Electronic Colloquium on Computational Complexity, Report No. 76 (2026)

Polynomial Identity Testing for Read-4 Arithmetic Formulas

Nimrod Kaplan * Amir Shpilka*

Abstract

We present the first algorithms for polynomial identity testing (PIT) of read-4 arithmetic formulas in the non-multilinear setting. Specifically, we give a polynomial-time PIT algorithm in the whitebox model and a quasi-polynomial-time algorithm in the blackbox model. Since our techniques are based on proving *hardness of representation* results, we extend our algorithms to *orbits* of read-4 formulas under the action of the affine linear group. Prior to our work, no subexponential white- or blackbox algorithms were known for this class of formulas. All our results hold over any field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = 0$ or $\mathrm{char}(\mathbb{F}) \geqslant 5$.

Prior work addressed only restricted cases. Anderson, van Melkebeek, and Volkovich (*Computational Complexity*, 2015) studied *multilinear* read-k formulas, giving a polynomial-time whitebox PIT algorithm and a quasi-polynomial-time blackbox algorithm. Without the multilinearity restriction, Mahajan, Rao, and Sreenivasaiah (*TCS*, 2014) gave polynomial-time whitebox algorithms for read-2 and read-3 formulas, Prakriya (*Doctoral Thesis*, 2019) obtained quasi-polynomial-time blackbox PIT algorithm for both read-2 and read-3 formulas. Independently, Shamir (*Master's Thesis*, 2022) obtained a quasi-polynomial-time blackbox PIT algorithm for read-2 formulas. For *bounded-depth* read-k formulas, Agrawal, Saha, Saptharishi, and Saxena (*SICOMP*, 2016) obtained a polynomial-time blackbox algorithm in the non-multilinear case. The running time of their algorithm is $n^{k^{2^\Delta}}$ for read-k, depth-Δ formulas, and hence it is applicable only to constant depth.

Partial derivatives are a central tool in the study of deterministic PIT for bounded-read formulas. However, for non-multilinear RkF, differentiation may increase the number of reads. To address this, we develop new structural results that ensure "nice behavior" of derivatives. Specifically, we introduce a new *Fragmentation Lemma* that reduces the PIT problem for general RkFs to simpler models via differentiation. In addition, we define the notion of *dominating degree patterns* and show that, in certain cases, taking partial derivatives with respect to these patterns preserves the read count.

*This research was funded by the European Union (ERC, EACTP, 101142020). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

ISSN 1433-8092

Contents

1 Introduction

Arithmetic circuits are directed acyclic graphs that compute multivariate polynomials using the arithmetic operations $+$ and $\times$. Similarly, arithmetic formulas are circuits whose underlying graph is a tree. These are the most common models for studying the complexity of computing polynomials.

Polynomial Identity Testing (PIT) is a central problem in algebraic complexity theory that asks, given an arithmetic circuit or formula, whether it computes the identically zero polynomial. The problem is studied in two main settings. In the *whitebox* model, the circuit is explicitly given to the algorithm. In the *blackbox* model the algorithm can only make query access to the circuit. That is, it can only evaluate the polynomial at chosen inputs. Solving PIT in the blackbox model is equivalent to constructing a *hitting set* for the class of polynomials under consideration. A hitting set consists of a set of inputs such that any nonzero polynomial in the class evaluates to a nonzero value on at least one point in the set.

If randomness is allowed, the polynomial identity lemma (also known as the Ore-DeMillo-Lipton-Schwartz-Zippel lemma [Ore22, DL78, Sch80, Zip79]) guarantees a randomly chosen input from a sufficiently large grid will, with high probability, provide a nonzero evaluation point for the circuit. The main challenge in the field, and indeed, one of the most fundamental open problems in theoretical computer science, is to find a *deterministic* algorithm. PIT is the most general and well-studied algebraic derandomization problem, with wide range of applications, including primality testing [AB03, AKS04], parallel algorithms for perfect matchings [Lov79, FGT19, ST17], linear matroid intersection [GT17], derandomization of Noether normalization lemma [FS13, Mul17] and even the celebrated proof that IP equals PSPACE [LFKN92, Sha92].

Another reason for the importance of the PIT problem is its deep connection to circuit lower bounds. A series of works [HS80, KI04, Agr05, DSY10, CKS19, KS19, GKSS22, KST23, And20] established that derandomizing PIT is essentially equivalent to proving strong lower bounds for arithmetic circuits, in various settings. Furthermore, depth reduction results show that solving PIT even for highly restricted models such as depth-3 and depth-4 circuits [AV08, Koi12, Tav15, GKKS13] would have major consequences.

Given its importance and the absence of deterministic algorithms in the general case, much research has focused on restricted models of computation. There is a rich body of work on PIT for small-depth circuits, read-once algebraic branching programs, bounded-read formulas, read-once determinants, and other restricted settings. For broader surveys on PIT and its connections, see [Sax09, SY10, Sap21, Sax14, DG24].

One natural approach to derandomizing PIT for arithmetic formulas is to restrict the number of times each variable is read at the leaves. This line of work began with Shpilka and Volkovich [SV15], who provided a quasi-polynomial-size hitting set and a polynomial-time whitebox algorithm for read-once formulas, i.e., formulas in which each variable appears at most once. They also showed how to reconstruct (learn) such formulas deterministically [SV14]. Minahan and Volkovich completely settled the case of read-once formulas by providing a polynomial-sized hitting set. The works [SV14, SV15] also introduced what is now known as the SV-generator, a k-wise independent polynomial map (see Subsection 2.3).

Anderson, van Melkebeek, and Volkovich significantly extended this line of work by studying *read-k multilinear formulas* (multilinear RkF for short),[1] for constant k. These are formulas in which

[1] We later refer to their model as *structurally-multilinear* to distinguish it from cases where multilinearity arises only through cancellations of nonmultilinear terms. In the literature this is often called *syntactically-multilinear*, but in our setting we wish to capture this "syntactic" behavior variable-wise and therefore use a different term.

each variable appears in at most k-leaves, and each gate computes a multilinear polynomial. They presented a polynomial-time whitebox algorithm and a quasi-polynomial-size hitting set. Notably, their result also implies a $2^{(1-\varepsilon)n}$-size hitting set for general multilinear formulas (without any restriction on the number of reads) of size $O(n)$.

As the results of [AvMV15] only hold for multilinear formulas, attempts were made to remove this restriction. Mahajan, Rao, and Sreenivasaiah [MRS14] studied read-2 and read-3 formulas, and gave polynomial-time whitebox PIT algorithms for them. Prakriya [Pra19] gave quasi-polynomial-size hitting sets for read-2 and read-3 formulas. Independently, Shamir [Sha22] gave quasi-polynomial-size hitting sets for read-2 formulas.

In a different direction, Anderson et al. [AFS+18] studied read-k algebraic branching programs (Rk-ABPs, for short). An ABP is a model that subsumes formulas and is believed to be weaker than circuits but stronger than formulas. Any RkF can be represented as an Rk-ABP, meaning that each variable labels at most k edges (for more on ABPs see [SY10, Sap21]). In general though, in an ROABP each variable can appear on multiple edges, as long as they all appear in the *same layer*. Thus, it is a much stronger model than read-once formulas. Anderson et al. gave a subexponential-time whitebox[2] PIT algorithm for Rk-ABPs. Specifically, their algorithm runs in time $2^{\tilde{O}(n^{1-1/2^{k-1}})}$, and yields an algorithm with similar running time for RkFs.

Other than these results, the problem of derandomizing PIT for RkFs remains widely open, both in the whitebox and blackbox settings.

In this work, we design a deterministic quasi-polynomial-time blackbox PIT algorithm for R4Fs. Before our result, no nontrivial deterministic blackbox or whitebox PIT algorithms were known for R4Fs, apart from the slightly subexponential Rk-ABP algorithm of [AFS+18].

1.1 Our Results

All our results are stated in terms of *independent polynomial maps*, a key tool that allows us to reduce the number of variables while preserving nonzeroness. The formal definition and properties of such maps are given in Subsection 2.3. We denote with G_m a polynomial map with seed length m.[3]

The main technical contribution of this work is a quasi-polynomial-time blackbox algorithm for R4Fs. The class R4F is of particular interest, as four is the minimum number of reads required for a non-structural[4] variable to appear. These variables may cause "massive cancellation" which is the core difficulty in verifying polynomial identities.

For a formula F over the variables $\mathbf{x}$ we denote with $p_F(\mathbf{x})$ the polynomial computed at the root of the formula.

All our results hold over any field $\mathbb{F}$ such that $\mathrm{char}(\mathbb{F}) = 0$ or $\mathrm{char}(\mathbb{F}) \geqslant 5$. This restriction arises because our argument uses partial derivatives of order at most four, which may vanish identically over fields of smaller characteristic. Henceforth, we omit explicit mention of the underlying field. Note, however, that constructing a hitting set may require passing to an extension field if $|\mathbb{F}|$ is too small.

Theorem 1.1 (Main blackbox result)**.** *Let $p_F(\mathbf{x})$ be an n-variate polynomial computable by an* R4F F*, over a field $\mathbb{F}$. There exists an absolute constant $c_{1.1}$, such that $p_F(\mathbf{x}) \not\equiv 0$ if and only if*

$$p_F \circ G_{9\log n + c_{1.1}} \not\equiv 0\,.$$

[2]The algorithm uses its whitebox access only to determine the order in which the variables are read. Such a model is sometimes referred to in the PIT literature as *grey-box*.

[3]Explicit constructions such as the SV-generator [SV15] and the RFE-generator [HMM24] are known.

[4]The definition of a structural variable appears in Definition 2.4.

We also present a polynomial-time whitebox algorithm for the same model.

Theorem 1.2 (Main whitebox result)**.** *Let* F *be an* n*-variate* R4F *formula over a field* $\mathbb{F}$*. There exists an algorithm that, given whitebox access to* F*, checks in time poly*(n) *whether* $p_F \equiv 0$*.*

For R3Fs, we obtain a constant-factor improvement over the hitting set introduced in [Pra19] (and also compared to Theorem 1.1)

Theorem 1.3. *Let* $F \in$ R3F *be nonzero. Then, there exists a constant* $c_{1.3}$

$$p_F \circ \mathcal{G}_{2\log n + c_{1.3}} \not\equiv 0\,.$$

We next give a constant-factor improvement in the seed length over the result of [Sha22], and an additive constant improvement in the seed length over the result of [Pra19] (and also compared to Theorem 1.3).

Theorem 1.4. *Let* $F \in$ R2F *be nonzero. Then*

$$p_F \circ \mathcal{G}_{\log n + 4} \not\equiv 0\,.$$

Having established deterministic hitting sets for bounded-read formulas, we next extend these results to orbit classes. I.e., we compose polynomials with *full-rank affine transformations*. The key observation is that hardness of representation results for the base class carry over to the corresponding orbit class. This allows us to "lift" results from the backbone class (the class whose being composed with the affine transformation) to the orbit class. In all results below, $N \leqslant n$ denotes the number of variables of the backbone formulas.

We begin by generalizing the results for sums of ROFs from [SV15] and [BGV23].

Theorem 1.5. *Let* $F_1(\mathbf{y}), F_2(\mathbf{y}), \ldots, F_k(\mathbf{y}) \in$ ROF*, and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$*. Then*

$$\sum_{i=1}^{k} p_{F_i}(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad \sum_{i=1}^{k} p_{F_i}(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{\log N + 3k} \not\equiv 0\,.$$

If $k = 2$ *then we can compose with* $\mathcal{G}_{\log N + 3}$ *instead of* $\mathcal{G}_{\log N + 6}$*.*

We next prove the result for orbits of structurally multilinear polynomials. Our proof closely follows the argument of [AvMV15], except that we use our version of the Fragmentation Lemma rather than the version introduced in that work.

Theorem 1.6. *Let* $F(\mathbf{y}) \in$ RkF *be a structurally multilinear formula, and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$*. Then, for* $r_k \coloneqq r_{2,2k} = k^{o(k)}$*,*

$$p_F(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{r_k + k\log N} \not\equiv 0\,.$$

Finally, we lift our results for general read-2, read-3, and read-4 formulas to their corresponding orbit classes.

Theorem 1.7. *Let* $F(\mathbf{y}) \in$ R2F *and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$*. Then*

$$p_F(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{3\log N + 5} \not\equiv 0.$$

Theorem 1.8. *Let* $F(\mathbf{y}) \in \mathsf{R3F}$ *and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. *Then, there exists a constant* $c_{1.8}$

$$p_F(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{6\log N + c_{1.8}} \not\equiv 0\,.$$

Theorem 1.9. *Let* $F(\mathbf{y}) \in \mathsf{R4F}$ *and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. *Then, there exists a constant* $c_{1.9}$

$$p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{12\log N + c_{1.9}} \not\equiv 0\,.$$

To the best of our knowledge, prior to this work, no nontrivial blackbox PIT algorithm was known for any of these classes.

An important tool in our proofs, and indeed in most previous work on PIT for bounded-read formulas, is the use of partial derivatives and the analysis of the resulting simplified formula. While this approach is effective for *multilinear* formulas, this is no longer the case for formulas with non-multilinear structure; in fact, differentiation may result in a more complex formula than the original one.

In this work, we introduce several structural tools that enable the identification of partial derivatives that do achieve the desired simplification. First, we define two key sets of gates: the $\mathsf{R}(k-1)\mathsf{F}$**-frontier** (Definition 3.1), which identifies "maximal" $\mathsf{R}(k-1)\mathsf{F}$ subformulas, and the set $\mathcal{F}^{(2)}$ (Definition 3.2), which captures "minimal" RkF subformulas. Additionally, we define the notion of **stable variables** (Definition 3.5). Differentiating with respect to a stable variable of a gate guarantees that the resulting structure above that gate is simplified.

These tools are combined and applied in our **Fragmentation Lemma**, which provides a method for simplifying the structure of a general RkF via differentiation. This lemma generalizes previous variants from earlier work and, in comparison, requires fewer partial derivatives to achieve the desired simplification.

Moreover, we introduce the concept of **dominating degree patterns** (Definition 6.2). A degree pattern is a partial exponent vector $\mathbf{e} \in \mathbb{N}^J$ defined over a domain $J \subseteq [n]$. We say $\mathbf{e}$ is *dominating* in a polynomial f if f contains a monomial whose exponent vector coincides with $\mathbf{e}$ on J, while any monomial in f with a strictly higher degree in one variable of J must necessarily have a strictly lower degree in another. We then demonstrate that taking partial derivatives with respect to such dominating degree patterns preserves the read-count of all other variables, provided each variable in the pattern is read at most twice and that $|J| \leqslant 3$.

1.2 Comparison to previous work

There are several works concerning bounded-read formulas with more than one read per variable.

Shpilka and Volkovich [SV15] studied sums of k ROFs. They gave an $n^{O(k)}$ whitebox algorithm and a hitting set of size $n^{O(\log n + k)}$. Their result is actually more flexible and allows to substitute univariable polynomials $f_i(x_i)$ for the x_i-s. This work also introduced the SV-generator and the hardness of representation technique (HoR for short). In a beautiful work, Minahan and Volkovich proved that a 1-independent map hits ROFs. Using that they obtained an $n^{O(k)}$ blackbox PIT algorithm for sum of k ROFs.

We note that a sum of k ROFs is naturally a *multilinear read-k formula*, but of a very restricted structure. Anderson, van Melkebeek and Volkovich presented a polynomial-time whitebox algorithm and a quasi-polynomial-size hitting set for the general model of multilinear read-k formulas. We later explain what the difficulty is in extending their techniques to the non-multilinear case.

Going beyond multilinearity, Mahajan, Rao, and Sreenivasaiah [MRS14] studied read-2 and read-3 formulas, without multilinearity restrictions, and gave polynomial-time whitebox PIT algorithms for them. In his Ph.D. thesis, Prakriya [Pra19] gave quasi-polynomial-size hitting sets

for read-2 and read-3 formulas. A central component of his approach is a reduction from black-box PIT for general RkFs to black-box PIT for $\sum^4 R(k-1)F$. This tool serves as an alternative to our fragmentation lemma. The main difference is the target model, while Prakriya's reduction targets $\sum^4 R(k-1)F$, our lemma reduces to certain derivatives of $\sum^2 R(k-1)F$. Furthermore, Prakriya's also introduced a generalization to the "shattering" lemma from [AvMV15]. Independently, Shamir, in his M.Sc. thesis [Sha22], obtained quasi-polynomial-size hitting sets for read-2 formulas.

A different multi-read restricted model was considered by Mahajan, Rao and Sreenivasaiah [MRS16]. Specifically, they considered what we term as $\sum^2 \prod \mathsf{ROF}$. That is, a sum of two terms, each is a product of ROFs. For this model they gave a whitebox PIT algorithm. We note that while this model can allow unbounded number of reads, it is a much easier model, from the PIT perspective. Indeed, in the white box model it is quite simple to find irreducible factors of ROFs and then all we have to do is normalize each factor and compare multiplicities and evaluation at a nonzero point common to all the ROFs.

Bisht, Gupta and Volkovich [BGV23] gave a polynomial-time blackbox PIT algorithm for $\sum^2 \prod \mathsf{ROF}$. In the same work, they also developed polynomial-time and quasi-polynomial-time blackbox algorithms for the classes of sums of three powers of ROFs and sums of three powers of multilinear RkFs, respectively. Thus, these models allow more than 2 reads, and they don't compute multilinear polynomials, yet there is a very strong restriction on their structure.

To summarize, prior work either dealt with multilinear bounded-read formulas, or bounded-read formulas with a very restricted structure. In particular, no whitebox or blackbox PIT was known for sums of two R2Fs, which is a very restricted form of R4Fs.

Algebraic branching programs (ABPs) are layered graphs with a source and sink nodes, in which variables label edges that go between layers and the polynomial computed by the ABP is the sum over all source-sink paths of the product of the labels of the edges in the path. ABPs form an important model of computation that is (believed to be) stronger than formulas and weaker than circuits. In this model, a read-once ABPs is an ABP in which each variable can appear on edges between two specific adjacent layers. This model is considerably stronger than ROFs as an ROF can be computed by an ABP in which a variable labels only a single edge.

Raz and Shpilka [RS05] provided a polynomial time whitebox PIT algorithm for ROABPs. This immediately implies such an algorithm for read-once formulas (ROFs), although in the case of formulas the problem is much easier. In the blackbox setting, Forbes and Shpilka [FS12] gave the first quasi-polynomial time blackbox algorithm for ROABPs. This started a long line of work on read-once models of ABPs – the usual ROABP model, commutative-ROABP, any-order OABPs, bounded-width ROABPs, sums of ROABPs etc. See [DG24] for a recent survey on PIT that contains many of the state of the art results on ROABPs.

We note that according to the basic definition, ROABPs compute multilinear polynomials. However, similar to [SV15], one can allow univariate polynomials to label edges instead of just variables. This results in non-multilinear computations, that are still as structured as ROABPs.

ABPs that read each variable more than once (i.e., in more than one layer), were first considered by [GKST17] who studied PIT for *sum* of ROABPs. They gave a a whitebox PIT algorithm for the sum of k ROABPs that runs in time $w^{O(2^k)} \cdot (nd)^{O(k)}$, where d is an upper bound on the univariates degree, k is the number of summands, and w is the maximal width (i.e. size of layer) of any of the ROABPs in the sum. In the blackbox setting they gave a hitting set of size $(ndw)^{k2^k \log(ndw)}$ for the model. This model allows reading variables in different layers, and allows non-multilinearity by enabling univariates on edges, yet its underlying structure still comes form that of a multilinear

sum of ROABPs. In that respect, this model can be seen as the ABP analog of the sum-of-ROFs results of [SV15, MV18] .

Other restricted models allowing multiple reads considered bounded depth formulas. There is a vast literature on PIT for bounded depth models, but as we are concerned with formulas without any depth restriction we refer the readers to the surveys [Sax09, SY10, Sax14, DG24] as well as to Forbes' thesis [For14].

Finally, the aforementioned work of [AFS+18] considered unrestricted read-k ABPs, which is a considerably stronger model than read-k formulas. They constructed a gray-box PIT algorithm running in time $2^{\tilde{O}(n^{1-1/2^{k-1}})}$ for this model. This is a stronger model than what we consider, but we achieve a polynomial time whitebox algorithm (compared to theirs slightly sub-exponential algorithm) and a quasi-polynomial-size hitting set, whereas their algorithm needs to know the order of the reads so it does not work in a blackbox model, and of course, it is also slightly sub-exponential, compared to our quasi-polynomial-size hitting set.

To conclude, besides the slightly sub-exponential PIT algorithm of [AFS+18], PIT algorithms for multi-read ABPs dealt with substitutions of univariates to models based on multilinear polynomials.

Another line of research, motivated by the fact that simple polynomials, when composed with full-rank affine transformations, form dense subsets in broader and more interesting classes. This line of work started with the work of Bringmann, Ikenmeyer and Zuiddam [BIZ18] who proved that orbits of width-2 ROABPs are dense in the class of poly-size formulas. Subsequent work, considered orbits of classes such as ROFs [MS21], sparse polynomials [MS21, ST24] and ROABPs [ST24, BG21]. Of particular relevance is the work by Medini and Shpilka [MS21] who proved that orbits of ROFs are dense within poly-size arithmetic formulas. They also gave hitting sets and reconstruction algorithms for polynomials in such orbits. As Forbes and Shpilka [FS18] noted, if these hitting sets could be made *robust*[5] then they would imply hitting set for *general arithmetic formulas*. Thus, exploring hitting sets for such simple classes is an extremely interesting question, with the goal being developing that would eventually lead to robust hitting sets. We note that in the aforementioned work [AvMV15], the authors allowed variables to be replaced with arbitrary sparse polynomial (of polynomial degree), but required that the children of every multiplication node in the substituted multilinear read-k formula are variable disjoint. Thus, this allows more flexibility in the sense that more general polynomials than linear functions are being substituted, however there is the strong restriction of remaining multilinear. In contrast, when we compose, even multilinear read-k formulas with a linear transformation, the result is unlikely to remain multilinear.

1.3 Proof overview and techniques

As mentioned earlier, the main difficulty in analyzing read-4 formulas is that variables cease to behave *structurally*. For read-3 or fewer read formulas, every variable occurs in a "consistent" way: if a node v depends on a variable x and u is a descendant of v, then the degree of x in the subformula at u cannot exceed its degree at v. For read-4 formulas this is no longer the case, and this loss of structural behavior is the source of most of the technical difficulties we face.

One reason this loss of structure is problematic is that previous algorithms crucially relied on taking partial derivatives of the formula. In the multilinear setting, differentiating with respect to a variable only simplifies the formula, while not increasing the number of reads. This is captured

[5] We say that a hitting set $\mathcal{H}$ is *robust* for a class $\mathcal{C}$ if there exists a constant $c > 0$ and a fixed norm $\|\cdot\|$ on $\mathbb{C}[\mathbf{x}]$ such that for every $f \in \mathcal{C}$ there exists $\boldsymbol{\alpha} \in \mathcal{H}$ satisfying $|f(\boldsymbol{\alpha})| \geqslant c \cdot \|f\|$.

by the *fragmentation lemma* of [AvMV15]. However, once variables can appear with higher degrees, differentiation may destroy the formula's bounded-read property. For example, if a subformula contains terms such as $(Ax+B)(Cx+D)$ and the quadratic term ACx^2 cancels out, then the formula depends on x with degree 2 at some gates and degree 1 at others. Differentiating twice with respect to x yields zero, while differentiating once produces a term like $AD + BC$, which may involve many more reads than the original subformula. Thus, the derivative can increase the read count beyond 4, breaking the inductive structure that underlies previous analyses.

It is not surprising that cancellations lie at the heart of the difficulty, since they are precisely what make PIT challenging in general. Many efficient computations for example, the Determinant polynomial, rely on intricate patterns of cancellation that obscure the underlying algebraic structure.

At a high level, and omitting many technical details, our approach is to isolate and control these cancellations. We first show that if a formula has only few structural variables, then we can reduce to the extreme case in which *no* variable is structural. Furthermore, we can assume that all cancellations occur at the top gate. This reduction yields formulas of the form $\sum^2 \mathsf{R2F}$, where the resulting polynomial is multilinear even though each $\mathsf{R2F}$ has degree 2 in every variable. We refer to such formulas as *totally non-structural*. Analyzing the patterns of cancellation that arise in this setting constitutes the main technical component of our work.

We shall now expand more on the ideas that we use to handle these obstacles.

The first tool that we need is a *fragmentation lemma*. Fragmentation lemmas, as termed in [AvMV15], aim to reduce the problem of hitting the class RkF to that of hitting the class $\sum^2 \mathsf{R}(k-1)\mathsf{F}$. Such lemmas were instrumental to the works of [AvMV15] and [Sha22]. In our case, by concentrating at the first gates which are read-4, i.e., gates whose children are read-3 but they themselves are read-4, we prove that by taking an appropriate partial derivative with respect to a variable that appears only at one of these first read-4 gates, simplifies the formula, thus achieving *fragmentation*.

Applying this procedure to $\mathsf{R2F}$s reduces the formula to a $\sum^2 \mathsf{ROF}$s, which is hit by $\mathcal{G}_3$, which is how we obtain Theorem 1.4. In a similar vain, to prove Theorem 1.3, we again use our fragmentation lemma to reduce the PIT problem for $\mathsf{R3F}$s to PIT for an $\mathsf{R3F}$s of the form $\sum^2 \mathsf{R2F}$. If the sum is multilinear, then by Corollary 2.6 it is structurally multilinear, and the result of [AvMV15] applies directly. Otherwise, some variable has degree 2, implying that its degrees in the two summands are unbalanced. In this case, taking the second-order partial derivative with respect to that variable reduces the problem to hitting a nonzero $\mathsf{R2F}$.

1.3.1 Proof Overview of the Blackbox Algorithm (Theorem 1.1)

The following sketch (Figure 1) outlines the structure of the proof and the dependencies among its key components. We refer back to this diagram in later sections to track our progress throughout the proof. As we proceed, we highlight boxes in green to indicate components whose proofs have been completed, and we color a box in orange to indicate the component currently under consideration (see e.g., Figure 2).

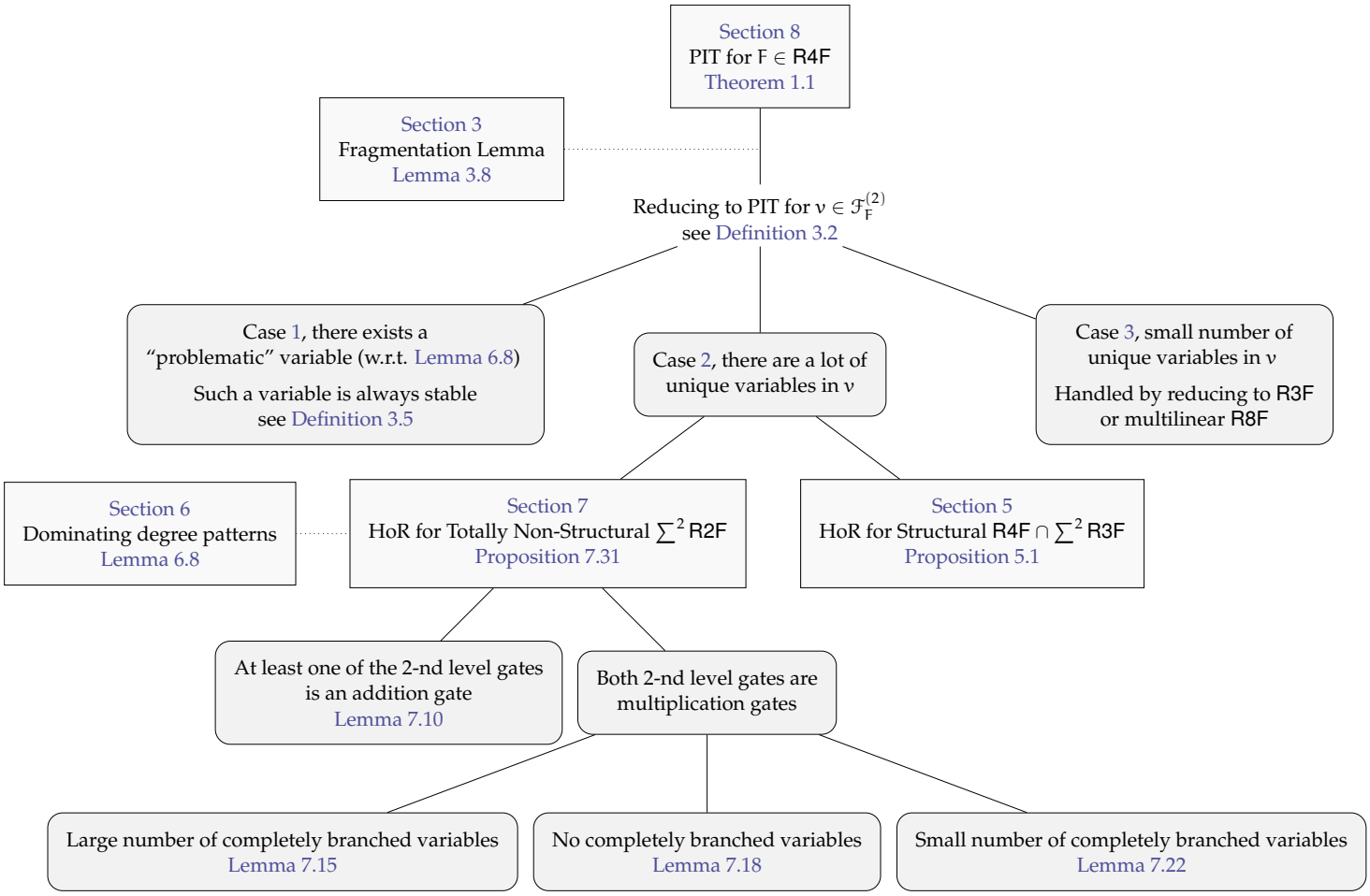

Figure 1: Structure of the proof of Theorem 1.1

The fragmentation procedure discussed above allows us to focus our attention on a gate that is read-4 but that both its children are read-3.

If ν is a multiplication gate then a hitting set for read-3 formulas hits ν. When ν is an addition gate then we are faced with the question of whether there are many or few non-structural variables.

An idea that considerably simplifies the analysis is the use of *generic assignments* to a subset of the variables. A generic assignment is an assignment that does not alter any algebraic property[6] of the polynomials under consideration.

Recall, that in order to apply the *hardness of representation* technique, we need to prove that a formula of a certain structure, cannot compute any polynomial divisible by the monomial $\mathcal{P}_n = \prod_{i=1}^{n} x_i$. This technique has been successfully employed in several prior works on bounded read formulas [SV15, AvMV15, BGV23].

In our analysis we combine this approach with generic assignments. Given a generic assignment we can *restrict* a formula to a subset of its variables, by substituting values to the other variables according to the assignment and simplifying the resulting formula. Observe that if the monomial $\mathcal{P}_n$ divides the original polynomial, then $\mathcal{P}_S = \prod_{i \in S} x_i$ must divide the restricted one. Hence, it is enough to prove hardness of representation results for the restricted polynomial.

[6] An algebraic property is one that can be expressed by a finite set of polynomial equalities.

The main advantage of restrictions is that it greatly simplifies the formula and enables us to reduce the analysis to arguing about polynomials in few variables.[7]

For example, to restrict our attention to p_v we may restrict the variables not appearing in its subtree.

We stress that the generic assignment is only used for the analysis, and not for the construction of the hitting set.

The difficult case, as explained, is when there are many non-structural variables, but let us first explain what we do in the mostly structural case. From now on we assume that our formula F is a structural $\sum^2 \mathsf{R2F}$ that computes a polynomial of degree 2 in each variable.

The structural case: To address the structural case, we introduce the notion of *splitness* (see Definition 5.7) and show that it implies hardness of representation. Splitting a polynomial $h = f + g$ with respect to variables x_i and x_j corresponds to transforming h into a polynomial of the form $\widetilde{h} = f_i f_j + g_i g_j$, where f_i and g_i do not depend on x_j, and f_j and g_j do not depend on x_i. We achieve this by taking a partial derivative with respect to a carefully chosen variable x_t.

More precisely, the three variables i, j, t depend on each other, and on the formula computing h, and we find them using combinatorial analysis of labeled trees, relying on the well-known Erdős–Szekeres theorem.

We then prove the main point which is that a split polynomial cannot be divisible by a high degree monomial. Going back to h we obtain the same result and hence conclude hardness of representation in the structural case.

Let us now switch gears and consider the non-structural case.

Totally non-structural case: Recall that here we assume that F is a $\sum^2 \mathsf{R2F}$ such that each of its $\mathsf{R2F}$ children is of degree 2 in each variable, yet all such monomials cancel out and the polynomial p_v itself is multilinear. Let us denote by $(o)_L$ and $(o)_R$ the left and right children of o, the root of F, respectively.

The property above implies that the first common gate (fcg) of the two leaves labeled by any x_i in $(o)_L$ must be a multiplication gate, and similarly in $(o)_R$. This implies that if, say $(o)_L$, is an addition gate, then there must exist $x_i, x_j \in \mathbf{x}$ whose first common gate is $(o)_L$. Taking a partial derivative with respect to these two variables reduces the problem to PIT for $\mathsf{R2F}$s.

Hence, the challenging case occurs when both $(o)_L$ and $(o)_R$ are a multiplication gates. To handle this case, we introduce the notion of *branched variables*. A variable x_i *branches* in a product gate u if each child of u computes a linear polynomial in x_i. A variable that branches in the two subformulas computed by the children of the root is called *completely branched*, while one that branches in only one child is *partially branched*.

Note that if x_1 completely branches in F then F computes a polynomial of the form $(a_1x_1 + b_1)(a_2x_1 + b_2) + (a_3x_1 + b_3)(a_4x_1 + b_4)$, where the a_is and b_js are subformulas of F.

We further divide the analysis into three cases:

1. there exists a "large" set of variables that branch completely in F;
2. every variable that branches in F branches only partially;
3. the set of completely branched variables is "small".

[7] For us, "few" can be 5^8 many variables, but in most cases we will eventually reduce to polynomials on at most 4 variables.

To handle the first case, we first strengthen, in an appropriate sense, [BGV23, Theorem 1]. This theorem provides a PIT algorithm for the class $\sum^2 \prod \mathsf{ROF}$. However, the proof does not proceed via a hardness of representation argument, which is what we require, since we already used a generic assignment in order to reduce to the $\sum^2 \mathsf{R2F}$ case. Accordingly, we first establish such a hardness of representation result in Lemma 7.13.

Now assume there exists a set $S \subseteq \mathbf{x}$ such that every variable in S branches completely in F. Observe that if we further restrict to the variables in S then the formula simplifies further and becomes a $\sum^2 \prod^2 \mathsf{ROF}$. Therefore, if S is large enough to apply the hardness of representation result from Lemma 7.13, the proof is complete. This is indeed the case whenever $|S| \geqslant 5$.

The second case is more delicate and relies on exploiting the structural constraints imposed by the definition of partial branching. The high level idea is that we first prove that if hardness of representation does not hold, i.e., that a high degree monomial divides p_F, then some variable must branch in one of the children of the root (Lemma 7.9). Assume then that x_1 branches in $(o)_L$. Since it does not completely branch, this means that in $(o)_R$ it appears in exactly one child (i.e., a grandchild of the root). We exploit this structure to eventually prove that p_F cannot be divisible by high degree monomial.

In the third case, we present a procedure that attempts to reduce the setting to one of the two previous cases by applying generic assignments, to get rid of the few variables that completely branch. Note, however, that once we assign a value to a variable, we change the structure of the formula, and new completely branched variables may appear. We show that if this is indeed the case, and it repeats itself through several iterations of fixing completely branched variables, then this outs a very rigid structure on the formula, which allows us to prove that the resulting polynomial cannot be divisible by $\mathcal{P}_n$.

1.3.2 Proof overview of our whitebox algorithm (Theorem 1.2)

We follow the high-level approach of the whitebox algorithm introduced in [AvMV15]. Let S_4 denote the set of read-4 variables in a formula $F \in \mathsf{R4F}$. Our algorithm proceeds iteratively: in each iteration, it removes at least one variable from S_4 without introducing any new variables into it, while preserving nonzeroness of F. After at most n iterations, the resulting formula is an $\mathsf{R3F}$. We then verify its identity using the whitebox algorithm for $\mathsf{R3F}$s from [MRS14].

To achieve this, we follow the cases handled in the proof of Theorem 1.1, and remove variables by following its arguments.

1.3.3 Proof overview of our orbit results

All of our orbit results follow from the simple observation that hardness of representation results for polynomial classes that are closed under translations and have low individual degree imply PIT not only for the backbone class[8], but also for the orbit class itself.

Assume we are concerned with an m-hard[9] polynomial class $\mathcal{C}$. Then every polynomial $f(\mathbf{y}) \in \mathcal{C}$ contains a monomial whose support size is at most $m - 1$.[10] Since the action of every nonzero matrix $\mathbf{A} \in \mathrm{GL}_n(\mathbb{F})$ on $\mathbb{F}[\mathbf{y}]$ induces a degree-preserving isomorphism, if f has low

[8] $\mathcal{C}$ denotes the backbone class corresponding to the orbit class $\mathcal{C}^{\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})}$.

[9] No monomial of support m divides any nonzero polynomial in this class, and the class is closed under zero substitutions. For the formal definition and further discussion, see Subsection 2.6.

[10] This is a well-known fact, formalized in Observation 2.38.

degree, then $f(\mathbf{A}\mathbf{x})$ must contain a monomial of small support. We extend this observation to the action of $\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$ by first translating our polynomials by an appropriate vector of scalars.

This observation immediately lifts a large number of known results for backbone classes to their corresponding orbit classes, as demonstrated in Section 9.

1.4 Organization

The paper is organized as follows. Section 2 introduces the notation, definitions, and prior results that we use throughout this work. In Section 3, we prove our Fragmentation Lemma and define several important sets of gates. Section 4 presents our results for R2Fs and R3Fs.

Our work on blackbox PIT for R4Fs begins in Section 5, where we prove a hardness of representation result for structural R4Fs that are sums of two R3Fs. In Section 6, we introduce the notion of a *dominating degree pattern*, define elementary sets of polynomials, and prove an extension of Lemma 2.11 for R2Fs. This extension allows us to hit the set of polynomials required by our PIT algorithms. In Section 7, we prove a hardness of representation result for totally non-structural sums of two R2Fs. Section 8 contains our main results, both blackbox and whitebox, for R4Fs. Finally, in Section 9, we show how hardness of representation results translate into PIT algorithms for orbit classes, and we prove all corresponding results for these classes.

2 Preliminaries

2.1 Notations

We use the symbol $\coloneqq$ to denote definitions. For natural numbers $n \leqslant m \in \mathbb{N}$, we write $[n] \coloneqq \{1, \ldots, n\}$ and $[m : n] \coloneqq \{m, \ldots, n\}$. We use $\binom{[n]}{k}$ to denote the collection of all subsets of $[n]$ of size k.

Unless otherwise specified, all polynomials in this work are considered over the set of variables $\{x_1, \ldots, x_n\}$, which we denote by $\mathbf{x}$. For a subset $J \subseteq [n]$, we write $\mathbf{x}_J$ for the subvector of variables indexed by J (in their natural order). For $I \subseteq [n]$, we denote $\mathbf{x}_{-I} \coloneqq \mathbf{x}_{[n]\setminus I}$, and for a singleton $\{i\} \subseteq [n]$, we write $\mathbf{x}_{-i} \coloneqq \mathbf{x}_{[n]\setminus\{i\}}$.

Polynomials. Let $p \in \mathbb{F}[\mathbf{x}]$. When used without parentheses, $p \coloneqq p(\mathbf{x})$ by default. We say that p *depends on* a variable x_i if there exist vectors $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{F}^n$ that differ only in their ith coordinate such that $p(\boldsymbol{\alpha}) \neq p(\boldsymbol{\beta})$. In the blackbox setting, we always assume, without loss of generality, that the polynomials under consideration depend on all variables in $\mathbf{x}$. For two polynomials $p, q \in \mathbb{F}[\mathbf{x}]$, we write $p \sim q$ to denote that p and q are *associates*, i.e., there exists $\alpha \in \mathbb{F}^\times$ such that $p = \alpha q$.

We use the following notation:

- $\mathrm{var}(p) \coloneqq \{x_i \mid p \text{ depends on } x_i\}$.
- For a monomial M, $\mathrm{coeff}_p(M)$ denotes the coefficient of M in p.
- $\mathrm{mon}(p)$ denotes the set of monomials with nonzero coefficients in p.
- For a monomial $M \in \mathbb{F}[\mathbf{x}]$, the *support* of M, denoted $\mathrm{Supp}(M)$, is the set of variables on which M depends.
- $\deg(p)$ denotes the total degree of p, and for each $x_i \in \mathbf{x}$, $\deg_{x_i}(p)$ denotes the degree of x_i in p.

- Let $J \subseteq [n]$ and $\boldsymbol{\alpha} \in \mathbb{F}^J$. We denote by

$$p|_{\mathbf{x}_J = \boldsymbol{\alpha}} \quad \text{(or, equivalently, } p|_{J \leftarrow \boldsymbol{\alpha}}\text{)}$$

  the polynomial obtained by substituting $x_i = \alpha_i$ for each $i \in J$. When $J = \{i\}$, we abuse notation and write $p_{x_i = \alpha}$ (equivalently, $p|_{i \leftarrow \alpha}$).

- For $x \in \mathbf{x}$ and $d \in \mathbb{N}$, we write

$$\partial_{x^d}(p) \coloneqq \frac{\partial^d p}{(\partial x)^d}\,.$$

  More generally, for $I \subseteq [n]$ with $I = \{i_1, \ldots, i_{|I|}\}$ and $\mathbf{d} \in \mathbb{N}^I$, we define

$$\partial_{I^{\mathbf{d}}}(p) \coloneqq \partial_{x_{i_1}^{d_{i_1}}} \cdots \partial_{x_{i_{|I|}}^{d_{i_{|I|}}}}(p)\,.$$

- We denote

$$\partial_{x_i^{\deg}}(p) \coloneqq \partial_{x_i^{\deg_{x_i}(p)}}(p)$$

  to mean the derivative of p taken with respect to x_i as many times as the degree of x_i in p.

- A vector $\boldsymbol{\alpha}$ is called a *zero* of p if $p(\boldsymbol{\alpha}) = 0$, and we write $\boldsymbol{\alpha} \in Z(p)$.

Multilinear monomials. For $J \subseteq [n]$ and $n \in \mathbb{N}$, we write

$$\mathcal{P}_J \coloneqq \prod_{i \in J} x_i \quad \text{and} \quad \mathcal{P}_n \coloneqq \mathcal{P}_{[n]} = \prod_{i=1}^{n} x_i.$$

2.2 Bounded Read Formulas

Definition 2.1 (Arithmetic formula). Arithmetic formula is a binary tree where each leaf is labeled with a variable $x_i \in \{x_1, \ldots, x_n\}$ and each internal node, called a gate, has an operation $+, \times$. Additionally, each internal node v is labeled with some $\alpha_v, \beta_v \in \mathbb{F}$.

The formula is evaluated recursively as follows: Each leaf ℓ labeled with x_i computes $p_\ell = \alpha_\ell x_i + \beta_\ell$. Each gate v with an operation $\mathrm{op} \in \{+, \times\}$, computes $\alpha_v(p_{(v)_L} \,\mathrm{op}\, p_{(v)_R}) + \beta_v$, where $p_{(v)_L}$ and $p_{(v)_R}$ are the polynomials computed by it's left and right children respectively.

For an algebraic formula F, we denote by p_F the polynomial computed by F.

We denote by o_F the output gate of F, if F is clear from context we simply denote it's root by o. ◊

The notations for polynomials extend to formulas in the natural way. In particular, we define:

- $\mathrm{var}(F) \coloneqq \mathrm{var}(p_F)$, the set of variables on which F depends.

- For $x_i \in \mathbf{x}$ and $\alpha \in \mathbb{F}$, we write $F|_{x_i = \alpha}$ to denote the formula obtained by substituting $x_i = \alpha$ in F. This extends naturally to any subset $I \subseteq [n]$, as in the polynomial case.

We denote by $\mathrm{Read}_{x_i}(F)$ the number of leaves of F labeled by x_i. We sometimes refer to this number as the *number of reads with respect to* x_i.

The notation $v \in F$ indicates that v is a gate that appears in the formula F. For $v \in F$, we let F_v denote the subformula rooted at the gate v, and we write $p_v \coloneqq p_{F_v}$ for the polynomial computed by F_v. We also set $\mathrm{var}(v) \coloneqq \mathrm{var}(p_v)$, the set of variables on which v depends.

We refer to the left and right children of a gate v as $(v)_L$ and $(v)_R$, respectively. To refer to deeper descendants of v, we use sequences of ''L''s and ''R''s separated by semicolons. For example, the left child of the right child of v is denoted $(v)_{R,L}$, its left child is $(v)_{R,L,L}$, and so on. For a gate v, we refer to $(v)_L$ and $(v)_R$ as its *children*, and say that they are *siblings*. Similarly, the children of $(v)_L$ and $(v)_R$ are the *grandchildren* of v.

For two gates $v, u \in F$, we say that v and u are *disjoint* if neither lies on the path from the other to the root o. The *first common gate* (fcg) of v and u, denoted $\mathrm{fcg}(v,u)$, is the deepest node that has both v and u in its subtree. Equivalently, $\mathrm{fcg}(v,u) = w$ if $v \in F_{(w)_L}$ and $u \in F_{(w)_R}$, or vice versa. The notion of the first common gate extends naturally to any number of gates.

For a formula F, a gate $v \in F$, and a constant $\gamma \in \mathbb{F}$, we denote by $F|_{v=\gamma}$ the formula obtained by replacing the gate v with the constant γ and simplifying the resulting formula accordingly.

Definition 2.2 (Read-k formula)**.** A read-k formula is an arithmetic formula in which each variable $x_i \in \mathbf{x}$ labels at most k leaves, i.e., $\mathrm{Read}_{x_i}(F) \leqslant k$. A polynomial is a read-k polynomial if it can be computed by a read-k formula. We denote by RkF the class of read-k formulas. For read-once formulas, we use the standard notation ROF.

The class $\sum^m \mathsf{RkF}$ consists of all the polynomials that can be decomposed into a sum of m RkP polynomials.

The class $\prod^k \mathsf{ROF}$ consists of all RkF that factor as products of ROFs. We similarly define the class $\sum^m \prod^k \mathsf{ROF}$. ◊

The following is a simple observation.

Observation 2.3. *Let* $G \in \mathbb{F}[x_1, x_2]$ *be multilinear. Then* G *is a* ROP*.*

The next two definitions are taken from [Sha22] with slight changes.

Definition 2.4 (Structural variable, Structural formula)**.** For some algebraic formula F, a variable $x_i \in \mathrm{var}(F)$ is said to be structural if for each pair of gates $v, u \in F$, such that v is an ancestor of u, it holds that $\deg_{x_i}(v) \geqslant \deg_{x_i}(u)$.

A formula F is structural if every $x_i \in \mathrm{var}(F)$ is structural. ◊

The next observation implies that we can always assume that a read-3 formula is structural. We omit the simple proof.

