Hitting Sets for Orbits of Circuit Classes and Polynomial Families

Chandan Saha
Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India

Bhargav Thankey
Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India

Abstract

The orbit of an n-variate polynomial $f(\mathbf{x})$ over a field $\mathbb{F}$ is the set $\{f(A\mathbf{x}+\mathbf{b}) : A \in \mathrm{GL}(n,\mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^n\}$. In this paper, we initiate the study of explicit hitting sets for the *orbits* of polynomials computable by several natural and well-studied circuit classes and polynomial families. In particular, we give quasi-polynomial time hitting sets for the orbits of:

1. Low-individual-degree polynomials computable by *commutative ROABPs*. This implies quasi-polynomial time hitting sets for the orbits of the *elementary symmetric polynomials*.
2. Multilinear polynomials computable by *constant-width ROABPs*. This implies a quasi-polynomial time hitting set for the orbits of the family $\{\mathsf{IMM}_{3,d}\}_{d\in\mathbb{N}}$, which is complete for arithmetic formulas.
3. Polynomials computable by *constant-depth, constant-occur formulas*. This implies quasi-polynomial time hitting sets for the orbits of *multilinear depth-4 circuits with constant top fan-in*, and also polynomial-time hitting sets for the orbits of the *power symmetric* and the *sum-product polynomials*.
4. Polynomials computable by *occur-once formulas*.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases Hitting Sets, Orbits, ROABPs, Rank Concentration

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2021.50

Category RANDOM

Related Version *Full Version*: `https://eccc.weizmann.ac.il/report/2021/015/`

Acknowledgements We thank Rohit Gurjar, Ankit Garg, Neeraj Kayal, and Vishwas Bhargava for several stimulating discussions at the onset of this work.

1 Introduction

Polynomial identity testing (PIT) is a fundamental problem in arithmetic circuit complexity. PIT is the problem of deciding if a given arithmetic circuit computes an identically zero polynomial. It is one of the few natural problems in BPP (in fact, in co-RP) for which we do not know of deterministic polynomial-time algorithms. A probabilistic polynomial-time algorithm for PIT follows from the DeMillo-Lipton-Schwartz–Zippel lemma [15, 71, 78]. PIT has connections to other interesting problems like perfect matching [19, 41, 49, 53, 75], the linear matroid intersection [33, 55], and the maximum rank matrix completion [33, 54]. The deterministic primality testing algorithm in [4] derandomizes a particular instance of PIT over a ring [2]. Also, multivariate polynomial factorization for general circuits can be efficiently reduced to PIT and factoring univariate polynomials [37, 38, 48]. Moreover, derandomizing PIT or the black-box version of PIT[1] is essentially equivalent to proving arithmetic circuit lower bounds.

[1] An algorithm for the black-box PIT problem takes as input black-box access to a circuit. The algorithm cannot "see" the circuit but can query it at any point [1, 34, 36, 57]. The black-box PIT problem for a circuit class $\mathcal{C}$ is also known as the problem of constructing *hitting sets* for $\mathcal{C}$

In the past two decades, PIT algorithms and hitting set constructions have been studied for various restricted classes/models of circuits. Bounding the read of every variable is a natural restriction that has received a lot of attention. Two constant-read models, viz. *read-once oblivious algebraic branching programs* (ROABPs) and *constant-read* (more generally, *constant-occur*) *formulas*. These models are quiet powerful and capture many interesting circuit classes. A polynomial-time PIT algorithm and a quasi-polynomial time hitting set construction for ROABPs are known [3, 23, 61]. A quasi-polynomial time hitting set construction for multilinear constant-read formulas was given by [10]. [5] obtained polynomial-time constructible hitting sets for constant-depth, constant-occur formulas.

Hitting sets for orbits. In this paper, we study hitting set constructions for the *orbits* of ROABPs and constant-occur formulas. The orbit of a polynomial f is the set of polynomials obtained by applying invertible affine transformations on the variables of f, i.e., by replacing the variables of f with linearly independent affine forms. The orbit of a circuit class is the union of the orbits of the polynomials computable by the circuits in the class. Our reasons for studying hitting sets for the orbits of ROABPs and constant-occur formulas are threefold:

1. *The power of orbit closures:* The set of affine projections of an n-variate polynomial $f(\mathbf{x})$ over a field $\mathbb{F}$ is $\text{aproj}(f) := \{f(A\mathbf{x}+\mathbf{b}) \ : \ A \in \mathbb{F}^{n\times n} \text{ and } \mathbf{b} \in \mathbb{F}^n\}$; the orbit of f is the set $\text{orb}(f) = \{f(A\mathbf{x}+\mathbf{b}) \ : \ A \in \text{GL}(n,\mathbb{F}) \text{ and } \mathbf{b} \in \mathbb{F}^n\} \subseteq \text{aproj}(f)$. Affine projections of polynomials computable by polynomial-size ROABPs or constant-occur formulas have great expressive power. All polynomials computed by algebraic branching programs and arithmetic formuals are affine projections of polynomial width ROABPs and constant ROABPs, respectively. Similarly, all polynomials computed by depth-3 arithmetic circuits (which are quiet powerful [7, 30, 46, 76, 77]) and arithmetic formulas are affine projections of read once formulas. The orbit of f being a mathematically interesting subset of $\text{aproj}(f)$, it is natural to ask if we can construct efficient hitting sets for the orbits of the above-mentioned circuit classes. Moreover, $\text{orb}(f)$ is not "much smaller" than $\text{aproj}(f)$, as the latter is contained in the *orbit closure* of f if $\text{char}(\mathbb{F}) = 0$ (see the full version [64] for more details).

2. *Geometry of the circuit classes:* Consider an n-variate polynomial $f \in \mathbb{R}[\mathbf{x}]$ and let $\mathbb{V}(f)$ be the variety (i.e., the zero locus) of f. The geometry of $\mathbb{V}(f)$ is preserved by any rigid transformation on $\mathbb{R}^n$. Computation of a set $\mathcal{H} \subseteq \mathbb{R}^n$ that is not contained in $T(\mathbb{V}(f))$, for every rigid transformation T, would have to be "mindful" of the geometry of $\mathbb{V}(f)$ and oblivious to the choice of the coordinate system. Computing such an $\mathcal{H}$ is exactly the problem of constructing a hitting set for the polynomials $\{f(R\mathbf{x}+\mathbf{b}) \ : \ R \in O(n,\mathbb{R}) \text{ and } \mathbf{b} \in \mathbb{R}^n\}$. We can generalize the problem slightly by replacing $R \in O(n,\mathbb{R})$ with $A \in \text{GL}(n,\mathbb{R})$. A hitting set for ROABPs or constant-occur formulas does not immediately give a hitting set for $\{f(A\mathbf{x}+\mathbf{b}) \ : \ A \in \text{GL}(n,\mathbb{R}) \text{ and } \mathbf{b} \in \mathbb{R}^n\}$, as the definitions of an ROABP and a constant-occur formula are tied to the choice of the coordinate system. It is thus natural to ask if there is anything special about the geometry of $\mathbb{V}(f)$ which can facilitate efficient constructions of hitting sets for $\text{orb}(f)$.

3. *Strengthening existing techniques:* Finally, it is worth investigating whether the techniques used to design hitting sets for ROABPs and constant-occur formulas can be applied or strengthened or combined to give hitting sets for the orbits of these circuit classes.

1.1 The models

Unless otherwise stated, we will assume that polynomials are over a field $\mathbb{F}$. Read Once Algebraic Branching Programs (ROABPs) are the read once versions of Algebraic Branching Programs defined by Nisan [56]. While Nisan defined ABPs using directed graphs, we use the following equivalent and conventional definition of an ROABP.

▶ **Definition 1** (ROABP [23]). *An n-variate, width-w read-once oblivious algebraic branching program (ROABP) is a product of the form $\mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot \mathbf{1}$, where $\mathbf{1}$ is the $w \times 1$ vector of all ones, and for every $i \in [n]$, $M_i(x_i)$ is a $w \times w$ matrix whose entries are in $\mathbb{F}[x_i]$.*

▶ **Definition 2** (Commutative ROABP). *An n-variate, width-w* commutative *ROABP is an n-variate, width-w ROABP $\mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot \mathbf{1}$, where for all $i, j \in [n]$, $M_i(x_i)$ and $M_j(x_j)$ commute with each other.*

A polynomial f is *s-sparse* if it has at most s monomials with non-zero coefficients; these monomials will be referred to as the *monomials of f*. A degree-d s-sparse polynomial can be computed by a depth-2 circuit of size sd as well as by a width-s commutative ROABP.

▶ **Definition 3** (Occur-k formula [5]). *An occur-k formula is a rooted tree whose leaves are labelled by s-sparse polynomials and whose internal nodes are sum ($+$) gates or product-power ($\times^{\wedge}$) gates. Each variable appears in at most k of the sparse polynomials that label the leaves. The edges feeding into a $+$ gate are labelled by field elements and have* 1 *as* edge weights*, whereas the edges feeding into a $\times^{\wedge}$ gate have natural numbers as edge weights. A leaf node computes the s-sparse polynomial that labels it. A $+$ gate with inputs from nodes that compute $f_1, ..., f_m$ and with the corresponding input edge labels $\alpha_1, ..., \alpha_m$, computes $\alpha_1 f_1 + \cdots + \alpha_m f_m$. A $\times^{\wedge}$ gate with inputs from nodes that compute $f_1, ..., f_m$ and with the corresponding input edge weights $e_1, ..., e_m$, computes $f_1^{e_1} \cdots f_m^{e_m}$. The formula computes the polynomial that is computed by the root node.*

The size *of an occur-k formula is the weighted sum of all the edges in it (i.e., an edge is counted as many times as its edge weight) plus the sizes of the depth-2 circuits computing the s-sparse polynomials at the leaves. The* depth *of an occur-k formula is equal to the depth of the underlying tree plus* 2*, to account for the depth of the circuits computing the sparse polynomials at the leaves.*

Read-k formulas have been studied intensely in the literature (see Section 1.4). Occur-k formulas generalize read-k formulas in two ways – the leaves are labelled by arbitrary sparse polynomials instead of just variables, and powering gates are included along with the usual sum and product gates. These generalizations help make the occur-k model complete [2], and capture other interesting circuit classes (such as multilinear depth-4 circuits with constant top fan-in [39, 65]) and polynomial families (such as the power symmetric polynomials). Besides, unlike some prior work [10, 39, 65], there is no restriction of multilinearity on the model. We will identify the variable set $\mathbf{x} = \{x_1, \ldots, x_n\}$ with the column vector $(x_1\ x_2\ \cdots\ x_n)^T$.

▶ **Definition 4** (Orbits of polynomials). *Let $f(\mathbf{x})$ be an n-variate polynomial over a field $\mathbb{F}$. The* orbit *of f, denoted by* $\mathrm{orb}(f)$*, is the set $\{f(A\mathbf{x}) : A \in \mathrm{GL}(n, \mathbb{F})\}$. The orbit of a set of polynomials $\mathcal{C}$, denoted by* $\mathrm{orb}(\mathcal{C})$*, is the union of the orbits of the polynomials in $\mathcal{C}$.*

[2] For example, the power symmetric polynomial $x_1^n + \ldots + x_n^n$ cannot be computed by a read-k formula for any $k < n$, but it can be computed by an occur-once formula.

The results we present in this paper hold even if we define the orbit of an n-variate polynomial f as $\text{orb}(f) = \{f(A\mathbf{y} + \mathbf{b}) \ : \ |\mathbf{y}| = m \geq n, \ A \in \mathbb{F}^{n \times m}$ has rank n, and $\mathbf{b} \in \mathbb{F}^n\}$. However, we work with this slightly conventional definition of $\text{orb}(f)$ for simplicity of exposition, and because the proofs in the general setting are nearly the same as the proofs we present here. By the "orbit of a circuit class $\mathcal{C}$", we mean the union of the orbits of the polynomials computable by the circuits in the class $\mathcal{C}$. Our main results are efficient constructions of hitting sets for the orbits of commutative ROABPs and constant-width ROABPs (under low individual degree restriction), and the orbits of constant-depth constant-occur formulas and occur-once formulas.

1.2 Our results

▶ **Definition 5** (Hitting set). *Let $\mathcal{C}$ be a set of n-variate polynomials. A set of points $\mathcal{H} \subseteq \mathbb{F}^n$ is a hitting set for $\mathcal{C}$ if for every non-zero $f \in \mathcal{C}$, there is a point $\mathbf{a} \in \mathcal{H}$ such that $f(\mathbf{a}) \neq 0$.*

By a "T-time hitting set", we mean that the hitting set can be computed in T time. The *individual degree* of a monomial is the largest of the exponents of the variables that appear in it. The individual degree of a polynomial is the largest of the individual degrees of its monomials. We are now ready to state our results.

▶ **Theorem 6** (Hitting sets for the orbits of commutative ROABPs with low individual degree). *Let $\mathcal{C}$ be the set of n-variate polynomials with individual degree at most d that are computable by width-w commutative ROABPs. If $|\mathbb{F}| > n^2 d$, then a hitting set for* $\text{orb}(\mathcal{C})$ *can be computed in $(nd)^{O(d \log w)}$ time.*

We say an n-variate polynomial $f(x_1, x_2, \ldots, x_n)$ can be expressed as a sum of s products of univariates if $f = \sum_{i \in [s]} \prod_{j \in [n]} f_{i,j}(x_j)$, where each $f_{i,j}(x_j)$ is a univariate polynomial in x_j. This model is subsumed by commutative ROABPs and has found important applications in several other works [30, 63, 66]. The above theorem implies a $nd^{O(d \log s)}$ time hitting set for this model. As the elementary symmetric polynomials and low individual degree sparse polynomials are special cases of low individual degree sum of products of univariates, we also get quasi-polynomial hitting sets for these models. It turns out though that for the particular case of sparse polynomials it is possible to remove the individual degree restriction from the above theorem. This is due to an independent and simultaneous work by [51]. We state their result next.

▶ **Theorem 7** (Hitting sets for the orbits of sparse polynomials [51]). *Let $\mathcal{C}$ be the set of n-variate, s-sparse polynomials of degree at most d. If $|\mathbb{F}| > nd$ and* $\text{char}(\mathbb{F}) = 0$ *or* $> d$, *then a hitting set for* $\text{orb}(\mathcal{C})$ *can be computed in $(nd)^{O(\log s)}$ time.*

The above theorem plays a basic role in the proofs of Theorem 9 and Theorem 10. There, we apply the algebraic independence based analysis from [5, 11] and the Shpilka-Volkovich (SV) generator based argument from [73], respectively, to reduce to the case of constructing hitting sets for the orbits of sparse polynomials. While in the original version of our work [64] we applied Theorem 6 in the base case of the proofs of Theorem 9 and 10, here we plug-in Theorem 7 in the base case. This helps us forgo the low individual degree restriction that was present in these theorems in the original version.

