$\mathrm{Res}^{x_i}_{d_1,\ldots,d_n}(f_1, \ldots, f_n) = \gamma \prod_{\boldsymbol{\alpha} \in V(f_1,\ldots,f_n)} (x_i - \alpha_i)^{m(\boldsymbol{\alpha})},$$

where $V(f_1, \ldots, f_n) \subseteq \overline{\mathbb{F}}^n$ is the set of common zeroes of $f_1, \ldots, f_n$ in the algebraic closure $\overline{\mathbb{F}}^n$ and $m(\boldsymbol{\alpha}) \in \mathbb{N}$ is the multiplicity of $\boldsymbol{\alpha}$ in $V(f_1, \ldots, f_n)$.

Note that $\mathrm{Res}^{x_i}_{d_1,\ldots,d_n}(f_1, \ldots, f_n)$ is an inhomogeneous resultant. By definition, it is equal to the resultant of the forms obtained by homogenising $f_1, \ldots, f_n$, while considering x_i to have degree 0. This is a resultant over the field $\mathbb{F}(x_i)$. Definition 4.2.4 allows us to reduce this to a computation of a resultant over $\mathbb{F}$, which will be crucial in our inductive construction of the circuit for the resultant.

In a similar way, we can compute all the hidden variable resultants, to get polynomials that vanish on the projection of $V(f_1, \ldots, f_n)$ in each coordinate. We call these hidden variable resultants $r_1(x_1), \ldots, r_n(x_n)$. Generically, $V(f_1, \ldots, f_n)$ has $\prod d_i$ points, and the projections to each coordinate is injective, so each r_i has degree $\prod d_i$. Further, since the projection to each coordinate is generic, the x_1 coordinate of a point in $V(f_1, \ldots, f_n)$ completely determines the point within this set. Equivalently, there are polynomials $p_2, \ldots, p_n$ such that for any point $(\alpha_1, \ldots, \alpha_n)$ in $V(f_1, \ldots, f_n)$ we have $\alpha_i = p_i(\alpha_1)$. We have $V(f_1, \ldots, f_n) = V(r_1(x_1), x_2 - p_2(x_1), \ldots, x_n - p_n(x_1))$. Such a set of polynomials is called a shape basis.

A generalisation of the shape lemma was given in [ABRW96], where instead of having to write the coordinates as polynomial functions of the first coordinate, they allowed the coordinates to be written as rational functions. In other words, a generalised shape basis for $V(f_1, \ldots, f_n)$ is a set of univariate polynomials $p_2, \ldots, p_n, q_2, \ldots, q_n$ such that

$$V(f_1, \ldots, f_n) = V(r_1(x_1), q_2(x_1)x_2 - p_2(x_1), \ldots, q_n(x_1)x_n - p_n(x_1)).$$

If the above holds, then we can write $x_i = p_i/q_i$ on $V(f_1, \ldots, f_n)$. [ABRW96] showed that generalised shape basis can be computed using hidden variable resultants as follows. We introduce a new variable y, and compute the hidden variable resultant of $f_1(x_1 + yx_2, x_2, \ldots, x_n), \ldots, f_n(x_1 + yx_2, \ldots, x_n)$ with respect to x_1. This is a polynomial in two variables, x_1 and y, call this polynomial $r_2(x_1, y)$. Picking $p_2 \coloneqq (\partial_y r_2(x_1, y))|_{y=0}$ and $q_2 = \partial_{x_1} r_1(x_1)$ gives us a generalised shape basis. Once we have a generalised shape basis, we can use the algorithm from the $n = 1$ case to compute product of evaluations.

Each of the above hidden variable resultants is a resultant in a function field over $\mathbb{F}$. By Definition 4.2.4, these resultants can be computed using circuits for the resultant over $\mathbb{F}$ itself, therefore we can do this task by induction. The one caveat is that by induction, we can only give a

bound on the number of $\mathbb{F}(x_1)$ operations or $\mathbb{F}(x_1, y)$ operations. To correctly count the number of $\mathbb{F}$ operations, we must pay a multiplicative factor in the size of the hidden resultants.

We now formalise the above discussion.

Theorem 4.2.11. *Let $n, d_0, \ldots, d_n \in \mathbb{N}$ be natural numbers. Let $d := \max(2, d_0, \ldots, d_n)$. Let $F_0, \ldots, F_n \in \mathbb{F}[x_0, \ldots, x_n]$ be homogeneous forms of degrees $d_0, \ldots, d_n$, respectively. Assume that $F_0, \ldots, F_n$ are generic. Then there is a piecewise arithmetic circuit of size $d^{O(n^2)}$ and depth $O(n)$ that computes the resultant $\mathrm{Res}_{d_0,\ldots,d_n}(F_0, \ldots, F_n)$.*

Proof. We will first describe the circuit, and then bound its size and depth. We proceed by induction on n. In the case where $n = 1$, this is precisely the result of Theorem 4.2.6, so in what remains we assume $n \geq 2$. For $i \in [n]$, let

$$\begin{aligned} \overline{F}_i(\mathbf{x}) &:= F_i(0, x_1, \ldots, x_n) \\ f_i(\mathbf{x}) &:= F_i(1, x_1, \ldots, x_n). \end{aligned}$$

Because $F_0, \ldots, F_n$ are generic, they satisfy the hypothesis of Theorem 4.2.9, so we have

$$\mathrm{Res}_{d_0,\ldots,d_n}(F_0, \ldots, F_n) = \mathrm{Res}_{d_1,\ldots,d_n}(\overline{F}_1, \ldots, \overline{F}_n)^{d_0} \prod_{\boldsymbol{\alpha} \in V(f_1,\ldots,f_n)} f_0(\boldsymbol{\alpha}).$$

The system $\overline{F}_1, \ldots, \overline{F}_n$ is generic, so by induction we can compute $\mathrm{Res}_{d_1,\ldots,d_n}(\overline{F}_1, \ldots, \overline{F}_n)$. All that remains is to compute the product on the right-hand side of the Poisson formula above.

Let $V_1 \subseteq \mathbb{F}$ be the multiset obtained by projecting $V(f_1, \ldots, f_n)$ onto the first coordinate axis. Define

$$r(x_1) := \mathrm{Res}^{x_1}_{d_1,\ldots,d_n}(f_1, \ldots, f_n),$$

where the resultant on the right-hand side above treats $f_1, \ldots, f_n$ as polynomials in the ring $\mathbb{F}(x_1)[x_2, \ldots, x_n]$. While this is a resultant over the ring $\mathbb{F}(x_1)$, the universality of the resultant (Definition 4.2.4) implies that we can compute $r(x_1)$ using the resultant $\mathrm{Res}^{\mathbb{F}}_{d_1,\ldots,d_n}$. We can then compute the coefficients of r using Lemma 2.2.3. Lemma 4.2.10 implies that the set of roots of the polynomial r is precisely V_1.

Fix a choice of $i \in \{2, 3, \ldots, n\}$. Let y be a fresh variable. For each $i \in \{0, 1, \ldots, n\}$, define the polynomials

$$\begin{aligned} G_i(\mathbf{x}, y) &= F_i(x_0, x_1 - yx_j, x_2, \ldots, x_n) \\ \overline{G}_i(\mathbf{x}, y) &= F_i(0, x_1 - yx_j, x_2, \ldots, x_n) \\ g_i(\mathbf{x}, y) &= F_i(1, x_1 - yx_j, x_2, \ldots, x_n), \end{aligned}$$

which we regard as elements of $\overline{\mathbb{F}(y)}[\mathbf{x}]$. Observe that the zero set $V(g_1, \ldots, g_n)$ is given by

$$V(g_1, \ldots, g_n) = \{(\alpha_1 + y\alpha_j, \alpha_2, \ldots, \alpha_n) \,|\, \alpha \in V(f_1, \ldots, f_n)\}.$$

By the same reasoning above, the polynomial

$$r_j(x_1, y) := \mathrm{Res}^{x_1}_{d_1,\ldots,d_n}(g_1, \ldots, g_n)$$

can be factorized as

$$\begin{aligned} r_j(x_1, y) &= \gamma(y) \prod_{\beta \in V(g_1,\ldots,g_n)} (x_1 - \beta_1) \\ &= \gamma(y) \prod_{\alpha \in V(f_1,\ldots,f_n)} (x_1 - \alpha_1 - y\alpha_j) \end{aligned}$$

for some $\gamma(y) \in \overline{\mathbb{F}(y)}$. Note that $g_i(\mathbf{x}, 0) = f_i(\mathbf{x})$, so we have

$$\begin{aligned}\gamma(0) \prod_{\alpha \in V(f_1,\ldots,f_n)} (x_1 - \alpha_1) &= r_j(x_1, 0) \\ &= \operatorname{Res}^{x_1}(g_1(\mathbf{x}, 0), \ldots, g_n(\mathbf{x}, 0)) \\ &= \operatorname{Res}^{x_1}(f_1, \ldots, f_n) \\ &= r_1(x_1).\end{aligned}$$

Let

$$\begin{aligned}h_j(x_1) &\coloneqq \left(\frac{\partial}{\partial y} r_j(x_1, y)\right)_{y=0} \\ &= \left(\gamma'(y) \prod_{\alpha \in V} (x_1 - \alpha_1 - y\alpha_j) + \gamma(y) \left[\sum_{\alpha \in V} -\alpha_j \prod_{\beta \in V : \beta \neq \alpha} (x_1 - \beta_1 - y\beta_j)\right]\right)_{y=0} \\ &= \gamma'(0) \prod_{\alpha \in V} (x_1 - \alpha_1) + \gamma(0) \left[\sum_{\alpha \in V} -\alpha_j \prod_{\beta \in V : \beta \neq \alpha} (x_1 - \beta_1)\right],\end{aligned}$$

where V denotes $V(f_1, \ldots, f_n)$. Observe that if we evaluate h_j at $\mathbf{x} = \alpha$ for a point $\alpha \in V(f_1, \ldots, f_n)$, we obtain

$$\begin{aligned}h_j(\alpha_1) &= -\gamma(0)\alpha_j \prod_{\beta \in V(f_1,\ldots,f_n) : \beta \neq \alpha} (\alpha_1 - \beta_1) \\ &= -\alpha_j r'(\alpha_1).\end{aligned}$$

Thus, for every point $\alpha \in V(f_1, \ldots, f_n)$, we have the equality

$$\alpha_j = \frac{-h_j(\alpha_1)}{r'(\alpha_1)}.$$

For each $i \in [n]$, we can compute the coefficients of $g_i(\mathbf{x}, y)$ via Lemma 2.2.3. We then compute the resultant $r_j(x_1, y)$ recursively. The degree of the resultant $\operatorname{Res}_{d_1,\ldots,d_n}$ is bounded by nd^n. Each coefficient of the polynomials g_i is a polynomial in x_1 and y of degree at most d. It follows that the degree of $r_j(x_1, y)$ as a polynomial in x_1 and y is bounded by $nd^{O(n)} \leq d^{O(n)}$. We can then interpolate the coefficients of r_j, and hence compute the coefficients of $h_j(x_1)$.

