An Almost Optimal Rank Bound for Depth-3 Identities*

Nitin Saxena † C. Seshadhri ‡

Abstract

We study the problem of polynomial identity testing for depth-3 circuits of degree d and top fanin k. The rank of any such identity is essentially the minimum number of independent variables present. Small bounds on this quantity imply fast deterministic identity testers for these circuits. Dvir & Shpilka (STOC 2005) initiated the study of the rank and showed that any depth-3 identity (barring some uninteresting corner cases) has a rank of $2^{O(k^2)}(\log d)^{k-2}$. We show that the rank of a depth-3 identity is at most $O(k^3 \log d)$. This bound is almost tight, since we also provide an identity of rank $\Omega(k \log d)$.

Our rank bound significantly improves (dependence on k exponentially reduced) the best known deterministic black-box identity tests for depth-3 circuits by Karnin and Shpilka (CCC 2008). Our techniques also shed light on the factorization pattern of nonzero depth-3 circuits: the rank of linear factors of a simple, minimal and nonzero depth-3 circuit (over any field) is at most $O(k^3 \log d)$.

The novel feature of this work is a new notion of maps between sets of linear forms, called *ideal matchings*, used to study depth-3 circuits. We prove interesting structural results about depth-3 identities using these techniques. We believe that these ideas may lead to the goal of a deterministic polynomial time identity test for these circuits.

1 Introduction

Polynomial identity testing (PIT) ranks as one of the most important open problems in the intersection of algebra and computer science. We are provided an arithmetic circuit that computes a polynomial $p(x_1, x_2, \cdots, x_n)$ over a field $\mathbb{F}$, and we wish to test if p is identically zero. In the black-box setting, the circuit is provided as a black-box and we are only allowed to evaluate the polynomial p at various domain points. The main goal is to devise a deterministic polynomial time algorithm for PIT. Kabanets and Impagliazzo [KI04] and Agrawal [Agr05] have shown connections between deterministic algorithms for identity testing and circuit lower bounds, emphasizing the importance of this problem.

The first randomized polynomial time PIT algorithm, which was a black-box algorithm, was given (independently) by Schwartz [Sch80] and Zippel [Zip79]. Randomized algorithms that use less randomness were given by Chen & Kao [CK00], Lewin & Vadhan [LV98], and

*A preliminary version was presented in the IEEE Conference on Computational Complexity in 2009 [SS09].

†Hausdorff Center for Mathematics, Bonn 53115, Germany. E-mail: ns@hcm.uni-bonn.de

‡IBM Almaden Research Center, San Jose - 95120, USA. E-mail: csesha@us.ibm.com

Agrawal & Biswas [AB03]. Klivans & Spielman [KS01] observed that even for depth-3 circuits for bounded top fanin, deterministic identity testing was open. Progress towards this was first made by Dvir & Shpilka [DS06], who gave a quasi-polynomial time algorithm, although with a doubly-exponential dependence on the top fanin. The problem was resolved by a polynomial time algorithm given by Kayal & Saxena [KS07], with a running time exponential in the top fanin. Some special cases of PIT for depth-4 circuits have been dealt in by Arvind & Mukhopadhyaya [AM07], Saxena [Sax08], and Volkovich & Shpilka [SV09]. Why is progress restricted to small depth circuits? Agrawal and Vinay [AV08] recently showed that an efficient black-box identity test for depth-4 circuits will actually give a quasi-polynomial black-box test for circuits of *all depths*.

For deterministic black-box testing, the first results were given by Karnin and Shpilka [KS08]. Based on results in [DS06], they gave an algorithm for depth-3 circuits having a quasi-polynomial running time (with a doubly-exponential dependence on the top fanin). One of the consequences of our result will be a significant improvement in the running time of their deterministic black-box tester.

This work focuses on depth-3 circuits. A structural study of depth-3 identities was initiated in [DS06] by defining the *rank* of *simple* and *minimal* identities. A depth-3 circuit C over a field $\mathbb{F}$ is:

$$C(x_1, \ldots, x_n) = \sum_{i=1}^{k} T_i$$

where T_i (*a multiplication term*) is a product of d_i linear functions $\ell_{i,j}$ over $\mathbb{F}$. For the purposes of identity testing, there is a simple procedure (Lemma 3.5 in [DS06]) that converts any such circuit C into an "equivalent" one where all $\ell_{i,j}$'s are linear *forms* (i.e. linear polynomials with a zero constant coefficient) and hence $d_1 = \cdots = d_k =: d$. We give details about this *homogenization* procedure towards the end in Section 2.7. Henceforth, we will just assume that C in homogeneous. Such a circuit is referred to as a $\Sigma\Pi\Sigma(k,d)$ circuit, where k is the *top fanin* of C and d is the *degree* of C. We call C a $\Sigma\Pi\Sigma$-*identity*, if C is an identically zero $\Sigma\Pi\Sigma$-circuit. This means that all coefficients of C (on expanding out) are zero. We give a few definitions from [DS06].

Definition 1. **[Simple Circuits]** *C is a* simple *circuit if there is no nonzero linear form dividing all the T_i's.*

[Minimal Circuits] *C is a* minimal *circuit if for every proper subset $S \subset [k]$, $\sum_{i \in S} T_i$ is nonzero.*

[Rank of a circuit] *The* rank *of the circuit, $rank(C)$, is defined as the rank of the linear forms $\ell_{i,j}$'s viewed as n-dimensional vectors over $\mathbb{F}$.*

Can all the forms $\ell_{i,j}$ be independent, or must there be relations between them? The rank can be interpreted as the minimum number of variables that are required to express C. There exists a linear transformation converting the n variables of the circuit into $rank(C)$ independent variables. A trivial bound on the rank (for any $\Sigma\Pi\Sigma$-circuit) is kd, since that is the total number of linear forms involved in C. The rank is a fundamental property of a $\Sigma\Pi\Sigma(k,d)$ circuit and it is crucial to understand how large this can be for identities. A substantially smaller rank bound than kd shows that identities do not have as many "degrees of freedom" as general circuits. This leads to deterministic *black-box* identity tests[1]. Furthermore, the techniques used to prove rank bounds show us structural properties of identities that may suggest directions to resolve PIT for $\Sigma\Pi\Sigma(k,d)$ circuits.

Dvir and Shplika [DS06] proved that the rank is bounded by $2^{O(k^2)}(\log d)^{k-2}$, and this bound is translated to a $\text{poly}(n)\exp(2^{O(k^2)}(\log d)^{k-1})$ time black-box identity tester by Karnin and Shpilka [KS08]. Note that when k is larger than $\log d$, these bounds are trivial.

Our present understanding of $\Sigma\Pi\Sigma(k,d)$ identities is very poor when k is larger than a constant. We present the first result in this direction.

Theorem 2 (Main Theorem)**.** *The rank of a simple and minimal* $\Sigma\Pi\Sigma(k,d)$ *identity (over any field) is* $O(k^3 \log d)$.

This gives an exponential improvement on the previously known dependence on k, and is strictly better than the previous rank bound for every $k > 3$. We also give a simple construction of identities with rank $\Omega(k \log d)$ in Section 3, showing that the above theorem is almost optimal. Note that this construction is over special fields, and does not work for all fields. We can interpret the main theorem as saying that any simple and minimal $\Sigma\Pi\Sigma(k,d)$ identity can be expressed using $O(k^3 \log d)$ independent variables. One of the most interesting features of this result is a novel technique developed to study depth-3 circuits. We introduce the concepts of *ideal matchings* and *ordered matchings*, that allow us to analyze the structure of depth-3 identities. These matchings are studied in detail to get the rank bound. Along the way we develop a theory of ideal matchings, viewing a matching as a fundamental map between two sets of linear forms.

Why are the simplicity and minimality restrictions required? Take the non-simple $\Sigma\Pi\Sigma(2,d)$ identity $(x_1x_2\cdots x_d) - (x_1x_2\cdots x_d)$. This has rank d. Similarly, we can take the non-minimal $\Sigma\Pi\Sigma(4,d+1)$ identity $(y_1y_2\cdots y_d)(x_1-x_1) + (z_1z_2\cdots z_d)(x_2-x_2)$ that has rank $(2d+2)$. In some sense, these restrictions only ignore identities that are composed of smaller identities.

Recent developments: Subsequent to the publication of the conference version of this paper, there have been some very exciting developments on this problem. The question of rank bounds for depth-3 identities where the underlying field is $\mathbb{Q}$ or $\mathbb{R}$ has been of special interest. Dvir & Shpilka [DS06] conjectured that the rank of these identities is at most $\text{poly}(k)$. Note that the lower bound constructions in this paper are over finite fields, and hence do not contradict this. Kayal and Saraf [KS09] made significant progress towards this by proving a rank bound of $2^{O(k \log k)}$ for this case. Observe that the rank bound is independent of d. With the results of Karnin & Shpilka [KS08], this implies the existence of polynomial time black-box PIT algorithms for bounded fanin depth-3 circuits (over $\mathbb{Q}$ or $\mathbb{R}$). Nonetheless, the dependence of k was still very large.

Rank bounds were further improved in a later work of the authors [SS10]. Their result finally resolves the conjecture of Dvir & Shpilka, by proving a rank bound of $O(k^2)$ for $\mathbb{Q}$ and $\mathbb{R}$. For general fields, a rank bound of $O(k^2 \log d)$ was proven, thereby improving on the main result of this paper. The main technique of ideal matchings, first introduced in this paper, was heavily used in the results of [SS10]. Combined with the insights of Kayal & Saraf [KS09], the authors were able to prove improved bounds.

[1]We usually do not get a *truly* polynomial time algorithm, i.e., one whose running time depends polynomially on n, d, and k.

This current paper still has techniques that are not subsumed by the developments of [SS10]. Namely, Theorem 4 gives a rank bound for the linear factors of a *nonzero* depth-3 circuit.

1.1 Consequences

Apart from being an interesting structural result about $\Sigma\Pi\Sigma$ identities, we can use the rank bound to get nice algorithmic results. Our rank bound immediately gives faster deterministic black-box identity testers for $\Sigma\Pi\Sigma(k,d)$ circuits. A direct application of Lemma 4.10 in [KS08] to our rank bound gives an exponential improvement in the dependence of k compared to previous black-box testers (that had a running time of $\mathrm{poly}(n)\exp(2^{O(k^2)}(\log d)^{k-1})$).

Theorem 3. *There is a deterministic black-box identity tester for $\Sigma\Pi\Sigma(k,d)$ circuits (over any field) that runs in poly$(n, d^{k^3 \log d})$ field operations.*

The above black-box tester is now much closer in complexity to the best *non* black-box tester known (*poly*(n, d^k) time by [KS07]). Our black-box tester runs in subexponential time when $k = d^{o(1)}$, while the previous black-box testers were unable to handle even $k = \log d$.

Although it is not immediate from Theorem 2, our technique also provides an interesting algebraic result about polynomials computed by simple, minimal, and nonzero $\Sigma\Pi\Sigma(k,d)$ circuits[2]. Consider such a circuit C that computes a polynomial $p(x_1, \cdots, x_n)$. Let us factorize p into $\prod_i q_i$, where each q_i is a nonconstant and irreducible polynomial. We denote by $L(p)$ the set of *linear factors* of p (that is, $q_i \in L(p)$ iff $q_i | p$ is linear).

Theorem 4. *If p is computed by a simple, minimal, nonzero $\Sigma\Pi\Sigma(k,d)$ circuit (over any field) then the rank of $L(p)$ is at most $O(k^3 \log d)$.*

This property has been recently used by Shpilka and Volkovich (Section 7.1 in [SV09]) to give an alternate black-box PIT algorithm of a similar time complexity.

1.2 Organization

Section 2 contains the proof of our main theorem. We give some preliminary notation in Section 2.1 before explaining an intuitive picture of our ideas (Section 2.2). We then explain our main tool of *ideal matchings* (Section 2.3) and prove some useful lemmas about them. We move to Section 2.4 where the concepts of *ordered matchings* and *simple parts of circuits* are introduced. We motivate these definitions and then prove some easy facts about them. We are now ready to tackle the problem of bounding the rank. We describe our proof in terms of an iterative procedure in Section 2.5. Everything is put together in Section 2.6 to bound the rank. A simple construction of identities with rank $\Omega(k \log d)$ is provided in Section 3.

Throughout the paper, various claims will be labeled as "facts". We believe that the proofs of these facts are not important. We suggest that the reader skip these proofs on initial readings.

[2] Here we can also consider circuits where the different terms in C have different degrees. The parameter d is then an upper bound on the degree of C.

2 Rank Bound

Our technique to bound the rank of $\Sigma\Pi\Sigma$ identities relies mainly on two notions - *form-ideals* and *matchings* by them - that occur naturally in studying a $\Sigma\Pi\Sigma$ circuit C. Using these tools we perform a surgery on the circuit C and extract out smaller circuits and smaller identities. Before explaining our basic idea we need to develop a small theory of *ideal matchings*. We need the definitions of the *gcd and simple* parts of a *subcircuit* [DS06].

We set down some preliminary definitions before giving an imprecise, yet intuitive explanation of our idea and an overall picture of how we bound the rank.

2.1 Preliminaries

We will denote the set $\{1,\ldots,n\}$ by $[n]$. For any set S, $\#S$ denotes the size of the set. We use $\mathcal{R}$ to denote the ring of polynomials $\mathbb{F}[x_1,\ldots,x_n]$, and $\mathbb{F}^* := \mathbb{F}\setminus\{0\}$. As mentioned earlier, a $\Sigma\Pi\Sigma(k,d)$ circuit is homogenous with top fanin k and degree d.

