On the Structure and Complexity of Symbolic Proofs of Polynomial Identities*

Iddo Tzameret†

Tel Aviv University, Tel Aviv
69978, Israel

Abstract

A *symbolic proof* for establishing that a given arithmetic formula Φ computes the zero polynomial (or equivalently, that two given arithmetic formulas compute the same polynomial) is a sequence of formulas, starting with Φ and deriving the formula 0 by means of the standard polynomial-ring axioms applied to any subformula.

Motivated by results in proof complexity and algebraic complexity, we investigate basic structural and complexity characterizations of symbolic proofs of polynomial identities. Specifically, we introduce fragments of symbolic proofs named *analytic symbolic proofs*, enjoying a natural property: a symbolic proof is analytic if one cannot introduce arbitrary new formulas throughout the proof (that is, formulas computing the zero polynomial which do not originate, in a precise manner, from the initial arithmetic formula).

We establish exponential lower bounds on the lengths of analytic symbolic proofs operating with depth-3 arithmetic formulas, under a certain regularity condition on the structure of proofs (roughly, mimicking a tree-like structure). The hard instances are explicit and rely on small formulas for the symmetric polynomials.

Key words and phrases: Algebraic proof systems, proof complexity, arithmetic circuits, polynomial identity testing, equational systems, lower bounds

2000 Mathematics subject classification: 68Q17, 68Q15, 03F20

1 Introduction

Let $\mathbb{F}$ be a field (say, the complex numbers) and let Φ be an arithmetic formula in the input variables $x_1, \ldots, x_n$, computing a polynomial in the ring of polynomials $\mathbb{F}[x_1, \ldots, x_n]$. A *symbolic operation* is any transformation of a subformula in Φ into another subformula, by means of the standard polynomial-ring axioms (expressing associativity and commutativity of both addition and multiplication, distributivity of multiplication over addition, equalities involving only field elements and the laws for the 0 and 1 elements in the field). This paper deals with the following basic question:

*Independently of this paper, Pavel Hrubeš recently investigated close problems concerning equational systems (Hrubeš (2008)). The current paper is about to be merged with Hrubeš (2008).

†tzameret@cs.tau.ac.il. This work was carried out in partial fulfillment of the requirements for the Ph.D. degree and was supported by The Israel Science Foundation (grant no. 250/05).

How many symbolic operations one needs to perform on Φ in order to make sure that Φ computes the zero polynomial?

To this end we define the notion of *symbolic proofs of polynomial identities* as follows: Assume that the arithmetic formula Φ computes the zero polynomial, then a *symbolic proof* of this fact is a sequence of arithmetic formulas, where the first formula is Φ, the last formula is the formula 0, and every formula in the sequence (excluding the first one) is derived from the (immediate) previous formula in the sequence by a symbolic operation. In this paper we are interested in the lengths of symbolic proofs of polynomial identities, namely, the number of proof-lines in such proof sequences.

1.1 Background and Motivation

The problem of deciding whether a given arithmetic circuit (or formula) over some field computes the zero polynomial – namely, the *polynomial identity testing* problem (PIT, for short) – is of great importance in algebraic complexity theory, and complexity theory in general. It is known that when the underlying field is big enough there is an efficient *probabilistic* procedure for deciding whether an arithmetic circuit computes the zero polynomial (cf. Schwartz (1980); Zippel (1979)). However, not much is known about the complexity of *deterministic* algorithms for this problem. Devising an efficient deterministic procedure, or even demonstrating (non-deterministic) sub-exponential witnesses, for the polynomial identity testing problem is a fundamental open problem.

The importance and apparent difficulty in finding an efficient deterministic procedure (or sub-exponential non-deterministic witnesses for that matter) for PIT led researchers to several different directions. On the one hand, there is a growing body of work dedicated to establishing efficient deterministic procedures for PIT when arithmetic circuits are replaced by more restrictive models of computing polynomials (cf. Raz and Shpilka (2005); Dvir and Shpilka (2006); Kayal and Saxena (2007); Karnin and Shpilka (2007); Shpilka and Volkovich (2008)). On the other hand, in a somewhat more logical vein, evidence or "justifications" for the apparent empirical difficulty in finding efficient deterministic algorithms for PIT were discovered in Kabanets and Impagliazzo (2004) (see also, Dvir et al. (2008)).

In this paper we propose a different direction of research, relevant both to the polynomial identity testing problem as well as to *proof complexity* (namely, the field that studies the sizes of symbolic proofs – and especially of propositional proofs). Instead of studying algorithms for the PIT we shall concentrate on *proofs*, and further restrict our study to *symbolic* proofs of polynomial identities, that is, proof sequences that manipulate formulas and which have clear and natural structure (besides the fact that they can be efficiently recognized). On the one hand, the choice to study proofs instead of algorithms gives the model more strength (in comparison to algorithms), as one can use non-determinism. On the other hand, we will restrict severely the "reasoning" allowed in these proofs, and this will in turn enable us to demonstrate exponential-size lower bounds on certain proofs of polynomial identities. Apparently, this lower bound is the first of its kind.

Connections to proof complexity. Research in proof complexity usually focuses on proof systems for the collection of propositional tautologies. In this setting, a proof of a propositional tautology τ written in the standard De-Morgan langauge (that is, the language that includes propositional variables and the logical connectives $\wedge, \vee, \neg$) is typically a sequence of formulas (in the same language), such that each formula in the sequence is either an axiom or is derived

from prior formulas in the sequence by applying some simple (and logically sound) inference rules, and where the terminal formula in the sequence is τ.

In this paper we shall not deal with proof systems for the collection of propositional tautologies; Instead, we are interested in proof systems for the collection of formulas computing the zero polynomial over a given fixed field. Nevertheless, the way we define (fragments of) symbolic proofs of polynomial identities is inspired by notions from proof complexity (that is, the notion of cut-free or analytic proofs); and moreover, one of the main motivations in studying such systems comes from research in algebraic propositional proof systems (see below for more details on this).

1.1.1 Motivations

As discussed above, research into the complexity of symbolic proofs of polynomial identities is directed, among others, to achieve better understanding of the polynomial identity testing problem: although it is reasonable to assume that there *are* polynomial size witnesses (or "proofs") of polynomial identities, lower bounds on certain symbolic proofs of polynomial identities might lead to better understanding of the structure of proofs needed in order to efficiently prove polynomial identities. On the applicative level, our work can be regarded as a contribution to the understanding of the efficiency of symbolic manipulation systems (like symbolic mathematical software tools). Nevertheless, a concrete motivation for the study of symbolic proofs of polynomial identities comes from the realm of algebraic proof systems, as we now explain.

Algebraic proof systems. Algebraic proof systems, which are proof systems operating with multivariate polynomials over a fixed field, attracts much attention in proof complexity theory. Such algebraic proofs usually demonstrate that a collection of polynomial equations, derived from the clauses of an unsatisfiable formula in conjunctive normal form (CNF), has no $0,1$ solutions over the fixed field – in which case the systems are called *algebraic propositional proof systems*.

In a typical algebraic proof system, the fact that the a given collection of polynomial equations has no solutions over the base field is proved by using basic (and sound) algebraic inference rules (for example, from two polynomial equations $p = 0$ and $q = 0$, we can deduce $\alpha \cdot p + \beta \cdot q$, where α and β are some elements of $\mathbb{F}$; and from $p = 0$ we can deduce $q \cdot p = 0$, for any polynomial q). A sequence of polynomials that starts from the initial polynomial equations, follows the algebraic inference rules of the system and terminates with (the unsatisfiable equation) $1 = 0$, is an algebraic proof establishing the unsatisfiability of the initial polynomial equations (over the base field).

One of the most natural complexity measures for algebraic proofs is their *algebraic complexity*, namely, the number of symbols it takes to write down an algebraic proof when every proof-line is written as an arithmetic formula (or circuit). When this complexity-measure is considered, an algebraic proof is a sequence of *arithmetic formulas* such that each formula in the sequence computes a formal multivariate polynomial (that is, an element of the polynomial-ring $\mathbb{F}[x_1, \ldots, x_n]$, where $\mathbb{F}$ is the base field and $x_1, \ldots, x_n$ are the formal variables of the system). Note however, that for every polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ there is *no unique* arithmetic formula computing p. Thus, each polynomial in the algebraic proof can be written in more than one way as an arithmetic formula. This means that such algebraic proof systems are *semantic* proof systems (and not syntactic), in the sense that the inference of new polynomials from previous ones, via the algebraic inference rules, is a semantic inference of polynomials from preceding ones, rather than a syntactic inference of formulas from preceding formulas (for instance, the two inference

rules mentioned above are semantic in the sense that every root of p in $\mathbb{F}$ is also a root of $q \cdot p$ in $\mathbb{F}$ [for every polynomial q]; and every common root of p and q in $\mathbb{F}$ is also a root of $\alpha \cdot p + \beta \cdot q$, for any $\alpha, \beta \in \mathbb{F}$).

It stems from the aforesaid, that algebraic proofs operating with arithmetic formulas as described above might not necessarily be recognizable in polynomial-time (in the sizes of the proofs): because no polynomial-time procedure for the polynomial identity testing problem is known, no known polynomial-time procedure can verify that a proof-line in an algebraic proof was derived correctly from preceding lines in the proof (for instance, the polynomial p might be written as two completely different arithmetic formulas in a proof-line consisting of p and in its legitimate consequence proof-line consisting of $q \cdot p$). Algebraic proof systems of this nature (alas operating with *multilinear* arithmetic formulas, instead of general arithmetic formulas) where investigated in Raz and Tzameret (2008a,b).

Nevertheless, it is sometime preferable to turn to algebraic proofs that are polynomial-time recognizable. The most natural choice here is to join together the underlying semantic algebraic proof system that operates with arithmetic formulas over a field (similar to that mentioned above), with a symbolic proof system for establishing polynomial identities. This can be achieved, for instance, in the following simple manner: A *syntactic algebraic proof* is defined now to be a sequence of *arithmetic formulas* in which each proof-line is either (i) an initial formula; or (ii) was derived from a previous formula in the sequence by one of the derivation rules pertaining to the symbolic proof system (expressing the polynomial-ring axioms, applicable on any subformula); or (iii) was derived by the following inference rules that correspond to the rules of the underlying algebraic proof system: from the *formulas* φ and ψ derive the *formula* $\alpha \times \varphi + \beta \times \psi$ (for α, β field elements); and from the formula φ derive the formula $\psi \times \varphi$ for any arithmetic formula ψ over the base field (the symbol $\times$ stands for the product gate).

Precisely this kind of natural syntactic algebraic propositional proof systems operating with arithmetic formulas were mentioned in Buss et al. (1996/97) and were explicitly introduced in Grigoriev and Hirsch (2003). Understanding the complexity of symbolic proofs of polynomial identities is thus essential in order to understand such syntactic algebraic propositional proof systems operating with arithmetic formulas. Moreover, establishing super-polynomial lower bounds on symbolic proofs of polynomial identities might yield a super-polynomial separation between semantic and syntactic algebraic (propositional) proof systems. For instance, the short algebraic proofs of "hard tautologies" demonstrated in Raz and Tzameret (2008a,b) use in an essential way the fact that the algebraic proof systems are semantic, and it is not known whether such short proofs exist for corresponding syntactic algebraic proof systems.

1.2 The Basic Model: Analytic Symbolic Proofs

Recall the underlying model of symbolic proofs of polynomial identities illustrated above. We now explain the fragment of symbolic proofs we shall study here. With analogy to traditional research in proof complexity (as well as classical proof theory and automated proofs), we will consider symbolic proof systems that enjoy a useful property, analogous to some extent with the so-called *subformula property* in standard (propositional or predicate sequent calculus) proofs. The subformula property states that every formula that appears in a proof sequence π of T also appears in T. Intuitively, this means that the prover is forbidden from introducing notions not already present in the statement to be proved. Proofs having the subformula property are sometime called *analytic* (or *cut-free* in the framework of the sequent calculus), and we shall adopt this term here.

Accordingly, we will introduce a proof system for the set of arithmetic formulas computing

the zero polynomial, called *analytic symbolic proofs*, in which the following (relaxed form of the) subformula property holds: if π is a proof sequence that intends to establish that the formula Φ computes the zero polynomial, then every subformula that appears in some proof-line in π is "originated" from the initial formula to be proved. More formally, this means that for every proof-line and every monomial (with its coefficient) that is syntactically computed in the proof-line, the same monomial is also syntactically computed in the initial proof-line (see Section 3 for more details on this and for the definition of syntactic computation of monomials).

The analytic criterion thus implies, for instance, that one cannot add arbitrary formulas computing the zero polynomial in the proof (for example, one cannot get from the proof-line φ to the proof-line $\varphi + f - f$, where f is some arbitrarily chosen arithmetic formula). The (analytic) proof system we introduce, is a natural proof system since, first, symbolic manipulations of polynomial formulas according to the polynomial-ring axioms is something familiar to every high-school student; and second, the restriction to analytic proofs forbids only "ingenious" steps as illustrated above (that is, adding a formula $f - f$, and then using in some clever way this f to efficiently derive the formula 0 in the system).

1.3 Results

The main technical contribution of this paper is an exponential-size lower bound on analytic symbolic proofs of certain hard formulas computing the zero polynomial, where proofs operate with depth-3 formulas and conform to a certain regularity condition on the structure of proofs.

The hard formulas we provide are based on small depth-3 formulas for the elementary symmetric polynomials (see Equation (2)). We establish a lower bound rate of $2^{\Omega(\sqrt{\ell})}$, where ℓ is the number of distinct variables in the initial hard formulas.

The regularity condition intends to keep the following condition: once a proof-line $A \times (B + C)$ is transformed into the proof-line $A \times B + A \times C$, in no way the two formulas $A \times B$ and $A \times C$, as well as any other two formulas that originate (among others) from $A \times B$ and $A \times C$ (in a manner made precise), be united together again into a product formula by means of the distributivity rule. For instance, in our case, after $A \times (B + C)$ was broken into the two sums $A \times B + A \times C$, these two sums ($A \times B$ and $A \times C$) cannot be united together again into a product formula by means of the 'backward' distributivity rule, to yield $A \times (B + C)$, once more.

Techniques. Our lower bound follows by a structural analysis of symbolic proofs, and specifically, by tracing the "paths" in which monomials and subformulas "move" along the proof. Some basic algebraic properties of the small depth-3 formulas of the elementary symmetric polynomials are also exploited in the lower bound proof.

To some extent, the exponential-size lower bounds on analytic regular symbolic proofs we establish can be compared to the exponential-size lower bounds on cut-free tree-like propositional sequent-calculus proofs established by Statman (Statman (1978)). In particular, the *analyticity* of symbolic proofs of polynomial identities is analogous to the cut-freeness of proofs in Statman (1978), and the *regularity* of symbolic proofs roughly stands for the the tree-likeness of the proofs in Statman (1978).