Observation 2.5. *Let* $F \in \mathsf{RkF}$*. Let* $x_i \in \mathbf{x}$ *such that* $\mathit{Read}_F(x_i) \leqslant 3$*. Then, there exists* $F' \in \mathsf{RkF}$ *such that* $p_F = p_{F'}$*,* x_i *is structural in* F' *and for every* $x_j \in \mathbf{x}$*,* $\mathit{Read}_{x_j}(F') \leqslant \mathit{Read}_{x_j}(F)$*.*

Corollary 2.6. *Let* $F \in \mathsf{R3F}$*. Then, there exists a structural* $F' \in \mathsf{RkF}$ *such that* $p_F = p_{F'}$ *and for every* $x_i \in \mathbf{x}$ *:* $\mathit{Read}_{x_i}(F') \leqslant \mathit{Read}_{x_i}(F)$*.*

Definition 2.7 (Unvisited gates)**.** Let F be a formula and $g \in F$. We denote by $\mathrm{Unv}_F(g)$ the set of unvisited children of multiplication gates along the path from g to the root of F. Likewise, we denote by $\mathrm{Unv}^+_F(g)$ the set of unvisited children of addition gates along the same path. When F is clear from context, we omit the subscript and simply write $\mathrm{Unv}(g)$ and $\mathrm{Unv}^+(g)$. ◊

The next simple lemma will be applied repeatedly throughout this work.

Lemma 2.8 ([Sha22, Lemma 2.12, restated])**.** *Let* $F \in \mathsf{RkF}$*,* $x_i \in \mathbf{x}$ *and let* u *be the first common gate (fcg) to all the leaves of* F *which are labeled with* x_i*. Then, for any* $d \in \mathbb{N}$*:*

$$\partial_{x_i^d} p_F = \partial_{x_i^d}(p_u) \prod_{v \in \mathit{Unv}(u)} p_v \,.$$

Lemma 2.9 ([SV14, Implicit in the proof of Lemma 3.12])**.** *Let* F *be a* ROF*. Then* p_{F} *is reducible if and only if* $o^{op} = \times$ *and the additive constant of* o *is zero.*

Lemma 2.10. *Let* F *be a* ROF*. Each irreducible factor of* p_{F} *is, up to multiplication by a scalar, computed at one of the gates of* F*.*

Proof. Let F be a ROF with irreducible factorization $p_{\mathsf{F}} = \prod_{i=1}^{m} p_i$. We prove the claim by induction on m.

If $m = 1$, the claim follows immediately since the output gate o computes p_{F}. Assume the claim holds for ROFs with smaller values of m. Since $m > 1$, p_{F} is reducible, and by Lemma 2.9, we have $\mathsf{F} = \mathsf{F}_1 \times \mathsf{F}_2$ for some ROFs F_1 and F_2. The claim then follows by applying the induction hypothesis to F_1 and F_2. □

The next lemma shows the connection between structural variables and partial derivatives. In particular, it shows that taking partial derivatives with respect to any structural variable, then by Lemma 2.8 the complexity of the underlying algebraic formula is not increased. For completeness, we provide the proof in Appendix A.

Lemma 2.11 ([Sha22, Implicit in Lemma 2.7])**.** *Let* $\mathsf{F} \in$ RkF *and let* $x_i \in var(\mathsf{F})$ *be a structural variable. Set* $p_{\mathsf{F}'} \coloneqq \partial_{x_i^{\deg}} p_{\mathsf{F}}$*. Then* $p_{\mathsf{F}'}$ *is an* RkP *computable by an* RkF F' *such that:*

1. F' *consists entirely of disjoint subformulas of* F*, and*
2. *for every* $x_j \in \mathbf{x}_{-i}$*, we have* $Read_{x_j}(\mathsf{F}') \leqslant Read_{x_j}(\mathsf{F})$*.*

Remark 2.12. Since we will use it throughout this paper, we won't repeatedly mention the generalized multiplication rule, and will just write

$$\partial_{x_i^{\deg}} p_u \sim \partial_{x_i^{\deg}} p_{(u)_L} \partial_{x_i^{\deg}} p_{(u)_R}$$

whenever u is a multiplication gate.

Definition 2.13 (Structurally Multilinear Formula)**.** We say that a formula F is structurally multilinear, if for every gate $v \in \mathsf{F}$, p_v is multilinear. ◊

2.3 Independent Polynomial Maps and the SV-Generator

Definition 2.14 (Independent Polynomial Maps)**.** A polynomial map

$$\mathcal{G}(y_1, \ldots, y_t, z) : \mathbb{F}^{t+1} \to \mathbb{F}^n$$

is called a *1-independent polynomial map* if for every $x_i \in \mathbf{x}$ there exists an assignment $\boldsymbol{\alpha}_i \in \mathbb{F}^t$ to the variables $y_1, \ldots, y_t$ such that the i-th coordinate of $\mathcal{G}(\boldsymbol{\alpha}_i, z)$ equals z, and all other coordinates are 0. The assignment $\boldsymbol{\alpha}_i$ is called the i-*th selector assignment*.

For $k \geqslant 1$, a polynomial map

$$\mathcal{G}(y_1, \ldots, y_{kt}, z_1, \ldots, z_k) : \mathbb{F}^{k(t+1)} \to \mathbb{F}^n$$

is called a k-*independent polynomial map* if it can be written as a sum of k variable-disjoint 1-independent polynomial maps.

We denote k-independent polynomial maps by $\mathcal{G}_k$ (with t implicit). The $\mathbf{y}$-variables are called *selector variables*. The parameter k is the *seed length* of $\mathcal{G}_k$, and n is the *stretch*. ◊

Explicit constructions of independent polynomial maps are known for every $k, n \in \mathbb{N}$. A canonical example is the *SV-Generator* from [SV15]. For $n, k \in \mathbb{N}$, we denote by

$$\mathcal{G}_k : \mathbb{F}^{2k} \to \mathbb{F}^n$$

the k-independent SV-Generator. Unless stated otherwise, n will be implicit in the notation; when needed, we write $\mathcal{G}_{n,k}$ to emphasize the output stretch. The power of the SV-Generator was extensively studied in [HMM24], which also provided a characterization of its vanishing ideal.

The usefulness of independent polynomial maps is illustrated by the following result of [MV18].

Theorem 2.15 ([MV18])**.** *Let $F \in \mathsf{ROF}$ be nonzero. Then $F \circ \mathcal{G}_1 \not\equiv 0$.*

We list below elementary properties of independent polynomial maps, used implicitly throughout the paper.

Fact 2.16.

1. $\mathcal{G}_{n,k}|_{y_k \leftarrow \alpha_i} = \mathcal{G}_{n,k-1} + z_k \cdot e_i$, *where e_i is the i-th standard basis vector and α_i is the i-th reviving assignment.*
2. $\mathcal{G}_{n,k_1} + \mathcal{G}_{n,k_2} = \mathcal{G}_{n,k_1+k_2}$, *where $\mathcal{G}_{n,k_1}$ and $\mathcal{G}_{n,k_2}$ are defined on disjoint sets of input variables of sizes $2k_1$ and $2k_2$, respectively.*
3. *For every $k' \leqslant k$, $Img(\mathcal{G}_{n,k'}) \subseteq Img(\mathcal{G}_{n,k})$.*

Definition 2.17 (Hitting set generator)**.** A polynomial map $\mathcal{H} = (\mathcal{H}_1, \ldots, \mathcal{H}_n) : \mathbb{F}^k \to \mathbb{F}^n$ is called a *hitting set generator* (HSG) for a circuit class $\mathcal{C}$ if for every nonzero n-variate polynomial $f \in \mathcal{C}$ we have $f \circ \mathcal{H} \not\equiv 0$. For brevity, we often refer to a hitting set generator simply as a *generator*. ◊

A hitting set generator $\mathcal{H}$ for a class $\mathcal{C}$ is said to *hit* $\mathcal{C}$; equivalently, $\mathcal{H}$ *hits* every nonzero n-variate polynomial $f \in \mathcal{C}$

Observation 2.18. *Let $f = \prod_{i=1}^{m} f_i$ be a nonzero polynomial such that $f_i \in \mathcal{C}$ for every $i \in [m]$, and let $\mathcal{H}$ be a generator for $\mathcal{C}$. Then $f \circ \mathcal{H} \not\equiv 0$.*

The following is an important property of any independent polynomial map.

Observation 2.19. *Let $f \in \mathbb{F}[\mathbf{x}]$ be a polynomial and $\mathcal{G} : \mathbb{F}^t \to \mathbb{F}^n$ be some polynomial map such that $\mathcal{G} = (\mathcal{G}^1(\boldsymbol{w}), \ldots, \mathcal{G}^n(\boldsymbol{w}))$. Define $\mathcal{H} : F^{t+2} \to F^n$ as $\mathcal{H} = \mathcal{G}(\boldsymbol{w}) + \mathcal{G}_1(y, z)$. Let $x_i \in \mathbf{x}$. By Fact 2.16:*

$$f(\mathcal{H})|_{y=\alpha_i, z=x_i-\mathcal{G}^i(\boldsymbol{w})} = f\left(\mathcal{G}^1(\boldsymbol{w}), \ldots, \mathcal{G}^{i-1}(\boldsymbol{w}), x_i, \mathcal{G}^{i+1}(\boldsymbol{w}), \ldots, \mathcal{G}^n(\boldsymbol{w})\right)$$

where α_i is the i-th reviving assignment.

This leads to the following corollary, a standard result appearing throughout the literature on independent polynomial maps.

Corollary 2.20. *Let $f \in \mathbb{F}[\mathbf{x}]$ be a nonzero polynomial, and suppose that some $M \in mon(f)$ satisfies $|\mathrm{Supp}(M)| \leqslant k$ for some $k \in \mathbb{N}$. Then*

$$f \circ \mathcal{G}_k \not\equiv 0\,.$$

Proof. Let $M' \in \mathrm{mon}(f)$ be a monomial of minimal support size and assume without loss of generality that $\mathrm{Supp}(M') = \{x_1, x_2, \ldots, x_{k'}\}$. By Observation 2.19, there exists an assignment $\boldsymbol{\alpha}$ to the variables of $\mathcal{G}_{k'}$ such that

$$f \circ \mathcal{G}_{k'}(\boldsymbol{\alpha}) = f(x_1, x_2, \ldots, x_{k'}, 0, 0, \ldots, 0) \sim M' \not\equiv 0\,.$$

Since $k' \leqslant k$, Fact 2.16 implies that $f \circ \mathcal{G}_k \not\equiv 0$, as claimed. □

The next theorem is improved in Section 4.

Theorem 2.21 ([Sha22, Theorem 5.2]). *Let* $F \in \mathsf{R2F}$ *be nonzero. Then*

$$p_F \circ \mathcal{G}_{6+6\log n} \not\equiv 0.$$

2.4 Preserving Assignments

In this section we introduce three types of assignments that preserve structural properties of polynomials. The first two have appeared in prior work on PIT for bounded-read formulas. The third, introduced here, extends these ideas to non-multilinear polynomials and can be viewed as a generalization of justifying assignments (the first type).

2.4.1 Justifying Assignments

Justifying assignments were first defined in [SV15]; related variants appeared earlier in [HH91, BHH95]. Since then, they have been used in many works in the area (see, e.g., [SV10, MV18, BGV23]).

Definition 2.22 (Justifying assignment). Let $f \in \mathbb{F}[\mathbf{x}]$ be a nonzero polynomial and $\boldsymbol{\alpha} \in \mathbb{F}^n$. Then, $\boldsymbol{\alpha}$ is called a *justifying assignment* of f (equivalently, f is said to be $\boldsymbol{\alpha}$*-justified*) if the following properties are satisfied:

1. For every $I \subseteq \mathrm{var}(f)$, $\mathrm{var}(f|_{I \leftarrow \boldsymbol{\alpha}_I}) = \mathrm{var}(f) \setminus I$,
2. $f(\boldsymbol{\alpha}) \neq 0$. ◊

Proposition 2.23 ([SV15, Properties of justifying assignments]). *Let* F *be an algebraic formula. The following hold:*

1. *If* $\boldsymbol{\alpha} \in \mathbb{F}^n$ *satisfies*
$$\partial_{x_i} p_F(\boldsymbol{\alpha}) \neq 0 \quad \textit{for all } x_i \in \mathbf{x},$$
then $\boldsymbol{\alpha}$ *is a justifying assignment for* F *[SV15, Prop. 2.10].*
2. *The class of* **0***-justified* ROF*s is closed under taking partial derivatives [SV15, Lemma 3.11].*[11]
3. *If* F *is a* **0***-justified multiplicative* ROF,[12] *then for each* $i \in [n]$ *there is at most one* $\beta \neq \alpha_i$ *and one* $x_j \in \mathbf{x}_{-i}$ *such that* $p_F|_{x_i=\beta, x_j=0} \equiv 0$*. Moreover, if such* β *and* x_j *exist, the sibling of the leaf labeled* x_i *is* x_j *[SV15, Lemma 3.13].*[13]

2.4.2 Irreducibility preserving assignments

We next define the *commutator* of a polynomial, a notion that has appeared in [SV10, SV14, BGV23]. Our definition is specialized to multilinear polynomials and will suffice for our purposes; for the more general definition applicable to arbitrary polynomials, see [SV10].

Definition 2.24 (Commutator). Let $p \in \mathbb{F}[\mathbf{x}]$ and $i, j \in [n]$. The commutator of p with respect to x_i and x_j is

$$\Delta_{i,j} p \coloneqq p|_{x_i=1,\, x_j=1} \cdot p|_{x_i=0,\, x_j=0} - p|_{x_i=0,\, x_j=1} \cdot p|_{x_i=1,\, x_j=0}.$$ ◊

[11] [SV15, Lemma 3.11] is stated for weakly **0**-justified ROF, but as noted there it extends to **0**-justified ROFs.
[12] A multiplicative ROF is an ROF with no addition gates.
[13] This statement is implicit in the proof of [SV15, Lemma 3.13].

The following lemma connects commutators of multilinear polynomials with irreducibility:

Lemma 2.25 ([SV10, Lemma 4.6])**.** *Let $f \in \mathbb{F}[\mathbf{x}]$ be a non-constant multilinear polynomial and $i \neq j \in [n]$. Then $f = g \cdot h$ for some $g, h \in \mathbb{F}[\mathbf{x}]$ with $x_i \notin var(h)$ and $x_j \notin var(g)$ if and only if $\Delta_{i,j} f \equiv 0$.*

We use commutators only for ROPs, where the commutator itself is a product of ROPs; the next lemma follows from [SV14, Lemma 3.14] and its proof.

Lemma 2.26 ([SV14, Lemma 3.14, simplified])**.** *Let F be a* ROF *and let $i \neq j \in [n]$ be such that*

$$\partial_{x_i, x_j} p_F \not\equiv 0.$$

Then, let $v = fcg(\ell_i, \ell_j)$ and let $\gamma \in \mathbb{F}$ be the additive constant of the gate v. Denote by $F' = F|_{v=\gamma}$ the ROF *obtained by substituting the constant γ instead of the gate v in F. Then,*

$$\Delta_{i,j} p_F = p_{F'} \cdot \partial_{x_i, x_j} p_F \,.$$

We next define assignments that ensure that the irreducibility of a polynomial is maintained.

Definition 2.27 (Irreducibility-preserving assignment, [BGV23, Definition 2.4])**.** Let $f \in \mathbb{F}[\mathbf{x}]$ and $\boldsymbol{\sigma} \in \mathbb{F}^n$. We say that $\boldsymbol{\sigma}$ is an *irreducibility-preserving assignment* for f (equivalently, f is $\boldsymbol{\sigma}$*-irreducible*) if for every proper subset $I \subsetneq [n]$, the restricted polynomial $f|_{I \leftarrow \boldsymbol{\sigma}_I}$ is non-constant and irreducible, and moreover $f(\boldsymbol{\sigma}) \neq 0$. ◊

Remark 2.28. In [BGV23], an irreducibility-preserving assignment is defined by requiring $f|_{\mathbf{x}_I = \boldsymbol{\sigma}_I}$ to be irreducible, without explicitly excluding constants. This difference reflects a variation in conventions: [BGV23] treats constant polynomials as reducible, whereas we do not. Accordingly, to match our convention and to ensure the validity of the forthcoming claim, we require $f|_{\mathbf{x}_I = \boldsymbol{\sigma}_I}$ to be both non-constant and irreducible.

Claim 2.29 ([BGV23, Claim 2.5])**.** *Let $f \in \mathbb{F}[\mathbf{x}]$ and $\boldsymbol{\alpha} \in \mathbb{F}^n$. If f is $\boldsymbol{\alpha}$-irreducible then f is $\boldsymbol{\alpha}$-justified.*

The next claim follows from Lemma 2.25.

Claim 2.30 ([BGV23, Claim 2.23])**.** *Let $f \in \mathbb{F}[\mathbf{x}]$ be a multilinear polynomial and $\boldsymbol{\alpha} \in \mathbb{F}^n$ such that for every $i \neq j \in [n]$: $\Delta_{i,j}(f)(\boldsymbol{\alpha}) \neq 0$ and $f(\boldsymbol{\alpha}) \neq 0$. Then, f is $\boldsymbol{\alpha}$-irreducible.*

2.4.3 Degree Preserving Assignments

The following definition can be viewed as a generalization of **0**-justification to the setting of non-multilinear polynomials.

Definition 2.31 (Degree-preserving assignment)**.** Let $f \in \mathbb{F}[\mathbf{x}]$ and let $\boldsymbol{\alpha} \in \mathbb{F}^n$. We say that $\boldsymbol{\alpha}$ is a *degree-preserving assignment* for f (equivalently, f is $\boldsymbol{\alpha}$*-preserved*) if the following conditions hold:

1. For every $I \subseteq [n]$ and every $x_j \in \mathbf{x}_{-I}$, $\deg_{x_j}(f|_{\mathbf{x}_I = \boldsymbol{\alpha}_I}) = \deg_{x_j}(f)$.
2. $f(\boldsymbol{\alpha}) \neq 0$. ◊

It is not hard to see that every degree-preserving assignment is also a justifying assignment.

A useful property of degree-preserving assignments, which does not generally hold for justifying assignments of non-multilinear polynomials, is the following.

Lemma 2.32. *Let $f \in \mathbb{F}[\mathbf{x}]$ be $\boldsymbol{\alpha}$-preserved, and let $q \in \mathbb{F}[\mathbf{x}]$ be a factor of f. Then $\boldsymbol{\alpha}$ is also a degree-preserving assignment for q.*

Proof. Let $p \in \mathbb{F}[\mathbf{x}]$ with $f = p \cdot q$. Since $f(\boldsymbol{\alpha}) \neq 0$, it follows that $q(\boldsymbol{\alpha}) \neq 0$. Now fix $I \subset [n]$ and $x_j \in \mathbf{x}_{-I}$. To see that $\deg_{x_j}(q) = \deg_{x_j}(p|_{I \leftarrow \boldsymbol{\alpha}_I})$, observe that

$$\deg_{x_j}(p) + \deg_{x_j}(q) = \deg_{x_j}(f) = \deg_{x_j}(f|_{I \leftarrow \boldsymbol{\alpha}_I}) = \deg_{x_j}(p|_{I \leftarrow \boldsymbol{\alpha}_I}) + \deg_{x_j}(q|_{I \leftarrow \boldsymbol{\alpha}_I}). \qquad \square$$

Recall that Proposition 2.23(1) shows that if $\boldsymbol{\alpha} \in \mathbb{F}^n$ is a common nonzero of $\{\partial_{x_i} f\}_{i \in [n]}$, then f is $\boldsymbol{\alpha}$-justified. The situation for degree-preserving assignments is analogous: the relevant set of polynomials to hit is given by the partial derivatives with respect to the degree, namely $\{\partial_{x_i^{\deg}} f\}_{i \in [n]}$.

Lemma 2.33. *Let $f \in \mathbb{F}[\mathbf{x}]$ and $\boldsymbol{\alpha} \in \mathbb{F}^n$ such that $\partial_{x_i^{\deg}} f(\boldsymbol{\alpha}) \neq 0$ for every $i \in [n]$. Then $\boldsymbol{\alpha}$ is a degree-preserving assignment for f.*

Proof. Let $I \subset [n]$ and $x_j \in \mathbf{x}_{-I}$. Since partial differentiation with respect to x_j commutes with substitution of other variables, and since $\partial_{x_i^{\deg}} f(\boldsymbol{\alpha}) \neq 0$, we have

$$\partial_{x_j^{\deg}}\left(f|_{I \leftarrow \boldsymbol{\alpha}_I}\right) = \left(\partial_{x_j^{\deg}} f\right)|_{I \leftarrow \boldsymbol{\alpha}_I} \not\equiv 0,$$

by choice of $\boldsymbol{\alpha}$. Therefore, $\deg_{x_j}(f|_{I \leftarrow \boldsymbol{\alpha}_I}) \geqslant \deg_{x_j}(f)$, which implies equality. $\square$

We will use the following simple fact implicitly throughout the work.

Fact 2.34. *If $\boldsymbol{\alpha} \in \mathbb{F}^n$ is a degree-preserving (resp. justifying, irreducibility-preserving) assignment for $f \in \mathbb{F}[\mathbf{x}]$, then $f(\mathbf{x} + \boldsymbol{\alpha})$ is $\mathbf{0}$-preserved (resp. $\mathbf{0}$-justified, $\mathbf{0}$-irreducible).*

2.5 Resultant

Given two polynomials $f, g \in \mathbb{F}[\mathbf{x}, y]$, we can treat f, g as polynomials over $\mathbb{F}(\mathbf{x})[y]$ and define $\gcd_y(f, g)$ to be the unique monic polynomial (in y) of maximal degree that divides both f and g.

Let $f(y) = \sum_{i=0}^{d} a_i y^i$ and $g(y) = \sum_{j=0}^{e} b_j y^j$, where each $a_i, b_j \in \mathbb{F}[\mathbf{x}]$. The *Sylvester matrix* of f and g with respect to y, which we denote as $M_y^{\mathrm{syl}}(f, g)$ is the following $(d + e) \times (d + e)$ matrix:

$$M_y^{\mathrm{syl}}(f, g) = \begin{bmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \ddots & \vdots & b_1 & b_0 & \ddots & \vdots \\ \vdots & a_1 & \ddots & 0 & \vdots & b_1 & \ddots & 0 \\ a_d & \vdots & \ddots & a_0 & b_e & \vdots & \ddots & b_0 \\ 0 & a_d & \ddots & a_1 & 0 & b_e & \ddots & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_d & 0 & 0 & \cdots & b_e \end{bmatrix}.$$

Definition 2.35 (Resultant)**.** For polynomials $f, g \in \mathbb{F}[\mathbf{x}]$, the resultant $\mathrm{Res}_y(f, x) \in \mathbb{F}[\mathbf{x}]$ is defined as the determinant of the Sylvester matrix $M_y^{\mathrm{syl}}(f, g)$. That is, $\mathrm{Res}_y(f, g) = \det(M_y^{\mathrm{syl}}(f, g))$. $\lozenge$

We will use the following properties of the resultant:

Fact 2.36 (See [GCL92], [vzGG99], [CLO15])**.** *Let $f, g \in \mathbb{F}[\mathbf{x}]$. Then,*

1. *$\gcd_y(f, g) \neq 1$ if and only if $Res_y(f, g) \equiv 0$. That is, f and g have a non-trivial factor that depends on the variable y (i.e., $\deg_y(\gcd(f, g)) > 0$) if and only if the resultant with respect to y of f and g is the identically zero polynomial.*
2. *Let $\boldsymbol{\alpha} \in \mathbb{F}^n$. If $\deg_y(f) = \deg_y(f|_{\mathbf{x}=\boldsymbol{\alpha}})$ and $\deg_y(g) = \deg_y(g|_{\mathbf{x}=\boldsymbol{\alpha}})$, then $Res_y(f, g)|_{\mathbf{x}=\boldsymbol{\alpha}} = Res_y(f|_{\mathbf{x}=\boldsymbol{\alpha}}, g|_{\mathbf{x}=\boldsymbol{\alpha}})$.*

2.6 Hardness of Representation

We begin by defining the notion of hardness of representation. In previous works, two equivalent but distinct definitions were given. In this work, we adopt the definition of [SV15], presented here using the notation of [BGV23].

Definition 2.37 (Hardness of Representation)**.** A polynomial $p \in \mathbb{F}[\mathbf{x}]$ is said to be m-*hard* if no monomial with support size at least m divides p. ◊

The following simple observation, noted in earlier works [BGV23, SV15, AvMV15], states that a polynomial that remains m-hard under any $\mathbf{0}$-substitution must contain a monomial whose support is of size at most $m-1$.

Observation 2.38. *Let $n, m \in \mathbb{N}$, and let $f \in \mathbb{F}[\mathbf{x}]$. Assume that for every $J \subseteq [n]$, the polynomial $f|_{J \leftarrow \mathbf{0}}$ is m-hard. Then f contains a monomial $M \in mon(f)$ with $|\mathrm{Supp}(M)| < m$. In particular, $f \circ \mathcal{G}_{m-1} \not\equiv 0$.*

Proof. Assume for contradiction that every monomial in $\mathrm{mon}(f)$ has support size at least m. Let $M \in \mathrm{mon}(f)$ be a monomial of minimal support size, and set $I = \mathrm{Supp}(M)$. Then $\mathcal{P}_I \mid f|_{[n]\setminus I \leftarrow \mathbf{0}}$, contradicting the assumption that $f|_{[n]\setminus I \leftarrow \mathbf{0}}$ is m-hard.

This establishes the first claim. The fact that $f \circ \mathcal{G}_{m-1} \not\equiv 0$ follows from Corollary 2.20. □

Weaker versions of our next lemma appeared in [SV15, Sha22].

Lemma 2.39. *Let $f \in \mathbb{F}[\mathbf{x}]$. Suppose there exist $I \subseteq [n]$ and $\mathbf{d} \in \mathbb{N}^I$ such that $\partial_{I^{\mathbf{d}}} f(\boldsymbol{\sigma}) \neq 0$ for some $\boldsymbol{\sigma} \in \mathbb{F}^n$. Then the polynomial $g(\mathbf{x}) = f(\mathbf{x} + \boldsymbol{\sigma})$ is $(|I|+1)$-hard.*

Proof. Assume, towards a contradiction, that $f(\mathbf{x}+\boldsymbol{\sigma})$ is not $(|I|+1)$-hard. Without loss of generality, suppose $1 \notin I$ and that $x_1 \mid f(\mathbf{x}+\boldsymbol{\sigma})$. Then there exists $h \in \mathbb{F}[\mathbf{x}]$ with $f(\mathbf{x}+\boldsymbol{\sigma}) = x_1 h$. Thus

$$0 \neq \partial_{I^{\mathbf{d}}} f(\boldsymbol{\sigma}) = \left(\partial_{I^{\mathbf{d}}} f(\mathbf{x}+\boldsymbol{\sigma})\right)\big|_{\mathbf{x}=\mathbf{0}} = \left(\partial_{I^{\mathbf{d}}}(x_1 h)\right)\big|_{\mathbf{x}=\mathbf{0}} = \left(x_1 \cdot \partial_{I^{\mathbf{d}}} h\right)\big|_{\mathbf{x}=\mathbf{0}} = 0. \qquad \square$$

The following is another simple and useful lemma.

Lemma 2.40. *Let $f \in \mathbb{F}[\mathbf{x}]$. Suppose there exist $I \subseteq [n]$ and $\mathbf{d} \in \mathbb{N}^I$ such that for some $\boldsymbol{\sigma} \in \mathbb{F}^n$ $\partial_{I^{\mathbf{d}}} f(\boldsymbol{\sigma}) \neq 0$. Then, $f(\mathbf{x}+\boldsymbol{\sigma}) \circ \mathcal{G}_{|I|} \not\equiv 0$.*

Proof. We prove the claim for $|I| = 1$, the statement for bigger sets then follows by using the argument inductively.

Assume for some $i \in [n]$ and $d \in \mathbb{N}$ we have that $\partial_{x_i^d} f(\boldsymbol{\sigma}) \not\equiv 0$. Let

$$f(\mathbf{x}) = \sum_{j=0}^{\deg_{x_i}(f)} x_i^j \cdot g(\mathbf{x}_{-i}).$$

Using the assumption that $\partial_{x_i^d} f(\boldsymbol{\sigma}) \not\equiv 0$ we get for $c_{j,d} = j \cdot (j-1) \cdots (j-d)$

$$\partial_{x_i^d} \sum_{j=0}^{\deg_{x_i}(f)} x_i^j \cdot g_j(\boldsymbol{\sigma}) = \sum_{j=d}^{\deg_{x_i}(f)} c_{j,d} \cdot x_i^{j-d} \cdot g_j(\boldsymbol{\sigma}) \not\equiv 0.$$

This implies that there exists $j' \geqslant d$ (although $j' \geqslant 0$ suffices) such that $g_{j'}(\boldsymbol{\sigma}) \not\equiv 0$.

Using Observation 2.19 to revive x_i we get

$$f(\mathcal{G}_1|_{y=\alpha_i, z=x_i-\sigma_i} + \boldsymbol{\sigma}) = \sum_{\deg_{x_i}(f) \geqslant j > j'} x^j g_j(\boldsymbol{\sigma}_{-i}) + x^{j'} g_{j'}(\boldsymbol{\sigma}_{-i}) + \sum_{j' > j \geqslant 0} x^j g_j(\boldsymbol{\sigma}_{-i}) \not\equiv 0$$

where the last inequality is due to linear independence of different monomials. □

We now turn to several useful known results concerning hardness of representation for bounded read formulas. We begin with the main hardness result of [SV15].

Theorem 2.41 ([SV15, Theorem 6.1]). *Let* $0 \not\equiv F \in \sum^k \mathsf{ROF}$ *and assume every summand is* **0***-justified. Then* F *is* $3k$*-hard.*

We begin with results for structurally multilinear polynomials, studied in [AvMV15].

Before moving to the lemma, the set of non-vanishing polynomials used in it appears frequently in our proofs. Therefore, we provide it with a dedicated definition.

Definition 2.42 ($\mathcal{Q}_{F_1,F_2,\ldots,F_m}$). Let $F_1, F_2, \ldots, F_m$, be algebraic formulas such that F_i is an $\mathsf{Rk}_i\mathsf{F}$ and let $k = \sum_{i\in[m]} k_i$. Define

$$b_{m,k} \coloneqq (k - m + 1) \cdot 4k \cdot (k+1)^2 .$$

Then we define

$$\mathcal{Q}_{F_1,F_2,\ldots,F_m} \coloneqq \left\{ \partial_{Y^{\deg}}(p_w)\big|_{Z\leftarrow \mathbf{0}} \;\middle|\; w \in \bigcup_{i\in[m]} F_i,\ Z \subseteq [n],\ Y \subseteq [n]\setminus Z,\ |Z \cup Y| \leqslant b_{m,k} \right\}.$$

In words, F_w is a subformula of some F_i, and $Z, Y \subseteq [n]$ are disjoint subsets satisfying $|Z \cup Y| \leqslant b_{m,k}$. ◊

Lemma 2.43 ([AvMV15, Lemma 5.3]). *Let* $F = \gamma + \sum_{i\in[m]} F_i$ *be a structurally multilinear formula, where* $\gamma \in \mathbb{F}$ *and each* $F_i \in \mathsf{Rk}_i\mathsf{F}$. *Set* $k = \sum_{i\in[m]} k_i$. *Let* $\boldsymbol{\sigma} \in \mathbb{F}^n$ *be a common nonzero to* $\mathcal{Q}_{F_1,F_2,\ldots,F_m}$. *Then*

$$p_F(\mathbf{x} + \boldsymbol{\sigma}) \quad \textit{is } r_{m,k}\textit{-hard},$$

for $r_{m,k} \coloneqq \left(8k \cdot (k+1)^2\right)^{k-m+1}$.

Theorem 2.44 ([AvMV15, Theorem 6.3]). *Let* $0 \not\equiv F \in \mathsf{RkF}$ *be structurally multilinear. Then, for some function* $r_k = k^{O(k)}$

$$p_F \circ \mathcal{G}_{r_k + k\log n} \not\equiv 0 .$$

Remark 2.45. In the above theorem, the function r_k is the same as the one used in Theorem 1.6.

The model $\sum^2\prod\mathsf{ROF}$ was studied in [BGV23], where techniques based on resultants were developed. These techniques will play an important role in our analysis of *non-structural* R4Fs. We begin with the simpler case of $\sum^2\mathsf{ROF}$, for which a polynomial-time PIT algorithm was given in [SV15], and subsequently improved in [BGV23].

Lemma 2.46. ([BGV23, Fact 3.9]) *Let* p_1, p_2 *be* **0***-justified* ROPs. *If* $p_1 + p_2 \neq 0$ *then* $p_1 + p_2$ *is* 3*-hard.*

Using this lemma, we prove the next result, which we then use throughout this work.

Corollary 2.47. *Let* $F_1, F_2 \in \mathsf{ROF}$ *such that* $p_{F_1} + p_{F_2} \not\equiv 0$. *Then* $(p_{F_1} + p_{F_2}) \circ \mathcal{G}_3 \not\equiv 0$.

Proof. Denote by S the set of nonzero polynomials in $\{\partial_{x_i} p_{F_j}\}_{i\in[n],\, j\in\{1,2\}}$. Lemma 2.11 implies that $S \subseteq \mathsf{ROF}$. By Theorem 2.15, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\mathcal{G}_1)$ such that $p(\boldsymbol{\alpha}) \neq 0$ for every $p \in S$. From Proposition 2.23(1) we get that both $F_1(\mathbf{x} + \boldsymbol{\alpha})$ and $F_2(\mathbf{x} + \boldsymbol{\alpha})$ are **0**-justified. Finally, Lemma 2.46 yields that $(F_1 + F_2)(\mathbf{x} + \boldsymbol{\alpha})$ is 3-hard, and the claim follows from Observation 2.38. □

We shall need the following hardness of representation result for the resultant of two ROFs.

Lemma 2.48 ([BGV23, Lemma 4.2]). *Let* $a, b \in \mathbb{F}[\mathbf{x}]$ *be* **0***-irreducible* ROPs *and assume* $n \geqslant 3$. *Let* $x_i \in [n]$ *be such that* $Res_{x_i}(a, b) \not\equiv 0$. *Then* $Res_{x_i}(a, b)$ *is 3-hard.*

2.7 Generic Assignments

In our proofs we shall often rely on *generic* assignments to the variables. Intuitively, one may think of a generic assignment as one that maintains the nonzeroness of all polynomials naturally associated with the given polynomial, such as its partial derivatives, commutators, and related constructs.

Definition 2.49 (Generic assignment)**.** An assignment $\boldsymbol{\alpha} \in \mathbb{F}^n$ is said to be *generic* with respect to a finite set of polynomials $\{f_i\}$, if $\boldsymbol{\alpha}$ is not a zero of any of the polynomials. ◊

In other words, there is a Zariski open set from which any assignment is "good". Whenever we will need to use generic assignments there will always be a well defined finite set of polynomials that the assignment should be generic with respect to them. This set will always be clear from the context.

The following are simple observations that we will use throughout this work:

Observation 2.50. *Let $f \in \mathbb{F}[\mathbf{x}]$ be* **0***-preserved. Then there is $\boldsymbol{\alpha} \in Z(f)$, which is nonzero in every coordinate.*

Proof. Let $\boldsymbol{\tau} \in \mathbb{F}^{[n-1]}$ be a generic assignment to all variables of f except x_n. Such an assignment is nonzero in every coordinate, and it preserves the **0**-justifiability property of f. Hence the restricted polynomial $f|_{[n-1]\leftarrow\boldsymbol{\tau}}$ remains **0**-justified. In particular, $x_n \nmid f$.

Let $\delta \in \mathbb{F}$ be a zero of $f|_{[n-1]\leftarrow\boldsymbol{\tau}}$. Then $(\boldsymbol{\tau}, \delta)$ is a zero of f that is nonzero in every coordinate, as required. □

Since this is the first time we consider generic assignments, we spell out explicitly the set of polynomials that $\boldsymbol{\tau}$ must not vanish on. Namely, $\boldsymbol{\tau}$ should not be a zero of any $f_i(\mathbf{x}) = x_i$, and it must also not annihilate the coefficient of x_n in f. From now on, we will specify the set of polynomials in the genericity assumption only when the setting is somewhat more involved.

Observation 2.51. *Let $f \in \mathbb{F}[\mathbf{x}]$ and $\boldsymbol{\tau} \in \mathbb{F}^n$ be a generic assignment. Then, $\boldsymbol{\tau}$ is degree preserving.*

It is also easy to see that generic assignment does not "ruin" any of the types of assignments defined in Subsection 2.4, as they all rely on nonzeroness of certain polynomials.

Lemma 2.52. *Let $f \in \mathbb{F}[x]$ be* **0***-justified (or* **0***-irreducible or* **0***-preserved). Then, a generic assignment $\boldsymbol{\alpha} \in \mathbb{F}^n$ satisfies that for every $I \subsetneq [n]$, $f|_{I\leftarrow\boldsymbol{\alpha}_I}$ is* **0***-justified (or* **0***-preserved). If $|I| \leqslant n-2$ then $f|_{I\leftarrow\boldsymbol{\alpha}_I}$ is* **0***-irreducible.*

Proof. The claim regarding **0**-justified and **0**-preserved polynomials is clear, as these properties are defined using non-vanishing of certain polynomials.

The claim about preserving irreducibility follows from the fact that if a polynomial is reducible then its coefficients satisfy a certain (finite) set of polynomials. See e.g. [Kal95]. □

We shall often use generic assignments to prove hardness of representation. The following lemma illustrates this approach.

Lemma 2.53. *Let $f \in \mathbb{F}[\mathbf{x}]$ be a nonzero polynomial, and let $\boldsymbol{\tau}$ be a generic assignment. Assume that one of the following holds:*

1. *For every $I \in \binom{[n]}{k_1+k_2}$ there exists a subset $S \subseteq I$ with $|S| \geqslant k_1$ such that $f\big|_{\mathbf{x}_{-S}=\boldsymbol{\tau}_{-S}}$ is k_1-hard. (Here the assignment is to the variables outside S.)*

2. *There exists a set* $B \subseteq [n]$ *with* $|B| \leqslant k_1$ *such that* $f\big|_{x_B=\tau_B}$ *is* k_2*-hard. (Here the assignment is to the variables in* B*.)*

Then f *is* (k_1+k_2)*-hard.*

Proof. Assume, for contradiction, that $\mathcal{P}_I \mid f$ for some $I \in \binom{[n]}{k_1+k_2}$. If (1) holds, let $S \subseteq I$ be as guaranteed. Then $\mathcal{P}_S \sim \mathcal{P}_I\big|_{x_{-S}=\tau_{-S}} \mid f\big|_{x_{-S}=\tau_{-S}}$, contradicting k_1-hardness. If (2) holds, then let B be the guaranteed set. We have $\mathcal{P}_{I\setminus B} \sim \mathcal{P}_I\big|_{x_B=\tau_B} \mid f\big|_{x_B=\tau_B}$, and since $|I \setminus B| \geqslant k_2$, this contradicts k_2-hardness. □

2.7.1 Restricting algebraic formulas by generic assignments

In this subsection we define the operation of restriction for algebraic formulas. This is obtained by substituting a generic assignment: Given an RkF F and a generic assignment $\tau \in \mathbb{F}^S$ for some $S \subseteq [n]$, we substitute the variables from S by their values in τ and simplify the formula by absorbing gates that have become constant into their parents. This process yields a new RkF formula, denoted F^τ_S, whose structure is derived from that of F.

The correspondence between the gates of F and those of the restricted formula F^τ_S is captured by a *gate map*

$$\mathcal{V}_S : F \longrightarrow F^\tau_S \cup \{\bot\},$$

where $\mathcal{V}_S(v)$ specifies the gate in F^τ_S that results from $v \in F$, and $\bot$ indicates that v vanishes (i.e., becomes constant) under the restriction. The map $\mathcal{V}_S$ depends only on F and the set S, and is identical for all generic assignments τ.

In addition, for each gate $u \in F^\tau_S$, we define its *origin gates* in F: the gates $v \in F$ such that $\mathcal{V}_S(v) = u$. Among these, the *origin upper gate* and *origin lower gate* are the highest and lowest such gates in F, respectively, giving rise to the origin maps

$$\mathcal{O}^F_S,\ \mathcal{L}^F_S : F^\tau_S \longrightarrow F.$$

The path in F from $\mathcal{L}^F_S(u)$ to $\mathcal{O}^F_S(u)$ is called the *restriction path* of u, and its length is the *restriction depth*, denoted $\mathcal{D}^F_S(u)$ (or simply $\mathcal{D}_S(u)$ when F is clear).

We next give a formal definition of this construction and establish several basic properties. Although the proofs are straightforward, they are included for completeness in Subsection A.1.

Definition 2.54 (Restriction of an Algebraic Formula and Origin Maps)**.** Let F be an algebraic formula on variables $x_1, \ldots, x_n$, and let $S \subseteq [n]$. For a generic assignment $\tau \in \mathbb{F}^S$, we define the *restriction* of F to S, denoted $F|^\tau_S$, together with the *restricted-gate map*

$$\mathcal{V}_S : F \longrightarrow F|^\tau_S \cup \{\bot\},$$

by the following recursion. Write $\overline{S} = [n] \setminus S$, and let $\alpha_v, \beta_v \in \mathbb{F}$ be the multiplicative and additive labels of each internal gate $v \in F$.

(1) Constant case. If $F_v\big|_{x_{\overline{S}}\leftarrow\tau_{\overline{S}}}$ is constant, then $F_v|^\tau_S$ is the constant *formula* computing this value, and we set $\mathcal{V}_S(v) = \bot$.

(2) No-reduction case. If both children remain non-constant after restriction,

$$F_{(v)_L}\big|_{x_{\overline{S}}\leftarrow\tau_{\overline{S}}} \notin \mathbb{F} \quad \text{and} \quad F_{(v)_R}\big|_{x_{\overline{S}}\leftarrow\tau_{\overline{S}}} \notin \mathbb{F},$$

then

$$F_v|^\tau_S \;=\; \alpha_v \cdot v^{\mathrm{op}}\Big(F_{(v)_L}|^\tau_S,\, F_{(v)_R}|^\tau_S\Big) \;+\; \beta_v, \qquad \mathcal{V}_S(v) \;=\; o_{F_v|^\tau_S}.$$

(3) Reduction case. Suppose exactly one child becomes constant after restriction. If $F_{(v)_R}\big|_{x_{\overline{S}} \leftarrow \boldsymbol{\tau}_{\overline{S}}} \notin \mathbb{F}$ but $F_{(v)_L}\big|_{x_{\overline{S}} \leftarrow \boldsymbol{\tau}_{\overline{S}}} \in \mathbb{F}$, then we set $\mathcal{V}_S(v) = \mathcal{V}_S((v)_R)$ and define

$$F_v|_S^{\boldsymbol{\tau}} = \begin{cases} \alpha_v F_{(v)_R}|_S^{\boldsymbol{\tau}} + \left(\alpha_v F_{(v)_L}\big|_{x_{\overline{S}} \leftarrow \boldsymbol{\tau}_{\overline{S}}} + \beta_v\right), & v^{op} = +, \\ \alpha_v \left(F_{(v)_L}\big|_{x_{\overline{S}} \leftarrow \boldsymbol{\tau}_{\overline{S}}}\right) \cdot \left(F_{(v)_R}|_S^{\boldsymbol{\tau}}\right) + \beta_v, & v^{op} = \times. \end{cases}$$

The symmetric case (left child non-constant, right child constant) is defined analogously.

Finally, set $F|_S^{\boldsymbol{\tau}} := F_o|_S^{\boldsymbol{\tau}}$.
When the assignment is clear, we write $F|_S$ and still view $\mathcal{V}_S$ as a map $F \to F|_S \cup \{\perp\}$.
For each gate $v \in F|_S^{\boldsymbol{\tau}}$, define two *origin maps*

$$\mathcal{O}_S^F, \mathcal{L}_S^F : F|_S^{\boldsymbol{\tau}} \longrightarrow F, \tag{1}$$

where $\mathcal{O}_S^F(v)$ (resp. $\mathcal{L}_S^F(v)$) is the highest (resp. lowest) gate $u \in F$ such that $\mathcal{V}_S(u) = v$. The *restriction path* of v is the path in F from $\mathcal{L}_S^F(v)$ to $\mathcal{O}_S^F(v)$; its length is the *restriction depth*, denoted $\mathcal{D}_S^F(v)$ (or simply $\mathcal{D}_S(v)$ when F is clear). ◊

Claim 2.55. *The procedure described in Definition 2.54 yields a well-formed algebraic formula $F|_S^{\boldsymbol{\tau}}$ (on the variables $\{x_i : i \in I\}$).*

Proof. The proof is by a simple induction on the structure of F. The only nontrivial case is where one of the children of the root F_o restricts to a constant. In this case, the fact that we restricted by a generic assignment implies that all leaves of that child are labeled by the variables in S and hence the restriction of the tree of the formula would also contract the entire child. □

Observation 2.56. *Let F be an algebraic formula, $S \subseteq [n]$, and $\boldsymbol{\tau} \in \mathbb{F}^S$. Let $v \in F|_S$ and $w = \mathcal{O}_S(v)$. Then*

1. *w is either o_F or the parent of w creates a new gate in $F|_S$,*
2. *w is the last gate that changes the constants in v, and*
3. *$p_{F_w|_S^{\boldsymbol{\tau}}} = p_v$.*

Lemma 2.57. *Let F be an algebraic formula, $S \subseteq [n]$ and $\boldsymbol{\tau} \in \mathbb{F}^S$. Let $v \in F|_S^{\boldsymbol{\tau}}$, $u = \mathcal{O}_S(v)$ and $w = \mathcal{L}_S(v)$. Denote by γ the additive constant of v and by δ the additive constant of w.*

1. *Let $G = F_u|_{w=\delta}$. Then $\gamma = p_{G|_S^{\boldsymbol{\tau}}}$.*
2. *$\left(F|_S^{\boldsymbol{\tau}}\right)|_{v=\gamma} = (F|_{w=\delta})|_S^{\boldsymbol{\tau}}$.*

2.8 Orbits

Given a matrix $\boldsymbol{A} \in \mathbb{F}^{n \times n}$ and variables $\boldsymbol{x}$, define

$$\boldsymbol{A}\boldsymbol{x} := \left(\sum_{i=1}^{n} A_{1,i}x_i, \sum_{i=1}^{n} A_{2,i}x_i, \ldots, \sum_{i=1}^{n} A_{n,i}x_i\right).$$

We denote by $\mathrm{GL}_n(\mathbb{F})$ the group of invertible $n \times n$ matrices over $\mathbb{F}$, and by $\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$ the group of invertible affine transformations—i.e., all maps defined by a pair $(\mathbf{A}, \boldsymbol{\beta})$, where $\mathbf{A} \in \mathrm{GL}_n(\mathbb{F})$ and $\boldsymbol{\beta} \in \mathbb{F}^n$, acting as

$$\mathbf{x} \longmapsto \mathbf{A}\mathbf{x} + \boldsymbol{\beta}\,.$$

Let $n \geqslant N \in \mathbb{N}$ and consider an N-variate polynomial $f(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$. The action of $(\mathbf{A}, \boldsymbol{\beta}) \in \mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$ on f is defined by

$$(\mathbf{A}, \boldsymbol{\beta}) \circ f(\mathbf{y}) = f(\mathbf{A}\mathbf{x} + \boldsymbol{\beta}) = f\left(\sum_{i=1}^{n} A_{1,i}x_i + \boldsymbol{\beta}_1, \sum_{i=1}^{n} A_{2,i}x_i + \boldsymbol{\beta}_2, \ldots, \sum_{i=1}^{n} A_{n,i}x_i + \boldsymbol{\beta}_n\right).$$

The orbit of the polynomial $f(\mathbf{y})$ under the action of $\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$, denoted $f^{\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})}$, is defined as

$$f^{\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})} \coloneqq \left\{f(\mathbf{A}, \boldsymbol{\beta}) \mid \mathbf{A} \in \mathrm{GL}_n(\mathbb{F}),\, \boldsymbol{\beta} \in \mathbb{F}^n\right\}.$$

Similarly, the *orbit* of a polynomial class $\mathcal{C}$ under the action of $\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$, denoted $\mathcal{C}^{\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})}$, is defined as

$$\mathcal{C}^{\mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})} \coloneqq \left\{f(\mathbf{A}, \boldsymbol{\beta}) \mid f \in \mathcal{C},\, \mathbf{A} \in \mathrm{GL}_n(\mathbb{F}),\, \boldsymbol{\beta} \in \mathbb{F}^n\right\}.$$

The action and orbits with respect to the group $\mathrm{GL}_n(\mathbb{F})$ are defined analogously by taking $\boldsymbol{\beta} = \mathbf{0}$.

