Electronic Colloquium on Computational Complexity, Report No. 115 (2022)

Polynomial Identity Testing via Evaluation of Rational Functions

Dieter van Melkebeek* Andrew Morgan*

August 17, 2022

Abstract

We introduce a hitting set generator for Polynomial Identity Testing based on evaluations of low-degree univariate rational functions at abscissas associated with the variables. Despite the univariate nature, we establish an equivalence up to rescaling with a generator introduced by Shpilka and Volkovich, which has a similar structure but uses multivariate polynomials in the abscissas.

We study the power of the generator by characterizing its vanishing ideal, i.e., the set of polynomials that it fails to hit. Capitalizing on the univariate nature, we develop a small collection of polynomials that jointly produce the vanishing ideal. As corollaries, we obtain tight bounds on the minimum degree, sparseness, and partition class size of set-multilinearity in the vanishing ideal. Inspired by an alternating algebra representation, we develop a structured deterministic membership test for the vanishing ideal. As a proof of concept, we rederive known derandomization results based on the generator by Shpilka and Volkovich and present a new application for read-once oblivious algebraic branching programs.

1 Overview

Polynomial identity testing (PIT) is the fundamental problem of deciding whether a given multivariate algebraic circuit formally computes the zero polynomial. PIT has a simple efficient randomized algorithm that only needs blackbox access to the circuit: Pick a random point and check whether the circuit evaluates to zero on that particular point.

Despite the fundamental nature of PIT and the simplicity of the randomized algorithm, no efficient deterministic algorithm is known—even in the white-box setting, where the algorithm has access to the description of the circuit. The existence of such an algorithm would imply long-sought circuit lower bounds [HS80; Agr05; KI04]. Conversely, sufficiently strong circuit lower bounds yield blackbox derandomization for all of BPP, the class of decision problems admitting efficient randomized algorithms with bounded error [NW94; IW97]. Although the known results leave gaps between the two directions, they suggest that PIT acts as a BPP-complete problem in the context of derandomization, and that derandomization of BPP can be achieved in a blackbox fashion if at all.

Blackbox derandomization of PIT for a class of polynomials $\mathcal{C}$ in the variables $x_1, \ldots, x_n$ is equivalent to the efficient construction of a substitution G that replaces each x_i by a low-degree polynomial in a small set of fresh variables such that, for every nonzero polynomial p from $\mathcal{C}$, $p(G)$ remains nonzero [SY10, Lemma 4.1]. We refer to G as a generator, the fresh variables are its seed,

*University of Wisconsin–Madison, Madison, WI, USA. {dieter,amorgan}@cs.wisc.edu. A preliminary version of this paper appeared in [vMM22]. Research partially supported by the US National Science Foundation through grant CCF-1838434.

ISSN 1433-8092

and say that G hits the class $\mathcal{C}$. If there are l seed variables, and if p and G have degree at most $n^{O(1)}$, then the resulting deterministic PIT algorithm for $\mathcal{C}$ makes $n^{O(l)}$ blackbox queries.

Much progress on derandomizing PIT has been obtained by designing such substitutions and analyzing their hitting properties for interesting classes $\mathcal{C}$. Shpilka and Volkovich [SV15] introduced a generator, by now dubbed the Shpilka–Volkovich generator or "SV generator" for short, and proved that it hits sums of a bounded number of read-once formulas for $l = O(\log n)$, later improved to $l = O(1)$ [MV18]. The generator for $l = O(\log n)$ has also been shown to hit multilinear depth-4 circuits with bounded top fan-in [KMS+13], multilinear bounded-read formulas [AvMV15], commutative read-once oblivious algebraic branching programs [FSS14], $\Sigma\mathrm{m}\bigwedge\Sigma\Pi^{O(1)}$ formulas [For15], circuits with locally-low algebraic rank in the sense of [KS17], and orbits of simple polynomial classes under invertible linear transformations of the variables [MS21]. The generator is an ingredient in other hitting set constructions, as well, notably constructions using the technique of low-support rank concentration [ASS13; AGK+15; GKS+17; GKS17; ST21; BG21]. It also forms the core of a "succinct" generator that hits a variety of classes including depth-2 circuits [FSV17].

Vanishing ideal. In this paper, we initiate a systematic study of the power of a generator G through the set of polynomials p such that $p(G)$ vanishes. The set has the algebraic structure of an ideal and is known as the *vanishing ideal* of G; we denote it by $\mathrm{Van}[G]$. Our technical contributions can be understood as precisely characterizing the vanishing ideal of the SV generator.

Characterizing the vanishing ideal facilitates two objectives:

Derandomization. A generator G hits a class $\mathcal{C}$ of polynomials if and only if $\mathcal{C}$ and $\mathrm{Van}[G]$ have at most the zero polynomial in common. If the characterization of $\mathrm{Van}[G]$ is incompatible with being computable within some resource bound, then G trivially hits the class $\mathcal{C}$ of polynomials that are computable within the bound. In other words, derandomization of PIT for $\mathcal{C}$ reduces to proving lower bounds for $\mathrm{Van}[G]$. Characterizing $\mathrm{Van}[G]$ yields explicit structure that makes the lower bounds more tractable.

More generally, given a characterization of $\mathrm{Van}[G]$, in order to derandomize PIT for a class $\mathcal{C}$ it suffices to design another generator G' that hits merely the polynomials in $\mathcal{C} \cap \mathrm{Van}[G]$. As G hits the remainder of $\mathcal{C}$, combining G with G' yields a generator for all of $\mathcal{C}$. In this way, one may assume for free additional structure about the polynomials in $\mathcal{C}$, namely that the polynomials moreover belong to $\mathrm{Van}[G]$. Characterizing the vanishing ideal of G makes this additional structure explicit.

Lower bounds. If we happen to know that G hits the class $\mathcal{C}$ of polynomials computable within some resource bound, then any expression for a nonzero polynomial in $\mathrm{Van}[G]$ yields an explicit polynomial that falls outside $\mathcal{C}$. Such a statement is often referred to as hardness of representation, and can be viewed as a lower bound in the model of computation underlying $\mathcal{C}$ (provided the polynomial can be computed in the model at all). Characterizing $\mathrm{Van}[G]$ makes explicit the polynomials to which the lower bound applies.

We will illustrate how to make progress on both objectives through our characterizations of the SV generator's vanishing ideal.

Rational function evaluations. Another contribution of our paper is the development of an alternate view of the SV generator, namely as evaluations of univariate rational functions of low degree. We would like to promote the perspective for its intrinsic appeal and its applicability. Among other benefits, it facilitates the study of the vanishing ideal.

The transition goes as follows. The SV generator takes as additional parameters a positive integer l and a choice of distinct field elements a_i for each of the original variables x_i, $i \in [n]$.

We refer to the elements a_i as *abscissas* and denote the generator for a given value of l by SV^l (suppressing the choice of abscissas). When $l = 1$, SV^1 takes as seed two fresh variables, y and z, and can be described succinctly in terms of the Lagrange interpolants L_i for the set of abscissas:

$$x_i \leftarrow z \cdot L_i(y) \doteq z \cdot \prod_{j \in [n] \setminus \{i\}} \frac{y - a_j}{a_i - a_j}. \tag{1}$$

By rescaling, the denominators on the right-hand side of (1) can be cleared, resulting in the following somewhat simpler substitution:

$$x_i \leftarrow z \cdot \prod_{j \in [n] \setminus \{i\}} (y - a_j). \tag{2}$$

The vanishing ideals of (2) and SV^1 are the same up to rescaling the variables to match the rescaling from (1) to (2).

More importantly, we apply the change of variables $z \leftarrow z' / \prod_{j \in [n]} (y - a_j)$, resulting in a substitution that now uses rational functions of the seed:

$$x_i \leftarrow \frac{z'}{y - a_i}. \tag{3}$$

The notion of vanishing ideal naturally extends to rational function substitutions. The change of variables from (2) to (3) establishes that any polynomial vanishing on (2) also vanishes on (3). The change of variables is invertible (the inverse is $z' \leftarrow z \prod_{j \in [n]} (y - a_j)$), so any polynomial vanishing on (3) also vanishes on (2). We conclude that the vanishing ideal of (3) is the same as that of SV^1 up to rescaling the variables.

Note that, for fixed y and z', (3) may be interpreted as first forming a univariate rational function $f(\alpha) = \frac{z'}{y-\alpha}$ (depending on y and z' but independent of i) and then substituting $x_i \leftarrow f(a_i)$. As y and z' vary, f ranges over all rational functions with numerator degree zero and denominator degree one. We denote (3) by RFE_1^0, where RFE is a short-hand for *Rational Function Evaluation*, 0 bounds the numerator degree, and 1 bounds the denominator degree.

As a generator, RFE_1^0 naturally generalizes to RFE_l^k for arbitrary $k, l \in \mathbb{N}$.

Definition 1 (RFE Generator). *Let $\mathbb{F}$ be a field and $\{x_1, \ldots, x_n\}$ a set of variables. The* Rational Function Evaluation Generator (RFE) *for $\mathbb{F}[x_1, \ldots, x_n]$ is parametrized by the following data:*

- *For each $i \in [n]$, a distinct* abscissa $a_i \in \mathbb{F}$.
- *A non-negative integer k, the* numerator degree*.*
- *A non-negative integer l, the* denominator degree*.*

The generator takes as seed a rational function $f \in \mathbb{F}(\alpha)$ such that f can be written as g/h for some $g, h \in \mathbb{F}[\alpha]$ with $\deg(g) \le k$, $\deg(h) \le l$, and $h(a_i) \ne 0$ for all $i \in [n]$. From f, it generates the substitution $x_i \leftarrow f(a_i)$ for each $i \in [n]$.

There are multiple ways to parametrize the seed of RFE_l^k using scalars, such as by specifying the coefficients, evaluations, or roots for each of the numerator and denominator. The flexibility to choose is a source of convenience. As is customary in the context of blackbox derandomization of PIT, we assume that $\mathbb{F}$ is sufficiently large, possibly by taking a field extension. We refer to appendix A for a discussion on different parametrizations as well as on how to obtain deterministic blackbox PIT algorithms from the generator and how large the underlying field $\mathbb{F}$ must be.

The connection between RFE_1^0 and SV^1 extends as follows. For higher values of l, SV^l is defined as the sum of l independent instantiations of SV^1. The same transformations as above relate SV^l and the sum of l independent instantiations of RFE_1^0. The latter in turn is equivalent to RFE_l^{l-1} by partial fraction decomposition. The conclusion is that SV^l is equivalent in power to RFE_l^{l-1}, up to variable rescaling. We refer to appendix B for a formal treatment.

For parameter values $k \neq l-1$, there is no SV generator that corresponds to RFE_l^k, but $\mathrm{SV}^{\max(k+1,l)}$ encompasses RFE_l^k (up to rescaling) and uses at most twice as many seed variables. Thus, the RFE-generator and the SV-generator efficiently hit the same classes of polynomials. However, RFE's simpler univariate dependence on the abscissas—as opposed to SV's multi-variate dependence—enables our approach for determining the vanishing ideal. The moral is that, even though polynomial substitutions are sufficient for derandomizing PIT, it nevertheless helps to consider rational substitutions. Their use may simplify analysis and arguably yield more elegant constructions.

Generating set. Our first result describes a small and explicit generating set for the vanishing ideal of RFE. It consists of instantiations of a single determinant expression.

Theorem 2 (generating set). *Let $k, l \in \mathbb{N}$, $\{x_i : i \in [n]\}$ be a set of variables, and let a_i for $i \in [n]$ be distinct field elements. The vanishing ideal of RFE_l^k over the given set of variables for the given choice of abscissas $(a_i)_{i\in[n]}$ is generated by the following polynomials over all choices of $k+l+2$ variable indices $i_1, i_2, \ldots, i_{k+l+2} \in [n]$:*

$$\mathrm{EVC}_l^k[i_1, i_2, \ldots, i_{k+l+2}] \doteq \det\begin{bmatrix} a_{i_j}^l x_{i_j} & a_{i_j}^{l-1} x_{i_j} & \ldots & x_{i_j} & a_{i_j}^k & a_{i_j}^{k-1} & \ldots & 1 \end{bmatrix}_{j=1}^{k+l+2}. \tag{4}$$

Moreover, for any fixed set $C \subseteq [n]$ of $k+1$ variable indices, when $i_1, i_2, \ldots, i_{k+l+2}$ range over all the choices of $k+l+2$ indices from $[n]$ such that $i_1 < \cdots < i_{k+l+2}$ and $C \subseteq \{i_1, \ldots, i_{k+l+2}\}$, the polynomials $\mathrm{EVC}_l^k[i_1, i_2, \ldots, i_{k+l+2}]$ form a generating set of minimum size.