2 Preliminaries

For a natural number m, we use $[m]$ to denote $\{1, \ldots, m\}$. For a graph G we write $|G|$ to denote the number of vertices in G. For an edge directed from u to v in G, we also say that u *points to* v.

2.1 Arithmetic Formulas

Definition 2.1 (Arithmetic formula) *Fix a field $\mathbb{F}$. An* arithmetic formula *is a labeled ordered*[1] *tree, with edges directed from the leaves to the root, and with fan-in at most two. Every leaf of the tree (namely, a node of fan-in 0) is labeled with either an input variable or a field element. Every other node of the tree is labeled with either $+$ or $\times$ (in the first case the node is a* plus gate *and in the second case a* product gate*). We assume that there is only one node of out-degree zero, called the* root*. An arithmetic formula computes a polynomial in the ring of polynomials $\mathbb{F}[x_1, \ldots, x_n]$ in the following way. A leaf just computes the input variable or field element that labels it. A plus gate computes the sum of polynomials computed by its incoming nodes. A product gate computes the product of the polynomials computed by its incoming nodes. The output of the formula is the polynomial computed by the root. The* depth *of a formula Φ is the maximal* number of alternations between plus and product gates *in a path from a leaf to the root of Φ (given a path p from the root to a leaf, the number of alternations between plus and product gates is the number of alternations between consecutive blocks of the same gate-labels). We say that an arithmetic formula has a* plus *or product gate at the root if the root of the formula is labeled with a plus or product gate, respectively. An arithmetic with a plus or product gate at the root is said to be a* plus formula *or* product formula*, respectively.*

Given an arithmetic formula Φ a *subformula* of Φ is any (not necessarily proper) subtree of Φ. We say that *an arithmetic formula φ occurs in an arithmetic formula φ'* if φ is a subformula of φ'. In this case we also say that φ' *contains* φ as a subformula.

Notational conventions. In this paper we deal exclusively with arithmetic formulas, and so we will often use the term "formulas" to mean arithmetic formulas.

Given two formulas Φ_1 and Φ_2, we write $\Phi_1 + \Phi_2$ and $\Phi_1 \times \Phi_2$ to designate the formulas with a plus (respectively, product) gate at the root and Φ_1, Φ_2 as its two children. We will also use parenthesis to designate explicitly the structure of formulas. For example, $(x_1 + x_2) + (x_3 + x_4)$ means that $(x_1 + x_2)$ and $(x_3 + x_4)$ are the two subformulas attached to the root gate (while, $(x_2 + x_3)$, for instance, is not a subformula of the main formula). Most often we will not care for the associativity and order of multiplication and addition, to the effect that we shall not write explicitly the parentheses in formulas, like in $\Phi_1 \times \Phi_2 \times \Phi_3$. Further, we write $\prod_{i \in I} \Phi_i$, where I is a set of indices, to mean the product formula (see Definition 2.1) whose products are all Φ_i (for $i \in I$), where we ignore the associativity of subformulas (formally, every product gate in this formula is still of fun-in 2; though this is not essential). Similarly, we will write $\sum_{i \in I} \Phi_i$ for the plus formula of all Φ_i ($i \in I$). Also, we will sometime abuse notation by identifying arithmetic formulas with the polynomials they compute.

We write $\Phi_1 \equiv \Phi_2$ if Φ_1 and Φ_2 are two syntactically equal formulas (equal as labeled trees; not to be confused with equality between polynomials). For a formula Φ and a node v in Φ, we write Φ_v to denote the subformula of Φ whose root is v. We write $\Phi\{\psi\}_v$ to denote the formula Φ where the subtree rooted by v in Φ is replaced[2] by ψ. We write $\Phi\{\psi\}$ (without explicitly displaying v) to mean that ψ is a subformula of Φ.

[1] This means that there is an order on the edges coming into a certain node.

[2] Note that Φ_v is identified here with the labeled tree rooted in v, and not with all the formulas that are equivalent to the labeled tree rooted in v. In other words, when we replace Φ_v by ψ in Φ, we only replace the subtree of Φ whose root is v, by the tree ψ.

2.1.1 Constant-depth formulas

We shall consider bounded-depth formulas. This means that there is an a priori constant d that bounds the number of alternations between plus and product gates in every path in the formula-graph. A formula Φ is said to be a $\Sigma\Pi\Sigma\ldots$ *formula* (where $\Sigma\Pi\Sigma\ldots$ has $d \geq 1$ symbols) if every path in Φ starting at the root and ending in the immediate ancestor of a leaf in the formula-graph of Φ is labeled with a block of (zero or more) consecutive plus gates followed by a block of (zero or more) consecutive product gates and so forth (d times). If moreover, a $\Sigma\Pi\Sigma\ldots$ (with $d \geq 1$ symbols) formula Φ contains a path starting at the root and ending in the immediate ancestor of a leaf that is labeled with a block of *one or more* consecutive plus gates followed by a block of *one or more* consecutive product gates and so forth (d times), then we say that Φ is a *proper* $\Sigma\Pi\Sigma\ldots$ *formula*. The definition of (proper) $\Pi\Sigma\Pi\ldots$ *formulas* is dual.

Comment 1 *When considering depth-3 formulas we shall slightly change the definition of depth (see Section 3.1) in order to conform to the standard definition of depth-3 arithmetic formulas.*

3 Analytic Symbolic Proofs of Polynomial Identities

Let us fix our underlying field $\mathbb{F}$. Unless otherwise stated, from this point on, all formulas will be arithmetic formulas computing polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. An arithmetic formula computes the zero polynomial if the polynomial computed at the root of the formula is the zero polynomial (e.g., the formula $x_1 + (-1 \times x_1)$). In this section we describe our underlying proof system, that is, *analytic symbolic proof systems for polynomial identities*. The system introduced here is complete and sound for the set of arithmetic formulas computing the zero polynomial. In other words, every formula computing the zero polynomial has a proof (completeness), and only formulas computing the zero polynomial have proofs (soundness).

Definition 3.1 (Derivation rule) *A* derivation rule *is a pair of formulas F, G, written as:*

$$(\star)\ \frac{F}{G}\,,$$

where $(\star)$ being the name *of the derivation rule. Let Φ be a formula and v a node in Φ. Assume that $\Phi_v \equiv F$ (that is, $\Phi \equiv \Phi\{F\}_v$). Then, given the formula Φ and the derivation rule $(\star)$, we can derive from Φ the formula $\Phi\{G\}_v$, in which case we say that $\Phi\{G\}_v$* was derived from Φ by the derivation rule $(\star)$ applied on v.

Notation: Let Φ be a formula and v a node in Φ, such that $\Phi_v \equiv F$, and suppose that $\Phi\{G\}_v$ was derived from Φ by the derivation rule $(\star)$ applied on v. Then we will say that the derivation rule was *applied in* or *on* ψ, in case ψ is a subformula of Φ that contains v. Further, in this case we call the formula G *the consequence of the application of rule* $(\star)$. We write

$$\frac{\Phi\{F\}_v}{\Phi\{G\}_v}\,(\star)\,,$$

to denote the above derivation rule application, where $(\star)$ designates the name of the rule that was applied in order to derive the formula in the lower-line from the formula in the upper-line. (It should be clear from the context that this latter notation is meant to denote a proof sequence [see Definition 3.3 below], and not the description of a derivation rule.)

The following definition formulates the standard polynomial-ring axioms, where "non-analytic" rules are kept out (see discussion below).

Definition 3.2 (Polynomial-ring analytic derivation rules) *The following rules are the* polynomial-ring analytic derivation rules *(where Q_1, Q_2, Q_3 range over all arithmetic formulas computing polynomials in $\mathbb{F}[x_1, \ldots, x_n]$):*

Zero element rules:

$$\frac{0 \times Q_1}{0} \qquad \frac{0 + Q_1}{Q_1}$$

Unit element rules:

$$\frac{1 \times Q_1}{Q_1} \qquad \frac{Q_1}{1 \times Q_1}$$

Scalar rules: let $\alpha, \alpha_1, \alpha_2$ be elements in $\mathbb{F}$.

$$\frac{\alpha_1 + \alpha_2}{\alpha} \text{ (where } \alpha = \alpha_1 + \alpha_2)$$

$$\frac{\alpha_1 \times \alpha_2}{\alpha} \text{ (where } \alpha = \alpha_1 \cdot \alpha_2)$$

$$\frac{\alpha}{\alpha_1 \times \alpha_2} \text{ (where } \alpha = \alpha_1 \cdot \alpha_2)$$

Commutativity:

$$\frac{Q_1 + Q_2}{Q_2 + Q_1} \qquad \frac{Q_1 \times Q_2}{Q_2 \times Q_1}$$

Associativity:

$$\frac{Q_1 + (Q_2 + Q_3)}{(Q_1 + Q_2) + Q_3} \qquad \frac{Q_1 \times (Q_2 \times Q_3)}{(Q_1 \times Q_2) \times Q_3}$$

Forward distributivity:

$$\frac{Q_1 \times (Q_2 + Q_3)}{(Q_1 \times Q_2) + (Q_1 \times Q_3)}$$

Backward distributivity:

$$\frac{(Q_1 \times Q_2) + (Q_1 \times Q_3)}{Q_1 \times (Q_2 + Q_3)}$$

Comment 2 *It is easy to show that by using the commutativity and associativity (to the left) rules in Definition 3.2, one can efficiently simulate the associativity to the right rules.*

Definition 3.3 (Analytic symbolic proofs of polynomial identities) *Let Φ and Φ' be two arithmetic formulas (computing the same polynomial in $\mathbb{F}[x_1, \ldots, x_n]$). An* analytic symbolic derivation of Φ' from Φ *is a sequence of arithmetic formulas $\psi_1, \ldots, \psi_m$ such that $\psi_1 \equiv \Phi$, $\psi_m \equiv \Phi'$, and for all $1 < i \leq m$, ψ_i is derived from ψ_{i-1} by applying the polynomial-ring analytic derivation-rules (applicable on any subformula) from Definition 3.2. If the initial formula Φ computes the zero polynomial and Φ' is the formula 0, then we call such a derivation of Φ' from Φ an* analytic symbolic proof of Φ.

The ***length*** *of an analytic symbolic proof (or derivation)* is defined to be the number of proof-lines in the proof (derivation, respectively).

We shall prove the completeness (and soundness) of analytic symbolic proofs in Section 5.

Discussion about analytic symbolic proofs and the subformula property. Recall that we aim at formulating *analytic* symbolic proofs (Section 1.2), that is, a system that enjoys a sort of subformula property. The intuitive interpretation of this property in our setting would be to prevent the prover from using "clever tricks" (or "detours") when proving polynomial identities, in the sense that one cannot introduce new algebraic formulas that might help in efficiently proving the identities (like, introducing new monomials or new formulas, later to be cut-off in the proof).

The subformula property usually states that every formula that appears in an analytic proof sequence π of T also appears in the terminal proof-line T. In our setting this should mean that the consequence (i.e., lower-line) in any rule may only contain subformulas already appearing in the premise (i.e., upper-line) of the rule. (Note that in a standard sequent calculus proof, the proof starts with the axioms and terminates in a tautology; this should be analogous, in our setting, to a symbolic proofs of an arithmetic formula computing the zero polynomial *taken backward*: one starts from the "axiom" formula 0 and develops the formula computing the zero-polynomial; thus, whereas the subformula criterion usually requires that every formula in an upper-line of a rule occurs as a subformula in the lower-line of the rule, we should require that every (sub)formula in a lower-line of a rule should appear in some sense in the upper-line of the rule.) However, in our system we cannot follow precisely this requirement, since the two distributivity rules might change the structure of subformulas (for instance, $(Q_1 \times Q_2)$ in the lower-line of the forward distributivity rule is not a subformula of the upper-line $Q_1 \times (Q_2 + Q_3)$). Nevertheless, analytic symbolic proofs keep a relaxed subformula property (a "sub-monomial property", so to speak), as we now explain.

We say that a monomial is *syntactically computed by an arithmetic formula* Φ if it occurs in the set of monomials that results when expanding all the monomials in Φ while using *no canceling of monomials* (no canceling of monomials occurs also in any gate of Φ, not just the root gate).[3] In analytic symbolic proofs we have the following "sub-monomial property": If π is an analytic symbolic proof and the formula Φ is a proof-line (*not a proper* subformula in a proof-line) in π, then every monomial that is syntactically computed by Φ either is syntactically computed by the initial proof-line (again, not a proper subformula of the initial proof-line), or is the sum of (two or more) monomials that are syntactically computed in the initial proof-line. This way, *the number of monomials syntactically computed by each proof-line does not increase along the proof sequence*. We shall not prove this statement formally here, and so this idea should only be kept as an intuition in the mind of the reader.

Consider the following three (sound) derivation rules (*absent* from our system):

(i) $\dfrac{Q_1}{Q_1 + 0}$; (ii) $\dfrac{0}{0 \times Q_1}$; (iii) $\dfrac{\alpha}{\alpha_1 + \alpha_2}$ (where $\alpha = \alpha_1 + \alpha_2$, for $\alpha_1, \alpha_2, \alpha \in \mathbb{F}$).

We explain now how the choice not to include the above three rules helps us in keeping the relaxed form of the subformula property in our proof system. First, given Φ, one cannot simply derive $\Phi + \Delta - \Delta$ from Φ, where Δ is any non-empty formula. Note that for this derivation to

[3] A monomial here means a product of variables *with its scalar coefficient*.

be possible, we would need the following derivation sequence that uses rules (i) and (ii) above:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Phi}{\Phi+0}\text{ apply rule (i)}}{\Phi+0\times\Delta}\text{ apply rule (ii)}}{\Phi+(1-1)\times\Delta}}{\Phi+1\times\Delta+(-1)\times\Delta}}{\Phi+\Delta+(-1)\times\Delta}\ .$$

Second, if we had the rule (iii) in our proof system it would be possible to add arbitrary number of new monomials that are syntactically computed by proof-lines throughout the proof, as the following example illustrates:

$$\dfrac{\dfrac{3\times(x_1\times x_2)}{(2+1)\times(x_1\times x_2)}\text{ apply rule (iii)}}{2\times(x_1\times x_2)+1\times(x_1\times x_2)}$$
$$\cdots$$

3.1 Depth-3 formulas and depth-3 analytic symbolic proofs

We now consider analytic symbolic proofs of polynomial identities operating with depth-3 arithmetic formulas. The standard definition of depth-3 (and specifically $\Sigma\Pi\Sigma$) arithmetic formulas includes all formulas that can be written as a sum of products of linear polynomials. In other words, according to the standard definition, a $\Sigma\Pi\Sigma$ formula Φ (in the variables $x_1,\ldots,x_n$) can be written as:

$$\Phi\equiv\sum_{i=1}^{m}\prod_{j=1}^{d_i}L_{ij}\,, \tag{1}$$

where the L_{ij}'s are linear polynomials in the variables $x_1,\ldots,x_n$.