3 Frontiers and Fragmentation

In this section we define the important notions of the RkF frontier and the set $\mathcal{F}^{(2)}$, consisting of the first gates above the frontier. We then present a lemma that simplifies an RkF via a well-chosen partial derivative, in the spirit of the fragmentation lemmas of [AvMV15, Sha22].[14]

Definition 3.1 (R$(k-1)$F Frontier)**.** Let F be an RkF. The R$(k-1)$F Frontier of F, denoted $\mathcal{F}_F$, is the set of the topmost gates that compute an R$(k-1)$P polynomial. That is, there is no other gate in the path from these gates to the root that computes an R$(k-1)$P polynomial. Whenever F is clear from context, we write $\mathcal{F}$ instead of $\mathcal{F}_F$. ◊

Definition 3.2 ($\mathcal{F}^{(2)}$)**.** Let F in RkF. Denote by $\mathcal{F}_F^{(2)}$ the set of all the gates in F which have two children in $\mathcal{F}_F$. ◊

Definition 3.3 (Unique set of a gate in an RkF, $U(v)$)**.** Let $F \in$ RkF and let $v \in F$. The *unique set* of v, denoted $U_F(v) \subseteq \mathbf{x}$, is the set of variables which F depends on and whose every leaf occurs in F_v. When F is clear from context, we write $U(v)$. ◊

Remark 3.4. In the blackbox model, we assume that every variable appearing in F is one on which F depends. Consequently, the distinction that $U(v)$ contains only such variables becomes relevant only in Subsection 8.2, where we design our whitebox algorithm.

Clearly, Definition 3.1 implies that for every $v \in \mathcal{F}^{(2)}$ we have $|U(v)| \geqslant 1$. One

The importance of these definitions stems from a simple observation that lies at the heart of our fragmentation results. If $x_i \in U(v)$ for some $v \in \mathcal{F}_F^{(2)}$, then taking the derivative with respect to x_i simplifies the formula to $\partial_{x_i^{\deg}} p_v$ multiplied by the product of the gates in $\mathrm{Unv}(v)$. Moreover,

[14] By *fragmentation* we refer to operations that simplify formulas belonging to a given class.

since $v \in \mathcal{F}_F^{(2)}$, it is either the sum or the product of two read-$(k-1)$ formulas, meaning that it possesses a somewhat simpler structure than F.

Our next definition extends the notion of the unique set of a gate v to capture all variables x_i for which taking the derivative with respect to x_i yields a structure similar to that obtained if x_i belonged to $U(v)$.

Definition 3.5 (Stable set of a gate in an RkF, $U^{(s)}(v)$)**.** Let $F \in \mathsf{RkF}$ and let $v \in F$. The *stable set* of v, denoted $U_F^{(s)}(v) \subseteq \mathbf{x}$, is the set of variables $x_i \in \mathrm{var}(v)$ such that

$$\partial_{x_i^{\deg}} p_F = \partial_{x_i^{\deg}} p_v \cdot f \quad \text{or} \quad \partial_{x_i^{\deg}} p_F = f\,,$$

where $f = \prod_{i=1}^{m} f_i$ is an RkP and the f_is are computed by distinct subformulas of gates in $\mathrm{Unv}_F(v)$. When F is clear from context, we write $U^{(s)}(v)$. ◊

The next observation follows immediately from Lemma 2.8.

Observation 3.6. *For every $F \in \mathsf{RkF}$ and $v \in F$, $U(v) \subseteq U^{(s)}(v)$.*

We next prove a fragmentation result showing that there exists a way to take a derivative with respect to some variable $x_i \in U^{(s)}(v)$, for a suitably chosen v, that simplifies the formula.

Lemma 3.7. *Let $F \in \mathsf{RkF}$ and set $t \coloneqq |\mathcal{F}_F^{(2)}|$. Then there exists a vertex $v \in \mathcal{F}_F^{(2)}$ such that for every $x_i \in U^{(s)}(v)$,*

$$\partial_{x_i^{\deg}} p_F \;=\; \left(\partial_{x_i^{\deg}} p_v\right) \cdot \prod_{j=1}^{m} p_{F_j} \quad \textit{or} \quad \partial_{x_i^{\deg}} p_F \;=\; \prod_{j=1}^{m} p_{F_j}$$

where each F_j is a subformula of F which is either an $\mathsf{R}(k-1)\mathsf{F}$ or with $|\mathcal{F}_{F_j}^{(2)}| \leqslant \frac{t}{2}$.

Proof. We first describe how to choose v. Starting from the root gate, repeatedly descend to the child whose subtree contains the largest number of gates from $\mathcal{F}_F^{(2)}$. This process terminates at some $v \in \mathcal{F}_F^{(2)}$, which is the required gate. Every vertex not chosen along this path has at most $t/2$ vertices of $\mathcal{F}_F^{(2)}$ in its subtree. The claim now follows from the definition of the stable set of v and from the choice of v. □

The usefulness of this lemma is demonstrated in the next claim, which shows how PIT for RkF can be reduced to PIT for simpler formulas.

Lemma 3.8. *Let F be a nonzero RkF and set $t \coloneqq |\mathcal{F}_F^{(2)}|$. Write $\mathcal{F}_F^{(2)} = \{v_j\}_{j=1}^{t}$, and let $I = \{i_j\}_{j=1}^{t}$ with $x_{i_j} \in U^{(s)}(v_j)$. Denote*

$$\mathcal{D}_I \coloneqq \left\{ \partial_{x_{i_j}^{\deg}} p_{v_j} : j = 1, \ldots, t \right\}.$$

If $\mathcal{H}$ is a hitting set for $\mathcal{D}_I \cup \mathsf{R}(k-1)\mathsf{F}$, then[15]

$$F \circ \left(\mathcal{H} + \mathcal{G}_{\log t + 1}\right) \not\equiv 0.$$

[15] Whenever we write a sum of generators, we mean they are defined over disjoint sets of variables.

Proof. If $t = 0$, then $F \in \mathsf{R}(k-1)\mathsf{F}$, and the claim follows from the assumption on $\mathcal{H}$. If $t = 1$, let $\mathcal{F}^{(2)} = \{v\}$ and $I = \{i\}$. By Definition 3.5 we have

$$\partial_{x_i^{\deg}} p_F = \left(\partial_{x_i^{\deg}} p_v\right) \cdot \prod_{j=1}^{m} p_{F_j} \quad \text{or} \quad \partial_{x_i^{\deg}} p_F = \prod_{j=1}^{m} p_{F_j},$$

where each F_j is an $\mathsf{R}(k-1)\mathsf{F}$ (as $t = 1$), hence it gets hit by $\mathcal{H}$. Since $\partial_{x_i^{\deg}} p_v \circ \mathcal{H} \not\equiv 0$, we get that $\partial_{x_i^{\deg}} p_F \circ \mathcal{H} \not\equiv 0$. The claim is then implied by Lemma 2.40.

Assume $t > 1$ and proceed by induction on t. Let v_1 be the gate guaranteed by Lemma 3.7. Then for $x_{j_1} \in U^{(s)}(v_1)$,

$$\partial_{x_{j_1}^{\deg}} p_F = \left(\partial_{x_{j_1}^{\deg}} p_{v_1}\right) \cdot \prod_{j=1}^{m} p_{F_j} \quad \text{or} \quad \partial_{x_{j_1}^{\deg}} p_F = \prod_{j=1}^{m} p_{F_j},$$

where each F_j is an RkF that is either an $\mathsf{R}(k-1)\mathsf{F}$, or it satisfies $|\mathcal{F}_{F_j}^{(2)}| \leqslant t/2$. If F_j is an $\mathsf{R}(k-1)\mathsf{F}$, then by the property of $\mathcal{H}$, $p_{F_j} \circ \mathcal{H} \not\equiv 0$. If $|\mathcal{F}_{F_j}^{(2)}| \leqslant t/2$, then by the induction hypothesis, $p_{F_j} \circ \left(\mathcal{H} + \mathcal{G}_{\log t/2+1}\right) \not\equiv 0$. Hence

$$\prod_{j=1}^{m} p_{F_j} \circ \left(\mathcal{H} + \mathcal{G}_{\log t}\right) \not\equiv 0.$$

By the assumption on $\mathcal{H}$, $\partial_{x_{j_1}} p_F \circ \mathcal{H} \not\equiv 0$, and thus

$$\left(\partial_{x_{j_1}^{\deg}} p_F\right) \circ \left(\mathcal{H} + \mathcal{G}_{\log t}\right) \not\equiv 0.$$

The claim now follows from Lemma 2.40. □

Observe that for every $F \in \mathsf{RkF}$, each $v \in \mathcal{F}^{(2)}$ satisfies $U_F(v) \neq \emptyset$, hence $|\mathcal{F}_F^{(2)}| \leqslant n$. Moreover, if equality holds, then F is a PROP (a ROF with univariate polynomials substituted at its leaves, see [SV15]). By [MV18], in this case $F \equiv 0 \iff F \circ \mathcal{G}_1 \equiv 0$. This yields the following weaker corollary.

Lemma 3.9. *In the setting of Lemma 3.8,* $F \circ \left(\mathcal{H} + \mathcal{G}_{\log n+1}\right) \not\equiv 0$.

We shall apply this lemma in Section 4 to reduce PIT for structural RkFs (see Definition 2.4) to PIT for structural $\sum^2 \mathsf{R}(k-1)\mathsf{F}$s (see Definition 2.2).

4 Blackbox PIT for R2F and R3F

In this section, we present the *first* hitting-set generator (HSG) with logarithmic seed length for R3Fs. We also slightly improve the HSG from Theorem 2.21 for R2Fs. To achieve this, we leverage the fact established in Corollary 2.6 that both R2Fs and R3Fs are structural bounded-read formulas. For such formulas, the polynomials in the set $\mathcal{D}_I$ of Lemma 3.8 possess a simple structure.

Observation 4.1. *Let* $F \in \mathsf{RkF}$ *be structural, and let* $\mathcal{H}$ *be a hitting-set generator that hits every polynomial that is a structural* RkF *of the form* $\sum^2 \mathsf{R}(k-1)\mathsf{F}$. *Then, for every* $v \in \mathcal{F}_F^{(2)}$ *and* $x_i \in U^{(s)}(v)$, *we have* $\partial_{x_i^{\deg}} p_v \circ \mathcal{H} \not\equiv 0$.

Proof. Since $v \in \mathcal{F}_F^{(2)}$, we have $F_{(v)_L}, F_{(v)_R} \in \mathsf{R}(k-1)\mathsf{F}$. We proceed by cases according to v^{op}.
Case 1: $v^{op} = \times$. We have

$$\partial_{x_i^{\deg}} p_v \sim \left(\partial_{x_i^{\deg}} p_{(v)_L}\right)\left(\partial_{x_i^{\deg}} p_{(v)_R}\right) \in \prod^{2} \mathsf{R}(k-1)\mathsf{F}\,,$$

where the inclusion on the right-hand side follows from Lemma 2.11. The claim then follows from the properties of $\mathcal{H}$.
Case 2: $v^{op} = +$. Let $d = \deg_{x_i}(p_v)$. We have

$$\partial_{x_i^{\deg}} p_v = \partial_{x_i^d} p_{(v)_L} + \partial_{x_i^d} p_{(v)_R}\,.$$

Set $p_L = \partial_{x_i^d} p_{(v)_L}$ and $p_R = \partial_{x_i^d} p_{(v)_R}$. Since $\deg_{x_i}(p_{(v)_L}), \deg_{x_i}(p_{(v)_R}) \leqslant \deg_{x_i}(p_v) = d$, for $j \in \{L, R\}$, p_j is either 0 or $\partial_{x_i^{\deg}} p_{(v)_j}$. Therefore, by Lemma 2.11, we have

$$\partial_{x_i^{\deg}} p_v \in \mathsf{RkF} \cap \sum^{2} \mathsf{R}(k-1)\mathsf{F}\,.$$

The claim again follows from the properties of $\mathcal{H}$. □

4.1 Improved PIT for R2F

Using Observation 4.1, and the fact that $\mathcal{G}_3$ hits $\sum^2 \mathsf{ROF}$, we get the next theorem with a constant factor improvement upon Theorem 2.21.

Theorem 1.4. *Let $F \in \mathsf{R2F}$ be nonzero. Then*

$$p_F \circ \mathcal{G}_{\log n+4} \not\equiv 0\,.$$

Proof. By Corollary 2.6, the formula F is structural. Using Observation 4.1 together with Corollary 2.47, we get that for every $v \in \mathcal{F}^{(2)}$ and every $x_i \in U(F_v)$,

$$\partial_{x_i^{\deg}} p_v \circ \mathcal{G}_3 \not\equiv 0.$$

The claim now follows from Lemma 3.8 and Observation 3.6. □

4.2 PIT for R3F

Recall the definition of the set $\mathcal{Q}$, see Definition 2.42.

Proposition 4.2. *Let $F \in \mathsf{R3F}$ of the form $\sum^2 \mathsf{R2F}$ be nonzero. Suppose that every nonzero polynomial in $\mathcal{Q}_{F_{(o)_L}, F_{(o)_R}}$ does not vanish at $\mathbf{0}$. Then p_F is $r_{2,4}$-hard, where $r_{2,4}$ is defined in Lemma 2.43.*

Proof. If F is multilinear, then it is structurally multilinear by Corollary 2.6. Therefore, by the assumption on $\mathcal{Q}_{F_{(o)_L}, F_{(o)_R}}$ and by Lemma 2.43, we conclude that p_F is $r_{2,4}$-hard.

Otherwise, choose a variable x_i of degree 2 in F. Since F is read-3, without loss of generality we may assume that $\deg_{x_i}((o)_L) = 2 > \deg_{x_i}((o)_R)$. Hence,

$$\partial_{x_i^{\deg}} p_F(\mathbf{0}) = \left(\partial_{x_i^2} p_{(o)_L} + \partial_{x_i^2} p_{(o)_R}\right)(\mathbf{0}) = \partial_{x_i^{\deg}} p_{(o)_L}(\mathbf{0}) \neq 0,$$

where the last inequality follows from the assumption on $\mathcal{Q}_{F_{(o)_L}, F_{(o)_R}}$. By Lemma 2.39, p_F is 2-hard, and since $r_{2,4} \geqslant 2$, the claim follows. □

Lemma 4.3. *Let* $F \in \mathsf{R3F}$ *of the form* $\sum^2 \mathsf{R2F}$ *be nonzero. Then,*

$$p_F \circ \mathcal{G}_{\log n + r_{2,4}+3} \not\equiv 0\,.$$

Proof. Observe that every polynomial in $\mathcal{Q}_{F_{(o)_L},F_{(o)_R}}$ is an $\mathsf{R2F}$ by Corollary 2.6 and Lemma 2.11. Therefore, by Theorem 1.4, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\mathcal{G}_{\log n+4})$ which is a common nonzero for all nonzero polynomials in $\mathcal{Q}_{F_{(o)_L},F_{(o)_R}}$. Proposition 4.2 then implies that $p_F(\boldsymbol{x}+\boldsymbol{\alpha})$ is $r_{2,4}$-hard.

By Observation 2.38, we have

$$p_F(\boldsymbol{x}+\boldsymbol{\alpha}) \circ \mathcal{G}_{r_{2,4}-1} \not\equiv 0\,.$$

The claim follows from Fact 2.16. □

We are now ready to prove the result for general $\mathsf{R3F}$s.

Theorem 1.3. *Let* $F \in \mathsf{R3F}$ *be nonzero. Then, there exists a constant* $c_{1.3}$

$$p_F \circ \mathcal{G}_{2\log n + c_{1.3}} \not\equiv 0\,.$$

Proof. By combining Observation 4.1 with Lemma 4.3, we obtain that $\mathcal{G}_{\log n + r_{2,4}+3}$ serves as a hitting-set generator for the polynomials

$$\left\{ \partial_{x_i^d} p_\nu : \nu \in \mathcal{F}_F^{(2)},\ i \in U(\nu) \right\}.$$

Therefore, by Lemma 3.9 and Observation 3.6,

$$p_F \circ \left(\mathcal{G}_{\log n + r_{2,4}+3} + \mathcal{G}_{\log n+1}\right) \not\equiv 0\,.$$

The claim follows from Fact 2.16. □

5 Structural Read-4 Formulas

In this section we prove a hardness of representation result for structural $\mathsf{R4F}$s which are also $\sum^2 \mathsf{R3F}$. The following is our main hardness of representation lemma. Recall the definition of $r_{m,k}$ from Lemma 2.43.

Proposition 5.1. *Let* F, G *be* $\mathsf{R3F}$*s such that the formula* $F+G$ *is a structural* $\mathsf{R4F}$*. Let* U *be a set containing all the read-4 variables in the formula* $F+G$*.*

Denote

$$\mathcal{T} := \left\{ \partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p\right) \;\middle|\; p \in \{p_F, p_G\},\ x_i, x_j \in \mathit{var}(p) \right\}.^{16}$$

Assume $U \neq \emptyset$ *and that no nonzero polynomial in* $\mathcal{Q}_{F,G} \cup \mathcal{T}$ *vanishes at* $\mathbf{0}$*.*

Then, if $n \geqslant r_{2,6} + 5^8$*, there exists* $t \in U$ *such that* $\mathcal{P}_{[n]\setminus\{t\}} \nmid \partial_{x_t^{\deg}}(F+G)$*.*

[16] Observe that $\partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p\right) \neq \partial_{x_i^{\deg} x_j^{\deg}} p$. On the left-hand side, we differentiate with respect to x_i as many times as its degree in $\partial_{x_j^{\deg}} p$, whereas on the right-hand side, we differentiate as many times as its degree in p.

This proposition corresponds to the orange segment in Figure 2:

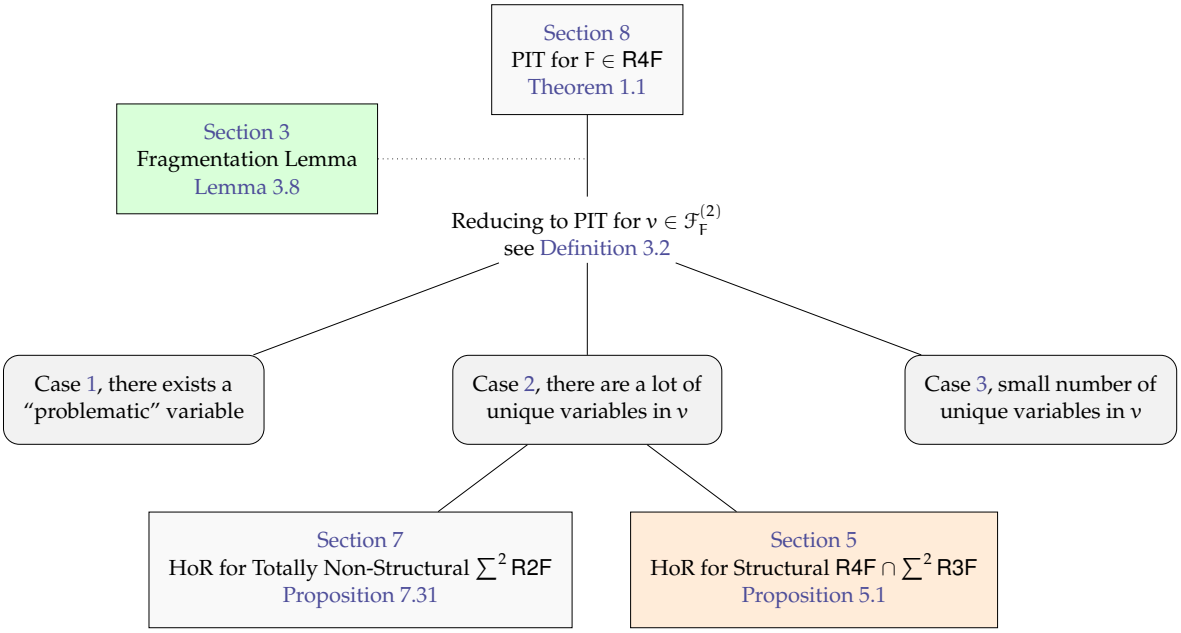

Figure 2: Our current position in the proof of Theorem 1.1

We need some preparatory work before proving Proposition 5.1.

5.1 Partial Split

In this section, we define the notion of a *partial split* of a formula. We then show that if a formula admits a partial split, it decomposes in a particular way that entails hardness of representation. The following lemma will be instrumental in establishing this result.

Lemma 5.2. *Let* $f_i, f_j, g_i, g_j, g' \in \mathbb{F}[\mathbf{x}]$ *such that* g' *is a* ROP. *Assume* $x_j \notin var(f_i) \cup var(g_i)$ *and* $x_i \notin var(f_j) \cup var(g_j)$. *Assume further that* $f_i f_j$ *and* $g_i g_j g'$ *are* **0***-preserved. Set* $F \coloneqq f_i f_j + g_i g_j g'$. *If* $n \geqslant 3$ *and* $\mathcal{P}_n \mid F$ *then for any* $\kappa \in [n] \setminus \{i, j\}$ *there exists* $p_\kappa \in \mathbb{F}[x_i, x_j, x_\kappa]$, *such that when restricting to the variables* $\{x_i, x_j, x_\kappa\}$ *we obtain*

$$F|_{\{i,j,\kappa\}} \;=\; p_\kappa \cdot \left(f_i' f_j' + \hat{g}\, g'\right),$$

where $f_i' \mid f_i$, $f_j' \mid f_j$, $\hat{g} \mid g_i g_j$, $\deg_\kappa(f_i') = \deg_\kappa(f_j') = \deg_\kappa(\hat{g}) = \deg_\kappa(g') = 1$, *and* f_i', f_j' *and* g' *are irreducible.*

Proof. Let $h \in \mathbb{F}[\mathbf{x}]$ satisfy $F = \mathcal{P}_n h$. We first note that if, for some $t \in [n]$, $\deg_t(f_i f_j) < \deg_t(g_i g_j g') = d$, then $\partial_{x_t^d} F = \partial_{x_t^d}(g_i g_j g') = \partial_{x_t^d}(\mathcal{P}_n h) \not\equiv 0$. Since $g_i g_j g'$ is **0**-preserved we get a contradiction by setting all variables in $[n] \setminus \{t\}$ to zero. Therefore $\mathrm{var}(f_i f_j) = \mathrm{var}(g_i g_j g') = [n]$ and the individual degrees are the same.

Next, we pick any $\{i, j, \kappa\}$ for $\kappa \in [n] \setminus \{i, j\}$ and restrict the polynomials to the variables $\{i, j, \kappa\}$. Observe that each of f_i, f_j depends on at most two variables. Additionally, we note that if $\gcd\left(f_i f_j, g_i g_j g'\right) = q(\mathbf{x}) \neq 1$ then since no variable divides $f_i f_j$ we must have $q(\mathbf{x}) \mid h$. Thus, dividing the equation by q does not change the situation (as factors of **0**-preserved polynomials are still **0**-preserved and so are factors of ROPs). Hence, from now on we shall assume that $\gcd\left(f_i f_j, g_i g_j g'\right) = 1$.

The last preparation step is by factoring out the irreducible factors that depend only on x_κ from $g_i g_j$. That is, we let $g_i', g_j', \tilde{g}$ be such that $g_i g_j = g_i' g_j' \tilde{g}$, each irreducible factor of g_i' involves x_i, each irreducible factor of g_j' involves x_j, and $\tilde{g}$ is a polynomial in $\mathbf{x}_\kappa$.

Claim 5.3. *It holds that both g_i', g_j' are constant polynomials.*

Proof. First, assume for a contradiction that g_j' is not a constant. Consider $F|_{x_i=0}$:

$$F|_{x_i=0} = f_i|_{x_i=0} f_j + g'_i|_{x_i=0} g'_j \tilde{g} g'|_{x_i=0} = (\mathcal{P}_n h)|_{x_i=0} \equiv 0. \tag{2}$$

It follows that $g_j' \mid f_j$, in contradiction to the assumption that their gcd is 1. □

If $\tilde{g}$ were constant, then (by the degree equality and since g' is a ROP) $f_i f_j$ would be multilinear. Hence x_κ appears in at most one of f_i, f_j, say in f_i. By Observation 2.3, f_i is a ROP; with f_j linear in x_j, the product $f_i f_j$ is a ROP. Therefore $\mathcal{P}_{\{i,j,\kappa\}}$ would divide a sum of two **0**-preserved ROPs, contradicting Lemma 2.46.

Next, consider (2) again. As $g_i' g_j' \in \mathbb{F}$ and $\gcd(f_j, \tilde{g}) = 1$, it must be the case that $\tilde{g} \mid f_i|_{x_i=0}$. Similarly we get $\tilde{g} \mid f_j|_{x_j=0}$. Recalling that the degrees in $f_i f_j$ and $\tilde{g} g'$ are equal we get

$$2\deg(\tilde{g}) \leqslant \deg_\kappa(f_i) + \deg_\kappa(f_j) = \deg(\tilde{g}) + \deg_\kappa(g') \leqslant \deg(\tilde{g}) + 1$$

and hence $1 \leqslant \deg_\kappa(\tilde{g}) \leqslant 1$. As $\tilde{g} \mid f_i|_{x_i=0}$, we see that $\deg_\kappa(f_i) \geqslant 1$ and similarly $\deg_\kappa(f_j) \geqslant 1$. By degree considerations we also have $\deg_\kappa(f_i) = \deg_\kappa(f_j) = \deg_\kappa(g') = 1$.

Claim 5.4. *The polynomials f_i, f_j and g' are irreducible.*

Proof. Assume towards a contradiction that f_i is reducible. Since f_i is linear in x_i and x_κ, any nontrivial factorization must split these variables, so we can write

$$f_i = f_i' \cdot f_\kappa',$$

where f_i' depends on x_i and f_κ' depends on x_κ. Consider $F|_{x_i=0}$:

$$F\big|_{x_i=0} = f_i'\big|_{x_i=0} f_\kappa' f_j + \tilde{g}\, g'\big|_{x_i=0} = \big(\mathcal{P}_n h\big)|_{x_i=0} \equiv 0.$$

By Lemma 2.32, f_i' is **0**-preserved, hence $f_i'\big|_{x_i=0}$ is nonzero. Therefore, either $\tilde{g} \mid f_\kappa'$ or $\tilde{g} \mid f_j$. In either case we contradict the hypothesis $\gcd(f_i f_j, \tilde{g} g') = 1$. Hence f_i must be irreducible. The argument for f_j is symmetric.

Now suppose g' is reducible, without loss of generality $g' = g_i' \cdot g_\kappa'$ where g_i' depends on x_i and g_κ' depends on x_κ. Consider $F|_{x_i=0}$:

$$F\big|_{x_i=0} = f_i\big|_{x_i=0} f_j + \tilde{g}\, g_i'\big|_{x_i=0} g_\kappa' = \big(\mathcal{P}_n h\big)\big|_{x_i=0} \equiv 0.$$

Arguing as before, this forces either $g_\kappa' \mid f_j$ or $\tilde{g} \mid f_j$, contradicting $\gcd(f_i f_j, \tilde{g} g') = 1$. Thus g' is irreducible. □

This concludes the proof of Lemma 5.2. □

The last lemma motivates the next definition.

Definition 5.5 (Partial Split)**.** Let $f, g \in \mathbb{F}[\mathbf{x}]$. We say that $t \in [n]$ partially splits f and g with respect to $i, j, \kappa \in [n]$ if

$$\partial_{x_t^{\deg}}(f + g) = f_i f_j + g_i g_j g'$$

such that

1. $x_i \notin \mathrm{var}(f_j) \cup \mathrm{var}(g_j)$ and $x_j \notin \mathrm{var}(f_i) \cup \mathrm{var}(g_i)$.
2. $g'|_{\{x_i,x_j,x_\kappa\}}$ is a ROP.
3. One of the following happens, $\kappa \notin \mathrm{var}(f_i)$, $\kappa \notin \mathrm{var}(f_j)$ or $\kappa \notin \mathrm{var}(g')$.

Since the roles of f and g are not symmetrical, we call f the *first polynomial* in the split. ◊

Corollary 5.6. *Let $f, g \in \mathbb{F}[\mathbf{x}]$ such that for some $t \in [n]$, t partially splits f and g with respect to some $i, j, \kappa \in [n]$. Moreover, assume that for every $x_j \in \mathbf{x}$ and $p \in \{f, g\}$, $\partial_{x_j^{\deg}} p$ is $\mathbf{0}$-preserved.*

Then,

$$\mathcal{P}_{[n]_{-t}} \nmid \partial_{x_t^{\deg}}(f+g)$$

Proof. Assume towards contradiction $\mathcal{P}_{[n]_{-t}} \mid \partial_{x_t^{\deg}}(f+g)$. By Definition 5.5 we get

$$\left(\partial_{x_t^{\deg}}(f+g)\right)\Big|_{\{i,j,\kappa\}} = f_i f_j + g_i g_j g'$$

for $f_i, g_i \in \mathbb{F}[x_i, x_\kappa]$, $f_j, g_j \in \mathbb{F}[x_j, x_\kappa]$ and a ROP g'. Moreover, by assumption $f_i f_j$ and $g_i g_j g'$ are $\mathbf{0}$-preserved.

Using Lemma 5.2 we get that f_i, f_j and g' all have an irreducible factor depending on x_κ, in contradiction to Property (3). □

A stronger version of a partial split is the next.

Definition 5.7 (Complete Split)**.** Let $f, g \in \mathbb{F}[\mathbf{x}]$. We say that $t \in [n]$ completely splits f and g with respect to $i, j \in [n]$ if

$$\partial_{x_t^{\deg}}(f+g) = f_i f_j + g_i g_j$$

such that $x_i \notin \mathrm{var}(f_j) \cup \mathrm{var}(g_j)$ and $x_j \notin \mathrm{var}(f_i) \cup \mathrm{var}(g_i)$. ◊

Remark 5.8. One can show that if t completely splits f and g and both have all of their derivatives being $\mathbf{0}$-preserved, then $\partial_{x_t^{\deg}}(f+g)$ is 2-hard.

Observation 5.9. *Let $f, g \in \mathbb{F}[\mathbf{x}]$ such that t completely splits f and g with respect to i and j. Then, for every $\kappa \in [n] \setminus \{i, j, t\}$, the variable t partially splits f and g with respect to i, j and κ.*

Proof. Choose $g' = 1$ and f_i, f_j, g_i, g_j according to the complete split. □

5.2 Splitting two R2Fs

In this subsection we focus on a specific type of R2Fs, which we call individually multiplicative. Figure 3 serves as an illustrative example for the core definitions presented in the following subsection.

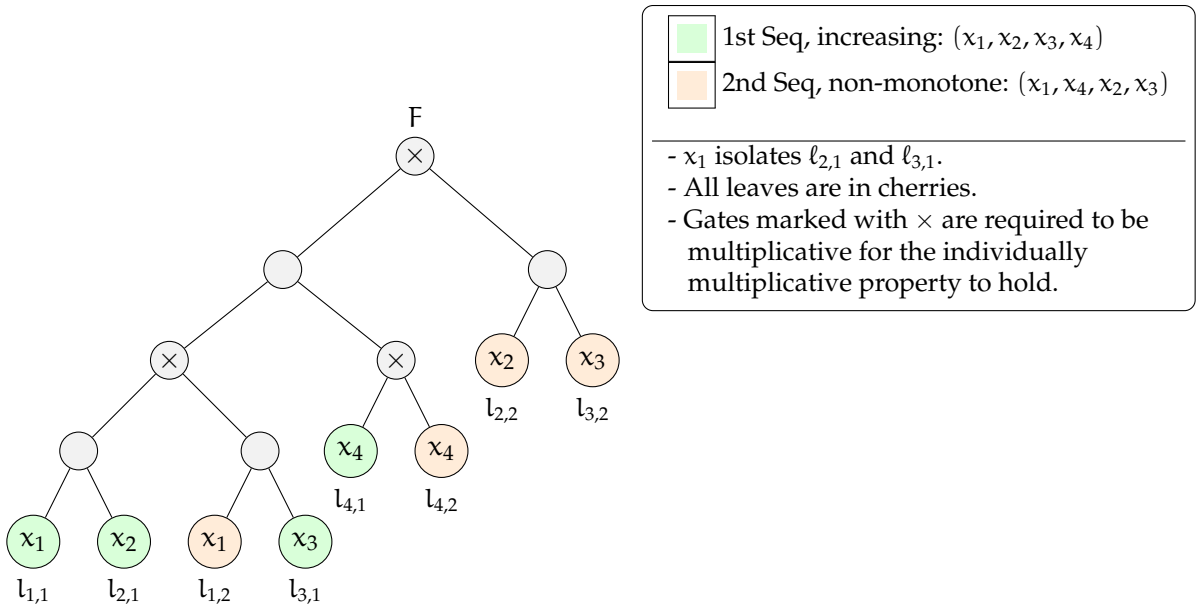

Figure 3: An individually multiplicative R2F

Definition 5.10 (Individually multiplicative)**.** Let F be an algebraic formula. We say that F is *individually multiplicative* if for every $x_i \in \mathbf{x}$, $\mathrm{Read}_F(x_i) = \deg_{x_i}(F)$. ◊

The important property of such formulas is the following.

Observation 5.11. *Let* F *be an individually multiplicative algebraic formula. Then, for every* $x_i \in \mathbf{x}$

$$\partial_{x_i^{\deg}} p_F = \prod_{j=1}^{m} p_{F_j},$$

where every F_j *is a subformula of* F.

The proof is straight forward using induction on the number of reads of any variable and Lemma 2.8 and we omit it.

Definition 5.12 (Read Sequence)**.** Let $F \in$ RkF.

The *first read sequence* of F is the sequence obtained by recording, from left to right in (the underlying tree of) F, the first occurrence of each variable; thus, every variable appears exactly once.

For $1 < i \leqslant k$, the i*-th read sequence* of F is defined recursively as the first read sequence of the tree obtained from F after removing all leaves that appeared in any of the previous read sequences, i.e., in the j-th read sequence for every $j \in [i-1]$.

For $m \in [k]$ and $x_i \in \mathbf{x}$, the m*-th read* of x_i is the leaf labeled by x_i which appears in the m-th read sequence of F. we denote it by $\ell^F_{i,m}$. Whenever F is clear from context we simply write $\ell_{i,m}$. ◊

Definition 5.13. For $F \in$ R2F, if its first read sequence is monotonically increasing (decreasing) and also the second read sequence is monotonically increasing (decreasing), we say that its read sequences are *the same*. If one read sequence is monotonically increasing and the other are monotonically decreasing we say that F's read sequences are *reversed*. ◊

Definition 5.14. Let F be an algebraic formula and ℓ a leaf in its underlying tree. Denote by $\mathrm{Unv}^{L}_{F}(\ell)$ the gates which are left siblings to nodes in the path from ℓ to o_F. Define $\mathrm{Unv}^{R}_{F}(\ell)$ similarly with respect to right siblings. ◊

Definition 5.15. Let F be an algebraic formula and let ℓ_1 and ℓ_2 be two leaves in its underlying tree. We say that ℓ_1 *lies to the left* of ℓ_2 if for some $w \in \mathrm{Unv}^{L}(\ell_2)$, ℓ_1 is a leaf of F_w. We say ℓ_1 *lies to the right* of ℓ_2 if such w exists in $\mathrm{Unv}^{R}(\ell_2)$. ◊

Definition 5.16 (Split)**.** Let F be an individually multiplicative algebraic formula and let $x_t \in \mathbf{x}$. Let

$$\partial_{x_t^{\deg}} p_F = \prod_{\kappa=1}^{m} p_{F_\kappa}$$

where every F_κ is a subformula of F (by Observation 5.11).

For two variables $x_i, x_j \in \mathbf{x}$ we say that t *splits* F *with respect to* i *and* j, if no F_κ depends on both x_i and x_j.

For two leaves $\ell_1, \ell_2 \in F$ we say that t *splits* F *with respect to* ℓ_1 *and* ℓ_2, if ℓ_1 and ℓ_2 are leaves in two distinct formulas of the set $\{F_\kappa\}_{\kappa=1}^{m}$. ◊

The following observation is simple but important, and we therefore state it explicitly. We shall use it implicitly in what follows.

Observation 5.17. *Let F be individually multiplicative, and let $\ell_i, \ell_j, \ell_t \in F$ be leaves such that ℓ_i lies to the left of ℓ_t and ℓ_j lies to the right of ℓ_t. Then, t splits F with respect to ℓ_i and ℓ_j.*

Observation 5.18. *Let F be individually multiplicative and assume it's m-th read sequence is monotonically increasing. Then for any $x_i, x_j, x_t \in \mathbf{x}$ such that $i < t < j$, t splits $\ell_{i,m}$ from $\ell_{j,m}$.*

Lemma 5.19. *Let $F \in \mathsf{R2F}$ be individually multiplicative such that its read sequences are reversed. Then, for every $x_i, x_t, x_j \in \mathbf{x}$ such that $i < t < j$, t splits F with respect to i and j.*

Proof. Without loss of generality, assume the first read sequence is monotonically increasing. By Observation 5.18, x_t splits $\ell_{i,1}$ from $\ell_{j,1}$ and $\ell_{i,2}$ from $\ell_{j,2}$, hence only the second read of x_i and first read of x_j may not be split. Since the second read is reversed, $\ell_{i,2}$ appears to the right of $\ell_{t,2}$ and $\ell_{j,2}$ appears to its left. By definition, this implies that $\ell_{j,1}$ also appears to the left of $\ell_{t,2}$ and hence it is split from $\ell_{i,2}$. □

Definition 5.20 (Isolate)**.** Let F be individually multiplicative and let $x_t \in \mathbf{x}$. Let

$$\partial_{x_t^{\deg}} p_F = \prod_{\kappa=1}^{m} p_{F_\kappa}$$

where every F_κ is a subformula of F.

We say that x_t isolates some leaf $\ell \in F$, if for some $\kappa \in [m]$, $F_\ell = F_\kappa$. ◊

Definition 5.21 (Cherry)**.** Let F be an algebraic formula. We say that two leaves $\ell_1, \ell_2 \in F$ are forming a cherry if ℓ_1 is the sibling of ℓ_2. ◊

The next observation is a direct consequence of Lemma 2.8.

Observation 5.22. *Let F be individually multiplicative and let $\ell \in F$ be some leaf of it. Let v be the sibling of ℓ, then every $x_t \in var(v)$ isolates ℓ.*

Remark 5.23. In what follows, when we write "we have a partial split of F_1 and F_2 with respect to $t = \rho$, $i = \iota$, $j = \nu$, and $\kappa = \mu$, with p_{F_1} as the first polynomial," we mean that the parameters ρ, ι, ν, μ correspond to t, i, j, κ, respectively, in Definition 5.5, and that p_{F_1} is taken as the first polynomial in the split.

The proof of the next lemma relies on the famous Erdős-Szekeres theorem.

Theorem 5.24 (Erdős-Szekeres theorem [ES35])**.** *Every sequence of length* $(r-1)(s-1)+1$ *contains either a monotonically increasing subsequence of length* r *or a monotonically decreasing subsequence of length* s*.*

Lemma 5.25 (R2F splitting)**.** *Let* $F_1, F_2 \in$ R2F *be individually multiplicative, and suppose that for every* $x_i \in \mathbf{x}$ *we have* $\deg_{x_i}(F_1) = \deg_{x_i}(F_2) = 2$*. Then, if* $n \geqslant 5^8$*, there exist* $i, j, t, \kappa \in [n]$ *such that* t *partially splits* F_1 *and* F_2 *with respect to* i, j*, and* κ*.*

Proof. We may assume, without loss of generality, that the first read sequence of F_1 is monotonically increasing. Since $n \geqslant 5^8$, by Theorem 5.24, there exists a set of 5 variables such that after restricting to it, the first and second read sequences of F_1 and F_2 are each monotone (either increasing or decreasing). Moreover, without loss of generality, assume these sequences are over $[5]$. We split into cases.

1. The read sequences are the same for both F_1 and F_2: Assume that the second read of x_1 and the first read of x_5 do not form a cherry in F_1. Then, by Observation 5.22, there exists some $t \in [5] \setminus \{1, 5\}$ such that t isolates the first read of x_5. Assume first that $t = 4$. We argue that we obtain a partial split with respect to t, $i = 1$, $j = 5$, and $\kappa = 3$, where the first tree is F_1.

   We first argue that t splits F_1 with respect to κ and j. By Observation 5.18, the first reads of x_j and x_κ are split, and the same holds for their second reads. Thus, it remains to show that t splits the second read of κ from the first read of x_j. This follows since t isolates the first read of x_j. Hence, Property (3) holds.

   Now, if t yields a complete split, we are done by Observation 5.9. Otherwise, by Observation 5.18, t splits $\ell_{i,1}^{F_2}$ from $\ell_{j,1}^{F_2}$ and $\ell_{i,2}^{F_1}$ from $\ell_{j,2}^{F_2}$. The fact that the split is not complete implies that there exists a gate $w \in F_2$ defining one of the sets in the split such that both $\ell_{j,1}$ and $\ell_{i,2}$ are leaves of F_w, and $x_t \notin \mathrm{var}(F_w)$. In particular, the first read of x_t is to the left of all the leaves in F_w, and its second read is to their right. Since the ordering of both reads is monotone, it follows that no variable is read twice between the two leaves $\ell_{t,1}$ and $\ell_{t,2}$. Consequently, F_w is a ROF. Thus, Property (2) holds. Property (1) is easy to verify, and we conclude that t, i, j, κ form a partial split of F_1 and F_2.

   Otherwise, the second read of x_1 and the first read of x_5 form a cherry in F_1. Observe that the argument above requires only four variables. Hence, if the first sibling of the cherry lies to its left, the previous argument yields a partial split when both formulas are restricted to $\{x_2, x_3, x_4, x_5\}$. Otherwise, we obtain a partial split for the labels $\{x_1, x_2, x_3, x_4\}$.

2. The read sequences are the same for F_1 and reversed for F_2: Assume, without loss of generality, that the sequences are monotonically increasing for F_1. We claim that $t = 2$, $i = 1$, $j = 4$, and $\kappa = 3$ partially split F_1 and F_2, with F_2 serving as the first tree. Indeed, by Lemma 5.19, t splits F_2 with respect to i and j, as well as with respect to i and κ. Hence, Property (3) holds.

   As before, in F_1, only $\ell_{i,2}$ and $\ell_{j,1}$ may remain unsplit. If they are split, we obtain a complete split. Otherwise, by the same reasoning as in the previous case, the subformula F_w containing both of these reads is a ROF, and thus Property (2) holds.

3. The read sequences are reversed both for F_1 and for F_2: By Lemma 5.19 2 completely splits both F_1 and F_2 with respect to 1 and 3. □

Remark 5.26. This lemma does not aim to optimize the label bound. A more careful and technical argument shows $n = 32$ suffices; we chose to present a simpler proof with weaker parameters.

5.3 Hardness of Representation for $\sum^2$ R3F which is also an R4F

We are now ready to prove Proposition 5.1. For ease of reading, we repeat it below.

Proposition 5.1. *Let* F, G *be* R3Fs *such that the formula* $F + G$ *is a structural* R4F. *Let* U *be a set containing all the read-4 variables in the formula* $F + G$.

Denote

$$\mathcal{T} \coloneqq \left\{ \partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p\right) \;\middle|\; p \in \{p_F, p_G\},\ x_i, x_j \in \mathit{var}(p) \right\}.^{17}$$

Assume $U \neq \emptyset$ *and that no nonzero polynomial in* $\mathcal{Q}_{F,G} \cup \mathcal{T}$ *vanishes at* $\mathbf{0}$.
Then, if $n \geqslant r_{2,6} + 5^8$, *there exists* $t \in U$ *such that* $\mathcal{P}_{[n]\setminus\{t\}} \nmid \partial_{x_t^{\deg}}(F + G)$.

Proof. Denote by S_2 the set of variables that have degree 2 in both F and G, and let $S_1 = [n] \setminus S_2$. We consider two cases:

1. $|S_2| \geqslant 5^8$**:** Restrict F and G by a generic assignment to the variables in S_2.

   Let $x_i, x_j \in S_2$ be such that $\partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p_F\right) \not\equiv 0$. By our assumption on the set $\mathcal{T}$, polynomials of this form do not vanish at $\mathbf{0}$, hence

   $$\partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p_F(\mathbf{0})\right) \neq 0\,.$$

   By the properties of generic assignments, it follows that

   $$\partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p_{F|_{S_2}}(\mathbf{0})\right) \neq 0\,.$$

   Therefore, for every $x_j \in S_2$, $\partial_{x_j^{\deg}} p_{F|_{S_2}}$ is $\mathbf{0}$-preserved. The same argument applies to G.

   Observe that $F|_{S_2}$ and $G|_{S_2}$ are individually multiplicative R2Fs, of degree 2 with respect to every variable. Hence, by Lemma 5.25, there exists t that partially splits $p_{F|_{S_2}}$ and $p_{G|_{S_2}}$. Therefore, $p_{F|_{S_2}}$ and $p_{G|_{S_2}}$ satisfy the conditions of Corollary 5.6, which implies that

   $$\mathcal{P}_{S_2\setminus\{t\}} \nmid \partial_{x_t^{\deg}}\left(p_{F|_{S_2}} + p_{G|_{S_2}}\right).$$

   Since partial differentiation with respect to x_t commutes with substitution of the other variables, we get

   $$\mathcal{P}_{S_2\setminus\{t\}} \nmid \left(\partial_{x_t^{\deg}}(p_F + p_G)\right)\big|_{S_2},$$

   which further implies that

   $$\mathcal{P}_{[n]_{-t}} \nmid \partial_{x_t^{\deg}}(p_F + p_G)\,.$$

   Finally, note that $S_2 \subseteq U$.

[17] Observe that $\partial_{x_i^{\deg}}\left(\partial_{x_j^{\deg}} p\right) \neq \partial_{x_i^{\deg} x_j^{\deg}} p$. On the left-hand side, we differentiate with respect to x_i as many times as its degree in $\partial_{x_j^{\deg}} p$, whereas on the right-hand side, we differentiate as many times as its degree in p.