The name "EVC" is a shorthand for "Elementary Vandermonde Circulation". Later we discuss a representation of polynomials using alternating algebra, with connections to notions from network flow. In this representation, polynomials in the vanishing ideal coincide with circulations, and instantiations of EVC are the elementary circulations.

We refer to the set C in Theorem 2 as a *core*. The core C plays a similar role as in a combinatorial sunflower except that, unlike the petals of a sunflower, the sets $\{i_1, \ldots, i_{k+l+2}\}$ do not need to be disjoint outside the core.

Example 3. Consider the special case where $k = 0$ and $l = 1$. The generator for RFE_1^0 when $i_1 = 1$, $i_2 = 2$, and $i_3 = 3$ is given by

$$\mathrm{EVC}_1^0[1,2,3] \doteq \begin{vmatrix} a_1 x_1 & x_1 & 1 \\ a_2 x_2 & x_2 & 1 \\ a_3 x_3 & x_3 & 1 \end{vmatrix} = (a_1 - a_2)x_1x_2 + (a_2 - a_3)x_2x_3 + (a_3 - a_1)x_3x_1.$$

For any fixed $i^* \in [n]$, the polynomials $\mathrm{EVC}_1^0[i_1, i_2, i_3]$ form a generating set of minimum size when $\{i_1, i_2, i_3\}$ ranges over all subsets of $[n]$ that contain $C = \{i^*\}$, and $i_1 < i_2 < i_3$. ◂

In general, the generators EVC_l^k are nonzero multilinear homogeneous polynomials of degree $l+1$ containing all multilinear monomials of degree $l+1$.

Each generating set of minimum size in Theorem 2 yields a Gröbner basis with respect to every monomial order that prioritizes the variables outside C. A Gröbner basis is a special basis that allows solving ideal-membership queries more efficiently as well as solving systems of polynomial equations [CLO13]. Computing Gröbner bases for general ideals is exponential-space complete [KM96; May97]. Theorem 2 represents a rare instance of a natural and interesting ideal for which we know a small and moreover explicit Gröbner basis.

To gain some intuition about dependencies between the generators EVC_l^k, note that permuting the order of the variables used in the construction of EVC_l^k yields the same polynomial or minus that polynomial, depending on the sign of the permutation. This follows from the determinant structure of EVC_l^k and is the reason why we need to fix the order of the variables in order to obtain a generating set of minimum size. More profoundly, the following relationship holds for every choice of $k+l+3$ indices $i_1, i_2, \ldots, i_{k+l+3} \in [n]$ and every univariate polynomial w of degree at most k:

$$\det\begin{bmatrix} w(a_{i_j}) & a_{i_j}^l x_{i_j} & a_{i_j}^{l-1} x_{i_j} & \ldots & x_{i_j} & a_{i_j}^k & a_{i_j}^{k-1} & \ldots & 1 \end{bmatrix}_{j=1}^{k+l+3} = 0. \tag{5}$$

The determinant in (5) vanishes because the first column of the matrix is a linear combination of the last $k+1$. A Laplace expansion across the first column allows us to write the determinant of the matrix as a linear combination of minors, and each minor is an instantiation of EVC_l^k. Since the determinant vanishes, (5) represents a linear dependency for every nonzero polynomial w of degree at most k. In fact, when $\{i_1, \ldots, i_{k+l+3}\}$ varies over subsets of $[n]$ containing a fixed core of size $k+1$, the equations (5) generate *all* linear dependencies among instances of EVC_l^k.

As corollaries to Theorem 2 we obtain the following tight bounds on $\mathrm{Van}[\mathrm{RFE}_l^k]$. The bounds hold for every way to choose the parameters in Definition 1.

- The minimum *degree* of a nonzero polynomial in $\mathrm{Van}[\mathrm{RFE}_l^k]$ equals $l+1$. This proves a conjecture by Fournier and Korwar [FK18] (additional partial results reported in [Kor21]) that there exists a polynomial of degree $l+1$ in $n = 2l+1$ variables that SV^l fails to hit. The conjecture follows because the generators for $\mathrm{Van}[\mathrm{SV}^l]$ have degree $l+1$ and use $2l+1$ variables.

  As none of the generators contain a monomial of support l or less, the same holds for every nonzero polynomial in $\mathrm{Van}[\mathrm{RFE}_l^k]$. This extends the known property that SV^l hits every polynomial that contains a monomial of support l or less [SV15].

- The minimum *sparseness*, i.e., number of monomials, of a nonzero polynomial in $\mathrm{Van}[\mathrm{RFE}_l^k]$ equals $\binom{k+l+2}{l+1}$. The generators EVC_l^k realize the bound as they exactly contain all multilinear monomials of degree $l+1$ that can be formed out of their $k+l+2$ variables.

  The claim that no nonzero polynomial in $\mathrm{Van}[\mathrm{RFE}_l^k]$ contains fewer than $\binom{k+l+2}{l+1}$ monomials requires an additional combinatorial argument. It is a (tight) quantitative strengthening of the known property that SV^l hits every polynomial with fewer than 2^l monomials [AvMV15; GKS+17; For15; FSV17]. Note that for $k = l-1$ we have that $\binom{k+l+2}{l+1} = \binom{2l+1}{l+1} = \Theta(2^{2l}/\sqrt{l})$.

- The minimum *partition class size* of a nonzero set-multilinear polynomial of degree $l+1$ in $\mathrm{Van}[\mathrm{RFE}_l^k]$ equals $k+2$. Set-multilinearity is a common restriction in works on derandomizing PIT and algebraic circuit lower bounds. A polynomial p of degree $l+1$ in a set of variables $\{x_1, \ldots, x_n\}$ is said to be set-multilinear if $[n]$ can be partitioned as $[n] = X_1 \sqcup X_2 \sqcup \cdots \sqcup X_{l+1}$ such that every monomial in p is a product $x_{i_1} \cdot x_{i_2} \cdot \cdots \cdot x_{i_{l+1}}$, where $i_j \in X_j$. Note that set-multilinearity implies multilinearity but not the other way around.

  As the generators EVC_l^k are not set-multilinear, it is not immediately clear from Theorem 2 whether $\mathrm{Van}[\mathrm{RFE}_l^k]$ contains nontrivial set-multilinear polynomials. However, a variation on the construction of the generators EVC_l^k yields explicit set-multilinear homogeneous polynomials in $\mathrm{Van}[\mathrm{RFE}_l^k]$ of degree $l+1$ where each X_j has size $k+2$. We denote them by ESMVC_l^k, where ESMVC stands for "Elementary Set-Multilinear Vandermonde Circulation". ESMVC_l^k contains all monomials of the form $x_{i_1} \cdot x_{i_2} \cdot \dots \cdot x_{i_{l+1}}$ with $i_j \in X_j$. For any variable partition $X_1 \sqcup X_2 \sqcup \dots \sqcup X_{l+1}$ with $|X_1| = \dots = |X_{l+1}| = k+2$, ESMVC_l^k is the only set-multilinear polynomial in $\mathrm{Van}[\mathrm{RFE}_l^k]$ with that variable partition, up to a scalar multiple.

Membership test. Our second characterization of the vanishing ideal of RFE can be viewed as a structured membership test. Given a polynomial p, there is a generic way to test whether p belongs to the vanishing ideal of any generator G, namely by symbolically substituting G into p and verifying that the result simplifies to zero. When G is a polynomial substitution, the well-known transformation of a generator into a deterministic blackbox PIT algorithm yields another test: verify $p(G) = 0$ for a sufficiently large set of substitutions into the seed variables [Ore22; DL78; Zip79; Sch80]. By clearing denominators, the same goes for rational substitutions like RFE_l^k. While these tests work, their genericity with respect to G implies that they cannot provide any G-specific insight into whether or why a given p belongs to the vanishing ideal of G. As we will argue, our structured membership test does.

Several prior papers demonstrated the utility of partial derivatives and zero substitutions in the context of derandomizing PIT using the SV generator, especially for syntactically multilinear models [SV15; KMS+13; AvMV15]. Building on the generating set of Theorem 2, we state a more structured test for membership in $\mathrm{Van}[\mathrm{RFE}_l^k]$ in terms of those operations.

Theorem 4 (membership test). *Let $k, l \in \mathbb{N}$, $X = \{x_i : i \in [n]\}$ be a set of variables, a_i for $i \in [n]$ be distinct field elements, and Z be a set of at least $n-k-l-1$ nonzero field elements. A multilinear polynomial $p \in \mathbb{F}[X]$ belongs to $\mathrm{Van}[\mathrm{RFE}_l^k]$ if and only if both of the following conditions hold:*

1. *There are no monomials of degree l or less, nor of degree $n-k$ or more, in p.*
2. *For all disjoint subsets $K, L \subseteq [n]$ with $|K| = k$ and $|L| = l$, and every $z \in Z$, $\left.\left(\frac{\partial p}{\partial L}\right)\right|_{K \leftarrow 0}$ evaluates to zero upon the following substitution for each $i \in \overline{K \cup L}$*

$$x_i \leftarrow z \cdot \frac{\prod_{j \in K}(a_i - a_j)}{\prod_{j \in L}(a_i - a_j)}. \tag{6}$$

A few technical comments regarding the statement: The first part of condition 1 in Theorem 4 reflects the known property that SV^l hits every multilinear polynomial that contains a monomial of degree l or less [SV15]. The second part expresses an analogous property, and together they imply that all multilinear polynomials on $n \le k+l+1$ variables are hit by RFE_l^k. In condition 2, $\left.\left(\frac{\partial p}{\partial L}\right)\right|_{K \leftarrow 0}$ denotes the polynomial obtained by taking the partial derivative of p with respect to every variable in L and setting all the variables in K to zero. The order of the operations does not matter, and the resulting polynomial depends only on variables in $\overline{K \cup L}$. The set Z serves only to separate the homogeneous components of p (cf. Theorem 16 and Proposition 17 in Section 4). One can just as well use a fresh variable z instead of using substitions from Z, in which case condition 2 becomes a single substitution for each K and L. The substitution (6) can be viewed as $x_i \leftarrow z \cdot f_{K,L}(a_i)$, where

$$f_{K,L}(\alpha) \doteq \frac{\prod_{j \in K}(\alpha - a_j)}{\prod_{j \in L}(\alpha - a_j)}$$

is a valid seed for RFE_l^k when substituted into the variables in $\overline{K \cup L}$.

All together, Theorem 4 can be understood as stating that a multilinear polynomial p is hit by RFE_l^k if and only if p has a monomial supported on few or all-but-few variables, or else there is a small set of zero substitutions, K, and a small set of partial derivatives, L, whose application to p leaves a polynomial that is nonzero after substituting $x_i \leftarrow z \cdot f_{K,L}(a_i)$. As we mentioned above, several prior papers demonstrated the utility of partial derivatives and zero substitutions in the context of derandomizing PIT using the SV generator. By judiciously choosing variables for those operations, these papers managed to simplify p and reduce PIT for p to PIT for simpler instances, resulting in efficient recursive algorithms. In Section 4, we develop a general framework for such algorithms and prove correctness directly from Theorem 4. Moreover, because Theorem 4 is a precise characterization, any argument that SV or RFE hits a class of multilinear polynomials can be converted into one within our framework, i.e., into an argument based on zero substitutions and partial derivatives. Thus, Theorem 4 shows that these tools harness the complete power of SV and RFE for multilinear polynomials.

Applications. We illustrate the utility of our characterizations of the vanishing ideal of RFE in the two directions mentioned before.

Derandomization. To start, we demonstrate how Theorem 4 yields an alternate proof of the result from [MV18] that SV^1—equivalently, RFE_1^0—hits every nonzero read-once formula F. Whereas the original proof hinges on a clever ad-hoc argument, our proof (described in Section 4) is entirely systematic and amounts to a couple straightforward observations in order to apply Theorem 4.

As a proof of concept of the additional power of our characterization for derandomization, we make progress in the model of read-once oblivious algebraic branching programs (ROABPs).