▶ **Theorem 8** (Hitting sets for the orbits of multilinear constant-width ROABPs). *Let $\mathcal{C}$ be the set of n-variate multilinear polynomials that are computable by width-w ROABPs. If $|\mathbb{F}| > n^{O(w^4)}$, then a hitting set for* $\text{orb}(\mathcal{C})$ *can be computed in $n^{O(w^6 \cdot \log n)}$ time.*

The theorem gives a quasi-polynomial time hitting set for $\mathrm{orb}(\mathsf{IMM}_{3,d})$, which is complete for the class of arithmetic formulas under affine projections (in fact, under p-projections) [12]. The set of affine projections of $\mathsf{IMM}_{2,d}$ is also quite rich, despite the fact that there are simple quadratic polynomials that are not in $\mathrm{aproj}(\mathsf{IMM}_{2,d})$ for *any* d [8, 62]. This is because hitting sets for $\mathrm{aproj}(\mathsf{IMM}_{2,d})$ give hitting sets for depth-3 circuits [62]. Moreover, $\overline{\mathrm{orb}(\mathsf{IMM}_{2,d})}$ captures the orbit closures of arithmetic formulas [13]. The above theorem implies a quasi-polynomial time hitting set for $\mathrm{orb}(\mathsf{IMM}_{2,d})$.

▶ **Theorem 9** (Hitting sets for the orbits of constant-depth, constant-occur formulas). *Let $\mathcal{C}$ be the set of n-variate, degree-D polynomials that are computable by depth-Δ, occur-k formulas of size s. Let $R := (2k)^{2\Delta \cdot 2^{\Delta}}$. If $\mathrm{char}(\mathbb{F}) = 0$ or $> (2ks)^{\Delta^3 R}$, then a hitting set for $\mathrm{orb}(\mathcal{C})$ can be computed in $(nRD)^{O(R(\log R + \Delta \log k + \Delta \log s) + \Delta R)}$ time. If the leaves are labelled by b-variate polynomials, then a hitting set for $\mathrm{orb}(\mathcal{C})$ can be computed in $(nRD)^{O(Rb + \Delta R)}$ time. In particular, if Δ and k are constants, then the hitting sets can be constructed in time $(nD)^{O(\log s)}$ and $(nD)^{O(b)}$, respectively.*

The above theorem gives quasi-polynomial hitting sets for the orbits of two other interesting models viz. multilinear depth-4 circuits with constant top fan-in and the class of polynomials $C(f_1, \ldots, f_m)$, where C is a low-degree circuit and $f_1, \ldots, f_m$ are sparse polynomials with bounded transcendence degree [11]. The theorem also yields polynomial-time hitting sets for the orbits of the power symmetric polynomial and the sum-product polynomial $\mathsf{SP}_{n,D} = \sum_{i \in [n]} \prod_{j \in [D]} x_{i,j}$. Prior to our work, [47] gave a polynomial-time hitting set for the orbit of power symmetric polynomials using a different argument.

▶ **Theorem 10** (Hitting sets for the orbits of occur-once formulas). *Let $\mathcal{C}$ be the set of n-variate, degree-D polynomials that are computable by occur-once formulas whose leaves are labelled by s-sparse polynomials. If $|\mathbb{F}| > nD$ and $\mathrm{char}(\mathbb{F}) = 0$ or $> D$, then a hitting set for $\mathrm{orb}(\mathcal{C})$ can be computed in $(nD)^{O(\log n + \log s)}$ time. If the leaves are labelled by b-variate polynomials, then a hitting set for $\mathrm{orb}(\mathcal{C})$ can be computed in $(nD)^{O(\log n + b)}$ time.*

The independent and concurrent work [51] gave (among other results) a quasi-polynomial time hitting set for the orbits of read-once formulas. We note that this result also follows from the second part of the above theorem which is already present in the original version of this work [64]. The proofs of Theorems 9 and 10 can be found in the full version [64].

1.3 Proof techniques

Let us briefly discuss the techniques that go into proving the above results.

Commutative ROABPs with low individual degree. Theorem 6 is proved by adapting the rank concentration by translation technique of [6] to work for the orbits of commutative ROABPs. Let $f = \mathbf{1}^T \cdot M_1(x_1) M_2(x_2) \cdots M_n(x_n) \cdot \mathbf{1}$ be a commutative ROABP and $F = M_1(x_1) M_2(x_2) \cdots M_n(x_n)$. For any $A \in \mathrm{GL}(n, \mathbb{F})$, let $g = f(A\mathbf{x})$ and $G = F(A\mathbf{x})$. Suppose that A maps x_i to a linear form $\ell_i(\mathbf{x})$ for every $i \in [n]$, and let $y_i = \ell_i(\mathbf{x})$. Then, $g = \mathbf{1}^T \cdot M_1(y_1) M_2(y_2) \cdots M_n(y_n) \cdot \mathbf{1}$ and $G = M_1(y_1) M_2(y_2) \cdots M_n(y_n)$. We show that if $g \neq 0$, then there exist *explicit* "low" degree polynomials $t_1(\mathbf{z}), \ldots, t_n(\mathbf{z})$, where $\mathbf{z}$ is a "small" set of variables, such that $g(x_1 + t_1(\mathbf{z}), \ldots, x_n + t_n(\mathbf{z}))$ has a "low" support monomial. This is done by proving that $G(x_1 + t_1(\mathbf{z}), \ldots, x_n + t_n(\mathbf{z}))$ has low support rank concentration over $\mathbb{F}(\mathbf{z})$ in the "$\mathbf{y}$-variables" (see Section 2.2 for the meaning of low support rank concentration.). That done, we use the assumption that f has low individual degree to argue that $g(x_1 + t_1(\mathbf{z}), \ldots, x_n + t_n(\mathbf{z}))$ also has a low support $\mathbf{x}$-monomial. This and the fact that $|\mathbf{z}|$ is small imply that $g(x_1 + t_1(\mathbf{z}), \ldots, x_n + t_n(\mathbf{z}))$, when viewed as a polynomial in $\mathbb{F}[\mathbf{x}, \mathbf{z}]$, has a low support monomial. Finally, we use the SV generator to hit g.

Our analysis differs from that in [6] at a crucial point: In [6], it was shown that $F(\mathbf{x}+\mathbf{t}) = M_1(x_1+t_1)M_2(x_2+t_2)\cdots M_n(x_n+t_n)$ has low support rank concentration over $\mathbb{F}(\mathbf{t})$ if the nonzeroness of every polynomial in a certain collection of polynomials – each in a "small" set of $\mathbf{t}$-variables – is preserved. As each polynomial in the collection has "few" $\mathbf{t}$-variables, a substitution $t_i \leftarrow t_i(\mathbf{z})$ that preserves its nonzeroness is relatively easy to construct. But the collection of polynomials that we need to preserve to show low support rank concentration for $G(\mathbf{x}+\mathbf{t})$ is such that every polynomial in the collection has potentially all the $\mathbf{t}$-variables. However, we are able to argue that each of these polynomials still has a low support $\mathbf{t}$-monomial. This then helps us construct a substitution $t_i \mapsto t_i(\mathbf{z})$ that preserves the nonzeroness of these polynomials.

Multilinear constant-width ROABPs. Theorem 8 is proved by combining the rank concentration by translation technique of [6] with the merge-and-reduce idea from [23] and [21]. Let $f = \mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot \mathbf{1}$ be a multilinear, width-w ROABP; here $M_i(x_i) \in \mathbb{F}^{w\times w}[x_i]$ for all $i \in [n]$. Also, let $F = M_1(x_1)M_2(x_2)\cdots M_n(x_n)$. For any $A \in \mathrm{GL}(n,\mathbb{F})$, let $g = f(A\mathbf{x})$ and $G = F(A\mathbf{x})$. For $i \in [n]$, suppose that A maps $x_i \mapsto \ell_i(\mathbf{x})$, where ℓ_i is a linear form, and let $y_i = \ell_i(\mathbf{x})$ and $\mathbf{y} = \{y_1,\ldots,y_n\}$. Then, $g = \mathbf{1}^T \cdot M_1(y_1)M_2(y_2)\cdots M_n(y_n)\cdot \mathbf{1}$ and $G = M_1(y_1)M_2(y_2)\cdots M_n(y_n)$. Much like in the case of commutative ROABPs, we show that if $g \neq 0$, then there exist explicit "low" degree polynomials $t_1(\mathbf{z}),\ldots,t_n(\mathbf{z})$, where $\mathbf{z}$ is a "small" set of variables such that $G(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$ has "low" support rank concentration in the "$\mathbf{y}$-variables". While in the rank concentration argument for commutative ROABPs the $\mathbf{x}$-variables were translated only once, here the translations can be thought of as happening sequentially and in stages. There will be $\lceil \log n \rceil$ stages with each stage also consisting of multiple translations. After the p-th stage, the product of any 2^p consecutive matrices in G will have low support rank concentration in the $\mathbf{y}$-variables. Thus, after $\lceil \log n \rceil$ stages, we will have low support rank concentration in the $\mathbf{y}$-variables for $G(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$.

As in the case of commutative ROABPs, we show that $G(\mathbf{x}+\mathbf{t})$ has low support rank concentration if each polynomial in a certain collection of non-zero polynomials in the $\mathbf{t}$-variables is kept non-zero by the substitution $t_i \mapsto t_i(\mathbf{z})$. However, in this case, it is trickier to show that these polynomials have low support $\mathbf{t}$-monomials. We do this by arguing that each such polynomial can be expressed as a ratio of a polynomial that contains a low support $\mathbf{t}$-monomial and a product of linear forms in the $\mathbf{t}$-variables.

Constant-depth, constant-occur formulas. We prove Theorem 9 by combining the algebraic independence based technique in [5] with Theorem 7. Let f be a constant-depth, constant-occur formula. We first show that it can be assumed without loss of generality that the top-most gate of f is a $+$ gate whose fan-in is upper bounded by the occur of f, say k. In [5], they were able to upper bound the top fan-in by simply translating a variable by 1 and subtracting the original formula. However, the same idea does not quite work here, because we have only access to a polynomial in the *orbit* of f. To upper bound the top fan-in, we show that there exists a variable x_i such that $\frac{\partial f}{\partial x_i}$ is a constant-depth, constant-occur formula with top fan-in bounded by k. Then, using the chain rule of differentiation, we show that one can construct a hitting set generator for $\mathrm{orb}(f)$ from a generator for $\mathrm{orb}\left(\frac{\partial f}{\partial x_i}\right)$; this means that we can shift our attention to $f' = \frac{\partial f}{\partial x_i}$, which we shall henceforth refer to as f.

Let $f = f_1 + \cdots + f_k$, $A \in \mathrm{GL}(n,\mathbb{F})$, $g = f(A\mathbf{x})$, $g = g_1 + \ldots + g_k$ where for all $i \in [k]$, $g_i = f_i(A\mathbf{x})$. It was shown in [5] that a homomorphism, which is faithful (see Definition 17) to $f_1,\ldots,f_k$, is a hitting set generator for f. In our case, this translates to 'a

homomorphism that is faithful to $g_1, \ldots, g_k$ is a hitting set generator for g '. [5] also showed that the problem of constructing a homomorphism ϕ that is faithful to $f_1, \ldots, f_k$ reduces to constructing a homomorphism ψ that preserves the determinant of a certain matrix. This matrix is an appropriate sub-matrix of the Jacobian of $f_1, \ldots, f_k$. Also, it was argued that its determinant is a product of sparse polynomials and so ψ was obtained from [45]. We use a similar argument, along with the chain rule, to show that the problem of constructing a homomorphism ϕ that is faithful to $g_1, \ldots, g_k$ reduces to constructing a homomorphism ψ that preserves the determinant of a sub-matrix of the same Jacobian *evaluated at* $A\mathbf{x}$. As this determinant is a product of polynomials in the orbit of sparse polynomials, we can use Theorem 7 to construct such a ψ.

Occur-once formulas. We prove Theorem 10 by building upon the arguments in [73] and linking it with Theorem 7. At first, we show two structural results for occur-once formulas. These lemmas are generalizations of similar structural results for read-once formulas shown in [73]. Much like in [73], the structural results help us show that for a "typical" occur-once formula f with a $+$ gate as the root node, there exists a variable x_i such that $\frac{\partial f}{\partial x_i}$ is a product of occur-once formulas, each of which has at most half as many non-constant leaves as f. We then use this fact to show that a hitting-set generator for $\text{orb}(f)$ can be constructed from a generator for $\text{orb}\left(\frac{\partial f}{\partial x_i}\right)$. [73] uses the derivatives of f in a similar way to show that a generator for f can be constructed from that for $\frac{\partial f}{\partial x_i}$ using the SV generator (see Definition 12). However, in our case, we want a generator for $\text{orb}(f)$ and not just for f. For this reason, we first use the chain rule for derivatives to relate the gradient of a $g \in \text{orb}(f)$ with that of f, and then argue that there exists a x_j such that a generator for $\text{orb}\left(\frac{\partial f}{\partial x_i}\right)$ is also a generator for $\frac{\partial g}{\partial x_j}$. Finally, we use this generator for $\frac{\partial g}{\partial x_j}$ to construct a generator for g. The argument then proceeds by induction on the number of non-constant leaves. In the base case, we need a hitting set generator for orbits of sparse polynomials which we get from Theorem 7.

1.4 Related work

We give a brief account of known results on PIT and hitting sets for arithmetic circuits. The results on hitting sets for the constant-read models are most relevant to our work here. However, for the sake of completeness, we mention a few other prominent results.

Constant-read models. [73] gave a polynomial-time PIT algorithm and a quasi-polynomial time hitting set construction for sums of constantly many *preprocessed* read-once formulas (PROFs). [52] later gave a polynomial time hitting set for the same model. [10] gave a quasi-polynomial time hitting set construction for multilinear *sparse-substituted* read-k formulas, wherein the leaves are replaced by sparse polynomials and every variable appears in at most k of the sparse polynomials. Observe that the models studied in all three works are special cases of constant occur formulas.

A polynomial-time PIT for ROABPs follows from the PIT algorithm for non-commutative formulas [61]. [23] gave quasi-polynomial time hitting sets for ROABPs, when the order of the variables is known. Building on the rank concentration by translation technique from [6] and the merge-and-reduce idea from [23], [21] gave a quasi-polynomial time hitting set construction for low individual degree ROABPs. Finally, [3] obtained a quasi-polynomial time constructible hitting set for ROABPs using a different and simpler method, namely *basis isolation*, which can be thought of as a generalization of the monomial isolation method in [45]. [32] designed hitting sets for sums of constantly many ROABPs

in quasi-polynomial time; they also gave a polynomial-time PIT algorithm for the same model. Recently, more efficient constructions of hitting sets for ROABPs have been obtained [27], sometimes under additional restrictions on the model such as commutativity and constant-width [31]. For read-k oblivious ABPs, [9] obtained a subexponential-time PIT algorithm.

Orbits and orbit closures. A polynomial-time hitting set for the *orbit* of the power symmetric polynomial $\mathsf{PSym}_{n,d} = x_1^d + \ldots + x_n^d$ was given by [47]. As that PSym is computable by a depth-2 occur-once formula, Theorem 9 subsumes this result. Our hitting-set construction is different from the one in [47] which involves the Hessian matrix, whereas the proofs here work with just the first order derivatives. Very recently and independent of our work, [51] gave quasi-polynomial time hitting sets for the orbits of sparse polynomials and read-once formulas. For the orbit closures of polynomials that are computable by low-degree, polynomial-size circuits (i.e., VP circuits), [24, 28] gave PSPACE constructions of hitting sets.