A *linear form* is a linear polynomial in $\mathcal{R}$. We will denote the set of all linear forms by $L(\mathcal{R})$:

$$L(\mathcal{R}) := \left\{ \sum_{i=1}^{n} a_i x_i \mid a_1,\ldots,a_n \in \mathbb{F} \right\}$$

There is a direct association between each linear form and a vector in $\mathbb{F}^n$. Hence, for any subset of forms in $L(\mathcal{R})$, we can naturally define a basis. Much of what we do shall deal with sets of linear forms, and various maps between them. A *list* L of linear forms is a multi-set of forms with an arbitrary order associated with them. The actual ordering is unimportant: we merely have it to distinguish between repeated forms in the list. One of the fundamental constructs we use are maps between lists, which could have many copies of the same form. The ordering allows us to define these maps unambiguously. All lists we consider will be finite.

Definition 5. [Multiplication term] *A* multiplication term f *is an expression in* $\mathcal{R}$ *given as (the product may have repeated* ℓ*'s):*

$$f := c \cdot \prod_{\ell\in S} \ell, \qquad \text{where } c \in \mathbb{F}^* \text{ and } S \text{ is a list of linear forms.}$$

The list of linear forms in f, $L(f)$, *is the list* S *of forms occurring in the product above.* $\#L(f)$ *is naturally called the* degree *of the multiplication term. For a list* S *of linear forms we define the* multiplication term of S, $M(S)$, *as* $\prod_{\ell\in S}\ell$ *or* 1 *if* $S=\phi$.

Definition 6. [Forms in a Circuit] *We will represent a* $\Sigma\Pi\Sigma(k,d)$ *circuit* C *as a sum of* k *multiplication terms of degree* d, $C=\sum_{i=1}^{k} T_i$. *The list of* linear forms occurring in C *is* $L(C) := \bigcup_{i\in[k]} L(T_i)$. *(Remark: for the purposes of this paper* T_i*'s are given in the "input" forming the circuit, and thus* $L(T_i)$ *is unambiguously defined.)*

Note that $L(C)$ *is a list of size exactly* kd. *The* rank of C, $rank(C)$, *is the number of linearly independent linear forms in* $L(C)$. *Clearly,* $0 \leqslant rank(C) \leqslant kd$.

2.2 Intuition

We set the scene, for proving the rank bound of a $\Sigma\Pi\Sigma(k,d)$ identity, by giving a combinatorial/graphical picture to keep in mind. Our circuits consist of k multiplication terms, and each term is a product of d linear forms. Think of there being k groups of d nodes, so each node corresponds to a form and each group represents a term[3]. We will incrementally construct a small basis for all these forms. This process will be described as some kind of a *coloring procedure*.

At any intermediate stage, we have a partial basis of forms. These are all linearly independent, and the corresponding nodes (we will use node and form interchangeably) are colored *red*. Forms not in the basis that are linear combinations of the basis forms (and are therefore in the span of the basis) are colored *green*. Once all the forms are colored, either green or red, all the red forms form a basis of *all forms*. The number of red forms is the rank of the circuit. When we have a partial basis, we carefully choose some uncolored forms and color them red (add them to the basis). As a result, some other forms get "automatically" colored green (they get added to the span). We "pay" only for the red forms, and we would like to get many green forms for "free". Note that we are trying to prove that the rank is $k^{O(1)} \log d$, when the total number of forms is kd. Roughly speaking, for every $k^{O(1)}$ forms we color red, we need to show that the number of green forms will *double*.

So far nothing ingenious has been done. Nonetheless, this image of coloring forms is very useful to get an intuitive and clear idea of how the proof works. The main challenge comes in choosing the right forms to color red. Once that is done, how do we keep an accurate count on the forms that get colored green? One of the main conceptual contributions of this work is the idea of *matchings*, which aid us in these tasks. Let us start from a trivial example. Suppose we have two terms that sum to zero, i.e. $T_1 + T_2 = 0$. By the unique factorization of polynomials, for every form $\ell \in T_1$, there is a unique form $m \in T_2$ such that $\ell = cm$, where $c \in \mathbb{F}^*$ (we will denote this by $\ell \sim m$). By associating the forms in T_1 to those in T_2, we create a *matching* between the forms in these two groups (or terms). This rather simple observation is the starting point for the construction of matchings.

Let us now move to $k = 3$, so we have a simple circuit $C \equiv T_1 + T_2 + T_3 = 0$. Therefore, there are no common factors in the terms. To get matchings, we will look at C modulo some forms in T_3. By looking at C modulo various forms in T_3, we reduce the fanin of C and get many matchings. Then we can deduce structural results about C. Similar ideas were used by Dvir and Shpilka [DS06] for their rank bound. Taking a form $q \in T_3$, we look at $C(\text{mod } q)$ which gives $T_1 + T_2 = 0(\text{mod } q)$. By unique factorization of polynomials modulo q, we get a *q-matching*. Suppose (ℓ, m) is an edge in this matching. In terms of the coloring procedure, this means that if q is colored and ℓ gets colored, then m must also be colored. At some intermediate stage of the coloring, let us choose an uncolored form $q \in T_3$. A key structural lemma that we will prove is that in the q-matching (between T_1 and T_2) any neighbor of a colored form *must be uncolored*. This crucially requires the simplicity of C. We will color q red, and thus all neighbors of the colored forms in $T_1 \cup T_2$ will be colored green. By coloring q red, we can double the number of colored forms. It is the various matchings (combined with the above property) that allow us to show an

[3]We have a different node for each appearance of a form.

exponential growth in the colored forms as forms in T_3 are colored red. By this process, we can color all forms by coloring at most $O(\log d)$ forms. Quite surprisingly, the above verbal argument can be formalized easily to prove that rank of a minimal, simple circuit with top fanin 3 is at most $(\log_2 d + 2)$. For this case of $k = 3$, the logarithmic rank bound was proven by Dvir & Shpilka [DS06], though they did not present the proof idea in this form. In particular, their rank bound grew to $(\log d)^2$ for $k = 4$.

The major difficulty arises when we try to push these arguments for higher values of k. In essence, the ideas are the same, but there are many technical and conceptual issues that arise. Let us go to $k = 4$. The first attempt is to take a form $q \in T_4$ and look at $C(\text{mod } q)$ as a fanin 3 circuit. Can we now simply apply the above argument recursively, and cover all the forms in $T_1 \cup T_2 \cup T_3$? No, the possible lack of simplicity in $C(\text{mod } q)$ blocks this simple idea. It may be the case that T_1, T_2 and T_3 have no common factors, but once we go modulo q, there could be many common factors! (For example, let $q = x_1$. Modulo q, the forms $x_1 + x_2$ and x_2 would be common factors.)

Instead of doing things recursively (both [DS06] and [KS07] used recursive arguments), we look at generating matchings iteratively. By performing a careful iterative analysis that keeps track of many relations between the linear forms we achieve a stronger bound for $k > 3$. We start with a form $\ell_1 \in T_1$, and look at $C(\text{mod } \ell_1)$. From $C(\text{mod } \ell_1)$, we remove all common factors. This common factor part we shall refer to as the *gcd* of $C(\text{mod } \ell_1)$, the removal of which leaves the *simple* part of $C(\text{mod } \ell_1)$. Now, we choose an appropriate form ℓ_2 from the simple part, and look at $C(\text{mod } \ell_1, \ell_2)$. We now choose an ℓ_3 and so on and so forth. For each ℓ that we choose, we decrease the top fanin by at least 1, so we will end up with a matching modulo the *ideal* $(\ell_1, \ell_2, ..., \ell_r)$, where $r \leq (k-2)$. We call these special ideals *form ideals* (as they are generated by forms), and the main structures that we find are matchings modulo form ideals. The coloring procedure will color the forms in the form ideal red. It is not as simple as the case of $k = 3$, since, at the very least, we have to deal with the simple and gcd parts. Many other problems arise, but we will explain them as and when we they arise. For now, it suffices to understand the overall picture and the concept of matchings among the linear forms in C.

We now start by setting some notation and giving some key definitions.

2.3 Ideal Matchings

In this subsection, we provide the necessary definitions and prove some basic facts about these matchings.

First, we discuss *similarity* and *form-ideals.*

Definition 7. *We give several definitions :*

- **[Ideal]** *Given a set of polynomials* $f_1, \ldots, f_r$*, the* ideal *generated by them is the set of polynomials* $\{\sum_i f_i g_i | g_i \in \mathcal{R}\}$.

- **[Similar forms]** *For any two polynomials* $f, g \in \mathcal{R}$*,* f *is* similar to g *if there exists* $c \in \mathbb{F}^*$ *such that* $f = cg$*. We say* f is similar to g modulo I*, for some ideal* I *of* $\mathcal{R}$*, if there is a* $c \in \mathbb{F}^*$ *such that* $f = cg$ *(mod* I*). We also denote this by* $f \sim g$ *(mod* I*) or say that* f *is* I*-similar to* g.

- **[Similar lists]** *Let* $S_1 = (a_1, \ldots, a_d)$ *and* $S_2 = (b_1, \ldots, b_d)$ *be two lists of linear forms with a bijection* π *between them.* S_1 *and* S_2 *are called* similar under π *if for all* $i \in [d]$, a_i *is similar to* $\pi(a_i)$. *Any two lists of linear forms are called* similar *if there exists such a* π. *Empty lists of linear forms are similar vacuously.*

  For any $\ell \in L(\mathcal{R})$ *we define the* list of forms in S_1 similar to ℓ *as the following list (unique upto ordering):*

$$simi(\ell, S_1) := (a \in S_1 \mid a \text{ is similar to } \ell)$$

  We call S_1, S_2 coprime lists *if* $\forall \ell \in S_1$, $\#simi(\ell, S_2) = 0$.

- **[Form-ideal]** *A* form-ideal I *is the ideal of* $\mathcal{R}$ *generated by some nonempty* $S \subseteq L(\mathcal{R})$. *Note that if* $I = \{0\}$ *then* $a \equiv b (mod\ I)$ *simply means that* $a = b$ *absolutely.*

- **[Span** $sp(S)$**]** *For any* $S \subseteq L(\mathcal{R})$, *we let* $sp(S) \subseteq L(\mathcal{R})$ *be the* linear span *of the linear forms in* S *over the field* $\mathbb{F}$.

- **[Orthogonal sets of forms]** *Let* $S_1, \ldots, S_m$ *be sets of linear forms for* $m \geq 2$. *We call* $S_1, \ldots, S_m$ orthogonal *if for all* $m' \in [m-1]$:

$$sp(\bigcup_{j \in [m']} S_j) \cap sp(S_{m'+1}) = \{0\}$$

  Let form-ideals $I_1, \ldots, I_m$ *be generated by sets* $S_1, \ldots, S_m$ *respectively. If* $S_1, \ldots, S_m$ *are orthogonal, then the form-ideals* $I_1, \ldots, I_m$ *are also orthogonal.*

We give a few simple facts based on these definitions. It will be helpful to have these explicitly stated.

Fact 8. *Let* U, V *be lists of linear forms and* I *be a form-ideal. If* U, V *are similar then their sublists* $U' := (\ell \in U \mid \ell \in sp(I))$ *and* $V' := (\ell \in V \mid \ell \in sp(I))$ *are also similar.*

Proof. If U, V are similar then for some $c \in \mathbb{F}^*$, $M(V) = cM(U)$. This implies:

$$M(V') \cdot M(V \setminus V') = cM(U') \cdot M(U \setminus U')$$

Since elements of $U \setminus U'$ are not in $sp(I)$, for any $\ell \in V'$, ℓ does not divide $M(U \setminus U')$. In other words $M(V')$ divides $M(U')$, and vice versa. Thus, $M(U'), M(V')$ are similar and hence by unique factorization in $\mathcal{R}$, lists U' and V' are similar. □

Fact 9. *Let* I_1, I_2 *be two orthogonal form-ideals of* $\mathcal{R}$ *and let* D *be a* $\Sigma\Pi\Sigma(k,d)$ *circuit such that* $L(D)$ *has all its linear forms in* $sp(I_1)$. *If* $D \equiv 0$ $(mod\ I_2)$ *then* $D = 0$.

Proof. As I_1, I_2 are orthogonal we can assume I_1 to be $\{\ell_1, \ldots, \ell_m\}$ and I_2 to be $\{\ell'_1, \ldots, \ell'_{m'}\}$ where the ordered set $V := \{\ell_1, \ldots, \ell_m, \ell'_1, \ldots, \ell'_{m'}\}$ has $(m + m')$ linearly independent linear forms. There exists an invertible linear transformation τ on $sp(\{x_1, \ldots, x_n\})$ that maps the elements of V bijectively, in that order, to $x_1, \ldots, x_{m+m'}$. On applying τ to the equation $D \equiv 0$ (mod I_2) we get:

$$\tau(D) \equiv 0 \ (\text{mod } x_{m+1}, \ldots, x_{m+m'}), \quad \text{where } \tau(D) \in \mathbb{F}[x_1, \ldots, x_m].$$

Obviously, this means that $\tau(D) = 0$ which by the invertibility of τ implies $D = 0$. □

We now come to the most important definition of this section. We motivated the notion of *ideal matchings* in the intuition section. Thinking of two lists of linear forms as two sets of vertices, a matching between them signifies some linear relationship between the forms modulo a form-ideal. Essentially, we have a matching by ideal I between two lists U and V if U is similar to V modulo I. We will be more precise in the following definition.