Theorem 5 (ROABP hitting property). *For integer $l \geq 1$, SV^l hits the class of polynomials computed by read-once oblivious algebraic branching programs of width less than $(l/3)+1$ that contain a monomial of degree at most $l+1$.*

To the best of our knowledge, Theorem 5 is incomparable to the known results for ROABPs [RS05; JQS09; JQS10; FS13; FSS14; AGK+15; AFS+18; GKS+17; GKS17; GG20; ST21; BG21]. Without the restriction that the polynomial has a monomial of degree at most $l+1$, Theorem 5 would imply a fully blackbox polynomial-time identity test for the class of constant-width ROABPs. No such test has been proven to exist at this time; prior work requires either quasipolynomial time or requires opening the blackbox, such as by knowing the order in which the variables are read.

With the restriction, hitting the class in Theorem 5 with a generator G amounts to proving a lower bound on the width of ROABPs that compute nonzero degree-$(l+1)$ polynomials in the vanishing ideal of G. For $G = \mathrm{SV}^l$, such a lower bound is interesting because there are nonzero degree-$(l+1)$ polynomials in the vanishing ideal, and, moreover, the vanishing ideal is generated by such polynomials (Theorem 2). That Theorem 5 holds suggests that the lower bound could hold for polynomials in $\mathrm{Van}[\mathrm{SV}^l]$ of all degrees, in which case one would fully derandomize PIT for constant-width ROABPs.

This stands in contrast to the case where $G = \mathrm{SV}^{l+1}$. It is well-known that SV^{l+1} hits every polynomial containing a monomial of support $l+1$ or less, and thus it hits the class in Theorem 5, irrespective of the restriction on ROABP width. Indeed, there are *no* nonzero polynomials of degree $l+1$ or less in the vanishing ideal of SV^{l+1}. While the lower bound necessary for SV^{l+1} to hit the class in Theorem 5 holds, it does so only vacuously—a phenomenon that has no hope of extending to higher degrees.

The method of proof of Theorem 5 diverges significantly from prior uses of the SV generator and therefore may be of independent interest. We elaborate on the method more when we discuss the techniques of this paper, but for now, we point out that most prior uses of the SV generator rely on combinatorial arguments, i.e., arguments that depend only on which monomials are present in the polynomials to hit. Theorem 5 necessarily goes beyond this because there is a polynomial in $\mathrm{Van}[\mathrm{SV}^l]$ of degree $l+1$ that has the same monomials as a polynomial computed by an ROABP of width 2, which by Theorem 5 is not in $\mathrm{Van}[\mathrm{SV}^l]$ for $l \geq 4$. Namely, any instance of ESMVC_l^{l-1} contains exactly all the monomials of the form $x_{i_1} \cdot x_{i_2} \cdot \dots \cdot x_{i_{l+1}}$ with $(i_1, \dots, i_{l+1}) \in X_1 \times \dots \times X_{l+1}$ for some disjoint sets X_j; the same goes for $\prod_j \sum_{i_j \in X_j} x_{i_j}$, which is computed by an ROABP of width 2.

Lower bounds. Our result for ROABPs also illustrates this direction. Our derandomization result for the class in Theorem 5 is equivalent to a lower bound of at least $(l/3)+1$ on the width of any ROABP computing a nonzero degree-$(l+1)$ polynomial in the vanishing ideal of SV^l, and in particular implies the lower bound for EVC_l^{l-1} and ESMVC_l^{l-1}. Other hardness of representation results follow in a similar manner from prior hitting properties of SV in the literature. The following lower bounds apply to computing both EVC_l^{l-1} and ESMVC_l^{l-1}:

- Any syntactically multilinear formula must have at least $\Omega(\log(l)/\log\log(l))$ reads of some variable [AvMV15, Theorem 6.3].
- Any sum of read-once formulas must have at least $\Omega(l)$ terms [MV18, Corollary 5.2].
- There exists an order of the variables such that any ROABP with that order must have width at least $2^{\Omega(l)}$ [FSS14, Corollary 4.3].
- Any $\Sigma \mathrm{m} \wedge \Sigma\Pi^{O(1)}$ formula must have top fan-in at least $2^{\Omega(l)}$ [For15]. See also [FSV18, Lemma 5.12].
- Lower bounds over characteristic zero for circuits with locally-low algebraic rank [KS17, Lemma 5.2].

Techniques. Many of our results ultimately require showing that, under suitable conditions, RFE hits a polynomial p. A recurring analysis fulfills this role in the proofs of Theorems 2, 4 and 5, as well as several of the other results. We take intuition from the analytic setting (e.g., $\mathbb{F} = \mathbb{R}$) and study the behavior of $p(\mathrm{RFE})$ as a function of the seed when the seed's zeroes and poles are near the abscissas of chosen variables of p. The behavior is dominated by the contributions of the monomials of p for which the variables with abscissas near zeros have minimal degree and the variables with abscissas near poles have maximal degree. Thus we may analyze a first approximation to $p(\mathrm{RFE})$ by "zooming in" on the contributions of the monomials in which the chosen variables have extremal degrees. If the first approximation is nonzero, then we can conclude that RFE hits p. We capture the technique in our Zoom Lemma (Lemma 14). Formal Laurent series can express the analytic intuition purely algebraically. We provide a proof from first principles that does not require any background in Laurent series and works over all fields.

Theorem 2 states the equality of two ideals, $\langle \mathrm{EVC}_l^k \rangle = \mathrm{Van}[\mathrm{RFE}_l^k]$, where $\langle \mathrm{EVC}_l^k \rangle$ denotes the ideal generated by all instantiations of EVC_l^k, and $\mathrm{Van}[\mathrm{RFE}_l^k]$ the vanishing ideal of RFE_l^k.

- The inclusion $\subseteq$ follows from linearizing the defining equations of RFE_l^k. The technique mirrors the use of resultants to compute implicit equations for rational plane curves. This is where the univariate dependence on the abscissas comes into play.
- To establish the inclusion $\supseteq$ we first show that every equivalence class of polynomials modulo $\langle \mathrm{EVC}_l^k \rangle$ contains a representative p whose monomials exhibit the combinatorial structure of a core. The structure allows us to apply the Zoom Lemma such that the zoomed-in

contributions of p have only a single monomial. A single monomial is nonzero at the evaluation in the Zoom Lemma, and we conclude that RFE_l^k hits p.

The proof of Theorem 4 also relies on the Zoom Lemma. Membership to the ideal is equivalent to the vanishing of all coefficients of the expansion. The proof can be viewed as determining a small number of coefficients sufficient to guarantee that their vanishing implies all coefficients vanish. The restriction to multilinear polynomials p allows us to express the zoomed-in contributions of p as the result of applying partial derivatives and zero-substitutions.

Theorem 5 makes use of the characterization of the minimum width of a read-once oblivious algebraic branching program computing a polynomial p as the maximum rank of the monomial coefficient matrices of p for various variable partitions [Nis91]. We reduce to the case where p is homogeneous of degree $l+1$, whence the monomial coefficient matrices have a block-diagonal structure. An application of the Zoom Lemma in the contrapositive yields linear equations between elements of consecutive blocks under the assumption that SV^l fails to hit p. When some block is zero, the equations yield a Cauchy system of equations on the rows or columns of its neighboring blocks; since Cauchy systems have full rank, we deduce severe constraints on the row-space/column-space of the neighboring blocks. A careful analysis turns this observation into a rank lower bound of at least $(l/3)+1$ for a well-chosen partition of the variables.

We point out that, in the preceding application, the Zoom Lemma is instantiated several times in parallel to form a large system of equations on the coefficients of p, and the whole system is necessary for the analysis. This stands in contrast to most prior work using SV, which uses knowledge of how p is computed to guide a search for a *single* fruitful instantiation of the Zoom Lemma.

Alternating algebra representation. The inspiration for several of our results stems from expressing the polynomials EVC_l^k using concepts from alternating algebra (also known as exterior algebra or Grassmann algebra). In fact, the relationship between Theorems 2 and 4 is based on the relationship $\partial^2 = 0$ from alternating algebra. Our original statement and proof of the theorem made use of that framework, but we managed to eliminate the alternating algebra afterwards. Still, as we find the perspective insightful and potentially helpful for future developments, we describe the connection briefly here and in more detail in Section 8. We explain the intuition behind Theorem 4 for the simple case where the degree of the polynomial p equals $l+1$. In that setting, belonging to the ideal generated by the polynomials EVC_l^k is equivalent to being in their linear span.

The alternating algebra A of a vector space V over a field $\mathbb{F}$ consists of the closure of V under an additional binary operation, referred to as "wedge" and denoted $\wedge$, which is bilinear, associative, and satisfies

$$v \wedge v = 0 \tag{7}$$

for every $v \in V$. This determines a well-defined algebra. When the characteristic of $\mathbb{F}$ is not 2, this can equivalently be understood as

$$v_1 \wedge v_2 = -(v_2 \wedge v_1) \tag{8}$$

for every $v_1, v_2 \in V$. In any case, for any $v_1, v_2, \ldots, v_k \in V$,

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k \tag{9}$$

is nonzero iff the v_i's are linearly independent, and any permutation of the order of the vectors in (9) yields the same element of A up to a sign. The sign equals the sign of the permutation, whence the name "alternating algebra." If V has a basis X of size n, then a basis for A can be formed by all 2^n expressions of the form (9) where the v_i's range over all subsets of X and are taken in some fixed order. Considering the elements of X as vertices, the basis elements of A can be thought of as the oriented simplices of all dimensions that can be built from X.

Anti-commutativity, the relation (8), arises naturally in the context of network flow, where X denotes the vertices of the underlying graph, and a wedge $v_1 \wedge v_2$ of level $k = 2$ represents one unit of flow from v_1 to v_2. Equation (8) reflects the fact that one unit of flow from v_1 to v_2 cancels with one unit of flow from v_2 to v_1. The adjacent levels $k = 1$ and $k = 3$ also have natural interpretations in the flow setting: v_1 (the element of A of the form (9) with $k = 1$) represents one unit of surplus flow at v_1 (the vertex of the graph), and $v_1 \wedge v_2 \wedge v_3$ abstracts an elementary circulation of one unit along the directed cycle $v_1 \to v_2 \to v_3 \to v_1$.

The different levels are related by so-called boundary maps. Boundary maps are linear transformations that map a simplex to a linear combination of its subsimplices of one dimension less. The maps are parametrized by a weight function $w : X \to \mathbb{F}$, and defined by

$$\partial_w : v_1 \wedge v_2 \wedge \cdots \wedge v_m \mapsto \sum_{i=1}^{m} (-1)^{i+1} w(v_i)\, v_1 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} \wedge \cdots \wedge v_m, \tag{10}$$

an expression resembling the Laplace expansion of a determinant along a column $[w(v_i)]_{i=1}^m$. In the flow setting, using $w \equiv 1$, applying ∂_1 to $v_1 \wedge v_2$ yields $v_2 - v_1$, the superposition of demand at v_1 and surplus at v_2 corresponding to one unit of flow from v_1 to v_2. Likewise, ∂_1 sends the abstract elementary cycle $v_1 \wedge v_2 \wedge v_3$ to the superposition of the three edge flows that make up the cycle. A linear combination p of terms (9) with $k = 2$ represents a valid circulation iff it satisfies conservation of flow at every vertex, which can be expressed as $\partial_1(p) = 0$, i.e., p is in the kernel of ∂_1. An equivalent criterion is for p to be the superposition of elementary circulations, which can be expressed as p being in the image of ∂_1. The relationship between the image and the kernel of boundary maps holds in general: For any linearly independent $w_0, \ldots, w_m$, it holds that

$$\mathrm{Im}\left(\partial_{w_m} \circ \partial_{w_{m-1}} \circ \cdots \circ \partial_{w_0}\right) = \bigcap_{i=0}^{m} \mathrm{Ker}\left(\partial_{w_i}\right). \tag{11}$$

(When $w_0, \ldots, w_m$ are linearly dependent, $\partial_{w_m} \circ \cdots \circ \partial_{w_0}$ is the zero map.)

In the context of RFE, the set X creates a vertex for each variable, and simplices correspond to multilinear monomials. The anti-commutativity of $\wedge$ coincides with the fact that swapping two arguments to EVC_l^k means swapping two rows in (4), which changes the sign of the determinant. Using the above boundary maps, the right-hand side of (4) can be viewed as $\partial_\omega(v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_{k+l+2}})$, where $\partial_\omega \doteq \partial_{w_k} \circ \partial_{w_{k-1}} \circ \cdots \circ \partial_{w_0}$ and $w_d(v_i) \doteq (a_i)^d$. By (11), this means that EVC_l^k is in the kernel of ∂_{w_d} for each $d \in \{0, 1, \ldots, k\}$, or equivalently, in the kernel of $\partial_{\widetilde{w}}$ for each $\widetilde{w} : X \to \mathbb{F}$ of the form $\widetilde{w}(v_i) = w(a_i)$ where w is a polynomial of degree at most k. This is precisely the condition (5). In fact, (11) implies that the linear span of the generators EVC_l^k consists exactly of the polynomials of degree $l+1$ in this kernel. The latter condition is precisely what condition 2 in Theorem 4 expresses.