2. $|S_2| < 5^8$**:** Let $t \in U$ be arbitrary, and define

$$f = \partial_{x_t^{\deg}}(p_F)\big|_{S_1} \quad \text{and} \quad g = \partial_{x_t^{\deg}}(p_G)\big|_{S_1}.$$

If there exists $x_j \in \mathrm{var}(f+g)$ such that $\deg_{x_j}(f+g) > 1$, then, since $j \notin S_2$, by definition its degrees in f and g differ. Without loss of generality, assume $\deg_{x_j}(f) < \deg_{x_j}(g)$. Then

$$\partial_{x_j^{\deg}}(f+g)(\mathbf{0}) = \partial_{x_j^{\deg}} g(\mathbf{0}) = \partial_{x_j^{\deg}}\big(\partial_{x_t^{\deg}}(p_G)\big)(\mathbf{0}) \neq 0,$$

by our assumption on the set $\mathcal{T}$.

Hence, by Lemma 2.39, we conclude that $(f+g)$ is 2-hard, as desired.

Otherwise, $f+g$ is a multilinear polynomial. Since $S_1 \sqcup S_2 = [n]$, we have $|S_1| > r_{2,6}$, and in particular $|S_1 \setminus \{t\}| \geqslant r_{2,6}$. By the definition of S_1 and by Lemma 2.11, both f and g are multilinear R3Ps. By Corollary 2.6, these are structurally multilinear R3Ps. As the nonzero polynomials in $Q_{F,G}$ do not vanish at $\mathbf{0}$, it follows that $\mathbf{0}$ satisfies the conditions of Lemma 2.43. Applying that lemma with $m = 2$ and $k = 6$, we obtain

$$\mathcal{P}_{S_1} \nmid f+g,$$

which implies that

$$\mathcal{P}_{[n]_{-t}} \nmid \partial_{x_t^{\deg}}(p_F + p_G). \qquad \square$$

6 Dominating degree patterns

In this section we introduce the notion of *dominating degree patterns* and discuss its importance. We begin by recalling notation and definitions introduced in [HMM24].

6.1 Notation and Definitions

Definition 6.1. (Degree pattern, taken from [HMM24]) Let $J \subseteq [n]$. A degree pattern with domain J is a J-indexed tuple $\mathfrak{e} \in \mathbb{N}^J$ of nonnegative integers.

For some degree pattern $\mathfrak{e}$ we denote by $\mathrm{Domain}(\mathfrak{e})$ the set of indices in the domain of $\mathfrak{e}$. Moreover, we denote by $\boldsymbol{x}^{\mathfrak{e}}$ the monomial $\prod_{j \in \mathrm{Domain}(\mathfrak{e})} x_j^{\mathfrak{e}_j}$.

A degree pattern $\mathfrak{d}$ with domain $J \subseteq [n]$ matches a monomial $m = \boldsymbol{x}^{\boldsymbol{e}}$ such that $\boldsymbol{e} \in \mathbb{F}^n$, if for every $j \in J$, $\deg_{x_j}(m) = \mathfrak{d}_j$.

We say that $\mathfrak{d}$ is in $p \in \mathbb{F}[\boldsymbol{x}]$ if $\mathfrak{d}$ matches some monomial in p.

Every polynomial $p \in \mathbb{F}[\boldsymbol{x}]$ can be written uniquely in the form

$$p = \sum_{\mathfrak{d} \in \mathbb{N}^J} p_{\mathfrak{d}} \boldsymbol{x}^{\mathfrak{d}}$$

where p_d is a polynomial over $\mathbb{F}$ that depends on the variables outside of J ($[n] \setminus J$). We refer to p_d as the coefficient of $\boldsymbol{x}^{\mathbf{d}}$ in p. $\Diamond$

The next definition is close to the definition of dominating monomials.

Definition 6.2. (Dominating degree pattern) Fix $J \subseteq [n]$ and let $\mathfrak{e}$ and $\mathfrak{d}$ be two degree patterns with domain J. We say that $\mathfrak{e}$ dominates $\mathfrak{d}$ if $\mathfrak{e}_i \geqslant \mathfrak{d}_i$ for every $i \in J$ and there exists $i_0 \in J$ such that $\mathfrak{e}_{i_0} > \mathfrak{d}_{i_0}$.

Moreover, we say $\mathfrak{e}$ is a dominating degree pattern in $F \in \mathbb{F}[x_1, \ldots, x_n]$ if its coefficient in F is nonzero and there is no degree pattern in F that dominates it. When J and F are obvious from context, we simply say that $\mathfrak{e}$ is dominating. ◊

Next, we define a new notation for taking partial derivatives.

Definition 6.3. Let F be a polynomial and $J \subset \mathrm{var}(F)$. We denote $\partial_{J^{\mathrm{dom}}} F$ as the set of nonzero partial derivatives, with respect to the dominating degree patterns with domain J:

$$\partial_{J^{\mathrm{dom}}} F \coloneqq \{\partial_{\mathfrak{e}} F \mid \mathrm{Domain}(\mathfrak{e}) = J \text{ and } \mathfrak{e} \text{ is dominating in } F\} . \qquad ◊$$

6.2 Two Elementary Sets

In this subsection we define two elementary sets of polynomials which will be used throughout our proofs.

Definition 6.4 ($\mathcal{A}_F$)**.** Let F be an algebraic formula and $t \in \mathbb{N}$. Define,

$$\mathcal{A}_F^t \coloneqq \bigcup_{I \in \binom{[n]}{\leqslant t}} \partial_{I^{\mathrm{dom}}} p_F ,$$

In words, this is the set obtained by taking all possible derivatives according to all dominating degree patterns of size at most t. ◊

Definition 6.5 ($\mathcal{E}_F^m$)**.** Let F be an algebraic formula and let $m \in \mathbb{N}$. Define:

$$\mathcal{E}_F^m \coloneqq \bigcup_{u \in F_{(o)_L} \cup F_{(o)_R}} \mathcal{A}_{F_u}^m. \qquad ◊$$

The utility of this definition is captured in the next claim.

Claim 6.6. *Let F be an algebraic formula. If no polynomial in $\mathcal{E}_F^1$ vanishes at $\mathbf{0}$ then both p_{F_L} and p_{F_R} are $\mathbf{0}$-preserved.*

Proof. This is an immediate consequence of Lemma 2.33. □

Remark 6.7. Note that if $t \leqslant t'$ then $\mathcal{A}_F^t \subseteq \mathcal{A}_F^{t'}$, and similarly if $m \leqslant m'$ then $\mathcal{E}_F^m \subseteq \mathcal{E}_F^{m'}$

6.3 Extension of Lemma 2.11 for R2Fs

In Proposition 7.31, we established a hardness of representation result for totally non-structural $\sum^2 \mathsf{R2F}$, under the assumption that for every $p \in \mathcal{E}_F^3$, for some $F \in \sum^2 \mathsf{R2F}$, we have $p(\mathbf{0}) \neq 0$. When proving Theorem 1.1, to ensure this property, we translate p_F by a common nonzero of the polynomials in $\mathcal{E}_F^3$. To perform this translation, we must be able to find such a common nonzero. The goal of this section is therefore to construct a hitting set for the polynomials in $\mathcal{E}^3$. For this, we prove Lemma 6.8, which strengthens Lemma 2.11 in the case of R2Fs.

Lemma 6.8. *Let $S \subseteq [n]$ be of size $|S| \leqslant 3$. Let $F \in \mathsf{RkF}$ be such that for every $i \in S$, $\mathrm{Read}_F(x_i) \leqslant 2$. Let $\mathfrak{e} \in \{1,2\}^S$ be a dominating degree pattern of F with domain S. Then, $\partial_{\mathfrak{e}} F \in \mathsf{RkF}$, and moreover it is composed entirely of disjoint subformulas of F.*

We prove this claim in Subsection 6.4.

The following is a simple corollary that follows from Lemma 6.8. We will not use it in our proofs, but we still find it interesting.

Lemma 6.9. *Let $\mathfrak{e}$ be a degree pattern with domain $J \subseteq [n]$ such that $|J| \leqslant 3$. Let $F \in \mathsf{R2F}$ and let $\mathcal{H}$ be a generator for n-variate $\mathsf{R2F}$s. Then*

$$\partial_{\mathfrak{e}} p_F \equiv 0 \iff \partial_{\mathfrak{e}} p_F \circ (\mathcal{H} + \mathcal{G}_{|J|}) \equiv 0.$$

Proof. In the case where $\partial_{\mathfrak{e}} p_F \equiv 0$ the claim is immediate. Assume therefore that $\partial_{\mathfrak{e}} p_F \not\equiv 0$.

If $\mathfrak{e}$ is a dominating degree pattern, then the claim follows directly from Lemma 6.8. Otherwise, let $\mathfrak{d}$ be a degree pattern that dominates $\mathfrak{e}$, and denote

$$I = \left\{ \mathfrak{d}' \;\middle|\; \mathfrak{d}' \in \{1,2\}^J,\ \ \mathfrak{d}'_i \geqslant \mathfrak{e}_i \ \text{ for all } i \in J, \text{ and } \ \mathfrak{d}' \neq \mathfrak{d} \right\}.$$

Then, using the notation of Definition 6.2, we have

$$\partial_{\mathfrak{e}} p_F = \partial_{\mathfrak{e}} \left(\sum_{\mathfrak{r} \in \{1,2\}^J} p_{\mathfrak{r}} \mathbf{x}^{\mathfrak{r}} \right) = c_{\mathfrak{d}}\, p_{\mathfrak{d}} \prod_{i \in J} x_i^{\mathfrak{d}_i - \mathfrak{e}_i} + \sum_{\mathfrak{r} \in I} c_{\mathfrak{r}}\, p_{\mathfrak{r}} \prod_{i \in J} x_i^{\mathfrak{r}_i - \mathfrak{e}_i},$$

where $c_{\mathfrak{d}} \in \mathbb{F} \setminus \{0\}$ and $p_{\mathfrak{d}} \in \mathbb{F}[\mathbf{x}_{-J}]$ and for each $\mathfrak{r} \in I$, $c_{\mathfrak{r}} \in \mathbb{F} \setminus \{0\}$ and $p_{\mathfrak{r}} \in \mathbb{F}[\mathbf{x}_{-J}]$.

Note that the monomial $\prod_{i \in J} x_i^{\mathfrak{d}_i - \mathfrak{e}_i}$ does not appear in the second summation on the right hand side. By Lemma 6.8 we know that

$$p_{\mathfrak{d}} \sim \partial_{\mathfrak{d}} p_F$$

is an R2P. Moreover, by assumption, it is nonzero. Hence $p_{\mathfrak{d}} \circ \mathcal{H} \not\equiv 0$. Therefore, by reviving the variables in J (see Observation 2.19), we obtain the desired result. □

6.4 Proof of Lemma 6.8

To ease the reading we restate it.

Lemma 6.8. *Let $S \subseteq [n]$ be of size $|S| \leqslant 3$. Let $F \in \mathsf{RkF}$ be such that for every $i \in S$, $\mathit{Read}_F(x_i) \leqslant 2$. Let $\mathfrak{e} \in \{1,2\}^S$ be a dominating degree pattern of F with domain S. Then, $\partial_{\mathfrak{e}} F \in \mathsf{RkF}$, and moreover it is composed entirely of disjoint subformulas of F.*

Proof. For $|S| = 1$, the claim follows directly from Lemma 2.11, noting that if $S = \{i\}$, then it must be the case that $\mathfrak{e}_i = \deg_{x_i}(p_F)$.

We now prove the claim for $|S| = 2, 3$. To ease notation, we assume that when $|S| = 2$, the variables indexed by S are y and z, and when $|S| = 3$, they are y, z, and w. We denote by $\mathbf{x}$ the remaining variables. We shall occasionally abuse notation and treat S itself as the corresponding set of variables.

We first note that if, for some variable in S, its degree in $\mathfrak{e}$ equals its degree in p_F, then, as in the case $|S| = 1$, Lemma 2.11 allows us to take the derivative with respect to its maximal degree and thereby reduce the problem to $|S| - 1$ variables. We may therefore assume that each variable in S appears nontrivially in F, and that for all such variables, their degree in $\mathfrak{e}$ is 1, while their degree in p_F is 2.

Let v be the *first common gate (fcg)* of all leaves labeled by variables in S. Denote by v_y, v_z, and v_w the fcg of the y-, z-, and w-leaves, respectively. (When $|S| = 2$, we ignore v_w.)

Claim 6.10. *Either Lemma 6.8 holds, or for some variable r which has its index in S, $v = v_r$.*

Proof. Assume, for contradiction, that this is not the case. Without loss of generality, assume that both v_y and v_w are descendants of $(v)_L$, and that v_z is a descendant of $(v)_R$.

If v is an addition gate, then the derivative with respect to $\mathfrak{e}$ is clearly the zero polynomial, and Lemma 6.8 trivially holds. Thus, assume that v is a multiplication gate. In this case, $(v)_L$ must contain the monomial yw, while $(v)_R$ contains z^2, contradicting the assumption that $\mathfrak{e}$ is dominating. □

We may therefore assume, without loss of generality, that $v = v_y$. Since $\deg_y(p_F) = 2$, the gate v must be a multiplication gate. We now divide the analysis into several cases, according to how many times each variable in S appears in each child of v. The following four cases exhaust all possible configurations, up to renaming of the variables and exchanging the children of v.

1. Each of v_z and v_w is either equal to $(v)_R$ or a descendant of it.
   This situation can occur only when $|S| = 3$, since otherwise the monomial yz^2 would appear in p_F. Observe that the degree pattern restricted to z and w must be dominating in $p_{(v)_R}$, as any monomial in z, w is multiplied by y and therefore cannot cancel out. Hence, the claim follows from the $|S| = 2$ case by first taking the derivative $\partial_{zw} p_F$ and noting that, in the resulting polynomial, the degree of y becomes 1, allowing us to apply Lemma 2.11.

2. v_z is a descendant of $(v)_R$, and v_w is a descendant of $(v)_L$.
   Again, this configuration can occur only when $|S| = 3$. In this case, the derivative with respect to the S-variables is zero. Indeed, if $p_{(v)_R}$ contained a monomial divisible by yz, then p_F would contain a monomial divisible by w^2yz, contradicting the assumption that the $(1,1,1)$ degree pattern is dominating. Hence, for some polynomials $a, a', b, b', c, c', d, d' \in \mathbb{F}[\mathbf{x}]$, we have

   $$p_{(v)_R} = az^2 + bz + cy + d, \qquad p_{(v)_L} = a'w^2 + b'w + c'y + d',$$

   and no monomial in p_v, and therefore in p_F, is divisible by wyz.

3. v_z is a descendant of $(v)_R$, and $v_y = v_w = v$.
   Assume $|S| = 3$ (the case $|S| = 2$ will appear in the next case and essentially follows from the analysis here). As before, $p_{(v)_L}$ cannot contain a monomial divisible by wy, since this would yield a monomial divisible by wyz^2. Thus, $u_{L,yw} \coloneqq \mathrm{fcg}_{(v)_L}(y, w)$ must be an addition gate. Similarly, no monomial in $p_{(v)_R}$ is divisible by yz^2 or wz^2. On the other hand, there must exist a monomial divisible by either yz or wz; otherwise, no monomial would contain all three variables.

   Define $u_{R,yz} = \mathrm{fcg}_{(v)_R}(y, z)$ as the first gate whose subformula contains both a y-leaf and a z-leaf. We claim that $u_{R,yz}$ is an addition gate. This follows from the fact that z^2 appears in some monomial of $p_{(v)_R}$ but yz^2 does not. The same holds for $u_{R,wz} = \mathrm{fcg}_{(v)_R}(w, z)$, defined analogously. Similarly, define $u_{R,yw}$.

   Up to symmetry, three subcases remain.

(a) $v_z = u_{R,yw}$. Then for some polynomials $a, \ldots, e' \in \mathbb{F}[\mathbf{x}]$,

$$\begin{aligned} p_{u_{R,yz}} &= ay + bz + c, \\ p_{u_{R,wz}} &= a'w + b'z + c', \\ p_{v_z} &= (d \cdot p_{u_{R,yz}} + e)(d' \cdot p_{u_{R,wz}} + e'), \\ p_{u_{L,yw}} &= a''y + b''w + c'', \end{aligned}$$

where all suformulas computing $a, a', a'', b, b', b'', d, d'$ are disjoint. Clearly, these subformulas are also disjoint from the (disjoint) subformulas

$$\left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right), \quad \left(\prod_{r \in \mathrm{Unv}_{F_{(u)_R}}(v_z)} p_r\right) \quad \text{and} \quad \left(\prod_{r \in \mathrm{Unv}_{F_{(u)_L}}(u_{L,zw})} p_r\right).$$

Thus,

$$\begin{aligned} \partial_{yzw} p_F &= \left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right) \left(\partial_{yz} p_{(v)_R} \partial_w p_{(v)_L} + \partial_{wz} p_{(v)_R} \partial_y p_{(v)_L}\right) \\ &= \left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right) \left(\left(\prod_{r \in \mathrm{Unv}_{F_{(v)_R}}(v_z)} p_r\right) dd'ab' \left(\prod_{r \in \mathrm{Unv}_{F_{(v)_L}}(u_{L,yw})} p_r\right) b'' \right. \\ &\quad \left. + \left(\prod_{r \in \mathrm{Unv}_{F_{(v)_R}}(v_z)} p_r\right) dd'ba' \left(\prod_{r \in \mathrm{Unv}_{F_{(v)_L}}(u_{L,yw})} p_r\right) a'' \right) \\ &= \left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right) \left(\prod_{r \in \mathrm{Unv}_{F_{(v)_R}}(v_z)} p_r\right) dd' \left(\prod_{r \in \mathrm{Unv}_{F_{(v)_L}}(u_{L,yw})} p_r\right) \left(ab'b'' + a'ba''\right) \end{aligned}$$

which, by disjointness, is an RkP.

(b) $u_{R,yw}$ is a descendant of $u_{R,yz} = u_{R,wz}$. Here,

$$\begin{aligned} p_{u_{R,yw}} &= ay + bw + c, \\ p_{u_{R,wz}} &= a' p_{u_{R,yw}} + b'z + c', \\ p_{v_z} &= (d p_{u_{R,wz}} + e)(d'z + e'), \\ p_{u_{L,yw}} &= a''y + b''w + c'', \end{aligned}$$

where the subformulas computing a'', b'', a, b are disjoint. Then

$$\begin{aligned}
\partial_{yzw} p_F &= \left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right) \left(\partial_y p_{(v)_L} \partial_{wz} p_{(v)_R} + \partial_w p_{(v)_L} \partial_{yz} p_{(v)_R}\right) \\
&= \left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right) \left(\left(\prod_{r \in \mathrm{Unv}_{(v)_L}(u_{L,yw})} p_r\right) a'' \left(\prod_{r \in \mathrm{Unv}_{(v)_R}(v_z)} p_r\right) d'da'b \right. \\
&\quad \left. + \left(\prod_{r \in \mathrm{Unv}_{(v)_L}(u_{L,yw})} p_r\right) b'' \left(\prod_{r \in \mathrm{Unv}_{(v)_R}(v_z)} p_r\right) d'da'a \right) \\
&= \left(\prod_{r \in \mathrm{Unv}_F(v)} p_r\right) \left(\prod_{r \in \mathrm{Unv}_{(v)_L}(u_{L,yw})} p_r\right) \left(\prod_{r \in \mathrm{Unv}_{(v)_R}(v_z)} p_r\right) dd'a'(a''b + b''a),
\end{aligned}$$

which is again an RkP.

(c) $u_{R,yz}$ is a descendant of $u_{R,yw}$. The argument is symmetric to the previous subcase. Since $v_z \neq u_{R,yw}$, we have that $u_{R,yw}$ is an addition gate, and the same calculation applies.

4. $v_y = v_z = v_w = v$.
When $|S| = 2$, both $p_{(v)_L}$ and $p_{(v)_R}$ must be linear in y, z over $\mathbb{F}(\mathbf{x})$, and the reasoning parallels the previous case. Assume $|S| = 3$. Without loss of generality, suppose that $p_{(v)_R}$ has total degree 2 in y, z, w over $\mathbb{F}(\mathbf{x})$ and that the monomial yz appears in it. If $p_{(v)_L}$ also has degree 2 in these variables, then it can contain only the monomial yz and be linear in w; the same holds for $p_{(v)_R}$. Otherwise, if $p_{(v)_L}$ is linear, then $p_{(v)_R}$ can contain at most one additional quadratic monomial (say wz). Indeed, $p_{(v)_R}$ is a ROF and if all fcg's where multiplication gates then it would also have contained the monomial yzw. Thus, one of the following must hold:

(a) $u_{R,wy} = u_{R,wz}$, $u_{L,wy} = u_{L,wz}$, and

$$\begin{aligned}
p_{u_{R,wy}} &= e(ay + b)(cz + d) + fw + g, \\
p_{u_{L,wy}} &= e'(a'y + b')(c'z + d') + f'w + g',
\end{aligned}$$

where a, c, f, a', c', f' are computed by disjoint subformulas. A direct calculation shows that the derivative is an RkF.

(b) $u_{R,wz} = u_{R,yz}$ and

$$p_{u_{R,wz}} = (d(ay + bw + c) + e)(fz + g) + h.$$

Then p_{v_L} must be linear in y, z, w. Regardless of the specific arrangement of gates, a direct computation shows that all yzw monomials share a common part, and the remainder consists of disjoint subformulas, so the derivative is again an RkF. An illustrative example (in which we dropped the additive terms in each gate) appears in Figure 4. One can see that to get ywz term we either take the $a'c'dbf$ path or the $d'daf$ path. □

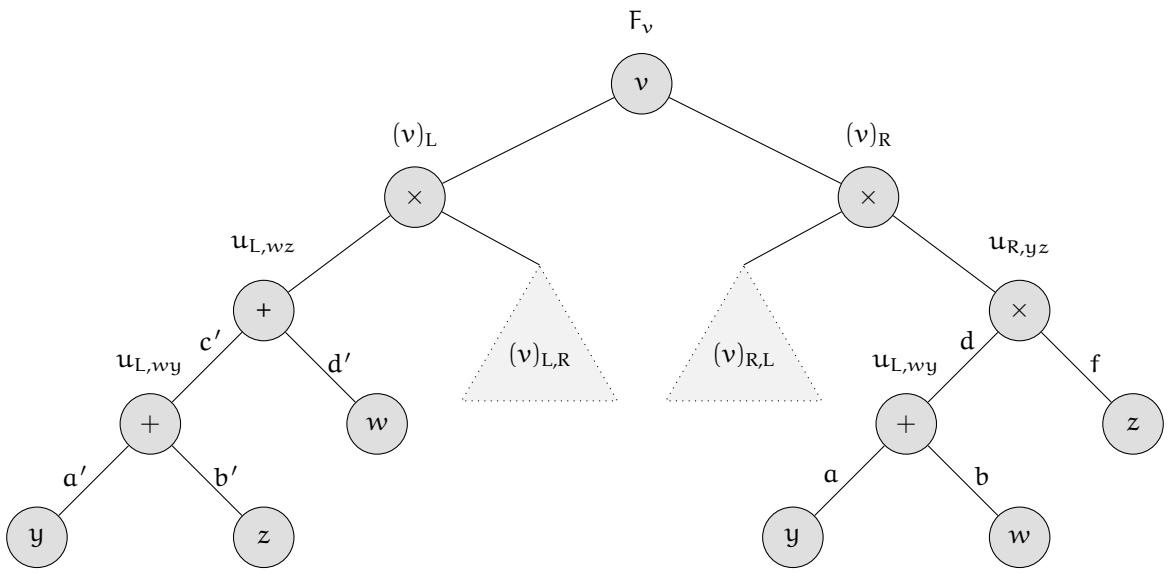

Figure 4: An illustrative structure for Case 4b, showing disjoint subformulas contributing to the derivative.

7 Totally Non-Structural $\sum^2$ R2F

In this section we prove a hardness of representation result for the totally non-structural case. By this we mean the case where all the variables are non-structural (see Definition 7.1 below). We also establish a hardness of representation result for the class $\sum^2 \prod^k$ ROF, for which a PIT algorithm was given in [BGV23]. This extension is crucial, as it enables the use of the method of generic assignments. The arguments in this part rely heavily on the structural constraints imposed by the totally non-structural setting.

To indicate our current position in the proof, we include Figure 5. In this section, we prove Proposition 7.31, whose proof corresponds to the subtree rooted at the orange box.

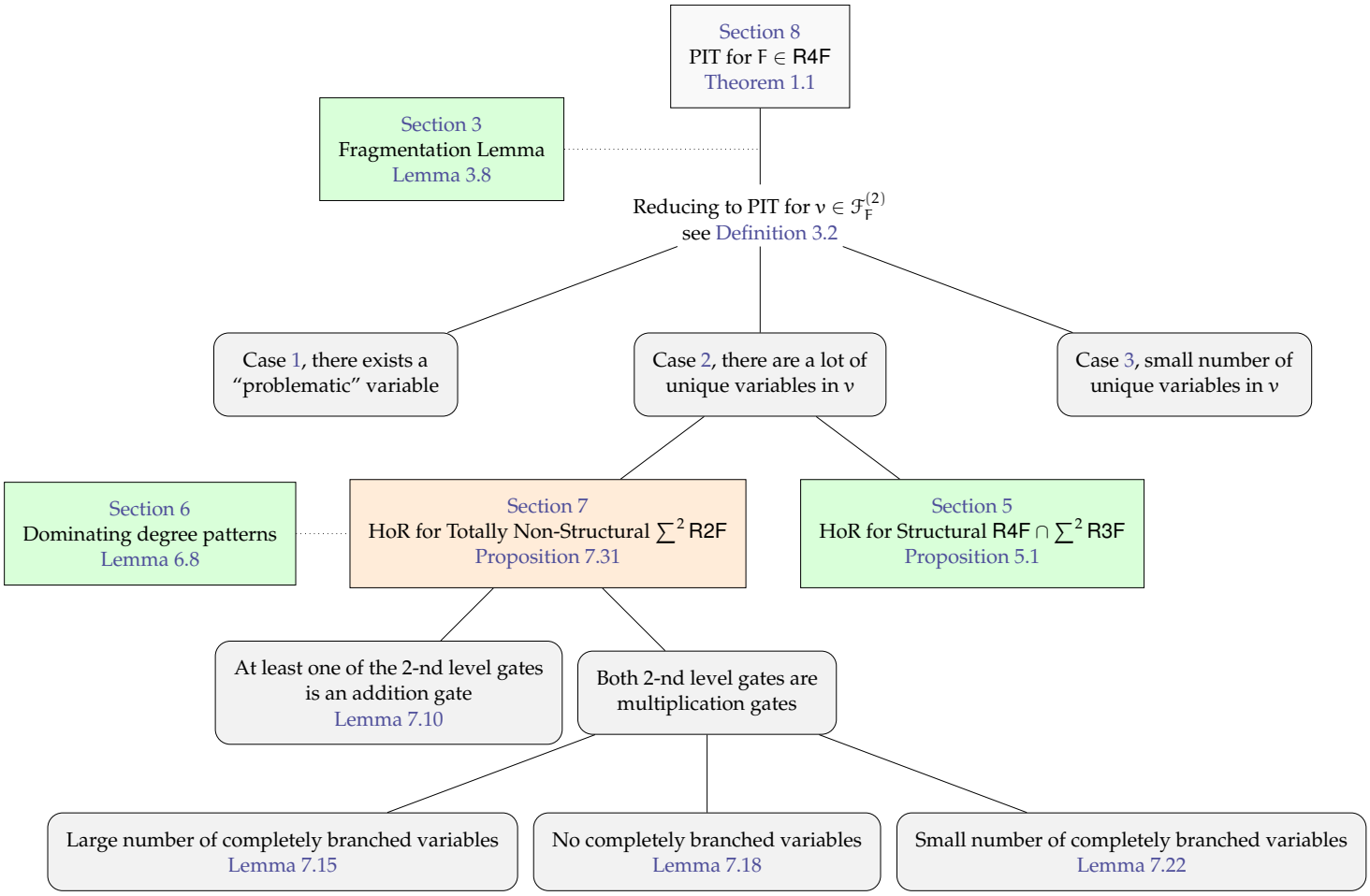

Figure 5: Our current position in the proof of Theorem 1.1

Definition 7.1 (Totally non-structural $\Pi\Sigma$kF)**.** We say that $F \in \Pi\Sigma$kF over the set of variables $[n]$ is totally non-structural, if for every $i \in [n]$ there are two gates ν and u in F, such that ν is a child of u and $\deg_{x_i}(p_\nu) > \deg_{x_i}(p_u)$. ◊

In the following subsections, we focus exclusively on totally non-structural $\sum^2 \Pi\Sigma$F formulas. Later, in Section 8, we combine the hardness of representation result established here with the corresponding result for structural $\Pi\Sigma\Pi$F formulas, which take the form of $\sum^2 \Pi\Sigma$F, to obtain a hardness of representation result for general $\Pi\Sigma\Pi$Fs. This is achieved using generic assignments, as outlined in Lemma 2.53.

7.1 Preliminaries

7.1.1 Structural definitions

We give some definitions and simple results concerning the structure of a totally non-structural $\Pi\Sigma\Pi$F and of totally non-structural $\sum^2 \Pi\Sigma$F.

Observation 7.2. *Let $F \in \Pi\Sigma\Pi$F be totally non-structural, and let $x_i \in \mathbf{x}$. Let u be the fcg of the four occurrences of x_i in F. Then, u must be an addition gate, and without loss of generality:*

$$\deg_{x_i}(p_u) = 1 \quad \textit{and} \quad \deg_{x_i}(p_{(u)_L}) = \deg_{x_i}(p_{(u)_R}) = 2.$$

Proof. First, observe that u cannot be a multiplication gate. Indeed, by Corollary 2.6, both its children $(u)_L$ and $(u)_R$ are structural. Therefore, if u were a multiplication gate, then F would be structural with respect to x_i, contradicting the assumption that F is totally non-structural. As the degree of x_i in p_F is at least 1, and since the individual degrees with respect to x_i of the children of u must be equal, the only possibility is that $\deg_{x_i}(p_{(u)_L}) = \deg_{x_i}(p_{(u)_R}) = 2$ and $\deg_{x_i}(p_u) = 1$, as claimed. □

Corollary 7.3. *Let $F \in \mathsf{R4F}$ be totally non-structural. Then p_F is multilinear.*

Corollary 7.4. *Let $F \in \mathsf{R4F}$ be totally non-structural, and let $i \in [n]$. Let u be the fcg of the four occurrences of x_i in F. Then, in each of the subtrees rooted at $(u)_L$ and $(u)_R$ the fcg of the x_i-leaves is a multiplication gate.*

We next consider properties of totally non-structural $\sum^2 \mathsf{R2F}$. The following is a restatement of Corollary 7.4.

Corollary 7.5. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula, and let $x_i \in \boldsymbol{x}$. If ℓ_1 and ℓ_2 are the two leaves labeled by x_i in some child of o, then $fcg^{op}(\ell_1, \ell_2) = \times$.*

In the rest of this subsection we assume both $(o)_L$ and $(o)_R$ are multiplication gates.

Definition 7.6 (Branching)**.** Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula. Suppose both $(o)_L$ and $(o)_R$ are multiplication gates. We say that a variable $x_i \in \boldsymbol{x}$ *branches* in $(o)_L$ (respectively, in $(o)_R$) if x_i appears at most once in $(o)_{L,L}$ and at most once in $(o)_{L,R}$. If x_i branches both in $(o)_L$ and in $(o)_R$, we say that x_i *branches completely* in F. If x_i branches only in one of the children of o then we say that it *partially branches* in F. If either happens, we simply say that x_i branches in F. ◊

Observe that if x_i branches in $(o)_L$ then

$$p_{(o)_L} = (ax_i + b)(cx_i + d),$$

where $a, b, c, d \in \mathbb{F}[\boldsymbol{x}_{-i}]$.

Definition 7.7 (Separated variables)**.** Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula.

We say that two variables $x_i, x_j \in [n]$ are *separated* in $(o)_L$ (respectively, in $(o)_R$) if, within $(o)_L$ (respectively, within $(o)_R$), each occurs in exactly one child and they occur in different children. ◊

Lemma 7.8. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula, and suppose that both $(o)_L$ and $(o)_R$ are multiplication gates. If, for some $\boldsymbol{\alpha} \in \mathbb{F}^n$, we have $p_F(\boldsymbol{x} + \boldsymbol{\alpha}) \sim \mathcal{P}_n$, then there do not exist indices $x_i, x_j \in \boldsymbol{x}$ such that x_i and x_j are separated in both $(o)_L$ and $(o)_R$.*

Proof. Since the statement concerns the structure of F, which is invariant under translations by constant vectors, we may assume without loss of generality that $p_F \sim \mathcal{P}_n$ to simplify the presentation.

Suppose, for the sake of contradiction, that there exist variables $x_i, x_j \in \boldsymbol{x}$ such that x_i and x_j are separated in both $(o)_L$ and $(o)_R$. Since $(o)_L$ and $(o)_R$ are multiplication gates, and each variable appears in only one child, the totally non-structural assumption implies that the subformulas rooted at $(o)_L$ and $(o)_R$ compute polynomials of the form

$$p_{(o)_L} = a_L \cdot (b_L x_i^2 + c_L x_i + d_L)(e_L x_j^2 + f_L x_j + g_L) + h_L,$$
$$p_{(o)_R} = a_R \cdot (b_R x_i^2 + c_R x_i + d_R)(e_R x_j^2 + f_R x_j + g_R) + h_R,$$

where $a_R, a_L, \ldots, h_R, h_L \in \mathbb{F}[\mathbf{x}_{-\{i,j\}}]$. Since multiplying a_R by $b_R e_R$ and rescaling $(b_R x_i^2 + c_R x_i + d_R)$ by b_R^{-1} and $(e_R x_j^2 + f_R x_j + g_R)$ by e_R^{-1} yields an equivalent polynomial, we may assume without loss of generality that $b_R = b_L = e_R = e_L = 1$. Note that this normalization may cause some coefficients to become rational functions rather than polynomials. Abusing notation, we write

$$p_{(o)_L} = a_L \cdot (x_i^2 + c_L x_i + d_L)(x_j^2 + f_L x_j + g_L) + h_L,$$
$$p_{(o)_R} = a_R \cdot (x_i^2 + c_R x_i + d_R)(x_j^2 + f_R x_j + g_R) + h_R,$$

where $a_R, a_L, \ldots, h_R, h_L \in \mathbb{F}(\mathbf{x}_{-\{i,j\}})$. Let $\zeta \in \mathbb{F}$ be such that[18]

$$p_F = p_{(o)_L} - p_{(o)_R} + \zeta.$$

Since $p_F \sim \mathcal{P}_n$, it must hold that at least one of a_L, a_R is nonzero (otherwise x_i and x_j would not appear in p_F). As no variable appears in $\mathcal{P}_n$ with degree greater than one, the following coefficients in p_F must vanish:

1. Coefficient of $x_i^2 x_j^2$: $\quad a_L - a_R = 0 \quad \Rightarrow a_L = a_R \neq 0.$
2. Coefficient of $x_i^2 x_j$: $\quad a_L f_L - a_R f_R = 0$, and by (1) this implies $f_L = f_R$.
3. Coefficient of $x_i x_j^2$: $\quad a_L c_L - a_R c_R = 0$, which implies $c_L = c_R$.

These equations imply that the coefficient of $x_i x_j$ also vanishes:

$$a_L c_L f_L - a_R c_R f_R = 0\,.$$

This contradicts the assumption that $p_F \sim \mathcal{P}_n$. □

Lemma 7.9. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula, and suppose that both $(o)_L$ and $(o)_R$ are multiplication gates. If, for some $\boldsymbol{\alpha} \in \mathbb{F}^n$, we have $p_F(\mathbf{x} + \boldsymbol{\alpha}) \sim \mathcal{P}_n$, then there exists a variable $x_i \in \mathbf{x}$ that branches in F.*

Proof. Assume, for the sake of contradiction, that no variable branches in F.

Pick any pair of variables $x_i \in \mathrm{var}((o)_{L,L})$ and $x_j \in \mathrm{var}((o)_{L,R})$. Since no variable branches, x_i and x_j are separated in $(o)_L$.

By Lemma 7.8, x_i and x_j cannot also be separated in $(o)_R$. Hence, at least one of them must appear in both children of $(o)_R$ (i.e., branch in $(o)_R$), or both must appear in the same child. If one of them branches in $(o)_R$, then we are done. Otherwise, assume without loss of generality that both x_i and x_j occur only in $(o)_{R,L}$.

Pick any variable $x_t \in \mathrm{var}((o)_{R,R})$ (such a variable exists since each gate computes a non-constant polynomial). Since x_t does not branch, applying Lemma 7.8 to the pair (x_i, x_t) implies that x_t must lie in the same child of $(o)_L$ as x_i, namely $(o)_{L,L}$. Similarly, applying the lemma to (x_j, x_t) implies that x_t must lie in $(o)_{L,R}$ as well.

Therefore, x_t appears in both children of $(o)_L$, contradicting the assumption that no variable branches in F. □

[18] We may assume, without loss of generality, that we subtract the two polynomials rather than add them.

7.2 Hardness of Representation for Totally Non-Structural $\sum^2$ R2F

We begin by proving a hardness of representation result for the case where at least one of the two R2Fs in the sum has an addition gate at it's root. Recall the definitions in Subsection 6.2.

Lemma 7.10. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula. Suppose that $(o)_L$ is an addition gate. Moreover, suppose that no nonzero polynomial in $\mathcal{A}^2_{F_{(o)_L}} \cup \mathcal{A}^2_{F_{(o)_R}}$ vanishes at $\mathbf{0}$.*

Then, if $n \geqslant 3$, we have

$$\mathcal{P}_n \nmid p_F \,.$$

Proof. Assume towards contradiction that the claim does not hold and $\mathcal{P}_n \mid p_F$. From Observation 7.2 we get that for every $x_i \in \mathbf{x}$, $\deg_{x_i}(p_F) \leqslant 1$. Therefore, $p_n \sim p_F$.

Using Corollary 7.4 we get that the fcg of the two leaves labeled with x_i is a multiplication gate both in $(o)_L$ and in $(o)_R$. Therefore, because $(o)_L$ is an addition gate, the sets of variables in the subformulas rooted at $(o)_{L,L}$ and $(o)_{L,R}$ must be disjoint.

Without loss of generality, assume $x_n \in \mathrm{var}(p_{(o)_{L,L}})$ and $x_{n-1} \in \mathrm{var}(p_{(o)_{L,R}})$. Thus, $x_n \notin \mathrm{var}(p_{(o)_{L,R}})$ and $x_{n-1} \notin \mathrm{var}(p_{(o)_{L,L}})$. We have

$$\partial_{x_n x_{n-1}} p_{(o)_L} = \partial_{x_n x_{n-1}}(p_{(o)_{L,L}} + p_{(o)_{L,R}}) = 0 \,.$$

As $p_F = p_{(o)_L} + p_{(o)_R}$, we get:

$$\partial_{x_n x_{n-1}} p_F = \partial_{x_n x_{n-1}}(p_{(o)_L} + p_{(o)_R}) = \partial_{x_n x_{n-1}} p_{(o)_R} \,.$$

Moreover, since $\mathcal{P}_n \sim p_F$, it follows that:

$$\partial_{x_n x_{n-1}} p_{(o)_R} = \partial_{x_n x_{n-1}} p_F \sim \partial_{x_n x_{n-1}} \mathcal{P}_n = \mathcal{P}_{[n-2]}.$$

Therefore,

$$\partial_{x_n x_{n-1}} p_{(o)_R} \sim \mathcal{P}_{[n-2]}.$$

By the assumption, $\partial_{x_n x_{n-1}} p_{(o)_R}$ does not vanish at $\mathbf{0}$, in contradiction. □

The next three claims Lemma 7.15, Lemma 7.18 and Lemma 7.22 are the main technical contribution of this part. These establish a hardness of representation result for totally non-structural $\sum^2$ R2F formulas, under the assumption that both children of the root are multiplication gates.

Lemma 7.15 handles the case where there exists a "large" set of variables that branch completely. Lemma 7.18 addresses the case in which no variable branches completely. Then Lemma 7.22 handles the case where there exists a "small" set of variables that completely split, by attempting to reduce the problem to any of the previously mentioned cases.

In these lemmas we use restriction of algebraic formulas (Definition 2.54). Moreover, the next property of generic assignments is used implicitly.

Observation 7.11. *Let F be a totally non-structural algebraic formula and let $\boldsymbol{\tau} \in \mathbb{F}^n$ be a generic assignment. Then, for any set $S \subseteq [n]$, $F|^{\boldsymbol{\tau}}_S$ is also totally non-structural.*

Proof. Let $x_i \in \mathbf{x}$ and let $v, u \in F$, such that v is a child of u and $\deg_{x_i}(p_v) > \deg_{x_i}(p_u)$. As a generic assignment to variables other than x_i won't affect its the degree the claim follows. □

Before proving the main hardness of representation results, we first establish such a result for the class $\sum^2 \prod^k \mathsf{ROF}$ (recall Definition 2.2).

Remark 7.12. In [BGV23], the authors showed that $\mathcal{G}_3$ hits the class of $\sum^2 \prod \mathsf{ROF}$. While [BGV23] presented a PIT algorithm for this class, their result does not provide any corresponding hardness of representation result, which is crucial for us, as explained in Subsubsection 1.3.1.

Lemma 7.13. *Let* $F_1, F_2 \in \prod^k \mathsf{ROF}$ *such that* $F_i = \prod_{j=1}^{m_i} f_{i,j}$*, where all* $f_{i,j}$ *(for* $i \in [2]$*,* $j \in [m_i]$*) are irreducible. Moreover, assume these irreducible factors are* **0***-irreducible and* **0***-preserved (recall Definition 2.27). Then* $F_1 + F_2$ *is* $(2k+1)$*-hard.*

Proof. Suppose, for the sake of contradiction, that the claim does not hold. Without loss of generality, assume that $\mathcal{P}_{2k+1} \mid F_1 + F_2$, and let $h \in \mathbb{F}[\mathbf{x}]$ be such that

$$F_1 + F_2 = \mathcal{P}_{2k+1} \cdot h\,.$$

Observe that any common factors of F_1 and F_2 are **0**-preserved. Hence, no variable divides these factors, and factoring them out does not affect the correctness of the lemma. We therefore assume, without loss of generality, that F_1 and F_2 are coprime.

Then, for every $i \in [2k+1]$, we have

$$(F_1 + F_2)\big|_{x_i=0} = \mathcal{P}_{2k+1}\big|_{x_i=0} \cdot h\big|_{x_i=0} \equiv 0\,.$$

Thus, $F_1|_{x_i=0} = -F_2|_{x_i=0}$. This implies that, for each $i \in [2k+1]$, the sets of irreducible factors of $F_1|_{x_i=0}$ and $F_2|_{x_i=0}$ must be identical (up to scalar multiples).

Without loss of generality, assume that for all $j \in [m_2]$ we have $|\mathrm{var}(f_{1,1})| \geqslant |\mathrm{var}(f_{2,j})|$.

Suppose there exists $i \in [2k+1]$ such that $x_i \notin \mathrm{var}(f_{1,1})$. Then, for any $j \in [m_2]$ with $x_i \in \mathrm{var}(f_{2,j})$, we have $|\mathrm{var}(f_{1,1})|_{x_i=0}| > |\mathrm{var}(f_{2,j})|_{x_i=0}|$, and therefore $f_{1,1}|_{x_i=0}$ and $f_{2,j}|_{x_i=0}$ are not similar. If $x_i \notin \mathrm{var}(f_{2,j})$, then $f_{2,j} = f_{2,j}|_{x_i=0}$, and again the two cannot be similar since $f_{1,1}$ and $f_{2,j}$ are coprime. This contradicts the assumption that the irreducible factors of $F_1|_{x_i=0}$ and $F_2|_{x_i=0}$ match. Hence, we must have $[2k+1] \subseteq \mathrm{var}(f_{1,1})$.

For each $i \in [2k+1]$, let j_i be such that $f_{1,1}|_{x_i=0} \sim f_{2,j_i}|_{x_i=0}$. Define the multiset

$$S \coloneqq \{f_{2,j_i} : i \in [2k+1]\}\,.$$

We now argue that any polynomial in S appears at most twice. Indeed, if $f_{1,1}|_{x_i=0} \sim f_{2,j_i}|_{x_i=0}$, then using the fact that these irreducible polynomials are **0**-preserved and applying the second item of Fact 2.36, we obtain that for each $j \neq i$, $x_i \mid \mathrm{Res}_{x_j}(f_{1,1}, f_{2,j_i})$. Since $f_{1,1}$ and f_{2,j_i} are coprime, we have $\mathrm{Res}_{x_j}(f_{1,1}, f_{2,j_i}) \neq 0$. Moreover, as they are also **0**-irreducible, Lemma 2.48 implies that, for every j, $\mathrm{Res}_{x_j}(f_{1,1}, f_{2,j_i})$ is 3-hard. Hence, there can be at most two variables $i, i' \neq j$ that divide $\mathrm{Res}_{x_j}(f_{1,1}, f_{2,j_i})$, and the claim follows.

By **0**-preservation, each polynomial in S must depend on all variables in $\mathrm{var}(f_{1,1})$ except possibly one, and any polynomial that appears twice in S must depend on all of $\mathrm{var}(f_{1,1})$.

Now suppose that for some $t \in \{0, 1, \ldots, k\}$ the first t polynomials in S each appear only once. Assume, without loss of generality, that in the i-th polynomial of S only the i-th variable may be missing. Then x_{t+1} appears in those t polynomials. By the observation above, x_{t+1} also appears in every other polynomial in S, and since each polynomial appears at most twice, there are at least $\left\lceil \frac{2k+1-t}{2} \right\rceil$ such distinct polynomials. Therefore, the total number of occurrences of x_{t+1} in F_2 is at least

$$t + \left\lceil \frac{2k+1-t}{2} \right\rceil \;\geqslant\; t + \frac{2k+1-t}{2} \;=\; k + \frac{t+1}{2} \;>\; k,$$

which contradicts the assumption that F_2 is read-k. □

For the next lemma we shall use the following definition:

Definition 7.14 ($\mathcal{C}_F$)**.** Let F be an algebraic formula. For any pair of gates $u, v \in F$ such that u is an ancestor of v, let $\gamma_v \in \mathbb{F}$ denote the additive constant of gate v. Define

$$\mathcal{C}_u \coloneqq \left\{ p_{F_u|_{v=\gamma_v}} \;\middle|\; v \in F_u \right\},$$

In words, $\mathcal{C}_u$ is obtained by replacing, for each descendant v of u, the subtree rooted at v with a leaf labeled by the scalar γ_v and then reducing the formula, in a similar fashion to Definition 2.54.

We also define

$$\mathcal{C}_F \coloneqq \bigcup_{u \in F_{(o)_L} \cup F_{(o)_R}} \mathcal{C}_u \,.$$

When the formula F is clear from context, we simply write $\mathcal{C}$. ◊

The reason for this definition will become clear later. In the meantime, we note that these are precisely the polynomials that appear in Lemma 2.26.

For the proof of the next lemma we shall use the definitions, claims and observations related to Definition 2.54.