Definition 10. [Ideal matchings] *Let U, V be lists of linear forms and I be a form-ideal. An* ideal matching π between U, V by I *is a bijection π between lists U, V such that: $\forall \ell \in U$, $\exists \ell' \in V$, such that $\pi(\ell) = \ell'$ and $\ell' = c\ell + v$ for some $c \in \mathbb{F}^*$ and $v \in sp(I)$. We also call this an I-*matching between U and V. *The matching π is called* trivial *if U, V are similar.*

Two sublists $U' \subseteq U$ and $V' \subseteq V$ are similar under π *if: $\forall \ell \in U'$, $\pi(\ell)$ is similar to ℓ and $\pi(\ell) \in V'$.*

An I-matching is orthogonal *to an I'-matching (both between U and V) if the form-ideals I and I' are orthogonal.*

An I*-matching π between multiplication terms f, g* is the one that matches $L(f), L(g)$. (For convenience, we will just say "matching" instead of "ideal matching".) Note that since π is a bijection and $c \neq 0$, π^{-1} can be viewed as a matching between V, U by I.

The following is an easy fact about matchings.

Fact 11. *Let π be a matching between lists of linear forms U, V by I and let $U' \subseteq U$, $V' \subseteq V$ be similar sublists. Then there exists a matching π' between U, V by I such that U' and V' are similar under π'.*

Proof. We begin with a matching π such that U' is not similar V' under π. We will convert this into a matching $\widetilde{\pi}$. The number of forms in U' matched to similar forms in V' increases. By repeating the process, we will eventually get a matching π' that matches U' to V'.

Let $\ell' \in U'$ be such that $\pi(\ell') = d'\ell' + v'$ (for some $d' \in \mathbb{F}^*$ and $v' \in sp(I)$) is not in V' or is not similar to ℓ'. Since U' and V' are similar, $\#simi(\ell', U') = \#simi(\ell', V')$. So it cannot be the case that all forms similar to ℓ' in V' are matched to similar forms in U'. Hence, there is a form equal to $\alpha\ell'$ in V' (for some $\alpha \in \mathbb{F}^*$) with the following property. The matching π maps some $\ell \in U$ to $\alpha\ell'$ in V' such that *either* $\ell \notin U'$ or ℓ is not similar to $\alpha\ell'$. There exists some $d \in \mathbb{F}^*$ and $v \in sp(I)$ such that $\pi(\ell) = d\ell + v = \alpha\ell'$.

Now we define a new matching $\widetilde{\pi}$ by flipping the images of ℓ and ℓ' under π, i.e., define $\widetilde{\pi}$ to be the same as π on $U \setminus \{\ell, \ell'\}$ and: $\widetilde{\pi}(\ell) \overset{V}{:=} \pi(\ell')$ and $\widetilde{\pi}(\ell') \overset{V}{:=} \pi(\ell)$. Note that $\widetilde{\pi}$ inherits the bijection property from π and it is an I-matching because: $\widetilde{\pi}(\ell') = \alpha\ell'$ for $\alpha \in \mathbb{F}^*$ and more importantly,

$$\widetilde{\pi}(\ell) = \pi(\ell') = d'\ell' + v' = d'\left(\frac{d\ell + v}{\alpha}\right) + v' = \left(\frac{dd'}{\alpha}\right)\ell + \left(\frac{d'v}{\alpha} + v'\right)$$

The form $(\frac{d'v}{\alpha} + v')$ is clearly in $sp(I)$. Thus, we have obtained now a matching $\widetilde{\pi}$ between U, V by I such that the $\ell' \in U'$ is similar to $\widetilde{\pi}(\ell') \in V'$.

Since $\ell \notin U'$ or ℓ is not similar to ℓ', the number of forms in U' that are matched to similar forms in V' increases. This completes the proof. □

We are ready to present the most important lemma of this section. The following lemma shows that there cannot be too many matchings between two nonsimilar lists of linear forms. It is at the heart of our rank bound proof and the reason for the logarithmic dependence of the rank on the degree. Similar ideas were used by Dvir & Shpilka to prove lower bounds for 2-query locally decodable codes (Corollary 2.9 of [DS06]).

Lemma 12. *Let U, V be lists of linear forms each of size $d > 0$ and $I_1, \ldots, I_r$ be orthogonal form-ideals such that for all $i \in [r]$, there is a matching π_i between U, V by I_i. If $r > (\log_2 d + 2)$ then U, V are similar lists.*

Before giving the proof, let us first put it in the context of our overall approach. In the sketch that we gave for $k = 3$, at each step, we were generating orthogonal matchings between two terms. For each orthogonal matching we got, we colored one linear form red (added one form to our basis) and *doubled* the number of green forms (doubled the number of forms in the circuit that are in the span of the basis). This showed that there is a logarithmic-sized basis for all $L(C)$. If we take the contrapositive of this, we get that there *cannot* be too many orthogonal matchings between two (nonsimilar) lists of forms. For dealing with larger k, it will be convenient to state things in this way.

Proof. Let $U_1 \subseteq U$ be a sublist such that: there exists a sublist $V_1 \subseteq V$ similar to U_1 for which $U' := U \setminus U_1$ and $V' := V \setminus V_1$ are coprime lists. Let U', V' be of size d'. If $d' = 0$ then U, V are indeed similar and we are done already. So assume that $d' > 0$. By the hypothesis and Fact 11, for all $i \in [r]$, there exists a matching π'_i between U, V by I_i such that: U_1, V_1 are similar under π'_i *and* π'_i is a matching between U', V' by I_i. Our subsequent argument will only consider the latter property of π'_i for all $i \in [r]$.

Intuitively, it is best to think of the various π'_is as bipartite matchings. The graph $G = (U', V', E)$ has vertices labelled with the respective form. For each π'_i and each $\ell \in U'$, we add an (undirected) edge tagged with I_i between ℓ and $\pi'_i(\ell)$. There may be many tagged edges between a pair of vertices[4]. We call $\pi'_i(\ell)$ the I_i-neighbor of ℓ (and vice versa, since the edges are undirected). Abusing notation, we use *vertex* to refer to a form in $U' \cup V'$. We will denote $\bigcup_{j \leq i} I_i$ by J_i.

We will now show that there cannot be more than $(\log_2 d + 2)$ such perfect matchings in G. The proof is done by following an iterative process that has r phases, one for each I_i. This is essentially the coloring process that we described earlier. We maintain a partial basis for the forms in $U' \cup V'$ which will be updated iteratively. This basis is kept in the set B. Note that although we only want to span $U' \cup V'$, we will use forms in the various I_i's for spanning.

We start with empty B and initialize by adding some $\ell \in U'$ to B. In the ith round, we will add all forms in I_i to B. All forms of $U' \cup V'$ in $sp(\{\ell\} \cup J_i)$ are now spanned. We then proceed to the next round. To introduce some colorful terminology, a *green* vertex is one that is in the set $sp(B)$ (a form in $(U' \cup V') \cap sp(B)$). Here is a nice fact : at the end of a round, the number of green vertices in U' and V' are the same. Why? All forms of I_1 are in B, at the end of any round. Let vertex v be green, so $v \in sp(B)$. The I_1-neighbor of v is a linear combination of v and I_1. Therefore, this neighbor of v is in $sp(B)$ and is colored green. This shows that the number of green vertices in U' is equal to the number of those in V'.

[4]It can be shown, using the orthogonality of the I_i's, that an edge can have at most *two* distinct tags.

Let $i_0 \in [r]$ be the least index such that $\{\ell\}$, $I_1, \dots, I_{i_0}$ are not orthogonal, if it does not exist then set $i_0 := r + 1$. We have the following claim.

Claim 13. *The sets $\{\ell\}$, $I_1, \dots, I_{i_0-1}$ are orthogonal and the sets:*

$$\{\ell\} \cup J_{i_0}, I_{i_0+1}, \dots, I_r$$

are orthogonal.

Proof of Claim 13. The ideals $\{\ell\}$, $I_1, \dots, I_{i_0-1}$ are orthogonal by the minimality of i_0.

As $I_1, \dots, I_{i_0}$ are orthogonal and $\{\ell\}$, $I_1, \dots, I_{i_0}$ are not orthogonal, we deduce that $\{\ell\} \in sp(J_{i_0})$. Thus, $\{\ell\} \cup sp(J_{i_0}) = sp(J_{i_0})$ which is orthogonal to the sets $I_{i_0+1}, \dots, I_r$ by the orthogonality of $I_1, \dots, I_r$. □

We now show that the green vertices at least double in $(r-2)$ rounds.

Claim 14. *For $i \notin \{1, i_0\}$, the number of green vertices doubles in the ith round.*

Proof of Claim 14. Let ℓ' be a green vertex, say in U', at the end of the $(i-1)$th round ($B = \{\ell\} \cup J_{i-1}$). Consider the I_i-neighbor of ℓ'. This is in V' and is equal to $(c\ell' + v)$ where $c \in \mathbb{F}^*$ and v is a *nonzero* element in $sp(I_i)$ (this is because U', V' are coprime). If this neighbor is green, then v would be a linear combination of two green forms, implying $v \in sp(B)$. But by Claim 13, I_i is orthogonal to B, implying $v \in sp(B) \cap sp(I_i) = \{0\}$ which is a contradiction. Therefore, the I_i-neighbor of any green vertex is *not* green. Because we have an I_i-matching, the number of such neighbors is the number of green vertices. On adding I_i to B, all these neighbors will become green. This completes the proof. □

We started off with one green vertex ℓ, and U', V' each of size d'. This doubling can happen at most $\log_2 d'$ times, implying that $(r-2) \le \log_2 d'$.

□

Remark: The bound of $r = \log_2 d + 2$ is achievable by lists of linear forms inspired by Section 3. Fix an odd s and define:

$$U := \{(b_1x_1 + \cdots + b_{s-1}x_{s-1} + x_s) \mid b_1, \dots, b_{s-1} \in \{0,1\} \text{ s.t. } b_1 + \cdots + b_{s-1} \text{ is even}\}$$

$$V := \{(b_1x_1 + \cdots + b_{s-1}x_{s-1} + x_s) \mid b_1, \dots, b_{s-1} \in \{0,1\} \text{ s.t. } b_1 + \cdots + b_{s-1} \text{ is odd}\}$$

It is easy to see that over $\mathbb{F}_2$, $\#U = \#V = 2^{s-2}$ and for all $i \in [s-1]$, there is a matching between U, V by (x_i), furthermore, there is a matching by $(x_1 + \cdots + x_{s-1} + 2x_s)$. Thus there are $(\log_2 |U| + 2)$ many orthogonal matchings between these nonsimilar U, V; showing that Lemma 12 is tight.

2.4 Ordered Matchings and Simple Parts of Circuits

Before we delve into the definitions and proofs, let us motivate them by an intuitive explanation.

2.4.1 Intuition

Our main goal is to deal with the case $k > 3$. The overall picture is still the same. We keep updating a partial basis S for $L(C)$. This process goes through various *rounds*, each round consisting of *iterations*. At the end of each round, we obtain a form-ideal I that is orthogonal to S. In the first iteration of a round, we start by choosing a form ℓ_1 in $L(T_1)$ that is not in $sp(S)$, and adding it to I. We look at the circuit $C(\text{mod } \ell_1)$ in the next iteration. Every T_i that ℓ_1 divides become zero, and the remaining terms "survive". The top fanin has decreased by at least 1, so we have a smaller circuit to deal with. We would like to proceed to the next iteration with this circuit. The major obstacle to proceeding is that our circuit is not simple any more, because there *can* be common factors among multiplication terms modulo ℓ_1. Note how this seems to be a difficulty, since it appears that our matchings will not give us a proper handle on these common factors. Suppose that form v is now a common factor. That means, in every surviving term, there is a form that is v modulo ℓ_1. So these forms can be ℓ_1-matched to each other! We have converted the obstacle into some kind of a partial matching, which we can hopefully exploit.

Let us go back to $C(\text{mod } \ell_1)$. Let us remove all common factors from this circuit. This stripped down identity circuit is the *simple* part, denoted by $sim(C \text{ mod } \ell_1)$. Note that every surviving term is present in this circuit. The removed portion, called the *gcd* part, is referred to as $gcd(C \text{ mod } \ell_1)$. By the above observation, the *gcd* part has ℓ_1-matchings. All the forms in the *gcd* part are *not* similar to ℓ_1 and can be matched mod ℓ_1. (This is because if a form is similar to ℓ_1, then the term containing that form vanishes.) Having (somewhat) dealt with $gcd(C \text{ mod } \ell_1)$ by finding I-matchings, let us focus on the smaller circuit $sim(C \text{ mod } \ell_1)$

We try to find an $\ell_2 \in L(sim(C \text{ mod } \ell_1))$ that is not in $sp(S \cup \{\ell_1\})$. Suppose we can find such an ℓ_2. Then, we add ℓ_2 to I and proceed to the next iteration. In a given iteration, we start with a form-ideal I, and a circuit $sim(C \text{ mod } I)$. We find a form $\ell \in L(sim(C \text{ mod } I)) \backslash sp(S \cup I)$. We add ℓ to I (for convenience, let us set $I' = I \cup \{\ell\}$) and look at the $C(\text{mod } I')$. We now have new terms in the *gcd* part, which we can match through I'-matchings. As we observed earlier, all the terms that have forms in I' are removed, so the terms we match here are all nonzero modulo I'. We remove the *gcd* part to get $sim(C \text{ mod } I')$, and go to the next iteration with I' as the new I. When does this stop? If there is no ℓ in $L(sim(C \text{ mod } I)) \backslash sp(S \cup I)$, then this means that all of $L(sim(C \text{ mod } I))$ is in our current span. So we happily stop here with all the matchings obtained from the *gcd* parts. Also, if the fan-in reaches 2, then we can imagine that the whole circuit is itself in the *gcd* portion. At each iteration, the fan-in goes down by at least one, so we can have at most $(k-2)$ iterations in a round, hence the I in any round is generated by at most $(k-2)$ forms. When we finish a round obtaining an ideal I, there are some multiplication terms in C that are nonzero modulo I *after* the *gcd* parts in the various iterations are removed from these terms. These we shall refer to as constituting the *blocking subset* of $[k]$, for that round.