Further research. In this paper we propose to investigate the power of generators through characterizations of their vanishing ideals, develop such characterizations for the generators SV and RFE, and initiate a systematic study of the repercussions. Completing the study entails going over classes $\mathcal{C}$ of polynomials of interest, checking whether they intersect nontrivially with the vanishing ideal, and, if so, checking whether the intersection can be hit by other generators. One specific target is the elimination of the degree restriction from Theorem 5.

In addition to Shpilka–Volkovich, other specific generators for which we suggest to characterize their vanishing ideals include Klivans–Spielman and Gabizon–Raz, as well as generators based on isolating weight assignments find recurring use in the PIT literature. Other possibilities may include the RFE generator with pseudorandom abscissas, or higher-level work that relates the vanishing ideal of a combination of generators to the vanishing ideals of the constituent generators.

Along a related line of thought, the generator SV^l is the canonical example of an l-wise independent generator in the algebraic setting. Understanding the power such generators more broadly should lead to useful insights for derandomizing PIT. This work demonstrates explicit polynomials like EVC_l^{l-1} and ESMVC_l^{l-1} that are not automatically hit by l-wise independence; indeed they are not hit by SV^l. Is there a deeper underlying reason related to l-wise independence?

Lastly, we remark that our derivation of EVC from RFE in Section 2 can be abstracted to construct a polynomial p in $\mathrm{Van}[G]$ for any generator G, provided that G substitutes rational functions for which the numerators and denominators belong to linear spaces whose dimensions sum to less than n. As the resulting polynomial p is the determinant of an $n \times n$ or smaller matrix, such a generator G cannot derandomize PIT for models that compute such determinants. One may view this phenomenon as a possible barrier for derandomization. Our derivation works for the RFE generator, but does not directly apply to the SV generator, even though the two are equivalent in derandomization power. It would be useful to have sufficient conditions under which a generator and *all* its reparametrizations avoid the barrier.

Organization. We construct the generating set for the vanishing ideal (Theorem 2) in Section 2, followed by the Zoom Lemma in Section 3. The ideal membership test (Theorem 4) is developed in Section 4. We present the results on sparseness in Section 5, and the ones on set-multilinearity in Section 6. Background on ROABPs and our result on derandomizing PIT for ROABPs (Theorem 5) are covered in Section 7. We end our paper in Section 8 with a further discussion of the alternating algebra representation. The appendix contains some details and a formal treatment of the relationship between RFE and SV.

2 Generating Set

In this section, we establish Theorem 2, our characterization of the vanishing ideal of RFE in terms of an explicit generating set. For every $k, l \in \mathbb{N}$, we develop a template, EVC_l^k, for constructing polynomials that belong to the vanishing ideal of RFE_l^k such that all instantiations collectively generate the vanishing ideal.

The template can be derived in the following fashion. Fix any seed f of RFE_l^k, and write it as $f = g/h$ where $g(\alpha) = \sum_{d=0}^k g_d \alpha^d$ and $h(\alpha) = \sum_{d=0}^l h_d \alpha^d$ are respectively polynomials of degree k and l. For any $i \in [n]$, the polynomial $x_i - g(a_i)/h(a_i) \in \mathbb{F}[x_1, \ldots, x_n]$ vanishes by definition at $\mathrm{RFE}_l^k(f)$. While this polynomial varies with f, it does so uniformly. Specifically, after rescaling to $h(a_i)x_i - g(a_i)$, the polynomial depends only *linearly* on the coefficients of g and h. This uniformity can be exploited to construct a polynomial that vanishes at $\mathrm{RFE}_l^k(f)$ but that now is *independent* of f. Since f is arbitrary, the constructed polynomial belongs to the vanishing ideal of RFE_l^k.

The construction begins by expressing the vanishing of each $h(a_i)x_i - g(a_i)$ at $\mathrm{RFE}_l^k(f)$ as the following system of equations. Abbreviating

$$\begin{aligned} \vec{g} &\doteq \begin{bmatrix} g_k & g_{k-1} & \ldots & g_1 & g_0 \end{bmatrix}^\top \\ \vec{h} &\doteq \begin{bmatrix} h_l & h_{l-1} & \ldots & h_1 & h_0 \end{bmatrix}^\top, \end{aligned}$$

we write

$$\begin{bmatrix} a_i^l x_i & a_i^{l-1} x_i & \ldots & x_i & a_i^k & a_i^{k-1} & \ldots & 1 \end{bmatrix}_{i \in [n]} \cdot \begin{bmatrix} \vec{h} \\ -\vec{g} \end{bmatrix} = 0. \tag{12}$$

Written this way, (12) has the form of a homogeneous system of linear equations. There is one equation for each $i \in [n]$ and one unknown for each of the $k+l+2$ parameters of the seed f. The system's coefficient matrix has no dependence on f. For any f, substituting $\mathrm{RFE}_l^k(f)$ into $x_1, \ldots, x_n$ yields a system that has a nontrivial solution, namely the vector in (12).

Consider, then, the determinant of the square subsystem of (12) formed by any $k+l+2$ rows. It is a polynomial in $\mathbb{F}[x_1, \ldots, x_n]$. Because the coefficient matrix in (12) is independent of f, the determinant is independent of f. Because the subsystem has a nonzero solution after substituting $\mathrm{RFE}_l^k(f)$ for any f, the determinant vanishes after substituting $\mathrm{RFE}_l^k(f)$ for any f. We conclude that the determinant belongs to the vanishing ideal of RFE_l^k.

Recalling that the determinant for the subsystem consisting of rows $i_1, \ldots, i_{k+l+2}$ is identically $\mathrm{EVC}_l^k[i_1, i_2, \ldots, i_{k+l+2}]$, we have established:

Proposition 6. *For every $k, l \in \mathbb{N}$ and every $i_1, i_2, \ldots, i_{k+l+2} \in [n]$, $\mathrm{EVC}_l^k[i_1, \ldots, i_{k+l+2}] \in \mathrm{Van}[\mathrm{RFE}_l^k]$.*

Before moving on, we point out the following properties.

Proposition 7. *If any of $i_1, \ldots, i_{k+l+2}$ coincide, then $\mathrm{EVC}_l^k[i_1, \ldots, i_{k+l+2}]$ is zero. Otherwise, it is nonzero, multilinear, and homogeneous of total degree $l+1$, and every multilinear monomial of degree $l+1$ in $x_{i_1}, \ldots, x_{i_{k+l+2}}$ appears with a nonzero coefficient. EVC_l^k is skew-symmetric in that, for any permutation π of $i_1, \ldots, i_{k+l+2}$,*

$$\mathrm{EVC}_l^k[i_1, \ldots, i_{k+l+2}] = (-1)^{\mathrm{sign}(\pi)} \cdot \mathrm{EVC}_l^k[\pi(i_1), \ldots, \pi(i_{k+l+2})].$$

The coefficient of $x_{i_1} \cdot \cdots \cdot x_{i_{l+1}}$ in $\mathrm{EVC}_l^k[i_1, \ldots, i_{k+l+2}]$ is the product of Vandermonde determinants

$$\begin{vmatrix} a_{i_1}^l & \cdots & 1 \\ \vdots & \ddots & \vdots \\ a_{i_{l+1}}^l & \cdots & 1 \end{vmatrix} \begin{vmatrix} a_{i_{l+2}}^k & \cdots & 1 \\ \vdots & \ddots & \vdots \\ a_{i_{l+k+2}}^k & \cdots & 1 \end{vmatrix}.$$

Proof. All the assertions to be proved follow from elementary properties of determinants, that Vandermonde determinants are nonzero unless they have duplicate rows, and the following computation: After plugging in 1 for $x_{i_1}, \ldots, x_{i_{l+1}}$, and 0 for $x_{i_{l+2}}, \ldots, x_{i_{l+k+2}}$, the determinant has the form

$$\begin{vmatrix} a_{i_1}^l & \cdots & 1 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i_{l+1}}^l & \cdots & 1 & * & \cdots & * \\ 0 & \cdots & 0 & a_{i_{l+2}}^k & \cdots & 1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{i_{l+k+2}}^k & \cdots & 1 \end{vmatrix},$$

which equals the product of Vandermonde matrices in the statement. $\square$

Proposition 6 shows that the polynomials $\mathrm{EVC}_l^k[i_1, \ldots, i_{k+l+2}]$ belong to the vanishing ideal of RFE_l^k. To prove that they collectively generate the vanishing ideal, we use a two-phase approach:

1. We show that every polynomial is equal, modulo the ideal $\langle \mathrm{EVC}_l^k \rangle$ generated by the instantiations of EVC_l^k, to a polynomial with a particular combinatorial structure (Lemma 9).

2. We then show that every nonzero polynomial with that structure is hit by RFE_l^k (Lemma 11).

Together, these show that every polynomial in the vanishing ideal of RFE_l^k is equal, modulo $\langle \mathrm{EVC}_l^k \rangle$, to the zero polynomial. We conclude that the ideals coincide, i.e., the vanishing ideal is generated by instantiations of EVC_l^k.

The combinatorial structure bridging the two phases is that the polynomial is cored.

Definition 8 (cored polynomial). *For $c, t \in \mathbb{N}$, a polynomial p is said to be (c,t)-*cored *if there exists a set of at most c variables, called the* core*, such that every monomial of p depends on at most t variables outside the core.*

Lemma 9 formalizes the first phase of our approach.

Lemma 9. *Let $k, l \in \mathbb{N}$ and let C be a $(k+1)$-subset of $[n]$. Let $\langle \mathrm{EVC}_l^k[C] \rangle$ be the ideal generated by the polynomials $\mathrm{EVC}_l^k[S]$ where S ranges over all $(k+l+2)$-subsets of $[n]$ that contain C. Modulo $\langle \mathrm{EVC}_l^k[C] \rangle$ every polynomial is equal to a $(k+1,l)$-cored polynomial with core indexed by C*

Proof. Fix k, l, and C as in the statement. Every monomial m can be uniquely factored as $m_0 m_1$, where m_0 is supported on variables indexed by C, and m_1 is supported on variables indexed by $\overline{C}$. Call m_1 the *non-core* of m. We show the following:

Claim 10. *Modulo $\langle \mathrm{EVC}_l^k[C] \rangle$, every monomial with more than l variables in its non-core is equivalent to a linear combination of monomials that all have non-cores of lower degree.*

This lets us prove Lemma 9 as follows. For any polynomial p, Claim 10 implies that we may, without changing p modulo $\langle \mathrm{EVC}_l^k[C] \rangle$, eliminate any monomial in p that violates the $(k+1,l)$-cored condition, while possibly introducing monomials with lower non-core degree. Thus we can systematically eliminate all monomials that violate the cored condition by eliminating them in order of decreasing non-core degree. After that, p is $(k+1,l)$-cored with core indexed by C, and the lemma follows.

It remains to show Claim 10. Let m be a monomial with more than l variables in its non-core. Let $L \subseteq [n]$ index a set of $l+1$ of the variables in the non-core; let m' be the product of the variables indexed by L; and let m'' be the complementary factor of m, i.e., the monomial that satisfies $m = m'm''$. The union of L and C has size exactly $k+l+2$; we set $q \doteq \mathrm{EVC}_l^k[L \cup C]$, where the variables in $L \cup C$ are ordered arbitrarily. By Proposition 7, m' appears in q, and every other monomial in q has lower non-core degree than m'. It follows that m appears with nonzero coefficient in $m'' \cdot q$, and every other monomial in $m'' \cdot q$ has lower non-core degree than m. Since ideals are closed under multiplication by any other polynomial, $m'' \cdot q$ is in $\langle \mathrm{EVC}_l^k[C] \rangle$. The desired equivalence now follows by expressing the left-hand side of the equation $m'' \cdot q \equiv 0 \pmod{\langle \mathrm{EVC}_l^k[C] \rangle}$ as a sum of monomials and then isolating m. □

The second phase of our approach is formalized in Lemma 11.