Lemma 7.15. *Let $F \in \sum^2 \mathsf{R2F}$ be totally non-structural. Assume that both $(o)_L$ and $(o)_R$ are multiplication gates. Moreover, suppose there exists a set of variables S with $|S| \geqslant 5$ such that every $x_i \in S$ branches completely in F (recall Definition 7.6).*

If no nonzero polynomial $g \in \mathcal{C} \cup \mathcal{E}^2$ (recall Definition 6.5) vanishes at $\mathbf{0}$, then $\mathcal{P}_n \nmid p_F$.

Proof. Let $\boldsymbol{\tau} \in \mathbb{F}^{[n]\setminus S}$ be a generic assignment. Since every variable in S branches completely in F, we have $F|^{\boldsymbol{\tau}}_S \in \sum^2 \prod^2 \mathsf{ROF}$. From this point onward, we suppress $\boldsymbol{\tau}$ in the notation.

Remark 6.7 shows that $\mathcal{E} \subseteq \mathcal{E}^2$. Hence Claim 6.6 implies that both $p_{(o)_L}$ and $p_{(o)_R}$ are $\mathbf{0}$-preserved. By the properties of generic assignments, this further implies that both $p_{(o_{F|_S})_L}$ and $p_{(o_{F|_S})_R}$ are $\mathbf{0}$-preserved. Therefore, by Lemma 2.32, any factor of these polynomials is also $\mathbf{0}$-preserved. Consequently, $\gcd(p_{(o_{F|_S})_L}, p_{(o_{F|_S})_R})$ cannot be divisible by any variable, and factoring it out does not affect the correctness of the lemma. We may thus assume, without loss of generality, that

$$\gcd(p_{(o_{F|_S})_L}, p_{(o_{F|_S})_R}) = 1 \,.$$

For $\rho \in \{L, R\}$, let $\{h_{\rho,j}\}_{j\in[m_\rho]}$ be the irreducible factors of $p_{(o_{F|_S})_\rho}$. The next lemma (applied with $F' = F$, $S' = S$, and $\boldsymbol{\tau}' = \boldsymbol{\tau}$) shows that each $h_{\rho,j}$ is $\mathbf{0}$-irreducible. We state it in a more general form, as it will also be used in later sections.

Claim 7.16. *Let $F' \in \sum^2 \mathsf{R2F}$ be such that no nonzero polynomial $g \in \mathcal{C}_{F'} \cup \mathcal{E}^2_{F'}$ vanishes at $\mathbf{0}$. Then, for $\rho \in \{L, R\}$, $S' \subset [n]$, and a generic assignment $\boldsymbol{\tau}'$, if $(o_{F'|^{\boldsymbol{\tau}'}_{S'}})_\rho$ is a product of ROFs, then its irreducible factors are $\mathbf{0}$-irreducible.*

Proof. Let $\{h_{\rho,j}\}_{j\in[m_\rho]}$ be the irreducible factors of $p_{(o_{F'|^{\boldsymbol{\tau}'}_{S'}})_\rho}$, and let $h \in \{h_{\rho,j}\}_{j\in[m_\rho]}$ be an arbitrary factor. By Lemma 2.10 h is computed by some subgate of $(o_{F'|^{\boldsymbol{\tau}'}_{S'}})_\rho$, let this gate be v_h.

Recall the definition of a commutator from Definition 2.24. For every $i \neq j \in \mathrm{var}(h)$, let ℓ_i and ℓ_j denote the unique leaves labeled by x_i and x_j respectively in $(F'|^{\boldsymbol{\tau}'}_{S'})_{v_h}$. Let $u = \mathrm{fcg}_{(F'|^{\boldsymbol{\tau}'}_{S'})_{v_h}}(\ell_i, \ell_j)$ and let $\gamma_u \in \mathbb{F}$ be the additive constant of u. Then, by Lemma 2.26 we have

$$\Delta_{i,j} h = p_{(F'|^{\boldsymbol{\tau}'}_{S'})_{v_h}|_{u=\gamma_u}} \cdot \partial_{x_i,x_j} h \,.$$

Since h is irreducible, by Lemma 2.25 we get that $\Delta_{i,j} h \not\equiv 0$. Hence $p_{(F'|^{\tau'}_{S'})_{v_h}|_{u=\gamma_u}} \not\equiv 0$ and $\partial_{x_i,x_j} h \not\equiv 0$.

Let $w = \mathcal{O}^{F'}_{S'}(v_h)$ (recall Definition 2.54). By Lemma 2.57(2), there exist a gate $r \in F'_w$ with an additive constant $\gamma_r \in \mathbb{F}$ such that

$$\left(F'_w|_{r=\gamma_r}\right)|^{\tau'}_{S'} = (F'|^{\tau'}_{S'})_{v_h}|_{u=\gamma_u} \not\equiv 0\,. \tag{3}$$

Therefore, $F'_w|_{r=\gamma_r} \in \mathcal{C}$ and is nonzero, hence it does not vanish at $\mathbf{0}$. This implies, by the properties of a generic assignment, and (3), that $p_{(F'|^{\tau'}_{S'})_{v_h}|_{u=\gamma_u}}$ does not vanish at $\mathbf{0}$.

Since $\partial_{i,j} p_w \in \mathcal{E}^2_{F'}$, it does not vanish at $\mathbf{0}$. By Observation 2.56(3) we have that $p_{F'_w|^{\tau'}_{S'}} = h$. Hence by the properties of a generic assignment $\partial_{i,j} h$ does not vanish at $\mathbf{0}$.

We conclude that $\mathbf{0}$ is a nonzero of $\Delta_{i,j} h = p_{F'_{v_h}|_{u=\gamma_u}} \cdot \partial_{i,j} h$. Hence, by applying Claim 2.30 and the generality of i and j, we get that h is $\mathbf{0}$-irreducible. □

The claim above shows that each factor $h \in \{h_{\rho,j}\}_{\rho\in\{L,R\},\, j\in[m_\rho]}$ is $\mathbf{0}$-irreducible. It is $\mathbf{0}$-preserved by the properties of generic assignments and Lemma 2.32. The claim follows from Lemma 7.13 and the fact that $|S| \geqslant 5$. □

We now consider the second case, where no variable branches completely in F. Before establishing the corresponding hardness of representation result, we first prove the following claim.

Claim 7.17. *Let $F \in \mathsf{R2F}$ be over the variable set $[3]$. Suppose that for every $i \in [3]$, we have $x_i^2 \in mon(p_F)$, and that every other monomial in p_F with individual degree 2 in some variable has coefficient 0. Then $x_1x_2x_3 \notin mon(p_F)$.*

Proof. For $i \in [3]$, relabel the leaves labeled by x_i as $x_{i,0}$ and $x_{i,1}$. By the assumption of the claim, $fcg^{op}(x_{i,0}, x_{i,1}) = \times$ for all i.

Assume, for a contradiction, that $x_1x_2x_3$ appears in the original polynomial. It follows that, without loss of generality, $x_{1,0}x_{2,0}x_{3,0}$ appears in the relabeled polynomial. Thus, for $i \neq j$, we must have $fcg^{op}(x_{i,0}, x_{j,0}) = \times$. Denote $v_{i,j} = fcg(x_{i,0}, x_{j,0})$. Assume, without loss of generality, that $v_{1,2}$ is the deepest among the $v_{i,j}$. Further, assume that $x_{1,0} \in (v_{1,2})_L$ and $x_{2,0} \in (v_{1,2})_R$. We claim that $u_1 = v_{1,2}$.

Indeed, let $u_1 = fcg(x_{1,0}, x_{1,1})$. Then u_1 cannot be a descendant of $(v_{1,2})_L$, as this would imply that the monomial $x_1^2x_2$ appears in p_F. Similarly, $v_{1,2}$ cannot be a descendant of u_1. Thus, it must be that $u_1 = v_{1,2}$. A similar argument shows that $u_2 = v_{1,2}$.

Since $v_{1,2}$ is the deepest among the $v_{i,j}$ and $u_1 = v_{1,2}$, it follows that either $v_{1,2} = v_{2,3}$ or $v_{1,2}$ is a descendant of $v_{1,3}$. The latter case cannot occur, as it would imply that the monomial $x_1^2x_3$ appears in p_F. Thus, all $v_{i,j}$ and all u_i coincide and are equal to the same gate v. Observe that both $p_{(v)_L}$ and $p_{(v)_R}$ must be linear functions (otherwise we would obtain a contradiction). This implies, however, that $\deg(p_F) = 2$, a contradiction. □

Lemma 7.18. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula. Suppose that both $(o)_L$ and $(o)_R$ are multiplication gates, and that no variable branches completely in F. Moreover, Suppose that no polynomial in $\mathcal{A}^3_{F_{(o)_L}} \cup \mathcal{A}^3_{F_{(o)_R}}$ vanishes at $\mathbf{0}$.*

Then, if $n \geqslant 5$, we have:

$$\mathcal{P}_n \nmid p_F.$$

Proof sketch. We give a sketch of the proof and postpone the full proof to Subsubsection 7.3.1

Assume for contradiction that $\mathcal{P}_n \mid p_F$. Therefore, Corollary 7.3 implies that $p_F \sim \mathcal{P}_n$.

As o is an addition gate and any additive constants at $(o)_L$ or $(o)_R$ can be pushed up to the root and absorbed into the additive constant of o. Hence, we can assume w.l.o.g. that $(o)_L$ and $(o)_R$ have no additive constants. Let $\zeta \in \mathbb{F}$ be the additive constant of o. Dentoe

$$\tilde{p}_F \coloneqq p_F - \zeta.$$

By Lemma 7.9, there exists $x_i \in [n]$ which branches in F. Let J be the set of variables that branch in F. To ease the notation we denote $i = 1$ from now on, but the analysis holds for any $x_j \in J$. Assume w.l.o.g., that x_1 branches in $(o)_L$. The assumption that no variable branches completely in F, implies that x_1 does not branch in $(o)_R$. Hence, all occurrences of x_1 in $(o)_R$ are confined to a single subformula. Without loss of generality, assume that x_1 does not appear in $(o)_{R,L}$.

Let $p \coloneqq p_{(o)_{R,L}}$, and factor it as $p = \prod_{j=1}^{m} p_j$, where each p_j is irreducible. Using Lemma 2.33 we get that p is $\mathbf{0}$-preserved and then, by Lemma 2.32 we get that every irreducible factor of it is also $\mathbf{0}$-preserved.

Since x_1 branches in $(o)_L$, we can write:

$$\begin{aligned}
p_{(o)_{L,L}} &= a_1 x_1 + b_1, \quad p_{(o)_{L,R}} = a_2 x_1 + b_2, \\
p_{(o)_L} &= (a_1 x_1 + b_1)(a_2 x_1 + b_2) = a_1 a_2 x_1^2 + (a_1 b_2 + a_2 b_1) x_1 + b_1 b_2, \\
p_{(o)_R} &= p \cdot (c_2 x_1^2 + c_1 x_1 + c_0), \\
p_F &= p_{(o)_L} + p_{(o)_R} + \zeta,
\end{aligned} \tag{4}$$

for some $a_1, a_2, b_1, b_2, c_0, c_1, c_2 \in \mathbb{F}[\mathbf{x} \setminus \{x_1\}]$.

Since $p_F \sim \mathcal{P}_n$ by assumption, we have the identities

$$p \cdot c_2 = -a_1 a_2 \quad \text{and} \quad p \cdot c_0 = -(b_1 b_2 + \zeta). \tag{5}$$

We argue that a_1 and a_2 are $\mathbf{0}$-preserved. Indeed, to see this, observe that $a_1 a_2 = \partial_{x_1^{\deg}} p_{(o)_L}$. This means that for every $x_i \in x_{-1}$, $\partial_{x_i^{\deg}} a_1 a_2 \in \mathcal{A}_{F_{(o)_L}}$. Hence, by Lemma 2.33 we get that $a_1 a_2$ is $\mathbf{0}$-preserved. The claim now follows from Lemma 2.32. The next claim is stated for x_1, but a similar claim holds for every $x_j \in J$ and the appropriate p.

Claim 7.19. *If Lemma 7.18 does not hold then var*$(p) = \mathbf{x}_{-1}$

The following claim is an important ingredient in the proof of Claim 7.19.

Claim 7.20. *If Lemma 7.18 does not hold, and var*$(p) \neq \mathbf{x}_{-1}$*, then every* $x_i \in$ *var*$((o)_{R,L}) \cup$ *var*$((o)_{R,R})$ *must appear once in* $(o)_{L,L}$ *and once in* $(o)_{L,R}$.

Since no variable branches completely, it follows that every variable in $\mathrm{var}((o)_{R,L})$ is read-twice in $(o)_{R,L}$. In particular, this implies that

$$\mathrm{var}((o)_{R,L}) \cap \mathrm{var}((o)_{R,R}) = \emptyset. \tag{6}$$

This last equation together with Claim 7.20 lead to a contradiction from which we conclude that Claim 7.19 holds. Relying on it we further obtain

Claim 7.21. *If Lemma 7.18 does not hold then no other variable except* x_1 *branches in* F.

Claim 7.19 and (6) imply that $\mathrm{var}(p_{(o)_{R,R}}) = \{1\}$. Now, for each $i \in \{L, R\}$, define

$$S_i \coloneqq \mathrm{var}(p_{o_{L,i}}) \setminus \{1\}.$$

Assume without loss of generality that $|S_L| \geqslant |S_R|$. Since $n \geqslant 5$, it follows that $|S_L| \geqslant 2$.

Let $\boldsymbol{\tau} \in \mathbb{F}^{S_R}$ be a generic assignment to the variables in S_R, and let $\delta \in \mathbb{F}$ be a root of the univariate polynomial $p_{(o)_{L,R}}|_{S_R \leftarrow \boldsymbol{\tau}}$. Then, for some nonzero $\alpha \in \mathbb{F} \setminus \{0\}$ and some $\gamma \in \mathbb{F}$, we obtain:

$$\begin{aligned}\alpha \cdot \mathcal{P}_{[n]\setminus \mathrm{var}(p_{(o)_{L,R}})} &= \left(p_{(o)_L} + p \cdot p_{(o)_{R,R}} + \zeta\right)\Big|_{S_R \leftarrow \boldsymbol{\tau},\, x_1 = \delta} \\ &= \left(p_{(o)_{L,L}} \cdot p_{(o)_{L,R}}\right)\Big|_{S_R \leftarrow \boldsymbol{\tau},\, x_1 = \delta} + p|_{S_R \leftarrow \boldsymbol{\tau}} \cdot p_{(o)_{R,R}}|_{x_1 = \delta} + \zeta \\ &= \gamma \cdot p|_{S_R \leftarrow \boldsymbol{\tau}} + \zeta.\end{aligned}$$

However, the left-hand side is not $\mathbf{0}$-preserved, while the right-hand side is, in contradiction. This concludes the proof of Lemma 7.18. □

Lemma 7.22. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula. Suppose that both $(o)_L$ and $(o)_R$ are multiplication gates. Let S_1 be the set of completely branched variables in F, and assume that $1 \leqslant |S_1| \leqslant 4$. Furthermore, assume that no nonzero polynomial in $\mathcal{C} \cup \mathcal{E}^3$ vanishes at $\mathbf{0}$.*

Then, if $n \geqslant 23$, we have

$$\mathcal{P}_n \nvdash p_F.$$

Proof sketch. Our strategy is to find a small set of variables, make a generic assignment to them, and obtain a modified formula that satisfies the conditions of Lemma 7.10 or Lemma 7.18, and pull the conclusion back to F.

Denote $F_1 = F$ and let $\boldsymbol{\tau}$ be a generic assignment to all the variables. Consider the following process: Let S_i be the set of completely branched variables of F_i. At the ith step we substitute $\boldsymbol{\tau}_{S_i}$ to S_i, and let $F_{i+1} \coloneqq F_i\big|_{S_i \leftarrow \boldsymbol{\tau}_{S_i}}$. We terminate the process if one of the following conditions hold, where o_i is the root of F_i:

1. either $(o_i)_L^{\mathrm{op}} = +$, or $(o_i)_R^{\mathrm{op}} = +$,
2. $\left|\bigcup_{j=1}^{i} S_j\right| \geqslant 18$, or
3. $S_i = \emptyset$, i.e., no variable branches completely in F_i.

We now prove that if the process stopped at the mth formula, F_m, then $\mathcal{P}_{V_m} \nvdash p_{F_m}$. We show it by analyzing the different stopping conditions. Denote $S \coloneqq \bigcup_{j \in [m]} S_j$.

Case 1. As we did not halt at the previous step, and as $S_m = \emptyset$, we have that $|S| < 18$. Since $\mathcal{A}^2_{F_{(o)_L}} \cup \mathcal{A}^2_{F_{(o)_R}} \subseteq \mathcal{E}^3$, and each polynomial in $\mathcal{A}^2_{F_{(o_m)_L}} \cup \mathcal{A}^2_{F_{(o_m)_R}}$ is obtained from a generic assignment to a polynomial in $\mathcal{A}^2_{F_{(o)_L}} \cup \mathcal{A}^2_{F_{(o)_R}}$, by item 3 of Observation 2.56, it follows that every nonzero polynomial in $\mathcal{A}^2_{F_{(o_m)_L}} \cup \mathcal{A}^2_{F_{(o_m)_R}}$ does not vanish at $\mathbf{0}$. By Lemma 7.10 we get that F_m is 3-hard. Since $n \geqslant 23$ and $|S| < 18$, we have

$$\mathcal{P}_{[n]\setminus S} \nvdash F_m,$$

which implies

$$\mathcal{P}_n \nvdash p_F.$$

Let us postpone the analysis of Case 2, and prove the statement assuming that the process stopped because of Case 3.

Case 3. Since $n \geqslant 23$ and $|S| < 18$, it follows that $|\mathrm{var}(F_i)| > 23 - 18 = 5$. As in the analysis of Case 1, every nonzero polynomial in $\mathcal{A}^3_{F_{o_{iL}}} \cup \mathcal{A}^3_{F_{o_{iR}}}$ does not vanish at $\mathbf{0}$. Therefore, by Lemma 7.18, we conclude that F_i is 5-hard, which implies the claim in the same manner as in Case 1.

The analysis of Case 2 takes most of the proof so we highlight the important steps and give the full proof in Subsubsection 7.3.2.

Case 2. From properties of generic assignments (see Observation 7.11), we know that for every $j \in [m]$, p_{F_j} is totally non-structural. Consequently, for every $t \in V_j$, $\deg_{x_t}(p_{(o_j)_R}) = 2$.

Definition 7.23. Let $\rho \in \{L, R\}$ and let $t \in [m]$. We denote with $\ell_{t,\rho,l}$ and $\ell_{t,\rho,r}$ the two leaves labeled by x_t in o_ρ. Let $v_{t,\rho} = \mathrm{fcg}(\ell_{t,\rho,l}, \ell_{t,\rho,r})$. Similarly, we denote the left and right leaves in $(o_j)_\rho$ by $\ell_{t,j,\rho,l}$ and $\ell_{t,j,\rho,r}$, respectively and set

$$v_{t,j,L} = \mathrm{fcg}_{(o_j)_L}(\ell_{t,j,L,l}, \ell_{t,j,L,r}) \quad \text{and} \quad v_{t,j,R} = \mathrm{fcg}_{(o_j)_R}(\ell_{t,R,l}, \ell_{t,R,r})\,.$$

the fcg of the two leaves labeled by x_t in $(o_j)_L$ and $(o_j)_R$, respectively. ◊

We note that by simple properties of generic assignments, it holds that

$$\mathcal{L}^F_{[j-1]}(v_{t,j,\rho}) = v_{t,\rho} \quad \text{and} \quad \mathcal{L}^F_{[j-1]}(\ell_{t,j,\rho,l}) = \ell_{t,\rho,l}\,. \tag{7}$$

Thus, for brevity of presentation, we omit the index j from both $v_{t,j,\rho}$ and $\ell_{t,j,\rho,l}$.

Assume for a contradiction that $\mathcal{P}_n \mid p_F$. Corollary 7.3 implies that we actually have $\mathcal{P}_n \sim p_F$. The next claim shows that in the process above, we get that at every step at most one variable completely branches.

Corollary 7.3 implies that we actually have $\mathcal{P}_n \sim p_F$.

Claim 7.24. $|S_j| = 1$ *for every* $j \in [m]$, *and in particular* $m = 18$. *Furthermore, in each such* F_j *one of the grandchildren of the root is supported on* S_j.

From now on we assume without loss of generality that $S_j = \{x_j\}$ for every $j \in [m]$.

To obtain an even more defined structure, note that by the pigeonhole principle, in at least half of the rounds of the process, the grandchild of the root that computes a univariate polynomial at that round is a child of $(o_i)_L$. Without loss of generality (by rotating the children of $(o_i)_R$ if necessary), we may assume that this univariate polynomial is computed by $(o_i)_{L,L}$. Denote this set of rounds by $\mathcal{I}$.

We now focus only on these rounds and apply the generic assignment to all variables that were constructed in the other rounds, i.e., to the variables x_j such that $j \in S \setminus \mathcal{I}$. By the properties of generic assignments, we observe that if we were to rerun the process from the beginning, the first $|\mathcal{I}|$ variables collected would be exactly those with indices in $\mathcal{I}$.

Thus, from now on we assume, without loss of generality, that $|S| \geqslant m/2 = 18/2 = 9$ and that, at the ith step, the polynomial computed at $(o_i)_{L,L}$ is univariate in x_i. Denote $m' = m/2 = 9$ and $n' = n - m/2 = 23 - 18/2 = 14$.

Corollary 7.25. *For every* $j \in [m']$ *it holds that* $\mathrm{var}\left((o_j)_{L,R}\right) = [j : n']$.

Proof. As a generic restriction keeps any non-structural variable non-structural, and since exactly one variable branches completely at each step of the process it follows that $(o_j)_L$ contains all the variables in V_j. □

Let $\delta_j \in \mathbb{F} \setminus \{0\}$ be such that $p_{(o_j)_{L,L}} \sim (x_j - \delta_j)$. Clearly, we can assume w.l.o.g., that

$$p_{(o_j)_{L,L}} = (x_j - \delta_j)\,. \tag{8}$$

As before, we can "push" the additive constants in $(o_j)_L$ and $(o_j)_R$ to o_j and assume both $(o_j)_L$ and $(o_j)_R$ are pure multiplication gates. Thus, for $t \in V_j \setminus \{j\}$ we have

$$p_{(o_j)_L}|_{x_j=\delta_j} = 0\,, \tag{9}$$

$$(x_j - \delta_j) \mid \partial_{x_t} F_{(o_j)_L}\,. \tag{10}$$

Let α_j be such that

$$p_{F_j} = p_{(o_j)_L} + p_{(o)_R} + \alpha_j. \tag{11}$$

For $j \in [m']$, we compute the quotient and remainder of dividing $p_{(o_j)_{R,L}}$ and $p_{(o_j)_{R,R}}$ by $(x_j - \delta_j)$:

$$p_{(o_j)_{R,L}} = a_{j,L}\,(x_j - \delta_j) + b_{j,L} \qquad \text{and} \qquad p_{(o_j)_{R,R}} = a_{j,R}\,(x_j - \delta_j) + b_{j,R}\,. \tag{12}$$

We have,

$$\begin{aligned} p_{F_j} &= (x_j - \delta_j)p_{(o_j)_{L,R}} + \left(a_{j,L}\,(x_j - \delta_j) + b_{j,L}\right)\left(a_{j,R}\,(x_j - \delta_j) + b_{j,R}\right) + \alpha_j \\ &= (x_j - \delta_j)\Big(p_{(o_j)_{L,R}} + a_{j,L}a_{j,R}\,(x_j - \delta_j) + (a_{j,L}b_{j,R} + a_{j,R}b_{j,L})\Big) + b_{j,L}b_{j,R} + \alpha_j\,. \end{aligned} \tag{13}$$

Hence,

$$\mathcal{P}_{[j:n']}|_{x_j=\delta_j} \sim p_{F_j}|_{x_j=\delta_j} = b_{j,L}b_{j,R} + \alpha_j. \tag{14}$$

The following claim and corollary imply that, upon restricting F to the variables indexed by $[m'-2]$, $(o)_R$ becomes a product of two ROFs.

Claim 7.26. *Under our contradiction assumption it holds that* $s := \min(|var((o)_{R,L})|, |var((o)_{R,R})|) \geqslant m'-2$.

The idea of the proof is the following. Assume that the claim does not hold and, without loss of generality, $|\mathrm{var}((o)_{R,L})| < m'-2$. We find two indices t, i such that $i \notin \mathrm{var}((o)_{R,L})$ and $t \notin (v_{i,L})_L$. A simple calculation then shows that when assigning $\boldsymbol{\tau}$ to the variables in $[s] \setminus \{t, i\}$ and then taking a derivative according to $x_{n'}$, we obtain a polynomial of the form $(x_t - \beta)\cdot a + (x_i - \delta_i)\cdot b \sim \mathcal{P}_{[n']\setminus([s]\setminus\{i,t,n'\}}$, with $\delta_t, \delta_i \neq 0$, which is a contradiction.

Corollary 7.27. *The restriction of* F *to the variables indexed by* $[m'-2]$ *results in a formula* $F' = F_{[m'-2]}$ *such that* $(o_{F'})_R$ *is a product of two* ROFs.

Proof. In Claim 7.26 we proved that each child of $(o)_R$ has at least $m'-2$ many variables. By the fact that the first $m'-2$ variables branch in our process, it follows that $[m'-2] \subseteq \mathrm{var}((o)_R) \cap \mathrm{var}((o)_L)$. As $(o)_R$ is an R2F, and a multiplication gate, it must be the case that $(o)_R|_{[m'-2]}$ is a product of two ROFs. □

From now on, we consider $F' = F|_{[m'-2]}$. Observe that, had we run our process on F' instead of F, we would obtain the same set of branched variables (excluding those that were restricted). Furthermore, by the properties of generic assignments, F' satisfies the same assumptions as F in the statement of the claim. For readability, we continue to use the notation F instead of F', keeping in mind that $\mathrm{var}(F) = [m'-2]$ and that by Corollary 7.27, $(o_F)_R$ is a product of two ROFs.

Claim 7.28. *For every* $j \in [m'-5]$, $\alpha_j \neq 0$.

The idea is that the assumption in the statement of Lemma 7.22 and Claim 7.16, imply that each of the irreducible factors of $(o)_R$ is also **0**-irreducible. Analyzing (14) we conclude that $\alpha_j \neq 0$. Since $\mathcal{P}_{[j+1:m'-2]} - \alpha_j$ is irreducible when $m'-2 > j+1$, the same equation yields the following corollary.

Corollary 7.29. *For every* $j \in [m'-7]$, $b_{j,R} \in \mathbb{F}$ *or* $b_{j,L} \in \mathbb{F}$.

We next prove that $x_1 - \delta_1$ divides many of the leaf-root paths in o_R. This also let us conclude whether $b_{j,R} \in \mathbb{F}$ or $b_{j,L} \in \mathbb{F}$.

Claim 7.30. *One of the following two symmetrical cases hold:*

$$(x_1-\delta_1) \mid \prod_{u \in Unv(F_{v_{2,R}})_L(\ell_{t,R,l})} p_u \quad \text{and} \quad b_{2,L} \in \mathbb{F}\text{, or}$$
$$(x_1-\delta_1) \mid \prod_{u \in Unv(F_{v_{2,R}})_R(\ell_{2,R,r})} p_u \quad \text{and} \quad b_{2,R} \in \mathbb{F}.$$

Sketch of proof. Consider the first and second derivatives of $F|_{x_1=\delta_1}$ according to x_2. It is not hard to prove that

$$0 \not\equiv \partial_{x_2}\mathcal{P}_{[2:m'-2]} \sim \partial_{x_2} p_{(o)_R}|_{x_1=\delta_1} \quad \text{and} \quad 0 \equiv \partial_{x_2^2}\mathcal{P}_{[2:m'-2]} \sim \partial_{x_2^2} p_{(o)_R}|_{x_1=\delta_1}. \tag{15}$$

From Lemma 2.8 and definition of $v_{t,R}$ we conclude that

$$x_1-\delta_1 \mid \prod_{u \in \mathrm{Unv}_{F_{(v_{t,R})_L}}(\ell_{t,R,l})} p_u \quad \text{or} \quad x_1-\delta_1 \mid \prod_{u \in \mathrm{Unv}_{F_{(v_{t,R})_R}}(\ell_{t,R,r})} p_u. \qquad \square$$

We next prove that if it is the case that $x_1 - \delta_1 \mid \prod_{u \in \mathrm{Unv}_{F_{(v_{t,R})_L}}(\ell_{t,R,l})} p_u$ then $b_{1,L} \in \mathbb{F}$ (the other case is analogous). This is shown by inspection of $\partial_{x_2}\mathcal{P}_{[m'-2]}\big|_{x_1=\delta_1}$ and, relying on Lemma 2.8, (9), (11), we obtain a contradiction when $b_{1,R} \in \mathbb{F}$.

We are now ready to conclude the proof. By Claim 7.30, (12) and (14), we obtain that for some $\beta \in \mathbb{F}$,

$$b_{1,R} = \beta\mathcal{P}_{[2:m'-2]} - \alpha_1 \quad \text{and} \quad p_{F_{(o)_{R,R}}} = a_{1,R}(x_1-\delta_1) + \beta\mathcal{P}_{[2:m'-2]} - \alpha_1.$$

We have

$$\partial_{x_2} p_{(o_1)_{R,R}} = (\partial_t a_{1,R})(x_1-\delta_1) + \beta\mathcal{P}_{[3:m'-2]}. \tag{16}$$

By Corollary 7.27, $F_{o_R} \in \prod^2 \mathsf{ROF}$. In particular, $F_{(o)_{R,R}}$ is a ROF.

Observe that $a_{1,R} = \partial_1 p_{(o_1)_{R,R}}$, and it does not involve x_1. As the derivative of an ROF is an ROF (this was proved in [SV15] and it is a special case of Lemma 2.8), we get that both $(\partial_t a_{1,R})(x_1-\delta_1)$ and $\partial_t p_{(o_1)_{R,R}}$ are ROFs.

Recalling that every preserving assignment is also a justifying assignment (recall Definition 2.31), we get form Proposition 2.23(2) that both $\partial_t p_{(o_1)_{R,R}}$ and $a_{1,R} = \partial_1 p_{(o_1)_{R,R}}$ are **0**-justified ROFs. As $a_{1,R}$ does not involve x_1 and $\delta_j \neq 0$, we similarly get that $(\partial_t a_{1,R})(x_1-\delta_1)$ is a **0**-justified ROF.

Hence, (16) implies that $P_{V\setminus\{t\}}$ is a sum of two **0**-justified ROFs. Since

$$|[3 : m'-2]| = m'-4 > 2$$

this contradicts Lemma 2.46. $\square$

The following result follows almost immediately.

Proposition 7.31. *Let* $F \in \sum^2 \mathsf{R2F}$ *be totally non-structural. Suppose that no nonzero polynomial in* $\mathcal{C} \cup \mathcal{E}^3$ *vanishes at* $\mathbf{0}$*. Then, if* $n \geqslant 23$*, we have* $\mathcal{P}_n \nmid p_F$*.*

Proof. First, assume that one of the children of the root of F is an addition gate. By definition, $\mathcal{A}^2_{F_{(o)_L}} \cup \mathcal{A}^2_{F_{(o)_R}} \subseteq \mathcal{E}^3$. Therefore, every nonzero polynomial in $\mathcal{A}^2_{F_{(o)_L}} \cup \mathcal{A}^2_{F_{(o)_R}}$ does not vanish at $\mathbf{0}$. Hence, by Lemma 7.10, we obtain that p_F is 3-hard.

Otherwise, both children of the root are multiplication gates. Suppose there are at least five variables that branch completely in F. Since $\mathcal{E}^2 \subseteq \mathcal{E}^3$, every polynomial in $\mathcal{C} \cup \mathcal{E}^2$ does not vanish at $\mathbf{0}$. Therefore, by Lemma 7.15, we obtain that p_F is 5-hard.

Otherwise, there are at most four variables that branch completely in F. If at least one variable branches completely in F, then Lemma 7.22 implies the claim. Otherwise, the claim follows from Lemma 7.18. □

7.3 Missing proofs

In this section we give the proofs that we've sketched or omitted from the previous subsections.

7.3.1 Proof of Lemma 7.18

Proof of Lemma 7.18. We provide the missing details for the proof sketch given earlier. For convenience, we briefly recall some of the notation from the proof sketch. We assume for contradiction that $\mathcal{P}_n \mid p_F$, and hence $p_F \sim \mathcal{P}_n$. We also assume that $(o)_L$ or $(o)_R$ have no additive constants and set

$$\tilde{p}_F \coloneqq p_F - \zeta,$$

where $\zeta \in \mathbb{F}$ be the additive constant of o. Let J denote the set of variables that branch in F. To simplify notation, we assume that $x_1 \in J$, but the analysis holds for any $x_j \in J$.

We assume w.l.o.g., that x_1 branches in $(o)_L$, but not in $(o)_R$. Hence, all occurrences of x_1 in $(o)_R$ are confined to a single subformula. Without loss of generality, assume that x_1 does not appear in $(o)_{R,L}$. Let $p \coloneqq p_{(o)_{R,L}}$, and factor it as $p = \prod_{j=1}^{m} p_j$, where each p_j is irreducible. Using Lemma 2.33 we get that p is $\mathbf{0}$-preserved and then, by Lemma 2.32 we get that every irreducible factor of it is also $\mathbf{0}$-preserved. Since x_1 branches in $(o)_L$, we can write:

$$\begin{aligned}
p_{(o)_{L,L}} &= a_1 x_1 + b_1, \quad p_{(o)_{L,R}} = a_2 x_1 + b_2, \\
p_{(o)_L} &= (a_1 x_1 + b_1)(a_2 x_1 + b_2) = a_1 a_2 x_1^2 + (a_1 b_2 + a_2 b_1) x_1 + b_1 b_2, \\
p_{(o)_R} &= p \cdot (c_2 x_1^2 + c_1 x_1 + c_0), \\
p_F &= p_{(o)_L} + p_{(o)_R} + \zeta,
\end{aligned} \tag{17}$$

for some $a_1, a_2, b_1, b_2, c_0, c_1, c_2 \in \mathbb{F}[\mathbf{x} \setminus \{x_1\}]$. The overall structure of the formula F is therefore:

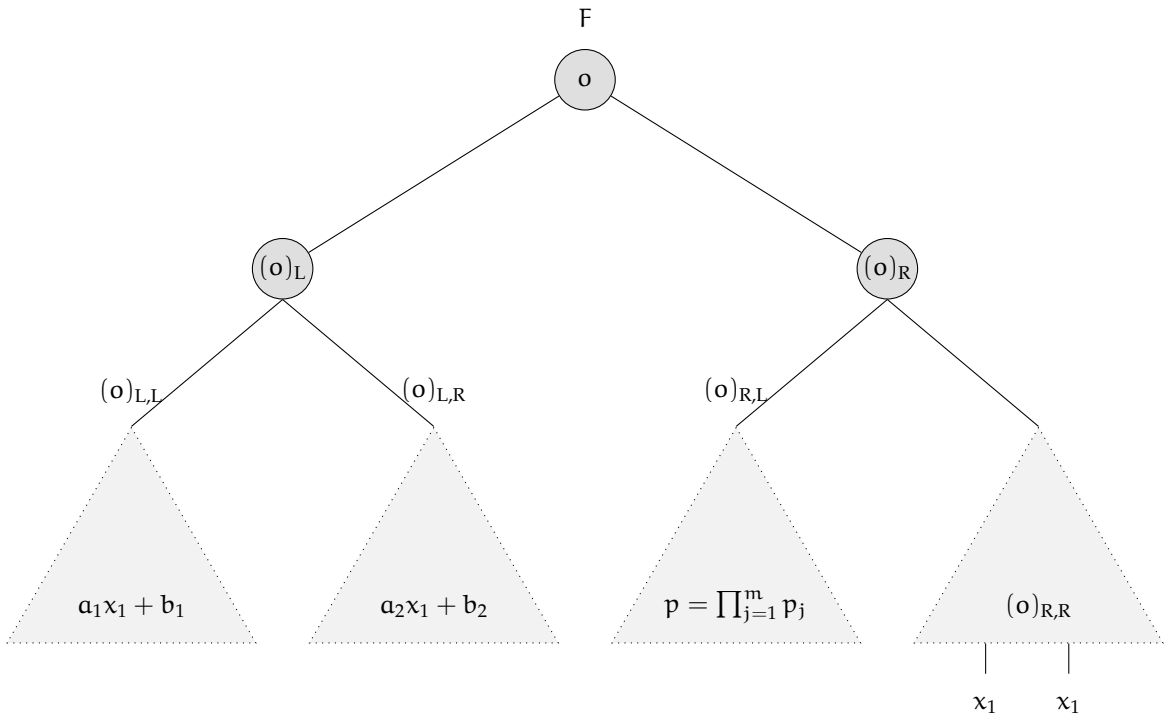

Figure 6: Structure of the formula in the proof of Lemma 7.18.

Since $p_F \sim \mathcal{P}_n$ by assumption, we have the identities

$$p \cdot c_2 = -a_1 a_2 \quad \text{and} \quad p \cdot c_0 = -(b_1 b_2 + \zeta). \tag{18}$$

We argue that a_1 and a_2 are $\mathbf{0}$-preserved. Indeed, to see this, observe that $a_1 a_2 = \partial_{x_1^{\deg}} p_{(o)_L}$. This means that for every $x_i \in x_{-1}$, $\partial_{x_i^{\deg}} a_1 a_2 \in \mathcal{A}_{F_{(o)_L}}$. Hence, by Lemma 2.33 we get that $a_1 a_2$ is $\mathbf{0}$-preserved. The claim now follows from Lemma 2.32. The next claim is stated for x_1, but a similar claim holds for every $x_j \in J$ and the appropriate p.

Claim 7.19. *If Lemma 7.18 does not hold then var*$(p) = \mathbf{x}_{-1}$

Proof. Assume, for the sake of contradiction, that $\mathrm{var}(p) \neq \mathbf{x}_{-1}$; that is, there exists some variable, other than x_1, that appears only in $(o)_{R,R}$.

Claim 7.20. *If Lemma 7.18 does not hold, and var*$(p) \neq \mathbf{x}_{-1}$*, then every* $x_i \in \mathit{var}((o)_{R,L}) \cup \mathit{var}((o)_{R,R})$ *must appear once in* $(o)_{L,L}$ *and once in* $(o)_{L,R}$.

Proof. Suppose, for contradiction, that this is not the case. Without loss of generality, suppose $x_2 \in \mathrm{var}(p_1)$ and that x_2 is read-twice in $(o)_{L,L}$. Since x_2 does not appear in $(o)_{L,R}$, (18) implies that $p_1 \mid a_1$. Let $g \in \mathbb{F}[\mathbf{x}]$ be such that $a_1 = p_1 g$.

Define $S \coloneqq \mathrm{var}(p_1) \setminus \{2\}$. Let $\boldsymbol{\tau} \in \mathbb{F}^S$ be a generic assignment. By the fact that p_1 is $\mathbf{0}$-preserved and the fact that $\boldsymbol{\tau}$ is a generic assignment it follows that $p_1|_{S \leftarrow \boldsymbol{\tau}}$ is $\mathbf{0}$-preserved. Hence, by Observation 2.50 there exists $\delta \in \mathbb{F} \setminus \{0\}$ which is a root of $p_1|_{S \leftarrow \boldsymbol{\tau}}$. Then, for some $\alpha \in \mathbb{F} \setminus \{0\}$ and $\eta \in \mathbb{F}$,

we obtain the following equation:

$$\begin{aligned}
\alpha \cdot \mathcal{P}_{[n]\setminus \mathrm{var}(p_1)} &= \left(p_{(o)_L} + p \cdot p_{(o)_{R,R}} + \zeta\right)\Big|_{S\leftarrow \boldsymbol{\tau},\, x_2=\delta} \\
&= \left(p_{(o)_L} + \zeta\right)\Big|_{S\leftarrow \boldsymbol{\tau},\, x_2=\delta} \\
&= \left(a_1 a_2 x_1^2 + (a_1 b_2 + a_2 b_1)x_1 + b_1 b_2 + \zeta\right)\Big|_{S\leftarrow \boldsymbol{\tau},\, x_2=\delta} \\
&= \left(p_1 g a_2 x_1^2 + (p_1 g b_2 + a_2 b_1)x_1 + b_1 b_2 + \zeta\right)\Big|_{S\leftarrow \boldsymbol{\tau},\, x_2=\delta} \\
&= \left(a_2 b_1 x_1 - p_1 \cdot c_0\right)\Big|_{S\leftarrow \boldsymbol{\tau},\, x_2=\delta} \\
&= a_2|_{S\leftarrow \boldsymbol{\tau}} \cdot \eta x_1.
\end{aligned}$$

Since by our assumption $\mathrm{var}(p_1) \neq x_{-1}$, the monomial $\mathcal{P}_{[n]\setminus \mathrm{var}(p_1)}$ depends on at least one variable other than x_1. On the other hand, a_2 is nonzero and $\mathbf{0}$-preserved, and $\boldsymbol{\tau}$ is a generic assignment, so $a_2|_{S\leftarrow \boldsymbol{\tau}}$ is a nonzero, $\mathbf{0}$-preserved polynomial. Hence, $\mathrm{var}(a_2|_{S\leftarrow \boldsymbol{\tau}}) = [n] \setminus (\mathrm{var}(p_1) \cup \{x_1\})$. Thus, there is some $x_i \in [n]\setminus(\mathrm{var}(p_1) \cup \{x_1\})$ that divides $a_2|_{S\leftarrow \boldsymbol{\tau}}$, in contradiction to it being $\mathbf{0}$-preserved. This concludes the proof of Claim 7.20. □

Thus, every variable in $\mathrm{var}(p)$ branches in $(o)_L$. Combined with the assumption that no variable branches completely, it follows that every variable in $\mathrm{var}(p)$ is read-twice in $(o)_{R,L}$. In particular, this implies that

$$\mathrm{var}((o)_{R,L}) \cap \mathrm{var}((o)_{R,R}) = \emptyset. \tag{19}$$

Since $n \geqslant 5$ and the two variable sets above are disjoint, it must be the case that one of the following holds: $|\mathrm{var}(p_{(o)_{R,R}})| \geqslant 3$ or $|\mathrm{var}(p_{(o)_{R,L}})| \geqslant 3$. Observe that if the latter holds (and thus $|\mathrm{var}(p_{(o)_{R,R}})| < n-1$), then by applying a symmetrical argument now with respect to a variable in $\mathrm{var}(p_{(o)_{R,L}})$ instead of x_1, we get that every variable in $\mathrm{var}(p_{(o)_{R,R}})$ branches in $(o)_L$.

For ease of exposition, and without loss of generality, we assume that $|\mathrm{var}(p_{(o)_{R,R}})| \geqslant 3$, and continue using the original notation $p = p_{(o)_{R,L}}$. We thus have that every variables in $(o)_R$ branches in $(o)_L$.

Claim 7.32. *Let $x_i \in \mathrm{var}(p_{(o)_{R,R}})$. Then, either there exists a monomial $m \in \mathrm{mon}(p_{(o)_{R,R}})$ such that $x_i \in \mathrm{Supp}(m)$ and $|\mathrm{Supp}(m)| \geqslant 3$, or p_F is 4-hard.*

Proof. Assume such a monomial does not exist. Let m be any multilinear monomial with support size 3 such that $\mathrm{Supp}(m) \subseteq \mathrm{var}(p_{(o)_{R,R}})$ and $x_i \in \mathrm{Supp}(m)$. By (19), m does not appear in $p_{(o)_R}$ as it does not appear in $p_{(o)_{R,R}}$. We get that

$$0 \not\equiv \partial_m \mathcal{P}_n \sim \partial_m p_F = \partial_m p_{(o)_L} + \partial_m p_{(o)_R} = \partial_m p_{(o)_L}.$$

Therefore, $\partial_m p_{(o)_L} \in \mathcal{A}^3_{\Gamma_{(o)_L}}$ which implies that $\partial_m p_{(o)_L}(\mathbf{0}) \neq 0$. By Lemma 2.39 we get that p_F is 4-hard. □

If p_F is 4-hard then Lemma 7.18 holds, so let us assume that there exists such a monomial m. We next show that we may assume w.l.o.g., that m has individual degree 2 in some variable $x_i \in \mathrm{var}(m)$ and any degree in some other variable $x_j \in \mathrm{var}(m)$. Indeed, otherwise, all such monomials in $p_{(o)_{R,R}}$ are multilinear. Then, by Claim 7.17, there exist $i, j \in \mathrm{var}(p_{(o)_{R,R}})$ such that

$$\partial_{x_i^2 x_j} p_{(o)_{R,R}} \not\equiv 0.$$

By (19) x_i does not appear in p, and by the discussion following Claim 7.20 it also branches in $(o)_L$. Therefore, we can repeat the entire argument for $x_i \notin \mathrm{var}(p)$ instead of $x_1 \notin \mathrm{var}(p)$. Thus, we can assume without loss of generality that $i = 1$ and $j = 2$. In particular, $\deg_{x_2}(c_2) > 0$. From (18) we obtain

$$\deg_{x_2}(a_1 a_2) = \deg_{x_2}(c_2) > 0.$$

Assume without loss of generality that $\deg_{x_2}(a_1) > 0$. Since p is of individual degree either 2 or 0 with respect to every variable, $p \mid a_1 a_2$, and every variable in $\mathrm{var}(p)$ branches in $F_{(o)_L}$, it must be the case that there exists a non-constant irreducible factor of p which divides a_1, and another which divides a_2 (by the branching condition, these factors have individual degree at most 1 in every variable). Thus, $q_1 := \gcd(p, a_1) \neq 1$, $q_2 := \gcd(p_2, a_2) \neq 1$ and $p = q_1 q_2$.

Moreover, let g_1 be the factor of c_2 dividing a_1, and g_2 the one which divides a_2. Then we have the factorizations:

$$a_1 = q_1 g_1, \qquad a_2 = q_2 g_2.$$

This implies, together with the fact that a_1 and a_2 are $\mathbf{0}$-preserved and Lemma 2.32 that q_1, q_2, g_1, g_2 are all $\mathbf{0}$-preserved.

Let $S_q \subseteq \mathrm{var}(q_2)$ and $S_g \subseteq \mathrm{var}(g_1)$. Since $\mathrm{var}(p) \cap \mathrm{var}(c_2) = \emptyset$, it follows that $S_q \cap S_g = \emptyset$. Moreover, since q_2 and g_1 are $\mathbf{0}$-preserved, by Observation 2.50, there exist assignments $\boldsymbol{\beta} \in \mathbb{F}^{S_q}$ and $\boldsymbol{\gamma} \in \mathbb{F}^{S_g}$ such that $\boldsymbol{\beta}$ is a root of q_2 and $\boldsymbol{\gamma}$ is a root of g_1 and both with all coordinates nonzero. Thus, a_1, a_2, p all vanish under the restriction $S_q \leftarrow \boldsymbol{\beta}$, $S_g \leftarrow \boldsymbol{\gamma}$. By (18) we get that the same restriction also makes $b_1 b_2 + \zeta$ vanish. We thus obtain from (17)

$$\mathcal{P}_n\big|_{S_q \leftarrow \boldsymbol{\beta},\, S_g \leftarrow \boldsymbol{\gamma}} \equiv 0.$$

However, the left-hand side is nonzero, since both $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ have nonzero entries. This yields the required contradiction and concludes the proof of Claim 7.19. □

The proof so far gives the following more general corollary.