Now consider the rational function

$$f(x_1) \coloneqq f_0\left(x_1, \frac{-h_2(x_1)}{r'(x_1)}, \ldots, \frac{-h_n(x_1)}{r'(x_1)}\right).$$

Observe that

$$\prod_{\alpha \in V(f_1,\ldots,f_n)} f(\alpha_1) = \prod_{\alpha \in V(f_1,\ldots,f_n)} f_0\left(\alpha_1, \frac{-h_2(\alpha_1)}{r'(\alpha_1)}, \ldots, \frac{-h_n(\alpha_1)}{r'(\alpha_1)}\right) = \prod_{\alpha \in V(f_1,\ldots,f_n)} f_0(\alpha_1, \alpha_2, \ldots, \alpha_n),$$

which is what we want to compute.

Write $f(x_1) = p(x_1)/q(x_1)$ where $\gcd(p, q) = 1$. Observe that the definition of f implies that $q(x_1)$ is a power of $r'(x_1)$. By applying Strassen's elimination of divisions to the resulting circuit for $f(x_1)$, we obtain circuits to compute $\hat{p}(x)$ and $\hat{q}(x)$ where $f = \hat{p}/\hat{q}$. The degrees of $\hat{p}$ and $\hat{q}$ are

bounded by $d^{O(n)}$, so we can interpolate out their coefficients using Lemma 2.2.3. We can then run the algorithm of Andrews and Wigderson [AW24] to remove the GCD from $\hat{p}$ and $\hat{q}$, resulting in the coefficients of p and q.

Now that we have the coefficients of $p(x_1)$ and $q(x_1)$, we compute the resultants

$$\begin{aligned}\operatorname{Res}(p(x_1), r(x_1)) &= \prod_{\alpha \in V(f_1,\dots,f_n)} p(\alpha_1) \\ \operatorname{Res}(q(x_1), r(x_1)) &= \prod_{\alpha \in V(f_1,\dots,f_n)} q(\alpha_1),\end{aligned}$$

using Theorem 4.2.6. Because q is a power of r' and r is squarefree, we have $\operatorname{Res}(q(x_1), r(x_1)) \neq 0$. This allows us to take the quotient of these two values, which is precisely

$$\prod_{\alpha \in V(f_1,\dots,f_n)} f(\alpha_1) = \prod_{\alpha \in V(f_1,\dots,f_n)} f_0(\alpha_1, \alpha_2, \dots, \alpha_n)$$

as desired.

We now analyze the size and depth of the resulting piecewise circuit to compute the resultant. Letting $S(n,d)$ denote the size of the piecewise circuit. To compute the network for $n+1$ polynomials, we perform a total of $n+1$ hidden variable resultants. Each of these require circuits of size $S(n-1,d)$. For n of these, we have to interpolate out the coefficients. By Lemma 2.2.3, doing so requires circuits of size $O(d^{O(n)}S(n-1,d))$, since the hidden variable resultants have degrees at most $d^{O(n)}$. With all of these at hand, we can write down a circuit for f, and interpolating the coefficients of $\hat{p}, \hat{q}$ and performing the division elimination incurs a multiplicative increase of size by a factor of $d^{O(n)}$. Finally, the GCD algorithm to compute the coefficients of p, q and the algorithm for the univariate resultants cost an additional $d^{O(n)}$, since these are being performed on inputs of degree $d^{O(n)}$. Combining these, we have the recursive bound

$$S(n,d) \leq d^{O(n)}S(n-1,d) + d^{O(n)},$$

which solves to $S(n,d) \leq d^{O(n^2)}$. Letting $\Delta(n,d)$ denote the depth of the piecewise circuit, we have the recursive bound

$$\Delta(n,d) \leq \Delta(n-1,d) + O(1),$$

which solves to $\Delta(n,d) \leq O(1)$. □

Using branching and division elimination, we can immediately deduce algebraic circuits for the resultant that work on all inputs, and not just generic ones.

Corollary 4.2.12. *Let $n, d_0, \dots, d_n \in \mathbb{N}$ be natural numbers. Let $d := \max(2, d_0, \dots, d_n)$. Let $F_0, \dots, F_n \in \mathbb{F}[x_0, \dots, x_n]$ be homogeneous forms of degrees $d_0, \dots, d_n$, respectively. Then there is an algebraic circuit of size $d^{O(n^2)}$ and depth $O(n)$ that computes the resultant $\operatorname{Res}_{d_0,\dots,d_n}(F_0, \dots, F_n)$.*

Proof. The piecewise circuit constructed in Theorem 4.2.11 correctly computes the resultant on a dense set. By Proposition 2.2.7, we can obtain an algebraic circuit with divisions that computes the resultant everywhere, and finally by Theorem 2.2.2 we can eliminate divisions, and obtain a division free algebraic circuit that correctly computes the resultant everywhere. □

4.2.3 Resultant systems

The resultant $\mathrm{Res}_{d_0,\ldots,d_n}$ is a polynomial that detects when a system of $n+1$ homogeneous equations in $n+1$ variables has a solution in $\mathbb{P}^n$. More generally, one can ask if there are similar algebraic criteria that detect when systems of m homogeneous equations in $n+1$ variables have solutions in $\mathbb{P}^n$.

Fix a choice of natural numbers $d_0, \ldots, d_{m-1} \in \mathbb{N}$. For $d \in \mathbb{N}$, let $M_{n,d}$ be the set of homogeneous degree-d monomials in $n+1$ variables. Let $\mathbf{u} = \{u_{i,\boldsymbol{\alpha}} \mid i \in \{0, \ldots, m-1\}, \boldsymbol{\alpha} \in M_{n,d}\}$ be a set of variables, corresponding to the coefficients of m homogeneous forms of degrees $d_0, \ldots, d_n$, respectively. Each specialisation of the variables $\mathbf{u}$ to elements $\boldsymbol{\gamma} \in \overline{\mathbb{F}}^{|\mathbf{u}|}$ gives us a system of m forms of degrees $d_0, \ldots, d_{m-1}$.

The fundamental theorem of elimination theory implies that the set of specialisations that give rise to satisfiable systems is a closed subset of $\overline{\mathbb{F}}^{|\mathbf{u}|}$. When $m = n$, that is, when we are in the square case, Definition 4.2.3 tells us that this closed subset is defined by a single equation, that is, it has codimension 1. In general, the closed subset will have higher codimension, and will be defined by a set of polynomials. A set of polynomials defining this closed subset is called a resultant system. We formalise this in the following definition.

Definition 4.2.13 ([IK93, Section 20.3.1])**.** A resultant system for m homogeneous forms of degrees $d_0, \ldots, d_{m-1}$ in $n+1$ variables is a set of polynomials $g_1, \ldots, g_r \in \mathbb{Z}[\mathbf{u}]$ satisfying the following condition: for every $\boldsymbol{\gamma} \in \overline{\mathbb{F}}^{|\mathbf{u}|}$, the forms obtained by specialising $\mathbf{u}$ to $\boldsymbol{\gamma}$ have a root in $\mathbb{P}^n$ if and only if $g_i(\boldsymbol{\gamma}) = 0$ for $1 \le i \le r$. ◊

We note that when $m \le n$, every set of m homogeneous forms has a common root, therefore set consisting of just the zero polynomial is a resultant system. For the rest of this section, we will focus on the case when $m > n$. When discussing resultant systems, there is a new parameter of interest, namely the number of polynomials in the resultant system, which we call its size. There is a natural trade-off between the size of resultant systems, the degree of the polynomials in the system, and the circuit complexity of the polynomials in the system. In this work, we focus on constructing resultant systems where the polynomials can be computed in constant depth.

It is instructive to consider the familiar special case of linear systems. Recall the set up from the previous section. We have a collection of homogeneous linear forms $L_0, \ldots, L_m \in \mathbb{F}[x_0, \ldots, x_n]$ given by $L_i(\mathbf{x}) = \sum_{j=0}^{n} \ell_{i,j} x_j$. The system of equations $L_0(\mathbf{x}) = \cdots = L_m(\mathbf{x}) = 0$ has a nonzero solution exactly when the matrix of coefficients of $L_0, \ldots, L_m$ (we denote this matrix by L) is not full rank.

When $m \ge n+1$, the system has a solution if and only if L has rank at most n, which holds if and only if every minor of L of size $(n+1) \times (n+1)$ vanishes. This immediately gives us a resultant system of size $\binom{m}{n+1}$ for such equations, namely the determinant of every maximal minor of L. This is not satisfactory, since the size of this system is exponential in the input parameters.

A smaller resultant system can be constructed in this case using the notion of rank condensers. If we multiply the matrix L by a random matrix of size $(n+1) \times m$ on the left, then with high probability the resulting matrix has rank the same as that of L, and the determinant of the resulting matrix can be used to decide if L has a solution. A rank condenser is basically a hitting set for the above randomised algorithm. They originate in the work of [GR08]. We state a specialised notion of these rank condensers that will suffice for us.