Lemma 11. *Suppose p is nonzero and $(k+1,l)$-cored. Then RFE_l^k hits p.*

We prove Lemma 11 from the Zoom Lemma in Section 3. Before moving on to that, let us argue how the "moreover" part of Theorem 2 also follows. The combination of Proposition 6, Lemma 9, and Lemma 11 shows that, for every core $C \subseteq \{x_1, \dots, x_n\}$ of $k+1$ variables, the instances of EVC_l^k that use all the variables in C generate the vanishing ideal. This generating set has minimum size because the generators are all homogeneous of the same degree, and each generator has a monomial that occurs in none of the other generators (namely the product of the variables outside C). This completes the proof of Theorem 2 modulo the proof of Lemma 11.

3 Zoom Lemma

Throughout the paper we make repeated use of a key technical tool, the Zoom Lemma. The lemma allows us to zoom in on the contributions of the monomials in a polynomial p that have prescribed degrees in a subset of the variables. We introduce the following terminology for prescribing degrees.

Definition 12 (degree pattern). *Let $J \subseteq [n]$. A* degree pattern *with domain J is a J-indexed tuple $d \in \mathbb{N}^J$ of nonnegative integers. A degree pattern d* matches *a monomial $m \in \mathbb{F}[x_1, \ldots, x_n]$ if, for every $j \in J$, m has degree exactly d_j in x_j. We say that d is in p if d matches some monomial in p.*

For any fixed J, every polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ can be written uniquely in the form

$$p = \sum_{d \in \mathbb{N}^J} p_d \cdot x^d$$

where $x^d \doteq \prod_{j \in J} x_j^{d_j}$ and p_d depends only on variables not indexed by J. We refer to p_d as the coefficient of d in p.

The notation p_d can be viewed as a generalization of the common one for the coefficient of degree d of a univariate polynomial p.

Our technique allows us to zoom in on the contributions of the coefficients p_d of degree patterns d with domain $J = K \cup L$ that satisfy the following additional constraint.

Definition 13 (extremal degree pattern). *Let $K, L \subseteq [n]$. A degree pattern $d^* \in \mathbb{N}^{K \cup L}$ is* (K, L)-extremal *in a polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ if the only degree pattern $d \in \mathbb{N}^{K \cup L}$ in p that satisfies*

(i) $d_j \leq d_j^*$ for all $j \in K$, and

(ii) $d_j \geq d_j^*$ for all $j \in L$

is d^* itself.

When K and L intersect, note that only degree patterns $d \in \mathbb{N}^{K \cup L}$ with $d_j = d_j^*$ for all $j \in K \cap L$ have any bearing on whether d^* is (K, L)-extremal.

The notion of extremality in Definition 13 is closely related to standard notions of minimality and maximality of tuples of numbers. A J-tuple d^* is minimal in a set D of such tuples if the only tuple $d \in D$ that satisfies $d_j \leq d_j^*$ for all $j \in J$, is d^* itself. A maximal tuple is defined similarly by replacing $\leq$ by $\geq$. Minimality is equivalently $(J, \varnothing)$-extremality, and maximality is equivalently $(\varnothing, J)$-extremality.

The above terminology lets us state our key technical lemma succinctly.

Lemma 14 (Zoom Lemma). *Let $K, L \subseteq [n]$, let $p \in \mathbb{F}[x_1, \ldots, x_n]$, and let $d^* \in \mathbb{N}^{K \cup L}$ be a degree pattern that is (K, L)-extremal in p. If the coefficient p_{d^*} has nonzero evaluation at the point*

$$x_i \leftarrow z \cdot \frac{\prod_{j \in K \setminus L} (a_i - a_j)}{\prod_{j \in L \setminus K} (a_i - a_j)} \qquad \forall i \in [n] \setminus (K \cup L) \tag{13}$$

for some $z \in \mathbb{F}$, then RFE_l^k hits p with $k = |K|$ and $l = |L|$.

Note that, since the coefficient p_{d^*} depends on no variable indexed by $K \cup L$, the result of substituting (13) into p_{d^*} is simply a scalar in $\mathbb{F}$. Also, whereas typical instantiations of the Zoom Lemma have K and L disjoint, this is not necessary for the lemma to hold.[1]

Let us first see how the Zoom Lemma allows us to complete the proof of Theorem 2.

Proof of Lemma 11 from the Zoom Lemma. Let $C \subseteq [n]$ denote a core of size at most $k+1$ for p. We construct subsets $K, L \subseteq [n]$ with $|K| \leq k$ and $|L| \leq l$, a degree pattern d^* with domain $K \cup L$ that is (K, L)-extremal in p, and a scalar $z \in \mathbb{F}$ such that p_{d^*} is nonzero at the point (13). The Zoom Lemma then implies that RFE_l^k hits p.

The construction consists of two steps. First, we pick $i^* \in C$ arbitrarily (we can assume without loss of generality that C is nonempty), set $K \doteq C \smallsetminus \{i^*\}$, and let d_+ be a degree pattern with domain K that matches a monomial in p and that is minimal among all such degree patterns. The existence of d_+ follows from the fact that p is nonzero. By construction, $|K| \leq (k+1) - 1 = k$ and p_{d_+} is nonzero.

Second, we pick a degree pattern d_- with domain $\overline{C}$ that matches a monomial in p_{d_+} and that is maximal among all such degree patterns. The existence of d_- follows from the fact that p_{d_+} is nonzero. Let L denote the set of indices $j \in \overline{C}$ on which d_- is positive. The hypothesis that C is a $(k+1, l)$-core for p implies that $|L| \leq l$. By construction, the restriction of d_- to the domain L is maximal among the degree patterns with domain L in p_{d_+}.

Note that K and L are disjoint, because $K \subseteq C$ and $L \subseteq \overline{C}$. We define d^* as the degree pattern with domain $K \cup L$ that agrees with d_+ on K and with d_- on L. The minimality and maximality properties of d_+ and d_- imply that d^* is (K, L)-extremal in p. The coefficient p_{d^*} is a nonzero polynomial that depends only on x_{i^*}. It follows that for all but finitely many $z \in \mathbb{F}$, substituting (13) into p_{d^*} yields a nonzero result. $\square$

Before giving a formal proof of the Zoom Lemma, we provide some intuition for the mechanism behind it, and we explain how the choice of the evaluation point (13) and the extremality requirement arise. We start with the special case where (i) $\ell = 0$, or equivalently $L = \varnothing$, and (ii) the degree pattern $d^* \in \mathbb{N}^K$ is zero in every coordinate, so x^{d^*} is the constant monomial 1. We can zoom in on p_{d^*} by setting all variables x_j for $j \in K$ to zero. The generator RFE_0^k allows us to do so by picking a seed f such that $f(a_j) = 0$ for all $j \in K$, namely

$$f(\alpha) \doteq z \cdot \prod_{j \in K} (\alpha - a_j) \tag{14}$$

for any $z \in \mathbb{F}$. The evaluation of p at $\mathrm{RFE}(f)$ coincides with the evaluation of p_{d^*} at $\mathrm{RFE}(f)$, which is precisely (13). If the evaluation is nonzero, then evidently RFE_0^k hits p, as desired.

In order to handle more general degree patterns $d^* \in \mathbb{N}^K$, we introduce a fresh parameter ξ_j for each $j \in K$, and replace a_j in (14) by $a_j - \xi_j$, i.e., we consider the seeds

$$\hat{f}(\alpha) \doteq z \cdot \prod_{j \in K} (\alpha - a_j + \xi_j), \tag{15}$$

for any $z \in \mathbb{F}$, where the hat indicates a dependency on the parameters. For each i, $\hat{f}(a_i)$ is a multivariate polynomial in the parameters, and $\mathrm{RFE}(\hat{f})$ applies the substitution $x_i \leftarrow \hat{f}(a_i)$ for each $i \in [n]$. The parametrization ensures that $\hat{f}(a_i)$ is divisible by ξ_i for $i \in K$ but not for $i \notin K$. More precisely,

$$\hat{f}(a_i) = \begin{cases} \hat{c}_i \cdot \xi_i & i \in K \\ \hat{c}_i & i \notin K \end{cases},$$

[1] In fact, allowing K and L to overlap is useful in Section 7 (see Proposition 40).

where $\hat{c}_i \doteq z \cdot \prod_{j \in K \smallsetminus \{i\}} (a_i - a_j + \xi_j)$ is a multivariate polynomial in the parameters ξ with nonzero constant term, namely $c_i \doteq z \cdot \prod_{j \in K \smallsetminus \{i\}} (a_i - a_j)$. For any monomial m with matching degree pattern $d \in \mathbb{N}^K$, we have

$$m(\mathrm{RFE}(\hat{f})) = m(\hat{c}) \cdot \xi^d = m_d(\hat{c}) \cdot \hat{c}^d \cdot \xi^d.$$

Here we see that, when $m(\mathrm{RFE}(\hat{f}))$ is expanded as a linear combination of monomials in the ξ_j, the combination contains only monomials divisible by ξ^d.

In the expansion of $p(\mathrm{RFE}(\hat{f}))$, the coefficient of ξ^{d^*}:

(a) has a contribution $m(c) = m_{d^*}(c) \cdot c^{d^*}$ from each monomial m in p that matches d^*, and

(b) may have contributions from other monomials m in p but only from those whose degree pattern on K is smaller than d^*, i.e., only if $\deg_j(m) \leq d^*_j$ for all $j \in K$.

By adding the contributions of all monomials m with degree pattern d^* we obtain

$$p_{d^*}(\mathrm{RFE}(\hat{f})) \cdot \mathrm{RFE}(\hat{f})^{d^*} = p_{d^*}(\hat{c}) \cdot \hat{c}^{d^*} \cdot \xi^{d^*}.$$

By properties (a) and (b) above, we conclude that the coefficient of the monomial ξ^{d^*} in $p(\mathrm{RFE}(\hat{f}))$:

(a') has a contribution of $p_{d^*}(c) \cdot c^{d^*}$ from the monomials matching d^*, and

(b') cannot have any additional contributions provided that there are no degree patterns on K in p that are smaller than d^*.

For a degree pattern d^* in p, condition (b') can be formulated as the minimality of d^* among the degree patterns on K in p, which is exactly the requirement that d^* is (K, L)-extremal in p for $L = \varnothing$. Under this condition we conclude that the coefficient of the monomial ξ^{d^*} in $p(\mathrm{RFE}(\hat{f}))$ equals $p_{d^*}(c) \cdot c^{d^*}$. Note that c^{d^*} is nonzero. Since $p_{d^*}(c)$ only depends on the components c_i for $i \in [n] \smallsetminus K$, and those components agree with (13), the coefficient of the monomial ξ^{d^*} in $p(\mathrm{RFE}(\hat{f}))$ is nonzero if and only if p^*_d is nonzero at the point (13). Thus, the hypotheses of the lemma imply that $p(\mathrm{RFE}(\hat{f})$ is a nonzero polynomial in the parameters ξ. It follows that a random setting of the parameters ξ yields a seed f' for RFE^k_0 such that $p(\mathrm{RFE}(f'))$ is nonzero. This shows that RFE^k_0 hits p.

The symmetric case $k = 0$ can be obtained from the case $l = 0$ by transforming $x_i \mapsto x_i^{-1}$ for each $i \in [n]$. The transformation maps a seed f for RFE^0_l into a seed $\tilde{f}$ for RFE^l_0, wherein the zeroes of $\tilde{f}$ come from the poles of f. Given a polynomial $p(x_1, \ldots, x_n)$, we similarly transform the variables and clear dominators to obtain the polynomial $\tilde{p}(x_1, \ldots, x_n) \doteq p(x_1^{-1}, \ldots, x_n^{-1}) \cdot x^g$, where g is any degree pattern with domain $[n]$ for which g_i is at least the degree of x_i in p for every $i \in [n]$. We apply the previous case of the Zoom Lemma to $\tilde{p}$ and obtain the new case of the Zoom Lemma for p. Note that a monomial with degree pattern $\tilde{d}$ in $\tilde{p}$ corresponds to a monomial with degree pattern $d = g - \tilde{d}$ in p. It follows that $\widetilde{d^*}$ is minimal in $\tilde{p}$ iff d^* is maximal in p, which is exactly the (K, L)-extremality requirement of the Zoom Lemma in the case where $K = \varnothing$.

The general case follows in a similar fashion, introducing parameters for the zeroes as well as the poles of the seed f, considering the monomial in those parameters with degree pattern determined by d^*, and clearing denominators.