Due to the syntactic nature of our symbolic proof system, we require that a field element $\alpha\in\mathbb{F}$ that multiplies a (polynomial computed by a) formula f is written as $\alpha\times f$, where the product gate $\times$ is written explicitly. This makes, in our setting, the polynomial in (1) to be a depth-4 formula, that is a $\Sigma\Pi\Sigma\Pi$ (the reason is that variables inside a linear polynomial L_{ij} might have coefficients, which makes L_{ij} a $\Sigma\Pi$ formula). In order to include polynomials of the form shown in Equation (1) in our depth-3 proof systems we define the following class of formulas.

Definition 3.4 ($\hat{\Sigma}$ and $\Sigma\Pi\hat{\Sigma}$ formulas) *A $\hat{\Sigma}$ formula is a $\Sigma\Pi\Sigma$ arithmetic formula (according to Definition 2.1 and Section 2.1.1), such that the bottom $\Pi\Sigma$ level may include only field elements or products of a single variable with a sum (of zero or more) field elements. (Accordingly a $\Sigma\Pi\hat{\Sigma}$ formula is a $\Sigma\Pi\Sigma\Pi\Sigma$ arithmetic formula where the bottom $\Sigma\Pi\Sigma$ level is $\hat{\Sigma}$.)*

Example: The formula $(2+4+1)\times x_1+3\times x_2+(1+2)\times x_3+1$, is a $\hat{\Sigma}$ formula. The formula 3×2 is *not* a $\hat{\Sigma}$ formula.

Thus, a $\hat{\Sigma}$ formula is a formula computing a linear form (we need to include sums of fields elements as coefficients and not just a single field element as a coefficient, since this will enable us to add two linear forms using only $\hat{\Sigma}$ formulas). Note that indeed any sum of products of linear polynomials can be computed by a $\Sigma\Pi\hat{\Sigma}$ formula.

We conclude that when dealing with depth-3 proof systems we will in fact assume that all formulas in the proofs are $\Sigma\Pi\hat{\Sigma}$ formulas.

Definition 3.5 (Depth-3 analytic symbolic proofs) *A* depth-3 analytic symbolic derivation (proof) *is an analytic symbolic derivation (proof, respectively) in which every proof-line is a* $\Sigma\Pi\hat{\Sigma}$ *formula.*

We shall prove the completeness of depth-3 analytic symbolic proofs in Section 5 (this is done by proving the completeness of a fragment of depth-3 analytic symbolic proofs).

Example: A typical application of the backward distributivity rule inside depth-3 symbolic proofs proceeds according to the following scheme:

$$\frac{\Delta + \prod_{j=1}^{d} L_j \times L' + \prod_{j=1}^{d} L_j \times L''}{\Delta + \prod_{j=1}^{d} L_j \times (L' + L'')},$$

where L', L'' and the L_j's are formulas of linear polynomials in the variables $x_1, \ldots, x_n$ and Δ is any possibly empty formula (note that the two occurrences of $\prod_{j=1}^{d} L_j$ in the upper-line must be *identical*).

(When exchanging the upper- and lower- lines, we get a typical application of the *forward* distributivity rule.)

4 The Structure of Symbolic Proofs

In this section we develop terminology and concepts for dealing with structural properties of symbolic proofs. The notions developed here are suitable mainly for dealing with small depth symbolic proofs, as the notions mainly take into account the top gates of the formulas in the proofs. This will suffice for stating and proving our main lower bounds.

We will identify a certain graph structure induced by symbolic proofs. The idea is to partition the graph into levels, each level corresponds to a proof-line. Each node in level i corresponds to one summand of the formula in the ith line of the proof. From each summand (that is, a vertex in the graph) there will be edge(s) directed to its "immediate ancestor(s)" in the previous line. The formal definition follows.

Definition 4.1 (Underlying graph of analytic symbolic proof) *Let* $\pi = (\Phi_1, \ldots, \Phi_m)$ *be an analytic symbolic proof. Define the* underlying directed acyclic graph of π *denoted* G_π *as follows. The graph* G_π *has* m *levels and edges may only go from level* i *to level* $i-1$*, for* $1 < i \leq m$ *(the vertices in level* 1 *have no outgoing edges). For any* $1 \leq i \leq m$*, write* Φ_i *as* $\varphi_1 + \ldots + \varphi_{\ell_i}$*, for some* $\ell_i \geq 1$*, where every* φ_j *is either a product formula or a formula containing only a single leaf. Then, the* i*th level of* G_π *consists of* ℓ_i *vertices, each vertex is labeled by a different summand* φ_j *from* Φ_i *and if* $1 \leq i < m$ *then the incoming edges of level* i *(from vertices in level* $i+1$*) are defined as follows (we shall sometime abuse notation by identifying a vertex in the graph* G_π *with its label and a level in the graph* G_π *with its corresponding proof-line in* π*):*

(i) In case no rule was applied on φ_j *in level* i*, for* $j \in [\ell_i]$ *(see notation after Definition 3.1), then* φ_j *appears (as a single vertex) in both level* i *and level* $i+1$*, and we put an incoming edge from* φ_j *in level* $i+1$ *to* φ_j *in level* i*.*

(ii) *Assume that a rule different from the forward distributivity rule was applied on φ_j in level i, for $j \in [\ell_i]$. Since φ_j is in a separate vertex in G_π and some rule was applied on φ_j (in the ith step in π) then it must be that φ_j is a product formula (and not a single leaf formula). It can be verified by a straightforward inspection of the derivation rules (Definition 3.2) that a consequence of any rule different from the forward distributivity rule is not a plus formulas, and so the consequence of φ_j in proof-line $i+1$ in π is (still) a product formula. This implies that there is a single vertex φ'_j in level $i+1$ which is the consequence of φ_j. We define φ_j in level i to have a single incoming edged from φ'_j in level $i+1$.*

(iii) *In case the forward distributivity rule was applied on φ_j in level i, for $j \in [\ell_i]$, on the root gate of φ_j (again, see notation after Definition 3.1), then the consequence of φ_j in level $i+1$ is a sum of two formulas denoted $\psi_0+\psi_1$. Thus (by definition of the vertices in G_π), level $i+1$ contains two vertices ψ_0 and ψ_1, and we define φ_j to have two incoming edges, one from ψ_0 and one from ψ_1.*

(iv) *In case the forward distributivity rule was applied in φ_j in level i, for $j \in [\ell_i]$, but not on the root gate of φ_j, then the consequence of φ_j in level $i+1$ must be a* product formula *denoted φ'_j. We define φ_j in level i to have a single incoming edged from φ'_j in level $i+1$.*

(v) *In case the backward distributivity rule was applied on φ_j and φ_{j+1} in level i, for $j \in [\ell_i - 1]$ (note that the backward distributivity rule must be applied on a plus formula, and so it must involve two vertices in level i of G_π), then the consequence of $\varphi_j + \varphi_{j+1}$ in level $i+1$ is a product formula denoted ψ. We define φ_j to have an incoming edge from ψ in level $i+1$.*

Notation: For a vertex v in G_π we denote by $\dot{v}$ the formula that labels v and by $\mathsf{level}(v)$ the level of G_π in which v occurs.

Example. The following derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{((x_1 \times x_2) + (2 \times (x_1 + (x_1 + 3)) + (-1 \times (2 \times (x_1 + 3) + x_1 \times (x_2 + 2))))}{((x_1 \times x_2) + (2 \times x_1) + (2 \times (x_1 + 3)) + (-1 \times (2 \times (x_1 + 3) + x_1 \times (x_2 + 2))))}}{(x_1 \times x_2) + (2 \times x_1) + (2 \times (x_1 + 3)) + (-1 \times (2 \times (x_1 + 3))) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times x_2) + (2 \times x_1) + (2 \times (x_1 + 3)) + ((-1 \times 2) \times (x_1 + 3)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times x_2) + (2 \times x_1) + (2 \times (x_1 + 3)) + (-2 \times (x_1 + 3)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times x_2) + (x_1 \times 2) + (2 \times (x_1 + 3)) + (-2 \times (x_1 + 3)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times (x_2 + 2)) + (2 \times (x_1 + 3)) + (-2 \times (x_1 + 3)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times (x_2 + 2)) + ((2 - 2) \times (x_1 + 3)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times (x_2 + 2)) + (0 \times (x_1 + 3)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(x_1 \times (x_2 + 2)) + (-1 \times (x_1 \times (x_2 + 2)))}}{(1 \times (x_1 \times (x_2 + 2))) + (-1 \times (x_1 \times (x_2 + 2)))}}{((1 - 1) \times (x_1 \times (x_2 + 2)))}}{(0 \times (x_1 \times (x_2 + 2)))}}{0}$$

has the corresponding graph structure shown in Figure 1 (we ignore the associativity rule applications).

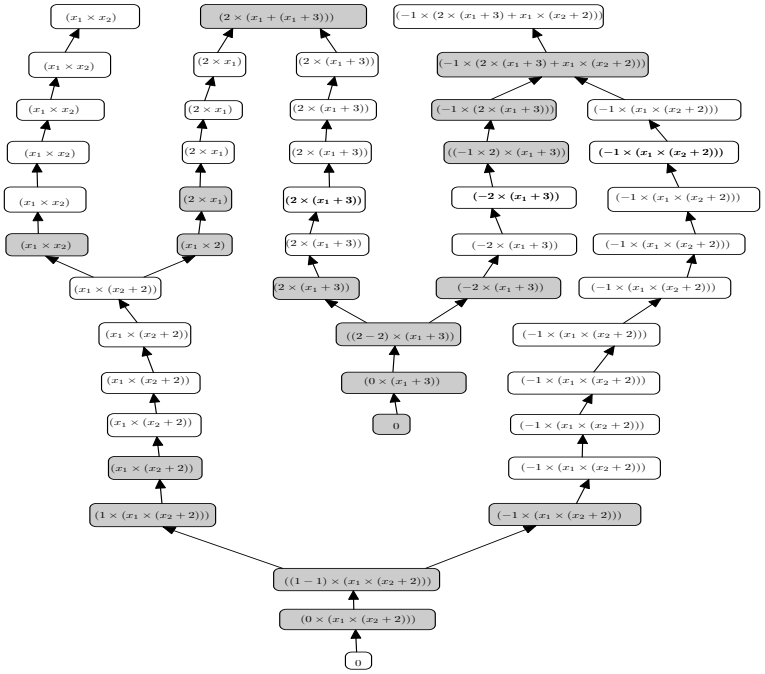

Figure 1: Underlying graph of an analytic symbolic proof (this proof is also regular [see Definition 4.4]). A single shaded vertex in a level means that a derivation rule is applied on this vertex (that is, an application rule is applied on one of the gates [formally, nodes] in the formula labeling the vertex). Two shaded vertices in a level means that the backward distributivity rule is applied in π on the plus gate (in the corresponding proof-line in π) that has the two shaded vertices as its children.

4.1 Regular Analytic Symbolic Proofs

We define here a fragment of analytic symbolic proofs, called *regular analytic symbolic proofs*, which mimics to some extent a tree-like structure on analytic symbolic proofs. We shall use the following two definitions for that purpose.

Atomic formulas are essentially formulas computing single monomials:

Definition 4.2 (Atomic and non-atomic formulas) *A formula Φ is* atomic *if it has the form $\phi_1 \times \cdots \times \phi_k$, for $k \geq 1$, where each ϕ_i ($1 \leq i \leq k$) is either a variable or a sum of one or more field elements; otherwise, Φ is said to be* non-atomic.

Example: The formulas $(1 + 2) \times x_2 \times x_1$ and $1 \times 3 \times x_1$ are atomic formulas, as well as the formula 0 and the formula x_1. The formula $x_2 + 3$ is a non-atomic formula.

Definition 4.3 *Given a proof π and its underlying graph G_π let G' be the subgraph of G_π induced by considering only the non-atomic vertices in G_π, and let v be a vertex in G'. Then, $\mathbb{T}_v(G_\pi)$ is defined to be the* directed *subgraph of G' induced by all the vertices (in G') reachable from v via a directed-path in G' (including v itself).*

The idea of the regularity condition of analytic symbolic proofs we are about to define is to guarantee that if a forward distribution rule is applied on $A \times (B + C)$, which breaks the formula into the sum of two terms $A \times B$ and $A \times C$, then the two formulas $A \times B$ and $A \times C$, as well

as any other two formulas that originate from $A \times B$ and $A \times C$ (possibly also originating by other formulas), might not be united together again by means of the backward distributivity rule (in fact, this rule is the only rule that can "unite" two separate summands into a product formula) .

In the terminology of the graph structure of analytic symbolic proofs (Definition 4.1) the regularity condition means that a vertex v in a proof-graph cannot have two edge-disjoint (directed) paths starting in v and leading to the same vertex.

Definition 4.4 (Regular analytic symbolic proofs) *Let π be an analytic symbolic proof. We say that π is a* regular analytic symbolic proof *(or* regular proof *for short), if for every (non-atomic) vertex v in G_π, the subgraph $\mathbb{T}_v(G_\pi)$ is a* tree.[4]

In other words, analytic symbolic proofs are regular if for every (non-atomic) vertex w in their underlying proof-graph there are no two distinct directed paths originating in w that reach a common vertex (different from w).

Example: Figure 1 illustrates a regular analytic symbolic proof. One can verify that for every (non-atomic) vertex w in the graph there are no two distinct directed paths originating in w that reach a common vertex.

Comment 3 *In the definition of regular proofs we need to specifically refer to non-atomic vertices (through the definition of $\mathbb{T}_v(G_\pi)$) for the following reason. If for* no *vertex v in G_π does one get by two distinct (directed) paths from v to a vertex corresponding to an initial summand (in the initial formula), then some formulas computing the zero polynomial might not be provable. Consider for example the formula $(x_1 + 1) \times (x_1 - 1) - x_1 \times x_1 - 1$. This formula computes the zero polynomial. However, the first term (from the left) is a product formula that when expanded, contains the two canceling monomials $x_1 \times -1$ and $1 \times x_1$. Thus, in order to reach the formula 0 we must enable these two terms to cancel each other (that is, to derive the (sub-)formula 0 from them). Hence, this amounts to a vertex labeled $((1-1) \times x_1)$ in G_π from which two distinct paths lead to the same initial summand $(x_1 + 1) \times (x_1 - 1)$.*

We use the following simple structural proposition in the next sections. Essentially, the proposition states that if there is a derivation starting with $\Delta + \Theta$ and terminates in $\Delta' + \varphi$, and φ in the terminal line was "originated" only from Θ in the initial line, then there is a derivation starting in Θ and terminating in $\Delta'' + \varphi$ for some possibly empty formula Δ''.