Definition 4.2.14. Let $a > b$ be integers. A rank condenser for $a \times b$ matrices is a set of matrices $M_1, \ldots, M_t$ of size $b \times a$ such that for every $a \times b$ matrix N with $\operatorname{rank} N = b$, there is an i such that $\operatorname{rank} M_i N = b$. ◊

Forbes and Shpilka [FS12] and Forbes and Guruswami [FG15] proved that explicit rank condensers of quadratic size exist.

Lemma 4.2.15 ([FS12; FG15]). *Let $\omega \in \mathbb{F}$ be an element of multiplicative order n, and let $\alpha_1, \ldots, \alpha_t \in \mathbb{F}$ be a collection of distinct elements. Suppose M_i is the matrix with $(M_i)_{j,k} = \left(\omega^j \alpha_i\right)^k$. If $t > a(b-a)$ then the above collection of matrices is a rank condenser in the sense of Definition 4.2.14.*

Returning to our setting, given a matrix L of size $m \times (n+1)$, we can create a resultant system by constructing a rank condenser of size $O(mn)$ using Lemma 4.2.15, multiplying L by every matrix in this rank condenser, and computing all the determinants. This gives us a resultant system of quadratic size. Since each polynomial is a determinant, it can be computed by an arithmetic branching program of size polynomial in n, but cannot be computed by a constant depth circuit of size polynomial in n. We also remark that the matrices of rank n have codimension $m-n$ in the space of all $m \times (n+1)$ matrices ([BV88, Proposition 1.1]), therefore there does not exist any set of less than $m-n$ equations on the coefficients of a linear system that detect the existence of a non-zero solution. This shows that the above construction is almost optimal in terms of the size of the system.

We now return to the general case. There are two well known constructions of resultant systems in the literature. The first of these is obtained by using an effective projective Nullstellensatz to obtain an effective version of the main theorem of elimination theory. This construction of resultant systems was given in [IK93]. The projective Nullstellensatz states that a given set of forms $F_0, \ldots, F_{m-1}$ do not have a common zero in $\mathbb{P}^n$ if and only if there exists an integer D such that the ideal containment $(x_0, \ldots, x_n)^D \subseteq (F_0, \ldots, F_{m-1})$ holds. If $\mathbf{x}^{\boldsymbol{\alpha}}$ is a monomial of degree D such that $\mathbf{x}^{\boldsymbol{\alpha}} \in (F_0, \ldots, F_{m-1})$, then by homogeneity there exists forms $G_0, \ldots, G_{m-1}$ of degrees $D - d_0, \ldots, D - d_{m-1}$ such that $\mathbf{x}^{\boldsymbol{\alpha}} = \sum F_i G_i$. Therefore, the above ideal containment is equivalent to the surjectivity of the following map:

$$\mathbb{F}[x_0, \ldots, x_n]_{D-d_0} \times \cdots \times \mathbb{F}[x_0, \ldots, x_n]_{D-d_{m-1}} \to \mathbb{F}[x_0, \ldots, x_n]_D$$

defined by $(G_0, \ldots, G_{m-1}) \to \sum F_i G_i$. This condition is a linear algebraic condition in the coefficients of $F_0, \ldots, F_{m-1}$, and therefore gives rise to a resultant system. An effective version of the projective Nullstellensatz [Laz81] states that it suffices to check the above for $D = 1 + \sum_{i=0}^{n-1}(d_i - 1)$. Using this bound, the problem of deciding if $F_0, \ldots, F_{m-1}$ have a common root reduces to checking if a matrix with $\binom{D}{n}$ rows and $\overline{D} := \sum_{i=0}^{m-1} \binom{D-d_i}{n}$ columns is full rank. The total number of maximal minors of this matrix is very large, therefore in order to check that it is full rank, we can use the rank condensers as discussed above. [4] All the above steps are algebraic in the coefficients of $F_0, \ldots, F_{m-1}$, therefore this gives us a resultant system with $\mathrm{poly}\left(\overline{D}\right)$ many polynomials, each of degree $\binom{D}{n} + 1$. Each polynomial is a determinant of a matrix of size $\binom{D}{n}$.

The second construction of resultant systems is obtained by reducing to the square case. We also sketch this construction, since the resultant system we construct is an improved version of this construction. A full proof can be found for example in [Lan02, Section 4, Chapter 9]. For this discussion we assume that $d_0 = \cdots = d_{m-1}$. The main idea behind the constructions is the following: If the forms $F_0, \ldots, F_{m-1}$ are satisfiable, then any set of forms $P_0, \ldots, P_{n-1}$ obtained by taking linear combinations of $F_0, \ldots, F_{m-1}$ is satisfiable, therefore $\mathrm{Res}(P_0, \ldots, P_{n-1})$ will be zero. On the other hand, if $F_0, \ldots, F_{m-1}$ are unsatisfiable, then if $P_0, \ldots, P_{n-1}$ is a random collection of linear combinations of $F_0, \ldots, F_{m-1}$ then $P_0, \ldots, P_{n-1}$ is unsatisfiable with high probability, therefore $\mathrm{Res}(P_0, \ldots, P_{n-1})$ will be nonzero.

[4] In [IK93], this step is performed using algorithms for fast parallel linear algebra, which gives slightly worse bounds.

To actually construct the resultant system based on the above idea, we need to derandomise the second step. To this end, we consider a sufficient condition on coefficients $\nu_{i,j}$ such that the forms $P_i := \sum_j \nu_{i,j} F_j$ are unsatisfiable if $F_0, \dots, F_{m-1}$ are unsatisfiable. Consider the map $\varphi : \mathbb{P}^n \to \mathbb{P}^{m-1}$ with coordinate functions $F_0, \dots, F_{m-1}$. Since the forms are unsatisfiable, φ is a morphism, and the image has dimension n in $\mathbb{P}^{m-1}$. Suppose $y_0, \dots, y_{m-1}$ are the generators of the homogeneous coordinate ring of $\mathbb{P}^{m-1}$. For a given choice of coefficients $\nu_{i,j}$ we consider the map $\psi_{\boldsymbol{\nu}} : \mathbb{P}^{m-1} \to \mathbb{P}^n$ with coordinate functions $\psi_i := \sum_j \nu_{i,j} y_j$. The composed map $\psi_{\boldsymbol{\nu}} \circ \varphi$ has coordinate functions $P_0, \dots, P_{n-1}$. This composed map is a morphism if and only if the common zeroes of $\sum_j \nu_{i,j} y_j$ are disjointed from $\mathrm{im}(\varphi)$, and this condition is equivalent to the unsatisfiability of $P_0, \dots, P_{n-1}$. This will be the sufficient condition on the coefficients $\nu_{i,j}$ that we work with.

We can also deduce a bound on the probability that a random choice of coefficients chosen from a set satisfies the above sufficient condition as follows. Fix a subset $S \subset \mathbb{F}$. By [BCS97, Theorem 8.48], the variety $\mathrm{im}(\varphi)$ has degree at most d_0^{n+1}. Fix a point in $\mathrm{im}(\varphi)$. The probability that $\sum_j \nu_{0,j} y_j$ vanishes at p is at most $1/|S|$, therefore, with probability $1 - 1/|S|$, the intersection of $\mathrm{im}(\varphi)$ with $\sum_j \nu_{0,j}$ is a variety of dimension $n-1$ and degree d_0^{n+1}. We fix a point from each component of the intersection that does not lie in more than one component, with probability at least $1 - d_0^{n+1}/|S|$, the linear form $\sum_j \nu_{1,j} y_j$ does not vanish at any of these points, and the intersection of $\mathrm{im}(\varphi)$ with $V(\sum_j \nu_{0,j} y_j)$ and $V(\sum_j \nu_{1,j} y_j)$ is a variety of dimension $n-2$ and degree at most d_0^{n+1}. We can repeat this process n times, and take a union bound. As long as $|S| > (n+2) d_0^{n+1}$, for any set of unsatisfiable forms $F_0, \dots, F_{m-1}$, there exists a choice of $\nu_{i,j} \in S$ such that the $n+1$ forms $\sum_j \nu_{i,j} F_j$ are also unsatisfiable. This gives us our second resultant system: we just consider every possible choice of $\nu_{i,j}$ from a fixed S that is big enough (there are $(n+2)^{m(n+1)} d_0^{(n+1)^2 m}$ many such choices), and add to our system the polynomial $\mathrm{Res}\left(\sum_{i,j} \nu_{i,j} F_j\right)$. Each polynomial in the system is a polynomial in the coefficients of $F_0, \dots, F_{m-1}$, and by the above argument, this a valid resultant system.

An alternate way of phrasing the above discussion is as follows. We can start with forms $F_0, \dots, F_{m-1}$, and take $n+1$ many symbolic linear combinations $G_i := \sum y_{i,j} F_j$. We then compute the resultant $\mathrm{Res}(G_0, \dots, G_{n-1})$, treating the G_i as polynomials in $x_0, \dots, x_n$ with coefficients in $\mathbb{F}(\mathbf{y})$. This is a polynomial in $\mathbb{F}[\mathbf{y}]$. Each coefficient is a polynomial in the variables $\mathbf{u}$, and this set of polynomials is a resultant system. The construction in the previous paragraph is essentially the same resultant system as this, where we use evaluation interpolation to extract the coefficients.

Our improvement on this construction comes from the following simple observation.

Lemma 4.2.16. *Suppose $V \subset \mathbb{P}^{m-1}$ is a projective variety of dimension r. There exists elements $\gamma_1, \dots, \gamma_{r+1} \in \mathbb{F}$ with the following property. If we define $L_i := \sum_{j=0}^{n} \gamma_i^j x_j$ and $H_i := V(L_i)$ then the intersection $V \cap H_1 \cap \cdots \cap H_{r+1}$ is empty.*

Proof of Lemma 4.2.16. From each irreducible component of V, we pick a point p_i that does not lie in any other component of V. Consider the polynomial $U = \sum_i z^i x_i$ in the ring $\mathbb{F}[x_0, \dots, x_n, z]$. If we evaluate at p_i, we get $U(p_i, z)$, which is a nonzero polynomial in $\mathbb{F}[z]$, since p_i has at least one non-zero coordinate. If we pick γ_1 a non-root of $\prod_i U(p_i, z)$, then the linear form $L_1 := \sum \gamma^i x_i$ is non vanishing at every p_i. Therefore $V \cap V(L_1)$ has dimension $r-1$. The proof is complete by induction. □

In the framework above, this says that instead of having to pick each $\nu_{i,j}$ at random (which is a total of $O(mn)$ elements), we can just pick $n+1$ elements γ_i and use their powers as the coefficients of the linear forms. In this sense, we have essentially constructed a psuedorandom generator for the above randomised algorithm that stretches $n+1$ bits to $O(mn)$ bits, and has degree d. Because the

dependence of the system size is exponential in the number of variables but only polynomial in the degree, overall this will give us a smaller resultant system.