The way we prove rank bounds is by invoking Lemma 12. Each round constructs a new orthogonal form ideal. At the end of a round, we have a set S, which is a partial basis. If S does not cover all of $L(C)$, then we use the above process (of iterations) to generate a form-ideal I orthogonal to S. We then add I to S, and repeat this process until all of $L(C)$ is covered. Essentially, we argue that if this process takes too many rounds, there exist two terms that have too many orthogonal matchings between them. This violates Lemma 12. Hence, we are able to cover $L(C)$ with a relatively small set of forms, leading to the rank bound.

2.4.2 Definitions

We start with looking at the particular kind of matchings that we get. Take two terms T_a and T_b that survive a round, where we find the form-ideal I generated by $\{\ell_1, \ell_2, \cdots, \ell_r\}$. At the end of the first iteration, we add ℓ_1 to I. No form in $L(T_a) \cup L(T_b)$ can be $0 (\text{mod } \ell_1)$. We match some forms in T_a to T_b via ℓ_1-matchings. They are removed, and then we proceed to the next iteration. We now match some forms via $sp(\{\ell_1, \ell_2\})$ matchings and none of these forms are in this span. So in each iteration, the forms that are matched (and then removed) are non-zero mod the partial I obtained by that iteration. We formalize this as an *ordered matching*.

Definition 15. [Ordered matching] *Let U, V be lists of linear forms and I be a form-ideal generated by an ordered set $\{v_1, \dots, v_i\}$ of i linearly independent linear forms. A matching π between U, V by I is called an* ordered I-matching *if:*

Let v_0 be zero. For all $\ell \in U$, $\pi(\ell) = (c\ell + w)$ where $c \in \mathbb{F}^*$, and $w \in sp(v_0, \dots, v_j)$ for some j satisfying $\ell \notin sp(v_0, \dots, v_j)$.

First, a small explanation about ordered matchings. We will stick to the notation in Definition 15. For convenience, let $sp_j := sp(v_0, \cdots, v_j)$ and $\text{diff}_j := sp_j \backslash sp_{j-1}$. We can partition all the ordered matching edges into groups corresponding to $\text{diff}_1, \text{diff}_2, \cdots$. In the group of edges corresponding to diff_i, all forms do not belong to sp_j, for any $j < i$. We can think of this matching being constructed group by group: first, the diff_1 edges are inserted, then the diff_2 edges, so on and so forth. There is an order in which this matching is created.

We add the zero element v_0, just to deal with similar forms in U and V. Note that the inverse bijection π^{-1} is also an ordered matching between V, U by I. It is also easy to see that if π_1 and π_2 are ordered matchings between lists U_1, V_1 and lists U_2, V_2 respectively by the same ordered form-ideal I then their *disjoint union*, $\pi_1 \sqcup \pi_2$, is an ordered matching between lists $U_1 \cup U_2$, $V_1 \cup V_2$ by I.

Let $\pi(\ell) = d\ell + w$, where $w \in sp_j$ but $\ell \notin sp_j$. We show that the constant d is unique. If there were two such different constants, say d and d', then both $(\pi(\ell) - d\ell)$ and $(\pi(\ell) - d'\ell)$ would be in sp_j implying that $(d - d')\ell \in sp_j$. That contradicts $\ell \notin sp_j$. Thus for a fixed ℓ and an ordered matching π, d is uniquely determined. Keeping the notation above, we define:

Definition 16. [Scaling factor] *Let π be on ordered matching between U and V. For each $\ell \in U$, let d_ℓ be the unique constant such that $\pi(\ell) = d_\ell \ell + w$, where $w \in sp_j$ but $\ell \notin sp_j$. The* scaling factor *of π, denoted by $sc(\pi)$, is set to be $\prod_{\ell \in U} d_\ell$. For empty U, $sc(\pi)$ is set to be* 1.

Definition 17. [Subcircuits and regular circuits] *For non-empty $Q \subseteq [k]$, the* subcircuit *C_Q of a $\Sigma\Pi\Sigma(k,d)$ circuit C is the sum $\sum_{j \in Q} T_j$. For a form-ideal I we call C_Q* regular *mod I if $\forall q \in Q$, $T_q \not\equiv 0$ (mod I). We will denote the constant factor in the multiplication term T_q by $\alpha_q \in \mathbb{F}^*$, thus $T_q = \alpha_q M(L(T_q))$.*

We are now ready to define the *gcd* and *sim* parts of a subcircuit. Although the ideas are quite simple and intuitive, we have to be careful in dealing with constant factors. Much of this notation has been introduced for rigorous definitions. Take a subcircuit C_Q that is a regular identity modulo I. A maximal list of forms, say U, that divides T_q, for all $q \in Q$, is called the *gcd* of $C_Q(\text{mod } I)$. In every T_q, there is a list U_q of forms that are I-similar to U. Therefore, we have I-matchings between U and U_q. This is the *gcd data of C_Q modulo I*, and represents the various matchings that we will later exploit. If we remove U_q from each T_q, then (by accounting for constants carefully) we get a simple (mod I) identity, the *sim* part of $C_Q(\text{mod } I)$. We formalize this below.

Fact 18. *Let C_Q be regular modulo I. Let q_1 be the smallest index*[5] *in Q. Let U be a maximal sublist of $L(T_{q_1})$ such that $M(U)$ divides T_q modulo I for all $q \in Q$.*

- *$M(U)$ is a gcd of the polynomials $\{T_q \mid q \in Q\}$ modulo the ideal I.*
- *For all $q \in Q$, there exists a sublist $U_q \subseteq L(T_q)$ such that there exists an ordered I-matching π_q between U, U_q.*

Proof. This is a fairly direct consequence of unique factorization in $\mathcal{R}/I$. Since I is a form-ideal, under an appropriate linear transformation, we can assume that I is generated by the independent linear forms $x_1, x_2, \cdots, x_r$. So $\mathcal{R}/I$ is exactly the polynomial ring $\mathbb{F}[x_{r+1}, \cdots, x_n]$, and therefore enjoys the unique factorization property. Hence, any polynomial in $\mathcal{R}$ that is nonconstant modulo I uniquely factors modulo the ideal I into polynomials irreducible modulo I.

Each of the terms T_q is non-zero modulo I and is just a product of linear forms (which are irreducible factors) modulo I. So, $M(U)$ is a gcd of the T_q's. There must be some sublist $U_q \subseteq L(T_q)$ and a $c_q \in \mathbb{F}^*$ such that $M(U_q) \equiv c_q \cdot M(U) \pmod{I}$. By unique factorization, we have an I-matching between every U_q and U. No form in U_q can be zero modulo I, so this is an ordered matching. □

Given C_Q and I, there are many possibilities to choose the lists U and $\{U_q \mid q \in Q\}$. But they are all uniquely determined upto similarity modulo the ideal I and that will be good enough for our purposes. So we choose them in some way, say the lexicographically smallest one (unless specified otherwise). Also, by the definition of scaling factor of a matching, π_q satisfies: $\forall q \in Q$, $M(U_q) \equiv sc(\pi_q) \cdot M(U) \pmod{I}$. Using the gcd of $C_Q(\text{mod } I)$, we can extract out a smaller circuit from C_Q that we call the simple part.

Definition 19. [gcd and sim parts] *(We use the definition of α_q from Definition 17 and those of U, U_q, π_q from Fact 18.) The* gcd data of C_Q modulo I *is the following set of $\#Q$ matchings:*

$$\overline{gcd}(C_Q \ mod \ I) := \{(\pi_q, U, U_q) \mid q \in Q\} \tag{1}$$

The gcd of $C_Q(\text{mod } I)$ *is just $gcd(C_Q \ mod \ I) := M(U)$. The* simple part of C_Q mod I *is the circuit:*

$$sim(C_Q \ mod \ I) := \sum_{q \in Q} sc(\pi_q)\alpha_q \cdot M(L(T_q) \setminus U_q)$$

[5]It is not necessary to choose the smallest index. Any fixed index will do.

Before a round, we have a partial basis S. At the end of a round, we produce a form-ideal I that is orthogonal to S. We call this a *useful ideal*. Let $Q \subset [k]$ be such that $\forall q \in Q$, T_q survives (is non-zero) modulo I. This is called the *blocking subset*. For each such q, there are a list of forms $V_q \subset L(T_q)$ that are mutually matched via ordered I-matchings (these are really a collection of *gcd* datas). This is called the *matching data*. Even after we remove V_q from each term T_q (carefully accounting for constants, as explained above), we still have an identity modulo I. All forms of this identity are in $sp(S \cup I) \setminus sp(I)$, since we assume that the round has ended. These forms are not in $sp(I)$ because the corresponding terms survive modulo I. Furthermore, we can ensure (as shown later) that all V_q's can be made disjoint to $sp(S \cup I) \setminus sp(I)$. Therefore, this round partitions the $L(T_q)$ into V_q and $L(T_q) \cap (sp(S \cup I) \setminus sp(I))$ (for all $q \in Q$). These end-of-a-round properties motivate the following definition.

Definition 20. [Useful ideals, blocking subsets, and matching data] *Let $C = \sum_{j \le k} T_j$, $T_j = \alpha_j\ M(L(T_j))$. The set $S \subseteq L(\mathcal{R})$ and I is an ordered form-ideal orthogonal to S. We call I* useful in C wrt S *if $\exists Q \subset [k]$, $1 < \#Q < k$ with the following properties :*

For all $q \in Q$, let V_q be $L(T_q) \setminus (sp(S \cup I) \setminus sp(I))$. (Therefore, $L(T_q) \setminus V_q \subset sp(S \cup I) \setminus sp(I)$.)

- *There exists a list of linear forms V such that for all $q \in Q$, there is an ordered I-matching τ_q between V, V_q.*
- *The circuit $\sum_{q \in Q} sc(\tau_q)\alpha_q \cdot M(L(T_q) \setminus V_q)$ is a regular identity modulo I.*

Such a Q is a blocking subset *of C, S, I. By* matching data *of C, S, I, Q we refer to the set:*

$$mdata(C, S, I, Q) := \{(\tau_q, V, V_q) \mid q \in Q\}$$

We will call $mdata(C, S, I, Q)$ trivial *if the lists V_q, $q \in Q$, are all mutually similar.*

From the matching data, we will exploit the fact that for each pair $q_1, q_2 \in Q$, there is an ordered I-matching between V_{q_1} and V_{q_2}. Nonetheless, for convenience, we will represent these $\#Q$ matchings via V.

2.4.3 Basic facts

In this subsection, we prove some basic facts about ordered matchings, scaling factors and *gcd* and *sim* parts of a circuit. These facts are not difficult to prove, but it will be helpful later to have them. Again, we remind the reader that these proofs do not help in understanding the main result.

We begin by a simple statement of how extending an ideal preserves the ordered nature of matchings.

Fact 21. *Suppose an ideal I is generated by an ordered set S of linear forms. Let $S' = S \cup \{\ell\}$ be an ordered set of linear forms with ℓ at the end of the ordering. If π is an ordered I-matching (between lists U and V), then it is an ordered I'-matching (where I' is generated by S').*

Proof. Let $\pi(\ell) = c\ell + w$. We use the earlier notation of sp_j (resp. sp'_j) for the span of the first j forms in the set S (resp. S'). Since π is an ordered I-matching, for some index j, $w \in sp_j$ and $\ell \notin sp_j$. Note that I' preserves the ordering of I. So for j such that sp_j is defined, $sp_j = sp'_j$. Hence, π is also an ordered I'-matching. $\square$

From the definition of scaling factor, it is easy to see that ordered matchings have inverses and also a union.

Fact 22. *Let π_1 and π_2 be ordered I-matchings between lists U_1, V_1 and lists U_2, V_2 respectively. Then π_1^{-1} is an ordered I-matching such that $sc(\pi_1^{-1}) = sc(\pi_1)^{-1}$. The disjoint union $\pi_1 \sqcup \pi_2$ is also an ordered I-matching such that $sc(\pi_1 \sqcup \pi_2) = sc(\pi_1) \cdot sc(\pi_2)$.*

Proof. For some $\ell \in U$, let $\ell' := \pi_1(\ell) = c\ell + w$. Since I is an ordered matching, for some j, $w \in sp_j$ and $\ell \notin sp_j$. This implies that ℓ' is also not in sp_j. We have $\pi_1^{-1}(\ell') = \ell = \ell'/c - w/c$. Since $w/c \in sp_j$ and $\ell' \notin sp_j$, π_1^{-1} is an ordered I-matching. The contribution to the scaling factor in π_1^{-1} is $1/c$, which is the reciprocal of the corresponding contribution to $sc(\pi_1)$. This shows $sc(\pi_1^{-1}) = sc(\pi_1)^{-1}$.