Constant-depth models. The polynomial-time hitting set construction for depth-2 circuits (i.e., sparse polynomials) in [45] is one of the widely used results in black-box PIT. [16] gave a quasi-polynomial time PIT algorithm for depth-3 circuits with constant top fan-in. Later [44] improved the complexity to polynomial-time. Using ideas developed in [16], and [25], [40, 43, 70] gave polynomial-time constructible hitting sets for depth-3 circuits with constant top fan-in over $\mathbb{Q}$. Ultimately, a combination of ideas from the [44] and [25] led to a polynomial-time hitting set construction for the same model over any field [69,70]. Meanwhile, [42, 66] gave polynomial-time PIT for depth-3 powering circuits. Using ideas from [44] and [66], [63] gave polynomial-time PIT for the sum of a depth-3 circuit with constant top fan-in and a *semi-diagonal* circuit (which is a special kind of a depth-4 circuit). [62] showed that polynomial-time PIT (resp. hitting sets) for aproj($\mathsf{IMM}_{2,d}$) implies polynomial-time PIT (resp. hitting sets) for depth-3 circuits.

A quasi-polynomial time hitting set for set-multilinear depth-3 circuits with known variable-partition was given by [22]. Independently and simultaneously, [6] gave a quasi-polynomial time hitting set for set-multilinear depth-3 circuits with *unknown* variable-partition (and more generally, for constant-depth *pure* formulas [58]) using a different technique, namely *rank concentration by translation*. Set-multilinear depth-3 circuits (in fact, pure formulas) form a subclass of ROABPs. [14] gave subexponential-time hitting sets for multilinear depth-3 and depth-4 formulas (and more generally, for constant-depth multilinear regular formulas) by reducing the problem to constructing hitting sets for ROABPs. For multilinear depth-4 circuits with constant top fan-in, [39] gave a quasi-polynomial time hitting set. This was improved to a polynomial-time hitting set in [65]. Multilinear depth-4 circuits with constant top fan-in form a subclass of depth-4 constant-occur formulas. [5] gave a unifying method based on algebraic independence to design polynomial-time hitting sets for both depth-3 circuits with constant top fan-in and constant-depth, constant-occur formulas. A generalization of depth-3 powering circuits to depth-4 is sums of powers of constant degree polynomials; [20] gave a quasi-polynomial time hitting set for this model. Recently, a sequence of work [59, 60, 72] led to a polynomial-time hitting set for depth-4 circuits with top fan-in at most 3 and bottom fan-in at most 2 via a resolution of a conjecture of [11, 29] on the algebraic rank of the factors appearing in such circuits.

Edmonds' model. An important special case of PIT is the following problem: given $f = \det(A_0 + \sum_{i\in[n]} x_i A_i)$, where $A_i \in \mathbb{F}^{n\times n}$ is a rank-1 matrix for every $i \in [n]$ and $A_0 \in \mathbb{F}^{n\times n}$ is an arbitrary matrix, check if $f = 0$ [17]. This case of PIT, played an instrumental

role in devising fast parallel algorithms for several problems such as perfect matching, linear matroid intersection and maximum rank matrix completion [19, 33, 41, 49, 53–55, 75]. A polynomial-time PIT for this model is known [18, 26, 35, 50, 54]. [33] gave a quasi-polynomial time hitting set via a certain derandomization of the Isolation Lemma [53].

We refer the reader to the surveys [67, 68, 74] for more details on some of the results and the models mentioned above.

2 Preliminaries

▶ **Definition 11** (Hitting set generator). *Let $\mathcal{C}$ be a set of n-variate polynomials and $t \in \mathbb{N}$. A polynomial map $\mathcal{G} : \mathbb{F}^t \to \mathbb{F}^n$ is a hitting set generator for $\mathcal{C}$ if $\forall f \in \mathcal{C} \setminus \{0\}$, we have $f \circ \mathcal{G} \neq 0$.*

We say the number of variables of $\mathcal{G}$ is t, and the degree of $\mathcal{G}$ – denoted by $\deg(\mathcal{G})$ – is the maximum of the degrees of the n polynomials that define $\mathcal{G}$. We will denote the t-variate polynomial $f \circ \mathcal{G}$ by $f(\mathcal{G})$. By treating a matrix $A \in \mathbb{F}^{n \times n}$ as a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^n$, we will denote the polynomial map $A \circ \mathcal{G}$ by $A\mathcal{G}$ and the t-variate polynomial $f \circ A\mathcal{G}$ by $f(A\mathcal{G})$. If the defining polynomials of $\mathcal{G}$ have degree d_0 and the degree of the polynomials in $\mathcal{C}$ is at most D, then the degree of $f(\mathcal{G})$ is at most $d_0 D$. Thus, if we are given the defining polynomials of $\mathcal{G}$, then we can construct a hitting set for $\mathcal{C}$ in time $\text{poly}(n, (d_0 D)^t)$ using the Schwartz-Zippel lemma, provided also that $|\mathbb{F}| > d_0 D$.

2.1 The Shpilka-Volkovich generator

▶ **Definition 12** (The Shpilka-Volkovich hitting set generator [73]). *Assume that $|\mathbb{F}| \geq n$ and let $\alpha_1, ..., \alpha_n$ be distinct elements of $\mathbb{F}$. For $i \in [n]$, let $L_i(y) := \prod_{j \in [n], j \neq i} \frac{y - \alpha_j}{\alpha_i - \alpha_j}$ be the i-th Lagrange interpolation polynomial. Then, for $t \in \mathbb{N}$, the Shpilka-Volkovich (SV) generator $\mathcal{G}_t^{SV} : \mathbb{F}^{2t} \to \mathbb{F}^n$ is defined as $\mathcal{G}_t^{SV} := \left(\mathcal{G}_t^{(1)}, ..., \mathcal{G}_t^{(n)}\right)$ where, $\mathcal{G}_t^{(i)}(y_1, ..., y_t, z_1, ..., z_t) = \sum_{k=1}^{t} L_i(y_k) \cdot z_k$.*

Notice that $\deg\left(\mathcal{G}_t^{(i)}\right) = n$, and $\mathcal{G}_{t+1|(y_{t+1}=\alpha_i)}^{SV} = \mathcal{G}_t^{SV} + \mathbf{e}_i \cdot z_{t+1}$, where $\mathbf{e}_i$ is the i-th standard basis vector of $\mathbb{F}^n$. Thus, $\text{Img}\left(\mathcal{G}_t^{SV}\right) \subseteq \text{Img}\left(\mathcal{G}_{t+1}^{SV}\right)$ and, continuing in this manner, $\text{Img}\left(\mathcal{G}_t^{SV}\right) \subseteq \text{Img}\left(\mathcal{G}_{t'}^{SV}\right)$ for any $t' \geq t$.

▶ **Observation 13.** *Let $f \in \mathbb{F}[\mathbf{x}]$ be a non-zero polynomial that depends on only b of the $\mathbf{x}$ variables, and $g \in \text{orb}(f)$. Then, g has a monomial of support at most b and $g(\mathcal{G}_b^{SV}) \neq 0$.*

The above observation is proved in the full version [64]. The following observation, which allows us to construct a hitting set generator for f from a hitting set generator for $\frac{\partial f}{\partial x_i}$ is used crucially in the proofs of Theorems 9 and 10 and is proved in the full version [64].

▶ **Observation 14.** *Let $f \in \mathbb{F}[\mathbf{x}]$ be an n-variate, degree d polynomial, and for some $m \in \mathbb{N}$, let $\mathcal{G} : \mathbb{F}^m \to \mathbb{F}^n$ be a polynomial map of degree at most d'. If $|\mathbb{F}| > dd'$ and there is an $i \in [n]$ such that $\frac{\partial f}{\partial x_i}(\mathcal{G}) \neq 0$, then $f(\mathcal{G} + \mathcal{G}_1^{SV})$ is not a constant.*

2.2 Low support rank concentration

Let F be a polynomial in $\mathbf{x}$-variables with coefficients from $\mathbb{K}^{w \times w}$, where $\mathbb{K}$ is a field and $w \in \mathbb{N}$. For an $m \in \mathbb{N}$, we say that F has *support-m rank concentration* over $\mathbb{K}$ if the coefficient of every monomial in F is in the $\mathbb{K}$-span of the coefficients of the monomials of support at most m in F. Support of a monomial $\mathbf{x}^{\boldsymbol{\alpha}}$ will be denoted as $\text{Supp}\left(\mathbf{x}^{\boldsymbol{\alpha}}\right)$. We prove the below observation in the full version [64].

▶ **Observation 15.** *Let* $f = \mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot \mathbf{1} \in \mathbb{F}[\mathbf{x}]$ *be computable by an ROABP of width* w*, and* $F = M_1(x_1)M_2(x_2)\cdots M_n(x_n)$*. For an* $m \in \mathbb{N}$ *and* $t_1(\mathbf{z}),\ldots,t_n(\mathbf{z}) \in \mathbb{F}[\mathbf{z}]$*, where* $\mathbf{z}$ *is a set of variables different from* $\mathbf{x}$*, suppose that* $F(\mathbf{x}+\mathbf{t}(\mathbf{z})) := M_1(x_1+t_1(\mathbf{z}))M_2(x_2+t_2(\mathbf{z}))\cdots M_n(x_n+t_n(\mathbf{z})) \in \mathbb{F}(\mathbf{z})^{w\times w}[\mathbf{x}]$ *has support-*m *rank concentration over* $\mathbb{F}(\mathbf{z})$*. Then,* $f(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$*, when viewed as a polynomial in* $\mathbf{x}$*-variables with coefficients from* $\mathbb{F}[\mathbf{z}]$*, has an* $\mathbf{x}$*-monomial of support at most* m*, provided* $f \neq 0$*.*

2.3 Algebraic rank and faithful homomorphisms

We say that polynomials $f_1,\ldots,f_m \in \mathbb{F}[\mathbf{x}]$ are algebraically independent over $\mathbb{F}$, if they do not satisfy any non-trivial polynomial equation over $\mathbb{F}$, i.e., for any $p \in \mathbb{F}[y_1,\ldots,y_m]$, $p(f_1,\ldots,f_m) = 0$ only if $p = 0$. For $\mathbf{f} = (f_1,\ldots,f_m)$, the transcendence degree (i.e., the algebraic rank) of $\mathbf{f}$ over $\mathbb{F}$ is the cardinality of any maximal algebraically independent subset of $\{f_1,\ldots,f_m\}$ over $\mathbb{F}$. The notion of algebraic rank is well defined as algebraic independence satisfies the matroid properties.

For $\mathbf{f} = (f_1,\ldots,f_m) \in \mathbb{F}[\mathbf{x}]^m$, let $J_{\mathbf{x}}(\mathbf{f})$ denote the Jacobian matrix of $\mathbf{f}$. The following well-known lemma relates the transcendence degree of $\mathbf{f}$ over $\mathbb{F}$ – denoted by $\text{tr-deg}_{\mathbb{F}}(\mathbf{f})$ – to the rank of the Jacobian.

▶ **Lemma 16** (The Jacobian criterion)**.** *Let* $\mathbf{f} = (f_1,\ldots,f_m) \in \mathbb{F}[\mathbf{x}]^m$ *be a tuple of polynomials of degree at most* D *and* $\text{tr-deg}_{\mathbb{F}}(\mathbf{f}) = r$*. If* $\text{char}(\mathbb{F}) = 0$ *or* $\text{char}(\mathbb{F}) > D^r$*, then* $\text{tr-deg}_{\mathbb{F}}(\mathbf{f}) = \text{rank}_{\mathbb{F}(\mathbf{x})}\ J_{\mathbf{x}}(\mathbf{f})$*.*

▶ **Definition 17** (Faithful homomorphisms)**.** *A homomorphism* $\phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}]$ *is said to be* faithful *to* $\mathbf{f} = (f_1,\ldots,f_m) \in \mathbb{F}[\mathbf{x}]^m$ *if* $\text{tr-deg}_{\mathbb{F}}(\mathbf{f}) = \text{tr-deg}_{\mathbb{F}}(\phi(\mathbf{f}))$*.*

▶ **Lemma 18** (Theorem 2.4 in [5])**.** *If a homomorphism* $\phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}]$ *is faithful to* $\mathbf{f} = (f_1,\ldots,f_m) \in \mathbb{F}[\mathbf{x}]^m$ *, then for any* $p \in \mathbb{F}[y_1,\ldots,y_m]$*,* $p(\mathbf{f}) = 0$ *if and only if* $p(\phi(\mathbf{f})) = 0$*.*

The following lemma was proved in [5, 11].

▶ **Lemma 19** (Lemma 2.7 of [5])**.** *Let* $\mathbf{f} = (f_1,...,f_m)$ *be a tuple of polynomials of degree at most* D*,* $\text{tr-deg}_{\mathbb{F}}(\mathbf{f}) \leq r$*, and* $\text{char}(\mathbb{F}) = 0$ *or* $> D^r$*. Let* $\psi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}]$ *be a homomorphism such that* $\text{rank}_{\mathbb{F}(\mathbf{x})} J_{\mathbf{x}}(\mathbf{f}) = \text{rank}_{\mathbb{F}(\mathbf{z})}\psi(J_{\mathbf{x}}(\mathbf{f}))$*. Then, the map* $\phi : \mathbb{F}[\mathbf{x}] \to \mathbb{F}[\mathbf{z}, t, y_1, ..., y_r]$ *that, for all* $i \in [n]$*, maps* $x_i \to \left(\sum_{j=1}^{r} y_j t^{ij}\right) + \psi(x_i)$ *is faithful to* $\mathbf{f}$*.*

We will need the following observation in our proofs. It is proved in the full version [64].

▶ **Observation 20.** *Let* $\mathbf{f} = (f_1,\ldots,f_m) \in \mathbb{F}[\mathbf{x}]^m$ *be a tuple of polynomials with* $\text{tr-deg}_{\mathbb{F}}(\mathbf{f}) = r$*. For any* $A \in \text{GL}(n,\mathbb{F})$*, let* $g_i = f_i(A\mathbf{x})$ $\forall i \in [m]$ *and* $\mathbf{g} = (g_1,\ldots,g_m)$*. Then,* $\text{tr-deg}_{\mathbb{F}}(\mathbf{g}) = r$*.*

3 Hitting sets for the orbits of commutative ROABPs

The strategy. (Recap) Let $f = \mathbf{1}^T\cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot\mathbf{1}$ be a width-w commutative ROABP; here $M_i(x_i) \in \mathbb{F}^{w\times w}[x_i]$ for all $i \in [n]$. Also, let $F = M_1(x_1)M_2(x_2)\cdots M_n(x_n)$. For any $A \in \text{GL}(n,\mathbb{F})$, let $g = f(A\mathbf{x})$ and $G = F(A\mathbf{x})$. For $i \in [n]$, suppose that A maps $x_i \mapsto \ell_i(\mathbf{x})$, where ℓ_i is a linear form, and let $y_i = \ell_i(\mathbf{x})$ and $\mathbf{y} = \{y_1,\ldots,y_n\}$. Then, $g = \mathbf{1}^T\cdot M_1(y_1)M_2(y_2)\cdots M_n(y_n)\cdot\mathbf{1}$ and $G = M_1(y_1)M_2(y_2)\cdots M_n(y_n)$. We will show that if $g \neq 0$, then there exist explicit "low" degree polynomials $t_1(\mathbf{z}),\ldots,t_n(\mathbf{z})$, where $\mathbf{z}$ is a "small" set of variables such that $g(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$ has a "low" support monomial. This

will be done by proving that $G(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$ has low support rank concentration in the "$\mathbf{y}$-variables". Applying Observation 15, we will get that $g(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$ has a low support $\mathbf{y}$-monomial. This will then imply that $g(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$ has a low support $\mathbf{x}$-monomial, provided f has low individual degree. Finally, we will plug in the SV generator to preserve the non-zeroness of g. More precisely, we will prove the following theorem at the end of Section 3.2.