Proof of the Zoom Lemma. Fix K, L, p, d^*, and z as in the lemma statement, and let $k = |K|$ and $l = |L|$. Let ξ_j for each $j \in K$ and η_j for each $j \in L$ be fresh indeterminates. We denote by $\widehat{\mathbb{F}}$ the field of rational functions in those indeterminates with coefficients in $\mathbb{F}$, and by V the subset of elements that, when written in lowest terms, have denominators with nonzero constant terms. Let

$\Phi : V \to \mathbb{F}$ map each element of V to the result of substituting $\xi_j \leftarrow 0$ for each $j \in K$ and $\eta_j \leftarrow 0$ for each $j \in L$. The result is always well-defined.

Define $\hat{f} \in \widehat{\mathbb{F}}(\alpha)$ as follows:

$$\hat{f}(\alpha) \doteq z \cdot \frac{\prod_{j \in K}(\alpha - a_j + \xi_j)}{\prod_{j \in L}(\alpha - a_j + \eta_j)}.$$

The substitution $\mathrm{RFE}(\hat{f})$ effects $x_i \leftarrow \hat{f}(a_i) \in \widehat{\mathbb{F}}$ for each $i \in [n]$. We claim that $p(\mathrm{RFE}(\hat{f}))$ is nonzero. This suffices to conclude that RFE_l^k hits p, because substituting ξ_j and η_j by a random scalar from $\mathbb{F}$ transforms $\hat{f}$ into a seed f' such that, with high probability, f' is a valid seed for RFE_l^k and $p(\mathrm{RFE}(f')) \neq 0$. Henceforth we show that $p(\mathrm{RFE}(\hat{f})) \neq 0$.

For each $i \in [n]$, there exists $\hat{c}_i \in V$ with $\Phi(\hat{c}_i) \neq 0$ such that

$$\hat{f}(a_i) = \begin{cases} \hat{c}_i \cdot \frac{\xi_i}{\eta_i} & i \in K \cap L \\ \hat{c}_i \cdot \xi_i & i \in K \setminus L \\ \hat{c}_i \cdot \frac{1}{\eta_i} & i \in L \setminus K \\ \hat{c}_i & i \notin K \cup L \end{cases}, \tag{16}$$

namely

$$\hat{c}_i = z \cdot \frac{\prod_{j \in K \setminus \{i\}}(a_i - a_j + \xi_j)}{\prod_{j \in L \setminus \{i\}}(a_i - a_j + \eta_j)}.$$

For $i \notin K \cup L$, $\Phi(\hat{c}_i)$ is moreover the value substituted into x_i by (13).

Let D denote the set of all degree patterns $d \in \mathbb{N}^{K \cup L}$ that match a monomial in p. We have that

$$p = \sum_{d \in D} p_d \cdot x^d. \tag{17}$$

For $d \in D$, define $\hat{q}_d$ to be the result of substituting $x_i \leftarrow \hat{c}_i$ into p_d for each $i \in [n]$.

Combining (16) and (17), we obtain

$$p(\mathrm{RFE}(\hat{f})) = \sum_{d \in D} \hat{q}_d \cdot \hat{c}^d \cdot \frac{\xi^{d|_K}}{\eta^{d|_L}}, \tag{18}$$

where $d|_K$ and $d|_L$ respectively are the restrictions of d onto the domains K and L respectively. Fix any function $\psi : [k+l] \to K \cup L$ such that ψ establishes a bijection between $\{1, \ldots, k\}$ and K and establishes a bijection between $\{k+1, \ldots, k+l\}$ and L. For $j \in \{1, \ldots, k\}$, let ζ_j be an alias for $\xi_{\psi(j)}$, and for $j \in \{k+1, \ldots, k+l\}$, let ζ_j be an alias for $\eta_{\psi(j)}$. For each $d \in \mathbb{N}^{K \cup L}$, define a corresponding $\delta \in \mathbb{Z}^{k+l}$ given by $\delta_j = d_{\psi(j)}$ for $j \in \{1, \ldots, k\}$ and $\delta_j = -d_{\psi(j)}$ for $j \in \{k+1, \ldots, k+l\}$. Let $\Delta \subseteq \mathbb{Z}^{k+l}$ consist of the δ corresponding to each $d \in D$. Finally, for each $d \in D$ with corresponding $\delta \in \Delta$, define $\hat{c}_\delta \doteq \hat{q}_d \cdot \hat{c}^d$, capturing the first two factors in the d-th term of (18). Rewritten in this notation, (18) becomes

$$\sum_{\delta \in \Delta} \hat{c}_\delta \cdot \prod_{k=1}^{k+l} \zeta_j^{\delta_j}. \tag{19}$$

Our hypothesis that d^* is (K, L)-extremal in p says that the only $d \in D$ such that $d_j \le d_j^*$ for every $j \in K$ and $d_j \ge d_j^*$ for every $j \in L$, is $d = d^*$. Translated into a condition on the element $\delta^* \in \Delta$ corresponding to d^*, the hypothesis says that δ^* is minimal in Δ. Our other hypothesis states that p_{d^*} does not vanish upon substituting (13). As (13) equates to substituting $x_i \leftarrow \Phi(\hat{c}_i)$ for $i \notin K \cup L$, this hypothesis equivalently states that $\Phi(\hat{q}_{d^*})$ is nonzero. Since for each $j \in K \cup L$ we have $\Phi(\hat{c}_j) \neq 0$, we conclude that $\Phi(\hat{c}_{\delta^*}) \neq 0$. That $p(\mathrm{RFE}(\hat{f}))$ is nonzero now follows from the next proposition. $\square$

Proposition 15. *Let $\widehat{\mathbb{F}} = \mathbb{F}(\zeta_1, \ldots, \zeta_r)$ be the field of rational functions in indeterminates $\zeta_1, \ldots, \zeta_r$, let $V \subseteq \widehat{\mathbb{F}}$ consist of the rational functions whose denominator has nonzero constant term, and let $\Phi : V \to \mathbb{F}$ be the function that maps each rational function in V to its value after substituting $\zeta_j \leftarrow 0$ for all $j \in [r]$. Let*

$$s = \sum_{\delta \in \Delta} \hat{c}_\delta \cdot \prod_{j=1}^{r} \zeta_j^{\delta_j}$$

where $\Delta \subseteq \mathbb{Z}^r$ is some finite set, and we have $\hat{c}_\delta \in V$ for every $\delta \in \Delta$. If there exists $\delta^* \in \Delta$ that is minimal in Δ and for which $\Phi(\hat{c}_{\delta^*}) \neq 0$, then $s \neq 0$.

Proof. By clearing denominators, we may assume without loss of generality that, for every $\delta \in \Delta$ and every $j \in [r]$, $\delta_j \geq 0$, and that, for every $\delta \in \Delta$, $\hat{c}_\delta$ is a *polynomial* in $\zeta_1, \ldots, \zeta_r$. In this case, all quantities in the sum for s are polynomials in $\zeta_1, \ldots, \zeta_r$. The minimality hypothesis on δ^* implies that the coefficient of $\prod_{j=1}^{r} \zeta_j^{\delta_j^*}$ in the monomial expansion of s is precisely the constant coefficient of $\hat{c}_{\delta^*}$, and the hypothesis $\Phi(\hat{c}_{\delta^*}) \neq 0$ asserts that this coefficient is nonzero. □

4 Membership Test

In this section we develop the structured membership test for the vanishing ideal $\mathrm{Van}[\mathrm{RFE}_l^k]$ given in Theorem 4. We start by observing that it suffices to establish the following simpler version of Theorem 4 for the case where p is homogeneous.

Theorem 16. *A nonzero homogeneous multilinear polynomial p in the variables $x_1, \ldots, x_n$ belongs to $\mathrm{Van}[\mathrm{RFE}_l^k]$ if and only if both of the following conditions hold:*

1. *The degree of p satisfies $l < \deg(p) < n - k$.*
2. *For all disjoint subsets $K, L \subseteq [n]$ with $|K| = k$ and $|L| = l$, $\left(\frac{\partial p}{\partial L}\right)\Big|_{K \leftarrow 0}$ evaluates to zero upon substituting for each $i \in [n] \setminus (K \cup L)$*

$$x_i \leftarrow \frac{\prod_{j \in K}(a_i - a_j)}{\prod_{j \in L}(a_i - a_j)}. \tag{20}$$

To see why the general case reduces to the homogeneous case, we make use of the following property, well-known in the context of SV. We include a proof for completeness.

Proposition 17. *For any polynomial p, p vanishes upon substituting* RFE *if and only if every homogeneous part of p vanishes upon substituting* RFE*.*

Proof. For any seed f for RFE and any scalar z, the rescaled substitution $z \cdot \mathrm{RFE}(f)$ is in the range of RFE, namely as $\mathrm{RFE}(z \cdot f)$. It follows (provided that $\mathbb{F}$ is sufficiently large) that $p(\mathrm{RFE})$ vanishes if and only if $p(\zeta \cdot \mathrm{RFE})$ vanishes, where ζ is a fresh indeterminate. We now consider the expansion of $p(\zeta \cdot \mathrm{RFE})$ as a polynomial in ζ. With $p^{(d)}$ as the degree-d homogeneous part of p, we have

$$p(\zeta \cdot \mathrm{RFE}) = \sum_d p^{(d)}(\zeta \cdot \mathrm{RFE}) = \sum_d \zeta^d \cdot p^{(d)}(\mathrm{RFE}).$$

The coefficient of ζ^d, $p^{(d)}(\mathrm{RFE})$, has no dependence on ζ. We deduce that $p(\zeta \cdot \mathrm{RFE})$ is the zero polynomial if and only if $p^{(d)}(\mathrm{RFE})$ vanishes for every d. □

Here is how Theorem 4 follows from Theorem 16.

Proof of Theorem 4 from Theorem 16. For each $d \in \mathbb{N}$, let $p^{(d)}$ be the degree-d homogeneous part of p. By Proposition 17, $p \in \mathrm{Van}[\mathrm{RFE}_l^k]$ if and only if every $p^{(d)} \in \mathrm{Van}[\mathrm{RFE}_l^k]$. Condition 1 of Theorem 16 applied to each $p^{(d)}$ implies that condition 1 in Theorem 4 is necessary. In order to establish Theorem 4, it remains to consider polynomials p such that $p^{(d)} = 0$ for $d \leq l$ and $d \geq n-k$, and show that condition 2 of Theorem 4 holds for p if and only if, for every d with $l < d < n-k$, condition 2 of Theorem 16 holds for $p^{(d)}$.

Fix K, L, Z as in the statements of Theorem 4 and Theorem 16. Let $\lambda \in \mathbb{F}^Y$ where $Y = [n] \setminus (K \cup L)$ be the point (20), and note that (6) is $z \cdot \lambda$ for $z \in Z$. Let ζ be a fresh indeterminate. We have

$$\left(\frac{\partial p}{\partial L}\right)\bigg|_{K \leftarrow 0} (\zeta\lambda) = \sum_{l<d<n-k} \left(\frac{\partial p^{(d)}}{\partial L}\right)\bigg|_{K \leftarrow 0} (\zeta\lambda) = \sum_{l<d<n-k} \zeta^{d-l} \left(\frac{\partial p^{(d)}}{\partial L}\right)\bigg|_{K \leftarrow 0} (\lambda).$$

This is a polynomial in ζ, say $q(\zeta)$. Because q factors as ζ times a polynomial with degree at most $n-k-l-2$, it vanishes for some $n-k-l-1$ distinct nonzero choices for z—namely Z—if and only if for every d the coefficient of ζ^{d-l} vanishes. Substituting $\zeta \leftarrow z$ into q coincides with evaluating $\left(\frac{\partial p}{\partial L}\right)\big|_{K \leftarrow 0}$ at $z\lambda$; meanwhile, for each d the coefficient of ζ^{d-l} coincides with evaluating $\left(\frac{\partial p^{(d)}}{\partial L}\right)\big|_{K \leftarrow 0}$ at λ. Theorem 4 follows. $\square$

It remains to prove Theorem 16. We once again make use of the Zoom Lemma. Note that for multilinear polynomials and disjoint K and L, $\left(\frac{\partial p}{\partial L}\right)\big|_{K \leftarrow 0}$ coincides with the coefficient p_{d^*} where d^* is the degree pattern with domain $K \cup L$ with 0 in the positions indexed by K and 1 in the positions indexed by L. Moreover, since p is multilinear, the condition that d^* be (K, L)-extremal in p is automatically satisfied: The only multilinear monomial m supported in $K \cup L$ with $\deg_{x_i}(m) \leq d_i^* = 0$ for all $i \in K$ and $\deg_{x_i}(m) \geq d_i^* = 1$ for all $i \in L$ is $m = x^{d^*}$. This leads to the following specialization of the Zoom Lemma for multilinear polynomials with disjoint K and L:

Lemma 18. *Let $K, L \subseteq [n]$ be disjoint, and let $p \in \mathbb{F}[x_1, \ldots, x_n]$ be a multilinear polynomial. If $\left(\frac{\partial p}{\partial L}\right)\big|_{K \leftarrow 0}$ is nonzero at the point*

$$x_i \;\leftarrow\; z \cdot \frac{\prod_{j \in K} (a_i - a_j)}{\prod_{j \in L} (a_i - a_j)} \qquad \forall i \in [n] \setminus (K \cup L), \tag{21}$$

for some $z \in \mathbb{F}$, then RFE_l^k hits p with $k = |K|$ and $l = |L|$.