Consider the disjoint union $\pi' = \pi_1 \sqcup \pi_2$. For $\ell \in U_1$, $\pi'(\ell) = \pi_1(\ell) = c\ell + w$. Again, for some j, $w \in sp_j$ and $\ell \notin sp_j$. We have a similar statement for $\ell \in U_2$. So π' is an ordered matching. In π', we have all the edges in the union of π_1 and π_2. Hence the scaling factor $sc(\pi_1 \sqcup \pi_2)$ is just the product $sc(\pi_1) \cdot sc(\pi_2)$. $\square$

The following fact shows that ordered matchings can also be composed.

Fact 23. *Let π_1 and π_2 be ordered matchings between U_1, V and V, U_2 respectively by the same ordered form-ideal $I = \{v_1, \ldots, v_i\}$. Then the naturally defined composite matching $\pi_2\pi_1$ is also an ordered matching between U_1, U_2 by I. Furthermore, $sc(\pi_2\pi_1) = sc(\pi_1) \cdot sc(\pi_2)$.*

Proof. Consider a linear form $\ell \in U_1$. There exists $c_1 \in \mathbb{F}^*$ and $\alpha_1 \in sp_{j_1}, \ell \notin sp_{j_1}$ such that $\pi_1(\ell) = c_1\ell + \alpha_1$. Also, there exists $c_2 \in \mathbb{F}^*$ and $\alpha_2 \in sp_{j_2}$, $\pi_1(\ell) \notin sp_{j_2}$ such that $\pi_2(\pi_1(\ell)) = c_2(c_1\ell + \alpha_1) + \alpha_2$. Let $j = \max\{j_1, j_2\}$. Obviously, $(c_2\alpha_1 + \alpha_2) \in sp_j$. If $\ell \in sp_j$ then as $\ell \notin sp_{j_1}$ we deduce that $j = j_2 > j_1$, thus $\ell \in sp_{j_2}$, implying $\pi_1(\ell) = c_1\ell + \alpha_1 \in sp_{j_2}$, which is a contradiction. Therefore, $\ell \notin sp_j$. This proves that the composite bijection $\pi_2\pi_1$ is an ordered matching.

The contribution from the image of $\ell \in U_1$ to $sc(\pi_2\pi_1)$ is c_1c_2 while the corresponding contributions of $\ell \in U_1$ to $sc(\pi_1)$ is c_1 and of $\pi_1(\ell) \in V$ to $sc(\pi_2)$ was c_2. Thus, $sc(\pi_2\pi_1) = sc(\pi_1) \cdot sc(\pi_2)$. $\square$

The scaling factor characterizes the ratio of $M(U)$ and $M(V)$ when U, V are similar.

Fact 24. *Let π be an ordered matching between lists U, V of linear forms, by an ordered form-ideal $I = \{v_1, \ldots, v_i\}$. If U and V are similar, then $M(V) = sc(\pi) \cdot M(U)$. Thus, all ordered matchings between a given pair of similar lists, have the same scaling factor.*

Proof. The proof idea is identical to the one seen in Fact 11. Starting from an ordered matching π, we will convert it into a matching $\widetilde{\pi}$. The number of forms matched to similar forms will increase, and the scaling factor will remain the same.

Let $\ell \in U$ be such that $\pi(\ell) = d\ell + v$ is not similar to ℓ, where $d \in \mathbb{F}^*$, $v \in sp_j$ and $\ell \notin sp_j$. Since V is similar to U, $\#simi(\ell, U) = \#simi(\ell, V)$. So there exists a form similar to ℓ in V that is matched to a non-similar form in U. Let this form be $c\ell \in V$, for some $c \in \mathbb{F}^*$. As π is an ordered matching, it maps some $\ell' \in U$ to $c\ell$ in V, satisfying: $\pi(\ell') = d'\ell' + v' = c\ell$, where $d' \in \mathbb{F}^*$, $v' \in sp_{j'}$, and $\ell' \notin sp_{j'}$.

Now we define a new matching $\widetilde{\pi}$ by flipping the images of ℓ and ℓ' under π, i.e., define $\widetilde{\pi}$ to be the same as π on $U \setminus \{\ell, \ell'\}$ and: $\widetilde{\pi}(\ell) \overset{V}{:=} \pi(\ell')$ and $\widetilde{\pi}(\ell') \overset{V}{:=} \pi(\ell)$. The matching $\widetilde{\pi}$ is an ordered matching because: $\widetilde{\pi}(\ell) = c\ell$ for $c \in \mathbb{F}^*$ and more importantly $\widetilde{\pi}(\ell') = d\ell + v = d(\frac{d'\ell'+v'}{c}) + v = (\frac{dd'}{c})\ell' + (\frac{dv'}{c} + v)$. Let $j^* := \max\{j, j'\}$. Obviously, $(\frac{dv'}{c} + v) \in sp_{j^*}$. If $j^* = j'$, we are done, because we already know that $\ell' \notin sp_{j'}$. If $j^* = j$ and $\ell' \in sp_j$, then $c\ell = d'\ell' + v'$ is in sp_j (contradiction).

We have obtained now an ordered I-matching $\widetilde{\pi}$ between U, V where the number of forms mapped to a similar form has strictly increased. Observe that $sc(\pi)$ had a unique contribution of d, d' from the images of ℓ, ℓ' respectively while $sc(\widetilde{\pi})$ has a corresponding contribution of c, $(\frac{dd'}{c})$. On all the other elements of U, $\widetilde{\pi}$ is the same as π. Thus, we have that $sc(\widetilde{\pi}) = sc(\pi)$.

The above process will yield an ordered matching π' in at most $\#U$ many iterations, such that U, V are similar under π' and $sc(\pi') = sc(\pi)$. But this means that, for all $\ell \in U$, $\pi'(\ell) = \lambda_\ell \ell$, for some $\lambda_\ell \in \mathbb{F}^*$. The contribution by ℓ to $sc(\pi')$ is λ_ℓ. This implies that $M(V) = sc(\pi') \cdot M(U)$ and hence $M(V) = sc(\pi) \cdot M(U)$. □

We move on to facts about the *gcd* and *sim* parts of a circuit.

Fact 25. *If C_Q is a regular mod I subcircuit of C then:*

$$C_Q \equiv gcd(C_Q \ mod\ I) \cdot sim(C_Q \ mod\ I) \ (mod\ I)$$

Additionally, if C_Q is an identity modulo I then $sim(C_Q \ mod\ I)$ is a simple identity modulo I.

Proof. Recall that $C_Q = \sum_{q \in Q} T_q$ and the *gcd* data $\overline{gcd}(C_Q \text{mod } I)$ is $\{(\pi_q, U, U_q) \mid q \in Q\}$. Now $T_q = \alpha_q M(U_q) \cdot M(L(T_q) \setminus U_q)$ and $M(U_q) \equiv sc(\pi_q) \cdot M(U)$ (mod I), where $M(U)$ is $gcd(C_Q \text{mod } I)$. Thus,

$$\begin{aligned} C_Q &\equiv \sum_{q \in Q} \alpha_q sc(\pi_q) M(U) \cdot M(L(T_q) \setminus U_q) \ (\text{mod } I) \\ &\equiv gcd(C_Q \text{ mod } I) \cdot sim(C_Q \text{ mod } I) \ (\text{mod } I) \end{aligned}$$

This proves the first part. Assume now that $C_Q \equiv 0 (\text{mod } I)$ which means $sim(C_Q \text{ mod } I) \equiv 0 (\text{mod } I)$. If it is not a simple identity mod I, then there is an $\ell' \in L(sim(C_Q \text{ mod } I))$ such that, $\forall q \in Q$, $\ell' \mid M(L(T_q) \setminus U_q)$ mod I. Then, $M(U)$ cannot be the gcd of the polynomials $\{T_q \mid q \in Q\}$ modulo the ideal (I) (contradiction). □

When $I = \{0\}$ we write $\overline{gcd}(C_Q)$, $gcd(C_Q)$ and $sim(C_Q)$ instead of $\overline{gcd}(C_Q \text{ mod } I)$, $gcd(C_Q \text{ mod } I)$ and $sim(C_Q \text{ mod } I)$ respectively. We collect here some properties of $sim(C_Q)$ that would be directly useful in our rank bound proof.

Fact 26. *Let $\ell \in L(\mathcal{R})^*$ and C_Q be a subcircuit of C. Then $\#simi(\ell, L(sim(C_Q))) > 0$ iff $\exists q_1, q_2 \in Q$ such that $\#simi(\ell, L(T_{q_1})) \neq \#simi(\ell, L(T_{q_2}))$.*

Proof. Recall that $\#simi(\ell, L(T_q))$ is the highest power of ℓ that divides T_q. Thus, if $\#simi(\ell, L(T_q))$ is the same, say r, for all $q \in Q$ then the highest power of ℓ dividing $gcd(C_Q)$ is also r. Therefore, for all $q \in Q$, the polynomial $\frac{T_q}{gcd(C_Q)}$ is coprime to ℓ. By definition of the simple part of C_Q this means that $\#simi(\ell, L(sim(C_Q))) = 0$.

Conversely, for any $\ell \in L(\mathcal{R})^*$, $\exists q_1, q_2 \in Q$ such that $\#simi(\ell, L(T_{q_1})) > \#simi(\ell, L(T_{q_2}))$. It follows that $\frac{T_{q_1}}{gcd(C_Q)}$ cannot be coprime to ℓ. This implies that $\#simi(\ell, L(sim(C_Q))) > 0$. □

Fact 27. *Let $S \subseteq L(\mathcal{R})$ and $Q_2 \subseteq Q_1 \subseteq [k]$. If $L(sim(C_{Q_1}))$ has all its linear forms in $sp(S)$, then all the linear forms in $L(sim(C_{Q_2}))$ are also in $sp(S)$.*

Proof. For an arbitrary $\ell \in L(sim(C_{Q_2}))$, by Fact 26, there are $q_1, q_2 \in Q_2$ such that $\#simi(\ell, L(T_{q_1})) \neq \#simi(\ell, L(T_{q_2}))$. As $q_1, q_2 \in Q_1$, we can again apply Fact 26 to deduce that $\#simi(\ell, L(sim(C_{Q_1}))) > 0$. Therefore $\ell \in sp(S)$. □

Fact 28. *Let $S \subseteq L(\mathcal{R})$ and $Q_1, Q_2 \subseteq [k]$ such that $Q_1 \cap Q_2 \neq \phi$. If $L(sim(C_{Q_1}))$ and $L(sim(C_{Q_2}))$ have all their linear forms in $sp(S)$ then all the linear forms in $L(sim(C_{Q_1 \cup Q_2}))$ are also in $sp(S)$.*

Proof. Take $q_0 \in Q_1 \cap Q_2$ and an arbitrary $\ell \in L(sim(C_{Q_1 \cup Q_2}))$. By Fact 26, there are $q_1, q_2 \in Q_1 \cup Q_2$ such that $\#simi(\ell, L(T_{q_1})) \neq \#simi(\ell, L(T_{q_2}))$.

If q_1, q_2 are in the same set (wlog, in Q_1), then Fact 26 tells us that $\#simi(\ell, L(sim(C_{Q_1}))) > 0$, trivially implying that $\ell \in sp(S)$. Now assume wlog that $q_1 \in Q_1, q_2 \in Q_2$. For some $i \in \{1, 2\}$, $\#simi(\ell, L(T_{q_0})) \neq \#simi(\ell, L(T_{q_i}))$. Therefore, by Fact 26, $\ell \in sp(S)$. □

2.5 Getting Useful Form-ideals

Given a set S of linear forms that does not span all of $L(C)$, we can find a form-ideal that is useful wrt S. As we mentioned earlier, in a *round* we start with S, and end up with a useful I through various iterations. We will formally describe this process below.

An iteration starts with a partial I, and a simple regular identity E in the ring $\mathcal{R}/I$, which has multiplication terms with indices in $[k]$. At least one of the forms in E is *not* in $sp(S \cup I)$. At the beginning of the first iteration, E is set to C and I is $\{0\}$.

A SINGLE ITERATION

1. Let ℓ be a form in E that is not in $sp(S \cup I)$.
2. Add ℓ to I (add ℓ to end of the ordered set generating I).
3. Consider E modulo I and let Q be the subset of indices of nonzero multiplication terms.
4. Let U be the *gcd* of $E(\text{mod } I)$, and let the gcd data be $\overline{gcd} = \{(\pi_q, U, U_q) \mid q \in Q\}$.
5. If the fanin, $|Q|$, of $E(\text{mod } I)$ is 2, stop the round.
6. If all forms in $sim(E \text{ mod } I)$ are contained in $sp(S \cup I)$, stop the round. Otherwise, set E to be $sim(E \text{ mod } I)$ and go to the next iteration.

Lemma 29. *Let C be a simple $\Sigma\Pi\Sigma(k,d)$ identity in $\mathcal{R}$. Suppose $S \subseteq L(\mathcal{R})$ and $L(C)\backslash sp(S)$ is non-empty. Then the above procedure finds a form-ideal I useful in C wrt S.*

Proof. As discussed before in the intuition, we generate I in one round and the proof will be done by induction on the number of iterations in this round. For convenience, we set the end of the zero iteration to be the beginning of the round. We will prove the following claim:

Claim 30. *Consider the end of some iteration. There exists a list V of forms such that: for all q in the current Q, there is a list $V_q \subseteq L(T_q)$ that has an ordered I-matching τ_q to V. Furthermore, the circuit $sim(E \bmod I)$ is an identity and a multiple of $\sum_{q\in Q} sc(\tau_q)\alpha_q \cdot M(L(T_q) \setminus V_q)$*

Proof of Claim 30. This is proven by induction on the iterations. At the end of the zero iteration, E is just C and $I = \{0\}$. By the simplicity of C, $sim(E \bmod I)$ is just C, and $Q = [k]$. So all the V_q's can be taken just empty.