▶ **Theorem 21.** *Let f be an n-variate polynomial with individual degree at most d that is computable by a width-w commutative ROABP. If $|\mathbb{F}| \geq n$, then $\mathcal{G}^{SV}_{(2\lceil \log w^2\rceil (d+1)+1)}$ is a hitting set generator for* $\mathrm{orb}(f)$.

Notations and conventions. In the analysis, we will treat $t_1(\mathbf{z}),\ldots,t_n(\mathbf{z})$ as formal variables $\mathbf{t}=(t_1,\ldots,t_n)$ while always keeping in mind the substitution map $t_i \mapsto t_i(\mathbf{z})$. For $i\in[n]$, let $r_i=\ell_i(\mathbf{t})$. For $S\subseteq[n]$, define $\mathbf{r}_S=\{r_i : i\in S\}$. The $\mathbb{F}$-linear independence of $\ell_1,\ldots,\ell_n$ allows us to treat $\mathbf{y}$ and $\mathbf{r}$ as sets of formal variables. Notice that in this notation, $G(\mathbf{x}+\mathbf{t}) = M_1(y_1+r_1)M_2(y_2+r_2)\cdots M_n(y_n+r_n)$. Let $\mathbb{A}$ denote the matrix algebra $\mathbb{F}^{w\times w}$. For $i\in[n]$, let $M_i(y_i)=\sum_{e_i=0}^{d} u_{i,e_i}y_i^{e_i}$, where $u_{i,e_i}\in\mathbb{A}$ and $M_i(y_i+r_i)=\sum_{b_i=0}^{d} v_{i,b_i}y_i^{b_i}$, where $v_{i,b_i}\in\mathbb{A}[r_i]\subset\mathbb{A}[\mathbf{t}]$. As f is a commutative ROABP, $M_1(y_1),\ldots,M_n(y_n)$ commute with each other and hence u_{i,e_i} and u_{j,e_j} also commute for $i\neq j$. The following observation, which we prove in the full version [64], implies that v_{i,e_i} and v_{j,e_j} also commute for $i\neq j$.

▶ **Observation 22.** *For every $i\in[n]$ and $b_i,e_i\in\{0,\ldots,d\}$, $v_{i,b_i}=\sum_{e_i=0}^{d}\binom{e_i}{b_i}\cdot r_i^{e_i-b_i}\cdot u_{i,e_i}$ and $u_{i,e_i}=\sum_{b_i=0}^{d}\binom{b_i}{e_i}\cdot(-r_i)^{b_i-e_i}\cdot v_{i,b_i}$, where $\binom{a}{b}=0$ if $a<b$.*

For a set $S=\{i_1,i_2,\ldots,i_{|S|}\}\subseteq[n]$, where $i_1<i_2<\ldots<i_{|S|}$, the vector $(b_{i_1},b_{i_2},\ldots,b_{i_{|S|}})$ will be denoted by $(b_i : i\in S)$. Let $\mathrm{Supp}\,(\mathbf{b})$ denote the support of the vector $\mathbf{b}$ which is defined as the number of non-zero elements in it.Define the parameter $m:=2\left\lceil\log w^2\right\rceil+1$.

3.1 The goal: low support rank concentration

We set ourselves the goal of proving that there exist explicit degree-n polynomials $t_1(\mathbf{z}),\ldots,t_n(\mathbf{z})$, where $|\mathbf{z}|=2m$, such that $G(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))=M_1(y_1+r_1)M_2(y_2+r_2)\cdots M_n(y_n+r_n)\in\mathbb{A}[r_1,\ldots,r_n][\mathbf{y}]$ has support-$(m-1)$ rank concentration over $\mathbb{F}(\mathbf{z})$ in the $\mathbf{y}$-variables. We will show in this and the next section that this happens if all polynomials in a certain collection of non-zero polynomials $\left\{h_S(\mathbf{r}_S) : S\subseteq\binom{[n]}{m}\right\}\subseteq\mathbb{F}[r_1,\ldots,r_n]$, remain non-zero under the substitution $t_i\mapsto t_i(\mathbf{z})$. The following lemma, proved in the full version [64], will help us achieve this goal.

▶ **Lemma 23.** *Let $G,\mathbf{t},\mathbf{z},\mathbf{y}$ and $\mathbf{r}_S$ be as defined above. Suppose that the following two conditions are satisfied:*

1. *For every $S\subseteq\binom{[n]}{m}$ and $(b_i : i\in S)\in\{0,\ldots,d\}^m$, there is a non-zero polynomial $h_S(\mathbf{r}_S)$ such that $h_S(\mathbf{r}_S)\cdot\prod_{i\in S}v_{i,b_i}\in\mathbb{F}[\mathbf{t}]$-span $\left\{\prod_{i\in S}v_{i,b'_i} \;:\; \mathrm{Supp}\,(b'_i : i\in S)<m\right\}$.*
2. *There exists a substitution $t_i\mapsto t_i(\mathbf{z})$ that keeps $h_S(\mathbf{r}_S)$ non-zero for all $S\subseteq\binom{[n]}{m}$.*

Then, for every $\mathbf{b}=(b_i : i\in[n])\in\{0,\ldots,d\}^n$,

$$\prod_{i\in[n]}v_{i,b_i}\in\mathbb{F}(\mathbf{z})\text{-span}\left\{\prod_{i\in[n]}v_{i,b'_i} \;:\; \mathrm{Supp}\,(b'_i : i\in[n])<m\right\},$$

and $G(x_1+t_1(\mathbf{z}),\ldots,x_n+t_n(\mathbf{z}))$ has support-$(m-1)$ rank concentration over $\mathbb{F}(\mathbf{z})$ in $\mathbf{y}$-variables.

3.2 Achieving rank concentration

We will now see how to satisfy conditions 1 and 2 of Lemma 23 such that $\deg_{\mathbf{r}_S}(h_S(\mathbf{r}_S)) \leq md^{m+1}$, $t_i(\mathbf{z})$ is an explicit degree-n polynomial, and $|\mathbf{z}| = 2m$. Assume wlog that $S = [m]$. For $\mathbf{b} = (b_1, \ldots, b_m)$ and $\mathbf{e} = (e_1, \ldots, e_m)$ in $\{0, \ldots, d\}^m$, define $\binom{\mathbf{b}}{\mathbf{e}} := \prod_{i\in[m]} \binom{b_i}{e_i}$, where, $\binom{b_i}{e_i} = 0$ if $b_i < e_i$. Also, let $v_{\mathbf{b}} := \prod_{i\in[m]} v_{i,b_i}$ and $u_{\mathbf{e}} := \prod_{i\in[m]} u_{i,e_i}$. Define $\mathbf{r} := (-r_1, \ldots, -r_m)$, $\mathbf{r}^{\mathbf{b}} := \prod_{i\in[m]}(-r_i)^{b_i}$ and $\mathbf{r}^{-\mathbf{e}} := \prod_{i\in[m]}(-r_i)^{-e_i}$. We now define some vectors and matrices by fixing an arbitrary order on the elements of $\{0, \ldots, d\}^m$.

Let $V := (v_{\mathbf{b}} : \mathbf{b} \in \{0, \ldots, d\}^m)$ and $U := (u_{\mathbf{e}} : \mathbf{e} \in \{0, \ldots, d\}^m)$; V is a row vector in $\mathbb{A}[\mathbf{r}]^{(d+1)^m}$ whereas U is a row vector in $\mathbb{A}^{(d+1)^m}$. Let $C := \mathrm{diag}(\mathbf{r}^{\mathbf{b}} : \mathbf{b} \in \{0, \ldots, d\}^m)$ and $D := \mathrm{diag}(\mathbf{r}^{-\mathbf{e}} : \mathbf{e} \in \{0, \ldots, d\}^m)$; both C and D are $(d+1)^m \times (d+1)^m$ diagonal matrices. Let M be a $(d+1)^m \times (d+1)^m$ matrix whose rows and columns are indexed by $\mathbf{b} \in \{0, \ldots, d\}^m$ and $\mathbf{e} \in \{0, \ldots, d\}^m$ respectively. The entry of M indexed by $(\mathbf{b}, \mathbf{e})$ contains $\binom{\mathbf{b}}{\mathbf{e}}$. We now make the following claim which is proved in the full version [64].

▷ Claim 24. Let U, V, C, M and D be as defined above. Then, $U = VCMD$.

In [6], a very similar equation was called the *transfer equation* and we will refer to $U = VCMD$ by the same name. Let $F := \{\mathbf{b} \in \{0, \ldots, d\}^m : \mathrm{Supp}(\mathbf{b}) = m\}$; clearly, $|F| = d^m$. Also, let us call the set of all vectors $(n_{\mathbf{e}} : \mathbf{e} \in \{0, \ldots, d\}^m) \in \mathbb{F}^{(d+1)^m}$ for which $\sum_{\mathbf{e}\in\{0,\ldots,d\}^m} n_{\mathbf{e}} u_{\mathbf{e}} = 0$ the *null space* of U. Then, we have the following lemma.

▶ **Lemma 25.** *There are vectors $\{\mathbf{n}_{\mathbf{b}} : \mathbf{b} \in F\}$ in the null space of U such that the following holds: Let N be the $(d+1)^m \times d^m$ matrix whose rows are indexed by $\mathbf{e} \in \{0, \ldots, d\}^m$ and whose columns are indexed by $\mathbf{b} \in F$ and whose column indexed by $\mathbf{b}$ is $\mathbf{n}_{\mathbf{b}}$. Then, the square matrix $[CMDN]_F$ is invertible, where $[CMDN]_F$ is the sub-matrix of $CMDN$ consisting of only those rows of $CMDN$ that are indexed by $\mathbf{b} \in F$.*

We need the value of m in the proof of the lemma which is given in Appendix A. For now, observe that $\det([CMDN]_F) \in \mathbb{F}[\mathbf{r}]$: Every entry of $[CMDN]_F$ is a $\mathbb{F}$-linear combination of some entries of the matrix CMD. The entry of CMD indexed by $(\mathbf{b}, \mathbf{e})$ is $\binom{\mathbf{b}}{\mathbf{e}} \cdot \mathbf{r}^{\mathbf{b}} \cdot \mathbf{r}^{-\mathbf{e}}$, which is non-zero only if $b_i \geq e_i$ for all $i \in [m]$. In this case, $\mathbf{r}^{\mathbf{b}} \cdot \mathbf{r}^{-\mathbf{e}}$ is a monomial in the $\mathbf{r}$-variables. Thus, $\det([CMDN]_F)$ – which is a polynomial in the entries of $[CMDN]_F$ – is a polynomial in the $\mathbf{r}$-variables. This observation leads to the following corollary of the above lemma, which immediately gives a way to satisfy condition 1 of Lemma 23.

▶ **Corollary 26.** *Let $h(\mathbf{r}) := \det([CMDN]_F)$. Then, for every $\mathbf{b} \in F$,*

$$h(\mathbf{r}) \cdot v_{\mathbf{b}} \in \mathbb{F}[\mathbf{t}]\text{-span}\left\{v_{\mathbf{b}'} : \mathbf{b}' \in \{0, \ldots, d\}^m \text{ and } \mathrm{Supp}\left(\mathbf{b}'\right) < m\right\}.$$

The above corollary is proved in the full version [64]. The following claim about $h(\mathbf{r})$ gives us a way to satisfy condition 2 of Lemma 23. It's proof can be found in the full version [64].

▷ Claim 27. The polynomial $h(\mathbf{r})$, when viewed as a polynomial in the $\mathbf{t}$-variables after setting $r_i = \ell_i(\mathbf{t})$, has a $\mathbf{t}$-monomial of support at most m.

By substituting $\mathcal{G}_m^{SV}$ for $\mathbf{t}$, the polynomial $h(\mathbf{r})$ remains non-zero, satisfying condition 2. The number of variables in $\mathcal{G}_m^{SV}$, i.e., $|\mathbf{z}| = 2m$ and its degree is n. The proofs of Theorems 21 and 6 using Lemma 23 can be found in Appendix A.

4 Hitting sets for the orbits of multilinear constant-width ROABPs

The strategy. (Recap) Let $f = \mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot \mathbf{1}$ be a multilinear, width-w ROABP; here $M_i(x_i) \in \mathbb{F}^{w\times w}[x_i]$ for all $i \in [n]$. Also, let $F = M_1(x_1)M_2(x_2)\cdots M_n(x_n)$. For any $A \in \mathrm{GL}(n,\mathbb{F})$, let $g = f(A\mathbf{x})$ and $G = F(A\mathbf{x})$. For $i \in [n]$, suppose that A maps $x_i \mapsto \ell_i(\mathbf{x})$, where ℓ_i is a linear form, and let $y_i = \ell_i(\mathbf{x})$ and $\mathbf{y} = \{y_1,\ldots,y_n\}$. Then, $g = \mathbf{1}^T \cdot M_1(y_1)M_2(y_2)\cdots M_n(y_n)\cdot \mathbf{1}$ and $G = M_1(y_1)M_2(y_2)\cdots M_n(y_n)$. Just like in the previous section, we will show that if $g \neq 0$, then there exist explicit "low" degree polynomials $t_1(\mathbf{z}),\ldots,t_n(\mathbf{z})$, where $\mathbf{z}$ is a "small" set of variables such that $G(x_1 + t_1(\mathbf{z}),\ldots,x_n + t_n(\mathbf{z}))$ has "low" support rank concentration in the "$\mathbf{y}$-variables". While in the rank concentration argument in the previous section the $\mathbf{x}$-variables were translated only once, here the translations can be thought of as happening sequentially and in stages. There will be $\lceil \log n\rceil$ stages with each stage also consisting of multiple translations. After the p-th stage, the product of any 2^p consecutive matrices in G will have low support rank concentration in the $\mathbf{y}$-variables. Thus, after $\lceil \log n\rceil$ stages, we will have low support rank concentration in the $\mathbf{y}$-variables for $G(x_1 + t_1(\mathbf{z}),\ldots,x_n + t_n(\mathbf{z}))$.