Proposition 1 *Let Θ and Δ be two formulas, and assume that φ is a product formula (or a single leaf formula) such that there is a regular analytic symbolic derivation π starting from $\Delta+\Theta$ and deriving $\Delta'+\varphi$, where Δ' is any possibly empty formula. Denote by G_π the underlying graph of π, and let v be the vertex in the maximal level of G_π that is labeled with φ. Let $\mathbf{T}$ be the set of vertices of $\mathbb{T}_v(G_\pi)$, let $\mathbf{A}$ be the set of all the vertices in level 1 of G_π, and let $\mathbf{B}$ be the set of all the vertices that are labeled with summands of Θ in the first level of G_π. If $(\mathbf{T} \cap \mathbf{A}) \subseteq \mathbf{B}$, then there is a regular analytic symbolic derivation π'* starting from Θ alone *and deriving $\Delta'' + \varphi$, where Δ'' is any possibly empty formula.*

Proof: Starting from $\mathbb{T}_v(G_\pi)$, we construct a new graph G' that is a legitimate underlying graph of a regular analytic symbolic proof of $\Delta'' + \varphi$, where Δ'' is any possibly empty formula (we ignore the order of additions).

[4]The intention here, of course, is that $\mathbb{T}_v(G_\pi)$ is a tree also when considered as an undirected graph, that is, when replacing every directed edge in $\mathbb{T}_v(G_\pi)$ with an undirected one.

Let $G' := \mathbb{T}_v(G_\pi)$.

Step I: For every $u \in \mathbf{B} \setminus \mathbf{T}$, we add to every level $1 \leq j \leq \mathsf{level}(v)$ in[5] G' a new copy of u, and put an edge from u in level j to u in level $(j-1)$ in G' (for $1 < j \leq \mathsf{level}(v)$).

Step II: For all vertices w in $\mathbb{T}_v(G_\pi)$ such that $w \neq v$, we do the following. Let $\ell = \mathsf{level}(w)$. By definition of $\mathbb{T}_v(G_\pi)$ there is a vertex $u \in \mathbf{T}$ such that w has an incoming edge from u in $\mathbb{T}_v(G_\pi)$ (since there is a directed path from v to w in $\mathbb{T}_v(G_\pi)$). Assume that there is a vertex $r \in G_\pi \setminus \mathbf{T}$ such that w has an incoming edge from r. Then,

(i) we add the vertex r to level $\ell + 1$ in G'; and

(ii) we add an edge from r to w in G'; and

(iii) we add to every level $j > \ell + 1$ in G' a new copy of r, and put an edge from r in level j to r in level $(j-1)$ in G'.

We claim that G' is the underlying graph of a regular analytic symbolic derivation starting from Θ and deriving $\Delta'' + \varphi$, for some possibly empty formula Δ''. Note that the first level of G' consists precisely of the summands of Θ, and that the last level of G' contains the vertex v. Moreover, G' conforms to the regularity conditions: no vertex in G' has two distinct (non trivial) directed paths that reach the same vertex in G'.

Thus, we only need to show that G' constitutes a legal analytic symbolic derivation. To this end, it suffices to show that every level $j > 1$ in G' corresponds to a proof-line that is derivable from the previous level $j-1$, and that the edges from vertices in level j to the vertices in level $j-1$ in G' are legal (that is, conform to the definition of an underlying graph of an analytic symbolic proof):

Let w be a vertex in level $j > 1$ in G', and consider the following cases:

1. If $w \notin \mathbf{T}$ and w was added in Step I to G', then w was added to G' along with a single outgoing edge that points into another copy of w in level $j-1$ in G', and so we are done.

2. If $w \in \mathbf{T}$, then, by definition, there is an edge from w to another vertex $u \in \mathbf{T}$ and hence u is in G'. In case u has fan-in 1 in G_π we are done.

   Otherwise, there is a (unique) vertex $r \neq w$ in G_π such that r has an edge e pointing to u in G_π. By the construction of G' (Step II, items (i) and (ii) above), both r and the edge e are in G', and so both w and r and their outgoing edges correspond to a legal derivation. This takes care of all the vertices in level j in G' that are not in $\mathbf{T}$, *as well as all the vertices r added to G' in level j, in Step II, item (i) above.*

3. If $w \notin \mathbf{T}$ and w was added to G' in Step II, item (iii) above, then w has a single outgoing edge that points to a copy of w in level $j-1$ of G', and we are done.

■

Comment 4 *By inspecting the proof of Proposition 1, it is evident that the statement of this proposition also holds for regular analytic symbolic proofs* operating with depth-3 formulas. *We shall use this fact in the sequel.*

5 Completeness

Recall that when dealing with depth-3 formulas computing the zero polynomial, we assume that the formulas are in fact $\Sigma\Pi\hat{\Sigma}$ formulas (Definition 3.4).

[5] Clearly, every vertex in $\mathbb{T}_v(G_\pi)$ retain the same level it had in G_π.

Theorem 2 (Completeness of regular analytic depth-3 symbolic proofs) *Regular analytic depth-3 symbolic proofs are* complete *in the sense that every depth-3 formula computing the zero polynomial has such a proof; and* sound *in the sense that only formulas computing the zero polynomial have such proofs.*

Proof sketch: Soundness stems from the soundness of the derivation rules: every derivation rule that is applied on some formula Φ yields a formula computing the same polynomial as Φ. We turn to proving completeness.

We first expand all the (formulas of) monomials (including their coefficients, and with no cancelations applied anywhere) in each summand in the initial $\Sigma\Pi\hat{\Sigma}$ formula. This step requires *only the successive application of the forward distributivity rule*. In particular, assume that $\Phi \equiv \psi_1 + \ldots + \psi_k$, for some $k \geq 1$, is a $\Sigma\Pi\hat{\Sigma}$ formula computing the zero polynomial, where each ψ_i is a $\Pi\hat{\Sigma}$ formula. Then, applying the forward distributivity rule on the root product gate of a summand ψ_i (for $i \in [k]$), *yields a new* $\Sigma\Pi\hat{\Sigma}$ *formula*.

Next we derive the 0 formula from every group of canceling monomials. Thus, we arrive at a sum of 0's formulas. This can be further reduced into a single 0 by the zero element rule.

By simple inspection of the underlying graph of the above proof it is easy to see that the graph obtained is *regular*: Every non-atomic vertex has only a single (directed) path that ends at the vertex labeled with a summand from the initial line. Thus, every non-atomic vertex v trivially induces a tree, via $\mathbb{T}_v(G_\pi)$, in the underlying graph, which precisely means that the proof is regular. ∎

In the same manner we can get the completeness and soundness properties for unrestricted depth regular symbolic proofs for the set of all formulas that compute the zero polynomial.

Theorem 3 *Analytic symbolic proofs are sound and complete for the set of formulas computing the zero polynomial.*

Proof: Soundness stems from the soundness of the derivation rules. For completeness, we first expand all the (formulas of) monomials (including their coefficients, and with no cancelations applied anywhere) in each summand in the initial formula (and without taking care for the depth of formulas). Then we proceed as in the proof of Theorem 2. ∎

6 Lower Bounds

In this section we demonstrate exponential-size lower bounds on the length of regular analytic symbolic proofs operating with depth-3 formulas of certain polynomial identities. The hard instances will be built from small depth-3 formulas for the elementary symmetric polynomials over large enough fields (the. construction of such formulas is due to Ben-Or [cf. Shpilka and Wigderson (2001)]).

6.1 Hard Formulas

We shall consider from now on all arithmetic formulas to compute polynomials over the field of complex numbers $\mathbb{C}$. The following depth-3 formula will serve as our hard instance.

$$\begin{aligned}\mathsf{Sym}_n :\equiv r_0 \times (x_1 + b_0) \times (x_2 + b_0) \times \cdots \times (x_n + b_0) \quad &+ \\ r_1 \times (x_1 + b_1) \times (x_2 + b_1) \times \cdots \times (x_n + b_1) \quad &+ \\ \cdots \quad &\\ r_n \times (x_1 + b_n) \times (x_2 + b_n) \times \cdots \times (x_n + b_n) \quad &-1\,,\end{aligned} \tag{2}$$

where $r_0, \ldots, r_n$ are some $n+1$ complex numbers and $b_0, b_1, \ldots, b_n$ are $n+1$ *distinct* non-zero complex numbers (each of which is different from 1).

The jth summand in Sym_n, for $0 \leq j \leq n$, is a $\Pi\Sigma$ formula (and hence specifically a $\Pi\hat{\Sigma}$ formula) and is denoted $A_{n,j}$. In other words,

$$A_{n,j} \equiv r_j \times (x_1 + b_j) \times (x_2 + b_j) \times \ldots \times (x_n + b_j)\,, \tag{3}$$

and so

$$\mathsf{Sym}_n \equiv \sum_{j=0}^{n} A_{n,j} \; - 1\,.$$

Proposition 4 *There exist complex numbers $r_0, \ldots, r_n$ and $n+1$ distinct nonzero complex numbers $b_0, \ldots, b_n$ (each of which is different from 1), so that Sym_n computes the zero polynomial.*

The proof of Proposition 4 is given in Section 6.5.

6.2 Key Lemma: Lower Bounds on Deriving One Formula from Another

In this subsection we show a lower bound on the length of regular analytic symbolic derivations operating with $\Sigma\Pi\hat{\Sigma}$ formulas needed to derive a certain formula from another formula. Specifically, we provide two formulas Φ and Ψ, such that starting from Φ one cannot efficiently derive any formula that *contains* Ψ *as a subformula*. This will then facilitate us in the next subsection (when proving lower bounds on the length of proofs of formulas computing the zero polynomial, that is, on the derivation length needed to reach the formula 0). Note that the task of deriving a certain formula from a different formula, and the task of proving that a formula computes the zero polynomial (by deriving the zero formula) are not necessarily identical tasks since our derivation rules are asymmetric (Definition 3.2).[6]

We work in depth-3 from now on, and specifically with $\Sigma\Pi\hat{\Sigma}$ formulas.

Definition 6.1 (Derivable subformulas) *Let Φ be a formula. The collection of all formulas Ψ for which there is an analytic symbolic derivation of Ψ from Φ are said to be the formulas* derivable from Φ. *If B is the set of all formulas derivable from Φ then the set of all* subformulas *of formulas in B is called the set of* subformulas derivable from Φ, *denoted* $\mathsf{deriv}(\Phi)$.

Example: Let Φ be the Σ formula $(x_i + b_1)$. Then, the derivable formulas from Φ are, e.g., $(x_i + b_1)$, $(b_1 + x_i)$ and all $\hat{\Sigma}$ formulas $\prod_{i \in I} c_j \times \left(\prod_{k \in K} a_j \times x_i + \prod_{j \in J} b_j\right)$, where $\prod_{i \in I} c_i =$

[6] To see this, observe that deriving ψ_1 from ψ_2 means that we can derive $\psi_1 - \psi_1$ from $\psi_2 - \psi_1$, and thus we can derive 0 from $\psi_2 - \psi_1$. On the other hand, deriving 0 from $\psi_2 - \psi_1$ does not imply that we can derive ψ_1 from ψ_2 (note that we cannot start from ψ_2, add $-\psi_1 + \psi_1$ to yield $\psi_2 - \psi_1 + \psi_1$, and then derive 0 from the first two summands to yield ψ_1; this is because we would need a rule to introduce the formula $-\psi_1 + \psi_1$ in the first step, and we do not have such a rule in analytic symbolic proofs).

$\prod_{k\in K} a_k = 1$ and $\prod_{j\in J} b_j = b_1$. Note that $(x_i + b_2 + b_3)$ is *not* derivable from $(x_i + b_1)$, even when $b_2 + b_3 = b_1$ is true in the field; this stems from the definition of our derivation rules (Definition 3.2). Further, x_i, b_1 and $(x_i + b_1)$, for instance, are in $\mathsf{derive}(x_i + b_1)$, while $(x_i + b_0) \notin \mathsf{derive}(x_i + b_1)$ (in case $b_0 \neq b_1$) (see Proposition 9 for a proof).

The following definition is basically the transitive closure of a formula under "forward" and "backward" applications of all rules excluding the two distributivity rules (unless the distributivity rules are applied inside $\hat{\Sigma}$ formulas).

Definition 6.2 (Simple descendants and simple ancestors) *Given a formula Ψ we define $\mathsf{Cl}(\Psi)$ as the smallest set of formulas containing Ψ and satisfying the following: Assume that Δ is a formula and v is some node in the (tree of) Δ, then:*

(i) *If $\Delta\{\phi + 0\}_v \in \mathsf{Cl}(\Psi)$, then $\Delta\{\phi\}_v \in \mathsf{Cl}(\Psi)$;*

(ii) *If $\Delta\{\alpha_1 + \alpha_2\}_v \in \mathsf{Cl}(\Psi)$, then $\Delta\{\alpha\}_v \in \mathsf{Cl}(\Psi)$ for $\alpha, \alpha_1, \alpha_2 \in \mathbb{F}$, such that $\alpha = \alpha_1 + \alpha_2$;*

(iii) $\Delta\{\phi \times 1\}_v \in \mathsf{Cl}(\Psi)$ *iff* $\Delta\{\phi\}_v \in \mathsf{Cl}(\Psi)$ *;*

(iv) $\Delta\{\alpha_1 \times \alpha_2\}_v \in \mathsf{Cl}(\Psi)$ *iff* $\Delta\{\alpha\}_v \in \mathsf{Cl}(\Psi)$ *, for $\alpha, \alpha_1, \alpha_2 \in \mathbb{F}$ such that $\alpha = \alpha_1 \times \alpha_2$;*

(v) $\Delta\{\phi_2 \circ \phi_1\}_v \in \mathsf{Cl}(\Psi)$ *iff* $\Delta\{\phi_1 \circ \phi_2\}_v \in \mathsf{Cl}(\Psi)$ *, where $\circ \in \{+, \times\}$;*

(vi) $\Delta\{(\phi_1 \circ \phi_2) \circ \phi_3\}_v \in \mathsf{Cl}(\Psi)$ *iff* $\Delta\{\phi_1 \circ (\phi_2 \circ \phi_3)\}_v \in \mathsf{Cl}(\Psi)$ *, where $\circ \in \{+, \times\}$;*

(vii) $\Delta\left\{\left(\sum_{j\in J} \alpha_j + \sum_{k\in K} \alpha_k\right) \times x_i\right\}_v \in \mathsf{Cl}(\Psi)$ *(where $|J|, |K| \geq 1$ and for all $j \in J$ and $k \in K$, $\alpha_j, \alpha_k \in \mathbb{F}$) iff* $\Delta\left\{\left(\sum_{j\in J} \alpha_j\right) \times x_i + \left(\sum_{k\in K} \alpha_k\right) \times x_i\right\}_v \in \mathsf{Cl}(\Psi)$ *;*

(viii) $\Delta\left\{\left(\sum_{j\in J} \alpha_j\right) \times (x_i + x_j)\right\}_v \in \mathsf{Cl}(\Psi)$ *iff*
$\Delta\left\{\left(\sum_{j\in J} \alpha_j\right) \times x_i + \left(\sum_{j\in J} \alpha_j\right) \times x_j\right\}_v$ *, where $\alpha_j \in \mathbb{F}$ (for all $j \in J$) and $|J| \geq 1$.*

Whenever $\Psi' \in \mathsf{Cl}(\Psi)$ we call Ψ' a simple descendant of *Ψ , and Ψ a* simple ancestor of *Ψ'. Given a formula Ψ we denote by $\mathsf{Cl}^-(\Psi)$ the set of all simple ancestors of Ψ, that is, the set of all Φ such that $\Psi \in \mathsf{Cl}(\Phi)$.*

Note: Observe that $\mathsf{Cl}^-(\Psi)$ is closed under the specified derivation rules when they are applied "*backward*" on the formulas in $\mathsf{Cl}^-(\Psi)$. Also note that the only items in Definition 6.2 that are asymmetric are items (i) and (ii).