Theorem 4.2.17. *For any $d_0, \dots, d_{m-1} \in \mathbb{N}$, a resultant system for forms with degrees $d_0, \dots, d_{m-1}$ in $n+1$ variables can be computed by a circuit of depth $O(n)$ and size $d^{O(n^2)}(mn)^{O(n)}$, where $d = \max(2, d_0, \dots, d_{m-1})$. This system consists of at most $d^{O(n^2)}(mn)^{O(n)}$ polynomials, each of degree at most $(n+1)d_0^n$.*

Proof. We can assume without loss of generality that $d_0 \geq d_1 \geq \cdots \geq d_{m-1}$. The resultant system is a set of polynomials in the variables $u_{i,\boldsymbol{\alpha}}$. For convenience, we use G_i to denote the form $\sum u_{i,\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$. For any $\boldsymbol{\gamma}$, we use $G_i(\boldsymbol{\gamma})$ to denote the specialisation of G_i to $\boldsymbol{\gamma}$. A system of forms $F_0, \dots, F_{m-1}$ has a solution in $\mathbb{P}^n$ if and only if the system $x_j F_i$ with $0 \leq i \leq m-1$ and $0 \leq j \leq n$ has a solution in $\mathbb{P}^n$. This holds because any point in $\mathbb{P}^n$ has at least one non-zero coordinate, therefore for a fixed i, a common root of $x_0 F_i, \dots, x_n F_i$ has to be a root of F_i. Let $H_0, \dots, H_{t-1}$ be the set of forms $x_j^{d_0 - d_i} G_i$ for all $0 \leq i \leq m-1$ and $0 \leq j \leq n$. The forms H_i are all of degree d_0 in the variables $\mathbf{x}$. By the above argument, for every $\boldsymbol{\gamma}$, the system $G_0(\boldsymbol{\gamma}), \dots, G_{m-1}(\boldsymbol{\gamma})$ is satisfiable if and only if the system $H_0(\boldsymbol{\gamma}), \dots, H_{t-1}(\boldsymbol{\gamma})$ is satisfiable.

Now let $y_0, \dots, y_n$ be new variables, and let $P_i := \sum y_i^j H_j$. The polynomials P_i are homogeneous in $\mathbf{x}$. Suppose $P_i = \sum_{\boldsymbol{\alpha}} P_{i,\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$. Each $P_{i,\boldsymbol{\alpha}}$ is a polynomial in $y_i, \mathbf{u}$, and there is a $\mathsf{AC}^0_{\mathbb{F}}$ circuit that computes every $P_{i,\boldsymbol{\alpha}}$ given the inputs $\mathbf{y}, \mathbf{u}$. We now compose this circuit with the circuit for $\mathrm{Res}_{d_0,\dots,d_0}$ whose existence was proved in Corollary 4.2.12. The resulting circuit computes a polynomial $R \in \mathbb{Z}[\mathbf{y}, \mathbf{u}]$, that is of degree td_0^n in y_i. We can write $R = \sum_{\boldsymbol{\beta}} R_{\boldsymbol{\beta}} \mathbf{y}^{\beta}$, where the $\boldsymbol{\beta}$ range over all monomials of degree at most td_0^n in each variable. We claim that the set of polynomials $R_{\boldsymbol{\beta}} \in \mathbb{Z}[\mathbf{u}]$ is a resultant system for forms of degrees $d_0, \dots, d_{m-1}$. Before we go on to prove this, we observe that the number of polynomials $R_{\boldsymbol{\beta}}$ is at most $d_0^{n^2+n}((n+1)m)^{n+1}$, and there is a multi-output circuit of size $(m(n+1))^{O(n)} d^{O(n^2)}$ and depth $O(n)$ that computes these polynomials by Lemma 2.2.3. Here we used the fact that $t \leq (m(n+1))$.

We start with the easy direction of the claim. First suppose $\boldsymbol{\gamma}$ is such that $G_0(\boldsymbol{\gamma}), \dots, G_{m-1}(\boldsymbol{\gamma})$ is satisfiable. In this case the set of forms $H_0(\boldsymbol{\gamma}), \dots, H_{t-1}(\boldsymbol{\gamma})$ is also satisfiable. The polynomials $P_0(\boldsymbol{\gamma}), \dots, P_n(\boldsymbol{\gamma})$ are then satisfiable for every specialisation of the variables $\mathbf{y}$. For any specialisation of $\mathbf{y}$ to $\boldsymbol{\nu}$, the resultant of the forms $P_0(\boldsymbol{\gamma}, \boldsymbol{\nu}), \dots, P_n(\boldsymbol{\gamma}, \boldsymbol{\nu})$ is the same as the specialisation $R(\boldsymbol{\gamma}, \boldsymbol{\nu})$. Therefore $R(\boldsymbol{\gamma}, \boldsymbol{\nu}) = 0$ for every $\boldsymbol{\nu}$, which implies that $R_{\boldsymbol{\beta}}(\boldsymbol{\gamma}) = 0$ for every $\boldsymbol{\beta}$.

Now suppose $\boldsymbol{\gamma}$ is such that $G_0(\boldsymbol{\gamma}), \dots, G_{m-1}(\boldsymbol{\gamma})$ is not satisfiable. In this case the set of forms $H_0(\boldsymbol{\gamma}), \dots, H_{t-1}(\boldsymbol{\gamma})$ is also not satisfiable. To complete the proof, it suffices to find a specialisation $\boldsymbol{\nu}$ of $\mathbf{y}$ such that $P_0(\boldsymbol{\gamma}, \boldsymbol{\nu}), \dots, P_n(\boldsymbol{\gamma}, \boldsymbol{\nu})$ is unsatisfiable, in this case $R(\boldsymbol{\gamma}, \boldsymbol{\nu})$ must be non-zero, therefore $R_{\boldsymbol{\beta}}(\boldsymbol{\gamma})$ must be non-zero for at least one $\boldsymbol{\beta}$. Since the forms $H_0(\boldsymbol{\gamma}), \dots, H_{m-1}(\boldsymbol{\gamma})$ are unsatisfiable, they define a morphism $\mathbb{P}^n \to \mathbb{P}^{t-1}$. The image of this morphism has dimension n, therefore by Lemma 4.2.16, we can find $n+1$ linear forms $L_i = \sum_{j=0}^{t-1} \nu_i^j z_j$ such that the linear subspace defined by these forms does not intersect the image of the morphism. Equivalently, the forms $\sum_{j=0}^{t-1} \nu_i^j H_j(\boldsymbol{\gamma})$ are unsatisfiable. This is the desired specialisation $\boldsymbol{\nu}$. □

We make one final improvement to the above system, that applies when the number of polynomials m is larger than $d_0^{n+1} + n$. This improvement will come from applying a technique from [Guo24], and will allow us to get systems of size $d^{O(n^2)} m^{O(1)} n^{O(n)}$ instead of $d^{O(n^2)}(mn)^{O(n)}$. We note that in isolation this technique will not give us systems of size polynomial in m. We require the following technical lemma.

Proposition 4.2.18 ([Har13, Corollary 18.12])**.** *Suppose $V \subset \mathbb{P}^{m-1}$ is an irreducible variety of dimension r and degree D, and suppose* span(V) *is the smallest linear subspace of $\mathbb{P}^{m-1}$ containing*

V. Then V has codimension at most $D-1$ in $\operatorname{span}(V)$. In particular, $\operatorname{span}(V)$ is a linear subspace of dimension at most $D-1+r$.

Corollary 4.2.19. *For any $d_0, \dots, d_{m-1} \in \mathbb{N}$, a resultant system for forms with degrees $d_0, \dots, d_{m-1}$ in $n+1$ variables can be computed by a circuit of depth $O(n)$ and size $d^{O(n^2)} m^{O(1)} n^{O(n)}$, where $d = \max(2, d_0, \dots, d_{m-1})$. This system consists of at most $d^{O(n^2)} m^{O(1)} n^{O(n)}$ polynomials, each of degree at most $(n+1)d_0^n$.*

Proof. As in the proof of Theorem 4.2.17, if the degrees d_i are not all the same, we consider the forms $x_j^{d-d_i} F_i$, this increases m by a factor of at most n, and any resultant system for this set is a system for the original forms. Therefore we can assume that every forms has degree d.

In the case when $m < d_0^{n+1} + n - 1$, we just use the system from Theorem 4.2.17, it already has the required size bounds. Assume $m > a := d_0^{n+1} + n - 1$. We use Lemma 4.2.15 to construct a rank condenser for matrices of size $m \times a$, such a family has size $O(ma)$. For each matrix M_i in this rank condenser, we consider the system of a forms $M_i \mathbf{F}$ (where $\mathbf{F}$ is a column vector of the input forms) and apply Theorem 4.2.17 to construct a resultant system for this set of forms. This system consists of $d^{O(n^2)}(an)^{O(n)}$ many polynomials in $\mathbf{u}$. Taking a union of these systems over all elements of the rank condenser gives us a system whose size we can bound by $d^{O(n^2)} m^{O(1)} n^{O(n)}$. These systems also have the claimed circuit size and depth bounds.