Notations and conventions. Much like in the previous section, we will first translate the $\mathbf{x}$-variables by the $\mathbf{t}$-variables and then substitute the $\mathbf{t}$-variables by low degree polynomials in a small set of variables. We will translate the $\mathbf{x}$-variables by $\lceil \log n\rceil$ groups of $\mathbf{t}$-variables, $\mathbf{t}_1,\ldots,\mathbf{t}_{\lceil \log n\rceil}$. For all $p \in \lceil \log n\rceil$, the group $\mathbf{t}_p$ will have $\mu := w^2 + \lceil \log w^2\rceil$ sub-groups of $\mathbf{t}$-variables, $\mathbf{t}_{p,1},\ldots,\mathbf{t}_{p,\mu}$. For all $p \in \lceil \log n\rceil$ and $q \in [\mu]$, $\mathbf{t}_{p,q} := \{t_{p,q,1},\ldots,t_{p,q,n}\}$. Thus, finally the translation will look like $x_i \to x_i + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} t_{p,q,i}$ for all $i \in [n]$. Finally, we will substitute the $\mathbf{t}$-variables as $t_{p,q,i} \mapsto s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(i)}$, where $\beta_{p,q}(i)$ will be fixed later in the analysis. Let $r_{p,q,i} := \ell_i(\mathbf{t}_{p,q})$; notice that for all $i \in [n]$, y_i is translated as $y_i \to y_i + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} \ell_i(\mathbf{t}_{p,q}) = y_i + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} r_{p,q,i}$.

For the purpose of analysis, we will think of the translation as happening sequentially in the order $\mathbf{t}_{1,1},\ldots,\mathbf{t}_{1,\mu},\mathbf{t}_{2,1},\ldots,\mathbf{t}_{2,\mu},\ldots,\mathbf{t}_{n,1},\ldots\mathbf{t}_{n,\mu}$, i.e., we will first translate by $\mathbf{t}_{1,1}$, then by $\mathbf{t}_{1,2}$, and so on. We denote the order thus imposed on the set $\{(p,q) : p \in [\lceil \log n\rceil], q \in [\mu]\}$ by $\prec$.

For a set $S = \{i_1, i_2, \ldots, i_{|S|}\} \subseteq [n]$, where $i_1 < i_2 < \ldots < i_{|S|}$, the vector $(b_{i_1}, b_{i_2}, \ldots, b_{i_{|S|}})$ will be denoted by $(b_i : i \in S)$. Let $\mathrm{Supp}(\mathbf{b})$ denote the support of the vector $\mathbf{b}$ which is defined as the number of non-zero elements in it. The inductive argument given on the next two subsections is inspired by the "merge-and-reduce" idea from [21, 23].

4.1 Low support rank concentration: an inductive argument

In this and the next sections, we will prove the following lemma. Let $\mathbb{A} := \mathbb{F}^{w\times w}$.

▶ **Lemma 28.** *There exist $\{\beta_{p,q}(i) : p \in [\lceil \log n\rceil], q \in [\mu], i \in [n]\} \subset \mathbb{Z}_{\geq 0}$, such that when we treat $G\left(x_1 + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(1)}, \ldots, x_n + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(n)}\right)$, as a polynomial in the $\mathbf{y}$-variables over $\mathbb{A}[r_{p,q,i} : p \in [\lceil \log n\rceil], q \in [\mu], i \in [n]]$, has support-$\mu$ rank concentration in $\mathbf{y}$-variables over $\mathbb{F}(s_{p,q}, z_{p,q} : p \in [\lceil \log n\rceil], q \in [\mu])$. The $\beta_{p,q}(i)$s can be found in time $n^{O(w^4)}$ and each $\beta_{p,q}(i) \leq n^{O(w^4)}$.*

We will prove this lemma by induction on (p,q). Let us call $\beta_{p,q}(i)$s *efficiently computable and good* if they can be found in time $n^{O(w^4)}$ and each $\beta_{p,q}(i) \leq n^{O(w^4)}$. Precisely, the induction hypothesis is as follows.

Induction hypothesis. Just before translating by $\mathbf{t}_{p^*,q^*}$-variables, assume that there exist efficiently computable and good $\{\beta_{p,q}(i) : (p,q) \prec (p^*,q^*)\}$ such that the product of any 2^{p^*} consecutive matrices in

$$G\left(x_1 + \sum_{(p,q)\prec(p^*,q^*)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(1)}\ ,\ldots,\ x_n + \sum_{(p,q)\prec(p^*,q^*)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(n)}\right)$$

has support-$(2\mu-(q^*-1))$ rank concentration over $\mathbb{F}\left(s_{p,q}, z_{p,q} : (p,q)\prec(p^*,q^*)\right)$ in $\mathbf{y}$-variables.

Base case. In the base case, $(p^*,q^*)=(1,1)$. Observe that we can assume that $w\geq 2$; if $w=1$, then g is a product of univariates and the existence of a polynomial time hitting set follows from Observation 13. For any $w\geq 2$, $2\leq 2\mu$. As a product of any two consecutive matrices in G has support $2\leq 2\mu$ rank concentration in the $\mathbf{y}$-variables over $\mathbb{F}$, the base case is satisfied.

Induction step. We need to show that there exist $\{\beta_{p^*,q^*}(i) : i\in[n]\}$ which are efficiently computable and good, such that after translating by $\mathbf{t}_{p^*,q^*}$ and substituting $t_{p^*,q^*,i}\to s_{p^*,q^*}\cdot z_{p^*,q^*}^{\beta_{p^*,q^*}(i)}$, the product of any 2^{p^*} consecutive matrices in

$$G\left(x_1 + \sum_{(p,q)\preccurlyeq(p^*,q^*)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(1)}\ ,\ldots,\ x_n + \sum_{(p,q)\preccurlyeq(p^*,q^*)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(n)}\right)$$

has support-$(2\mu-q^*)$ rank concentration in the $\mathbf{y}$-variables over $\mathbb{F}\left(s_{p,q}, z_{p,q} : (p,q)\preccurlyeq(p^*,q^*)\right)$. If $q^*<\mu$, then this would mean that the induction hypothesis holds immediately before translation by $\mathbf{t}_{p^*,q^*+1}$. Otherwise, if $q^*=\mu$, then the following easy-to-verify observation implies that the induction hypothesis holds immediately before translation by $\mathbf{t}_{p^*+1,1}$.

▶ **Observation 29.** *Suppose that $\{\beta_{p,q}(i) : (p,q)\preccurlyeq(p^*,\mu)\}$ are such that the product of any 2^{p^*} consecutive matrices in*

$$G\left(x_1 + \sum_{(p,q)\preccurlyeq(p^*,\mu)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(1)}\ ,\ldots,\ x_n + \sum_{(p,q)\preccurlyeq(p^*,\mu)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(n)}\right)$$

has support-μ rank concentration in $\mathbf{y}$-variables over $\mathbb{F}\left(s_{p,q}, z_{p,q} : (p,q)\preccurlyeq(p^*,\mu)\right)$. Then the product of any 2^{p^*+1} consecutive matrices in

$$G\left(x_1 + \sum_{(p,q)\preccurlyeq(p^*,\mu)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(1)}\ ,\ldots,\ x_n + \sum_{(p,q)\preccurlyeq(p^*,\mu)} s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(n)}\right)$$

has support-2μ rank concentration in the $\mathbf{y}$-variables over $\mathbb{F}\left(s_{p,q}, z_{p,q} : (p,q)\preccurlyeq(p^*,\mu)\right)$.

Simplifying notations for the ease of exposition. By focusing on the induction step, we will henceforth denote $\mathbb{F}\left(s_{p,q}, z_{p,q} : (p,q)\prec(p^*,q^*)\right)$ by $\mathbb{F}$, and for all $i\in[n]$,

$$M_i\left(y_j + \sum_{(p,q)\prec(p^*,q^*)} \ell_i\left(s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(1)},\ldots,s_{p,q}\cdot z_{p,q}^{\beta_{p,q}(n)}\right)\right)$$

by $M_i(y_i)$, $t_{p^*,q^*,i}$ by t_i, $r_{p^*,q^*,i}$ by r_i, s_{p^*,q^*} by s, z_{p^*,q^*} by z and $\beta_{p^*,q^*}(i)$ by $\beta(i)$.

Without loss of generality, we shall consider the product $M_1(y_1 + r_1) \cdots M_m(y_n + r_m)$ of the first $m = 2^{p^*}$ matrices. Our goal is to show that there exist efficiently computable and good $\{\beta(i) : i \in [m]\}$ such that after substituting $t_i \to s \cdot z^{\beta(i)}$, the above product has support-$(2\mu - q^*)$ rank concentration in the $\mathbf{y}$-variables over $\mathbb{F}(s, z)$ *assuming that* $M_1(y_1) \cdots M_m(y_m)$ *has support-*$(2\mu - (q^* - 1))$ *rank concentration in the* $\mathbf{y}$*-variables over* $\mathbb{F}$.

4.2 Details of the induction step

Recalling some notations. Before we show how to achieve rank concentration, let us recall some notations defined in Section 3. While in Section 3, the individual degree is d, here the individual degree is 1 and so, we modify the definitions accordingly. $\mathbb{A}$ is used to denote the matrix algebra $\mathbb{F}^{w \times w}$. For $i \in [m]$, $M_i(y_i) = \sum_{e_i=0}^{1} u_{i,e_i} y_i^{e_i}$, where $u_{i,e_i} \in \mathbb{A}$ and $M_i(y_i + r_i) = \sum_{b_i=0}^{1} v_{i,b_i} y_i^{b_i}$, where $v_{i,b_i} \in \mathbb{A}[r_i] \subset \mathbb{A}[\mathbf{t}]$. For $\mathbf{b} = (b_1, \ldots, b_m)$ and $\mathbf{e} = (e_1, \ldots, e_m)$ in $\{0,1\}^m$, $\binom{\mathbf{b}}{\mathbf{e}} := \prod_{i \in [m]} \binom{b_i}{e_i}$. Also, $v_{\mathbf{b}} := \prod_{i \in [m]} v_{i,b_i}$ and $u_{\mathbf{e}} := \prod_{i \in [m]} u_{i,e_i}$. Moreover, $\mathbf{r} := (-r_1, \ldots, -r_m)$, $\mathbf{r}^{\mathbf{b}} := \prod_{i \in [m]} (-r_i)^{b_i}$ and $\mathbf{r}^{-\mathbf{e}} := \prod_{i \in [m]} (-r_i)^{-e_i}$. Let $\mathbf{t} := (t_1, \ldots, t_n)$.

The following vectors and matrices are defined by fixing an arbitrary order on the elements of $\{0,1\}^m$. $V := (v_{\mathbf{b}} : \mathbf{b} \in \{0,1\}^m)$ and $U := (u_{\mathbf{e}} : \mathbf{e} \in \{0,1\}^m)$; V is a row vector in $\mathbb{A}[\mathbf{r}]^{2^m}$ whereas U is a row vector in $\mathbb{A}^{2^m}$. Both $C := \mathrm{diag}(\mathbf{r}^{\mathbf{b}} : \mathbf{b} \in \{0,1\}^m)$ and $D := \mathrm{diag}(\mathbf{r}^{-\mathbf{e}} : \mathbf{e} \in \{0,1\}^m)$ are $2^m \times 2^m$ diagonal matrices. Finally, M is a $2^m \times 2^m$ numeric matrix whose rows and columns were indexed by $\mathbf{b} \in \{0,1\}^m$ and $\mathbf{e} \in \{0,1\}^m$, respectively. The entry of M indexed by $(\mathbf{b}, \mathbf{e})$ contains $\binom{\mathbf{b}}{\mathbf{e}}$. The proof of the following transfer equation is same as the proof of Claim 24.

▷ Claim 30. Let U, V, C, M and D be as defined above. Then, $U = VCMD$.

Let $F := \{\mathbf{b} \in \{0,1\}^m : \mathrm{Supp}(\mathbf{b}) > 2\mu - q^*\}$. Also, recall that the the *null space* of U is the set of all vectors $(n_{\mathbf{e}} : \mathbf{e} \in \{0,1\}^m) \in \mathbb{F}^{2^m}$ for which $\sum_{\mathbf{e} \in \{0,1\}^m} n_{\mathbf{e}} u_{\mathbf{e}} = 0$. We have the following lemma.

▶ **Lemma 31.** *There are vectors* $\{\mathbf{n}_{\mathbf{b}} : \mathbf{b} \in F\}$ *in the null space of* U *such that the following holds: Let* N *be the* $2^m \times |F|$ *matrix whose rows are indexed by* $\mathbf{e} \in \{0,1\}^m$ *and whose columns are indexed by* $\mathbf{b} \in F$ *and whose* $\mathbf{b}$*-th column is* $\mathbf{n}_{\mathbf{b}}$*. Then, the square matrix* $[CMDN]_F$ *is invertible, where* $[CMDN]_F$ *is the sub-matrix of* $CMDN$ *consisting of only those rows of* $CMDN$ *that are indexed by* F*. Also,* $\det([CMDN]_F) \in \mathbb{F}[\mathbf{r}] \subset \mathbb{F}[\mathbf{t}]$ *can be expressed as the ratio of a polynomial in* $\mathbb{F}[\mathbf{t}]$ *that contains a monomial of degree at most* $2w^2\mu$ *in the* $\mathbf{t}$*-variables and a product of linear forms in* $\mathbb{F}[\mathbf{t}]$.

The proof of this lemma, which uses the value of μ, is given in the full version [64]. We now complete the induction step using this lemma. As $\det([CMDN]_F)$ is a polynomial in $\mathbb{F}[\mathbf{r}]$ we get the following corollaries.

▶ **Corollary 32.** *Let* $h(\mathbf{r}) := \det([CMDN]_F)$*. Then, for every* $\mathbf{b} \in F$*,*

$$h(\mathbf{r}) \cdot v_{\mathbf{b}} \in \mathbb{F}[\mathbf{t}]\text{-span}\left\{v_{\mathbf{b}'} : \mathbf{b}' \in \{0,1\}^m \text{ and } \mathrm{Supp}\left(\mathbf{b}'\right) \leq 2\mu - q^*\right\}. \tag{1}$$

Proof. Same as the proof of Corollary 26. ◀

▶ **Corollary 33.** *Suppose* $\{\beta(i) : i \in [n]\}$ *are such that the substitution* $t_i \mapsto s \cdot z^{\beta(i)}$ *keeps all non-zero polynomials in* $\mathbb{F}[\mathbf{t}]$ *containing a monomial of degree at most* $2w^2\mu$ *in the* $\mathbf{t}$*-variables non-zero. Then, the product* $M_1(y_1 + r_1) \cdots M_m(y_m + r_m)$ *has support-*$(2\mu - q^*)$ *rank concentration in the* $\mathbf{y}$*-variables over* $\mathbb{F}(s, z)$ *after substituting* $t_i \to s \cdot z^{\beta(i)}$.

Proof. Multiply both sides of (1) by $(h(\mathbf{r}))^{-1}$ after substituting $t_i \mapsto s \cdot z^{\beta(i)}$. ◀

The following claim, proved in the full version [64], allows us to compute $\{\beta(i) : i \in [n]\}$ efficiently.

▷ Claim 34. There exist $\{\beta(i) : i \in [n]\}$ such that the substitution $t_i \mapsto s \cdot z^{\beta(i)}$ keeps all non-zero polynomials in $\mathbb{F}[\mathbf{t}]$ containing a monomial of degree at most $2w^2\mu$ in the $\mathbf{t}$-variables non-zero. Moreover, we can find all the $\beta(i)$ in time $n^{O(w^4)}$ and each $\beta(i) \leq n^{O(w^4)}$.

This completes the induction step. Lemma 28 and Theorem 8 are proved in Appendix B.