Corollary 7.33. *If some variable x_i branches in $(o)_L$ and it does not appear in $(o)_{R,L}$ (equivalently $(o)_{R,R}$) then $\mathrm{var}((o)_{R,L}) = [n] \setminus \{i\}$ ($\mathrm{var}((o)_{R,R}) = [n] \setminus \{i\}$). The same holds when we switch the roles of $(o)_L$ and $(o)_R$.*

We next show that unless Lemma 7.18 holds, x_1 is the only variable that branches in F.

Claim 7.21. *If Lemma 7.18 does not hold then no other variable except x_1 branches in F.*

Proof. Assume for a contradiction that Lemma 7.18 does not hold and there exists at least one more variable which branches in some child of o. Let it be x_2.

Claim 7.34. *It must be the case that x_2 branches in $(o)_R$.*

Proof. Assume for a contradiction that x_2 branches in $(o)_L$.

By the fact that no variable branches completely and the assumption that $\mathrm{var}(p) = [n] \setminus \{1\}$ we get that x_2 is read-twice in $(o)_{R,L}$.

If there exists one more variable which is read twice in $(o)_{R,L}$, then we would get a contradiction by Corollary 7.33, switching the roles of $(o)_{R,L}$ and $(o)_{R,R}$. Therefore, assume this does not happen, i.e. every variable in $[n] \setminus \{1, 2\}$ branches in $(o)_R$.

Since no variable branches completely, every variable in $[n] \setminus \{1, 2\}$ is read twice in some child of $(o)_L$. Since $n \geqslant 5$, there are at least 3 such variables. Therefore, one of the children of $(o)_L$ contains two variables which are read-twice in it and branch in $(o)_R$. Repeating the argument above, we get a contradiction. □

By the preceding claim we conclude that unless Lemma 7.18 holds, no other variable branches in either child of o.

Since no variable branches in both children, we may assume without loss of generality that x_2 is read-twice in $(o)_{L,L}$. Moreover, if $\mathrm{var}(p_{(o)_{L,R}}) \neq [n] \setminus \{2\}$ then by a claim analogous to Claim 7.19 we obtain that Lemma 7.18 holds.

Therefore we may further assume $\mathrm{var}(p_{(o)_{L,R}}) = [n] \setminus \{2\}$. Since no other variable branches in either child, we conclude that x_3 is read-twice in $(o)_{L,R}$.

Recall that by (18), $pc_2 = -a_1a_2$. As $x_2, x_3 \in \mathrm{var}(p)$, but they belong to different children of $(o)_L$, they appear in two distinct irreducible factors of p. Let these factors be p_1, p_2 and assume that $x_2 \in p_1$, $x_3 \in p_2$, $x_2 \notin p_2$ and $x_3 \notin p_1$. Consequently, $p_1 \mid a_1$ and $p_2 \mid a_2$. Let $g_1, g_2 \in \mathbb{F}[\mathbf{x}]$ be such that $p_1g_1 = a_1$ and $p_2g_2 = a_2$.

Denote $S_1 \coloneqq \mathrm{var}(p_1) \setminus \{x_2\}$, $S_2 \coloneqq \mathrm{var}(p_2) \setminus \{x_3\}$ and $S = S_1 \cup S_2$. Let $\boldsymbol{\tau} \in \mathbb{F}^S$ be a generic assignment to the variables in S. Let $\delta_1, \delta_2 \in \mathbb{F}$ be roots to the univariate polynomials $p_1|_{S \leftarrow \boldsymbol{\tau}}, p_2|_{S \leftarrow \boldsymbol{\tau}}$ respectively. Since p_1, p_2 are factors of p they are $\mathbf{0}$-preserved by Lemma 2.32, therefore $\delta_1, \delta_2 \neq 0$. We get the following contradiction:

$$\begin{aligned} 0 \not\equiv \alpha \mathcal{P}_{[n]\setminus(\mathrm{var}(p_1)\cup\mathrm{var}(p_2))} &= \left(p_{(o)_L} + p \cdot p_{(o)_{R,R}} + \zeta\right)|_{S \leftarrow \boldsymbol{\tau}, x_2=\delta_1, x_3=\delta_2} \\ &= \left(-pc_2x_1^2 + (a_1b_2 + a_2b_1)x_1 + b_1b_2 + \zeta\right)|_{S \leftarrow \boldsymbol{\tau}, x_2=\delta_1, x_3=\delta_2} \\ &= \left(-pc_2x_1^2 + (p_1g_1b_2 + p_2g_2b_1)x_1 - pc_0\right)|_{S \leftarrow \boldsymbol{\tau}, x_2=\delta_1, x_3=\delta_2} \equiv 0\,. \end{aligned}$$

Hence no variable besides x_1 branches in F. This concludes the proof of Claim 7.21, □

The previous observation, together with the fact that $\mathrm{var}(p) = [n] \setminus \{1\}$, implies that $\mathrm{var}(p_{(o)_{R,R}}) = \{1\}$. Now, for each $i \in \{L, R\}$, define

$$S_i \coloneqq \mathrm{var}(p_{o_{L,i}}) \setminus \{1\}.$$

Assume without loss of generality that $|S_L| \geqslant |S_R|$. Since $n \geqslant 5$, it follows that $|S_L| \geqslant 2$.

Let $\boldsymbol{\tau} \in \mathbb{F}^{S_R}$ be a generic assignment to the variables in S_R, and let $\delta \in \mathbb{F}$ be a root of the univariate polynomial $p_{(o)_{L,R}}|_{S_R \leftarrow \boldsymbol{\tau}}$. Then, for some nonzero $\alpha \in \mathbb{F} \setminus \{0\}$ and some $\gamma \in \mathbb{F}$, we obtain:

$$\begin{aligned} \alpha \cdot \mathcal{P}_{[n]\setminus\mathrm{var}(p_{(o)_{L,R}})} &= \left(p_{(o)_L} + p \cdot p_{(o)_{R,R}} + \zeta\right)\Big|_{S_R \leftarrow \boldsymbol{\tau},\, x_1=\delta} \\ &= \left(p_{(o)_{L,L}} \cdot p_{(o)_{L,R}}\right)\Big|_{S_R \leftarrow \boldsymbol{\tau},\, x_1=\delta} + p|_{S_R \leftarrow \boldsymbol{\tau}} \cdot p_{(o)_{R,R}}|_{x_1=\delta} + \zeta \\ &= \gamma \cdot p|_{S_R \leftarrow \boldsymbol{\tau}} + \zeta. \end{aligned}$$

However, the left-hand side is not $\mathbf{0}$-preserved, while the right-hand side is, in contradiction. This concludes the proof of Lemma 7.18. □

7.3.2 Proof of Lemma 7.22

In this subsection we give the missing details from the sketch given earlier. For convenience we repeat the statement of the relevant claims.

Lemma 7.22. *Let $F \in \sum^2 \mathsf{R2F}$ be a totally non-structural formula. Suppose that both $(o)_L$ and $(o)_R$ are multiplication gates. Let S_1 be the set of completely branched variables in F, and assume that $1 \leqslant |S_1| \leqslant 4$. Furthermore, assume that no nonzero polynomial in $\mathcal{C} \cup \mathcal{E}^3$ vanishes at $\mathbf{0}$.*

Then, if $n \geqslant 23$, we have

$$\mathcal{P}_n \nmid p_F.$$

Proof. We recall the process defined in the sketch.

Denote $F_1 = F$ and let $\boldsymbol{\tau}$ be a generic assignment to all the variables. Consider the following process: Let S_i be the set of completely branched variables of F_i. At the ith step we substitute $\boldsymbol{\tau}_{S_i}$ to S_i, and let $F_{i+1} \coloneqq F_i\big|_{S_i \leftarrow \boldsymbol{\tau}_{S_i}}$. We terminate the process if one of the following conditions hold, where o_i is the root of F_i:

1. either $(o_i)_L^{\mathrm{op}} = +$, or $(o_i)_R^{\mathrm{op}} = +$,
2. $\left|\bigcup_{j=1}^{i} S_j\right| \geqslant 18.$, or
3. $S_i = \emptyset$, i.e., no variable branches completely in F_i.

We now prove that if the process stopped at the mth formula, F_m, then for $V_m = \mathrm{var}(F_m)$, it holds that $\mathcal{P}_{V_m} \nmid p_{F_m}$. We show it by analyzing the different stopping conditions. Denote $S \coloneqq \bigcup_{j\in[m]} S_j$.

We already gave a full analysis of the claim in case we stopped due to Cases 1 or 3, so we proceed directly to the analysis of Case 2. For readability, we restate relevant definitions and claims that were given in the sketch, and we also repeat some arguments to improve the flow.

From properties of generic assignments (see Observation 7.11), we know that for every $j \in [m]$, p_{F_j} is totally non-structural. Consequently, for every $t \in V_j$, $\deg_{x_t}(p_{(o_j)_R}) = 2$.

Definition 7.23. Let $\rho \in \{L, R\}$ and let $t \in [m]$. We denote with $\ell_{t,\rho,l}$ and $\ell_{t,\rho,r}$ the two leaves labeled by x_t in o_ρ. Let $v_{t,\rho} = \mathrm{fcg}(\ell_{t,\rho,l}, \ell_{t,\rho,r})$. Similarly, we denote the left and right leaves in $(o_j)_\rho$ by $\ell_{t,j,\rho,l}$ and $\ell_{t,j,\rho,r}$, respectively and set

$$v_{t,j,L} = \mathrm{fcg}_{(o_j)_L}(\ell_{t,j,L,l}, \ell_{t,j,L,r}) \quad \text{and} \quad v_{t,j,R} = \mathrm{fcg}_{(o_j)_R}(\ell_{t,R,l}, \ell_{t,R,r})\,.$$

the fcg of the two leaves labeled by x_t in $(o_j)_L$ and $(o_j)_R$, respectively. ◊

We note that by simple properties of generic assignments, it holds that

$$\mathcal{L}^F_{[j-1]}(v_{t,j,\rho}) = v_{t,\rho} \quad \text{and} \quad \mathcal{L}^F_{[j-1]}(\ell_{t,j,\rho,l}) = \ell_{t,\rho,l}\,. \tag{20}$$

Thus, for brevity of presentation, we omit the index j from both $v_{t,j,\rho}$ and $\ell_{t,j,\rho,l}$.

Assume for a contradiction that $\mathcal{P}_n \mid p_F$. Corollary 7.3 implies that we actually have $\mathcal{P}_n \sim p_F$.

Claim 7.24. *$|S_j| = 1$ for every $j \in [m]$, and in particular $m = 18$. Furthermore, in each such F_j one of the grandchildren of the root is supported on S_j.*

Proof. We first establish this for S_1. Without loss of generality, assume that $|\mathrm{var}((o)_{L,L})|$ is minimal among the variable set sizes of the four grandchildren of the root. We claim that $\mathrm{var}((o)_{L,L}) = S_1$.

Since S_1 is the set of completely branched variables, the variables in S_1 appear in all four grandchildren of the root of F_1. Suppose, for contradiction, that there exists $x_t \in \mathrm{var}((o)_{L,L}) \setminus S_1$. Then $p_{(o)_{L,L}}|_{S_1 \leftarrow \boldsymbol{\tau}_{S_1}}$ is non-constant. Because $|\mathrm{var}((o)_{L,L})|$ was assumed minimal, the same argument applies to every other gate at the second level of F, implying that each of these gates remains non-constant. Hence, no new completely branched variable appears in F_1, contradicting the fact that we did not reach Case 2 in step $m-1$. We have thus established that $\mathrm{var}((o)_{L,L}) = S_1$.

Next, let $a_L(\boldsymbol{x}), b_L(\boldsymbol{x})$ and $a_R(\boldsymbol{x}), b_R(\boldsymbol{x})$ be such that

$$p_{(o)_{R,L}} = a_L\, p_{(o)_{L,L}} + b_L, \qquad p_{(o)_{R,R}} = a_R\, p_{(o)_{L,L}} + b_R. \tag{21}$$

Let α be such that

$$p_F = p_{(o)_L} + p_{(o)_R} + \alpha. \tag{22}$$

We then have

$$\begin{aligned} p_F &= p_{(o)_{L,L}} p_{(o)_{L,R}} + \left(a_L\, p_{(o)_{L,L}} + b_L\right)\left(a_R\, p_{(o)_{L,L}} + b_R\right) + \alpha \\ &= p_{(o)_{L,L}}\left(p_{(o)_{L,R}} + a_L a_R\, p_{(o)_{L,L}} + (a_L b_R + a_R b_L)\right) + b_L b_R + \alpha. \end{aligned} \tag{23}$$

From the property that polynomials in $\mathcal{E}^3$ do not vanish on $\mathbf{0}$, and from Lemma 2.33, we deduce that $p_{(o)_{L,L}}$ is $\mathbf{0}$-preserved. Therefore, by Observation 2.50, and recalling that $\mathrm{var}((o)_{L,L}) = S_1$, there exists a zero $\boldsymbol{\beta} \in \mathbb{F}^{S_1}$ of $p_{(o)_{L,L}}$ in which every coordinate is nonzero. Substituting $\boldsymbol{\beta}$, we obtain

$$\mathcal{P}_n|_{S_1 \leftarrow \boldsymbol{\beta}} \sim p_F|_{S_1 \leftarrow \boldsymbol{\beta}} = b_L b_R + \alpha. \tag{24}$$

From this and (23), we conclude that for some $\delta \in \mathbb{F} \setminus \{0\}$,

$$\mathcal{P}_{[n]\setminus S_1}(\mathcal{P}_{S_1} + \delta) \sim p_F - b_L b_R - \alpha = p_{(o)_{L,L}}\left(p_{(o)_{L,R}} + a_L a_R\, p_{(o)_{L,L}} + (a_L b_R + a_R b_L)\right).$$

Since $\mathrm{var}(p_{(o)_{L,L}}) = S_1$, it follows that $p_{(o)_{L,L}} \sim \mathcal{P}_{S_1} + \delta$. Because $p_{(o)_{L,L}}$ is $\mathbf{0}$-preserved, we must have $|S_1| = 1$. Consequently, $p_{(o)_{L,L}}$ is an affine linear function in the variable in S_1.

The proof for $1 < j \leqslant m$ is completely analogous. □

From now on we assume without loss of generality that $S_j = \{x_j\}$ for every $j \in [m]$.

To obtain an even more defined structure, note that by the pigeonhole principle, in at least half of the rounds of the process, the grandchild of the root that computes a univariate polynomial at that round is a child of $(o_i)_L$. Without loss of generality (by rotating the children of $(o_i)_R$ if necessary), we may assume that this univariate polynomial is computed by $(o_i)_{L,L}$. Denote this set of rounds by $\mathcal{I}$.

We now focus only on these rounds and apply the generic assignment to all variables that were constructed in the other rounds, i.e., to the variables x_j such that $j \in S \setminus \mathcal{I}$. By the properties of generic assignments, we observe that if we were to rerun the process from the beginning, the first $|\mathcal{I}|$ variables collected would be exactly those with indices in $\mathcal{I}$.

Thus, from now on we assume, without loss of generality, that $|S| \geqslant m/2 = 18/2 = 9$ and that, at the ith step, the polynomial computed at $(o_i)_{L,L}$ is univariate in x_i. Denote $m' = m/2 = 9$ and $n' = n - m/2 = 23 - 18/2 = 14$.

Corollary 7.25. *For every $j \in [m']$ it holds that var $\left((o_j)_{L,R}\right) = [j : n']$.*

Let $\delta_j \in \mathbb{F} \setminus \{0\}$ be such that $p_{(o_j)_{L,L}} \sim (x_j - \delta_j)$. Clearly, we can assume w.l.o.g., that

$$p_{(o_j)_{L,L}} = (x_j - \delta_j)\,. \tag{25}$$

As before, we can "push" the additive constants in $(o_j)_L$ and $(o_j)_R$ to o_j and assume both $(o_j)_L$ and $(o_j)_R$ are pure multiplication gates. Thus, for $t \in V_j \setminus \{j\}$ we have

$$p_{(o_j)_L}|_{x_j = \delta_j} = 0\,, \tag{26}$$

$$(x_j - \delta_j) \mid \partial_{x_t} F_{(o_j)_L}\,. \tag{27}$$

Let α_j be such that

$$p_{F_j} = p_{(o_j)_L} + p_{(o)_R} + \alpha_j. \tag{28}$$

For $j \in [m']$, we compute the quotient and remainder of dividing $p_{(o_j)_{R,L}}$ and $p_{(o_j)_{R,R}}$ by $(x_j - \delta_j)$:

$$p_{(o_j)_{R,L}} = a_{j,L}(x_j - \delta_j) + b_{j,L} \qquad \text{and} \qquad p_{(o_j)_{R,R}} = a_{j,R}(x_j - \delta_j) + b_{j,R}. \tag{29}$$

We have,

$$\begin{aligned} p_{F_j} &= (x_j - \delta_j)p_{(o_j)_{L,R}} + \left(a_{j,L}(x_j - \delta_j) + b_{j,L}\right)\left(a_{j,R}(x_j - \delta_j) + b_{j,R}\right) + \alpha_j \\ &= (x_j - \delta_j)\Big(p_{(o_j)_{L,R}} + a_{j,L}a_{j,R}(x_j - \delta_j) + (a_{j,L}b_{j,R} + a_{j,R}b_{j,L})\Big) + b_{j,L}b_{j,R} + \alpha_j. \end{aligned} \tag{30}$$

Hence,

$$\mathcal{P}_{\{j,\ldots,n\}}|_{x_j=\delta_j} \sim p_{F_j}|_{x_j=\delta_j} = b_{j,L}b_{j,R} + \alpha_j. \tag{31}$$

We next prove that for all $j \in [m'-2]$, $\alpha_j \neq 0$. Moreover, upon restricting F to the variables indexed by $[m'-4]$, $(o)_R$ becomes a product of two ROFs. We first prove the following claims.

The next claim shows that the only variables that can appear between $v_{i,\rho}$ and o_ρ are in $[i-1]$.

Claim 7.35. *Let $i \in [m']$, $\rho \in \{L, R\}$ and $w \in Unv_{F_{o_\rho}}(v_{i,\rho})$. Then var$(w) \subseteq [i-1]$.*

Proof. Assume towards contradiction this is not the case and let $w \in \mathrm{Unv}_{F_{o_\rho}}(v_{i,\rho})$ with $\mathrm{var}(w) \nsubseteq [i-1]$. Let $i-1 < t \in \mathrm{var}(w)$. $t \neq i$ since by definition $v_{i,\rho}$ is the fcg of all of the leaves labeled by x_i in F_{o_ρ}. Since $\boldsymbol{\tau}$ is generic, we get that $w|_{[i-1]\leftarrow\boldsymbol{\tau}_{[i-1]}} \notin \mathbb{F}$. This and the fact that it is a sibling to one of the gates in the path from $v_{i,\rho}|_{[i-1]\leftarrow\boldsymbol{\tau}_{[i-1]}}$ to $(o_i)_\rho$ contradicts the fact that $x_i \in S_i$. □

By our assumption and Corollary 7.25, the child of $v_{i,L}$ that does not restrict to a univariate, $(v_{i,L})_R$, must contain all the variables x_j such that $j \notin [i]$. We next argue that its sibling, $(v_{i,L})_L$, that restricts to a univariate, sees only the variables in $[i]$. Furthermore, any two such gates $(v_{i,L})_L$ and $(v_{j,L})_L$ are disjoint (i.e., one is not a descendant of the other).

Claim 7.36. *Let $i < j \in [m']$. Then:*

1. *var$((v_{i,L})_L) \subseteq [i]$ and var$\left((v_{j,L})_L\right) \subseteq [j]$.*
2. *$(v_{i,L})_L$ and $(v_{i,L})_L$ are disjoint.*

Proof. Since $\boldsymbol{\tau}$ is a generic assignment, $x_n \in \mathrm{var}\left((v_{i,L})_R|_{[i-1]\leftarrow\boldsymbol{\tau}_{[i-1]}}\right)$ hence it must be the case that $u|_{[i-1]\leftarrow\boldsymbol{\tau}_{[i-1]}}$ is univariate in x_i. Since $\boldsymbol{\tau}$ is a generic assignment, $(v_{i,L})_L|_{[i-1]\leftarrow\boldsymbol{\tau}_{[i-1]}}$ depends on any variable in $\mathrm{var}((v_{i,L})_L) \setminus [i-1]$. Therefore, $\mathrm{var}((v_{i,L})_L) \subseteq [i]$, which implies item 1. The argument for j is similar.

Assume towards contradiction that $(v_{i,L})_L$ and $(v_{j,L})_L$ are not disjoint. Then, by item 1 we have $\mathrm{var}((v_{i,L})_L) \subseteq [i]$. Since $j \in \mathrm{var}\left((v_{j,L})_L\right)$ and $j > i$ it cannot be the case that $(v_{j,L})_L$ is a descendant of $(v_{i,L})_L$, hence $(v_{i,L})_L$ is a descendant of $(v_{j,L})_L$. But this contradicts Claim 7.35. □

The following claim implies that when restricting to the variables in $[18]$, o_R is a product of two ROF.

Claim 7.26. *Under our contradiction assumption it holds that $s := \min(|var((o)_{R,L})|, |var((o)_{R,R})|) \geqslant m' - 2$.*

Proof. Assume for a contradiction that $s < m' - 2$. Further, assume w.l.o.g. that $|\mathrm{var}((o)_{R,L})| \leqslant |\mathrm{var}((o)_{R,R})|$. Hence, since in the ith step $S_i = \{x_i\}$, we have $\mathrm{var}((o)_{R,L}) = [s]$.

By Claim 7.36 we get that $\{F_{(v_{i,L})_L}\}_{i \leqslant m'}$ are disjoint subformulas of $F_{(o)_L}$. This and the fact that $F_{(o)_L} \in \mathsf{R2F}$ imply that there exist $t \in \mathrm{var}((o)_{R,L})$ and $m' - 2 \leqslant i \leqslant m'$ such that $t \notin (v_{i,L})_L$. Without loss of generality, assume $i = m' - 2$. Denote $B \coloneqq [m' - 3] \setminus \{x_t\}$.

$$\begin{aligned}\partial_{x_{n'}} p_F|_{B \leftarrow \boldsymbol{\tau}_B} &= \partial_{x_{n'}}\left(p_{(o)_L} + p_{(o)_R} + \alpha\right)\Big|_{B \leftarrow \boldsymbol{\tau}_B} \\ &= \left(\partial_{x_{n'}}\left(p_{v_{n',L}}\right) \prod_{u \in \mathrm{Unv}_{F_{(o)_L}}(v_{n',L})} p_u + \partial_{x_{n'}}\left(p_{v_{n',R}}\right) \prod_{u \in \mathrm{Unv}_{F_{(o)_R}}(v_{n',R})} p_u\right)\Bigg|_{B \leftarrow \boldsymbol{\tau}_B}. \end{aligned} \tag{32}$$

Recall that $(o_{m'-2})_{L,L}|_{B \leftarrow \boldsymbol{\tau}_B}$ computes the polynomial $(x_t - \delta_t)$ and that $t \notin \mathrm{var}((o_{m'-2})_{L,L})$. These imply that

$$p_{(o_{m'-2})_{L,L}|_{B \leftarrow \boldsymbol{\tau}_B}} = p_{(o_{m'-2})_{L,L}} = (x_t - \delta_t).$$

Since $t \in \mathrm{var}((o)_{R,L})$ we have that $p_{(o)_{R,L}}|_{B \leftarrow \boldsymbol{\tau}_B}$ is a univariate polynomial in t. As every variable in $[s]$ branches in o_R, $p_{(o)_{R,L}}|_{B \leftarrow \boldsymbol{\tau}_B}$ is a univariate linear polynomial in x_t. Let $\beta_t \in \mathbb{F}$ such that $p_{(o)_{R,L}}|_{B \leftarrow \boldsymbol{\tau}_B} \sim (x_t - \beta_t)$.

Since $n' > s$ we get that $v_{n',L}$ is a subgate of $(o_{m'-2})_{L,R}$ and that $v_{n',R}$ is a subgate of $(o)_{R,R}$.

By the previous arguments, and using Lemma 2.8 we get the following equations.

$$\begin{aligned}(x_{m'-2} - \delta_{m'-2}) \;\Big|\; &\left(\partial_{x_{n'}}\left(p_{v_{n',L}}\right) \prod_{u \in \mathrm{Unv}_{F_{(o)_L}}(v_{n',L})} p_u\right)\Bigg|_{B \leftarrow \boldsymbol{\tau}_B}, \\ (x_t - \beta_t) \;\Big|\; &\left(\partial_{x_{n'}}\left(p_{v_{n',R}}\right) \prod_{u \in \mathrm{Unv}_{F_{(o)_R}}(v_{n',R})} p_u\right)\Bigg|_{B \leftarrow \boldsymbol{\tau}_B}. \end{aligned} \tag{33}$$

Note that $\beta_t \neq 0$ since $p_{(o)_{R,L}}$ is $\mathbf{0}$-preserved and $\boldsymbol{\tau}$ is a generic assignment. From (32) and (33), we obtain the following contradiction

$$\begin{aligned}0 \not\equiv \mathcal{P}_{[n']\setminus(B \cup \{x_t, x_{m'-2}, x_{n'}\})} &\sim \partial_{x_{n'}} \mathcal{P}_{n'}\Big|_{B \leftarrow \boldsymbol{\tau}_B,\, x_t = \beta_t,\, x_{m'-2} = \delta_{m'-2}} \\ &\sim \partial_{x_{n'}} p_F\Big|_{B \leftarrow \boldsymbol{\tau}_B,\, x_t = \beta_t,\, x_{m'-2} = \delta_{m'-2}} \\ &= \left(\partial_{x_{n'}}(p_{v_{n',L}}) \prod_{u \in \mathrm{Unv}_{F_{(o)_L}}(v_{n',L})} p_u + \partial_{x_{n'}}(p_{v_{n',R}}) \prod_{u \in \mathrm{Unv}_{F_{(o)_R}}(v_{n',R})} p_u\right)\Big|_{B \leftarrow \boldsymbol{\tau}_B,\, x_t = \beta_t,\, x_{m'-2} = \delta_{m'-2}} \equiv 0. \square \end{aligned}$$

We remind

Corollary 7.27. *The restriction of* F *to the variables indexed by* $[m' - 2]$ *results in a formula* $F' = F_{[m'-2]}$ *such that* $(o_{F'})_R$ *is a product of two* ROF*s.*

The proof of this claim was given in the proof sketch earlier.

From now on, we consider $F' = F|_{[m'-2]}$. Observe that, had we run our process on F' instead of F, we would obtain the same set of branched variables (excluding those that were restricted).

Furthermore, by the properties of generic assignments, F' satisfies the same assumptions as F in the statement of the claim. For readability, we continue to use the notation F instead of F', keeping in mind that $\mathrm{var}(F) = [m' - 2]$ and that by Corollary 7.27, $(o_F)_R$ is a product of two ROFs.

We next prove

Claim 7.28. *For every* $j \in [m' - 5]$, $\alpha_j \neq 0$.

Proof. Assume for a contradiction that $\alpha_j = 0$. By Corollary 7.27, when we restrict the formula to the variables indexed by $[m' - 2]$, $(o_j)_R$ factors to a product of two ROFs. By the assumption in the statement of Lemma 7.22 and Claim 7.16, it follows that each of its irreducible factors is also **0**-irreducible.

From (31) we get that if $\alpha_j = 0$ then $\mathcal{P}_{[j+1:m'-2]} \sim b_{j,L} \cdot b_{j,R}$. As $\deg(\mathcal{P}_{[j+1:m'-2]}) \geqslant 3$, we assume w.l.o.g. $\deg(b_{j,R}) \geqslant 2$, so it is a non-trivial multilinear monomial not involving x_j. We next prove that $p_{(o_j)_{R,R}}$ is irreducible.

Indeed, as $p_{(o_j)_{R,R}}$ is linear in x_j, it is reducible if and only if it is divisible by some factor of $a_{j,R}$. Thus, if $p_{(o_j)_{R,R}} = a'((x_j - \delta_j)a'' + b')$. Then, $a'b' = b_{j,R}$, which means that a' is a monomial. By **0**-preserveness we must have $a' \in \mathbb{F}$.

Since $b_{1,R}$ is a nonrtivial monomial, fixing any variable $x_i \in b_{1,R}$ to zero makes $p_{(o)_{R,R}}|_{x_i=0} = (a_{1,R}(x_1 - \delta_1) + b_{1,R})_{x_i=0} = a_{1,R}(x_1 - \delta_1)$. By **0**-irreducibility we conclude that $a_{1,R} \in \mathbb{F}$. However, in this case too we get a contradiction to **0**-preserveness as $\deg(b_{1,R}) \geqslant 2$, so setting any variable in it to **0** makes all the other variables (which are all different from x_1) disappear as well. □

Corollary 7.29. *For every* $j \in [m' - 7]$, $b_{j,R} \in \mathbb{F}$ *or* $b_{j,L} \in \mathbb{F}$.

Proof. By (31), for some $\beta \in \mathbb{F} \setminus \{0\}$, we have $b_{j,R}b_{j,L} = \beta\mathcal{P}_{[j+1:m'-5]} - \alpha_j$. Since $\beta\alpha_j \neq 0$ and $m' - 5 > j + 1$, it follows that $\beta\mathcal{P}_{[j+1:m'-5]} - \alpha_j$ is an irreducible polynomial, and hence, one among $b_{j,L}, b_{j,R}$ must be a scalar. □

Claim 7.30. *One of the following two symmetrical cases hold:*

$$(x_1 - \delta_1) \mid \prod_{u \in Unv(F_{v_{2,R}})_L(\ell_{t,R,l})} p_u \quad \text{and} \quad b_{2,L} \in \mathbb{F}, \text{ or}$$

$$(x_1 - \delta_1) \mid \prod_{u \in Unv(F_{v_{2,R}})_R(\ell_{2,R,r})} p_u \quad \text{and} \quad b_{2,R} \in \mathbb{F}.$$

Proof. Consider the first and second order derivatives according to x_2. From Lemma 2.8 the definition of $v_{2,R}$ (recall Definition 7.23) and (26) we obtain:

$$0 \not\equiv \partial_{x_2}\mathcal{P}_{[2:m'-2]} \sim \partial_{x_2}p_F|_{x_1=\delta_1} = \partial_{x_2}\left(p_{(o)_L} + p_{(o)_R+\alpha_1}\right)|_{x_1=\delta_1} = \partial_{x_2}\left(\gamma_1 + p_{(o)_R}|_{x_1=\delta_1}\right)$$

$$= \partial_{x_2}p_{(o)_R}|_{x_1=\delta_1} = \left(\prod_{u \in \mathrm{Unv}_{F_{(o)_R}}(v_{2,R})} p_u\right)|_{x_1=\delta_1} \cdot (\partial_{x_2}p_{v_{2,R}})|_{x_1=\delta_1}, \tag{34}$$

and

$$0 \equiv \partial_{x_2^2}\mathcal{P}_{[2:m'-2]} \sim \partial_{x_2^2} p_{(o_j)_R}|_{x_1=\delta_1} = \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(o)_R}}(v_{2,R})} p_u\right)|_{x_1=\delta_1} \cdot (\partial_{x_2^2} p_{v_{2,R}})|_{x_1=\delta_1} \tag{35}$$

$$= \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(o)_R}}(v_{2,R})} p_u\right)|_{x_1=\delta_1} \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(v_{2,R})_L}}(\ell_{2,R,l})} p_u\right)|_{x_1=\delta_1} \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(v_{2,R})_R}}(\ell_{2,R,r})} p_u\right)|_{x_1=\delta_1}.$$

Combining (34) and (35) we conclude that

$$(x_1-\delta_1) \mid \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(v_{2,R})_L}}(\ell_{2,R,l})} p_u\right)\left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(v_{2,R})_R}}(\ell_{2,R,r})} p_u\right)$$

as claimed. Assume now that $(x_1-\delta_1) \mid \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(v_{2,R})_L}}(\ell_{2,R,l})} p_u\right)$. By Lemma 2.8 this implies that

$$\partial_{x_2}\left(p_{(o)_{R,L}}\big|_{x_1=\delta_1}\right) \equiv 0\,.$$

We now prove that $b_{2,L} \in \mathbb{F}$. As $m'-7=2$, Corollary 7.29 shows that one of $b_{2,R} \in \mathbb{F}$ and $b_{2,L} \in \mathbb{F}$ is a scalar, let us assume for a contradiction that $b_{1,R} \in \mathbb{F}$. By (29) this implies that

$$\partial_{x_2}\left(p_{(o)_{R,R}}\big|_{x_1=\delta_j}\right) \equiv \partial_{x_2} b_{1R} \equiv 0.$$

Considering (26), (28), and the equalities above we obtain the following contradiction

$$\begin{aligned} 0 \not\equiv &\partial_{x_2}\mathcal{P}_{[m'-2]}\big|_{x_1=\delta_1} \sim \partial_{x_2}(p_F)\Big|_{x_1=\delta_1} = \partial_{x_2}\left(p_{(o)_L}|_{x_1=\delta_1} + \left(p_{(o)_{R,L}} p_{(o)_{R,R}}\right)|_{x_1=\delta_1} + \alpha_1\right) \\ =&\partial_{x_2}\left(p_{(o)_{R,L}}\big|_{x_1=\delta_1}\right)\cdot p_{(o)_{R,R}}\big|_{x_1=\delta_1} + p_{(o)_{R,L}}\big|_{x_1=\delta_1}\cdot \partial_{x_{1+1}}\left(p_{(o)_{R,R}}\big|_{x_1=\delta_1}\right) \equiv 0\,. \end{aligned}$$ □

The case $(x_1-\delta_1) \mid \left(\prod_{u\in \mathrm{Unv}_{\mathsf{F}_{(v_{2,R})_R}}(\ell_{2,R,r})} p_u\right)$ is analogous.

The argument given in the proof sketch concludes the proof of Lemma 7.22. □

8 PIT for R4F

8.1 Blackbox

In this section, we prove our main blackbox result, namely that for some seed length $k \in O(\log n)$, the generator $\mathcal{G}_k$ hits the class of R4Fs. As before, Figure 7 indicates what we have proved so far, and describes the case analysis handled in this section.

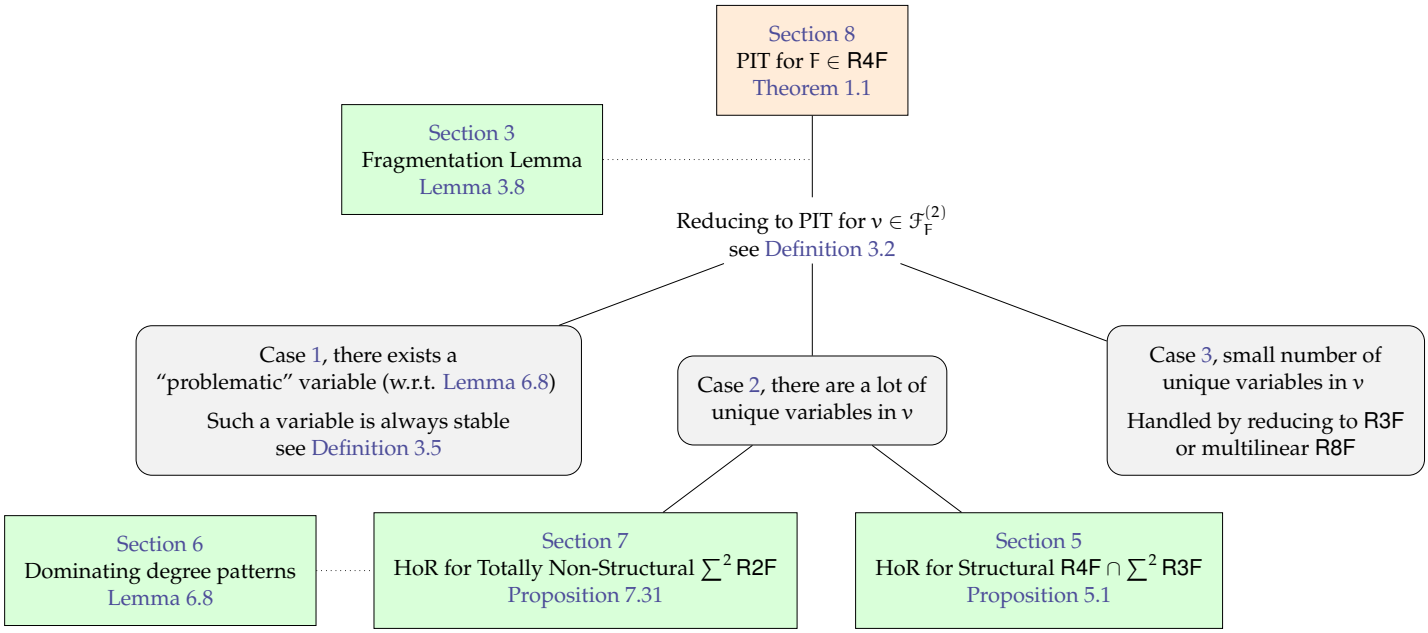

Figure 7: Our current position in the proof of Theorem 1.1

Since in the totally non-structural case taking partial derivatives does not preserve the structural form, we first establish the following lemma.

Lemma 8.1. *Let $F \in \sum^2$ R2F be totally non-structural. Suppose that no nonzero polynomial in $\mathcal{C} \cup \mathcal{E}^3$ vanishes at $\mathbf{0}$. Then, if $n \geqslant 23$, there exists $x_i \in \mathbf{x}$ such that $\mathcal{P}_{[n]\setminus\{-i\}} \nmid \partial_{x_i} p_F$.*

Proof. We first note that by Proposition 7.31, for every $\alpha, \beta \in \mathbb{F}$, $p_F \neq \alpha \mathcal{P}_n + \beta$. Hence, there exists $M \in \mathrm{mon}(F)$ such that $\mathrm{Supp}(M) \notin \{\emptyset, [n]\}$. Any $i \in \mathrm{Supp}(M)$ then satisfies the desired property. □

Theorem 1.1 (Main blackbox result). *Let $p_F(\mathbf{x})$ be an n-variate polynomial computable by an* R4F *F, over a field $\mathbb{F}$. There exists an absolute constant $c_{1.1}$, such that $p_F(\mathbf{x}) \not\equiv 0$ if and only if*

$$p_F \circ \mathcal{G}_{9\log n + c_{1.1}} \not\equiv 0\,.$$

Proof. Let $\nu \in \mathcal{F}_F^{(2)}$ (recall Definition 3.2).

Claim 8.2. *Let $\mathcal{H}$ be a hitting-set generator for* R3F*s. If $\nu^{op} = \times$, then for every $x_t \in U(\nu)$,*

$$\partial_{x_t^{\deg}} p_\nu \circ \mathcal{H} \not\equiv 0\,.$$

Proof. Let $x_t \in U(\nu)$. Then

$$\partial_{x_t^{\deg}} p_\nu = \left(\partial_{x_t^{\deg}} p_{(\nu)_L}\right)\left(\partial_{x_t^{\deg}} p_{(\nu)_R}\right) \in \prod^2 \text{R3F}\,.$$

The claim follows directly from the definition of $\mathcal{H}$. □

Assume from now on that $\nu^{op} = +$. Denote by $S_{(\nu,s)}$ and $S_{(\nu,ns)}$ the sets of structural and non-structural variables in ν, respectively. Observe that, since we consider read-4 formulas, every $x_i \in S_{(\nu,ns)}$ belongs to $U(\nu)$. We divide the analysis into three cases.

1. **There exists $x_t \in \mathbf{var}(v)$ such that, for some $G \in \{F_{(v)_L}, F_{(v)_R}\}$, we have $\mathbf{Read}_G(x_t) = 3$ and $\deg_{x_t}(G) > 1$.**

   Claim 8.3. $x_t \in U^{(s)}(v)$.

   Proof. Let u be the first common gate of all leaves labeled by x_t in F. Assume, without loss of generality, that $v \in (u)_L$. By Lemma 2.8,

   $$\partial_{x_t^{\deg}} p_F = \left(\prod_{w \in \mathrm{Unv}_F(u)} p_w \right) \partial_{x_t^{\deg}} p_u \,. \tag{36}$$

   Observe that if $u^{op} = +$, then by the choice of v and x_t, we have $\deg_{x_t}(p_F) = \deg_{x_t}(p_u) = \deg_{x_t}(p_v) > \deg_{x_t}(p_{(u)_R})$. In this case,

   $$\partial_{x_t^{\deg}} p_u = \partial_{x_t^{\deg}} p_{(u)_L} + \partial_{x_t^{\deg}} p_{(u)_R} = \partial_{x_t^{\deg}} p_{(u)_L} = \left(\prod_{w \in \mathrm{Unv}_{F_{(u)_L}}(v)} p_w \right) \partial_{x_t^{\deg}} p_v \,. \tag{37}$$

   On the other hand, if $u^{op} = \times$, then $\deg_{x_t}(p_F) = \deg_{x_t}(p_u) = \deg_{x_t}(p_v) + \deg_{x_t}(p_{(u)_R}) = \deg_{x_t}(p_v) + 1$. Let ℓ denote the leaf labeled by x_t in $(u)_R$. Then

   $$\begin{aligned} \partial_{x_t^{\deg}} p_u &= \left(\partial_{x_t^{\deg}} p_{(u)_L}\right)\left(\partial_{x_t^{\deg}} p_{(u)_R}\right) \\ &= \left(\partial_{x_t^{\deg}} p_v \prod_{w \in \mathrm{Unv}_{F_{(u)_L}}(v)} p_w \right) \left(\partial_{x_t} p_\ell \prod_{w \in \mathrm{Unv}_{F_{(u)_R}}(\ell)} p_w \right) \\ &\sim \partial_{x_t^{\deg}} p_v \left(\prod_{w \in \mathrm{Unv}_{F_{(u)_L}}(v)} p_w \right) \left(\prod_{w \in \mathrm{Unv}_{F_{(u)_R}}(\ell)} p_w \right) , \end{aligned} \tag{38}$$

   where we used the fact that $\partial_{x_t} p_\ell \in \mathbb{F}$.

   Combining (36), (37), and (38), it follows that $x_t \in U^{(s)}(v)$ (see Definition 3.5). □

   Claim 8.4. *In the setting of Case 1 Let $\mathcal{H}$ be a hitting-set generator for* R3Fs*. Then $\partial_{x_t^{\deg}} p_v \circ \mathcal{H} \not\equiv 0$.*

   Proof. Assume, without loss of generality, that $\mathrm{Read}_{F_{(v)_L}}(x_t) = 3$. Since $F \in$ R4F, it must be the case that $\deg_{x_t}(p_{(v)_R}) \leqslant 1$. Let $\deg_{x_t}(p_v) = d$. Then

   $$\partial_{x_t^{\deg}} p_v = \partial_{x_t^d} p_{(v)_L} + \partial_{x_t^d} p_{(v)_R} = \partial_{x_t^{\deg}} p_{(v)_L} \in \mathsf{R3F} \,.$$

   The claim follows directly from the definition of $\mathcal{H}$. □

   Assume, for the next two cases, that no such variable x_t exists in v.

2. **No x_t as in the first case exists and $|U(v)| \geqslant 23$.**

Claim 8.5. *Assume that no nonzero polynomial in* $\mathcal{Q}_{F_{(v)_L},F_{(v)_R}} \cup \mathcal{E}^3_{F_v} \cup \mathcal{C}_{F_v}$ *vanishes at* $\mathbf{0}$*, and that* $|U(v)| \geqslant 23$*. Then there exists a variable* $x_t \in U(v)$ *such that, if* $n' \coloneqq |var(\partial_{x_t^{\deg}} p_v)| \geqslant r_{2,6}+5^8+23$*, then*

$$\mathcal{P}_{var(p_v)\setminus\{t\}} \nmid \partial_{x_t^{\deg}} p_v(\mathbf{x}) .$$

Proof. We show how to identify such a variable x_t based on the type of gate v.

We consider two subcases:

(a) $|S_{(v,\mathbf{ns})}| \geqslant 23$**:** Observe that

$$p_{(v)_L}\big|_{S_{(v,ns)}} + p_{(v)_R}\big|_{S_{(v,ns)}}$$

is a totally non-structural $\sum^2$ R2F. By assumption, every nonzero polynomial in $\mathcal{E}^3_{F_v} \cup \mathcal{C}_{F_v}$ does not vanish at $\mathbf{0}$. By the properties of generic assignments, this continues to hold for the corresponding sets obtained after restricting to $S_{(v,ns)}$. Since $|S_{(v,ns)}| \geqslant 23$, Lemma 8.1 implies that there exists a variable x_t such that

$$\mathcal{P}_{S_{(v,ns)}} \nmid \partial_{x_t^{\deg}} p_v\big|_{S_{(v,ns)}} .$$

As $x_t \in S_{(v,ns)}$, it is clearly in $U(v)$.

(b) $|S_{(v,\mathbf{ns})}| < 23$**:** Since $S_{(v,ns)} \sqcup S_{(v,s)} = [n']$, it follows that $|S_{(v,s)}| > r_{2,6} + 5^8$.
Observe that $p_{(v)_L}\big|_{S_{(v,s)}}$ and $p_{(v)_R}\big|_{S_{(v,s)}}$ are R3Fs, and their sum is a structural R4F. Since $|U(v)| \geqslant 23$ and $|S_{(v,ns)}| < 23$, there exists at least one structural variable in $U(v)$. Any degree pattern $\mathbf{e}$ with domain $\{x_i, x_j\} \subseteq \mathbf{x}$, such that $\mathbf{e}_i = \deg_{x_i}(p_v)$ and $\mathbf{e}_j = \deg_{x_j}(\partial_{x_i^{\deg}} p_v)$, is a dominating degree pattern. Therefore, $\partial_{\mathbf{e}} p_{(v)_L}, \partial_{\mathbf{e}} p_{(v)_R} \in \mathcal{E}^3_{F_v}$.
As in the previous case, every polynomial in $\mathcal{Q}_{F_{(v)_L},F_{(v)_R}}$ and in $\mathcal{E}^3_{F_v}$, after restriction to $S_{(v,s)}$, does not vanish at $\mathbf{0}$. This satisfies the conditions of Proposition 5.1 applied to the restricted formulas, with the set U of that proposition corresponding to $U(v)\setminus S_{(v,ns)}$. That proposition guarantees the existence of $x_t \in U(v)$ such that

$$\mathcal{P}_{S_{(v,s)}\setminus\{t\}} \nmid \partial_{x_t^{\deg}} p_v\big|_{S_{(v,s)}} .$$

□

To obtain a claim analogous to Claim 8.4, we first show that the sets appearing in Claim 8.5 consist of read-3 polynomials.

Claim 8.6. $\mathcal{Q}_{F_{(v)_L},F_{(v)_R}} \cup \mathcal{E}^3_{F_v} \cup \mathcal{C}_{F_v} \subseteq$ R3P.

Proof. We consider each set separately.

(a) Let $p \in \mathcal{Q}_{F_{(v)_L},F_{(v)_R}}$. Since $v \in \mathcal{F}^{(2)}_F$, both $(v)_L$ and $(v)_R$ are R3Fs. The claim follows from Corollary 2.6 and Lemma 2.11.

(b) Let $p \in \mathcal{C}_{F_v}$. By definition, there exist $w, u \in F_{(v)}$, both descendants of the same child of v, such that $p = p_{F_u|_{w\leftarrow\gamma_w}}$, where $\gamma_w \in \mathbb{F}$ is the additive constant of w. Since substituting a scalar for a gate does not increase the number of reads, the claim follows.