We now show that the above system is a resultant system, which will complete the proof. The easy direction is still easy: if $F_0, \dots, F_{m-1}$ is solvable then so is $M_i \mathbf{F}$ for every i, therefore every polynomial in the resultant system vanishes at $F_0, \dots, F_{m-1}$. Suppose the system is not satisfiable. Then the map φ with coordinate functions F_i is a morphism, and the image has dimension n and degree d_0^{n+1}. Call this image W. By Proposition 4.2.18, the span of W has dimension at most $d_0^{n+1} + n - 1$, let this span be H. Let $H_i \subset \mathbb{P}^{m-1}$ be the linear subspace that is the common zero set of the linear forms whose coefficient vectors are the rows of M_i. By the definition of the rank condensers, there is at least one M_i, say M_1, such that $H \cap H_1 = \emptyset$. In this case, the map corresponding to $M_1 \mathbf{F}$ (which is the composition of φ with a linear map) is a morphism, whence the corresponding system is unsatisfiable. At least one of the equations coming from the resultant system of $M_1 \mathbf{F}$ has to be nonzero when evaluated at the coefficients of $F_0, \dots, F_{m-1}$ which completes the proof. □

Bibliography

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009 (cit. on pp. 1, 12).

[ABRW96] M.-E. Alonso, E. Becker, M.-F. Roy, and T. Wörmann. "Zeros, multiplicities, and idempotents for zero-dimensional systems". In: *Algorithms in algebraic geometry and applications (Santander, 1994)*. Vol. 143. Progr. Math. Birkhäuser, Basel, 1996, pp. 1–15. ISBN: 3-7643-5274-4. DOI: 10.1007/978-3-0348-9104-2_1 (cit. on p. 137).

[AF22] Robert Andrews and Michael A. Forbes. "Ideals, Determinants, and Straightening: Proving and Using Lower Bounds for Polynomial Ideals". In: *Proceedings of the 54th Annual ACM Symposium on Theory of Computing (STOC 2022)*. 2022, pp. 389–402. DOI: 10.1145/3519935.3520025 (cit. on p. 20).

[Agr05] Manindra Agrawal. "Proving Lower Bounds Via Pseudo-random Generators". In: *Proceedings of the 25th International Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2005)*. 2005, pp. 92–105 (cit. on p. 3).

[AH20a] Tigran Ananyan and Melvin Hochster. "Small subalgebras of polynomial rings and Stillman's conjecture". In: *Journal of the American Mathematical Society* 33.1 (2020), pp. 291–309 (cit. on pp. 22, 24, 25).

[AH20b] Tigran Ananyan and Melvin Hochster. "Strength conditions, small subalgebras, and Stillman bounds in degree ≤ 4". In: *Transactions of the American Mathematical Society* 373.7 (2020), pp. 4757–4806 (cit. on pp. 28, 34, 42, 47, 64, 69, 72, 76, 82, 85).

[AKS04] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. "PRIMES is in P". In: *Ann. of Math. (2)* 160.2 (2004), pp. 781–793. ISSN: 0003-486X. DOI: 10.4007/annals.2004.160.781 (cit. on p. 3).

[AO83] Leonard M Adleman and Andrew M Odlyzko. "Irreducibility testing and factorization of polynomials". In: *Mathematics of Computation* 41.164 (1983), pp. 699–709 (cit. on p. 115).

[AV08] M. Agrawal and Manindra. Vinay. "Arithmetic Circuits: A Chasm at Depth Four." In: *In 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28* (2008) (cit. on p. 3).

[AW24] Robert Andrews and Avi Wigderson. "Constant-Depth Arithmetic Circuits for Linear Algebra Problems". In: *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2024)*. 2024. DOI: 10.1109/FOCS61266.2024.00138. arXiv: 2404.10839 [cs.CC] (cit. on pp. 5, 9, 10, 134–136, 140).

[BCS97] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic complexity theory*. Vol. 315. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. With the collaboration of Thomas Lickteig. Springer-Verlag, Berlin, 1997, pp. xxiv+618. ISBN: 3-540-60582-7. DOI: 10.1007/978-3-662-03338-8 (cit. on pp. 2, 7, 9, 10, 143).

[BDWY13] Boaz Barak, Zeev Dvir, Avi Wigderson, and Amir Yehudayoff. "Fractional Sylvester-Gallai theorems". In: *Proceedings of the National Academy of Sciences* 110 (2013) (cit. on p. 15).

[BDYW11] Boaz Barak, Zeev Dvir, Amir Yehudayoff, and Avi Wigderson. "Rank Bounds for Design Matrices with Applications to Combinatorial Geometry and Locally Correctable Codes". In: *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*. STOC '11. San Jose, California, USA, 2011, pp. 519–528. ISBN: 9781450306911. DOI: 10.1145/1993636.1993705 (cit. on pp. 15–17, 37, 97).

[Ber84] Stuart J. Berkowitz. "On computing the determinant in small parallel time using a small number of processors". In: *Information Processing Letters* 18.3 (1984), pp. 147–150. ISSN: 0020-0190. DOI: 10.1016/0020-0190(84)90018-8 (cit. on pp. 113, 134).

[BKR+25] Somnath Bhattacharjee, Mrinal Kumar, Shanthanu Rai, Varun Ramanathan, Ramprasad Saptharishi, and Shubhangi Saraf. *Constant-depth circuits for polynomial GCD over any characteristic*. Electronic Colloquium on Computational Complexity (ECCC), Technical Report TR25-085. 2025 (cit. on pp. 135, 136).

[BM90] Peter Borwein and William OJ Moser. "A survey of Sylvester's problem and its generalizations". In: *Aequationes Mathematicae* 40.1 (1990), pp. 111–135 (cit. on p. 15).

[BM93] Dave Bayer and David Mumford. "What can be computed in algebraic geometry?" In: *arXiv preprint alg-geom/9304003* (1993) (cit. on pp. 4, 6, 111, 112).

[BMS13] Malte Beecken, Johannes Mittmann, and Nitin Saxena. "Algebraic independence and blackbox identity testing". In: *Information and Computation* 222 (2013), pp. 2–19 (cit. on pp. 19, 20).

[Bro87] W. Dale Brownawell. "Bounds for the Degrees in the Nullstellensatz". In: *Annals of Mathematics* 126.3 (1987), pp. 577–591. ISSN: 0003486X, 19398980. DOI: 10.2307/1971361 (cit. on pp. 5, 113).

[BS07] Peter Bürgisser and Peter Scheiblechner. "Differential forms in computational algebraic geometry". In: *Proceedings of the 2007 international symposium on Symbolic and algebraic computation*. 2007, pp. 61–68. DOI: https://doi.org/10.1145/1277548.1277558 (cit. on p. 112).

[BS09] Peter Bürgisser and Peter Scheiblechner. "On the complexity of counting components of algebraic varieties". In: *Journal of Symbolic Computation* 44.9 (2009), pp. 1114–1136 (cit. on p. 6).

[BS83] Walter Baur and Volker Strassen. "The complexity of partial derivatives". In: *Theoretical Computer Science* 22.3 (1983), pp. 317–330. ISSN: 0304-3975. DOI: 10.1016/0304-3975(83)90110-X (cit. on p. 8).

[Bür00a] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*. Springer-Verlag Berlin Heidelberg, 2000. DOI: 10.1007/978-3-662-04179-6 (cit. on pp. 2, 9).

[Bür00b] Peter Bürgisser. "Cook's versus Valiant's hypothesis". In: *Theoretical Computer Science* 235.1 (2000), pp. 71–88 (cit. on p. 2).

[BV88] Winfried Bruns and Udo Vetter. *Determinantal rings*. Vol. 1327. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1988, pp. viii+236. ISBN: 3-540-19468-1. DOI: 10.1007/BFb0080378 (cit. on p. 142).

[CG83] AL Chistov and D Yu Grigoriev. *Subexponential time solving systems of algebraic equations*. Tech. rep. LOMI preprint E-9-83, E-10-83, Steklov Institute, Leningrad, 1983 (cit. on p. 112).

[CKS19] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. "Closure Results for Polynomial Factorization". In: *Theory of Computing* 15.13 (2019). Preliminary version in the *33rd Annual Computational Complexity Conference (CCC 2018)*, pp. 1–34. DOI: 10.4086/toc.2019.v015a013 (cit. on p. 20).

[CLO05] David A. Cox, John Little, and Donal O'Shea. *Using Algebraic Geometry*. 2nd ed. Springer New York, NY, 2005. DOI: 10.1007/b138611 (cit. on pp. 133–137).

[CLO15] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms - an introduction to computational algebraic geometry and commutative algebra*. 4th ed. Undergraduate texts in mathematics. Springer, 2015. ISBN: 978-3-319-16720-6 (cit. on p. 10).

[CM06] Antonio Cafure and Guillermo Matera. "Improved explicit estimates on the number of solutions of equations over a finite field". In: *Finite Fields and Their Applications* 12.2 (2006), pp. 155–185 (cit. on p. 117).

[Csa76] L. Csanky. "Fast Parallel Matrix Inversion Algorithms". In: *SIAM Journal on Computing* 5.4 (1976), pp. 618–623. DOI: 10.1137/0205040 (cit. on pp. 5, 113, 134).

[DDS21] Pranjal Dutta, Prateek Dwivedi, and Nitin Saxena. "Deterministic Identity Testing Paradigms for Bounded Top-Fanin Depth-4 Circuits". In: *36th Computational Complexity Conference (CCC 2021)*. Ed. by Valentine Kabanets. Vol. 200. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021, 11:1–11:27. ISBN: 978-3-95977-193-1. DOI: 10.4230/LIPIcs.CCC.2021.11 (cit. on p. 20).

[DGOS18] Zeev Dvir, Ankit Garg, Rafael Oliveira, and József Solymosi. "Rank bounds for design matrices with block entries and geometric applications". In: *Discrete Analysis* 5.2018 (2018), pp. 1–24 (cit. on pp. 16, 96, 97).

[Dic14] Leonard Eugene Dickson. "The points of inflexion of a plane cubic curve". In: *The Annals of Mathematics* 16.1/4 (1914), pp. 50–66 (cit. on p. 15).

[DL78] Richard A. DeMillo and Richard J. Lipton. "A Probabilistic Remark on Algebraic Program Testing". In: *Inf. Process. Lett.* 7.4 (1978), pp. 193–195. DOI: 10.1016/0020-0190(78)90067-4 (cit. on p. 3).