5 Conclusion

In this paper, we have given efficient hitting sets for orbits of several well-studied circuit classes such as commutative ROABPs and constant-width ROABPs (under the low individual degree restriction), and constant-depth constant-occur formulas and occur-once formulas. In the process, we have obtained efficiently constructible hitting sets for the orbits of the elementary symmetric and power symmetric and sum-product polynomials as well as the iterated matrix multiplication polynomials of width-3, which is a complete family of polynomials for arithmetic formulas under p-projections. The hitting set problem for the orbits of these circuit classes and polynomial families is interesting as their affine projections capture much larger circuit classes and orbits are a natural and dense subset of the set of affine projections. However, the following questions still remain open:

- **Removing the low individual degree restriction.** The low individual degree restriction is natural as it subsumes the multilinear case. However, it would be ideal if we get rid of this limitation of our results. In particular, can we give an efficient hitting-set construction for the orbits of general commutative ROABPs and constant-width ROABPs?
- **Lower bound and hitting set for the orbits of ROABPs.** We would also like to remove the requirements of commutativity and constant-width from our results on hitting sets for the orbits of ROABPs. It is worth noting that an explicit hitting set for the orbits of ROABPs implies a lower bound for the same model computing some explicit polynomial [1]. To our knowledge, no explicit lower bound is known for the orbits of ROABPs. Can we prove such a lower bound first?
- **Hitting sets for the orbits of Det and IMM.** The determinant (Det) and the iterated matrix multiplication (IMM) polynomial families are complete for the class of algebraic branching programs under p-projections. Can we design efficiently constructible hitting sets for the orbits of Det and IMM?

References

1 Manindra Agrawal. Proving lower bounds via pseudo-random generators. In Ramaswamy Ramanujam and Sandeep Sen, editors, *FSTTCS 2005: Foundations of Software Technology and Theoretical Computer Science, 25th International Conference, Hyderabad, India, December 15-18, 2005, Proceedings*, volume 3821 of *Lecture Notes in Computer Science*, pages 92–105. Springer, 2005.

2 Manindra Agrawal and Somenath Biswas. Primality and identity testing via Chinese remaindering. *J. ACM*, 50(4):429–443, 2003. Conference version appeared in the proceedings of FOCS 1999.

3 Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-sets for ROABP and sum of set-multilinear circuits. *SIAM J. Comput.*, 44(3):669–697, 2015.

4 Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. PRIMES is in P. *Annals of Mathematics*, 160(2):781–793, 2004.

5 Manindra Agrawal, Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. Jacobian Hits Circuits: Hitting Sets, Lower Bounds for Depth-D Occur-k Formulas and Depth-3 Transcendence Degree-k Circuits. *SIAM J. Comput.*, 45(4):1533–1562, 2016. Conference version appeared in the proceedings of STOC 2012.

6 Manindra Agrawal, Chandan Saha, and Nitin Saxena. Quasi-polynomial hitting-set for set-depth-Δ formulas. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 321–330. ACM, 2013.

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

8 Eric Allender and Fengming Wang. On the power of algebraic branching programs of width two. *Comput. Complex.*, 25(1):217–253, 2016. Conference version appeared in the proceedings of ICALP 2011.

9 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. Conference version appeared in the proceedings of CCC 2016.

10 Matthew Anderson, Dieter van Melkebeek, and Ilya Volkovich. Deterministic polynomial identity tests for multilinear bounded-read formulae. *Comput. Complex.*, 24(4):695–776, 2015. Conference version appeared in the proceedings of CCC 2011.

11 Malte Beecken, Johannes Mittmann, and Nitin Saxena. Algebraic independence and blackbox identity testing. *Inf. Comput.*, 222:2–19, 2013. Conference version appeared in the proceedings of ICALP 2011.

12 Michael Ben-Or and Richard Cleve. Computing Algebraic Formulas Using a Constant Number of Registers. *SIAM J. Comput.*, 21(1):54–58, 1992. Conference version appeared in the proceedings of STOC 1988.

13 Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam. On algebraic branching programs of small width. *J. ACM*, 65(5):32:1–32:29, 2018. Conference version appeared in the proceedings of CCC 2017.

14 Rafael Mendes de Oliveira, Amir Shpilka, and Ben lee Volk. Subexponential Size Hitting Sets for Bounded Depth Multilinear Formulas. *Comput. Complex.*, 25(2):455–505, 2016. Conference version appeared in the proceedings of CCC 2015.

15 Richard A. DeMillo and Richard J. Lipton. A Probabilistic Remark on Algebraic Program Testing. *Inf. Process. Lett.*, 7(4):193–195, 1978.

16 Zeev Dvir and Amir Shpilka. Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits. *SIAM J. Comput.*, 36(5):1404–1434, 2007. Conference version appeared in the proceedings of STOC 2005.

17 Jack Edmonds. Systems of distinct representatives and linear algebra. *Journal of research of the National Bureau of Standards*, 71:241–245, 1967.

18 Jack Edmonds. Matroid intersection. In P.L. Hammer, E.L. Johnson, and B.H. Korte, editors, *Discrete Optimization I*, volume 4 of *Annals of Discrete Mathematics*, pages 39–49. Elsevier, 1979.

19 Stephen A. Fenner, Rohit Gurjar, and Thomas Thierauf. Bipartite perfect matching is in quasi-nc. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 754–763. ACM, 2016.

20 Michael A. Forbes. Deterministic divisibility testing via shifted partial derivatives. In Venkatesan Guruswami, editor, *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 451–465. IEEE Computer Society, 2015.

21 Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 867–875. ACM, 2014.

22 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.

23 Michael A. Forbes and Amir Shpilka. Quasipolynomial-time identity testing of non-commutative and read-once oblivious algebraic branching programs. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 243–252. IEEE Computer Society, 2013.

24 Michael A. Forbes and Amir Shpilka. A PSPACE construction of a hitting set for the closure of small algebraic circuits. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1180–1192. ACM, 2018.

25 Ariel Gabizon and Ran Raz. Deterministic extractors for affine sources over large fields. *Comb.*, 28(4):415–440, 2008. Conference version appeared in the proceedings of FOCS 2005.

26 James F. Geelen. Maximum rank matrix completion. *Linear Algebra and its Applications*, 288:211 – 217, 1999.

27 Zeyu Guo and Rohit Gurjar. Improved explicit hitting-sets for roabps. In Jaroslaw Byrka and Raghu Meka, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2020, August 17-19, 2020, Virtual Conference*, volume 176 of *LIPIcs*, pages 4:1–4:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

28 Zeyu Guo, Nitin Saxena, and Amit Sinhababu. Algebraic dependencies and PSPACE algorithms in approximative complexity. In Rocco A. Servedio, editor, *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, volume 102 of *LIPIcs*, pages 10:1–10:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

29 Ankit Gupta. Algebraic geometric techniques for depth-4 PIT & sylvester-gallai conjectures for varieties. *Electron. Colloquium Comput. Complex.*, 21:130, 2014. URL: `http://eccc.hpi-web.de/report/2014/130`.

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

31 Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Identity testing for constant-width, and any-order, read-once oblivious arithmetic branching programs. *Theory Comput.*, 13(1):1–21, 2017. Conference version appeared in the proceedings of CCC 2016.

32 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. Conference version appeared in the proceedings of CCC 2015.

33 Rohit Gurjar and Thomas Thierauf. Linear matroid intersection is in quasi-nc. *Comput. Complex.*, 29(2):9, 2020. Conference version appeared in the proceedings of STOC 2017.

34 Joos Heintz and Claus-Peter Schnorr. Testing polynomials which are easy to compute (extended abstract). In Raymond E. Miller, Seymour Ginsburg, Walter A. Burkhard, and Richard J. Lipton, editors, *Proceedings of the 12th Annual ACM Symposium on Theory of Computing, April 28-30, 1980, Los Angeles, California, USA*, pages 262–272. ACM, 1980.

35 Gábor Ivanyos, Marek Karpinski, and Nitin Saxena. Deterministic polynomial time algorithms for matrix completion problems. *SIAM J. Comput.*, 39(8):3736–3751, 2010.

36 Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Comput. Complex.*, 13(1-2):1–46, 2004. Conference version appeared in the proceedings of STOC 2003.

37 Erich Kaltofen. Factorization of polynomials given by straight-line programs. *Adv. Comput. Res.*, 5:375–412, 1989.

38 Erich Kaltofen and Barry M. Trager. Computing with Polynomials Given By Black Boxes for Their Evaluations: Greatest Common Divisors, Factorization, Separation of Numerators and Denominators. *J. Symb. Comput.*, 9(3):301–320, 1990. Conference version appeared in the proceedings of FOCS 1988.

39 Zohar Shay Karnin, Partha Mukhopadhyay, Amir Shpilka, and Ilya Volkovich. Deterministic Identity Testing of Depth-4 Multilinear Circuits with Bounded Top Fan-in. *SIAM J. Comput.*, 42(6):2114–2131, 2013. Conference version appeared in the proceedings of STOC 2010.

40 Zohar Shay Karnin and Amir Shpilka. Black box polynomial identity testing of generalized depth-3 arithmetic circuits with bounded top fan-in. *Comb.*, 31(3):333–364, 2011. Conference version appeared in the proceedings of CCC 2008.

41 Richard M. Karp, Eli Upfal, and Avi Wigderson. Constructing a perfect matching is in random NC. *Comb.*, 6(1):35–48, 1986. Conference version appeared in the proceedings of STOC 1985.

42 Neeraj Kayal. Algorithms for arithmetic circuits. *Electron. Colloquium Comput. Complex.*, 17:73, 2010. URL: `http://eccc.hpi-web.de/report/2010/073`.

43 Neeraj Kayal and Shubhangi Saraf. Blackbox polynomial identity testing for depth 3 circuits. In *50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA*, pages 198–207. IEEE Computer Society, 2009.

44 Neeraj Kayal and Nitin Saxena. Polynomial identity testing for depth 3 circuits. *Comput. Complex.*, 16(2):115–138, 2007. Conference version appeared in the proceedings of CCC 2006.

45 Adam R. Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, pages 216–223, 2001.

46 Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theor. Comput. Sci.*, 448:56–65, 2012.

47 Pascal Koiran and Mateusz Skomra. Derandomization and absolute reconstruction for sums of powers of linear forms. *CoRR*, abs/1912.02021, 2019. URL: `http://arxiv.org/abs/1912.02021`.

48 Swastik Kopparty, Shubhangi Saraf, and Amir Shpilka. Equivalence of polynomial identity testing and polynomial factorization. *Comput. Complex.*, 24(2):295–331, 2015. Conference version appeared in the proceedings of CCC 2014.

49 László Lovász. On determinants, matchings, and random algorithms. In Lothar Budach, editor, *Fundamentals of Computation Theory, FCT 1979, Proceedings of the Conference on Algebraic, Arithmetic, and Categorial Methods in Computation Theory, Berlin/Wendisch-Rietz, Germany, September 17-21, 1979*, pages 565–574. Akademie-Verlag, Berlin, 1979.

50 László Lovász. Singular spaces of matrices and their application in combinatorics. *Boletim da Sociedade Brasileira de Matemática - Bulletin/Brazilian Mathematical Society*, 20(1):87–99, 1989.

51 Dori Medini and Amir Shpilka. Hitting Sets and Reconstruction for Dense Orbits in VP_e and $\Sigma\Pi\Sigma$ Circuits. *CoRR*, abs/2102.05632, 2021. URL: `https://arxiv.org/abs/2102.05632`.

52 Daniel Minahan and Ilya Volkovich. Complete derandomization of identity testing and reconstruction of read-once formulas. *ACM Trans. Comput. Theory*, 10(3):10:1–10:11, 2018. Conference version appeared in the proceedings of CCC 2017.

53 Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. Matching is as easy as matrix inversion. *Comb.*, 7(1):105–113, 1987. Conference version appeared in the proceedings of STOC 1987.

54 K. Murota. Mixed matrices: Irreducibility and decomposition. In R. A. Brualdi, S. Friedland, and V. Klee, editors, *Combinatorial and Graph-Theoretical Problems in Linear Algebra. The IMA Volumes in Mathematics and its Applications, vol 50.*, pages 39–71. Springer, New York, NY, 1993.

55 H. Narayanan, Huzur Saran, and Vijay V. Vazirani. Randomized Parallel Algorithms for Matroid Union and Intersection, With Applications to Arboresences and Edge-Disjoint Spanning Trees. *SIAM J. Comput.*, 23(2):387–397, 1994. Conference version appeared in the proceedings of SODA 1992.

56 Noam Nisan. Lower Bounds for Non-Commutative Computation (Extended Abstract). In Cris Koutsougeras and Jeffrey Scott Vitter, editors, *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing, May 5-8, 1991, New Orleans, Louisiana, USA*, pages 410–418. ACM, 1991.

57 Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994. Conference version appeared in the proceedings of FOCS 1988.

58 Noam Nisan and Avi Wigderson. Lower Bounds on Arithmetic Circuits Via Partial Derivatives. *Computational Complexity*, 6(3):217–234, 1997. Conference version appeared in the proceedings of FOCS 1995.

59 Shir Peleg and Amir Shpilka. A generalized sylvester-gallai type theorem for quadratic polynomials. In Shubhangi Saraf, editor, *35th Computational Complexity Conference, CCC 2020, July 28-31, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 169 of *LIPIcs*, pages 8:1–8:33. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

60 Shir Peleg and Amir Shpilka. Polynomial time deterministic identity testing algorithm for $\Sigma^{[3]}\Pi\Sigma\Pi^{[2]}$ circuits via Edelstein-Kelly type theorem for quadratic polynomials. *CoRR*, abs/2006.08263, 2020.

61 Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Comput. Complex.*, 14(1):1–19, 2005. Conference version appeared in the proceedings of CCC 2004.

62 Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. The power of depth 2 circuits over algebras. In Ravi Kannan and K. Narayan Kumar, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2009, December 15-17, 2009, IIT Kanpur, India*, volume 4 of *LIPIcs*, pages 371–382. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2009.

63 Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. A case of depth-3 identity testing, sparse factorization and duality. *Comput. Complex.*, 22(1):39–69, 2013.

64 Chandan Saha and Bhargav Thankey. Hitting Sets for Orbits of Circuit Classes and Polynomial Families. *Electron. Colloquium Comput. Complex.*, 28:15, 2021. URL: `https://eccc.weizmann.ac.il/report/2021/015`.

65 Shubhangi Saraf and Ilya Volkovich. Black-Box Identity Testing of Depth-4 Multilinear Circuits. *Comb.*, 38(5):1205–1238, 2018. Conference version appeared in the proceedings of STOC 2011.

66 Nitin Saxena. Diagonal circuit identity testing and lower bounds. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part I: Tack A: Algorithms, Automata, Complexity, and Games*, volume 5125 of *Lecture Notes in Computer Science*, pages 60–71. Springer, 2008.

67 Nitin Saxena. Progress on polynomial identity testing. *Bull. EATCS*, 99:49–79, 2009.

68 Nitin Saxena. Progress on polynomial identity testing-ii. In M. Agrawal and V. Arvind, editors, *Perspectives in Computational Complexity*, volume 26 of *Progress in Computer Science and Applied Logic*, pages 131–146. Birkhäuser, Cham, 2014.

69 Nitin Saxena and C. Seshadhri. Blackbox Identity Testing for Bounded Top-Fanin Depth-3 Circuits: The Field Doesn't Matter. *SIAM J. Comput.*, 41(5):1285–1298, 2012. Conference version appeared in the proceedings of STOC 2011.