Comment 5 *The last two clauses (vii) and (viii) in Definition 6.2 intend to deal with distributivity rules applied inside $\hat{\Sigma}$ formulas only (and those distributivity rule applications whose consequence is a $\hat{\Sigma}$ formula).*

We shall use the following abbreviation (this is identical to Sym_n when excluding the first summand $A_{n,0}$).

$$\begin{aligned}\mathsf{Init}_n :\equiv\ & r_1 \times (x_1 + b_1) \times (x_2 + b_1) \times \cdots \times (x_n + b_1) + \\ & r_2 \times (x_1 + b_2) \times (x_2 + b_2) \times \cdots \times (x_n + b_2) + \\ & \cdots \\ & r_n \times (x_1 + b_n) \times (x_2 + b_n) \times \cdots \times (x_n + b_n) \\ & -1\,.\end{aligned} \tag{4}$$

We thus have, by Equation (3):

$$\mathsf{Init}_n \equiv \sum_{j=1}^{n} A_{n,j} - 1\,. \tag{5}$$

Definition 6.3 (Proper $\hat{\Sigma}$ formulas) *A proper $\hat{\Sigma}$ formula is a $\hat{\Sigma}$ formula which is a plus formula (that is, it has a plus gate at the root). A proper $\Pi\hat{\Sigma}$ formula is a proper $\Pi\Sigma$ formula where the Σ formulas in the bottom levels are replaced by proper $\hat{\Sigma}$ formulas. Similarly, a proper $\Sigma\Pi\hat{\Sigma}$ formula is a proper $\Sigma\Pi\Sigma$ formula where the Σ formulas in the bottom levels are replaced by proper $\hat{\Sigma}$ formulas.*

Note that every proper Σ formula (Section 2.1.1] is a proper $\hat{\Sigma}$ formula, while not every proper $\hat{\Sigma}$ formula is a Σ formula.

Lemma 5 (Key lower bound) *Let π be a regular analytic symbolic proof operating with $\Sigma\Pi\hat{\Sigma}$ formulas, and with the initial formula in π being Init_n, for some positive $n \in \mathbb{N}$. Let G_π be the corresponding graph of π. Assume that v is a vertex in G_π labeled with Φ, which is a simple ancestor of*

$$\psi \times \prod_{k=1}^{m} \Psi_k\,,$$

where ψ is any possibly empty formula, and for every $k \in [m]$, $i \in [n]$ and $j \in [n]$, Ψ_k is a proper $\hat{\Sigma}$ formula, such that $\Psi_k \notin \mathsf{deriv}(x_i + b_j)$. Then $|\mathbb{T}_v(G_\pi)| \geq 2^m$.

Proof: We go by induction on m. The idea is to build inductively and in a bottom-up fashion, starting with v, the tree $\mathbb{T}_v(G_\pi)$, by considering all the possible rules that can derive the formula Φ.

Base case: $m = 1$. The formula Φ is a simple ancestor of $\psi \times \Psi_1$, where ψ is some possibly empty formula[7] and Ψ_1 is a *proper* $\hat{\Sigma}$ formula that is not a subformula derivable from $(x_i + b_j)$, for all $j \in [n]$ and all $i \in [n]$. Observe that (for any positive $n \in \mathbb{N}$) every proper $\hat{\Sigma}$ formula that occurs in Init_n has the form $(x_i + b_j)$, for some $j \in [n]$ and $i \in [n]$. Thus, Φ is different from every subformula in the initial formula Init_n, and so the number of proof-lines in π is at least 2, and we are done.

Induction case: We have that Φ is a simple ancestor of $\psi \times \prod_{k=1}^{m} \Psi_k$, where ψ is some possibly empty formula, $m > 1$ and the Ψ_k's are proper $\hat{\Sigma}$ formulas. We shall use the following main technical lemma (whose proof is given in Section 6.4).

Lemma 6 (Technical lemma) *If $m > 1$ then (under the conditions stated in Lemma 5) there exists a vertex y in $\mathbb{T}_v(G_\pi)$, such that there is a path from v to y in $\mathbb{T}_v(G_\pi)$, and y has two outgoing edges to two (distinct) vertices u, w, so that: u and w are labeled with two simple ancestors of the following product formulas*

$$\psi_0 \times \prod_{k=1}^{m-1} \Psi'_k \qquad \textit{and} \qquad \psi_1 \times \prod_{k=1}^{m-1} \Psi''_k\,, \tag{6}$$

respectively, where ψ_0, ψ_1 are some possibly empty formulas, and for all $k \in [m-1]$, $i \in [n]$ and $j \in [n]$, Ψ'_k, Ψ''_k are some proper $\hat{\Sigma}$ formulas such that $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$ and $\Psi''_k \notin \mathsf{deriv}(x_i + b_j)$.

[7] Note that in this case (i.e., when $m = 1$), since Ψ_1 is a proper $\hat{\Sigma}$ formula, ψ is in fact a non-empty formula, as by definition of G_π, the vertex v cannot be labeled with a plus formula.

Given Lemma 6, we can then use the induction hypothesis on the two vertices u, w (whose existence is guaranteed by the lemma), showing that $|\mathbb{T}_u(G_\pi)| \geq 2^{k-1}$ and $|\mathbb{T}_w(G_\pi)| \geq 2^{k-1}$. By the regularity condition we know that $\mathbb{T}_u(G_\pi)$ and $\mathbb{T}_w(G_\pi)$ have no common vertices (note that $\dot{v}$ is a non-atomic formula [Definition 4.2] and so indeed $\mathbb{T}_v(G_\pi)$ is a tree [Definition 4.4]). Therefore, $|\mathbb{T}_v(G_\pi)| \geq |\mathbb{T}_u(G_\pi)| + |\mathbb{T}_w(G_\pi)| \geq 2^k$, which concludes the proof of Lemma 5. ∎

In the next section we shall need the following simple genralization of Lemma 5:

Corollary 7 *Let π be a regular symbolic proof operating with $\Sigma\Pi\hat{\Sigma}$ formulas, and with the initial formula in π being $\Delta+\mathsf{Init}_n$, for some $\Sigma\Pi\hat{\Sigma}$ formula Δ and some positive $n \in \mathbb{N}$. Let G_π be the corresponding graph of π. Assume that v is a vertex in G_π labeled with Φ, which is a simple ancestor of*

$$\psi \times \prod_{k=1}^{m} \Psi_k \,,$$

where ψ is any possibly empty formula, and for every $k \in [m]$, $i \in [n]$ and $j \in [n]$, Ψ_k is a proper $\hat{\Sigma}$ formula, such that $\Psi_k \notin \mathsf{deriv}(x_i + b_j)$. Let $\mathbf{T}$ be the set of vertices of $\mathbb{T}_v(G_\pi)$ and let $\mathbf{A}$ be the set of all vertices in level 1 of G_π that are subformulas in Δ (i.e., those corresponding to summands in Δ in the initial line). If $\mathbf{T} \cap \mathbf{A} = \emptyset$, then $|\mathbb{T}_v(G_\pi)| \geq 2^m$.

Proof: Immediate from Lemma 5 and Proposition 1. ∎

6.3 Main Lower Bound: Regular Depth-3 Proofs

In this subsection we demonstrate how to use the Key Lemma 5 in order to prove a lower bound on the *proof* length of Sym_n, that is, a lower bound on the size of derivation starting with Sym_n and terminating with the formula 0.

The main result of this section is the following:

Theorem 8 *Every regular analytic symbolic proof operating with $\Sigma\Pi\hat{\Sigma}$ formulas of Sym_n has length $2^{\Omega(n)}$.*

Comment 6 *The number of variables in Sym_n is $\ell = n \cdot (n+1)$, and so Theorem 8 gives a lower bound of $2^{\Omega(\sqrt{\ell})}$.*

The rest of this subsection is dedicated to the proof of Theorem 8. The proof idea goes as follows. We begin by considering the summand $A_{n,0}$ in the first line of π. We expand $A_{n,0}$ by applying the forward distributivity rule as much as we like (without using the backward distributivity rule, but possibly with applying other rules). In case we expand too much (i.e., exponentially many) summands, we are done. Otherwise, the expansion of $A_{n,0}$ yields a summand Φ which is a simple ancestor of $\psi \times \prod_{k=1}^{m} \Psi_k$, where m is big enough (that is, $m \geq n/2$), and Φ conforms to the conditions in Key Lemma 5 *excluding* the fact that the initial line of the proof is *not necessarily* Init_n. Since, by assumption, no forward distributivity rule applications are applied on Φ the *backward* distributivity must be applied on it sometime. We can then show that in this case there exists a vertex v in the underlying graph G_π of π that *fully* conforms to the conditions stated in Key Lemma 5 (more correctly, v fully conforms to the conditions stated in Corollary 7), which concludes the lower bound proof.

For every $k \in [n]$, we put

$$\Psi_k :\equiv (x_k + b_0) \,.$$

Proposition 9 *For all $k, i, j \in [n]$, $\Psi_k \notin \mathsf{deriv}(x_i + b_j)$.*

Proof: By the definition of the derivation rules (Definition 3.2), no rule can increase the number of plus gates in a formula, except for the forward distributivity rule $\dfrac{Q_1 \times (Q_2 + Q_3)}{(Q_1 \times Q_2) + (Q_1 \times Q_3)}$: every plus gate that occurs in (a substitution instance of) Q_1 in the upper-line $Q_1 \times (Q_2 + Q_3)$, appears twice in the lower-line $(Q_1 \times Q_2) + (Q_1 \times Q_3)$. Note that for this increase in the number of plus gates to happen, the upper-line $Q_1 \times (Q_2 + Q_3)$ must contain at least two plus gates. Thus, if Φ' is derived from Φ via an analytic symbolic derivation, and Φ contains only one plus gate, then Φ' contains at most one plus gate. We conclude that for all $i, j \in [n]$, any formula derivable from $(x_i + b_j)$ contains at most one plus gate.

Assume by a way of contradiction that there is a formula $\varphi \in \mathsf{deriv}(x_i + b_j)$, such that there exists $k \in [n]$ for which $(x_k + b_0)$ occurs inside φ. Since φ has only one plus gate, $\varphi \equiv \phi \times (x_k + b_0)$, for some possibly empty formula ϕ (when ignoring associativity of formulas and the order of multiplication). By the soundness of the proof system, we get that $\phi \times (x_k + b_0)$ and $(x_i + b_j)$ compute the same polynomial. But, by assumption, $b_j \neq b_0$ for all $j \in [n]$, and so for all $k \in [n]$, $(x_i + b_j)$ cannot be factored by $(x_k + b_0)$. ∎

Fix a regular depth-3 analytic symbolic proof π of Sym_n, and let G_π be the corresponding underlying graph of π. Note that the first line of π is $\mathsf{Sym}_n \equiv A_{n,0} + \Sigma_{j=1}^{n} A_{n,j} - 1$, and so the first level of G_π has a unique vertex labeled with (the product formula) $A_{n,0}$. Let $\mathcal{S}$ be the set of all vertices v in G_π such that $\dot{v}$ is non-atomic (Definition 4.2) and every (directed) path that originates in v terminates in the vertex that is labeled with $A_{n,0}$. Let

$$\mathcal{T} \text{ be the subgraph of } G_\pi \text{ induced by the vertices in } \mathcal{S}\,.$$

Proposition 10 *The graph $\mathcal{T}$ is a binary tree rooted at $A_{n,0}$ (where every vertex is directed from leaves toward the root).*[8]

Proof: This stems directly from the regularity condition on the structure of π. Formally, assume by a way of contradiction that there is a (possibly undirected) cycle C in $\mathcal{T}$, and let u be a vertex in C, where $\mathsf{level}(u)$ is the maximal level in G_π. Since u is in a cycle it has two edges adjacent to two distinct vertices w, v in C. But then the fan-out of u in G_π (considered as a directed graph) is 2, as both v and w must be in a smaller level than the level of u (by assumption on the maximality of $\mathsf{level}(u)$, and since no two vertices in the same level of a proof-graph are connected with an edge). Both v and w have a directed path leading to $A_{n,0}$ in G_π, and so u has two different paths leading to $A_{n,0}$, in contrast to the regularity condition which forbids this case (since, $\mathbb{T}_u(G_\pi)$ is a tree, rooted at u). This shows that $\mathcal{T}$ is a tree. To show that $\mathcal{T}$ is a binary tree, we only need to note that by definition, every vertex in G_π has fan-in at most 2. ∎

Recall the definition of a subformula in a proof-line being *the consequence* of some derivation rule (Notation after Definition 3.1). Also recall that a derivation rule is applied *outside* a subformula ψ if (in the terminology used in Notation after Definition 3.1) the vertex v is not in ψ.

[8] $\mathcal{T}$ is a tree also when considered as an undirected graph.

Proposition 11 *No vertex in $\mathcal{T}$ is (labeled with) a consequence of the backward distributivity rule applied outside a $\hat{\Sigma}$ formula.*[9]

Proof: By definition no vertex in G_π is labeled with a plus formula. Hence, all vertices in $\mathcal{T}$ are labeled with (not necessarily proper) $\Pi\hat{\Sigma}$ formulas (since the proof is restricted to $\Sigma\Pi\hat{\Sigma}$ formulas only). The upper-line in the backward distributivity rule is $(Q_1 \times Q_2) + (Q_1 \times Q_3)$.

Assume, by a way of contradiction, that there is a vertex v in $\mathcal{T}$, such that v is a consequence of the backward distributivity rule applied on (a substitution instance of) $(Q_1 \times Q_2) + (Q_1 \times Q_3)$ and $(Q_1 \times Q_2) + (Q_1 \times Q_3)$ is not a $\hat{\Sigma}$ formula. In case $(Q_1 \times Q_2)$ and $(Q_1 \times Q_3)$ appear in two distinct vertices u, w in G_π, then we get that v has two outgoing edges pointing to u and w. By the definition of $\mathcal{T}$, this means that both u and w have directed paths reaching the root of $\mathcal{T}$, which contradicts the regularity condition.