(c) Let $p \in \mathcal{E}^3_{F_\nu}$. Let $w \in F_\nu$ and let $\mathfrak{e}$ be the dominating degree pattern such that $p = \partial_{\mathfrak{e}} p_w$. Denote $I = \mathrm{Domain}(\mathfrak{e})$. For $|I| = 1$, we have $\partial_{\mathfrak{e}} p_\nu = \partial_{I^{\deg}} p_\nu$, and the claim follows from Lemma 2.11. For $1 < |I| \leqslant 3$, if there exists $j \in I$ such that $\mathrm{Read}_{x_j}(F_\nu) = 3$, then it must be the case that $\deg_{x_j}(p_\nu) = 1$ (otherwise we would be in the first case of the proof of Theorem 1.1). By Lemma 2.11, we have $\partial_{x_j} p_\nu$ is an R3P. Observe that

$$\mathfrak{d}_j = \begin{cases} 0, & i = j, \\ \mathfrak{e}_i, & \text{otherwise}, \end{cases}$$

is a dominating degree pattern in $\partial_{x_j^{\deg}} p_\nu$, and the claim follows (by induction). If no such j exists, then every variable in I is read at most twice in ν. By Lemma 6.8, the claim follows in this case as well. □

We can now prove the aforementioned claim.

Claim 8.7. *In the setting of Case 2, let x_t be chosen as in Claim 8.5, and let $\mathcal{H}$ be a hitting-set generator for* R3F*s. Then there exists $\boldsymbol{\alpha} \in Img(\mathcal{H})$ such that $\partial_{x_t^{\deg}} p_\nu(\mathbf{x} + \boldsymbol{\alpha})$ is $\left(r_{2,6} + 5^8 + 23\right)$-hard.*

Proof. By Claim 8.6, every polynomial in the sets $\mathcal{Q}_{F_{(g)_L}, F_{(g)_R}}$, $\mathcal{E}^3_{F_g}$, and $\mathcal{C}_{F_g}$ is an R3P. Therefore, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\mathcal{H})$ that hits all nonzero polynomials in these sets. By Claim 8.5, it follows that $\partial_{x_t^{\deg}} p_\nu(\mathbf{x} + \boldsymbol{\alpha})$ is $\left(r_{2,6} + 5^8 + 23\right)$-hard. □

3. **No x_t as in the first case exists and $|U(\nu)| < 23$.** From the fact that $\nu \in \mathcal{F}^{(2)}_F$, we have $U(\nu) \neq \emptyset$.

Claim 8.8. *Let $\mathcal{H}$ be a hitting-set generator for* R3F*s and for structurally multilinear* R8F*s. In the setting of Case 3,*

$$\partial_{x_t^{\deg}} p_\nu(\mathcal{H} + \mathcal{G}_{22}) \not\equiv 0\,.$$

Proof.

Claim 8.9. *Let $x_t \in U(\nu)$, and denote $I \coloneqq U(\nu) \setminus \{x_t\}$, $S \coloneqq [n] \setminus I$, and $f = \partial_{x_t^{\deg}}\left(p_\nu|_S\right)$. Then there exists $\boldsymbol{\alpha} \in Img(\mathcal{H})$ such that $f(\mathbf{x} + \boldsymbol{\alpha})$ is 2-hard.*

Proof. We prove the claim separately for the cases where f is multilinear and where f is non-multilinear. For the latter case, we first establish the following claim.

Claim 8.10. *If f is not multilinear, then for every $x_i \in var(f)$ with $\deg_{x_i}(f) > 1$, we have $\partial_{x_i^{\deg}} f \in$* R3P*.*

Proof. Let $x_i \in \mathrm{var}(f)$ such that $\deg_{x_i}(f) > 1$.

Claim 8.11. $\deg_{x_i}(p_\nu) = 2$.

Proof. Since $\nu \in \mathcal{F}^{(2)}_F$, the only other possibility is $\deg_{x_i}(p_\nu) = 3$. However, this would imply that for some $G \in \{F_{(\nu)_L}, F_{(\nu)_R}\}$, $\mathrm{Read}_G(x_i) = 3$ and $\deg_{x_i}(G) > 1$, contradicting the assumption that we are not in the first case. □

By Claim 8.11, we may assume without loss of generality that $\deg_{x_i}((v)_L) = 2$. Since $x_i \notin U(v)$, it must be the case that $\deg_{x_i}((v)_R) \leqslant 1$; otherwise, all occurrences of x_i would have been contained within v. Denote $\deg_{x_t}(p_v) = d$. Recall that $v^{op} = +$. We then have

$$\partial_{x_i^2} f = \partial_{x_i^2 x_t^d}(p_v|_S) = \partial_{x_t^d}\left(\partial_{x_i^2}(p_{(v)_L}|_S) + \partial_{x_i^2}(p_{(v)_R}|_S)\right) = \partial_{x_t^d x_i^2}(p_{(v)_L}|_S)\,. \tag{39}$$

We have

$$d \leqslant^{(*)} \deg_{x_t}\left(\partial_{x_i^2}(p_{(v)_L}|_S)\right) \leqslant \deg_{x_t}(p_v) = d\,,$$

where the inequality marked $(*)$ follows from the fact that $\partial_{x_i^2} f \not\equiv 0$ and from Equation (39). Hence $\deg_{x_t}\left(\partial_{x_i^2} p_{(v)_L}\right) = d$. The claim then follows from the facts that $F_{(v)_L}$ is an R3F (and therefore structural by Corollary 2.6) and by applying Lemma 2.11 twice. □

Using Claim 8.10, we obtain that if f is non-multilinear, i.e., a variable x_i as described in that claim exists, then there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\mathcal{H})$ such that $\partial_{x_i^2}\partial_{x_t} p_v(\boldsymbol{\alpha}) \not\equiv 0$. Therefore, by Lemma 2.39, we conclude that $\partial_{x_t} p_v$ is 2-hard.

If no such variable exists, then f is multilinear. Assume there exists $x_i \in \mathrm{var}(f)$ that is read 3 in some child of v. Since $x_i \notin U(v)$, it does not appear in the other child of v. Using an argument analogous to that in Claim 8.10, we can show that $\partial_{x_i} f \in$ R3P, and the claim follows in the same manner.

Hence, assume that every $x_i \in \mathrm{var}(f)$ is read at most twice in each child of v.

Claim 8.12. *In this case,* $f \in \sum^2$ R4P.

Proof. If x_t is structural in F_v, we are done by Lemma 2.11. Assume instead that it is non-structural, which implies that $\deg_{x_t}(p_v) = 1$. We first show that $\partial_{x_t}(p_{(v)_L}|_S) \in$ R4P; the claim for the other child then follows in the same way. By the definition of non-structurality, there exists $u \in (v)_L$ such that $\deg_{x_t}(u) = 2$. Using Lemma 2.8, we get

$$\begin{aligned}\partial_{x_t} p_{(v)_L}|_S &= \partial_{x_t} p_u|_S \prod_{w \in \mathrm{Unv}_{F_{(v)_L}}(u)} p_w|_S \\ &= \left(\partial_{x_t}(p_{(u)_L}|_S)\, p_{(u)_R}|_S + p_{(u)_L}|_S\, \partial_{x_t}(p_{(u)_R}|_S)\right) \prod_{w \in \mathrm{Unv}_{F_{(v)_L}}(u)} p_w|_S\,.\end{aligned}$$

In the polynomial on the right-hand side, we may set all variables not appearing in $\mathrm{var}(f)$ to 0 without changing the polynomial. The claim now follows since every $x_i \in \mathrm{var}(f)$ is read at most twice in any subformula of each child of v, and in u in particular. □

Therefore, by Claim 8.12, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\mathcal{H})$ that is a common nonzero for f, implying that $f(\boldsymbol{x} + \boldsymbol{\alpha})$ is 1-hard. □

By Claim 8.9 and Observation 2.38, we have $f(\mathcal{H} + \mathcal{G}_1) \not\equiv 0$.

Let $\boldsymbol{\tau} \in \mathbb{F}^n$ be the generic assignment satisfying $\partial_{x_t^{\deg}} p_v\big|_S^{\boldsymbol{\tau}} = f$. Recall that $S = [n] \setminus I$, so $f = \partial_{x_t^{\deg}} p_v\big|_{I \leftarrow \boldsymbol{\tau}_I}$. Since $|U(v)| < 23$, we have $|I| \leqslant 21$. Hence, by Observation 2.19, the assignment

$$\boldsymbol{\sigma} = \begin{cases} \boldsymbol{\tau}_i - (\mathcal{H} + \mathcal{G}_1)^{(i)}, & i \in I, \\ 0, & \text{otherwise,} \end{cases}$$

where $(\mathcal{H}+\mathcal{G}_1)^{(i)}$ denotes the ith coordinate of the generator $(\mathcal{H}+\mathcal{G}_1)$, lies in the image of $\mathcal{G}_{21}$. Therefore,

$$0 \not\equiv f(\mathcal{H}+\mathcal{G}_1) = \partial_{x_t^{\deg}} p_\nu\big|_{I\leftarrow\boldsymbol{\tau}_I}(\mathcal{H}+\mathcal{G}_1) = \partial_{x_t^{\deg}} p_\nu(\mathcal{H}+\mathcal{G}_1+\boldsymbol{\sigma})\,,$$

which implies that

$$\partial_{x_t^{\deg}} p_\nu(\mathcal{H}+\mathcal{G}_{22}) \not\equiv 0\,.$$

□

Combining Claim 8.4, Claim 8.7, and Claim 8.8, we conclude the proof of Theorem 1.1. By Theorem 1.3, there exists a constant $c_1 \in \mathbb{F}$ such that $\mathcal{G}_{2\log n+c_1}$ is a hitting-set generator for R3Fs. Moreover, by Theorem 2.44, $\mathcal{G}_{8\log n+r_8}$ is a hitting-set generator for structurally multilinear R8Fs. We therefore set $\mathcal{H} = \mathcal{G}_{8\log n+r_8+c_1}$.

Every $\nu \in \mathcal{F}_F^{(2)}$ falls into one of the three cases described above. If it falls into the setting of Case 1, then by Claim 8.4 and Observation 2.38, there exists $x_t \in U^{(s)}(\nu)$ such that

$$\partial_{x_t^{\deg}} p_\nu \circ \mathcal{H} \not\equiv 0\,.$$

If it falls into the setting of Case 2, then by Claim 8.7 and Observation 2.38, there exists $x_t \in U^{(s)}(\nu)$ such that

$$\partial_{x_t^{\deg}} p_\nu \circ \left(\mathcal{H}+\mathcal{G}_{r_{2,6}+5^8+23}\right) \not\equiv 0\,.$$

Lastly, if it falls into the setting of Case 3, then by Claim 8.8, there exists $x_t \in U^{(s)}(\nu)$ such that

$$\partial_{x_t^{\deg}} p_\nu \circ \left(\mathcal{H}+\mathcal{G}_{22}\right) \not\equiv 0\,.$$

Overall, for the constant $c = r_8 + c_1 + r_{2,6} + 5^8 + 23$, we obtain

$$\partial_{x_t^{\deg}} p_\nu \circ \mathcal{G}_{8\log n+c} \not\equiv 0$$

for any $\nu \in \mathcal{F}^{(2)}$ and some $x_t \in U^{(s)}(\nu)$. Together with Lemma 3.8, this implies the claim. □

8.2 Whitebox Algorithm

In this subsection, we present our polynomial-time whitebox algorithm and prove Theorem 1.2. For convenience, we restate the theorem here.

Theorem 1.2 (Main whitebox result)**.** *Let F be an n-variate* R4F *formula over a field $\mathbb{F}$. There exists an algorithm that, given whitebox access to F, checks in time poly(n) whether $p_F \equiv 0$.*

8.2.1 Definitions and Auxiliary Algorithms

The following definitions provide the whitebox analogues of $\mathrm{var}(F)$, the unique set $U(\nu)$, and the frontier $\mathcal{F}_F$.

Definition 8.13 ($\mathrm{occ}(F)$)**.** Let F be an algebraic formula. We denote by $\mathrm{occ}(F)$ the set of variables that label a leaf in F. ◊

Definition 8.14 ($U^{(w)}(\nu)$)**.** Let $F \in$ RkF and let $\nu \in F$. Define $U_F^{(w)}(\nu) \subseteq \boldsymbol{x}$ as the set of variables whose every leaf occurs in the subformula F_ν. When F is clear from context, we simply write $U^{(w)}(\nu)$. ◊

The only distinction between $U^{(w)}(v)$ and $U(v)$ is that the latter requires F to depend on the variables in $U(v)$, whereas $U^{(w)}(v)$ imposes no such requirement.

Definition 8.15 ($\mathcal{F}_F^{(w)}$)**.** Let $F \in \mathsf{RkF}$. Define $\mathcal{F}_F^{(w)}$ as the set of gates $v \in F$ for which there exists a variable in $\mathbf{x}$ that occurs in F_v exactly k times, and no descendant of v has this property. ◊

In our whitebox algorithm, we compute derivatives of formulas in the natural way.

Definition 8.16. Let F be an algebraic formula. For a variable x_t and an integer $d \geqslant 1$, the derivative $\partial_{x_t^d} F$ is defined as the formula obtained by recursively applying the usual derivative rules for addition and multiplication at every gate, starting at the output gate and proceeding toward the leaves. ◊

We use the whitebox algorithm for R3Fs developed in [MRS14], which we denote by $\mathrm{WB}_{\mathsf{R3F}}$, and the whitebox PIT algorithm for *structurally multilinear* R8Fs from [AvMV15], which we denote by $\mathrm{WB}_{\text{sm-R8F}}$. Both algorithms output '1' if their input is the zero polynomial, otherwise they output '0'.

Theorem 8.17 ([MRS14], Theorem 5, restated)**.** *Let $F \in \mathsf{R3F}$ and let n be the number of variables occurring in F. There is a polynomial time algorithm that, given whitebox access to F, decides whether $p_F \equiv 0$.*

Theorem 8.18 ([AvMV15], Theorem 1.1)**.** *Let F be a structurally multilinear RkF over n variables. There is an $n^{O_k(1)}$ time algorithm that, given whitebox access to F, decides whether $p_F \equiv 0$.*

Throughout our algorithm, we will need to determine the degree of a variable in an R3F at several points. Algorithm 1 describes how to compute $\deg_{x_t}(F)$ for $F \in \mathsf{R3F}$ and a variable x_t. It also removes redundant occurrences of x_t. Every polynomial identity is verified using Theorem 8.17.

Algorithm 1 $\mathrm{Deg}_{\mathsf{R3F}}(F, x_t)$

if $\mathrm{Read}_F(x_t) \leqslant 1$ **then**
  if $\mathrm{WB}_{\mathsf{R3F}}(\partial_{x_t} F) = 1$ **then** $F \leftarrow F|_{x_t=0}$
    return 0
  else return 1
  end if
end if
$d_L \leftarrow \mathrm{Deg}_{\mathsf{R3F}}(F_{(o)_L})$
$d_R \leftarrow \mathrm{Deg}_{\mathsf{R3F}}(F_{(o)_R})$
if $o^{\mathrm{op}} = \times$ **then return** $d_L + d_R$
else
  if $d_L \neq d_R$ **then return** $\max(d_L, d_R)$
  else if $\mathrm{WB}_{\mathsf{R3F}}(\partial_{x_t} F) = 1$ **then** $F \leftarrow F|_{x_t=0}$
    return 0
  else return 1
  end if
end if

Claim 8.19. *If $F \in \mathsf{R3F}$, then $\mathrm{Deg}_{\mathsf{R3F}}(F, x_t)$ computes correctly $\deg_{x_t}(F)$ in time $\mathrm{poly}(n)$.*

Proof. We prove the claim by induction on the size of F.

If $\mathrm{Read}_F(x_t) \leqslant 1$, the claim follows from the fact that Read, $\mathrm{WB}_{\mathsf{R3F}}$, and ∂_{x_t} can each be computed in time $\mathrm{poly}(n)$.

Otherwise, by the induction hypothesis, d_L and d_R are computed correctly in time $\mathrm{poly}(n)$. If $o^{\mathrm{op}} = \times$ or $d_L \neq d_R$, the claim follows immediately. Otherwise, since F is an $\mathsf{R3F}$, it must be the case that both d_L and d_R equal 1. Hence, by Corollary 2.6 and Lemma 2.11, $\partial_{x_t} F$ is an $\mathsf{R3F}$, and the claim follows since both ∂_{x_t} and $\mathrm{WB}_{\mathsf{R3F}}$ can be computed in time $\mathrm{poly}(n)$. □

8.2.2 The Algorithm

The algorithm is given in page 74.

Sketch of proof of Theorem 1.2. We give a sketch of the proof of Theorem 1.2. The missing details are given in Subsubsection 8.2.3.

We traverse the formula to identify the set S_4 of variables that occur exactly four times in F. We now define an algorithm, denoted by $\mathrm{WB}_{\mathsf{R4F}}(\cdot,\cdot)$, which takes as input a formula F and a set S, and operates iteratively. In each iteration, it aims to reduce the number of read-4 variables in F while preserving its identity (zeroness or nonzeroness). We run this algorithm on the input (F, S_4).

$\mathrm{WB}_{\mathsf{R4F}}(F, S_4)$:
If $S_4 = \emptyset$, then $F \in \mathsf{R3F}$, and we return the result of $\mathrm{WB}_{\mathsf{R3F}}(F)$. Otherwise, proceed to identify a gate $v \in \mathcal{F}_F^{(w)}$ (which by definition contains some variable from S_4).

If $v^{\mathrm{op}} = \times$: Using $\mathrm{WB}_{\mathsf{R3F}}$, we verify that both children of v are nonzero. If either child is zero, let γ_v be the additive constant of the gate v. We continue to the next iteration with $F|_{v=\gamma_v}$ and $S_4 \setminus U^{(w)}(v)$.

Otherwise, let $x_t \in U^{(w)}(v)$. If $\mathrm{Deg}_{\mathsf{R3F}}(F_{(v)_L}, x_t) \neq 0$ or $\mathrm{Deg}_{\mathsf{R3F}}(F_{(v)_R}, x_t) \neq 0$, then $\partial_{x_t} p_v \not\equiv 0$. We continue to the next iteration with $F|_{v=x_t}$ and $S_4 \setminus U^{(w)}(v)$. If p_F was not zero, then since x_t does not appear anywhere else in the formula, this property is maintained (see Claim 8.21).

Otherwise, $\partial_{x_t} p_v \equiv 0$. We continue to the next iteration with $F|_{x_t=0}$ and $S_4 \setminus \{x_t\}$.

From now on, we assume $v^{\mathrm{op}} = +$.

If a variable is both non-multilinear and read three times in some child of v: We check for this case in the natural way, i.e. for each $G \in \{F_{(v)_L}, F_{(v)_R}\}$ and $x_t \in \mathrm{occ}(G)$, we first verify whether $\mathrm{Read}_{x_t}(G) = 3$; if so, we then check whether $\mathrm{Deg}_{\mathsf{R3F}}(G, x_t) > 1$. If a variable satisfying both conditions exists, we proceed to the next iteration with $F|_{v=x_{t'}}$, where $x_{t'}$ is a new variable, and with the set $S_4 \setminus U^{(w)}(v)$. Noting that $\deg_{x_t}(G) \geqslant 2$ and at most 1 elsewhere, the identity of F is preserved by an argument similar to that of Claim 8.21.

Searching for a variable in $U(v)$: We proceed with two procedures, each designed to find a variable in $U(v)$ under different assumptions on its size. Observe that $U(v) \subseteq U^{(w)}(v)$, but some, or maybe all, variables in $U^{(w)}(v)$ may cancel out in p_v, such that F does not depend on them. To apply Claim 8.21 and proceed to the next iteration, we must ensure that there exists $x_t \in U^{(w)}(v)$ for which this does not happen. The first procedure finds such a variable when $1 \leqslant |U(v)| < 23$, and the second does so when $|U(v)| \geqslant 23$. If neither procedure identifies a variable in $U(v)$, then it must be that $U(v) = \emptyset$, in which case we can eliminate all variables in $U^{(w)}(v)$ while preserving the identity of F.

First procedure, designed for $1 \leqslant |U(v)| < 23$: For each $x_t \in U^{(w)}(v)$, let $I = U^{(w)}(v) \setminus \{x_t\}$. Substitute $\mathcal{G}_{I,21}$ for the variables in I and compute

$$H_t \coloneqq \partial_{x_t} F_v\big|_{I \leftarrow \mathcal{G}_{I,21}}.$$

Algorithm 2 $\mathrm{WB}_{\mathsf{R4F}}(F, S_4)$

if $S_4 = \emptyset$ **then return** $\mathrm{WB}_{\mathsf{R3F}}(F)$
end if
Find $v \in \mathcal{F}_F^{(w)}$.
if $v^{\mathrm{op}} = \times$ **then**
  if $\mathrm{WB}_{\mathsf{R3F}}\left(F_{(v)_L}\right) = 1$ or $\mathrm{WB}_{\mathsf{R3F}}\left(F_{(v)_R}\right) = 1$ **then return** $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{v=0}, S_4 \setminus U^{(w)}(v)\right)$
  else Find $x_t \in U^{(w)}(v)$ which occurs in v four times.
    if $\mathrm{Deg}_{\mathsf{R3F}}\left(F_{(v)_L}, x_t\right) \neq 0$ or $\mathrm{Deg}_{\mathsf{R3F}}\left(F_{(v)_R}, x_t\right) \neq 0$ **then return** $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{v=x_t}, S_4 \setminus U^{(w)}(v)\setminus\right)$ ▷ This implies that $\partial_{x_t^{\deg}} p_v \not\equiv 0$
    else return $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{x_t=0}, S_4 \setminus \{x_t\}\right)$
    end if
  end if
else ▷ $v^{\mathrm{op}} = +$
  for all $x_t \in \mathrm{var}(v)$ and $G \in \{F_{(v)_L}, F_{(v)_R}\}$ **do** ▷ Looking for a read-3 non-multilinear variable
    if $\mathrm{Read}_{x_t}(G) = 3$ and $\mathrm{Deg}_{\mathsf{R3F}}(G, x_t) > 1$ **then**
      Define $x_{t'} \notin \mathrm{var}(F)$ **return** $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{v=x_{t'}}, S_4 \setminus U^{(w)}(v)\right)$
    end if
  end for
  for all $x_t \in U^{(w)}(v)$ **do** ▷ Handling the case where $U(v) < 23$
    $I = U^{(w)}(v) \setminus \{x_t\}$
    $S = [n] \setminus I$
    $H = \partial_{x_t} F_v|_{I \leftarrow \mathcal{G}_{I,21}}$ ▷ We consider the variables of $\mathcal{G}_{I,21}$ as scalars
    for all $x_i \in \mathrm{occ}(H) \setminus \{x_t\}$ and $G \in \{F_{(v)_L}, F_{(v)_R}\}$ **do**
      if $\mathrm{Deg}_{\mathsf{R3F}}(G, x_i) = 2$ **then** $d = 2$
      else if $\mathrm{Read}_{x_i}(G) = 0$ **then** Let $G' \in \{F_{(v)_L}, F_{(v)_R}\} \setminus \{G\}$
        $d = \mathrm{Deg}_{\mathsf{R3F}}(G', x_i)$
      else $d = 0$
      end if
      if $d \neq 0$ **then**
        $P = \partial_{x_i^d} F_v|_{I \leftarrow \mathcal{G}_{I,21}}$
        if $\mathrm{Deg}_{\mathsf{R3F}}(P, x_t) > 0$ **then return** $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{v=x_t}, S_4 \setminus U^{(w)}(v)\right)$
        end if
      end if
    end for ▷ If we didn't return, then H is multilinear
    if $\mathrm{WB}_{\text{sm-R8F}}(H) = 1$ **then return** $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{v=x_t}, S_4 \setminus U^{(w)}(v)\right)$
    end if
  end for
  Calculate a common nonzero $\boldsymbol{\sigma}$ to the polynomials in $\mathcal{Q}_{F_{(v)_L}, F_{(v)_R}} \cup \mathcal{E}^3_{F_v} \cup \mathcal{C}_{F_v}$ ▷ Handling the case where $U(v) \geqslant 23$
  for all $x_t \in U^{(w)}(v)$ and $d \in \{1, 2\}$ **do**
    if $\mathrm{BB}\left(\partial_{x_t^d} p_v(\mathbf{x} + \boldsymbol{\sigma}) \circ \mathcal{G}_{r_{2,6}+5^8+23}, 2n\right) = 0$ **then return** $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{v=x_t}, S_4 \setminus U^{(w)}(v)\right)$
    end if
  end for
return $\mathrm{WB}_{\mathsf{R4F}}\left(F|_{U^{(w)}(v) \leftarrow \mathbf{0}}, S_4 \setminus U^{(w)}(v)\right)$
end if

We treat the variables of $\mathcal{G}_{I,21}$ as scalars, i.e. if $\mathbf{y}$ are the variables of $\mathcal{G}_{I,21}$, then in the following procedure we work over the field of rational functions, $\mathbb{F}(\mathbf{y})$.

Observe that if $x_t \in \mathrm{var}(p_\nu)$, and our assumption that $|U(\nu)| < 23$ is true then $|I| \leqslant 21$ and by Observation 2.19, we have $H \not\equiv 0$.

Handling assymetric degrees: For each x_i that occurs in H_t compute $d_L = \mathrm{Deg}_{\mathsf{R3F}}(F_{(\nu)_L}, x_i)$ and $d_R = \mathrm{Deg}_{\mathsf{R3F}}(F_{(\nu)_R}, x_i)$. If $d_L \neq d_R$, let $d = \max(d_L, d_R)$. Compute

$$P \coloneqq \partial_{x_i^d} F_\nu\big|_{I \leftarrow \mathcal{G}_{I,21}}.$$

Since the degrees are asymmetrical, $P \in \{\partial_{x_t^{\deg}} F_{(\nu)_L}, \partial_{x_t^{\deg}} F_{(\nu)_R}\}$. Then, by Corollary 2.6 and Lemma 2.11 we have $P \in \mathsf{R3F}$. Compute $\mathrm{Deg}_{\mathsf{R3F}}(P, x_t)$. If it is nonzero then $x_t \in U(\nu)$. In that case we continue to the next iteration with $F|_{\nu=x_t}$ and $S_4 \setminus U^{(w)}(\nu)$.

Claim 8.20. *If the degrees of every $x_i \in \mathit{occ}(H_t)$ in both children of ν are identical, then H_t is a structurally multilinear* R8F.

If the above procedure did not proceed to the next iteration, then by Claim 8.20, H_t is a structurally multilinear R8F. To find out if p_ν depends on x_t, under the assumption that $|U(\nu)| < 23$ we use $\mathrm{WB}_{\text{sm-R8F}}$ to test the identity of H_t. If H_t is nonzero, we continue to the next iteration with $F|_{\nu=x_t}$ and $S_4 \setminus U^{(w)}(\nu)$.

If for all $x_t \in U^{(w)}(\nu)$ we didn't proceed to the next iteration, then for all x_t, $H_t \equiv 0$. This implies that either $|U(\nu)| \geqslant 23$ or $U(\nu) = \emptyset$. We proceed to to check if the former holds.

Second procedure, designed for $|U(\nu)| \geqslant 23$: We handle this using Claim 8.5. To this end, we find a common nonzero $\boldsymbol{\sigma} \in \mathbb{F}^n$ for the nonzero polynomials in $\mathcal{Q}_{F_{(\nu)_L}, F_{(\nu)_R}} \cup \mathcal{E}^3_{F_\nu} \cup \mathcal{C}_{F_\nu}$.

Computing a common nonzero: We do so by computing all nonzero polynomials in that set. The computation of the polynomials in $\mathcal{Q}_{F_{(\nu)_L}}$ and $\mathcal{C}_{F_\nu}$ is straightforward, by definition. To compute the polynomials in $\mathcal{E}^3_{F_\nu}$, we use the nontrivial succinct formulas described in Lemma 6.8. To exclude zero polynomials, we use $\mathrm{WB}_{\mathsf{R3F}}$. Then, employing search-to-decision together with $\mathrm{WB}_{\mathsf{R3F}}$, we find $\boldsymbol{\sigma}$ as follows. By Claim 8.22, there are at most $\mathrm{poly}(n)$ polynomials in $\mathcal{Q}_{F_{(\nu)_L}, F_{(\nu)_R}} \cup \mathcal{E}^3_{F_\nu} \cup \mathcal{C}_{F_\nu}$. Since the individual degree with respect to any variable in each of these polynomials is at most 3, the individual degree of their product

$$\prod_{p \in \mathcal{Q}_{F_{(\nu)_L}, F_{(\nu)_R}} \cup \mathcal{E}^3_{F_\nu} \cup \mathcal{C}_{F_\nu}} p$$

is bounded by some $d = \mathrm{poly}(n)$. Therefore, for every coordinate $i \in [n]$, we are guaranteed that by evaluating this coordinate on at most $d+1$ different elements of $\mathbb{F}$, we can find a value $\gamma_i \in \mathbb{F}$ such that for every p in the above set, $p|_{x_i=\gamma_i} \not\equiv 0$. We verify the nonzeroness of each polynomial using $\mathrm{WB}_{\mathsf{R3F}}$. We find $\boldsymbol{\sigma} \in \mathbb{F}^n$ by doing so for every coordinate, then fixing it to γ_i.

After computing $\boldsymbol{\sigma}$, by Claim 8.5 we obtain that if $|U(\nu)| \geqslant 23$, then there exists $x_t \in U(\nu) \subseteq U^{(w)}(\nu)$ such that $\partial_{x_t} p_\nu(\mathbf{x} + \boldsymbol{\sigma})$ is $(r_{2,6} + 5^8 + 23)$-hard. Since $\deg_{x_t}(p_\nu) \in \{0, 1, 2\}$, we test for each $d \in \{1, 2\}$ the identity of $\partial_{x_t^d} p_\nu(\mathbf{x} + \boldsymbol{\sigma})$ by composing it with $\mathcal{G}_{(r_{2,6}+5^8+23)}$ and applying the trivial brute-force algorithm (In the procedure designed for the case $|U(\nu)| < 23$, we treated the variables of the generator as scalars. We now treat them as variables). Since the individual degree of p_ν is at most 2 and each coordinate of $\mathcal{G}$ (for any seed) has degree n, the test is performed for

an individual degree of at most $2n$ over $r_{2,6} + 5^8 + 23$ variables. Hence, each test requires at most $(2n+1)^{r_{2,6}+5^8+23}$ evaluations.

Concluding that $U(v) = \emptyset$: If both of the above procedures did not proceed to the next iteration, then $U(v) = \emptyset$, meaning that F does not depend on any variable in $U^{(w)}(v)$. We proceed to the next iteration with $F|_{U^{(w)}(v) \leftarrow 0}$ and $S_4 \setminus U^{(w)}(v)$. □

8.2.3 Missing proofs

In this section we present the proofs of the unproved claims in our whitebox algorithm.

Claim 8.21. *Assume that for some $x_t \in U^{(w)}(v)$ we have $\partial_{x_t} p_v \not\equiv 0$. Then, letting $F' = F|_{v=x_t}$, we have*

$$F \equiv 0 \iff F' \equiv 0\,.$$

Proof. If $F' \equiv 0$, then $0 \equiv F'|_{x_t = p_v} \equiv F$.

Otherwise, if $\prod_{w \in \mathrm{Unv}_F(v)} p_w \not\equiv 0$, then F depends on x_t. Indeed, along any addition gate on the path from v to o, only one child depends on x_t; hence such a gate cannot compute the zero polynomial. Therefore, if $F \equiv 0$, it follows that

$$\prod_{w \in \mathrm{Unv}_F(v)} p_w = \partial_{x_t} p_{F'} \equiv 0,$$

which proves the claim. □

Claim 8.20. *If the degrees of every $x_i \in occ(H_t)$ in both children of v are identical, then H_t is a structurally multilinear* R8F*.*

Proof. Since every variable in $\mathrm{occ}(v) \setminus U^{(w)}(v)$ is read 3, by our assumption its degrees in both children of v must be either one or zero. This, together with Corollary 2.6, implies that H_t is structurally multilinear with respect to all variables other than x_t. Moreover, for every variable whose degrees in both children are zero, the variable is redundant, and Algorithm 1 eliminates it. Otherwise, both degrees are one, which means that the variable appears in each child of v at most twice.

No variable can appear in any child of v with degree 3, since this would require it to be a variable of degree greater than one which is read three times in some child of v, contrary to the fact it was not found before. Hence $\deg_{x_t}(p_v) \leqslant 2$.

Let K be computed by some child of o_{H_t}. We claim that K is an R4F. Recall that $\deg_{x_t}(p_v) \leqslant 2$. If $\deg_{x_t}(K) = 1$, the claim is immediate, by Lemma 2.11. Otherwise, $\deg_{x_t}(K) = 2$. Let u be the first common gate in K of the two leaves labeled by t. By Lemma 2.8, we have

$$\partial_{x_t} K = \partial_{x_t} F_u \prod_{w \in \mathrm{Unv}_K(u)} F_w = \left(\partial_{x_t} F_{(u)_L} \cdot F_{(u)_R} + F_{(u)_L} \cdot \partial_{x_t} F_{(u)_R}\right) \prod_{w \in \mathrm{Unv}_K(u)} F_w\,,$$

which is an R4F since F_u is an R2F. Moreover, as can be seen in the equation above, it is structurally multilinear with respect to x_t. This, together with the first claim of this proof implies that H is structurally multilinear with respect to every variable.

□

To argue that our algorithm runs in time $\mathrm{poly}(n)$, the only nontrivial assertion is the following.

Claim 8.22. *Let $v \in \mathcal{F}^{(2)}_F$ and let the number of variables occurring in F be n. Then*

$$|\mathcal{Q}_{F_{(v)_L},F_{(v)_R}} \cup \mathcal{E}^3_{F_v} \cup \mathcal{C}_{F_v}| = \mathit{poly}(n)\,.$$

Proof. We bound the size of each of these sets separately. Let s be the size of F_v, which is at most $8n-1$.

- **Size of $\mathcal{Q}_{F_{(v)_L},F_{(v)_R}}$:** By Definition 2.42,

$$|\mathcal{Q}_{F_{(v)_L},F_{(v)_R}}| = \sum_{u \in F_{(v)_L} \cup F_{(v)_R}} \text{number of ways to choose the sets } Y \text{ and } Z \leqslant s\cdot 2\cdot \binom{n}{b_{2,6}} = \mathrm{poly}(n)\,.$$

- **Size of $\mathcal{E}^3_{F_v}$:** By Definition 6.5,

$$|\mathcal{E}^3_{F_v}| = \sum_{u \in F_{(o)_L} \cup F_{(o)_R}} |\mathcal{A}^3_u| \leqslant \sum_{u \in F_{(o)_L} \cup F_{(o)_R}} 3^2 \binom{n}{3} \leqslant 9sn^3 = \mathrm{poly}(n)\,.$$

- **Size of $\mathcal{C}_{F_v}$:** By Definition 7.14,

$$|\mathcal{C}_{F_v}| = \sum_{u \in F_{(o)_L} \cup F_{(o)_R}} |\mathcal{C}_u| \leqslant s^2 = \mathrm{poly}(n)\,.$$

□

9 Hardness of Representation Implies PIT for Orbits

In this section, we show how to translate known hardness of representation results for a backbone class into polynomial identity testing (PIT) results for the corresponding orbit class. This reduction follows from a simple but useful observation, which we formalize in Subsection 9.2.

9.1 Notation, Definitions, and Elementary Observations

In this section, we study orbits of bounded-read formulas under the action of the affine general linear group. The underlying formula will be referred to as the *backbone*. Throughout, the backbone variables are denoted by $\mathbf{y} = \{y_1, y_2, \ldots, y_N\}$, and the variables of the linear functions by $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$[19]. Unless stated otherwise, we follow the notation and conventions of [MS21].

For $f \in \mathbb{F}[\mathbf{x}]$ and $i \in \mathbb{N}$, we denote by $f^{[i]}$ the homogeneous component of f of total degree i.

Definition 9.1 (Directional Derivative)**.** [MS21, Definition 3.6] For $f \in \mathbb{F}[\mathbf{x}]$ and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n) \in \mathbb{F}^n$, the derivative of f in the direction $\boldsymbol{\delta}$ is defined as

$$\partial_{\boldsymbol{\delta}} f = \sum_{i=1}^{n} \delta_i\, \partial_{x_i} f. \qquad \lozenge$$

Observe that the usual calculus rules—the chain rule, the sum rule, and the product rule—continue to hold for directional derivatives in this setting.

[19] Note that $N \leqslant n$.

Definition 9.2 (Dual Set)**.** [MS21, Definition 3.7] A *dual set* for m linearly independent linear functions $\ell_1(\boldsymbol{x}), \ell_2(\boldsymbol{x}), \ldots, \ell_m(\boldsymbol{x})$ in $n \geqslant m$ variables is a collection of vectors $\{\boldsymbol{\delta}_i\}_{i\in[m]} \subset \mathbb{F}^n$ such that

$$\ell_i^{[1]}(\boldsymbol{\delta}_j) = \begin{cases} 1, & i = j, \\ 0, & \text{otherwise.} \end{cases}$$

(We say that linear functions are linearly independent if and only if their degree-1 homogeneous parts are linearly independent.)

The dual set corresponding to $(\boldsymbol{A}, \boldsymbol{\beta}) \in \mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$ is the dual set to the linear functions $\ell_i(\boldsymbol{x}) = (\boldsymbol{A}\boldsymbol{x})_i$. Equivalently, $\boldsymbol{\delta}_i$ is the i-th column of $\boldsymbol{A}^{-1}$. ◊

Lemma 9.3 ([MS21, Lemma 3.8])**.** *Let $(\boldsymbol{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$ with dual set $\mathcal{D} = \{\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_n\} \subset \mathbb{F}^n$, and let $g \in \mathbb{F}[\boldsymbol{y}]$. Then, for $f(\boldsymbol{x}) = g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta})$ it holds that*

$$\frac{\partial f}{\partial \boldsymbol{\delta}_i}(\boldsymbol{x}) = \frac{\partial g}{\partial y_i}(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}).$$

Lemma 9.4 ([MS21, Lemma 3.9])**.** *Let $f \in \mathbb{F}[\boldsymbol{x}]$ and let $\mathcal{H}$ be a polynomial map. Then, for any $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_k \in \mathbb{F}^n$,*

$$\partial_{\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_k} f \circ \mathcal{H} \not\equiv 0 \quad \Rightarrow \quad f \circ (\mathcal{G}_k + \mathcal{H}) \not\equiv 0.$$

Lemma 9.5 ([MS21, Implicit in Lemma 3.10])**.** *Let $g \in \mathbb{F}[\boldsymbol{x}]$ and let $(\boldsymbol{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. Let $S \subseteq [N]$ be a subset of size k and $\boldsymbol{\tau}_S \in \mathbb{F}^S$. Then, there exist $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{\beta}}) \in GL_n^{aff}(\mathbb{F})$, an assignment $\boldsymbol{\alpha} \in \mathbb{F}^k$ and linear functions $\boldsymbol{L} = L_1, L_2, \ldots, L_k \in \mathbb{F}[\boldsymbol{x}]$ such that*

$$g(\boldsymbol{A}(\boldsymbol{x} + \mathcal{G}_k(\boldsymbol{\alpha}, \boldsymbol{L}(\boldsymbol{x}))) + \boldsymbol{\beta}) = g|_{\boldsymbol{y}_S = \boldsymbol{\tau}_S}\left(\tilde{\boldsymbol{A}}\boldsymbol{x} + \tilde{\boldsymbol{\beta}}\right).$$

Theorem 9.6 ([MS21, Theorem 1.22])**.** *Let $0 \not\equiv F(\boldsymbol{y}) \in \mathsf{ROF}$ and let $(\boldsymbol{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. Then,*

$$p_F(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{\log N + 1} \not\equiv 0.$$

We next extend some of the definitions above for derivatives of higher orders.

Definition 9.7 ($\partial_{\boldsymbol{\delta}^{\deg}}$)**.** Let $F \in \mathbb{F}[\boldsymbol{y}]$, and let $(\boldsymbol{A}, \boldsymbol{\beta}) \in \mathrm{GL}_n^{\mathrm{aff}}(\mathbb{F})$ with dual set $\mathcal{D} = \{\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_n\} \subset \mathbb{F}^n$. Let $d_i \coloneqq \deg_{y_i}(F)$. We define

$$\partial_{\boldsymbol{\delta}_i^{\deg}} p_F \coloneqq \partial_{\boldsymbol{\delta}_i^{d_i}} p_F.$$ ◊

Lemma 9.8. *Let $F(\boldsymbol{y}) \in \mathsf{RkF}$ be structural, and let $(\boldsymbol{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$ with dual set $\mathcal{D} = \{\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_n\} \subset \mathbb{F}^n$. Then, for every $i \in [N]$, $\partial_{\boldsymbol{\delta}_i^{\deg}} p_F$ is computed by a formula in $\mathsf{RkF}^{GL_n^{aff}(\mathbb{F})}$.*

Proof. By Lemma 9.3, we have

$$\partial_{\boldsymbol{\delta}_i^{\deg}} p_F(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}) = \left(\partial_{y_i^{\deg}} p_F\right)(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}).$$

By Lemma 2.11, $\partial_{y_i^{\deg}} p_F$ is an RkP, which implies the claim. □

Observation 9.9. *Let $\mathcal{H}$ be a hitting-set generator for the class $\mathcal{C}^{GL_n^{aff}(\mathbb{F})}$. Then, for every $(\boldsymbol{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$, the composition $(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{H}$ is a hitting-set generator for $\mathcal{C}$.*

9.2 Translating Hardness to PIT for Orbits

The two key observations in what follows are that the action of matrices in $\mathrm{GL}_n(\mathbb{F})$ is a degree-preserving automorphism of $\mathbb{F}[\boldsymbol{x}]$, and that classes of bounded-read formulas are closed under translations.

Lemma 9.10. *Let $0 \not\equiv f(\boldsymbol{y}) \in \mathbb{F}[\boldsymbol{y}]$ be a polynomial with individual degrees at most d, and suppose that f contains a monomial whose support size is s. Then, for every $\boldsymbol{A} \in GL_n(\mathbb{F})$, we have*

$$f(\boldsymbol{A}\boldsymbol{x}) \circ \mathcal{G}_{d \cdot s} \not\equiv 0 \,.$$

Proof. Let $d_{\min}$ denote the minimal degree among all monomials of f. Clearly, $d_{\min} \leqslant d \cdot s$. Recall that for every $i \in \mathbb{N}$, $f^{[i]}$ denotes the homogeneous component of f of degree i, so that

$$f = \sum_{i=0}^{\deg(f)} f^{[i]}(\boldsymbol{y}) \,.$$

Since each linear function in $\boldsymbol{A}$ is homogeneous of degree 1, $f^{[i]}(\boldsymbol{A}\boldsymbol{x})$ is homogeneous of degree i. As the action of $\mathrm{GL}_n(\mathbb{F})$ on $\mathbb{F}[\boldsymbol{x}]$ is an automorphism and $f^{[d_{\min}]}(\boldsymbol{y})$ is nonzero, it follows that $f^{[d_{\min}]}(\boldsymbol{A}\boldsymbol{x})$ is also nonzero and has degree $d_{\min}$. Consequently,

$$f(\boldsymbol{A}\boldsymbol{x}) = \sum_{i=1}^{\deg(f)} f^{[i]}(\boldsymbol{A}\boldsymbol{x})$$

contains a monomial of degree $d_{\min}$ supported on at most $d_{\min}$ variables. Since $d_{\min} \leqslant d \cdot s$, Observation 2.19 implies that

$$f(\boldsymbol{A}\boldsymbol{x}) \circ \mathcal{G}_{d \cdot s} \not\equiv 0 \,. \qquad \square$$

The general outline we follow in most of the proofs in this section is summarized in the next lemma.

Lemma 9.11. *Let $0 \not\equiv f(\boldsymbol{y}) \in \mathcal{C}_1$ be a polynomial with individual degrees at most d, and assume that $\mathcal{C}_1$ is closed under translations. Moreover, suppose that for every nonzero polynomial in $\mathcal{C}_1$, there exists a set of polynomials $\mathcal{S} \subseteq \mathcal{C}_2$ such that translating f by a nonzero element of $\mathcal{S}$ makes it m-hard.*

Let $\mathcal{H}$ be a hitting-set generator for $\mathcal{C}_2^{GL_n^{aff}(\mathbb{F})}$. Then, for every $(\boldsymbol{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$ with $\boldsymbol{A} \neq \boldsymbol{0}$, we have

$$f(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \left(\mathcal{H} + \mathcal{G}_{d(m-1)}\right) \not\equiv 0 \,.$$

Proof. Let $\hat{f} = f(\boldsymbol{y} + \boldsymbol{\beta})$. Since $\mathcal{C}_1$ is closed under translations, $\hat{f} \in \mathcal{C}_1$. Moreover, $\hat{f}$ is nonzero because translation by $\boldsymbol{\beta}$ is an automorphism of $\mathbb{F}[\boldsymbol{y}]$. Let $\mathcal{S}$ be the corresponding set of polynomials associated with $\hat{f}$. By Observation 9.9, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\boldsymbol{A}\boldsymbol{x} \circ \mathcal{H})$ that is a common nonzero for all polynomials in $\mathcal{S}$. Since $\hat{f}(\boldsymbol{y} + \boldsymbol{\alpha})$ is m-hard, Observation 2.38 implies that $\hat{f}(\boldsymbol{y} + \boldsymbol{\alpha})$ contains a monomial with support size at most $m - 1$. Therefore, by Lemma 9.10, we have

$$\hat{f}(\boldsymbol{y} + \boldsymbol{\alpha})(\boldsymbol{A}\boldsymbol{x}) \circ \mathcal{G}_{d(m-1)} \not\equiv 0 \,.$$

This implies the desired result. $\square$

9.3 Sum of ROFs

We begin by showing how to make the backbone of a structural RkF **0**-preserved.

Lemma 9.12. *Let* $F_1(\mathbf{y}), F_2(\mathbf{y}), \ldots, F_m(\mathbf{y}) \in \mathsf{RkF}$ *be structural, and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$*. Let* $\mathcal{H}$ *be a hitting-set generator for* $\mathsf{RkF}^{GL_n^{aff}(\mathbb{F})}$*. Then, there exists* $\boldsymbol{\alpha} \in (\mathbf{A}\mathbf{x} + \boldsymbol{\beta}) \circ \mathcal{H}$ *such that each* $F_i(\mathbf{y} + \boldsymbol{\alpha})$ *is* **0***-preserved for every* $i \in [m]$*.*

Proof. By Lemma 2.11, we have

$$P \coloneqq \left\{ \partial_{y_i^{\deg}} F_j \;\middle|\; j \in [m],\, y_i \in \mathrm{var}(F_j) \right\} \subseteq \mathsf{RkF} \setminus \{0\}.$$

Hence, by the properties of the hitting-set generator $\mathcal{H}$, for every $p \in P$ we have

$$p(\mathbf{A}\mathbf{x} + \boldsymbol{\beta}) \circ \mathcal{H} \not\equiv 0.$$

This implies that there exists $\boldsymbol{\alpha} \in \mathrm{Img}\big((\mathbf{A}\mathbf{x}+\boldsymbol{\beta}) \circ \mathcal{H}\big)$ that is a common nonzero for all polynomials in P. The claim then follows from Lemma 2.33. □

Theorem 1.5. *Let* $F_1(\mathbf{y}), F_2(\mathbf{y}), \ldots, F_k(\mathbf{y}) \in \mathsf{ROF}$*, and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$*. Then*

$$\sum_{i=1}^{k} p_{F_i}(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad \sum_{i=1}^{k} p_{F_i}(\mathbf{A}\mathbf{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{\log N + 3k} \not\equiv 0.$$

If $k = 2$ *then we can compose with* $\mathcal{G}_{\log N + 3}$ *instead of* $\mathcal{G}_{\log N + 6}$*.*

Proof. Since we can remove every zero term from the sum, we may assume every polynomial F_i is nonzero. For every $i \in [k]$, define $\hat{F}_i(\mathbf{y}) \coloneqq F_i(\mathbf{y} + \boldsymbol{\beta})$. Each $\hat{F}_i$ remains nonzero, since translation by $\boldsymbol{\beta}$ is an automorphism of $\mathbb{F}[\mathbf{y}]$. Moreover, such a translation does not increase the number of reads of any variable.