70 Nitin Saxena and C. Seshadhri. From sylvester-gallai configurations to rank bounds: Improved blackbox identity test for depth-3 circuits. *J. ACM*, 60(5):33:1–33:33, 2013. Conference version appeared in the proceedings of FOCS 2010.

71 Jacob T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *J. ACM*, 27(4):701–717, 1980.

72 Amir Shpilka. Sylvester-gallai type theorems for quadratic polynomials. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 1203–1214. ACM, 2019.

73 Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. *Comput. Complex.*, 24(3):477–532, 2015. Conference versions appeared in the proceedings of STOC 2008 and APPROX-RANDOM 2009.

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

75 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.

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

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

78 Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation, EUROSAM '79, An International Symposiumon Symbolic and Algebraic Computation, Marseille, France, June 1979, Proceedings*, pages 216–226, 1979.

A Missing proofs from Section 3

A.1 Proof of Lemma 25

The entries of U, the columns of M, the rows and columns of D, and the rows of N are indexed by $\mathbf{e} \in \{0,\dots,d\}^m$. Impose an order $\prec$, say the lexicographical order, on the indices $\mathbf{e} \in \{0,\dots,d\}^m$ of U and the other three matrices. Pick the *minimal* basis of the space spanned by the entries of U according to this order, i.e., consider the entries of U in the order dictated by $\prec$ while forming the basis. Let $\mathcal{B} := \{\mathbf{e} \in \{0,\dots,d\}^m : u_{\mathbf{e}} \text{ is in the minimal basis of } U \text{ w.r.t. } \prec\}$.

Construction of the matrix N. The columns of N are indexed by $\mathbf{b} \in F$. We will now specify a set of column vectors $\{\mathbf{n_b} : \mathbf{b} \in F\}$ in the null space of U such that the column of N indexed by $\mathbf{b} \in F$ is $\mathbf{n_b}$. There are two cases for $\mathbf{b} \in F$:

Case 1: $\mathbf{b} \in F \setminus \mathcal{B}$. In this case, $u_{\mathbf{b}}$ is dependent on $\{u_{\mathbf{e}} : \mathbf{e} \in \mathcal{B} \text{ and } \mathbf{e} \prec \mathbf{b}\}$. Pick this dependence vector as $\mathbf{n_b}$.

Case 2: $\mathbf{b} \in F \cap \mathcal{B}$. Let there be p such $\mathbf{b}$, where $p \le |\mathcal{B}| \le w^2$. For a set $E \subseteq [m]$ and $\mathbf{b} \in \{0,\dots,d\}^m$, let $(\mathbf{b})_E$ denote the vector obtained by projecting $\mathbf{b}$ to the coordinates in E. Roughly speaking, the following claim which is proved in the full version [64] says that each of these p vectors has a "small signature" that differentiates it from the other $p-1$ vectors.

▷ Claim 35. There exists a way of numbering all $\mathbf{b} \in F \cap \mathcal{B}$ as $\mathbf{b}_1,\dots,\mathbf{b}_p$ and there exist non-empty sets $E_1,\dots,E_p \subseteq [m]$, each of size at most $\log p \le \log w^2$ such that for all $k \in [p-1]$,

$$(\mathbf{b}_k)_{E_k} \neq (\mathbf{b}_\ell)_{E_k} \quad \forall \ell \in \{k+1,\dots,p\} \tag{2}$$

We will call E_k the *signature* of $\mathbf{b}_k$ for $k \in [p]$. The following claim tells us that for each vector $\mathbf{b}_k$, there is a vector that is not in $\mathcal{B}$ and has support at most $m-1$, but agrees with $\mathbf{b}_k$ on its signature and so in some sense can be used as a proxy for $\mathbf{b}_k$.

▷ Claim 36. For every $k \in [p]$, there exists a vector $\mathbf{b}'_k \in \{0,\ldots,d\}^m \setminus (F \cup \mathcal{B})$ such that $(\mathbf{b}'_k)_{E_k} = (\mathbf{b}_k)_{E_k}$ and also $\mathbf{b}'_k$ and $\mathbf{b}_k$ agree on all locations where $\mathbf{b}'_k$ is non-zero.

A proof of the above claim is provided in the full version [64]. We will now use the above two claims to construct $\mathbf{n}_{\mathbf{b}_k}$ for all $k \in [p]$. We will use $\mathbf{b}'_k$ from Claim 36 as a proxy for $\mathbf{b}_k$. Notice that $u_{\mathbf{b}'_k}$ is dependent on $\{u_{\mathbf{e}} : \mathbf{e} \in \mathcal{B} \text{ and } \mathbf{e} \prec \mathbf{b}'_k\}$. Let this dependence vector be $\mathbf{n}_{\mathbf{b}_k}$. This completes the construction of N. We will now show that $[CMDN]_F$ is an invertible matrix.

$[CMDN]_F$ is invertible. As C is a diagonal matrix with non-zero entries, it is sufficient to show that $[MDN]_F = [M]_F DN$ is an invertible matrix, where $[M]_F$ is the sub-matrix of M consisting of only those rows of M that are indexed by $\mathbf{b} \in F$. The following claim lets us simplify the structure of $[M]_F$ so that it becomes easier to argue that $[M]_F DN$ is invertible.

▷ Claim 37. There is a row operation matrix $R \in \mathrm{GL}(d^m, \mathbb{F})$ with $det(R) = 1$ such that $R[M]_F$ has the following structure: The rows of $R[M]_F$ are indexed by $\mathbf{b} = (b_1,\ldots,b_m) \in F$ and its columns by $\mathbf{e} = (e_1,\ldots,e_m) \in \{0,\ldots,d\}^m$. Its entry indexed by $(\mathbf{b},\mathbf{e})$ is non-zero if and only if for all $i \in [m]$, $b_i = e_i$ if $e_i \neq 0$. All the non-zero entries of $R[M]_F$ are ± 1.

The above claim is proved in the full version [64]. Because of this claim, showing that $R[M]_F DN$ is invertible would suffice. Just like we did with M, we also impose the order $\prec$ on the columns of $R[M]_F$ that are indexed by $\mathbf{e} \in \{0,\ldots,d\}^m$. Recall that the rows of $R[M]_F$ and the columns of N are indexed by $\mathbf{b} \in F$. We order these indices as follows: we keep the indices $\mathbf{b} \in F \setminus \mathcal{B}$ before $\mathbf{b}_1,\ldots,\mathbf{b}_p$. We will treat $\mathbf{r}^{-\mathbf{e}}$ as a monomial in $(-r_1)^{-1},\ldots,(-r_m)^{-1}$ "variables" and impose the order $\prec$ on the monomials in these variables. Let $A := \{\mathbf{b} : \mathbf{b} \in F \setminus \mathcal{B}\} \cup \{\mathbf{b}'_1,\ldots,\mathbf{b}'_p\}$; notice that $|A| = |F|$. Also, the elements of A are ordered as the elements of F but with $\mathbf{b}'_k$ replacing $\mathbf{b}_k$ for $k \in [p]$. Then, from the Cauchy-Binet formula and the construction of the matrix N, $\det(R[M]_F DN)$ equals

$$\det\left([R[M]_F]_{\bullet,A}\right)[N]_A \cdot \prod_{\mathbf{e}\in A} \mathbf{r}^{-\mathbf{e}} + \text{ lower order monomials in the } (-r_1)^{-1},\ldots,(-r_m)^{-1}.$$

Here $[R[M]_F]_{\bullet,A}$ denotes the restriction of $R[M]_F$ to the columns indexed by $\mathbf{e} \in A$, and $[N]_A$ denotes the restriction of N to the rows indexed by $\mathbf{e} \in A$. Thus to show that $R[M]_F DN$ (and therefore $[CMDN]_F$) is invertible, the following two claims, both of which are proved in the full version [64], suffice.

▷ Claim 38. $[N]_A$ is an identity matrix.

▷ Claim 39. The matrix $[R[M]_F]_{\bullet,A}$ is an upper triangular matrix with 1 or -1 entries on the diagonal.

A.2 Proof of Theorem 21

Let $f = \mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n) \cdot \mathbf{1}$ be a width-w commutative ROABP having individual degree at most d; here $M_i \in \mathbb{F}^{w\times w}[x_i]$ for all $i \in [n]$. Also, let $F = M_1(x_1)M_2(x_2)\cdots M_n(x_n)$. For any $A \in \mathrm{GL}(n,\mathbb{F})$, let $g = f(A\mathbf{x})$ and $G = F(A\mathbf{x})$. Suppose that A maps $x_i \mapsto \ell_i(\mathbf{x})$ and let $y_i = \ell_i(\mathbf{x})$ for all $i \in [n]$. Then, $g = \mathbf{1}^T \cdot M_1(y_1)M_2(y_2)\cdots M_n(y_n)\cdot\mathbf{1}$ and

$G = M_1(y_1)M_2(y_2)\cdots M_n(y_n)$. In Sections 3.1 and 3.2, we have shown that $G\left(\mathbf{x}+\mathcal{G}_m^{SV}\right)$ has support-$(m-1)$ rank concentration (for $m = 2\left\lceil \log w^2\right\rceil + 1$) over $\mathbb{F}(\mathbf{z})$ in the $\mathbf{y}$-variables; the $\mathbf{z}$-variables are the variables introduced by the $\mathcal{G}_m^{SV}$ generator. From Observation 15, if $g(\mathbf{x}) \neq 0$, then $g\left(\mathbf{x}+\mathcal{G}_m^{SV}\right)$, when viewed as a polynomial over $\mathbb{F}[\mathbf{z}]$ in the $\mathbf{y}$-variables (this we can do as $g\left(\mathbf{x}+\mathcal{G}_m^{SV}\right) = \mathbf{1}^T \cdot G\left(\mathbf{x}+\mathcal{G}_m^{SV}\right)\cdot \mathbf{1}$, and $G\left(\mathbf{x}+\mathcal{G}_m^{SV}\right)$ can be viewed as a polynomial over $\mathbb{A}[\mathbf{z}]$ in the $\mathbf{y}$-variables), has a $\mathbf{y}$-monomial of support at most $m-1$. Let the $\mathbf{y}$-degree of this monomial be D'. As the individual degree of every $\mathbf{x}$-variable in f is at most d, the individual degree of every $\mathbf{y}$-variable in g is also at most d. Thus, $D' \leq (m-1)d$. As the homogeneous component of $g\left(\mathbf{x}+\mathcal{G}_m^{SV}\right)$ of $\mathbf{y}$-degree D' is non-zero, the homogeneous component of $g\left(\mathbf{x}+\mathcal{G}_m^{SV}\right)$ (now viewed as polynomial over $\mathbb{F}[\mathbf{z}]$ in the $\mathbf{x}$-variables) of $\mathbf{x}$-degree D' must also be non-zero, since $\ell_1,\ldots,\ell_n$ are linearly independent. This means that $g(\mathbf{x}+\mathcal{G}_m^{SV})$, when viewed as a polynomial over $\mathbb{F}[\mathbf{z}]$ in the $\mathbf{x}$-variables, has an $\mathbf{x}$-monomial of support (in fact, degree) at most $D' \leq (m-1)d$. Thus, $g\left(\mathcal{G}_{(m-1)d}^{SV}+\mathcal{G}_m^{SV}\right) \neq 0$. Now, it follows directly from the definition of the SV generator that $\mathcal{G}_{(m-1)d}^{SV}+\mathcal{G}_m^{SV} = \mathcal{G}_{m+(m-1)d}^{SV}$ and so $g\left(\mathcal{G}_{m+(m-1)d}^{SV}\right) \neq 0$. Replacing m by its value $2\left\lceil \log w^2\right\rceil + 1$ proves the theorem. Note that the SV generator needs $|\mathbb{F}| \geq n$.

A.3 Proof of Theorem 6

Let f be an n-variate polynomial computed by a width-w commutative ROABP of individual degree at most d, and $g \in \operatorname{orb}(f)$. Then, from Theorem 21, $g\left(\mathcal{G}^{SV}_{(2\lceil \log w^2\rceil (d+1)+1)}\right) \neq 0$ whenever $g \neq 0$. Now, $\mathcal{G}^{SV}_{(2\lceil \log w^2\rceil (d+1)+1)}$ has $2\left(2\left\lceil \log w^2\right\rceil (d+1)+1\right)$ variables, and is of degree n. So $g\left(\mathcal{G}^{SV}_{(2\lceil \log w^2\rceil (d+1)+1)}\right)$ also has $2\left(2\left\lceil \log w^2\right\rceil (d+1)+1\right)$ variables. Since the individual degree of f is at most d, the $\deg(f) = \deg(g) \leq nd$. So the degree of $g\left(\mathcal{G}^{SV}_{(2\lceil \log w^2\rceil (d+1)+1)}\right)$ is at most n^2d. Thus, as $|\mathbb{F}| > n^2d$, a hitting set for g can be computed in time $\left(n^2d+1\right)^{\left(2\lceil \log w^2\rceil (d+1)+1\right)} = (nd)^{O(d\log w)}$.

B Missing proofs from Section 4

B.1 Proof of Lemma 31

The entries of U, the columns of M, the rows and columns of D, and the rows of N are indexed by $\mathbf{e} \in \{0,1\}^m$. Impose the degree lexicographic order, denoted by $\prec_{\text{dlex}}$, on the indices $\mathbf{e} \in \{0,1\}^m$ of U and the other three matrices (by identifying $\mathbf{e}$ with an m-variate monomial) . Pick the *minimal* basis of the space spanned by the entries of U according to this order, i.e., consider the entries of U in the order dictated by $\prec_{\text{dlex}}$ while forming the basis. Let $\mathcal{B} := \{\mathbf{e} \in \{0,1\}^m : u_{\mathbf{e}} \text{ is in the minimal basis of } U \text{ w.r.t. } \prec_{\text{dlex}}\}$.

▶ **Observation 40.** *By the induction hypothesis, for every* $\mathbf{e} \in F \cap \mathcal{B}$, $\operatorname{Supp}(\mathbf{e}) = 2\mu - (q^* - 1)$.

Construction of the matrix N. The columns of N are indexed by $\mathbf{b} \in F$. We will now specify a set of column vectors $\{\mathbf{n_b} : \mathbf{b} \in F\}$ in the null space of U such that the column of N indexed by $\mathbf{b} \in F$ is $\mathbf{n_b}$. There are two cases for $\mathbf{b} \in F$:

Case 1: $\mathbf{b} \in F \setminus \mathcal{B}$. In this case, $u_{\mathbf{b}}$ is dependent on $\{u_{\mathbf{e}} : \mathbf{e} \in \mathcal{B} \text{ and } \mathbf{e} \prec_{\text{dlex}} \mathbf{b}\}$. Pick this dependence vector as $\mathbf{n_b}$.

Case 2: $\mathbf{b} \in F \cap \mathcal{B}$. Let there be p such $\mathbf{b}, \mathbf{b}_1, \ldots, \mathbf{b}_p$, where $p \leq |\mathcal{B}| \leq w^2$. For a set $E \subseteq [m]$ and $\mathbf{b} \in \{0,1\}^m$, let $(\mathbf{b})_E$ denote the vector obtained by projecting $\mathbf{b}$ to the coordinates

in E. Roughly speaking, the following claim, which is proved in the full version [64], says that each of these p vectors has a "small signature" that differentiates it from the other $p-1$ vectors.