Now, suppose that at the end of the ith iteration, we have the circuit E, the ideal I, and the subset Q. In the next iteration, we have $E' := sim(E \bmod I)$, the ideal $I' := I \cup \{\ell\}$, and $Q' \subset Q$, the subset of indices of non-zero terms in E' modulo I'. By the induction hypothesis, $sim(E \bmod I)$ is an identity. By Fact 25, the circuit $sim(E' \bmod I')$ is also an identity.

It remains to find an appropriate V_q's and ordered matchings. We have a list V and for every $q \in Q'$, an ordered I-matching τ_q between V, V_q. By Fact 21, τ_q is also an ordered I'-matching. Let U_q be the gcd (of the $(i+1)$th round) in T_q. Set $V_q' = V_q \cup U_q$, noting that this is a disjoint union. Consider the I'-matching π_q between U, U_q obtained in this iteration. No forms in these can be in $sp(I')$, since U is $gcd(E' \bmod I')$ and $q \in Q'$. Therefore, π_q is an ordered matching. We can take the disjoint union of these matchings to get an ordered I'-matching $\tau_q' = \tau_q \sqcup \pi_q$ between $V \cup U$ and V_q'. We set V' to be $V \cup U$. By the induction hypothesis, E' is a multiple of $\sum_{q\in Q} sc(\tau_q)\alpha_q \cdot M(L(T_q) \setminus V_q)$. The circuit $sim(E' \bmod I)$ is a multiple of $\sum_{q\in Q} sc(\tau_q)sc(\pi_q)\alpha_q \cdot M(L(T_q) \setminus V_q')$. Fact 22 tells us that $sc(\tau_q') = sc(\tau_q)sc(\pi_q)$, completing the proof. □

The number of iterations in a round is at most $(k-2)$. This is because after each iteration, the fanin of the circuit E goes down by at least 1. Therefore, there must be a last iteration (signifying the end of the round). Consider the end of the last iteration. If the fanin $|Q|$ of $E(\bmod I)$ is 2, then by unique factorization, $sim(E \bmod I)$ is empty. Whichever way the round ends, all the forms in $sim(E \bmod I)$ are in $sp(S \cup I)\backslash sp(I)$. By the previous claim, there is a list V such that for every surviving $q \in Q$, there is a sublist $V_q \subseteq L(T_q)$ and an ordered I-matching τ_q between V and V_q. By Fact 25, we have that $sim(E \bmod I))$ is a multiple of $\sum_{q\in Q} sc(\tau_q)\alpha_q \cdot M(L(T_q) \setminus V_q)$ and is an identity (in $\mathcal{R}/I$).

Let $V_q' := V_q \setminus (sp(S \cup I)\backslash sp(I))$ (similarly, define V'). Note that τ_q induces a matching τ_q' between V' and V_q'. Furthermore, $\sum_{q\in Q} sc(\tau_q')\alpha_q \cdot M(L(T_q) \setminus V_q')$ is a multiple of $E(\bmod I)$ and is regular (each term in the above sum is non-zero mod I). Thus, the form-ideal I is useful in C wrt S. □

To prove a rank bound for minimal and simple $\Sigma\Pi\Sigma(k,d)$ identity C, our plan is to start with $S = \phi$ and expand it round-by-round by adding the forms of a form-ideal, useful in C wrt S, to the current S. Trivially, such a process has to stop in at most kd iterations (over all rounds) but we intend to show that it actually terminate more rapidly. All the forms in $L(C)$ will be covered by a small set of form-ideals. To formalize this process we need the notion of a *chain of form-ideals.* This is just a concise representation of the matchings that we get from the various rounds.

Definition 31. [Chain of form-ideals] *Let C be a $\Sigma\Pi\Sigma(k,d)$ circuit. We define a* chain of form-ideals for *C to be the ordered set $\mathcal{T} := \{(C, S_1, I_1, Q_1), \ldots, (C, S_m, I_m, Q_m)\}$ where,*

- *For all $i \in [m]$, $S_i \subseteq L(\mathcal{R})$, I_i is a form-ideal orthogonal to S_i and $Q_i \subseteq [k]$.*
- *$S_1 = \phi$ and for all $2 \leq i \leq m$, $S_i = S_{i-1} \cup I_{i-1}$.*
- *For all $i \in [m]$, I_i is useful in C wrt S_i.*
- *For all $i \in [m]$, Q_i is a blocking subset of C, S_i, I_i.*

We will use $sp(\mathcal{T})$ to mean $sp(S_m \cup I_m)$ and $\#\mathcal{T}$ to denote m, the length of *$\mathcal{T}$. The chain $\mathcal{T}$ is* maximal *if $L(C) \subseteq sp(\mathcal{T})$.*

Note that by Lemma 29, if a chain $\mathcal{T}$ of length m is not maximal, then we can find a form-ideal I_{m+1} that is useful wrt $S_m \cup I_m$. This allows us to add a new $(C, S_{m+1}, I_{m+1}, Q_{m+1})$ to this chain. It is easy to construct a maximal chain for C, and the length of this can be used to bound the rank.

Fact 32. *Let C be a simple $\Sigma\Pi\Sigma(k,d)$ identity. Then there exists a maximal chain of form-ideals $\mathcal{T}$ for C. The rank of C is at most $(k-2)(\#\mathcal{T})$.*

Proof. We start with $S_1 = \phi$ and an $\ell \in L(C)$. By Lemma 29 there is a form-ideal I_1 (containing ℓ) useful in C wrt S_1 with blocking subset, say, Q_1. So we have a chain of form-ideals $\{(C, S_1, I_1, Q_1)\}$ to start with. Now if $L(C)$ has all its elements in $sp(S_1 \cup I_1)$ then the chain cannot be extended any further and we are done. Otherwise, we can again apply Lemma 29 to get a form-ideal I_2 useful in C wrt $S_2 := S_1 \cup I_1$ with blocking subset, say, Q_2. Thus, we have a longer chain of form-ideals $\{(C, S_1, I_1, Q_1), (C, S_2, I_2, Q_2)\}$ now. We keep repeating till we have a chain of length m where $L(C) \subseteq sp(S_m \cup I_m)$.

Note that $S_m \cup I_m = \bigcup_{i \leq m} I_m$. Each I_i is generated by at most $(k-2)$ forms, so there is a basis for $L(C)$ having at most $(k-2)m$ forms. □

We state a slightly stronger version of the main theorem of this paper.

Theorem 33. *If C is a simple and minimal $\Sigma\Pi\Sigma(k,d)$ identity then the length of any maximal chain of form-ideals for C is at most $\binom{k}{2}(\log_2 d + 3) + (k-1)$.*

This theorem with Fact 32 imply the main result, Theorem 2. We prove this theorem in the next section.

2.6 Counting All Matchings: Proof of Theorem 33

Let a maximal chain of form-ideals $\mathcal{T}$ for C be $\{(C, S_1, I_1, Q_1), \ldots, (C, S_m, I_m, Q_m)\}$. We will partition the elements of the chain into three types according to properties of the matchings that they represent. Each of these types will be counted separately.

We first set some notation before explaining the different types. Let the m matching data be:

$$mdata(C, S_i, I_i, Q_i) =: \{(\tau_{i,q}, V_i, V_{i,q}) \mid q \in Q_i\}$$

We will use $mdata_i$ as shorthand for the above. For all $q \in Q_i$, $V_{i,q}$ is a sublist of $L(T_q)$ and $\tau_{i,q}$ is an ordered matching between $V_i, V_{i,q}$ by I_i. By the definition of usefulness of form-ideal I_i we have that $V_{i,q}$ is disjoint to $sp(S_i \cup I_i) \setminus sp(I_i)$. Thus, $V_{i,q}$ can be partitioned into two sublists:

$$V_{i,q,0} := (\ell \in V_{i,q} \mid \ell \in sp(I_i)), \quad \text{and}$$
$$V_{i,q,1} := (\ell \in V_{i,q} \mid \ell \notin sp(S_i \cup I_i)).$$

and analogously V_i can be partitioned into two sublists $V_{i,0}$ and $V_{i,1}$. It is easy to see that these partitions induce a corresponding partition of $\tau_{i,q}$ as $\tau_{i,q,0} \sqcup \tau_{i,q,1}$, where $\tau_{i,q,0}$ (and $\tau_{i,q,1}$) is an ordered matching between $V_{i,0}$, $V_{i,q,0}$ (and $V_{i,1}$, $V_{i,q,1}$) by I_i.

Here are the three types of $mdata_i$'s:

1. **[Type 1]** There exist $q_1, q_2 \in Q_i$ such that $V_{i,q_1,1}$ is not similar to $V_{i,q_2,1}$.
2. **[Type 2]** There exist $q_1, q_2 \in Q_i$ such that V_{i,q_1} is not similar to V_{i,q_2}, but for all $r_1, r_2 \in Q_i$, $V_{i,r_1,1}$ and $V_{i,r_2,1}$ are similar.
3. **[Type 3]** For all $q_1, q_2 \in Q_i$, V_{i,q_1} is similar to V_{i,q_2}. In other words, $mdata_i$ is trivial.

We partition $[m]$ into sets N_1, N_2, N_3, which are the index sets for the $mdata$ of types $1, 2, 3$ respectively.

2.6.1 Bounding $\#N_1$ and $\#N_2$

The dominant term in Theorem 33 comes from $\#N_1$. If $\#N_1$ is large, then by an averaging argument, for some pair (a, b), we find many matchings between forms in T_a and T_b. These are all orthogonal matchings, but are defined on *different* sublists of $L(T_a)$ and $L(T_b)$. Nonetheless, we can find two dissimilar sublists that are matched too many times. Invoking Lemma 12 gives us the required bound.

Lemma 34. $\#N_1 \leq \binom{k}{2}(\log_2 d + 2)$.

Proof. For the sake of contradiction, let us assume $\#N_1 > \binom{k}{2}(\log_2 d + 2)$. For each $mdata_i$ $(i \in N_1)$, choose an unordered pair of indices $P_i = \{q_1, q_2\}$ such that $V_{i,q_1,1}$ and $V_{i,q_2,1}$ are not similar. As there can be only $\binom{k}{2}$ distinct pairs, by the pigeonhole principle, $s > (\log_2 d{+}2)$ of the pairs P_i are equal. Let $P_{i_1} = \cdots = P_{i_s} = \{a, b\}$ for $i_1 < \cdots < i_s \in N_1$. Now we will focus our attention solely on the ordered matchings $\mu_i := \tau_{i,b,1}\tau_{i,a,1}^{-1}$ between $V_{i,a,1}, V_{i,b,1}$ by I_i, for all $i \in \{i_1, \ldots, i_s\}$. The source of contradiction is the fact that all these matchings are also well defined on the 'last' pair of sublists $V_{i_s,a,1}, V_{i_s,b,1}$:

Claim 35. *For all $i \in \{i_1, \ldots, i_s\}$, μ_i induces an ordered matching between $V_{i_s,a,1}, V_{i_s,b,1}$ by I_i.*

Proof of Claim 35. The claim is true for $i = i_s$ so let $i < i_s$. The matching μ_i is an ordered I_i-matching between $V_{i,a,1}$, $V_{i,b,1}$. For $\ell \in V_{i_s,a,1}$, $\ell \notin sp(S_{i_s} \cup I_{i_s})$. Since $i < i_s$ and $L(T_a) \setminus V_{i,a,1} \subset sp(S_i \cup I_i)$, ℓ cannot be in $L(T_a) \setminus V_{i,a,1}$. Therefore, ℓ is in $V_{i,a,1}$. So μ_i maps ℓ to some element in $V_{i,b,1}$, showing μ_i is defined on the *domain* $V_{i_s,a,1}$.

So we know μ_i maps $\ell \in V_{i_s,a,1}$ to an element $\mu_i(\ell) \in V_{i,b,1}$. As μ_i is an I_i-matching, $\mu_i(\ell) = (c\ell + \alpha)$ for some $c \in \mathbb{F}^*$ and $\alpha \in sp(I_i) \subseteq sp(I_{i_s})$, thus $\mu_i(\ell) \notin sp(S_{i_s} \cup I_{i_s})$ (recall $\ell \notin sp(S_{i_s} \cup I_{i_s})$). Thus $\mu_i(\ell)$ cannot be in $L(T_b) \setminus V_{i_s,b,1}$ (which has all its elements in $sp(S_{i_s} \cup I_{i_s})$). Since $\mu_i(\ell) \in L(T_b)$, $\mu_i(\ell) \in V_{i_s,b,1}$.

Thus, μ_i maps an arbitrary $\ell \in V_{i_s,a,1}$ to $\mu_i(\ell) \in V_{i_s,b,1}$. In other words, μ_i induces an ordered matching between $V_{i_s,a,1}$, $V_{i_s,b,1}$ by I_i. □

This claim means that there are $s > (\log_2 d + 2)$ bipartite matchings between $V_{i_s,a,1}$, $V_{i_s,b,1}$ by orthogonal form-ideals $I_{i_1}, \ldots, I_{i_s}$ respectively. Lemma 12 implies that the lists $V_{i_s,a,1}, V_{i_s,b,1}$ are similar. This contradicts the definition of P_{i_s}. Thus, $\#N_1 \leq \binom{k}{2}(\log_2 d + 2)$. □

For dealing with $\#N_2$, we use a slightly different argument to get a better bound. We show that a Type 2 matching can involve a pair of terms at most once.