By Theorem 9.6, $\mathcal{G}_{\log N + 1}$ is a hitting set generator for orbits of ROFs. Hence, by Lemma 9.12, there exists $\boldsymbol{\alpha} \in (\mathbf{A}\mathbf{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{\log N + 1}$ such that every $p_{\hat{F}_i}(\mathbf{y} + \boldsymbol{\alpha})$ is **0**-preserved, and in particular, **0**-justified.

By Theorem 2.41, the sum $\sum_{i=1}^{k} \hat{F}_i$ is $3k$-hard. Therefore, by Observation 2.38, $\sum_{i=1}^{k} \hat{F}_i$ contains a monomial whose support size is at most $3k - 1$. Since each $\hat{F}_i$ is a ROF of individual degree at most 1, we can apply Lemma 9.10 to obtain

$$\sum_{i=1}^{k} p_{\hat{F}_i}(\mathbf{A}\mathbf{x} + \boldsymbol{\alpha}) \circ \mathcal{G}_{3k-1} \not\equiv 0.$$

As $\boldsymbol{\alpha} \in (\mathbf{A}\mathbf{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{\log N + 1}$, we get that

$$\sum_{i=1}^{k} p_{\hat{F}_i}(\mathbf{A}\mathbf{x}) \circ \mathcal{G}_{\log N + 3k} \not\equiv 0.$$

The claim follows since

$$\sum_{i=1}^{k} p_{\hat{F}_i}(\mathbf{A}\mathbf{x}) = \sum_{i=1}^{k} p_{F_i}(\mathbf{A}\mathbf{x} + \boldsymbol{\beta}).$$ □

For the special case of $\sum^2 \mathsf{ROF}$ using Lemma 2.46 instead of Theorem 2.41 gives the slight improvement mentioned.

9.4 Fragmentation Lemma for Orbits

In this subsection we prove a fragmentation lemma for orbits of RkFs. Its statement and proof closely follow those of Lemma 3.8.

Lemma 9.13. *Let* $F(\mathbf{y}) \in \mathsf{RkF}$*, and* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$ *with dual set* $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$*. Denote* $\mathcal{F}_F^{(2)} = \{\nu_j\}_{j=1}^t$*. For each* $j \in [t]$*, let* $y_{i_j} \in U^{(s)}(\nu_j)$ *and denote*

$$\mathcal{D}_I \coloneqq \left\{ \partial_{\delta_{i_j}^{\deg}}\left(p_{\nu_j}(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\right) : j = 1, \ldots, t \right\}.$$

Assume that $\mathcal{H}$ *is a hitting set generator for* $\mathcal{D}_I \cup \mathsf{R}(k-1)\mathsf{F}^{GL_n^{aff}(\mathbb{F})}$*. Then, if* $p_F(\mathbf{y}) \not\equiv 0$ *has individual degrees bounded by* $d \leqslant k$*, then*

$$p_F(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \left(\mathcal{H} + \mathcal{G}_{d(\log N+1)}\right) \not\equiv 0\,.$$

Proof. If $t = 0$, then $F \in \mathsf{R}(k-1)\mathsf{F}$, and the claim follows directly from the assumption on $\mathcal{H}$.

Case $t = 1$**.** Let $\mathcal{F}^{(2)} = \{\nu\}$ and $I = \{i\}$. By Lemma 9.3 and Definition 3.5 we have

$$\begin{aligned}\partial_{\delta_i^{\deg}}\left(p_F(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\right) &= \left(\partial_{y_i^{\deg}} p_F\right)(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \\ &= \left((\partial_{y_i^{\deg}} p_\nu)\cdot \prod_{j=1}^m p_{F_j}\right)(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \quad \text{or} \quad \partial_{\delta_i^{\deg}}\left(p_F(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\right) = \prod_{j=1}^m p_{F_j}(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\,,\end{aligned}$$

where each F_j is an $\mathsf{R}(k-1)\mathsf{F}$ and hence is hit by $\mathcal{H}$. Since $\partial_{y_i^{\deg}} p_\nu(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \mathcal{H} \not\equiv 0$, we get $\partial_{\delta_i^{\deg}}\left(p_F(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\right) \circ \mathcal{H} \not\equiv 0$. The claim then follows from Lemma 9.4.

Case $t > 1$**.** Proceed by induction on t. Let ν_1 be the gate guaranteed by Lemma 3.7. For $y_{j_1} \in U^{(s)}(\nu_1)$, we have

$$\partial_{y_{j_1}^{\deg}} p_F = (\partial_{y_{j_1}^{\deg}} p_{\nu_1})\cdot \prod_{j=1}^m p_{F_j} \quad \text{or} \quad \partial_{y_{j_1}^{\deg}} p_F = \prod_{j=1}^m p_{F_j}\,,$$

where each F_j is either an $\mathsf{R}(k-1)\mathsf{F}$, or an RkF satisfying $|\mathcal{F}_{F_j}^{(2)}| \leqslant t/2$.

If F_j is an $\mathsf{R}(k-1)\mathsf{F}$, then by the property of $\mathcal{H}$, $p_{F_j} \circ \mathcal{H} \not\equiv 0$. If $|\mathcal{F}_{F_j}^{(2)}| \leqslant t/2$, we get by the induction hypothesis,

$$p_{F_j}(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \left(\mathcal{H} + \mathcal{G}_{d(\log t/2+1)}\right) \not\equiv 0\,.$$

Therefore,

$$\prod_{j=1}^m p_{F_j}(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \left(\mathcal{H} + \mathcal{G}_{d\log t}\right) \not\equiv 0\,.$$

By Lemma 9.3 and the assumption on $\mathcal{H}$,

$$\left(\partial_{y_{j_1}^{\deg}} p_{\nu_1}\right)(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \mathcal{H} = \partial_{\delta_{j_1}^{\deg}}\left(p_{\nu_1}(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\right) \circ \mathcal{H} \not\equiv 0\,.$$

Thus,

$$\partial_{\delta_{j_1}^{\deg}}\left(p_F(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta})\right) \circ \left(\mathcal{H} + \mathcal{G}_{d\log t}\right) \overset{(*)}{=} \left(\partial_{y_{j_1}^{\deg}} p_F\right)(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \left(\mathcal{H} + \mathcal{G}_{d\log t}\right) \not\equiv 0\,,$$

where $(*)$ follows from Lemma 9.3. This, the fact that $\deg_{y_{j_1}}(F) \leqslant d$ and Lemma 9.4 imply that

$$p_F(\mathbf{A}\boldsymbol{x}+\boldsymbol{\beta}) \circ \left(\mathcal{H} + \mathcal{G}_{d\log t+d}\right) \not\equiv 0\,.$$

The claim then follows from the fact that $t \leqslant N$. □

Observation 9.14. *Let* $F(\mathbf{y}) \in \mathsf{RkF}$ *be structural, and* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$ *with dual set* $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$*. Let* $\mathcal{H}$ *be a hitting set generator that hits all polynomials that are structural* $\mathsf{RkF}^{GL_n^{aff}(\mathbb{F})}$ *of the form* $\sum^2 \mathsf{R}(k-1)\mathsf{F}^{GL_n^{aff}(\mathbb{F})}$*. Then, for every* $v \in \mathcal{F}_F^{(2)}$ *and every* $y_i \in U^{(s)}(v)$*, if* $p_F(\mathbf{y}) \not\equiv 0$ *then*

$$\partial_{\delta_i^{\deg}} p_v(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{H} \not\equiv 0\,.$$

The proof is analogous to that of Observation 4.1, and is therefore omitted.

9.5 Structurally Multilinear RkFs

In this subsection we use [AvMV15, Theorem 6.3] to obtain PIT for orbits of structurally multilinear RkFs.

Theorem 1.6. *Let* $F(\mathbf{y}) \in \mathsf{RkF}$ *be a structurally multilinear formula, and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$*. Then, for* $r_k := r_{2,2k} = k^{o(k)}$*,*

$$p_F(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{r_k + k\log N} \not\equiv 0\,.$$

Proof. Let $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$ be a dual set to $\mathbf{A}$. We proceed by induction on k. For $k = 1$, the argument follows from Theorem 9.6. Assume the statement holds for all smaller values of k.

Claim 9.15. *Let* $v \in \mathcal{F}_F^{(2)}$ *and* $y_i \in U(v)$*. We have*

$$\partial_{\delta_i^{\deg}}(p_v(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta})) \circ \mathcal{G}_{(r_{k-1} + (k-1)\log n} + (r_{2,2(k-1)} - 1)) \not\equiv 0\,.$$

Proof. If $v^{op} = \times$, then, since F is structurally multilinear, only one child of v depends on y_i. Assume without loss of generality this child is $(v)_L$, we have

$$\partial_{\delta_i}(p_v(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta})) = \partial_{\delta_i}\left(p_{(v)_L}(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta})\right) \cdot p_{(v)_R}(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta})\,.$$

By Lemma 9.8, and since $v \in \mathcal{F}^{(2)}$, both factors on the right-hand side are $\mathsf{R}(k-1)\mathrm{Ps}$. Hence, by the induction hypothesis, they are hit by $\mathcal{G}_{r_{k-1} + (k-1)\log n}$, so in this case, the claim holds.

Otherwise, $v^{op} = +$. By the definition of v as a gate in $\mathcal{F}_F^{(2)}$ and by Lemma 9.8, both $\partial_{\delta_i}(p_{(v)_L})$ and $\partial_{\delta_i}(p_{(v)_R})$ are computed by $\mathsf{R}(k-1)\mathsf{F}$ formulas. Let F_L and F_R denote the $\mathsf{R}(k-1)\mathsf{F}$ formulas computing these polynomials after translation by $\boldsymbol{\beta}$, i.e., $p_{F_L} = \partial_{\delta_i}(p_{(v)_L})(\mathbf{y} + \boldsymbol{\beta})$.

As substitutions do not increase the number of reads, we get by Lemma 2.11 that every $p \in \mathcal{Q}_{F_L, F_R}$ is computed by a structurally multilinear $\mathsf{R}(k-1)\mathsf{F}$. Hence, by the induction hypothesis and Observation 9.9, there exists a common nonzero $\boldsymbol{\alpha} \in \mathrm{Img}\left(\mathbf{A}\boldsymbol{x} \circ \mathcal{G}_{r_{k-1} + (k-1)\log n}\right)$ for all polynomials in $\mathcal{Q}_{F_L, F_R}$. Lemma 2.43 implies that

$$\partial_{y_i} p_v(\mathbf{y} + \boldsymbol{\beta} + \boldsymbol{\alpha}) = \left(p_{F_L} + p_{F_R}\right)(\mathbf{y} + \boldsymbol{\alpha})$$

is $r_{2,2(k-1)}$-hard, hence by Observation 2.38 contains a monomial whose support size is at most $r_{2,2(k-1)} - 1$. This, the fact that $\partial_{y_i} p_v$ is multilinear and Lemma 9.10 imply that

$$\partial_{y_i} p_v(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta} + \boldsymbol{\alpha}) \circ \mathcal{G}_{r_{2,2(k-1)} - 1} \not\equiv 0\,.$$

The claim then follows from the basic properties of independent polynomial maps. □

As p_F is multilinear, we get from Lemma 9.13 and Claim 9.15 that

$$p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{(r_{k-1} + k\log n + r_{2,2(k-1)})} \not\equiv 0\,.$$

The claim follows since $r_k \geqslant 2r_{k-1}$. □

9.6 PIT for Orbits of Read-2/3/4 Formulas

We follow the original proofs given in previous sections, applying lemmas Lemma 9.10 and Lemma 9.12 whenever needed.

Theorem 1.7. *Let* $F(\mathbf{y}) \in \mathsf{R2F}$ *and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. *Then*

$$p_F(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{3\log N+5} \not\equiv 0.$$

Proof. Let $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$ be a dual set to $\mathbf{A}$.

Since F is structural (Corollary 2.6) we get from Observation 9.14 (using Theorem 1.5) that for every $v \in \mathcal{F}^{(2)}$ and every $i \in [N]$ with $y_i \in U(F_v)$ we have

$$\partial_{\delta_i^{\deg}} p_v(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{\log N+3} \not\equiv 0.$$

The claim follows from Lemma 9.13 and the fact that each individual degree in an $\mathsf{R2F}$, is at most 2. □

Claim 9.16. *Let* $F(\mathbf{y}) \in \mathsf{R3F}$ *be of the form* $\sum^2 \mathsf{R2F}$. *Let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$ *with dual set* $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$. *If* $p_F(\mathbf{y}) \not\equiv 0$ *then*

$$p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{3\log N+2r_{2,4}+3} \not\equiv 0\,.$$

Proof. Let $\hat{F}(\mathbf{y}) = F(\mathbf{y} + \boldsymbol{\beta})$. This transformation maintains the nonzeroness of $\hat{F}$ and the number of reads with respect to each variable.

Corollary 2.6 and Lemma 2.11 imply that $\mathcal{Q}_{F(o_{\hat{F}})_L, F(o_{\hat{F}})_R} \subseteq \mathsf{R2F}$. Hence, by Observation 9.9 and Theorem 1.7, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\mathbf{A}\boldsymbol{x} \circ \mathcal{G}_{3\log N+5})$ that is a common nonzero for all polynomials in $\mathcal{Q}_{F(o_{\hat{F}})_L, F(o_{\hat{F}})_R}$. Consequently, Proposition 4.2 implies that $p_{\hat{F}}(\mathbf{y} + \boldsymbol{\alpha})$ is $r_{2,4}$-hard. By the structure of $\hat{F}$, its maximal individual degree is at most 2. Thus, from Observation 2.38 and Lemma 9.10 we obtain

$$p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta} + \boldsymbol{\alpha}) \circ \mathcal{G}_{2(r_{2,4}-1)} = p_{\hat{F}}(\mathbf{A}\boldsymbol{x} + \boldsymbol{\alpha}) \circ \mathcal{G}_{2(r_{2,4}-1)} \not\equiv 0\,.$$

The claim follows from Fact 2.16. □

Theorem 1.8. *Let* $F(\mathbf{y}) \in \mathsf{R3F}$ *and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. *Then, there exists a constant* $c_{1.8}$

$$p_F(\mathbf{y}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{6\log N+c_{1.8}} \not\equiv 0\,.$$

Proof. Let $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$ be the dual set of $\mathbf{A}$. By Corollary 2.6, the formula F is structural. Combining Observation 9.14 with Claim 9.16, we obtain that for every $v \in \mathcal{F}_F^{(2)}$ and every $y_i \in U(F_v)$,

$$\partial_{\delta_i^{\deg}} p_v(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{3\log N+2r_{2,4}+3} \not\equiv 0\,.$$

Since $F \in \mathsf{R3F}$, we have $\deg_{y_i}(p_F) \leqslant 3$ for all $y_i \in \mathbf{y}$. This bound, together with Lemma 9.13 and Observation 3.6, yields the claim. □

Theorem 1.9. *Let* $F(\mathbf{y}) \in \mathsf{R4F}$ *and let* $(\mathbf{A}, \boldsymbol{\beta}) \in GL_n^{aff}(\mathbb{F})$. *Then, there exists a constant* $c_{1.9}$

$$p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \not\equiv 0 \quad \Rightarrow \quad p_F(\mathbf{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{12\log N+c_{1.9}} \not\equiv 0\,.$$

Proof. Let $\mathcal{D} = \{\delta_1, \delta_2, \ldots, \delta_n\} \subset \mathbb{F}^n$ be the dual set corresponding to $\mathbf{A}$ and let $\mathcal{H}$ be a hitting-set generator for orbits of R3Fs and for orbits of structurally multilinear R8Fs. Set $\hat{F} = F(\mathbf{y} + \boldsymbol{\beta})$, this translation preserves both the nonzeroness of F and the read count of each variable. We proceed by following the structure of the proof for Theorem 1.1.

Let $\nu \in \mathcal{F}^{(2)}_{\hat{f}}$.

If $\nu^{op} = \times$, let $x_t \in U(\nu)$, by Claim 8.2 we get that for any hitting-set generator $\mathcal{H}'$ for R3Fs

$$\partial_{y_t^{deg}} p_\nu(\mathbf{y}) \circ \mathcal{H}' \not\equiv 0 .$$

Hence, by Lemma 9.3, and Observation 9.9 we get that

$$\partial_{\delta_t^{deg}} p_\nu(\mathbf{A}\mathbf{x}) \circ \mathcal{H} \not\equiv 0 .$$

Next, we assume $\nu^{op} = +$ and proceed by case analysis.

1. **There exists $x_t \in \mathbf{var}(\nu)$ such that, for some $G \in \{F_{(\nu)_L}, F_{(\nu)_R}\}$, we have $\mathbf{Read}_G(x_t) = 3$ and $\deg_{x_t}(G) > 1$.** By Claim 8.3, and Claim 8.4, there exists $y_t \in U^{(s)}(\nu)$ such that for any hitting-set generator $\mathcal{H}'$ for R3Fs,

   $$\partial_{y_t^{deg}} p_\nu(\mathbf{y}) \circ \mathcal{H}' \not\equiv 0 .$$

   Hence, by Lemma 9.3 and Observation 9.9,

   $$\partial_{\delta_t^{deg}} p_\nu(\mathbf{A}\mathbf{x}) \circ \mathcal{H} = \left(\partial_{y_t^{deg}} p_\nu\right)(\mathbf{A}\mathbf{x}) \circ \mathcal{H} \not\equiv 0 .$$

2. **No x_t as in the first case exists and $|U(\nu)| \geqslant 23$.** By Claim 8.7 and Observation 9.9 there exist $y_t \in U(\nu)$ and $\boldsymbol{\alpha} \in \mathrm{Img}\big((\mathbf{A}\mathbf{x}) \circ \mathcal{H}\big)$ such that $\partial_{y_t^{deg}} p_\nu(\mathbf{y} + \boldsymbol{\alpha})$ is $(r_{2,6} + 5^8 + 23)$-hard. From the fact that $F \in$ R4F, $\nu^{op} = +$ and a variable as in the first case does not exists, we get that the individual degree of p_ν with respect to any variable is at most two. Therefore, by Observation 2.38 and Lemma 9.10 we get that

   $$\partial_{y_t^{deg}} p_\nu(\mathbf{y} + \boldsymbol{\alpha}) \circ \mathcal{G}_{2(r_{2,6}+5^8+23)} \not\equiv 0 .$$

   Hence, by Lemma 9.3

   $$\partial_{\delta_t^{deg}} p_\nu(\mathbf{A}\mathbf{x}) \circ \left(\mathcal{H} + \mathcal{G}_{2(r_{2,6}+5^8+23)}\right) = \left(\partial_{y_t^{deg}} p_\nu\right)(\mathbf{A}\mathbf{x}) \circ \left(\mathcal{H} + \mathcal{G}_{2(r_{2,6}+5^8+23)}\right) \not\equiv 0 .$$

3. **No x_t as in the first case exists and $|U(\nu)| < 23$.** Since $\nu \in \mathcal{F}^{(2)}_{\hat{f}}$, we have $U(\nu) \neq \emptyset$. Let $y_t \in U(\nu)$ and define $S = U(\nu) \setminus \{y_t\}$ as in Claim 8.9. By the bound on the size of $U(\nu)$, we have $|S| \leqslant 21$. By Lemma 9.5, there exist an assignment $\boldsymbol{\alpha} \in \mathbb{F}^{21}$, linear functions $\mathbf{L}(\mathbf{x}) = L_1(\mathbf{x}), L_2(\mathbf{x}), \ldots, L_{21}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and $(\tilde{\mathbf{A}}, \tilde{\boldsymbol{\beta}}) \in \mathrm{GL}^{\mathrm{aff}}_n(\mathbb{F})$ such that

   $$\begin{aligned}
   \partial_{\delta_t^{deg}} p_\nu(\mathbf{A}(\mathbf{x} + \mathcal{G}_{21}(\boldsymbol{\alpha}, \mathbf{L}(\mathbf{x})))) &=^{(1)} \left(\partial_{y_t^{deg}} p_\nu\right)(\mathbf{A}(\mathbf{x} + \mathcal{G}_{21}(\boldsymbol{\alpha}, \mathbf{L}(\mathbf{x})))) \\
   &=^{(2)} \left(\partial_{y_t^{deg}} p_\nu\right)\big|_S(\tilde{\mathbf{A}}\mathbf{x} + \tilde{\boldsymbol{\beta}}) \\
   &=^{(3)} \left(\partial_{y_t^{deg}} (p_\nu|_S)\right)(\tilde{\mathbf{A}}\mathbf{x} + \tilde{\boldsymbol{\beta}}) ,
   \end{aligned}$$

where (1) follows from Lemma 9.3, (2) from Lemma 9.5, and (3) since $y_t \in S$.

Translating p_ν by $\tilde{\boldsymbol{\beta}}$ does not change the situation, i.e., the number of occurrences of any variable or its individual degree. By the definition of $\mathcal{H}$ and Observation 9.9, $\tilde{\mathbf{A}} \circ \mathcal{H}$ is a hitting-set generator for R3Fs and for structurally multilinear R8Fs. Hence, by Claim 8.9, there exists $\boldsymbol{\alpha} \in \mathrm{Img}(\tilde{\mathbf{A}} \circ \mathcal{H})$ such that $\left(\partial_{y_t^{\deg}(p_\nu|_S)}\right)(\mathbf{y} + \tilde{\boldsymbol{\beta}} + \boldsymbol{\sigma})$ is 2-hard. Since the individual degree of each variable in p_ν is at most 2, by Observation 2.38 and Lemma 9.10 we get

$$\left(\partial_{y_t^{\deg}(p_\nu|_S)}\right)(\tilde{\mathbf{A}}\boldsymbol{x} + \tilde{\boldsymbol{\beta}} + \boldsymbol{\sigma}) \circ \mathcal{G}_4 \not\equiv 0\,.$$

Overall, we obtain

$$\partial_{\delta_t^{\deg}} p_\nu(\boldsymbol{A}\boldsymbol{x}) \circ (\mathcal{H} + \mathcal{G}_{25}) \not\equiv 0\,.$$

To conclude, by Theorem 1.6 $\mathcal{G}_{8\log N + r_8}$ is a hitting-set generator for structurally multilinear R8Fs and by Theorem 1.8 there exists a constant c_1 such that $\mathcal{G}_{6\log N + c_1}$ is a hitting set generator for R3Fs. Hence we may take $\mathcal{H} = \mathcal{G}_{8\log N + r_8 + c_1}$. By the preceding case analysis, for every $\nu \in \mathcal{F}_{\hat{F}}^{(2)}$ there exists $t \in U^{(s)}(\nu)$ such that

$$\partial_{\delta_t^{\deg}} p_\nu(\boldsymbol{A}\boldsymbol{x}) \circ (\mathcal{H} + \mathcal{G}_{2(r_{2,6}+5^8+23)}) = \mathcal{G}_{8\log N + r_8 + c_1 + 2(r_{2,6}+5^8+23)} \not\equiv 0\,.$$

By Lemma 9.13, and since the individual degrees of $\hat{F}$ are at most 4, we have for the constant $c = r_8 + c_1 + 2(r_{2,6} + 5^8 + 23) + 4$

$$p_F(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}) \circ \mathcal{G}_{12\log N + c} = p_{\hat{F}}(\boldsymbol{A}\boldsymbol{x}) \circ \mathcal{G}_{12\log N + c} \not\equiv 0\,. \qquad \square$$

Bibliography

[AB03] Manindra Agrawal and Somenath Biswas. Primality and identity testing via chinese remaindering. *J. ACM*, 50(4):429–443, 2003. 1

[AFS+18] Matthew Anderson, Michael A. Forbes, Ramprasad Saptharishi, Amir Shpilka, and Ben Lee Volk. Identity testing and lower bounds for read-*k* oblivious algebraic branching programs. *ACM Trans. Comput. Theory*, 10(1):3:1–3:30, 2018. 2, 6

[Agr05] Manindra Agrawal. Proving lower bounds via pseudo-random generators. In Sundar Sarukkai and Sandeep Sen, editors, *FSTTCS 2005: Foundations of Software Technology and Theoretical Computer Science*, pages 92–105, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg. 1

[AKS04] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. Primes is in p. *Annals of Mathematics*, 160(2):781–793, September 2004. 1

[And20] Robert Andrews. Algebraic hardness versus randomness in low characteristic. In Shubhangi Saraf, editor, *35th Computational Complexity Conference, CCC 2020, July 28-31, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 169 of *LIPIcs*, pages 37:1–37:32. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 1

[AV08] Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, Philadelphia, PA, USA, October 25-28, 2008*, pages 67–75. IEEE Computer Society, 2008. 1

[AvMV15] Matthew Anderson, Dieter van Melkebeek, and Ilya Volkovich. Deterministic polynomial identity tests for multilinear bounded-read formulae. *Computational Complexity*, 24(4):695–776, Dec 2015. 2, 3, 5, 6, 7, 8, 10, 19, 20, 24, 72, 82

[BG21] Vishwas Bhargava and Sumanta Ghosh. Improved hitting set for orbit of roabps. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021)*, pages 30–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2021. 6

[BGV23] Pranav Bisht, Nikhil Gupta, and Ilya Volkovich. Towards Identity Testing for Sums of Products of Read-Once and Multilinear Bounded-Read Formulae. In Patricia Bouyer and Srikanth Srinivasan, editors, *43rd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2023)*, volume 284 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 9:1–9:23, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 3, 5, 8, 10, 16, 17, 19, 20, 42, 47

[BHH95] Nader Bshouty, Thomas Hancock, and Lisa Hellerstein. Learning arithmetic read-once formulas. *SIAM J. Comput.*, 24:706–735, 01 1995. 16

[BIZ18] Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam. On algebraic branching programs of small width. *J. ACM*, 65(5), August 2018. 6

[CKS19] Chi-Ning Chou, Mrinal Kumar, and Noam Solomon. Closure results for polynomial factorization. *Theory of Computing*, 15(1):1–34, 2019. 1

[CLO15] D.A. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Springer Cham, 2015. 18

[DG24] Pranjal Dutta and Sumanta Ghosh. Sigact news complexity theory column 121. *SIGACT News*, 55(2):53–88, June 2024. 1, 5, 6

[DL78] Richard A. Demillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, June 1978. 1

[DSY10] Zeev Dvir, Amir Shpilka, and Amir Yehudayoff. Hardness-randomness tradeoffs for bounded depth arithmetic circuits. *SIAM Journal on Computing*, 39(4):1279–1293, 2010. 1

[ES35] Pál Erdős and George Szekeres. A combinatorial problem in geometry. *Compos. Math.*, 2:463–470, 1935. 34

[FGT19] Stephen Fenner, Rohit Gurjar, and Thomas Thierauf. A deterministic parallel algorithm for bipartite perfect matching. *Communications of the ACM*, 62(3):109–115, 2019. 1

[For14] Michael Andrew Forbes. *Polynomial identity testing of read-once oblivious algebraic branching programs*. PhD thesis, Massachusetts Institute of Technology, 2014. 6

[FS12] Michael A. Forbes and Amir Shpilka. On identity testing of tensors, low-rank recovery and compressed sensing. In Howard J. Karloff and Toniann Pitassi, editors, *Proceedings*

of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012, pages 163–172. ACM, 2012. 5

[FS13] Michael A Forbes and Amir Shpilka. Explicit noether normalization for simultaneous conjugation via polynomial identity testing. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 527–542. Springer, 2013. 1

[FS18] Michael A. Forbes and Amir Shpilka. A pspace construction of a hitting set for the closure of small algebraic circuits. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, page 1180–1192, New York, NY, USA, 2018. Association for Computing Machinery. 6

[GCL92] K.O. Geddes, S.R. Czapor, and G. Labahan. *Algorithms for Computer Algebra*. Kluwer, 1992. 18

[GKKS13] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic Circuits: A Chasm at Depth Three . In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 578–587, Los Alamitos, CA, USA, October 2013. IEEE Computer Society. 1

[GKSS22] Zeyu Guo, Mrinal Kumar, Ramprasad Saptharishi, and Noam Solomon. Derandomization from Algebraic Hardness. *SIAM J. Comput.*, 51(2):315–335, 2022. 1

[GKST17] Rohit Gurjar, Arpita Korwar, Nitin Saxena, and Thomas Thierauf. Deterministic identity testing for sum of read-once oblivious arithmetic branching programs. *Comput. Complex.*, 26(4):835–880, 2017. 5

[GT17] Rohit Gurjar and Thomas Thierauf. Linear matroid intersection is in quasi-nc. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 821–830, 2017. 1

[HH91] T. R. Hancock and L. Hellerstein. Learning read-once formulas over fields and extended bases. In *Proceedings of the 4th Annual Workshop on Computational Learning Theory (COLT)*, pages 326–336, 1991. 16

[HMM24] Ivan Hu, Dieter van Melkebeek, and Andrew Morgan. Polynomial identity testing via evaluation of rational functions. *Theory of Computing*, 20(1):1–70, 2024. 2, 15, 36

[HS80] J. Heintz and C. P. Schnorr. Testing polynomials which are easy to compute (extended abstract). In *Proceedings of the Twelfth Annual ACM Symposium on Theory of Computing*, STOC '80, page 262–272, New York, NY, USA, 1980. Association for Computing Machinery. 1

[Kal95] Erich Kaltofen. Effective Noether irreducibility forms and applications. volume 50, pages 274–295. 1995. 23rd Symposium on the Theory of Computing (New Orleans, LA, 1991). 21

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

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012. 1

[KS19] Mrinal Kumar and Ramprasad Saptharishi. Hardness-randomness tradeoffs for algebraic computation. *Bulletin of EATCS*, 3(129), 2019. 1

[KST23] Mrinal Kumar, Ramprasad Saptharishi, and Anamay Tengse. Near-Optimal Bootstrapping of Hitting Sets for Algebraic Models. *Theory Comput.*, 19:1–30, 2023. 1

[LFKN92] Carsten Lund, Lance Fortnow, Howard Karloff, and Noam Nisan. Algebraic methods for interactive proof systems. *J. ACM*, 39(4):859–868, October 1992. 1

[Lov79] László Lovász. On determinants, matchings and random algorithms. volume 79, pages 565–574, 01 1979. 1

[MRS14] Meena Mahajan, B.V. Raghavendra Rao, and Karteek Sreenivasaiah. Monomials, multilinearity and identity testing in simple read-restricted circuits. *Theoretical Computer Science*, 524:90–102, 2014. 2, 4, 10, 72

[MRS16] Meena Mahajan, B. V. Raghavendra Rao, and Karteek Sreenivasaiah. Building above read-once polynomials: Identity testing and hardness of representation. *Algorithmica*, 76(4):890–909, Dec 2016. 5

[MS21] Dori Medini and Amir Shpilka. Hitting sets and reconstruction for dense orbits in $\mathrm{VP_e}$ and ΣΠΣ circuits. *CoRR*, abs/2102.05632, 2021. 6, 77, 78

[Mul17] Ketan Mulmuley. Geometric complexity theory v: Efficient algorithms for noether normalization. *Journal of the American Mathematical Society*, 30(1):225–309, 2017. 1

[MV18] Daniel Minahan and Ilya Volkovich. Complete derandomization of identity testing and reconstruction of read-once formulas. *ACM Trans. Comput. Theory*, 10(3), may 2018. 6, 15, 16, 26

[Ore22] Øystein Ore. Über höhere kongruenzen. *Norske Videnskaps-Akademi i Oslo. Forhandlinger (Proceedings of the Norwegian Academy of Science and Letters)*, 1922(12):1–8, 1922. Contains the original root bound for nonzero multivariate polynomials over finite fields, later used in polynomial identity testing (Schwartz–Zippel lemma). 1

[Pra19] Gautam Prakriya. Derandomizing isolation and polynomial identity testing. Doctoral thesis, 2019. 2, 3, 4

[RS05] Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Comput. Complex.*, 14(1):1–19, 2005. 5

[Sap21] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. 2021. 1, 2

[Sax09] Nitin Saxena. Progress on Polynomial Identity Testing. *Bull. EATCS*, 99:49–79, 2009. 1, 6

[Sax14] Nitin Saxena. Progress on polynomial identity testing-II. *Perspectives in Computational Complexity: The Somenath Biswas Anniversary Volume*, pages 131–146, 2014. 1, 6

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, October 1980. 1

[Sha92] Adi Shamir. IP = PSPACE. *J. ACM*, 39(4):869–877, 1992. 1

[Sha22] Nadav Shamir. Polynomial identity testing for read-2 formulas and linear separation between read-3 and read-2 multilinear formulas. Master thesis, 2022. 2, 3, 5, 7, 13, 14, 16, 19, 24

[ST17] Ola Svensson and Jakub Tarnawski. The Matching Problem in General Graphs Is in Quasi-NC. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 696–707. IEEE Computer Society, 2017. 1

[ST24] Chandan Saha and Bhargav Thankey. Hitting sets for orbits of circuit classes and polynomial families. *ACM Trans. Comput. Theory*, 16(3), September 2024. 6

[SV10] Amir Shpilka and Ilya Volkovich. On the relation between polynomial identity testing and finding variable disjoint factors. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors, *Automata, Languages and Programming*, pages 408–419, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. 16, 17

[SV14] Amir Shpilka and Ilya Volkovich. On reconstruction and testing of read-once formulas. *Theory of Computing*, 10(18):465–514, 2014. 1, 14, 16, 17

[SV15] Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. *Comput. Complex.*, 24(3):477–532, September 2015. 1, 2, 3, 4, 5, 6, 8, 15, 16, 19, 20, 26, 54

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Found. Trends Theor. Comput. Sci.*, 5(3–4):207–388, March 2010. 1, 2, 6

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Inf. Comput.*, 240:2–11, 2015. 1

[vzGG99] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, 1999. 18

[Zip79] Richard Zippel. *Probabilistic algorithms for sparse polynomials*, page 216–226. Springer Berlin Heidelberg, 1979. 1

A Missing proofs from Section 2

Proof of Lemma 2.11. Let $t = \mathrm{Read}_{x_i}(F)$. We prove the lemma by induction on t. For $t = 1$, the lemma follows from Lemma 2.8. For $t > 1$, Let $u \in F$ be the fcg of to all the leaves labeled by x_i in F. By Lemma 2.8 we have

$$\partial_{x_i^{\deg}} p_F = \partial_{x_i^{\deg}}(p_u) \prod_{v \in \mathrm{Unv}(u)} p_v \,.$$

Let $d = \deg_{x_i}(u)$, $d_L = \deg_{x_i}((u)_L)$ and $d_R = \deg_{x_i}((u)_R)$.

If $u^{op} = +$: We get by additivity

$$\partial_{x_i^{\deg}} p_u = \partial_{x_i^{d}}\left(p_{(u)_L}\right) + \partial_{x_i^{d}}\left(p_{(u)_R}\right).$$

Since x_i is structural in F, we get that $d_L, d_R \leqslant d$. Hence, we can apply the induction hypothesis to $\partial_{x_i^d}(p_{(u)_L})$ and $\partial_{x_i^d}(p_{(u)_R})$. These are either zero, or composed of disjoint subformulas of themselves. Since these subformulas are mutually disjoint and are disjoint of the formulas rooted in gates in $\mathrm{Unv}(u)$, the claim follows.

If $u^{op} = \times$: In this case, we have $d_L + d_R = d$. By the generalized multiplication rule,

$$\partial_{x_i^{\deg}} p_u = \sum_{j_L + j_R = d} \binom{d}{j_L} \partial_{x_i^{j_L}}(p_{(u)_L}) \cdot \partial_{x_i^{j_R}}(p_{(u)_R}) \sim \partial_{x_i^{d_L}}(p_{(u)_L}) \cdot \partial_{x_i^{d_R}}(p_{(u)_R}) ,$$

where the last equality holds since in every other summand, one of the multiplicands is zero. The claim then follows in the same way as in the previous case. □

A.1 Proofs for Subsubsection 2.7.1

In this section we give observations and missing proofs for Subsubsection 2.7.1. To ease the reading we repeat the statements of the relevant claims.

Observation A.1. *Let F be an algebraic formula and $S \subseteq [n]$. Then every $v \in F$ is either*

1. *reduced to a constant,*
2. *causes the creation of the gate $\mathcal{V}_S(v)$ in $F|_S$, or*
3. *affects the constants of $\mathcal{V}_S(v)$, which is the gate that was created by it's closest descendant that created a gate.*

Proof. Every gate $v \in F$ is called once and is assigned one of three cases. The three items are exactly the three cases described in Definition 2.54. □

Observation A.2. *Let F be an algebraic formula, $S \subseteq [n]$ and $v, u \in F$ be such that $\mathcal{V}_S(v) = \mathcal{V}_S(u) \neq \bot$. Then u and v are not disjoint.*

Proof. Let $v \in F$, the only case in Definition 2.54 where we set $\mathcal{V}_S(v)$ to a some different gate's restricted gate, is in Case (2.54). It can be observe that this gate is only one of v's children. □

Claim A.3. *Let F be an algebraic formula, let $S \subseteq [n]$ and let $\boldsymbol{\tau} \in \mathbb{F}^n$ be a generic assignment. Denote $I = [n] \setminus S$. Then,*

$$p_{F|_S^\tau} = p_F|_{I \leftarrow \boldsymbol{\tau}_I} .$$

Proof of Claim A.3. The proof is by induction on $|F|$. Let $\alpha, \beta \in \mathbb{F}$ be the multiplicative and additive constants of o.

If $|F| = 1$, then, for some $x_i \in \mathbf{x}$, we have $p_F = \alpha x_i + \beta$. If $i \in I$ then

$$p_F|_{I \leftarrow \boldsymbol{\tau}_I} = \alpha \boldsymbol{\tau}_i + \beta = p_{F|_S^\tau} .$$

Otherwise,

$$p_F|_{I \leftarrow \boldsymbol{\tau}_I} = \alpha x_i + \beta = p_{F|_S^\tau} .$$

Assume the claim holds for smaller sized formulas. In the following equations, we label equalities with $(*)$ to note these are implied by the induction hypothesis. If $p_{(o)_L}|_{I \leftarrow \boldsymbol{\tau}_I} \notin \mathbb{F}$ and $p_{(o)_R}|_{I \leftarrow \boldsymbol{\tau}_I} \notin \mathbb{F}$ then:

$$p_F|_{I \leftarrow \boldsymbol{\tau}_I} = \alpha \left(p_{(o)_L}|_{I \leftarrow \boldsymbol{\tau}_I} \mathbin{op} p_{(o)_R}|_{I \leftarrow \boldsymbol{\tau}_I} \right) + \beta =^* \alpha \left(p_{F_{(o)_L}|_S^\tau} \mathbin{op} p_{F_{(o)_R}|_S^\tau} \right) = p_{F|_S^\tau} .$$

Otherwise, assume $\gamma \coloneqq p_{(o)_L}|_{I\leftarrow\boldsymbol{\tau}_I} \in \mathbb{F}$.
If $o^{op} = +$:

$$\begin{aligned} p_F|_{I\leftarrow\boldsymbol{\tau}_I} &= \alpha\left(p_{(o)_L}|_{I\leftarrow\boldsymbol{\tau}_I} + p_{(o)_R}|_{I\leftarrow\boldsymbol{\tau}_I}\right) + \beta = \alpha\gamma + \alpha p_{(o)_R}|_{I\leftarrow\boldsymbol{\tau}_I} + \beta \\ &=^{*} \alpha\gamma + \alpha p_{F_{(o)_R}|_S^{\tau}} + \beta = p_{F|_S^{\tau}}\,. \end{aligned}$$

Otherwise $o^{op} = \times$:

$$\begin{aligned} p_F|_{I\leftarrow\boldsymbol{\tau}_I} &= \alpha\left(p_{(o)_L}|_{I\leftarrow\boldsymbol{\tau}_I} p_{(o)_R}|_{I\leftarrow\boldsymbol{\tau}_I}\right) + \beta = \alpha\gamma p_{(o)_R}|_{I\leftarrow\boldsymbol{\tau}_I} + \beta \\ &=^{*} \alpha\gamma p_{F_{(o)_R}|_S^{\tau}} + \beta = p_{F|_S^{\tau}}\,. \end{aligned}$$

□

Claim A.4. *The three functions $\mathcal{V}_S, \mathcal{O}_S$ and $\mathcal{L}_S$ are independent of the generic assignment $\boldsymbol{\tau}$.*

Proof. It is enough to prove this for $\mathcal{V}_S$, since the other two functions are only dependent on $\mathcal{V}_S$ and the formula the procedure is applied on.

$\mathcal{V}_S$ is independent of $\boldsymbol{\tau}$ by the property of generic assignments which is that a polynomial $p \in \mathbb{F}[\boldsymbol{x}]$ is restricted to a constant by a generic assignment to the set of variables S if and only if $\mathrm{var}(p) \subseteq S$. □

Observation A.5. *Let F be an algebraic formula, $S \subseteq [n]$ and $v \in F|_S$. And denote $u = \mathcal{O}_S(v)$ and $w = \mathcal{L}_S(v)$. Then,*

- *$\mathcal{D}_S(v) > 0$ iff $u \neq w$ iff $p_{(u)_L}|_S \in \mathbb{F}$ or $p_{(u)_R}|_S \in \mathbb{F}$ and the restricted polynomial of the other child is not a constant.*
- *$\mathcal{D}_S(v) = 0$ iff $u = w$ iff $p_{(u)_L}|_S \notin \mathbb{F}$ and $p_{(u)_R}|_S \in \mathbb{F}$.*

Proof. The left assertions in both items follows from the definition of $\mathcal{D}_S$. We move to handle the right assertions.

By Observation A.2, u is an ancestor of w. The value of $\mathcal{V}_S(u) \notin \{\bot, o_{F_u|_S}\}$ iff Case (2.54) is reached and this case is reached iff one of u's children is reduced to a constant by the restriction and the other one is not.

The second claim is implied by the fact that if Case (2.54) it does not result in a gate in $F|_S$. □

Observation A.6. *Let F be an algebraic formula, $S \subseteq [n]$, $v \in F|_S$ and $w = \mathcal{L}_S(v)$. Then*

1. *none of the children of w reduce to a constant by the restriction, and*
2. *v is created when the procedure handles w.*

Proof. Assume towards contradiction that any of the children of w is reduced to a constant by the restriction. Then, when the procedure arrives at w, we're either in Case (2.54), or we're in Case (2.54).

The former case is impossible, since this contradicts the fact that $\mathcal{V}_S(w) = v$. The latter case contradicts the definition of w as the deepest gate in the restriction path of v. This prove item 1 and moreover, implies that when the procedure is called on w, it enters Case (2.54).

Hence, by Observation A.1, v is created when the procedure handles w. □

Observation 2.56. *Let F be an algebraic formula, $S \subseteq [n]$, and $\boldsymbol{\tau} \in \mathbb{F}^S$. Let $v \in F|_S$ and $w = \mathcal{O}_S(v)$. Then*

1. *w is either o_F or the parent of w creates a new gate in $F|_S$,*

2. *w is the last gate that changes the constants in v, and*

3. $p_{F_w|_S^\tau} = p_v$.

Proof. If $w = o$, then since this is the last gate handled by the recursive procedure, both claims are trivial.

Otherwise, let u be the parent of w. Observe, that since $\mathcal{V}_S(w) \neq \perp$, it must be the case that $p_w|_S \notin \mathbb{F}$. Therefore, it must be the case that the procedure when called on u, enters either Case (2.54) or Case (2.54). Since $\mathcal{V}_S(u) \neq \mathcal{V}_S(w)$ the correct case is (2.54), this proves item 1.

By Observation A.1, the only gates which may change the constants after w are those in the path from w to o. By item 3, and the fact that u creates a new gate, they may only change the constants of $\mathcal{V}_S(u)$. This proves item 2.

By Claim A.3 we get that $p_{F_w|_S} = p_{F_w}|_S$. Now, item 2 implies that $p_{F_w|_S} = p_v$, item 3 follows from these two equalities. □

Lemma 2.57. *Let F be an algebraic formula, $S \subseteq [n]$ and $\boldsymbol{\tau} \in \mathbb{F}^S$. Let $v \in F|_S^\tau$, $u = \mathcal{O}_S(v)$ and $w = \mathcal{L}_S(v)$. Denote by γ the additive constant of v and by δ the additive constant of w.*

1. *Let $G = F_u|_{w=\delta}$. Then $\gamma = p_{G|_S^\tau}$.*

2. $\left(F|_S^\tau\right)|_{v=\gamma} = (F|_{w=\delta})|_S^\tau$.

Proof. We prove item 1 by induction on the restriction depth of v. Let $* = u^{op}$ and let $\alpha \in \mathbb{F}$ be u's multiplicative constant.

If $\mathcal{D}_S(v) = 0$, then $w = u$. By item 2, v was created when the procedure visited u, and by item 2 and Observation A.1 no gates other than u affect the additive constant of v. Therefore, by definition, $F_v = \alpha\left(F_{(u)_L}|_S * F_{(u)_R}|_S\right) + \delta$ and therefore $\gamma = \delta = p_{G|_S}$, as claimed.

Assume $\mathcal{D}_S(v) > 0$ and that the claim holds for any gate with smaller restriction depth. By Observation A.5 we may assume without loss of generality that $(u)_L$ reduces to a constant $\eta \in \mathbb{F}$ by the restriction. Hence w is a descendant of $(u)_R$. Let $\beta \in \mathbb{F}$ be the additive constant of u.

By Definition 2.54 $F_v = \alpha\eta * F_{(u)_R}|_S + \beta$. Hence, if δ' is the additive constant when the procedure arrived at $(u)_R$, then $\gamma = \alpha\eta * \delta' + \beta$.

By the definition of G, $F_{(u)_R}|_{w=\delta} = F_{(o_G)_R}$. By the induction hypothesis, $\delta' = p_{F_{(o_G)_R}|_S}$. Since by the construction of G, $F_{(o_G)_L} = F_{(u)_L}$, we then get

$$p_{G|_S} = \alpha(F_{(u)_L}|_S * F_{(o_G)_R}|_S) + \beta = \alpha\eta * \delta' + \beta = \gamma$$

and the induction is complete. This proves item 1.

Now, we have

$$(F|_S)|_{v=\gamma} = (F|_S)|_{v=(F_u|_{w=\delta})|_S} = (F|_{w=\delta})|_S$$

where the first equality follows by item 1 and the second equality follows from item 3. □

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il