▷ Claim 41. There exist sets $E_1, \ldots, E_p \subseteq [m]$, each of size $w^2 - 1$ such that for all $k \in [p]$,
1. $\mathrm{Supp}\left((\mathbf{b}_k)_{E_k}\right) = w^2 - 1$,
2. $(\mathbf{b}_k)_{E_k} \neq (\mathbf{b}_\ell)_{E_k}$ $\forall \ell \neq k$.

As before, we will call E_k the signature of $\mathbf{b}_k$. The following claim tells us that for each vector $\mathbf{b}_k$, there is a vector that is not in $\mathcal{B}$ and has support less than $2\mu - (q^* - 1)$, but agrees with $\mathbf{b}_k$ on its signature and so in some sense can be used as a proxy for $\mathbf{b}_k$.

▷ Claim 42. For every $k \in [p]$, there exists a vector $\mathbf{b}'_k \in \{0,1\}^m \setminus (F \cup \mathcal{B})$ such that $(\mathbf{b}'_k)_{E_k} = (\mathbf{b}_k)_{E_k}$ and also $\mathbf{b}'_k$ and $\mathbf{b}_k$ agree on all locations where $\mathbf{b}'_k$ is non-zero.

Proof. Similar to the proof of Claim 36. ◁

We will now use the above two claims to construct $\mathbf{n}_{\mathbf{b}_k}$ for all $k \in [p]$. We will use $\mathbf{b}'_k$ from Claim 42 as a proxy for $\mathbf{b}_k$. Notice that $u_{\mathbf{b}'_k}$ is dependent on $\{u_{\mathbf{e}} : \mathbf{e} \in \mathcal{B} \text{ and } \mathbf{e} \prec_{\mathrm{dlex}} \mathbf{b}'_k\}$. Let this dependence vector be $\mathbf{n}_{\mathbf{b}_k}$. This completes the construction of N. We will now show that $[CMDN]_F$ is invertible. In fact, we will show that $\det([CMDN]_F)$ is the ratio of a polynomial in $\mathbb{F}[\mathbf{t}]$ which contains a monomial of degree at most $2w^2\mu$ and a product of a bunch of non-zero linear forms in $\mathbb{F}[\mathbf{t}]$.

$[CMDN]_F$ is invertible. Let $[M]_F$ be the restriction of M to the rows indexed by F, and $[C]_F$ the restriction of C to the rows and columns indexed by F.

▶ **Observation 43.** *The matrix $[M]_F$ has the following structure: The rows of $[M]_F$ are indexed by $\mathbf{b} = (b_1, \ldots, b_m) \in F$ and its columns by $\mathbf{e} = (e_1, \ldots, e_m) \in \{0,1\}^m$. Its entry indexed by $(\mathbf{b}, \mathbf{e})$ is non-zero if and only if for all $i \in [m]$, $b_i = e_i$ if $e_i \neq 0$. All non-zero entries are 1.*

We order the indices $\mathbf{b} \in F$ as follows: Let $F_0 := \{\mathbf{b} \in F : \mathrm{Supp}(\mathbf{b}) > 2\mu - (q^* - 1)\}$ and $F_1 := \{\mathbf{b} \in F : \mathrm{Supp}(\mathbf{b}) = 2\mu - (q^* - 1)\}$. We first keep the $\mathbf{b} \in F_0$ in (descending) degree lexicographic order[3], followed by $\mathbf{b} \in F_1 \setminus \mathcal{B}$ in (reverse) lexicographic order[4], and then $\mathbf{b}_1, \ldots, \mathbf{b}_p$. Also, let $A := (F \setminus \mathcal{B}) \uplus \{\mathbf{b}'_1, \ldots, \mathbf{b}'_p\}$. Notice that $|A| = |F|$. Also, the elements of A are ordered as the elements of F but with $\mathbf{b}'_k$ replacing $\mathbf{b}_k$ for $k \in [p]$. For any $S \subseteq \{0,1\}^m$ of size $|S| = |F|$, let $[M]_{F,S}$ denote the restriction of $[M]_F$ to the columns indexed by $\mathbf{e} \in S$, and $[N]_S$ denote the restriction of N to the rows indexed by $\mathbf{e} \in S$. Now,

$$\begin{aligned}
\det([CMDN]_F) &= \det([C]_F)\det([M]_F DN) \\
&= \prod_{\mathbf{b}\in F} \mathbf{r}^{\mathbf{b}} \cdot \left(\sum_{\substack{S \subseteq A \uplus \mathcal{B} \\ |S| = |F|}} \det\left([M]_{F,S}\right) \cdot \det([N]_S) \cdot \prod_{\mathbf{e}\in S} \mathbf{r}^{-\mathbf{e}} \right) \\
&= \prod_{\mathbf{b}\in F} \mathbf{r}^{\mathbf{b}} \cdot \left(\sum_{\substack{S \subseteq A \uplus \mathcal{B} \\ |S| = |F|}} \det\left([M]_{F,S}\right) \cdot \det([N]_S) \cdot \prod_{\mathbf{e}\in S\cap A} \mathbf{r}^{-\mathbf{e}} \cdot \prod_{\mathbf{e}\in S\cap \mathcal{B}} \mathbf{r}^{-\mathbf{e}} \right)
\end{aligned}$$

[3] i.e., $\mathbf{b}$ comes before $\hat{\mathbf{b}}$ if $\mathrm{Supp}(\mathbf{b}) > \mathrm{Supp}(\hat{\mathbf{b}})$, or if $\mathrm{Supp}(\mathbf{b}) = \mathrm{Supp}(\hat{\mathbf{b}})$ and $\hat{\mathbf{b}} \prec_{lex} \mathbf{b}$.
[4] i.e., $\mathbf{b}$ comes before $\hat{\mathbf{b}}$ if $\hat{\mathbf{b}} \prec_{lex} \mathbf{b}$.

$$= \prod_{\mathbf{b}\in F} \mathbf{r}^{\mathbf{b}} \cdot \prod_{\mathbf{e}\in A\uplus\mathcal{B}} \mathbf{r}^{-\mathbf{e}} \cdot \left(\sum_{\substack{S\subseteq A\uplus\mathcal{B} \\ |S|=|F|}} \det\left([M]_{F,S}\right) \cdot \det([N]_S) \cdot \prod_{\mathbf{e}\in A\setminus S} \mathbf{r}^{\mathbf{e}} \cdot \prod_{\mathbf{e}\in\mathcal{B}\setminus S} \mathbf{r}^{\mathbf{e}} \right),$$

where the second equality follows from the Cauchy-Binet formula and the third equality from the fact that for any $S \not\subseteq A\uplus\mathcal{B}$, $\det([N]_S) = 0$. Now, notice that $\prod_{\mathbf{b}\in F} \mathbf{r}^{\mathbf{b}} \cdot \prod_{\mathbf{e}\in A\uplus\mathcal{B}} \mathbf{r}^{-\mathbf{e}}$ is the reciprocal of a product of non-zero linear forms in $\mathbf{t}$-variables, as $F \subseteq A \uplus \mathcal{B}$. We shall now prove that

$$\sum_{\substack{S\subseteq A\uplus\mathcal{B} \\ |S|=|F|}} \det\left([M]_{F,S}\right) \cdot \det([N]_S) \cdot \prod_{\mathbf{e}\in A\setminus S} \mathbf{r}^{\mathbf{e}} \cdot \prod_{\mathbf{e}\in\mathcal{B}\setminus S} \mathbf{r}^{\mathbf{e}} \tag{3}$$

has a $\mathbf{t}$-monomial of degree at most $w^2(2\mu - (q^* - 1))$.

▷ Claim 44. $[N]_A$ is an identity matrix.

Proof. Same as that of Claim 38. ◁

▷ Claim 45. The matrix $[M]_{F,A}$ is an upper triangular matrix with ones on the diagonal.

The proof of the above claim is provided in the full version [64].

▷ Claim 46. $\det\left([M]_{F,A}\right) \cdot \det([N]_A) \cdot \prod_{\mathbf{e}\in\mathcal{B}\setminus A} \mathbf{r}^{\mathbf{e}} = \prod_{\mathbf{e}\in\mathcal{B}} \mathbf{r}^{\mathbf{e}} \neq 0$ and has $\mathbf{t}$-degree at most $2w^2\mu$.

Proof. $\det\left([M]_{F,A}\right) \cdot \det([N]_A) \cdot \prod_{\mathbf{e}\in\mathcal{B}\setminus A} \mathbf{r}^{\mathbf{e}} = \prod_{\mathbf{e}\in\mathcal{B}} \mathbf{r}^{\mathbf{e}} \neq 0$ follows from Claims 44 and 45 and the fact that $A \cap \mathcal{B}$ is empty. For every $\mathbf{e} \in \mathcal{B}$, $\deg_{\mathbf{t}}(\mathbf{r}^{\mathbf{e}}) \leq 2\mu - (q^* - 1)$. So, $\deg_{\mathbf{t}}\left(\prod_{\mathbf{e}\in\mathcal{B}} \mathbf{r}^{\mathbf{e}}\right) \leq w^2 \cdot (2\mu - (q^* - 1)) \leq 2w^2\mu$, as $|\mathcal{B}| \leq w^2$. ◁

▷ Claim 47. For any $S \subseteq A \uplus \mathcal{B}$ such that $|S| = |F|$ and $\det([N]_S)$ is non-zero, there is a one-to-one correspondence between $A \setminus S$ and $S \cap \mathcal{B}$ such that if $\mathbf{e} \in A \setminus S$ corresponds to $\mathbf{e}' \in S \cap \mathcal{B}$, then $\mathbf{e}' \prec_{\text{dlex}} \mathbf{e}$.

The above claim, which is proved in the full version [64], implies that for every $S \in A\uplus\mathcal{B}$ of size $|F|$, either $\det\left([M]_{F,S}\right) \cdot \det([N]_S) \cdot \prod_{\mathbf{e}\in A\setminus S} \mathbf{r}^{\mathbf{e}} \cdot \prod_{\mathbf{e}\in\mathcal{B}\setminus S} \mathbf{r}^{\mathbf{e}}$ is 0, or $\prod_{\mathbf{e}\in\mathcal{B}} \mathbf{r}^{\mathbf{e}} \prec_{\text{dlex}} \prod_{\mathbf{e}\in A\setminus S} \mathbf{r}^{\mathbf{e}} \cdot \prod_{\mathbf{e}\in\mathcal{B}\setminus S} \mathbf{r}^{\mathbf{e}}$. Hence, $\prod_{\mathbf{e}\in\mathcal{B}} \mathbf{r}^{\mathbf{e}}$ is the smallest $\mathbf{r}$-monomial in the polynomial given in (3) w.r.t. $\prec_{\text{dlex}}$ order, and so, the homogeneous component of this polynomial that has the same $\mathbf{r}$-degree as that of $\prod_{\mathbf{e}\in\mathcal{B}} \mathbf{r}^{\mathbf{e}}$ survives. Now, from Claim 46 and the fact that $\ell_1, \ldots, \ell_n$ are linearly independent, the polynomial in (3) has a $\mathbf{t}$-monomial of degree $\leq 2w^2\mu$.

B.2 Proof of Lemma 28

So far we have proved that there exist $\{\beta_{p,q}(i) : p \in [\lceil \log n \rceil], q \in [\mu], i \in [n]]\}$, such that $G\left(x_1 + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} s_{p,q} \cdot z_{p,q}^{\beta_{p,q}(1)}, \ldots, x_n + \sum_{p\in\lceil \log n\rceil, q\in[\mu]} s_{p,q} \cdot z_{p,q}^{\beta_{p,q}(n)}\right)$ has support-μ rank concentration in the $\mathbf{y}$-variables over $\mathbb{F}\left(s_{p,q}, z_{p,q} : p \in [\lceil \log n \rceil], q \in [\mu]\right)$. Moreover, for each (p, q), we can find all $\beta_{p,q}(i)$ in time $n^{O(w^4)}$ and each $\beta_{p,q}(i) \leq n^{O(w^4)}$. However, since the algorithm that follows from [45] is oblivious, the $\beta_{p,q}(i)$ found for some fixed (p, q) can be used for all values of (p, q). This proves the lemma.

B.3 Proof of Theorem 8

Let $f = \mathbf{1}^T \cdot M_1(x_1)M_2(x_2)\cdots M_n(x_n)\cdot \mathbf{1}$ be a multilinear width-w ROABP; here $M_i(x_i) \in \mathbb{F}^{w\times w}[x_i]$ for all $i \in [n]$. Also, let $F = M_1(x_1)M_2(x_2)\cdots M_n(x_n)$. For any $A \in \mathrm{GL}(n,\mathbb{F})$, let $g = f(A\mathbf{x})$ and $G = F(A\mathbf{x})$. For $i \in [n]$, suppose that A maps $x_i \mapsto \ell_i(\mathbf{x})$, where ℓ_i is a linear form, and let $y_i = \ell_i(\mathbf{x})$ and $\mathbf{y} = \{y_1,\ldots,y_n\}$. Then, $g = \mathbf{1}^T \cdot M_1(y_1)M_2(y_2)\cdots M_n(y_n)\cdot \mathbf{1}$ and $G = M_1(y_1)M_2(y_2)\cdots M_n(y_n)$. Let $\mu = w^2 + \lceil \log w^2 \rceil$. From Lemma 28, there exist polynomials, say $t_1,\ldots,t_n$, in $\mathbb{F}\left[s_{p,q}, z_{p,q} : p \in [\lceil \log n \rceil], q \in [\mu]\right]$ of degree at most $n^{O(w^4)}$ such that $G(x_1+t_1,\ldots,x_n+t_n)$ has support-μ rank concentration in the $\mathbf{y}$-variables over $\mathbb{F}\left(\{s_{p,q}, z_{p,q}\}_{p,q}\right)$. Moreover, these polynomials can be computed in time $n^{O(w^4)}$. Suppose that $g \neq 0$. Then, from Observation 15, $g(x_1+t_1,\ldots,x_n+t_n)$ has a support-μ, $\mathbf{y}$-monomial when viewed as a polynomial over $\mathbb{F}\left[\{s_{p,q}, z_{p,q}\}_{p,q}\right]$ in the $\mathbf{y}$-variables. Since f is multilinear, as seen in the proof of Theorem 21, $g(x_1+t_1,\ldots,x_n+t_n)$ has a support-μ, $\mathbf{x}$-monomial. Thus, $g\left(\mathcal{G}_\mu^{SV} + (t_1,\ldots,t_n)\right) \neq 0$. Now, $g\left(\mathcal{G}_\mu^{SV} + (t_1,\ldots,t_n)\right)$ is a polynomial in $2\mu + \mu\cdot\lceil \log n\rceil$ variables over $\mathbb{F}$. Also, its degree is at most $n^{O(w^4)}$. So, if $|\mathbb{F}| > n^{O(w^4)}$, a hitting set for g can be computed in time $n^{O(w^4\cdot\mu\cdot\log n)} = n^{O\left(w^6\cdot\log n\right)}$. This, along with the time required to compute $t_1,\ldots,t_n$, still gives a $n^{O\left(w^6\cdot\log n\right)}$-time hitting set for g.