Otherwise, $(Q_1 \times Q_2) + (Q_1 \times Q_3)$ occurs in the label of some *single* node w. Since no single node is labeled with a plus formula, it must be that w is labeled with a product formula that contains $(Q_1 \times Q_2) + (Q_1 \times Q_3)$ as a subformula. This means that w contains a proper $\Pi\Sigma\Pi$ formula, where the middle $\Sigma\Pi$ formula is not a $\hat{\Sigma}$ formula (since $(Q_1 \times Q_2) + (Q_1 \times Q_3)$ is not a $\hat{\Sigma}$ formula). A contradiction to assumption. ∎

The following proposition exploits our restriction to depth-3 formulas.

Proposition 12 *Let π be an analytic symbolic proof operating with $\Sigma\Pi\hat{\Sigma}$ formulas. Let v be a node in G_π labeled with a $\Pi\hat{\Sigma}$ formula and let $\ell = \mathsf{level}(v)$. If the forward distributivity rule is applied in $\dot{v}$ (in the ℓth proof-line of π), and the rule is applied outside a $\hat{\Sigma}$ formula and further the consequence of the rule application is not a $\hat{\Sigma}$ formula, then v must have fan-in 2.*

Proof: The idea is that a forward distributivity rule cannot be applied "inside" a product formula (excluding the case when it is applied inside a $\hat{\Sigma}$ formula, or when the consequence of the rule application is a $\hat{\Sigma}$ formula), as this would result in a formula having depth > 3.

Formally, the formula $\dot{v}$ can be written as $L_1 \times \cdots \times L_k$, for some $k \geq 1$, where each L_i is a $\hat{\Sigma}$ formula. If $k = 1$ then the proposition trivially holds, since every application of the forward distributivity rule would be applied inside a $\hat{\Sigma}$ formula, in contrast to the assumption.

Otherwise, $k > 1$. Suppose that an application of the forward distributivity rule on some subformula occurring in $\dot{v}$ was performed outside a $\hat{\Sigma}$ formula. A vertex labeled with a formula on which the forward distributivity is applied either has fan-in 1 or 2 (and not 0). Assume by a way of contradiction that v has fan-in 1 and let u be the vertex that points to v. By the definition of the forward distributivity rule (Definition 3.2) $\dot{u}$ must be a product formula (since every vertex is labeled either with a product formula or a single leaf; and a single leaf cannot be the consequence of the forward distributivity rule) that resulted from distributing some subformula $\prod_{i \in I} L_i$ over some $\hat{\Sigma}$ formula $L_r \equiv \Delta_1 + \Delta_2$, for some non-empty $I \subseteq [k]$. This means that u is labeled with

$$\left(\prod_{i \in I} L_i \times \Delta_1 + \prod_{i \in I} L_i \times \Delta_2 \right) \times \prod_{j \in J} L_j \,,$$

for $J = ([k] \setminus I) \setminus \{r\}$, where J is non-empty (as otherwise, u was a plus formula). Therefore, u is a $\Pi\Sigma\Pi$ formula where the middle $\Sigma\Pi$ level consists of formulas that are not $\hat{\Sigma}$ formulas (since

[9] Here we identify the vertices in $\mathcal{T}$ with the corresponding vertices in G_π (where the latter correspond, as always, to subformulas occurring in π).

by assumption the consequence of the rule application $\left(\prod_{i\in I} L_i \times \Delta_1 + \prod_{i\in I} L_i \times \Delta_2\right)$ is not a $\hat{\Sigma}$ formula). This contradicts our assumption that all proof-lines are $\Sigma\Pi\hat{\Sigma}$ formulas. We then conclude that v must have two immediate children. ∎

A *simple path* in a binary tree is (the set of vertices that are included in) any path in the tree that starts in a vertex with a sibling (or else starts in the root) and goes down along the path (further away from the root) until it reaches a vertex that has two sons. Formally, in $\mathcal{T}$, we define simple paths as follows:

Definition 6.4 (Simple path) *Let P be the path $(v_1, v_2, \ldots, v_k)$, for $k \geq 1$, in the tree $\mathcal{T}$, beginning with the vertex v_1 and ending in v_k (where v_1 is the vertex closest to the root). If the following three conditions hold then the path P is said to be a* simple path *in $\mathcal{T}$.*

(i) v_1 points to no vertex (in which case v_1 is the root of $\mathcal{T}$), or v_1 points to a vertex v_0 such that v_0 has fan-in 2;

(ii) v_k has fan-in 2;

(iii) all vertices excluding v_k in the path have fan-in at most 1.

Proposition 13 *Let P be a simple path in $\mathcal{T}$, in which the first (that is, closest to the root) vertex is labeled with a formula from $\mathsf{Cl}\left(\psi \times \prod_{k\in K}(x_k + b_0)\right)$, for some possibly empty formula ψ, and some $K \subseteq [n]$. Then, for every vertex v in P, $\dot{v} \in \mathsf{Cl}\left(\psi \times \prod_{k\in K}(x_k + b_0)\right)$.*

Proof: Assume that P has more than one vertex (as otherwise the statement is trivial). Suppose that all the vertices in the simple path, excluding the first one, are not the consequences of applying one of the two distributivity rules. Thus, the lemma stems directly from the definition of a simple descendant (Definition 6.2).

Otherwise, there exists a vertex v in the simple path (different from the first vertex in P) that is a consequence of an application of one of the two distributivity rules.

By Proposition 11, v is not a consequence of the backward distributivity rule. Then, suppose that v is a consequence of the forward distributivity rule application. In case the forward distributivity rule was applied inside a $\hat{\Sigma}$ formula or the consequence of the forward distributivity rule application is a $\hat{\Sigma}$ formula, then again the lemma stems directly from the definition of a simple descendant. Otherwise, the forward distributivity rule application meets the conditions in Proposition 12, and so we obtain a vertex in the simple path P (that is, the vertex to which v points) that has fan-in 2, in contrast to the definition of simple paths. ∎

Proposition 14 *Let v be a vertex in $\mathcal{T}$ such that $\dot{v}$ is a simple descendant of $\psi \times \prod_{k\in K}(x_k + b_0)$, where ψ is any possibly empty formula and $K \subseteq [n]$, such that $|K| > 1$. Assume that v has two vertices u, w pointing to it. Then, u, w are the consequences of applying the forward distributivity rule on $\dot{v}$ (in π), and u, w are labeled with a simple descendant of $\psi' \times \prod_{k\in K'}(x_k + b_0)$ and $\psi'' \times \prod_{k\in K'}(x_k + b_0)$, respectively, where $K' \subseteq K$ and $|K'| \geq |K| - 1$, and ψ' and ψ'' are some two formulas.*

Proof: First note that the only derivation rule that can result in v having two sons is the forward distributivity rule. Thus the forward distributivity rule was applied on $\dot{v}$ in π.

Given an application of the forward distributivity rule, let us call the (substitution instance of) the subformula $(Q_2 + Q_3)$ (in the upper-line of the rule $Q_1 \times (Q_2 + Q_3)$), the *principal sum*. Thus, every application of the forward distributivity rule must break the principal sum into two subformulas Q_2 and Q_3 in the lower-line of the rule $Q_1 \times Q_2 + Q_1 \times Q_3$.

Under the conditions in the statement of the proposition, there are only two cases to consider. The first is that the principal sum of the distributivity rule applied on $\dot{v}$ is a sum $(x_j + b_0)$, for some $j \in K$;[10] the second, is that the principal sum is some subformula in ψ.

In the first case, we obtain that u, w are labeled with a simple descendant of $\psi' \times \prod_{k \in K'}(x_k + b_0)$ and $\psi'' \times \prod_{k \in K'}(x_k + b_0)$, respectively, where $K' = K \setminus \{j\}$ (and so $|K'| = |K| - 1$), and ψ' and ψ'' are some two formulas, which is precisely what we need to show.

In the second case, the principal sum that breaks into two subformulas is a part of ψ, and so both u, w are labeled each with a simple descendant of $\psi' \times \prod_{k \in K}(x_k + b_0)$, and $\psi'' \times \prod_{k \in K}(x_k + b_0)$, respectively, for some two formulas ψ', ψ'', which concludes the proof of the lemma. ∎

The following proposition is similar to Proposition 14.

Proposition 15 *Let w be a vertex in G_π (not necessarily in $\mathcal{T}$) and assume that w has fan-out 2 with two edges pointing to the vertices v and u. Suppose that $\dot{v}$ is a simple descendant of $\psi \times \prod_{k \in K}(x_k + b_0)$, where ψ is some possibly empty formula and $K \subseteq [n]$ such that $|K| > 1$. Then w is the consequence of applying the backward distributivity rule on v and u (in π; in fact the rule was applied on the plus gate that has $\dot{v}$ and $\dot{u}$ as its two sons), and u is labeled with a simple descendant of $\psi' \times \prod_{k \in K'}(x_k + b_0)$, where $K' \subseteq [n]$ and $|K'| \geq |K| - 1$, and ψ' is some possibly empty formula.*

Proof: This is similar to the proof of Proposition 14. We omit the details. ∎

Proposition 16 *If v is a vertex in $\mathcal{T}$, then $\dot{v}$ does not compute the zero polynomial.*

Proof: The root formula of $\mathcal{T}$ is $A_{n,0}$. Without loss of generality, we can assume that r_0 in $A_{n,0}$ is nonzero (there must be at least one nonzero r_i, for $0 \leq i \leq n$, as otherwise Sym_n would not compute the zero polynomial). Let P be (the set of vertices in) the path in $\mathcal{T}$ starting in the root and leading to the vertex v. We show by induction on the length of P that every vertex in P is labeled with a formula that can be written as follows (ignoring the order of multiplication and addition):

$$\prod_{k \in K} c_k \times \prod_{j \in J} x_j \times \prod_{i \in I}(d \times x_i + b_0)\,, \tag{7}$$

for three possibly empty sets of indices $J, I \subseteq [n]$ and K and where the c_k's are nonzero field elements and d is a possibly empty field element (and so the formula computes a nonzero polynomial).

The base case is immediate (it is the root formula $\mathcal{A}_{n,0}$). For the induction step, we consider all possible rule applications and show that they preserve the induction statement.

The backward distributivity rule is not applicable in P, by definition of $\mathcal{T}$, unless it is applied inside a $\hat{\Sigma}$ formula, which is impossible in this case.

Note that there are no coefficients multiplying the b_0's. This is because by definition of a $\hat{\Sigma}$ formula, such a formula does not contain a product of two or more field elements. Also, notice that every rule, different from the distributivity rules applied on the formula (7), keeps the induction statement (this can be verified by straightforward inspection). For the forward distributivity rule, it is easy to see that this rule may only transform a formula of the form (7) into two other formulas, each of the form (7). ∎

[10] Note that one cannot derive from $(x_j + b_0)$ any other formulas (ignoring the order) other than $(1 \times x_j + b_0)$ as any other formula will be a non $\hat{\Sigma}$ formula, which will result in $\dot{v}$ to be not a $\Sigma\Pi\hat{\Sigma}$ formula.

We now transform the tree $\mathcal{T}$ into a *full binary tree* (that is, a tree in which every vertex has either fan-in 2 or fan-in 0; a full binary tree might not be balanced, though). This is done by contracting all simple paths in $\mathcal{T}$ (that is, contracting all edges pertaining to a simple path). Formally, we perform the following process on $\mathcal{T}$. Let u be a vertex in $\mathcal{T}$ that has fan-in 1 (in $\mathcal{T}$) and denote by w the (single) vertex in $\mathcal{T}$ that points to u. Replace w with u (in other words, the edge from w to u is contracted and the edge(s) going into w now go into u [and u keeps its label]). Continue this process until there are no vertices of fan-in 1 in the graph. We thus obtain a full binary tree with the root being the vertex labeled with $A_{n,0}$. Denote by $\mathcal{T}'$ the graph just constructed.

Notation:

1. We identify the vertices common to both $\mathcal{T}$ and $\mathcal{T}'$.
2. The level of a vertex v in the full binary tree $\mathcal{T}'$ should not be confused with the level $\mathsf{level}(v)$ of the same vertex v in G_π. To avoid confusion we shall explicitly write in what follows *the level of v in $\mathcal{T}'$* when referring to the former measure.

Lemma 17 *For $0 \leq i \leq n-1$, let v be a vertex in the ith level of (the full binary tree) $\mathcal{T}'$ (if such a level exists in $\mathcal{T}'$). Then, v is a simple descendant of $\psi \times \prod_{k\in K}(x_k + b_0)$, where $K \subseteq [n]$ and $|K| \geq n - i$, and ψ is any possibly empty formula.*

Proof: Note that the root of $\mathcal{T}'$ is labeled with $A_{n,0}$ and that

$$A_{n,0} \equiv r_0 \times \prod_{i\in[n]}(x_i + b_0) \in \mathsf{Cl}\left(\psi \times \prod_{k\in K}(x_k + b_0)\right), \tag{8}$$

for $[n] = K$ and where $\psi = r_0$.

Consider the (not necessarily simple) path P in $\mathcal{T}$ (not in $\mathcal{T}'$) starting from the root and reaching v. By the definition of $\mathcal{T}'$, this path (formally, only the vertices of this path) consists in moving along i consecutive simple paths in $\mathcal{T}$ and then moving one edge further to v (if $i = 0$ then we start from the root of $\mathcal{T}$, and stay there).

By (8) and Proposition 13, all the vertices in the first simple path in P are labeled with elements of $\mathsf{Cl}\left(\psi \times \prod_{k\in K}(x_k + b_0)\right)$.

Starting in the last vertex of the first simple path we can move along P (further away from the root) to its son w. By definition the last vertex of every simple path has two sons (and so w has a sibling in $\mathcal{T}$). Hence, we can apply Proposition 14 to conclude that $\dot{w} \in \mathsf{Cl}\left(\psi' \times \prod_{k\in K}(x_k + b_0)\right)$, for some $|K| \subseteq [n]$ such that $|K| \geq n - 1$, and where ψ' is some possibly empty formula.

Using the same reasoning, after moving along P through i consecutive simple paths, and then moving one edge down to v, we conclude that $\dot{v} \in \mathsf{Cl}\left(\psi' \times \prod_{k\in K}(x_k + b_0)\right)$, where $K \subseteq [n]$ such that $|K| \geq n - i$, and ψ' is some possibly empty formula. ∎

Concluding the proof of Theorem 8. Let v be a *leaf* in $\mathcal{T}'$ having the minimal level ℓ *in the full binary tree* $\mathcal{T}'$.

Case 1: $\ell \geq n/2$. In this case $\mathcal{T}'$ is a full binary tree where all leaves are of level at least $n/2$. This means that the number of vertices in $\mathcal{T}'$ is at least as the number of vertices of a

complete binary tree with $n/2$ levels (a complete binary tree is a full and balanced binary tree). Thus, the number of vertices in $\mathcal{T}'$ is at least $2^{n/2}-1$. Since every vertex in $\mathcal{T}'$ appears in $\mathcal{T}$ and thus also appears in G_π, we get that $|G_\pi| = 2^{\Omega(n)}$.