Lemma 36. $\#N_2 \leq \binom{k}{2}$.

Proof. For the sake of contradiction, assume $\#N_2 > \binom{k}{2}$. For each $mdata_i$ ($i \in N_2$), let P_i be an unordered pair (q_1, q_2) such that V_{i,q_1} is not similar to V_{i,q_2}. Note that because $V_{i,q_1,1}$ is *similar* to $V_{i,q_2,1}$, it must be that $V_{i,q_1,0}$ is not similar to $V_{i,q_2,0}$. By the pigeonhole principle, at least two P_i's are the same. Suppose $P_{i_1} = P_{i_2} = \{a, b\}$ for $i_1 < i_2 \in N_2$.

We will show that the $V_{i_2,a,0} = V_{i_1,a,1} \cap sp(I_{i_2})$ and similarly, $V_{i_2,b,0} = V_{i_1,b,1} \cap sp(I_{i_2})$. This together with the similarity of $V_{i_1,a,1}$ and $V_{i_1,b,1}$ gives us (by Fact 8) that $V_{i_2,a,0}$ and $V_{i_2,b,0}$ are similar. This is a contradiction and thus, $\#N_2 \leq \binom{k}{2}$.

Let $\ell \in V_{i_2,a,0}$. By definition, $\ell \in sp(I_{i_2})$. We know that $L(T_a) \setminus V_{i_1,a,1}$ is contained in $sp(S_{i_1} \cup I_{i_1})$. Since $i_1 < i_2$, I_{i_2} is orthogonal to $sp(S_{i_1} \cup I_{i_1})$ and $\ell \notin L(T_a) \setminus V_{i_1,a,1}$. Since $\ell \in L(T_a)$, we get $V_{i_2,a,0} \subseteq V_{i_1,a,1} \cap sp(I_{i_2})$.

The usefulness of I_{i_2} implies that $V_{i_2,a,0} = L(T_a) \cap sp(I_{i_2})$. This gives $V_{i_1,a,1} \cap sp(I_{i_2}) \subseteq V_{i_2,a,0}$. □

2.6.2 Bounding $\#N_3$

This requires a different argument than the pigeonhole ideas used for $\#N_1$ and $\#N_2$. Observe that till now, we have not used the minimality of C. This is the crucial argument that uses this property of C.

Our final aim is to prove:

Lemma 37. $\#N_3 \leq (k-1)$

We shall use a combinatorial picture of how the chain of form-ideals connects the various multiplication terms through matchings. We will describe an evolving *forest* $\mathcal{F}$ and only deal with Type 3 $mdata_i$.

Initially, the forest $\mathcal{F}$ consists of k isolated vertices, each representing the k terms $T_1, \cdots, T_k$. We process each $mdata_i$ in increasing order of the i's, and update the forest $\mathcal{F}$ accordingly. We will refer to this as *adding* $mdata_i$ *to* $\mathcal{F}$. At any intermediate state, the forest $\mathcal{F}$ will be a collection of rooted trees with a total of k leaves.

We divide type 3 matchings into *internal* and *external* ones.

Definition 38. *Consider $\mathcal{F}$ when $mdata_i$ is processed. If all of Q_i belongs to a single tree in $\mathcal{F}$, then $mdata_i$ is called* internal. *Otherwise, it is called* external.

If $mdata_i$ is internal, $\mathcal{F}$ remains unchanged. While each time we encounter an external $mdata_i$, we update the forest $\mathcal{F}$ as follows. We create a new root node labelled with $mdata_i$ (abusing notation, we refer to $mdata_i$ as a node), and for any tree of $\mathcal{F}$ that contains a T_q, $q \in Q_i$, we make the root of this tree a child of $mdata_i$.

Fact 39. *The total number of external matchings is at most $(k-1)$.*

Proof. Note that each external $mdata_i$ reduces the number of trees in the forest $\mathcal{F}$ by at least one. Initially, $\mathcal{F}$ has k trees, and $\mathcal{F}$ always contains at least one tree. This completes the proof. □

It remains to count the number of internal matchings. Whenever we encounter an internal $mdata_i$, we can always associate it with some root $mdata_{i'}$ of the current $\mathcal{F}$ such that $i' < i$ and all of Q_i is in the tree rooted at $mdata_{i'}$.

Lemma 40. *If $mdata_i$ is internal, then the subcircuit C_{Q_i} is identically zero in $\mathcal{R}$. Therefore, by the minimality of C, no $mdata_i$ can be internal.*

This lemma with the previous fact immediately implies that $\#N_3 \leq (k-1)$. We now set the stage to prove this lemma. Take any Type 3 $mdata_i$. By the triviality of $mdata_i$, the lists in $\{V_{i,q} \mid q \in Q_i\}$ are mutually similar. By the usefulness of I_i, for all $q \in Q$, $L(T_q) \setminus V_{i,q} \subseteq sp(S_i \cup I_i) \setminus sp(I_i)$. Furthermore, $D_i := \sum_{q \in Q_i} sc(\tau_{i,q})\alpha_q \; M(L(T_q) \setminus V_{i,q})$ is a regular identity modulo I_i. Our aim is to remove the forms in D_i which are common factors (*not* mod I_i, but mod 0). This gives us a new circuit (quite naturally, that will turn out to be $sim(C_{Q_i})$) that is still an identity (mod I_i). In other words, start with the subcircuit C_{Q_i}, and remove all common factors from this subcircuit. This is expected to be both $sim(C_{Q_i})$ and an identity mod I_i.

Using this we will actually show that if $mdata_i$ is internal then $sim(C_{Q_i})$ is an identity (mod 0). Then we can multiply the common factors back, and C_{Q_i} would be an absolute identity (violating minimality of C). We proceed to show this rigorously. We have to carefully deal with field constants to ensure that $sim(C_{Q_i})$ is indeed a factor of D_i.

Claim 41. *For Type 3 $mdata_i$, the circuit $sim(C_{Q_i})$ is an identity mod I_i and has all its forms in $sp(S_i \cup I_i)$.*

Proof. Let the gcd data of D_i be:

$$\overline{gcd}(D_i) := \{(\pi_{i,q}, U_i, U_{i,q}) \mid q \in Q_i\}$$

where $U_{i,q}$ is a sublist of $L(T_q) \setminus V_{i,q}$ and $\pi_{i,q}$ is an ordered matching between $U_i, U_{i,q}$ by $\{0\}$. Note that this is *not* mod I_i, even though D_i is an identity only mod I_i.

By Facts 22 and 25 we can 'stitch' U's and V's to get:

- $\tau'_{i,q} := \tau_{i,q} \sqcup \pi_{i,q}$ is an ordered matching between $V'_i := V_i \cup U_i$, $V'_{i,q} := V_{i,q} \cup U_{i,q}$ by I_i.
- $D'_i := \sum_{q \in Q_i} sc(\tau'_{i,q})\alpha_q \; M(L(T_q) \setminus V'_{i,q})$, is a regular identity modulo I_i.

Fix an arbitrary element q_m in Q_i. We have that $\tau'_{i,q}\tau'^{-1}_{i,q_m}$ is an ordered I_i-matching between the similar lists $V'_{i,q_m}, V'_{i,q}$. By Fact 24, we can construct an ordered matching $\mu_{i,q}$ between V'_{i,q_m}, $V'_{i,q}$ by $\{0\}$, with scaling factor equal to $sc(\tau'_{i,q} \; \tau'^{-1}_{i,q_m}) = sc(\tau'_{i,q})/sc(\tau'_{i,q_m})$.

The way D'_i is constructed it is clear that D'_i is a simple circuit. This combined with the similarity of V'_{i,q_m}, $V'_{i,q}$ under $\mu_{i,q}$ implies that the following set of $\#Q_i$ matchings:

$$\left\{(\mu_{i,q}, V'_{i,q_m}, V'_{i,q}) \mid q \in Q_i\right\}$$

is a gcd data of C_{Q_i} modulo (0) and the corresponding simple part is:

$$\begin{aligned} sim(C_{Q_i}) &= \sum_{q \in Q_i} sc(\mu_{i,q})\alpha_q M(L(T_q) \setminus V'_{i,q}) \\ &= \sum_{q \in Q_i} \frac{sc(\tau'_{i,q})}{sc(\tau'_{i,q_m})}\alpha_q M(L(T_q) \setminus V'_{i,q}) \\ &= \frac{1}{sc(\tau'_{i,q_m})} \cdot D'_i \end{aligned}$$

Thus, $sim(C_{Q_i})$ is a regular identity mod I_i as well. Also, by the usefulness of I_i, $sim(C_{Q_i})$ has all its forms in $sp(S_i \cup I_i)$. This completes the proof. □

We now use the structure of $\mathcal{F}$ to show relationships between the various connected terms.

Claim 42. *At some stage, let $mdata_i$ be a root node of $\mathcal{F}$. Let X be a subset of the leaves of $mdata_i$. Then $L(sim(C_X))$ is a subset of $sp(S_i \cup I_i)$.*

Proof. Let the indices of all the external Type 3 $mdata$ be (in order) $i_1, i_2, \cdots$. We prove the claim by induction on the order in which $\mathcal{F}$ is processed. For the base case, let $i := i_1$. Consider $\mathcal{F}$ just after $mdata_i$ is added. The leaves of $mdata_i$ are all in Q_i. By Claim 41, $L(sim(C_{Q_i})) \subset sp(S_i \cup I_i)$. Any X is a subset of Q_i. By Fact 27, $L(sim(C_X)) \subset sp(S_i \cup I_i)$.

For the induction step, consider an external $mdata_i$. When this is processed, a series of trees rooted at $mdata_{j_1}, mdata_{j_2}, \cdots$ will be made children of $mdata_i$. Every j_r is less than i. Let Y_r denote the leaves of the tree $mdata_{j_r}$. Note that $Y_r \cap Q_i \neq \phi$. By the induction hypothesis, $L(sim(C_{Y_r}))$ is a subset of $sp(S_{j_r} \cup I_{j_r})$ $(\subset sp(S_i \cup I_i))$. Let Z_1 be $Q_i \cup Y_1$. By Fact 28 applied to $sim(C_{Y_1})$ and $sim(C_{Q_i})$, we have that $L(sim(C_{Z_1}))$ is in $sp(S_i \cup I_i))$. Let Z_2 be $Z_1 \cup Y_2$. We can apply the same argument to show that $L(sim(C_{Z_2}))$ is in $sp(S_i \cup I_i))$. With repeated applications, we get that for $Z = \bigcup_r Y_r$, $L(sim(C_Z)) \subset sp(S_i \cup I_i))$. Note that Z is the set of all leaves of the tree rooted at $mdata_i$. By Fact 27, $L(C_X) \subset sp(S_i \cup I_i)$, completing the proof. □

We are finally armed with all the tools to prove Lemma 40.

Proof. (of Lemma 40) Consider some internal $mdata_i$. All the elements of Q_i are leaves in the tree rooted at some $mdata_j$, for $j < i$. By Claim 42, $L(sim(C_{Q_i})) \subset sp(S_j \cup I_j)$. But by Claim 41, $sim(C_{Q_i}) \equiv 0 \pmod{I_i}$. Since I_i is orthogonal to $sp(S_j \cup I_j)$, Fact 9 tells us that $sim(C_{Q_i})$ is an identity (mod 0). Therefore, C_{Q_i} is an identity. □

2.7 Factors of a $\Sigma\Pi\Sigma(k,d)$ Circuit: Proof of Theorem 4

The ideal matching technique is quite robust and can be used to prove Theorem 4. Let C be a simple, minimal, nonzero circuit with top fanin k and degree d. This may not necessarily be homoegenous. There is a simple trick, given in Lemma 3.5 of [DS06], that converts C into a circuit computing a homogenous polynomial. For the sake of studying depth-3 circuits (especially, the rank of identities or the set of linear factors), this conversion makes a lot of the analysis cleaner. We give a simple extension of this lemma. The proof is almost identical to that for Lemma 3.5 of [DS06].

Lemma 43. *Given a simple and minimal depth-3 circuit C of degree d and top fanin k, there exists a simple and minimal $\Sigma\Pi\Sigma(k,d)$ circuit $\widehat{C}$ with the following property. The rank of linear factors of $\widehat{C}$ is the same as that of C.*

Proof. Let $C = \sum_{i\leq k} T_i$ be defined on the variables $x_1,\dots,x_n$. Let $T_i = \prod_{j\leq d_i} \ell_{i,j}$, where d_i is the degree of T_i. Abusing notation, we will use C to denote the generated polynomial. As usual, Greek letters are field constants. We let d denote the (formal) degree of C. We introduce a new variable y. We define $\widehat{C} := y^d \cdot C(x_1/y,\dots,x_n/y)$. Note that it can be viewed as a $\Sigma\Pi\Sigma(k,d)$ circuit and we look closely at its linear forms $\widehat{\ell}$. Consider a non-constant linear polynomial ℓ dividing some T_i. If $\ell = \sum_r \alpha_r x_r + \beta$, then the corresponding $\widehat{\ell} = \sum_r \alpha_r x_r + \beta y$. Consequently,

$$\widehat{C} = \sum_{i\leq k} y^{d-d_i} \prod_{j\leq d_i} \widehat{\ell}_{i,j} \quad (= \sum_{i\leq k} \widehat{T}_i)$$

The circuit $\widehat{C}$ is obviously simple. (There must be some $d_i = d$, so y cannot be a common factor.) Since $\widehat{T}_i$ evaluated at $y = 1$ is exactly T_i, $\widehat{C}$ is minimal. Let $C = \sum_{a=0}^{d} P_a$, where P_a is the homogenous part of degree a in C. Then $\widehat{C} = \sum_{a=0}^{d} y^{d-a} P_a$.