[DS07] Zeev Dvir and Amir Shpilka. "Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits". In: *SIAM J. Comput.* 36.5 (2007), pp. 1404–1434. ISSN: 0097-5397. DOI: 10.1137/05063605X. Preliminary version in the *37th Annual ACM Symposium on Theory of Computing (STOC 2005)* (cit. on pp. 4, 18).

[DSW14] Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. "Improved rank bounds for design matrices and a new proof of Kelly's theorem". In: *Forum of Mathematics, Sigma*. Vol. 2. Cambridge University Press. 2014 (cit. on pp. 16, 17, 37).

[Dub90] Thomas W Dubé. "The structure of polynomial ideals and Gröbner bases". In: *SIAM Journal on Computing* 19.4 (1990), pp. 750–773. DOI: https://doi.org/10.1137/0219053 (cit. on p. 112).

[EHV92] David Eisenbud, Craig Huneke, and Wolmer Vasconcelos. "Direct methods for primary decomposition". In: *Inventiones mathematicae* 110.1 (1992), pp. 207–235. DOI: https://doi.org/10.1007/BF01231331 (cit. on p. 112).

[Eis95] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Theory*. Springer-Verlag, New York, 1995 (cit. on p. 51).

[EK66] Michael Edelstein and Leroy M Kelly. "Bisecants of finite collections of sets in linear spaces". In: *Canadian Journal of Mathematics* 18 (1966), pp. 375–380 (cit. on p. 15).

[ESS19] Daniel Erman, Steven V. Sam, and Andrew Snowden. "Cubics in 10 variables vs. cubics in 1000 variables: uniformity phenomena for bounded degree polynomials". In: *Bull. Amer. Math. Soc. (N.S.)* 56.1 (2019), pp. 87–114. ISSN: 0273-0979,1088-9485. DOI: 10.1090/bull/1652 (cit. on p. 24).

[FG15] Michael A. Forbes and Venkatesan Guruswami. "Dimension Expanders via Rank Condensers". In: *Proceedings of the 19th International Workshop on Randomization and Computation (RANDOM 2015)*. Ed. by Naveen Garg, Klaus Jansen, Anup Rao, and José D. P. Rolim. Vol. 40. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015, pp. 800–814. ISBN: 978-3-939897-89-7. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2015.800 (cit. on p. 142).

[For24] Michael A. Forbes. "Low-Depth Algebraic Circuit Lower Bounds over Any Field". In: *39th Computational Complexity Conference (CCC 2024)*. Ed. by Rahul Santhanam. Vol. 300. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, 31:1–31:16. ISBN: 978-3-95977-331-7. DOI: 10.4230/LIPIcs.CCC.2024.31 (cit. on pp. 134, 135).

[FOV99] Hubert Flenner, Liam O'Carroll, and Wolfgang Vogel. *Joins and intersections*. Springer, 1999 (cit. on pp. 52–54).

[FS12] Michael A. Forbes and Amir Shpilka. "On identity testing of tensors, low-rank recovery and compressed sensing". In: *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC 2012)*. 2012, pp. 163–172 (cit. on p. 142).

[Ful89] William Fulton. *Algebraic Curves*. Addison Wesley Publishing Company, 1989 (cit. on p. 57).

[Gal44] Tibor Gallai. "Solution of problem 4065". In: *American Mathematical Monthly* 51 (1944), pp. 169–171 (cit. on p. 15).

[GG13] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. 3rd ed. Cambridge University Press, 2013 (cit. on pp. 1, 132).

[GH91] Marc Giusti and Joos Heintz. "Algorithmes–disons rapides–pour la décomposition d'une variété algébrique en composantes irréductibles et équidimensionnelles". In: *Effective Methods in Algebraic Geometry*. Springer. 1991, pp. 169–194 (cit. on p. 112).

[GHH+97] Marc Giusti, J Heintz, K Hägele, Jose E Morais, LM Pardo, and JL Montana. "Lower bounds for Diophantine approximations". In: *Journal of Pure and Applied Algebra* 117 (1997), pp. 277–317 (cit. on p. 2).

[GHS93] Marc Giusti, Joos Heintz, and Juan Sabia. "On the efficiency of effective Nullstellensätze". In: *Computational complexity* 3.1 (1993), pp. 56–95 (cit. on p. 2).

[GKKS13] A. Gupta, P. Kamath, N. Kayal, and R. Saptharishi. "Arithmetic Circuits: A Chasm at Depth Three." In: *In 54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013* (2013) (cit. on pp. 3, 4).

[GKP13] Bruno Grenet, Pascal Koiran, and Natacha Portier. "On the complexity of the multivariate resultant". In: *Journal of Complexity* 29.2 (2013), pp. 142–157. ISSN: 0885-064X. DOI: 10.1016/j.jco.2012.10.001 (cit. on p. 135).

[GKZ94] Israel M. Gelfand, Mikhail M. Kapranov, and Andrei V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants*. 1st ed. Birhaäuser Boston, MA, 1994, pp. x+523. ISBN: 978-0-8176-4771-1. DOI: 10.1007/978-0-8176-4771-1 (cit. on p. 133).

[GOPS23] Abhibhav Garg, Rafael Oliveira, Shir Peleg, and Akash Kumar Sengupta. "Radical Sylvester-Gallai Theorem for Tuples of Quadratics". In: *38th Computational Complexity Conference (CCC 2023)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. 2023, pp. 20–1 (cit. on pp. iv, 6, 22, 27–29, 34, 41, 46).

[GOS22] Abhibhav Garg, Rafael Oliveira, and Akash Kumar Sengupta. "Robust Radical Sylvester-Gallai Theorem for Quadratics". In: *38th International Symposium on Computational Geometry (SoCG 2022)*. Vol. 224. Leibniz International Proceedings in Informatics (LIPIcs). 2022, 42:1–42:13. ISBN: 978-3-95977-227-3. DOI: 10.4230/LIPIcs.SoCG.2022.42 (cit. on pp. iv, 22–24, 37, 41, 46).

[GOS25a] Abhibhav Garg, Rafael Oliveira, and Nitin Saxena. "Primes via zeros: interactive proofs for testing primality of natural classes of ideals". In: *STOC'25—Proceedings of the 57th Annual ACM Symposium on Theory of Computing*. ACM, New York, [2025] ©2025, pp. 1079–1086. ISBN: 979-8-4007-1510-5. DOI: 10.1145/3717823.3718167 (cit. on pp. iv, 6, 119).

[GOS25b] Abhibhav Garg, Rafael Oliveira, and Akash Sengupta. *Rank Bounds and PIT for $\Sigma^3\Pi\Sigma\Pi^d$ circuits via a non-linear Edelstein-Kelly theorem*. https://arxiv.org/abs/2504.14729. 2025 (cit. on pp. iv, 6, 22, 46, 48).

[GOS25c] Abhibhav Garg, Rafael Oliveira, and Akash Kumar Sengupta. "Uniform Bounds on Product Sylvester-Gallai Configurations". In: *41st International Symposium on Computational Geometry (SoCG 2025)*. Ed. by Oswin Aichholzer and Haitao Wang. Vol. 332. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025, 52:1–52:15. ISBN: 978-3-95977-370-6. DOI: 10.4230/LIPIcs.SoCG.2025.52 (cit. on pp. iv, 6, 22, 46, 56, 59, 70).

[GR08] Ariel Gabizon and Ran Raz. "Deterministic extractors for affine sources over large fields". In: *Combinatorica* 28 (2008), pp. 415–440. DOI: 10.1007/s00493-008-2259-3 (cit. on pp. 19, 141).

[GS86] Shafi Goldwasser and Michael Sipser. "Private coins versus public coins in interactive proof systems". In: *Proceedings of the eighteenth annual ACM symposium on Theory of computing*. 1986, pp. 59–68 (cit. on pp. 12, 116).

[GSS19] Zeyu Guo, Nitin Saxena, and Amit Sinhababu. "Algebraic dependencies and PSPACE algorithms in approximative complexity over any field". In: *Theory of Computing* 15.1 (2019), pp. 1–30 (cit. on p. 11).

[GTZ88] Patrizia Gianni, Barry Trager, and Gail Zacharias. "Gröbner bases and primary decomposition of polynomial ideals". In: *Journal of Symbolic Computation* 6.2-3 (1988), pp. 149–167. DOI: https://doi.org/10.1016/S0747-7171(88)80040-3 (cit. on p. 112).

[Guo24] Zeyu Guo. "Variety Evasive Subspace Families". In: *computational complexity* 33.2 (2024), p. 10. DOI: 10.1007/s00037-024-00256-1 (cit. on p. 144).

[Gup14] Ankit Gupta. "Algebraic Geometric Techniques for Depth-4 PIT & Sylvester-Gallai Conjectures for Varieties." In: *Electron. Colloquium Comput. Complex.* Vol. 21. 2014, p. 130 (cit. on pp. 4, 14, 19).

[Han65] Sten Hansen. "A generalization of a theorem of Sylvester on the lines determined by a finite point set". In: *Mathematica Scandinavica* 16.2 (1965), pp. 175–180 (cit. on pp. 15, 16).

[Har13] Joe Harris. *Algebraic geometry: a first course*. Vol. 133. Springer Science & Business Media, 2013. DOI: doi.org/10.1007/978-1-4757-2189-8 (cit. on p. 144).

[Har77] Robin Hartshorne. *Algebraic Geometry*. Springer-Verlag, 1977 (cit. on p. 54).

[Hei83] Joos Heintz. "Definability and fast quantifier elimination in algebraically closed fields". In: *Theoretical Computer Science* 24.3 (1983), pp. 239–277 (cit. on p. 10).

[Her26] Grete Hermann. "Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. (Unter Benutzung nachgelassener Sätze von K. Hentzelt)". In: *Mathematische Annalen* 95 (1926), pp. 736–788. DOI: 10.1007/BF01206635 (cit. on pp. 4, 112, 113).

[Hir83] Friedrich Hirzebruch. "Arrangements of lines and algebraic surfaces". In: *Arithmetic and geometry*. Springer, 1983, pp. 113–140 (cit. on p. 15).