In proving Theorem 16, we will apply Lemma 18 only to homogeneous polynomials, in which case we can take $z = 1$ without loss of generality. With that in mind, observe that (20) in Theorem 16 coincides with the substitution (21) from Lemma 18. So Theorem 16 amounts to saying that a homogeneous multilinear polynomial p is hit by RFE_l^k if and only if its degree is too low, its degree is too high, or else there is a way to apply Lemma 18 to prove that p is hit by RFE_l^k.

Proof of Theorem 16. Suppose that p has degree $d \leq l$. Set L to be the indices of the variables appearing in some monomial with nonzero coefficient in p, and set $K \leftarrow \varnothing$. $\left(\frac{\partial p}{\partial L}\right)\big|_{K \leftarrow 0}$ is a nonzero constant. Lemma 18 applies, concluding that RFE_d^0, and hence RFE_l^k, hits p.

Suppose now that p has degree $d \geq n-k$. Set K to be the indices of the variables *not* appearing in some monomial with nonzero coefficient in p, and set $L \leftarrow \varnothing$. $\left(\frac{\partial p}{\partial L}\right)\big|_{K \leftarrow 0}$ is a single monomial, namely the product of the variables indexed by $[n] \setminus (K \cup L)$. Lemma 18 applies. Since none of the substitutions in (21) is zero, we conclude that RFE_0^{n-d}, and hence RFE_l^k, hits p.

The remaining case is that p has degree d with $l < d < n-k$. We start by writing p as a multilinear element of $\mathrm{Van}[\mathrm{RFE}_l^k]$ plus a structured remainder term, as formalized in the following claim. The claim can be shown similarly to Lemma 9; we include a proof below.

Claim 19. *Let $l < d < n-k$. Every homogeneous degree-d multilinear polynomial can be written as $p_0 + r$ where p_0 and r are degree-d homogeneous multilinear polynomials, $p_0 \in \mathrm{Van}[\mathrm{RFE}_l^k]$, and r is $(d+k-l, l)$-cored.*

Let p_0, r be the result of applying the claim to p. The contrapositive of Lemma 18 implies that $\left(\frac{\partial p_0}{\partial L}\right)\Big|_{K\leftarrow 0}$ vanishes at (20) for every pair of disjoint subsets $K, L \subseteq [n]$ of respective sizes k and l. Since $\left(\frac{\partial p}{\partial L}\right)\Big|_{K\leftarrow 0} = \left(\frac{\partial p_0}{\partial L}\right)\Big|_{K\leftarrow 0} + \left(\frac{\partial r}{\partial L}\right)\Big|_{K\leftarrow 0}$, it follows that the evaluation of $\left(\frac{\partial p}{\partial L}\right)\Big|_{K\leftarrow 0}$ at (20) coincides with that of $\left(\frac{\partial r}{\partial L}\right)\Big|_{K\leftarrow 0}$. If it so happens that r is zero, then $p = p_0$ belongs to $\mathrm{Van}[\mathrm{RFE}_l^k]$ and vanishes at (20) for every pair of disjoint subsets $K, L \subseteq [n]$ of respective sizes k and l. Otherwise, when r is nonzero, we claim that it is possible to choose sets K and L so that $\left(\frac{\partial r}{\partial L}\right)\Big|_{K\leftarrow 0}$ is a single monomial:

Claim 20. *Let $l < d < n-k$. Let r be a nonzero degree-d homogeneous multilinear polynomial that is $(d+k-l, l)$-cored. There are disjoint sets $K, L \subseteq [n]$ with $|K| = k$ and $|L| = l$ so that $\left(\frac{\partial r}{\partial L}\right)\Big|_{K\leftarrow 0}$ is a single monomial.*

Since (20) substitutes a nonzero value into each variable, $\left(\frac{\partial r}{\partial L}\right)\Big|_{K\leftarrow 0}$ takes a nonzero value for the claimed choice of K and L. As previously discussed, this implies that $\left(\frac{\partial p}{\partial L}\right)\Big|_{K\leftarrow 0}$ takes a nonzero value at (20). Likewise, going through Lemma 18, we see that r does not belong to $\mathrm{Van}[\mathrm{RFE}_l^k]$. Since $r = p - p_0$, the same goes for p. □

We complete the argument by proving Claims 19 and 20. Claim 19 is similar to Lemma 9, and is obtained using a variant of polynomial division suited to multilinear polynomials:

Proof of Claim 19. Let $C \subseteq [n]$ have size $d+k-l$. Every multilinear monomial m factors uniquely as $m_0 m_1$ where m_0 and m_1 are multilinear monomials such that m_0 depends only on variables indexed by C and m_1 depends only on variables not indexed by C. Call m_1 the *non-core* of m. We show the following:

Claim 21. *Every multilinear monomial with more than l variables in its non-core is equivalent, modulo a multilinear element of* $\mathrm{Van}[\mathrm{RFE}_l^k]$, *to a linear combination of multilinear monomials that all have non-cores of lower degree.*

This lets us prove Claim 19 as follows. Claim 21 implies that, for any multilinear polynomial p, we may, without changing p modulo multilinear elements of $\mathrm{Van}[\mathrm{RFE}_l^k]$, eliminate any monomial in p that violates the $(d+k-l, l)$-cored condition, while possibly introducing multilinear monomials with lower non-core degree. Thus we can systematically eliminate all monomials that violate the cored condition by eliminating them in order of decreasing non-core degree. After that, p is $(d+k-l, l)$-cored (with core the variables indexed by C), and Claim 19 follows.

We now show Claim 21. Factor $m = m_0 m_1$ as above, and suppose there are more than l variables in m_1. Let L index some $l+1$ of the variables in m_1, let m' be their product, and let m'' satisfy $m = m'm''$. There are at most $d-l-1$ variables in m_0; let K be any $k+1$ elements of C that index variables not in m_0. Combined, L and K have size exactly $k+l+2$. Consider $q = \mathrm{EVC}_l^k[L \cup K]$, where the variables in $L \cup K$ are ordered arbitrarily. By Proposition 7, m' appears as a monomial in q; moreover, every other monomial in q has lower non-core degree. It follows that every monomial in $m''\cdot q$ either is m, or else has lower non-core degree. Moreover, every such monomial is multilinear and is supported on variables indexed by $K \cup L$, which is disjoint from the support of m''. As q is in $\mathrm{Van}[\mathrm{RFE}_l^k]$, rearranging the equation $m''\cdot q \equiv 0 \pmod{\mathrm{Van}[\mathrm{RFE}_l^k]}$ to isolate m gives the desired equivalence. □

Claim 20 is similar to the proof of Lemma 11:

Proof of Claim 20. Let $C \subseteq [n]$ index the variables that form a core for r. Recall that $l < d < n - k$. By shrinking C if need be, we can assume that there is a multilinear monomial m with nonzero coefficient in r that involves exactly l variables not indexed by C. Let L be the variables appearing in m that are not indexed by C. Now extend C to have size $d + k - l$ while remaining disjoint from L. There are precisely k variables indexed by C that do not appear in m; let K be this set. Since r is multilinear, homogeneous of degree d, and $(d + k - l, l)$-cored with core C, there is exactly one monomial with nonzero coefficient in r that is divisible by $\prod_{i\in L} x_i$ and by no variable in K; it is precisely m. It follows that $\left(\frac{\partial r}{\partial L}\right)\Big|_{K\leftarrow 0}$ is a single monomial. □

We conclude this section by detailing the connection between Theorem 4 and some prior applications of the SV generator.

Application to read-once formulas. We start with the theorem that SV^1 hits read-once formulas. The original proof in [MV18] goes by induction on the depth of F, showing that $F(\mathrm{SV}^1)$ is nonconstant whenever F is nonconstant, or, equivalently, that SV^1 hits $F + c$ for every $c \in \mathbb{F}$ whenever F is nonconstant. The inductive step consists of two cases, depending on whether the top gate is a multiplication gate or an addition gate. The case of a multiplication gate follows from the general property that the product of a nonconstant polynomial with any nonzero polynomial is nonconstant. The case of an addition gate, say $F = F_1 + F_2$, involves a clever analysis that uses the variable-disjointness of F_1 and F_2 to show that $F_1(\mathrm{SV}^1)$ and $F_2(\mathrm{SV}^1)$ cannot cancel each other out.

The case of an addition gate $F = F_1 + F_2$ alternately follows from Theorem 4 with $k = 0$ and $l = 1$ and the following two observations, each corresponding to one of the conditions in Theorem 4. Both observations are immediate because of the variable-disjointness of F_1 and F_2:

1. If at least one of F_1 of F_2 has a homogeneous component of degree 1 or at least n, then so does F.

2. If for $L = \{i\} \subseteq [n]$ at least one of the derivatives $\frac{\partial F_1}{\partial x_i}$ or $\frac{\partial F_2}{\partial x_i}$ is nonzero at some point (6), then the same goes for $\frac{\partial F}{\partial x_i}$.

In particular, under the hypothesis that $F_1 + c$ is hit by RFE_1^0 for all $c \in \mathbb{F}$, F_1 must violate one of the conditions of Theorem 4 besides the one that requires F_1 have no constant term. Similarly for F_2. By the above observations, any such violation is inherited by F, and the inductive step follows.

In the overview, we mentioned that we originally proved Theorem 4 from a perspective that carries a geometric interpretation. The case of an addition gate in the above proof takes a particularly clean form in that perspective, which we sketch now.

Recall from the overview that we can think of the variables as vertices, and multilinear monomials simplices made from those vertices. A multilinear polynomial is a weighted collection of such simplices with weights from $\mathbb{F}$. In this view, Theorem 4 translates to the following characterization: a weighted collection of simplices corresponds to a polynomial in the vanishing ideal of RFE_1^0 if and only if there are no simplices of zero, one, or all vertices, and the remaining weights satisfy a certain system of linear equations. Crucially, for each equation in the system, there is a vertex such that the equation reads only weights of the simplices *that contain that vertex*. Meanwhile, the sum of two variable-disjoint polynomials corresponds to taking the vertex-disjoint union of two weighted collections of simplices. It follows directly that if either term in the sum violates a requirement besides the "no simplex of zero vertices" requirement, then the sum violates the same requirement.

Zero-substitutions and partial derivatives. As mentioned in the overview, several prior papers demonstrated the utility of partial derivatives and zero substitutions in the context of derandomizing PIT using the SV generator, especially for syntactically multilinear models. By judiciously choosing variables for those operations, these papers managed to simplify p and reduce PIT for p to PIT for simpler instances, resulting in an efficient recursive algorithm. Such recursive arguments can be naturally reformulated to use Theorem 4, according to the following prototype.

Let $\mathcal{C}$ be a family of multilinear polynomials, such as those computable with some bounded complexity in some syntactic model. For the argument, we break up $\mathcal{C} = \bigcup_{k,l} \mathcal{C}_{k,l}$ such that for every k, l and $p \in \mathcal{C}_{k,l}$, at least one of the following holds:

- $k = l = 0$ and p is either zero or hit by RFE_0^0.
- $k > 0$ and there is a zero substitution such that the result is in $\mathcal{C}_{k-1,l}$.
- $l > 0$ and there is a derivative such that the result is in $\mathcal{C}_{k,l-1}$.

We also make the mild assumption that each $\mathcal{C}_{k,l}$ is closed under rescaling variables. With these hypotheses in place, we establish the following claim through direct applications of Theorem 4:

Claim 22. *Under the above hypotheses,* RFE_l^k *hits* $\mathcal{C}_{k,l}$ *for every* k, l.