Case 2: $\ell < n/2$. By Proposition 17:

$$\dot{v} \in \mathsf{CI}\left(\psi \times \prod_{k\in K}(x_k + b_0)\right), \tag{9}$$

where $K \subseteq [n]$ such that $|K| \geq n - \ell > n/2$ (and ψ is some possibly empty formula).

From now on, we will consider v as a vertex in G_π. By the definition of $\mathcal{T}$, there are only two cases that account for v being a leaf in $\mathcal{T}$: (i) the vertex v is of fan-in 0 in G_π; or (ii) there is a (unique) vertex w in G_π that has an out-going edge pointing to v and further there is a directed path in G_π starting from w and terminating in (the first level of G_π) in a vertex different from the root of $\mathcal{T}$ (i.e., different from the vertex labeled with $A_{n,0}$).

In case (i), v must be labeled with the formula 0 (this can be checked by inspection of the derivation rules [Definition 3.2]). But by Proposition 16, $\dot{v}$ does not compute the zero polynomial, and so we arrive at a contradiction.

In case (ii), the vertex w has fan-out 2 (since every directed path in G_π starting in v terminates in the root of $\mathcal{T}$ by definition of $\mathcal{T}$). Thus, w has an out-going edge that goes into v and another out-going edge that goes into a vertex we denote by u. Now, by (9) and Proposition 15 we conclude that u is labeled with a simple descendant of $\psi \times \prod_{k\in K}(x_k + b_0)$, where $K \subseteq [n]$ such that $|K| \geq n - \ell - 1 > n/2 - 1$ (and ψ is some possibly empty formula).

By the regularity condition we have that the trees $\mathbb{T}_v(G_\pi)$ and $\mathbb{T}_u(G_\pi)$ have no common vertices. This in turn means that the conditions of Corollary 7 hold for the vertex u (since, by regularity, u does not have a directed path that reaches $A_{n,0}$ in the initial line in π, and so every directed path beginning at u must terminate in the initial line in a summand pertaining to Init_n) which implies that $|\mathbb{T}_u(G_\pi)| > 2^{n/2-1}$, and so we conclude that $|G_\pi| = 2^{\Omega(n)}$.

6.4 Proof of Technical Lemma 6

Here we prove the main technical lemma. For the sake of convenience, we repeat it fully here:

Lemma 6 (Technical lemma) *Let π be a regular analytic symbolic proof operating with $\Sigma\Pi\hat{\Sigma}$ formulas, and with the initial formula in π being Init_n, for some positive $n \in \mathbb{N}$. Let G_π be the corresponding graph of π. Assume that v is a vertex in G_π labeled with Φ, which is a simple ancestor of*

$$\psi \times \prod_{k=1}^{m} \Psi_k\,,$$

where $m > 1$ and ψ is any possibly empty formula, and for every $k \in [m]$, $i \in [n]$ and $j \in [n]$, Ψ_k is a proper $\hat{\Sigma}$ formula, such that $\Psi_k \notin \mathsf{deriv}(x_i + b_j)$. Then, there exists a vertex y in $\mathbb{T}_v(G_\pi)$, such that there is a path from v to y in $\mathbb{T}_v(G_\pi)$, and y has two outgoing edges to two (distinct) vertices u, w, so that: u and w are labeled with two simple ancestors of the following product formulas

$$\psi_0 \times \prod_{k=1}^{m-1} \Psi'_k \qquad \text{and} \qquad \psi_1 \times \prod_{k=1}^{m-1} \Psi''_k\,, \tag{10}$$

respectively, where ψ_0, ψ_1 are some possibly empty formulas, and for all $k \in [m-1]$, $i \in [n]$ and $j \in [n]$, Ψ'_k, Ψ''_k are some proper $\hat{\Sigma}$ formulas such that $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$ and $\Psi''_k \notin \mathsf{deriv}(x_i + b_j)$.

Denote by $\mathscr{F}$ the set of simple ancestors of all formulas $\prod_{k=1}^{m}\Psi_k$ where each Ψ_k (from the statement of the lemma) is substituted by some Ψ'_k, such that Ψ_k is a derivable *sub*formula from Ψ'_k. Formally, we have:

$$\mathscr{F} := \left\{ \mathsf{Cl}^{-}\left(\prod_{k=1}^{m}\Psi'_k\right) \mid \forall k\in[m],\, \Psi_k \in \mathsf{deriv}(\Psi'_k) \right\}.$$

By the definitions of a simple closure (Definition 6.2) and of derivable subformulas (Definition 6.1), we have the following simple properties:

Fact 1 *1.* $\mathsf{deriv}(\cdot)$ *is transitive: if* $\phi_0 \in \mathsf{derive}(\phi_1)$ *and* $\phi_1 \in \mathsf{derive}(\phi_2)$*, then* $\phi_0 \in \mathsf{derive}(\phi_2)$*.*

2. For all $\phi \in \mathsf{Cl}^{-}(\Psi)$*,* $\Psi \in \mathsf{deriv}(\phi)$*. (On the other hand,* $\Psi \in \mathsf{deriv}(\phi)$ *does not necessarily imply* $\phi \in \mathsf{Cl}^{-}(\Psi)$ *(see Definition 6.1).)*

3. From the above two facts: If $\Psi_k \in \mathsf{deriv}(\Psi'_k)$*, then for all* $\phi \in \mathsf{Cl}^{-}(\Psi'_k)$ *it holds that* $\Psi_k \in \mathsf{deriv}(\phi)$*.*

Proposition 18 *Every formula in* $\mathscr{F}$ *is a simple ancestor of* $\prod_{k=1}^{m}\Psi'_k$*, where, for every* $k \in [m]$*,* $i \in [n]$ *and* $j \in [n]$*,* $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$*, and* Ψ'_k *contains at least one plus gate.*[11]

Proof: Assume by a way of contradiction that $\Psi'_k \in \mathsf{deriv}(x_i+b_j)$, for some $k \in [m]$, $i \in [n]$ and $j \in [n]$. By $\Psi_k \in \mathsf{deriv}(\Psi'_k)$, and by the transitivity of $\mathsf{deriv}(\cdot)$, we have that $\Psi_k \in \mathsf{deriv}(x_i+b_j)$, which contradicts the assumption on the Ψ_k's (in Lemma 6).

Further, since $\Psi_k \in \mathsf{deriv}(\Psi'_k)$, for every $k \in [m]$, and since Ψ_k is a proper $\hat{\Sigma}$ formula, every Ψ'_k must contain at least one plus gate (this can be verified by inspecting the definition of $\mathsf{deriv}(\cdot)$ and the derivation rules [Definition 3.2]). ∎

Proposition 19 *No member of* $\mathscr{F}$ *occurs as a subformula in the initial proof-line.*

Proof: Every formula in $\mathscr{F}$ must contain at least $m > 1$ products of ψ_k's, such that $k \in [m]$ and $\psi_k \in \mathsf{Cl}^{-}(\Psi'_k)$ and $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$, for all $i \in [n]$ and $j \in [n]$. By Item 3 in Fact 1 and by $\Psi_k \in \mathsf{deriv}(\Psi'_k)$ we get that $\Psi_k \in \mathsf{deriv}(\psi_k)$. By assumption (Lemma 6), $\Psi_k \notin \mathsf{deriv}(x_i + b_j)$, for all $i \in [n]$ and $j \in [n]$, and so $\psi_k \notin \mathsf{deriv}(x_i + b_j)$ (for all $i \in [n]$ and $j \in [n]$).

On the other hand, no formula in $\mathscr{F}$ contains products of (non empty) ψ_k's for which $\psi_k \notin \mathsf{deriv}(x_i + b_j)$, for all $i \in [n]$ and $j \in [n]$. This concludes the proof. ∎

Proposition 20 (Critical transition) *There is a directed path (with* 0 *or more edges) from* v *to some vertex* y *in* $\mathbb{T}_v(G_\pi)$*, where* $\mathsf{level}(v) \geq \mathsf{level}(y) > 1$*, such that:* $\dot{y}$ *contains a subformula which is an element of* $\mathscr{F}$*, while the preceding level* $\mathsf{level}(y) - 1$ *corresponds to a proof-line that does not contain a subformula which is an element of* $\mathscr{F}$*.*

(Note that $m > 1$ and so Φ and all formulas in $\mathscr{F}$ are product formulas, which means that every formula in $\mathscr{F}$ may occur [as a complete formula] only in a label of at most one vertex in each level.[12])

[11] When considering only $\Sigma\Pi\hat{\Sigma}$ formulas, the condition that Ψ'_k contains at least one plus gate may be replaced by the condition that Ψ'_k is a proper $\hat{\Sigma}$ formula (note that in the definition of $\mathscr{F}$ we have not restricted the depth of formulas).

[12] In other words, there are no two vertices s, t, at the same level and a formula $\Theta \in \mathscr{F}$, such that one subformula of Θ occurs in $\dot{s}$ and a different subformula of Θ occurs in $\dot{t}$.

Proof: Assume by a way of a contradiction that there is no such y. Note that $\dot{v}$ contains a subformula from $\mathscr{F}$, since clearly $\prod_{k=1}^{m}\Psi_k \in \mathscr{F}$. It is evident (by Definition 4.1) that every vertex in a proof-graph has a path originating in that vertex and terminating in the first level of the graph. Therefore, since there is no y that meets the conditions stated in the claim, every vertex in G_π, on the path from v to a vertex in the first level, must contain a subformula from $\mathscr{F}$. Thus, the initial proof line contains a subformula from $\mathscr{F}$, which contradicts Proposition 19. ∎

To conclude for now, let y be the vertex whose existence is guaranteed by Proposition 20, put $\ell = \mathsf{level}(y)$ and let $\mathcal{A}\{\Theta\}$ denote the formula that corresponds to the (whole) level ℓ, where

$$\Theta :\equiv \dot{y}$$

(note that since Θ is the label of a vertex, Θ is in fact just a summand in $\mathcal{A}$). Then we can write

$$\frac{\mathcal{A}\{\Theta'\}_t}{\mathcal{A}\{\Theta\}_t}\,(\star) \tag{11}$$

to denote the transformation made from level (i.e., proof-line) $\ell - 1$ to level (i.e., proof-line) ℓ, obtained by some derivation rule $(\star)$, where t is a node in $\mathcal{A}$ and:

$$\text{there exists a subformula } \theta \text{ in } \Theta \text{ such that } \theta \in \mathscr{F} \tag{12}$$

$$\text{there is no subformula } \theta \text{ in } \mathcal{A}\{\Theta'\}_t \text{ such that } \theta \in \mathscr{F}. \tag{13}$$

Proposition 21 *Both the formula Θ and the formula θ are proper $\Pi\hat{\Sigma}$ formulas (Definition 6.3).*

Proof: By definition, Ψ_k is a *proper* $\hat{\Sigma}$ formula, for every $k \in [m]$. Since $\Psi_k \in \mathsf{derive}(\Psi'_k)$ for all $k \in [m]$, Ψ'_k must contain some plus formula, for all $k \in [m]$ (this can be verified by inspection of the derivation rules [Definition 3.2]). Hence, since $m > 1$, $\theta \in \mathscr{F}$ must contain a product of at least two formulas, each of which contains a plus formula. Therefore, also the formula Θ must contain a product of at least two formulas, each of which contains a plus formula. Since Θ is a label of some vertex, it must be that Θ in this case is a product formula. And since we work with $\Sigma\Pi\hat{\Sigma}$ formulas, Θ must be a proper $\Pi\hat{\Sigma}$ formula. Further, since θ occurs in Θ (and θ contains a product of two proper $\hat{\Sigma}$ formulas), θ is also a proper $\Pi\hat{\Sigma}$ formula. ∎

In order to prove Lemma 6 we will demonstrate that y above is the vertex stated in the statement of this lemma. Specifically, we will consider all possible rules $(\star)$ that can be applied in (11) above, and conclude that there must be two outgoing edges from y into two vertices that meet the requirements of Lemma 6.

Notation: Assume, for example, that the rule $(\star)$ applied in (11) is $\dfrac{Q_1 + 0}{Q_1}$, and let φ be some formula. Then we shall say, for instance, that φ *occurs in the upper-line* $Q_1 + 0$, to mean that φ occurs in some *substitution instance* Δ of $Q_1 + 0$ (and that clearly Δ occurs in $\mathcal{A}\{\Theta'\}_t$ from (11)). In other words, when referring to occurrences in upper and lower lines of rules that are formulated with formulas in the variables Q_1, Q_2, Q_3, we are formally referring to *substitution instances* of these variables.

Definition 6.5 (Closure of $\mathscr{F}$ under derivation rules) *We say that $\mathscr{F}$ is* closed under some derivation rule $(\star)$ (or under certain instances of the derivation rules)*, if whenever a*

formula Δ is transformed via $(\star)$ (or via a certain instance of the derivation rule) into Δ', and Δ' contains a subformula in $\mathscr{F}$, then Δ also contains a subformula in $\mathscr{F}$. [13]

Case 1: The rule $(\star)$ is one of the following: commutativity, associativity, unit element rules, zero element rules or scalar rules; or the rule $(\star)$ is a distributivity rule applied in a $\hat{\Sigma}$ formula or the consequence of applying $(\star)$ is a $\hat{\Sigma}$ formula (the last two options correspond to the transformations made in the last two clauses in Definition 6.2). By the definition of $\mathscr{F}$ (and by the definition of a simple closure [Definition 6.2]), $\mathscr{F}$ is closed under these rules (Definition 6.5). Thus, proof-line $\ell-1$ contains a formula which is an element in $\mathscr{F}$, and we arrive at a contradiction with (13).

Case 2: The rule $(\star)$ is forward distributivity applied differently from that in Case 1 (that is, it is not applied inside a $\hat{\Sigma}$ formula and its consequence is not a $\hat{\Sigma}$ formula): $\dfrac{Q_1\times(Q_2+Q_3)}{(Q_1\times Q_2)+(Q_1\times Q_3)}$.

We consider the possible occurrences of the lower line $(Q_1\times Q_2)+(Q_1\times Q_3)$ in $\mathcal{A}\{\Theta\}_t$, and conclude that this case does not hold.

(i) There is no subformula of Θ that occurs in the lower line $(Q_1\times Q_2)+(Q_1\times Q_3)$.

Thus, proof-line $\ell-1$ contains Θ as a subformula, in contrast to (13).

(ii) (A substitution instance of) $(Q_1\times Q_2)+(Q_1\times Q_3)$ is a *proper* subformula of Θ.

By assumption (made in Case 2) $(Q_1\times Q_2)+(Q_1\times Q_3)$ is not a $\hat{\Sigma}$ formula. Note that $(Q_1\times Q_2)+(Q_1\times Q_3)$ is a proper $\Sigma\Pi$ formula when considered as a formula in the propositional variables Q_1,Q_2,Q_3 (and not necessarily as a substitution instance of these variables). Since $(Q_1\times Q_2)+(Q_1\times Q_3)$ is not a $\hat{\Sigma}$ formula and since Θ is a proper $\Pi\hat{\Sigma}$ formula (Proposition 21), we arrive at a contradiction.