[HS13] Marc Hindry and Joseph H Silverman. *Diophantine geometry: an introduction*. Vol. 201. Springer Science & Business Media, 2013 (cit. on p. 125).

[HS80] Joos Heintz and Claus-Peter Schnorr. "Testing polynomials which are easy to compute". In: *Proceedings of the twelfth annual ACM Symposium on Theory of Computing*. 1980, pp. 262–272 (cit. on pp. 3, 10).

[Huỳ86] Dũng T. Huỳnh. "A superexponential lower bound for Gröbner bases and Church-Rosser commutative Thue systems". In: *Inform. and Control* 68.1-3 (1986), pp. 196–206. ISSN: 0019-9958. DOI: 10.1016/S0019-9958(86)80035-3 (cit. on p. 112).

[Ier89] Doug Ierardi. "Quantifier Elimination in the Theory of an Algebraically-closed Field". In: *Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC 1989)*. Ed. by David S. Johnson. ACM, 1989, pp. 138–147. DOI: 10.1145/73007.73020 (cit. on p. 111).

[IK93] Douglas Ierardi and Dexter Kozen. "Parallel Resultant Computation". In: *Synthesis of Parallel Algorithms*. Ed. by John H. Reif. Morgan Kaufmann Publishers Inc., 1993, pp. 679–720 (cit. on pp. 141, 142).

[Jel05] Zbigniew Jelonek. "On the effective Nullstellensatz". In: *Invent. Math.* 162.1 (2005), pp. 1–17. DOI: 10.1007/s00222-004-0434-8 (cit. on pp. 5, 12, 113).

[Jou91] Jean-Pierre Jouanolou. "Le formalisme du résultant". In: *Adv. Math.* 90.2 (1991), pp. 117–263. ISSN: 0001-8708,1090-2082. DOI: 10.1016/0001-8708(91)90031-2 (cit. on p. 133).

[JS07] Gabriela Jeronimo and Juan Sabia. "Computing multihomogeneous resultants using straight-line programs". In: *J. Symbolic Comput.* 42.1-2 (2007), pp. 218–235. ISSN: 0747-7171,1095-855X. DOI: 10.1016/j.jsc.2006.03.006 (cit. on p. 135).

[Kal85] Erich Kaltofen. "Fast parallel absolute irreducibility testing". In: *Journal of Symbolic Computation* 1.1 (1985), pp. 57–67. DOI: https://doi.org/10.1016/S0747-7171(85)80029-8 (cit. on p. 2).

[Kal87] Erich Kaltofen. "Single-Factor Hensel Lifting and its Application to the Straight-Line Complexity of Certain Polynomials". In: *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*. 1987, pp. 443–452. DOI: 10.1145/28395.28443 (cit. on p. 2).

[Kal89] Erich Kaltofen. "Factorization of Polynomials Given by Straight-Line Programs". In: *Advances in Computing Research* 5 (1989), pp. 375–412 (cit. on p. 2).

[Kal95] Erich Kaltofen. "Effective Noether Irreducibility Forms and Applications". In: *Journal of Computer and System Sciences* 50.2 (1995), pp. 274–295. DOI: https://doi.org/10.1006/jcss.1995.1023 (cit. on pp. 121, 122, 125).

[KI04] Valentine Kabanets and Russell Impagliazzo. "Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds". In: *Computational Complexity* 13.1-2 (2004), pp. 1–46. DOI: 10.1007/s00037-004-0182-6 (cit. on p. 3).

[Koi12] Pascal Koiran. "Arithmetic circuits: the chasm at depth four gets wider". In: *Theoret. Comput. Sci.* 448 (2012), pp. 56–65. ISSN: 0304-3975,1879-2294. DOI: 10.1016/j.tcs.2012.03.041 (cit. on p. 3).

[Koi96] Pascal Koiran. "Hilbert's Nullstellensatz is in the polynomial hierarchy". In: *Journal of complexity* 12.4 (1996), pp. 273–286. DOI: https://doi.org/10.1006/jcom.1996.0019 (cit. on pp. 5, 113, 115, 116).

[Koi97] Pascal Koiran. "Randomized and deterministic algorithms for the dimension of algebraic varieties". In: *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*. FOCS '97. USA: IEEE Computer Society, 1997, p. 36. ISBN: 0818681977 (cit. on pp. 112, 116).

[Kol88] János Kollár. "Sharp effective Nullstellensatz". In: *J. Amer. Math. Soc.* 1.4 (1988), pp. 963–975. DOI: 10.2307/1990996 (cit. on pp. 5, 113).

[KP96] T. Krick and L. M. Pardo. "A computational method for Diophantine approximation". In: *Algorithms in algebraic geometry and applications (Santander, 1994)*. Vol. 143. Progr. Math. Birkhäuser, Basel, 1996, pp. 193–253. ISBN: 3-7643-5274-4 (cit. on pp. 113, 114).

[KPS01] Teresa Krick, Luis Miguel Pardo, and Martín Sombra. "Sharp estimates for the arithmetic Nullstellensatz". In: *Duke Mathematical Journal* 109.3 (2001), pp. 521–598. DOI: 10.1215/S0012-7094-01-10934-4 (cit. on pp. 122–125).

[KS07] Neeraj Kayal and Nitin Saxena. "Polynomial identity testing for depth 3 circuits". In: *Comput. Complexity* 16.2 (2007), pp. 115–138. ISSN: 1016-3328. DOI: 10.1007/s00037-007-0226-9. Preliminary version in the *21st Annual IEEE Conference on Computational Complexity (CCC 2006)* (cit. on p. 19).

[KS09] Neeraj Kayal and Shubhangi Saraf. "Blackbox polynomial identity testing for depth 3 circuits". In: *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)*. IEEE Computer Soc., Los Alamitos, CA, 2009, pp. 198–207. DOI: 10.1109/FOCS.2009.67 (cit. on pp. 4, 15, 19, 84, 109).

[KS11] Zohar S. Karnin and Amir Shpilka. "Black box polynomial identity testing of generalized depth-3 arithmetic circuits with bounded top fan-in". In: *Combinatorica* 31.3 (2011), pp. 333–364. ISSN: 0209-9683. DOI: 10.1007/s00493-011-2537-3. Preliminary version in the *23rd Annual IEEE Conference on Computational Complexity (CCC 2008)* (cit. on pp. 4, 19).

[KS19] Mrinal Kumar and Ramprasad Saptharishi. "Hardness-randomness tradeoffs for algebraic computation". In: *Bulletin of EATCS* 3.129 (2019) (cit. on p. 3).

[KT90] Erich Kaltofen and Barry M. Trager. "Computing with Polynomials Given By Black Boxes for Their Evaluations: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators". In: *J. Symb. Comput.* 9.3 (1990), pp. 301–320. DOI: 10.1016/S0747-7171(08)80015-6 (cit. on p. 2).

[KUW86] Richard M. Karp, Eli Upfal, and Avi Wigderson. "Constructing a perfect matching is in Random NC". In: *Combinatorica* 6.1 (1986), pp. 35–48. ISSN: 0209-9683. DOI: 10.1007/BF02579407. Preliminary version in the *17th Annual ACM Symposium on Theory of Computing (STOC 1985)* (cit. on p. 3).

[Lan02] Serge Lang. *Algebra*. third. Vol. 211. Graduate Texts in Mathematics. Springer-Verlag, New York, 2002, pp. xvi+914. ISBN: 0-387-95385-X. DOI: 10.1007/978-1-4613-0041-0 (cit. on p. 142).

[Las05] Emanuel Lasker. "Zur Theorie der moduln und Ideale". In: *Mathematische Annalen* 60.1 (1905), pp. 20–116 (cit. on p. 112).

[Laz81] Daniel Lazard. "Résolution des systèmes d'équations algébriques". In: *Theoret. Comput. Sci.* 15.1 (1981), pp. 77–110. ISSN: 0304-3975,1879-2294. DOI: 10.1016/0304-3975(81)90064-5 (cit. on p. 142).

[LO77] Jeffrey C Lagarias and Andrew M Odlyzko. "Effective versions of the Chebotarev density theorem". In: *Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975)*. Vol. 7. 1977, pp. 409–464 (cit. on p. 115).

[Lov79] László Lovász. "On determinants, matchings, and random algorithms". In: *Fundamentals of computation theory (Proc. Conf. Algebraic, Arith. and Categorical Methods in Comput. Theory, Berlin/Wendisch-Rietz, 1979)*. Vol. 2. Math. Res. Akademie-Verlag, Berlin, 1979, pp. 565–574 (cit. on p. 3).

[LST21] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. "Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits". In: *Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2021)*. 2021, pp. 804–814. DOI: 10.1109/FOCS52979.2021.00083 (cit. on pp. 20, 134–136).

[Mac94] Francis Sowerby Macaulay. *The algebraic theory of modular systems*. Cambridge University Press, 1994 (cit. on p. 133).

[Mar18] Daniel A. Marcus. *Number fields*. Second. Universitext. With a foreword by Barry Mazur. Springer, Cham, 2018, pp. xviii+203. ISBN: 978-3-319-90232-6; 978-3-319-90233-3. DOI: 10.1007/978-3-319-90233-3 (cit. on p. 118).

[Mel40] Eberhard Melchior. “Uber vielseite der projektiven ebene”. In: *Deutsche Math* 5 (1940), pp. 461–475 (cit. on p. 15).

[Mil20] James S. Milne. *Algebraic Number Theory (v3.08)*. Available at www.jmilne.org/math/. 2020 (cit. on p. 118).

[MM82] Ernst W Mayr and Albert R Meyer. “The complexity of the word problems for commutative semigroups and polynomial ideals”. In: *Advances in mathematics* 46.3 (1982), pp. 305–329 (cit. on pp. 4, 112, 113).

[MS13] Jason McCullough and Alexandra Seceleanu. “Bounding projective dimension”. In: *Commutative algebra*. Springer, New York, 2013, pp. 551–576. ISBN: 978-1-4614-5291-1; 978-1-4614-5292-8. DOI: 10.1007/978-1-4614-5292-8_17 (cit. on p. 24).