Suppose a non-constant linear polynomial ℓ' divides C. By a suitable linear transformation, we can assume that $\ell' = x_1 - \gamma$. Let us further express each P_a as a polynomial in x_1. So $P_a = \sum_{b=0}^{a} Q_{a,b} x_1^b$, where $Q_{a,b}$ is homogenous of degree $(a-b)$ (over variables $x_2,\dots,x_n$). Substituting γ for x_1 (in C) yields the zero polynomial. Collecting terms of the same degree, we get:

$$\sum_{a=0}^{d}\sum_{b=0}^{a} \gamma^b Q_{a,b} = 0 \quad \Longrightarrow \forall c \leq d, \sum_{a,b:a-b=c} \gamma^b Q_{a,b} = 0$$

We now show that the linear form $x_1 - \gamma y$ divides $\widehat{C}$. Setting x_1 to γy in $\widehat{C}$:

$$\begin{aligned}
\sum_{a=0}^{d} y^{d-a} \sum_{b=0}^{a} Q_{a,b} x_1^b &= \sum_{a=0}^{d} y^{d-a} \sum_{b=0}^{a} Q_{a,b} \gamma^b y^b \\
&= \sum_{a=0}^{d}\sum_{b=0}^{a} \gamma^b Q_{a,b} y^{d-(a-b)} \\
&= \sum_{c=0}^{d} y^{d-c} \sum_{a,b:a-b=c} \gamma^b Q_{a,b} = 0
\end{aligned}$$

If any linear polynomial divides C, the homogeneous version divides $\widehat{C}$, and vice-versa by setting $y = 1$. Hence, the rank of linear factors dividing $\widehat{C}$ is identical to that of C. $\square$

Consider a simple and minimal depth-3 circuit C with top fanin k and degree d, that computes polynomial p. We remind the reader of the definition of $L(p)$. Let us factorize p into $\prod_i q_i$, where each q_i is irreducible. Then $L(p)$ denotes the set of *linear factors* of p (that is, $q_i \in L(p)$ if q_i is linear).

By the lemma above, we convert our original circuit C into a simple and minimal $\Sigma\Pi\Sigma(k, d)$-circuit $\widehat{C}$. Let the polynomial computed by $\widehat{C}$ be $\widehat{p}$. For any $\widehat{q} \in L(\widehat{p})$, $\widehat{C} \equiv 0 \pmod{\widehat{q}}$. Therefore we can generate a form-ideal useful in $\widehat{C}$ involving $\widehat{q}$. Using these we can create a chain of form-ideals whose span contains $L(\widehat{p})$, and all our counting lemmas for the matchings of types $1, 2, 3$ will follow. As a result, we get a bound of $O(k^3 \log d)$ on the rank of $L(\widehat{p})$. By Lemma 43, this also holds for the polynomial p.

3 High Rank Identities

The following identity was constructed in [KS07]: over $\mathbb{F}_2$ (with $r \geqslant 2$),

$$
\begin{aligned}
C(x_1, \ldots, x_r) := \quad & \prod_{\substack{b_1,\ldots,b_{r-1} \in \mathbb{F}_2 \\ b_1+\cdots+b_{r-1} \equiv 1}} (b_1 x_1 + \cdots + b_{r-1} x_{r-1}) \\
& + \prod_{\substack{b_1,\ldots,b_{r-1} \in \mathbb{F}_2 \\ b_1+\cdots+b_{r-1} \equiv 0}} (x_r + b_1 x_1 + \cdots + b_{r-1} x_{r-1}) \\
& + \prod_{\substack{b_1,\ldots,b_{r-1} \in \mathbb{F}_2 \\ b_1+\cdots+b_{r-1} \equiv 1}} (x_r + b_1 x_1 + \cdots + b_{r-1} x_{r-1}) \qquad (2)
\end{aligned}
$$

It was shown that, over $\mathbb{F}_2$, C is a simple and minimal $\Sigma\Pi\Sigma$ zero circuit of degree $d = 2^{r-2}$ with $k = 3$ multiplication terms and $rank(C) = r = \log_2 d + 2$. Let $S_1(\overline{x})$, $S_2(\overline{x})$, $S_3(\overline{x})$ denote the three multiplication terms of C. We now build a high rank identity based on S_1, S_2, S_3. Our basic step is given by a composition lemma to construct identities of rank $\Omega(k)$ (first discussed in a personal communication [**?**]). The idea is that a simple, minimal $\Sigma\Pi\Sigma(k, d_1)$ identity can be combined with a $\Sigma\Pi\Sigma(3, d_2)$ one to generate a simple, minimal $\Sigma\Pi\Sigma(k+1, d_1+d_2)$ identity of higher rank. We mention that this lemma was used in [DS06] to generate the high rank identities referred to in their conclusion. Nonetheless, it has not been explicitly stated before.

Lemma 44. *Let $D := \sum_{j=1}^{k} T_j$ be a simple, minimal and zero $\Sigma\Pi\Sigma$ circuit over $\mathbb{F}_2$, defined over s variables $y_1, \ldots, y_s$. Let D have degree g, fanin k and rank s. The circuit C is defined over r variables $x_1, \ldots, x_r$ (as give in Equation 2). Define a new circuit over $\mathbb{F}_2$ using D and C:*

$$D' := \sum_{j=1}^{k-1} T_j \cdot S_1 - T_k \cdot S_2 - T_k \cdot S_3$$

Then D' is a simple, minimal and zero $\Sigma\Pi\Sigma$ circuit with degree $d' = g+d$, fanin $k' = k+1$ and rank $r' = s + r$.

Proof. Since C is an identity, we get that $S_2 + S_3 = -S_1$. Therefore,

$$D' = \big(\sum_{j=1}^{k-1} T_j\big)S_1 - T_k(S_2 + S_3) = \big(\sum_{j=1}^{k-1} T_j\big)S_1 + T_k S_1 = \big(\sum_{j=1}^{k-1} T_j\big)S_1 = 0$$

The terms T_j do not share any variables with S_ℓ ($\ell \in \{1, 2, 3\}$). Since D and C are simple, D' is also simple. Suppose D' is not minimal. We have some subset $P \subset [1, k-1]$ such that $C' := (\sum_{j\in P} T_j)S_1 - \alpha_2 T_k S_2 - \alpha_3 T_k S_3 = 0$, where $\alpha_2, \alpha_3 \in \{0, 1\}$. If both α_2 and α_3 are 1, then we get $(\sum_{j\in P} T_j)S_1 + T_k S_1 = 0$. Since D is minimal, P must be the whole set $[1, k-1]$. On the other hand, if both α_2, α_3 are 0, then $(\sum_{j\in P} T_j)S_1 = 0$ which is impossible as D is minimal. The only remaining possibility is (wlog) $(\sum_{j\in P} T_j)S_1 - T_k S_2 = 0$. As S_1 is coprime to S_2 and T_k, this is impossible. Therefore, D' is minimal.

It is easy to see the parameters of D': $k' = k + 1$ and $d' = g + d$. Because the T_j's do not share any variables with S_ℓ's, the rank $r' = s + r$. □

Family of High Rank Identities: We iteratively use Lemma 44 to generate circuit D_{i+1} from D_i. We initially state with $D_0 := C$. The circuit D_i has degree $d_i = (i + 1)d$, fanin $k_i = i + 3$, and rank $r_i = (i + 1)r = (i + 1)(\log_2 d + 2)$. So r_i relates to k_i, d_i as:

$$r_i = (k_i - 2)\left(\log_2 \frac{d_i}{k_i - 2} + 2\right).$$

Also it can be seen that if $d > i$ then $\frac{d_i}{k_i - 2} \geq \sqrt{d_i}$. Thus after simplification, we have for any $3 \leq i < d$, $r_i > \frac{k_i}{3} \cdot \log_2 d_i$. This gives us an infinite family of $\Sigma\Pi\Sigma(k, d)$ identities over $\mathbb{F}_2$ with rank $\Omega(k \log d)$.

Remark: The above family can be obtained over any field $\mathbb{F}$ of characteristic $p > 0$. The main idea is to generalize Equation (2) to:

$$\begin{aligned} C(x_1, \ldots, x_r) := & \prod_{\substack{b_1,\ldots,b_{r-1}\in\mathbb{F}_p \\ b_1+\cdots+b_{r-1}\equiv 1}} (b_1x_1 + \cdots + b_{r-1}x_{r-1}) \\ & + \prod_{\substack{b_1,\ldots,b_{r-1}\in\mathbb{F}_p \\ b_1+\cdots+b_{r-1}\equiv 0}} (x_r + b_1x_1 + \cdots + b_{r-1}x_{r-1}) \\ & - \prod_{\substack{b_1,\ldots,b_{r-1}\in\mathbb{F}_p \\ b_1+\cdots+b_{r-1}\equiv 1}} (x_r + b_1x_1 + \cdots + b_{r-1}x_{r-1}) \end{aligned} \tag{3}$$

It can be seen that the rank deteriorates with the characteristic as $(\log_p d + 2)$. On the other hand, the highest rank fanin-3 identity known over characteristic zero fields is of rank 4: $x_4(x_4+x_1+x_2)(x_4+x_2+x_3)(x_4+x_3+x_1) - (x_4+x_1)(x_4+x_2)(x_4+x_3)(x_4+x_1+x_2+x_3) + x_1x_2x_3(2x_4 + x_1 + x_2 + x_3)$.

4 Concluding Remarks

It would be very interesting to leverage the matching technique to design identity testing algorithms. By unique factorization, matchings can be easily detected in polynomial time, and it is also not hard to search for I-matchings involving a specific set of forms in I. We prove that depth-3 identities exhibit structural properties described by the ideal matchings. Can we reverse these theorems? In other words, can we show that certain collections of matchings are present iff C is an identity? This would lead to a polynomial time identity tester for *all* depth-3 circuits.

Acknowledgment

The authors would like to thank Neeraj Kayal for helpful discussions and anonymous referees for their suggestions on improving the presentation.

References

[AB03] M. Agrawal and S. Biswas. Primality and identity testing via chinese remaindering. *JACM*, 50(4):429–443, 2003.

[Agr05] M. Agrawal. Proving lower bounds via pseudo-random generators. In *Proceedings of the 25th Annual Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 92–105, 2005.

[AM07] V. Arvind and P. Mukhopadhyaya. The monomial ideal membership problem and polynomial identity testing. In *Proceedings of the 18th International Symposium on Algorithms and Computation (ISAAC)*, pages 800–811, 2007.

[AV08] Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *FOCS*, pages 67–75, 2008.

[CK00] Z. Chen and M. Kao. Reducing randomness via irrational numbers. *SIAM J. on Computing*, 29(4):1247–1256, 2000.

[DS06] Z. Dvir and A. Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. *SIAM J. on Computing*, 36(5):1404–1434, 2006.

[KI04] V. Kabanets and R. Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1):1–46, 2004.

[KS01] A. Klivans and D. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings of the 33rd Annual Symposium on the Theory of Computing (STOC)*, pages 216–223, 2001.

[KS07] N. Kayal and N. Saxena. Polynomial identity testing for depth 3 circuits. *Computational Complexity*, 16(2):115–138, 2007.

[KS08] Z. Karnin and A. Shpilka. Deterministic black box polynomial identity testing of depth-3 arithmetic circuits with bounded top fan-in. In *Proceedings of the 23rd Annual Conference on Computational Complexity (CCC)*, pages 280–291, 2008.

[KS09] N. Kayal and S. Saraf. Blackbox polynomial identity testing for depth 3 circuits. Technical Report TR09-032, ECCC, 2009.

[LV98] D. Lewin and S. Vadhan. Checking polynomial identities over any field: Towards a derandomization? In *Proceedings of the 30th Annual Symposium on the Theory of Computing (STOC)*, pages 428–437, 1998.

[Sax08] N. Saxena. Diagonal circuit identity testing and lower bounds. In *Proceedings of the 35th Annual International Colloquium on Automata, Languages and Programming (ICALP)*, pages 60–71, 2008.

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *JACM*, 27(4):701–717, 1980.

[SS09] N. Saxena and C. Seshadhri. An almost optimal rank bound for depth-3 identities. In *Proceedings of the 24th Annual Conference on Computational Complexity (CCC)*, pages 137–148, 2009.

[SS10] N. Saxena and C. Seshadhri. From Sylvester-Gallai configurations to rank bounds: Improved black-box identity test for depth-3 circuits. Technical Report TR10-013, ECCC, 2010. To appear in proceedings of FOCS 2010.

[SV09] A. Shpilka and I. Volkovich. Improved polynomial identity testing for read-once formulas. In *Proceedings of the 13th International Workshop on Randomization and Computation (RANDOM)*, 2009.

[Zip79] R. Zippel. Probabilistic algorithms for sparse polynomials. *Symbolic and algebraic computation*, pages 216–226, 1979.