(iii) The formula Θ is (a substitution instance of the whole formula) $(Q_1\times Q_2)+(Q_1\times Q_3)$.

This is contradictory to Θ being a product formula (Proposition 21).

(iv) The formula Θ is (a substitution instance of) $(Q_1\times Q_2)$.

Suppose that Q_2 is not a $\hat{\Sigma}$ formula. Then it ought to be a proper $\Pi\hat{\Sigma}$ (by Proposition 21). Thus, $Q_1\times(Q_2+Q_3)$ (from the upper-line) is a proper $\Pi\Sigma\Pi\hat{\Sigma}$ formula that appears in proof-line $\ell-1$ in π. This contradicts our assumption that all proof-lines are $\Sigma\Pi\hat{\Sigma}$ formulas.

Therefore, Q_2 is a $\hat{\Sigma}$ formula. Assume that Q_2 is $\hat{\Sigma}$ formula denoted Δ that occurs in $\theta\in\mathscr{F}$ (see (12)). Then, the upper line $Q_1\times(Q_2+Q_3)$ is just Θ with Q_3 added to one of the $\hat{\Sigma}$ products in it. Note that if $F\in\mathscr{F}$ and F' is the result of adding some formula to one of the $\hat{\Sigma}$ products occurring inside F, then $F'\in\mathscr{F}$; this is because for every two formulas Δ_1,Δ_2, $\Psi_k\in\mathsf{deriv}(\Delta_1)$ implies $\Psi_k\in\mathsf{deriv}(\Delta_1+\Delta_2)$ (by definition of $\mathsf{deriv}(\cdot)$).[14] Thus, the upper line $Q_1\times(Q_2+Q_3)$ still contains an element of $\mathscr{F}$ in contrast with (13).

(v) The formula Θ is $(Q_1\times Q_3)$. This is analogous to the previous sub-case (iv).

[13] Note that here we demand that $\mathscr{F}$ is closed under a derivation rule, if it is closed when the rule is applied "backward" on a formula in $\mathscr{F}$.

[14] Here we use the fact that $\mathsf{deriv}(\cdot)$ is defined as the derivable *sub*formulas, and not just derivable formulas.

(vi) The formula Θ is a substitution instance of one of Q_1 or Q_2 or Q_3.

Thus, Θ occurs also in the upper-line (and hence in line $\ell-1$ of the proof), which contradicts (13).

Case 3: The rule $(\star)$ is the backward distributivity $\dfrac{(Q_1\times Q_2)+(Q_1\times Q_3)}{Q_1\times(Q_2+Q_3)}$, applied differently from that in Case 1 (that is, it is not applied inside a $\hat{\Sigma}$ formula and its consequence is not a $\hat{\Sigma}$ formula[15]). Thus, $(Q_1\times Q_2)+(Q_1\times Q_3)$ is not a $\hat{\Sigma}$ formula.

We shall consider the possible occurrences of the lower-line $Q_1\times(Q_2+Q_3)$ in $\mathcal{A}\{\Theta\}_t$, and conclude that the upper-line $(Q_1\times Q_2)+(Q_1\times Q_3)$ constitutes in $\mathbb{T}_v(G_\pi)$ the two vertices u,w to which y points, and that u,w meet the conditions stated in the lemma (Lemma 6).

(i) There is no subformula of θ that occurs in the lower-line $Q_1\times(Q_2+Q_3)$.

Thus, proof-line $\ell-1$ contains θ as a subformula, in contrast to (13).

(ii) There is a only a *proper* subformula of θ that occurs in $Q_1\times(Q_2+Q_3)$ (that is, θ does not occur fully in $Q_1\times(Q_2+Q_3)$).

Recall that θ is a product formula (Proposition 21), and so we can write θ as $\Pi_{i\in I}\theta_i$ for some set of formulas θ_i, $i\in I$. The only possibility in the current case is that there is a partition of I into two nonempty subsets of indices $I=I_0\uplus I_1$, so that θ can be partitioned into two products; one is the product of all θ_i, $i\in I_0$, and the other is the product of all θ_i, $i\in I_1$ (we ignore the order in which the θ_i's occur in θ), and such that: either Q_1 or (Q_2+Q_3) or $Q_1\times(Q_2+Q_3)$ is the formula $\prod_{i\in I_0}\theta_i$ and the product of all θ_i, $i\in I_1$, *does not occur in the lower-line* $Q_1\times(Q_2+Q_3)$. Since $\theta\equiv\Pi_{i\in I}\theta_i$, it must be that $\prod_{i\in I_1}\theta_i$ *multiplies* $Q_1\times(Q_2+Q_3)$ inside $\mathcal{A}\{\Theta\}_t$. This means that in $\mathcal{A}\{\Theta'\}_t$, the formula $\prod_{i\in I_1}\theta_i$ multiplies $(Q_1\times Q_2)+(Q_1\times Q_3)$. Since $(Q_1\times Q_2)+(Q_1\times Q_3)$ is not a $\hat{\Sigma}$ formula, $\dot{y}$ contains a proper $\Pi\Sigma\Pi$ formula, where the middle $\Sigma\Pi$ level does not consist of only $\hat{\Sigma}$ formula, which contradicts our assumption that all proof-lines are $\Sigma\Pi\hat{\Sigma}$ formulas.

(iii) The formula θ occurs (fully) in $Q_1\times(Q_2+Q_3)$.

(1) θ occurs in one of Q_1,Q_2,Q_3. Thus, θ occurs also in the upper-line (and hence in line $\ell-1$ of the proof), which contradicts (13).[16]

(2) θ occurs in (Q_2+Q_3) and item (iii1) above does not hold. This is impossible since θ is a product formula (Proposition 21).

(3) θ occurs in $Q_1\times(Q_2+Q_3)$, and the above two items (iii1) and (iii2) do not hold.

Assume that there exists a proper subformula τ of θ such that $\tau\in\mathscr{F}$ and τ occurs fully in Q_1. Hence, the upper-line also contains τ and we arrive at a contradiction with (13).

Otherwise, we are left only with the following case to consider. Write θ as $\Pi_{i\in I}\theta_i$, for some $\hat{\Sigma}$ formulas θ_i, $i\in I$ (since θ is a proper $\Pi\hat{\Sigma}$ formula [Proposition 21], it is possible to write θ in this way). It must be that there exists a $j\in I$ such that

[15] In the case of the backward distributivity rule, these two options are the same, since the latter transformation is also applied inside a $\hat{\Sigma}$ formula.

[16] This subcase can be shown impossible to meet also due to the depth-3 restriction.

$\theta_j \equiv \Delta_0 + \Delta_1$, where Δ_0, Δ_1 are Q_2, Q_3, respectively; and Q_1 is

$$\left(\psi \times \prod_{i \in I \setminus \{j\}} \theta_i\right),$$

where ψ is some possibly empty formula. Therefore, $(Q_1 \times Q_2)$ and $(Q_1 \times Q_3)$ from the upper-line are:

$$\left(\psi \times \prod_{i \in I \setminus \{j\}} \theta_i\right) \times \Delta_0 \quad \text{and} \quad \left(\psi \times \prod_{i \in I \setminus \{j\}} \theta_i\right) \times \Delta_1,$$

respectively. Thus, we only need to show that these are precisely the two vertices u, w as stated in the lemma (Equation (10)). The proof of this is straightforward and we show it formally for the sake of completeness:

Claim 22 *The formula* $\left(\psi \times \prod_{i \in I \setminus \{j\}} \theta_i\right) \times \Delta_0$ *is a simple ancestor of the product formula* $\psi_0 \times \prod_{k \in K} \Psi'_k$, *where* ψ_0 *is any, possibly empty, formula and* $K \subseteq [m]$ *such that* $|K| = m - 1$ *and for every* $k \in K$, $i \in [n]$ *and* $j \in [n]$, Ψ'_k *is a proper* $\hat{\Sigma}$ *formulas such that* $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$.

(The proof of the right hand side formula $\left(\psi \times \prod_{k \in I \setminus \{j\}} \theta_i\right) \times \Delta_1$ is similar.)

Proof of claim: We know that $\Pi_{i \in I} \theta_i \in \mathscr{F}$, and so by Proposition 18 $\Pi_{i \in I} \theta_i$ is a simple ancestor of $\psi_0 \times \prod_{k=1}^{m} \Psi'_k$, where ψ_0 is some possibly empty formula and for every $k \in [m]$, $i \in [n]$ and $j \in [n]$, $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$, where Ψ'_k contains at least one plus gate. Since each Ψ'_k contains at least one plus gate and every proof-line is a $\Sigma\Pi\hat{\Sigma}$ formula then in fact each Ψ'_k is a proper $\hat{\Sigma}$ formula.

Therefore, $\left(\psi \times \prod_{k \in I \setminus \{j\}} \theta_i\right) \times \Delta_0$ constitutes a simple ancestor of $\psi_0 \times \prod_{k \in K} \Psi'_k$, where ψ_0 is some possibly empty formula and $k \in K$ where $K \subseteq [m]$ and $|K| = m - 1$, and for all $i \in [n]$ and $j \in [n]$ and $k \in K$, Ψ'_k is a proper $\hat{\Sigma}$ formulas such that $\Psi'_k \notin \mathsf{deriv}(x_i + b_j)$. ■

This concludes the proof of Lemma 6.

6.5 Proof of Proposition 4 (Existence of Hard Formulas)

The proof of Proposition 4 uses the standard interpolation formula (cf. Shpilka and Wigderson (2001)).

Consider the polynomial

$$\prod_{i=1}^{n} (x_i + t)$$

as a univariate polynomial of degree n in the indeterminate t over the field of complex numbers $\mathbb{C}$ (the x_i's are part of the coefficients). Then

$$\prod_{i=1}^{n} (x_i + t) = a_0 + a_1 t + \ldots + a_n t^n, \tag{14}$$

where the coefficients $a_0, \ldots, a_n$ may contain x_i variables. By considering the left hand side of (14), one observes that $a_n = 1$. We thus have

$$\begin{pmatrix} 1 & t & t^2 & \cdots & t^{n-1} & t^n \\ & \ddots & & & & \\ \vdots & & & & & \vdots \\ & & & & \ddots & \\ 1 & t & t^2 & \cdots & t^{n-1} & t^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ 1 \end{pmatrix} = \begin{pmatrix} \prod_{i=1}^{n}(x_i + t) \\ \vdots \\ \vdots \\ \vdots \\ \prod_{i=1}^{n}(x_i + t) \end{pmatrix} \tag{15}$$

Let $b_0, \ldots, b_n$ be $n+1$ distinct nonzero elements from $\mathbb{C}$. By Equation (15) we have:

$$\underbrace{\begin{pmatrix} 1 & b_0 & b_0^2 & \cdots & b_0^{n-1} & b_0^n \\ 1 & b_1 & b_1^2 & \cdots & b_1^{n-1} & b_1^n \\ & \ddots & & & & \\ \vdots & & & & & \vdots \\ & & & & \ddots & \\ 1 & b_n & b_n^2 & \cdots & b_n^{n-1} & b_n^n \end{pmatrix}}_{B} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ 1 \end{pmatrix} = \begin{pmatrix} \prod_{i=1}^{n}(x_i + b_0) \\ \prod_{i=1}^{n}(x_i + b_1) \\ \vdots \\ \\ \prod_{i=1}^{n}(x_i + b_n) \end{pmatrix}. \tag{16}$$

Consider the matrix B as defined in Equation (16). B is a Vandermonde matrix, and since the b_i's are all distinct elements from $\mathbb{C}$, by basic linear algebra, B has an inverse denoted B^{-1}, and so:

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ 1 \end{pmatrix} = B^{-1} \begin{pmatrix} \prod_{i=1}^{n}(x_i + b_0) \\ \prod_{i=1}^{n}(x_i + b_1) \\ \vdots \\ \\ \prod_{i=1}^{n}(x_i + b_n) \end{pmatrix} \tag{17}$$

Let $(r_0, \ldots, r_n)$ be the nth row in B^{-1}. Then by Equation (17) we get:

$$1 = r_0 \cdot \prod_{i=1}^{n}(x_i + b_0) + \ldots + r_n \cdot \prod_{i=1}^{n}(x_i + b_n) = \sum_{j=0}^{n} r_j \cdot \prod_{i=1}^{n}(x_i + b_j) \tag{18}$$

Notice that there must be at least one nonzero r_i $(0 \leq i \leq n)$.

Note that (18) is indeed a depth-3 formula with a plus gate at the root. Moving 1 from the left hand side to the right hand side in Equation (18), completes the proof of Proposition 4.

Acknowledgments

I have benefited greatly from discussions with Nachum Dershowitz, Pavel Hrubeš, Jan Krajíček, Ran Raz and Neil Thapen on issues close to this paper.

References

Samuel R. Buss, Russell Impagliazzo, Jan Krajíček, Pavel Pudlák, Alexander A. Razborov, and Jiří Sgall. Proof complexity in algebraic systems and bounded depth Frege systems with modular counting. *Computational Complexity*, 6(3):256–298, 1996/97. ISSN 1016-3328. 1.1.1

Zeev Dvir and Amir Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. *SIAM J. on Computing*, 36(5):1404–1434, 2006. 1.1

Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth circuits. In *Proceedings of the 40th Annual STOC*, pages 741–748, 2008. 1.1

Dima Grigoriev and Edward A. Hirsch. Algebraic proof systems over formulas. *Theoret. Comput. Sci.*, 303(1):83–102, 2003. ISSN 0304-3975. Logic and complexity in computer science (Créteil, 2001). 1.1.1

Pavel Hrubeš. Equational calculi. April 2008. Preprint. ∗, ∗

Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004. 1.1

Zohar Karnin and Amir Shpilka. Deterministic black box polynomial identity testing of depth-3 arithmetic circuits with bounded top fan-in. *ECCC Report TR07-042*, 2007. Revised version, to appear in CCC'08. 1.1

Neeraj Kayal and Nitin Saxena. Polynomial identity testing for depth 3 circuits. *Computational Complexity*, 16(2):115–138, 2007. 1.1

Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non commutative models. *Computational Complexity*, 14(1):1–19, 2005. 1.1

Ran Raz and Iddo Tzameret. The strength of multilinear proofs. *Computational Complexity*, 17(3):407–457, 2008a. 1.1.1

Ran Raz and Iddo Tzameret. Resolution over linear equations and multilinear proofs. *Ann. Pure Appl. Logic*, 155(3):194–224, 2008b. arXiv:0708.1529. 1.1.1

Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980. 1.1

Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. In *Proceedings of the 40th Annual STOC*, pages 507–516, 2008. 1.1

Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Comput. Complexity*, 10:1–27, 2001. 6, 6.5

Richard Statman. Bounds for proof-search and speed-up in the predicate calculus. *Annals of Mathematical Logic*, (15):225–287, 1978. 1.3

Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, pages 216–226. Springer-Verlag, 1979. 1.1