[MV05] Peter Bro Miltersen and N. V. Vinodchandran. “Derandomizing Arthur-Merlin games using hitting sets”. In: *Comput. Complexity* 14.3 (2005), pp. 256–279. ISSN: 1016-3328,1420-8954. DOI: 10.1007/s00037-005-0197-7 (cit. on p. 5).

[MVV87] Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. “Matching is as easy as matrix inversion”. In: *Combinatorica* 7.1 (1987), pp. 105–113. ISSN: 0209-9683. DOI: 10.1007/BF02579206. Preliminary version in the *19th Annual ACM Symposium on Theory of Computing (STOC 1987)* (cit. on p. 3).

[Nes85] Yu V Nesterenko. “On algebraic independence of algebraic powers of algebraic numbers”. In: *Sbornik: Mathematics* 51.2 (1985), pp. 429–454 (cit. on p. 122).

[Nes87] Yu V Nesterenko. “On a measure of the algebraic independence of the values of certain functions”. In: *Mathematics of the USSR-Sbornik* 56.2 (1987), p. 545 (cit. on p. 122).

[NW94] Noam Nisan and Avi Wigderson. “Hardness vs. randomness”. In: *J. Comput. System Sci.* 49.2 (1994), pp. 149–167. ISSN: 0022-0000. DOI: 10.1016/S0022-0000(05)80043-1 (cit. on p. 3).

[Ore21] Ø. Ore. *Über höhere Kongruenzen*. Norsk Matematisk Forenings skrifter. Grøndahl, 1921 (cit. on p. 3).

[OS22] Rafael Oliveira and Akash Sengupta. “Radical Sylvester-Gallai theorem for cubics”. In: *FOCS* (2022) (cit. on pp. 22, 24–28, 31, 32, 34, 41, 46, 49).

[OS24] Rafael Oliveira and Akash Kumar Sengupta. “Strong Algebras and Radical Sylvester-Gallai Configurations”. In: *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*. 2024, pp. 95–105 (cit. on pp. 16, 22, 31, 41–43, 46–49, 51, 59, 70, 81, 83, 93).

[Phi86] Patrice Philippon. “Critères pour l’indépendance algébrique”. In: *Inst. Hautes Études Sci. Publ. Math.* 64 (1986), pp. 5–52. ISSN: 0073-8301,1618-1913 (cit. on pp. 124, 133).

[Phi91] Patrice Philippon. “Sur des hauteurs alternatives. I”. In: *Mathematische Annalen* 289.1 (1991), pp. 255–283 (cit. on p. 122).

[Phi94] Patrice Philippon. “Sur des hauteurs alternatives. II”. In: *Annales de l’institut Fourier*. Vol. 44. 4. 1994, pp. 1043–1065 (cit. on p. 122).

[Phi95] Patrice Philippon. “Sur des hauteurs alternatives III”. In: *Journal de mathématiques pures et appliquées* 74.4 (1995), pp. 345–365 (cit. on pp. 122, 123).

[PS20] Shir Peleg and Amir Shpilka. "A Generalized Sylvester-Gallai Type Theorem for Quadratic Polynomials". In: *35th Computational Complexity Conference (CCC 2020)*. 2020, 8:1–8:33. ISBN: 978-3-95977-156-6. DOI: 10.4230/LIPIcs.CCC.2020.8 (cit. on pp. 22–24, 28, 31, 32, 41, 46, 85, 91).

[PS21] Shir Peleg and Amir Shpilka. "Polynomial Time Deterministic Identity Testing Algorithm for $\Sigma[3]\Pi\Sigma\Pi[2]$ Circuits via Edelstein–Kelly Type Theorem for Quadratic Polynomials". In: *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing (STOC 2021)*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 259–271. ISBN: 9781450380539 (cit. on pp. 17, 22–24, 41, 46).

[PS22] Shir Peleg and Amir Shpilka. "Robust Sylvester-Gallai Type Theorem for Quadratic Polynomials". In: *38th International Symposium on Computational Geometry, SoCG 2022*. Vol. 224. LIPIcs. 2022, 43:1–43:15. DOI: 10.4230/LIPIcs.SoCG.2022.43 (cit. on pp. 22–24, 41, 46).

[Ras99] Laila EM Rashid. "Application of Global Bertini Theorems". In: *Bulletin mathématique de la Société des Sciences Mathématiques de Roumanie* (1999), pp. 279–289 (cit. on p. 53).

[Sch07] Peter Scheiblechner. "On the Complexity of Counting Irreducible Components and Computing Betti Numbers of Algebraic Varieties". PhD thesis. University of Paderborn, 2007. DOI: https://doi.org/10.1007/s00037-009-0283-3 (cit. on pp. 6, 111).

[Sch80] Jacob T. Schwartz. "Fast Probabilistic Algorithms for Verification of Polynomial Identities". In: *J. ACM* 27.4 (1980), pp. 701–717. DOI: 10.1145/322217.322225 (cit. on p. 3).

[Sei74] Abraham Seidenberg. "Constructions in algebra". In: *Transactions of the American Mathematical Society* 197 (1974), pp. 273–313 (cit. on pp. 112, 113).

[Sei78] Abraham Seidenberg. "Constructions in a polynomial ring over the ring of integers". In: *American Journal of Mathematics* 100.4 (1978), pp. 685–703 (cit. on p. 112).

[Ser66] Jean-Pierre Serre. "Advanced problem 5359". In: *Amer. Math. Monthly* 73.1 (1966), p. 89 (cit. on p. 15).

[Shp20] Amir Shpilka. "Sylvester-Gallai type theorems for quadratic polynomials". In: *Discrete Analysis* (2020), p. 14492 (cit. on pp. 16, 17, 22–24, 31, 41, 46, 85).

[Sin16] Gaurav Sinha. "Reconstruction of Real Depth-3 Circuits with Top Fan-In 2". In: *31st Conference on Computational Complexity (CCC 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2016 (cit. on p. 15).

[SL96] P. Stevenhagen and H. W. Lenstra Jr. "Chebotarëv and his density theorem". In: *Math. Intelligencer* 18.2 (1996), pp. 26–37. ISSN: 0343-6993,1866-7414. DOI: 10.1007/BF03027290 (cit. on p. 114).

[SR94] Igor Rostislavovich Shafarevich and Miles Reid. *Basic algebraic geometry*. Vol. 2. Springer, 1994 (cit. on pp. 10, 11).

[SS11] Nitin Saxena and C. Seshadhri. "An almost optimal rank bound for depth-3 identities". In: *SIAM J. Comput.* 40.1 (2011), pp. 200–224. ISSN: 0097-5397. DOI: 10.1137/090770679. Preliminary version in the *24th Annual IEEE Conference on Computational Complexity (CCC 2009)* (cit. on pp. 4, 19).

[SS12] Nitin Saxena and C. Seshadhri. "Blackbox identity testing for bounded top-fanin depth-3 circuits: the field doesn't matter". In: *SIAM J. Comput.* 41.5 (2012), pp. 1285–1298. ISSN: 0097-5397. DOI: 10.1137/10848232. Preliminary version in the *43rd Annual ACM Symposium on Theory of Computing (STOC 2011)* (cit. on pp. 4, 19).

[SS13] Nitin Saxena and Comandur Seshadhri. "From Sylvester-Gallai configurations to rank bounds: Improved blackbox identity test for depth-3 circuits". In: *Journal of the ACM (JACM)* 60.5 (2013), pp. 1–33 (cit. on pp. 4, 15, 19).

[SS25] Shubhangi Saraf and Devansh Shringi. "Reconstruction of Depth 3 Arithmetic Circuits with Top Fan-In 3". In: *40th Computational Complexity Conference (CCC 2025)*. Ed. by Srikanth Srinivasan. Vol. 339. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025, 21:1–21:22. ISBN: 978-3-95977-379-9. DOI: 10.4230/LIPIcs.CCC.2025.21 (cit. on p. 15).

[Stacks] The Stacks Project Authors. *Stacks Project*. Open Source, 2015 (cit. on pp. 50, 58).

[Str73] Volker Strassen. "Vermeidung von Divisionen". In: *J. Reine Angew. Math.* 264 (1973), pp. 184–202. ISSN: 0075-4102 (cit. on p. 8).

[SY10] Amir Shpilka and Amir Yehudayoff. "Arithmetic Circuits: A survey of recent results and open questions". In: *Foundations and Trends in Theoretical Computer Science* 5.3-4 (2010), pp. 207–388. DOI: 10.1561/0400000039 (cit. on pp. 2, 7).

[Syl93] James Joseph Sylvester. "Mathematical question 11851". In: *Educational Times* 59.98 (1893), p. 256 (cit. on p. 15).

[Tav15] Sébastien Tavenas. "Improved bounds for reduction to depth 4 and depth 3". In: *Information and Computation* 240 (2015), pp. 2–11 (cit. on pp. 3, 4).

[Val79] Leslie G. Valiant. "Completeness Classes in Algebra". In: *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC 1979)*. Atlanta, Georgia, USA: Association for Computing Machinery, 1979, pp. 249–261. ISBN: 9781450374385. DOI: 10.1145/800135.804419 (cit. on p. 2).

[VSBR83] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. "Fast parallel computation of polynomials using few processors". In: *SIAM J. Comput.* 12.4 (1983), pp. 641–644. ISSN: 0097-5397. DOI: 10.1137/0212043 (cit. on p. 3).

[Wal00] Michel Waldschmidt. *Diophantine approximation on linear algebraic groups*. Vol. 326. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Transcendence properties of the exponential function in several variables. Springer-Verlag, Berlin, 2000, pp. xxiv+633. ISBN: 3-540-66785-7. DOI: 10.1007/978-3-662-11569-5 (cit. on pp. 119–122, 126).

[Zip79] Richard Zippel. "Probabilistic algorithms for sparse polynomials". In: *Proceedings of the International Symposium on Symbolic and Algebraic Computation, EUROSAM 1979*. 1979, pp. 216–226. DOI: 10.1007/3-540-09519-5_73 (cit. on p. 3).