Proof. The proof is by induction on k and l. The base case is $k = l = 0$, where the claim is immediate. When $k > 0$ or $l > 0$, our hypotheses are such that p either simplifies under a zero substitution $x_{i^*} \leftarrow 0$ or a derivative $\frac{\partial}{\partial x_{i^*}}$. We analyze each case separately. By condition 1 of Theorem 4, we may assume that p only has homogeneous parts with degrees in the range $l+1, \ldots, n-k-1$.

- If p simplifies under a zero substitution $x_{i^*} \leftarrow 0$, then let $p' \in \mathcal{C}_{k-1,l}$ be the simplified polynomial where moreover the remaining variables have been rescaled according to $x_i \leftarrow x_i \cdot (a_{i^*} - a_i)$. That is, write p as $p = qx_{i^*} + r$ where q and r are polynomials that do not depend on x_{i^*}, and set $p'(\ldots, x_i, \ldots) \doteq r(\ldots, x_i \cdot (a_{i^*} - a_i), \ldots)$. By induction, p' is hit by RFE_l^{k-1}. We apply Theorem 4 to p' with respect to the set of variables $\{x_1, \ldots, x_{i^*-1}, x_{i^*+1}, \ldots, x_n\}$ and k replaced by $k-1$. As p only has homogeneous parts with degrees in the range $l+1, \ldots, n-k-1$, so does p', and condition 1 of Theorem 4 fails. By condition 2, there must be $z \in Z$ and disjoint $K, L \subseteq [n] \setminus \{i\}$ with $|K| = k-1$ and $|L| = l$ so that substituting (6) yields a nonzero value. It follows directly that, with respect to the same z, $K' = K \cup \{i\}$, and the same L, the substitution (6) yields a nonzero value when applied to p.
- If p simplifies under a partial derivative $\frac{\partial}{\partial x_{i^*}}$, then a similar analysis works. Set $p' \in \mathcal{C}_{k,l-1}$ to be the simplification with variables rescaled according to $x_i \leftarrow x_i/(a_{i^*} - a_i)$. That is, write p as $p = qx_{i^*} + r$ where q and r are polynomials that do not depend on x_{i^*}, and set $p'(\ldots, x_i, \ldots) \doteq q(\ldots, x_i/(a_{i^*} - a_i), \ldots)$. By induction, p' is hit by RFE_{l-1}^k. We apply Theorem 4 to p' with respect to the set of variables $\{x_1, \ldots, x_{i^*-1}, x_{i^*+1}, \ldots, x_n\}$ and l replaced by $l-1$. As p' has homogeneous parts of degrees one less than p does, condition 1 of Theorem 4 fails. By condition 2, there is $z \in Z$ and disjoint $K, L \subseteq [n] \setminus \{i\}$ with $|K| = k$ and $|L| = l-1$ so that substituting (6) yields a nonzero value. It follows directly that, with respect to the same z, the same K, and $L' = L \cup \{i^*\}$, the substitution (6) yields a nonzero value when applied to p. $\square$

Theorem 4 tells us that derivatives and zero substitutions suffice to witness when a multilinear polynomial p is hit by SV or RFE. One can ask, if we know more information about p, can we infer *which* derivatives and zero substitutions form a witness? In some cases we know. For example, if p has a low-support monomial $x_1 \cdots x_l$, then it suffices to take derivatives with respect to each of $x_1, \dots, x_l$. On the other hand, consider that whenever two polynomials p and q are hit by SV, then so is their product pq. Given explicit witnesses for p and q, we do not know how to obtain an explicit witness for the product pq.

5 Sparseness

By Proposition 7, the generators EVC_l^k contain exactly $\binom{k+l+2}{l+1}$ monomials. The following result shows that no nonzero polynomial in the vanishing ideal of RFE_l^k has fewer monomials.

Lemma 23. *Suppose $p \in \mathbb{F}[x_1, \dots, x_n]$ is nonzero and has only s monomials with nonzero coefficients. Then, for any k, l such that $\binom{k+l+2}{l+1} > s$, RFE_l^k hits p.*

The tactic here is to show that, if p has too few monomials appearing in it, then there is a way to instantiate the Zoom Lemma wherein p_{d^*} is a single monomial and therefore does not vanish at (13).

Proof. For $i \in [n]$, we define two operations, $\downarrow_i$ and $\uparrow_i$, on nonempty sets of monomials. Given such a set M, $\downarrow_i(M)$ is the subset of M consisting of the monomials in which x_i appears with its least degree among all the monomials in M. We define $\uparrow_i$ similarly, except we select the monomials in which x_i appears with its highest degree. We make the following claim:

Claim 24. *For any nonempty set of monomials with fewer than $\binom{k+l+2}{l+1}$ monomials, there is a sequence of $\downarrow$ and $\uparrow$ operations, with at most k $\downarrow$ operations and at most l $\uparrow$ operations, such that the resulting set of monomials has exactly one element.*

The claim implies the lemma as follows. Let M be the set of monomials with nonzero coefficient in p. Apply the claim to M to get a sequence of $\downarrow$ and $\uparrow$ operations resulting in a single monomial m_0. Let K denote the indices used for the $\downarrow$ operations and L the indices used for the $\uparrow$ operations. Let d^* be the degree pattern with domain $K \cup L$ that matches m_0. By how the operators are defined, every monomial m in M satisfies either

- $\deg_{x_i}(m) > d_i^*$ for some $i \in K$ (m was removed by $\downarrow_i$),
- $\deg_{x_i}(m) < d_i^*$ for some $i \in L$ (m was removed by $\uparrow_i$), or
- $\deg_{x_i}(m) = d_i^*$ for every $i \in K \cup L$, in which case $m = m_0$.

Accordingly, d^* is (K, L)-extremal in p and the Zoom Lemma applies. As p_{d^*} is a single monomial, it does not vanish at (13). We conclude that p is hit by $\mathrm{RFE}_{|L|}^{|K|}$, and therefore by RFE_l^k.

It remains to prove Claim 24. We do this by induction on $|M|$. In the base case, $|M| = 1$, in which case the empty sequence suffices. Otherwise, $|M| > 1$, in which case there is a variable x_i that appears with at least two distinct degrees among monomials in M. The sets $\downarrow_i(M)$ and $\uparrow_i(M)$ are nonempty and disjoint. Since M has size less than $\binom{k+l+2}{l+1} = \binom{k+l+1}{l+1} + \binom{k+l+1}{l}$, either $\downarrow_i(M)$ has size less than $\binom{k+l+1}{l+1}$, or $\uparrow_i(M)$ has size less than $\binom{k+l+1}{l}$. Whichever is the case, the claim follows by applying the inductive hypothesis to it. ☐

6 Set-Multilinearity

Although the generators EVC_l^k provided by Theorem 2 are not set-multilinear, the vanishing ideal of RFE_l^k does contain set-multilinear polynomials. In this section, we construct some of degree $l+1$ with partition classes of size $k+2$. In fact, we argue that all set-multilinear polynomials in $\mathrm{Van}[\mathrm{RFE}_l^k]$ of degree $l+1$ are in the linear span of the ones we construct; we conclude that no polynomial of degree $l+1$ in $\mathrm{Van}[\mathrm{RFE}_l^k]$ can be set-multilinear with partitions of size less than $k+2$.

Our construction is a modification of the one for EVC_l^k.

Definition 25. *Let $k, l \in \mathbb{N}$ be parameters, and let $X_1, \ldots, X_{l+1} \subseteq [(l+1)(k+2)]$ be $l+1$ disjoint subsets of $k+2$ variables each. The polynomial* ESMVC_l^k *is an $(l+1)\times(l+1)$ determinant where each entry is itself a $(k+2)\times(k+2)$ determinant. We index the rows in the outer determinant by $i = 1, \ldots, l+1$, and the columns by $d = l, \ldots, 0$. In each (i,d)-th inner matrix, there is one row per $j \in X_i$; it is*

$$\begin{bmatrix} a_j^d x_j & a_j^k & a_j^{k-1} & \cdots & a_j^1 & a_j^0 \end{bmatrix}.$$

The name "ESMVC" is a shorthand for "Elementary Set-Multilinear Vandermonde Circulation".

Similar to EVC, the precise instantiation of ESMVC requires one to pick an order for the sets $X_1, \ldots, X_{l+1}$ (up to even permutations) and an order within each set (again up to even permutations). Changing any of those orders by an odd permutation causes the sign of ESMVC to flip, but otherwise it is unchanged.

Example 26. When $k = 1$ and $l = 2$, ESMVC uses three sets of three variables each. To help convey the structure of the determinant, we name the variable-sets $\{x_1, x_2, x_3\}$, $\{y_1, y_2, y_3\}$, and $\{z_1, z_2, z_3\}$, and denote the abscissa of x_i by a_i, the abscissa of y_i by b_i, and the abscissa of z_i by c_i. With this notation, ESMVC is the following:

$$\begin{vmatrix}
\begin{vmatrix} a_1^2x_1 & a_1^1 & a_1^0 \\ a_2^2x_2 & a_2^1 & a_2^0 \\ a_3^2x_3 & a_3^1 & a_3^0 \end{vmatrix} & \begin{vmatrix} a_1^1x_1 & a_1^1 & a_1^0 \\ a_2^1x_2 & a_2^1 & a_2^0 \\ a_3^1x_3 & a_3^1 & a_3^0 \end{vmatrix} & \begin{vmatrix} a_1^0x_1 & a_1^1 & a_1^0 \\ a_2^0x_2 & a_2^1 & a_2^0 \\ a_3^0x_3 & a_3^1 & a_3^0 \end{vmatrix} \\
\begin{vmatrix} b_1^2y_1 & b_1^1 & b_1^0 \\ b_2^2y_2 & b_2^1 & b_2^0 \\ b_3^2y_3 & b_3^1 & b_3^0 \end{vmatrix} & \begin{vmatrix} b_1^1y_1 & b_1^1 & b_1^0 \\ b_2^1y_2 & b_2^1 & b_2^0 \\ b_3^1y_3 & b_3^1 & b_3^0 \end{vmatrix} & \begin{vmatrix} b_1^0y_1 & b_1^1 & b_1^0 \\ b_2^0y_2 & b_2^1 & b_2^0 \\ b_3^0y_3 & b_3^1 & b_3^0 \end{vmatrix} \\
\begin{vmatrix} c_1^2z_1 & c_1^1 & c_1^0 \\ c_2^2z_2 & c_2^1 & c_2^0 \\ c_3^2z_3 & c_3^1 & c_3^0 \end{vmatrix} & \begin{vmatrix} c_1^1z_1 & c_1^1 & c_1^0 \\ c_2^1z_2 & c_2^1 & c_2^0 \\ c_3^1z_3 & c_3^1 & c_3^0 \end{vmatrix} & \begin{vmatrix} c_1^0z_1 & c_1^1 & c_1^0 \\ c_2^0z_2 & c_2^1 & c_2^0 \\ c_3^0z_3 & c_3^1 & c_3^0 \end{vmatrix}
\end{vmatrix}.$$

◀

Proposition 27. *For any $k, l \geq 0$ and variable-index sets $X_1, \ldots, X_{l+1}$ as in Definition 25,* ESMVC *is nonzero, homogeneous of degree $l+1$, and set-multilinear with respect to the partition $X_1 \sqcup \cdots \sqcup X_{l+1}$. Moreover, every monomial consistent with that appears with a nonzero coefficient.* ESMVC *is skew-symmetric with respect to the order of the sets $X_1, \ldots, X_{l+1}$, and the choice of order within each set, in that any permutation thereof changes the construction by merely multiplying by the sign of the permutation. When the sets are ordered as $X_1, \ldots, X_{l+1}$ and their members are ordered as*

$X_i = \{x_{i,1}, \ldots, x_{i,k+2}\}$ *for* $i = 1, \ldots, l+1$*, the coefficient of* $x_{1,1} \cdots x_{l+1,1}$ *is the product of Vandermonde determinants*

$$\begin{vmatrix} a_{1,1}^l & \cdots & a_{1,1}^0 \\ \vdots & \ddots & \vdots \\ a_{l+1,1}^l & \cdots & a_{l+1,1}^0 \end{vmatrix} \cdot \prod_{i=1}^{l+1} \begin{vmatrix} a_{i,2}^k & \cdots & a_{i,2}^0 \\ \vdots & \ddots & \vdots \\ a_{i,k+2}^k & \cdots & a_{i,k+2}^0 \end{vmatrix}.$$

Proof. All assertions to be proved follow from elementary properties of determinants, that Vandermonde determinants are nonzero unless they have duplicate rows, and the following computation: the result of plugging 1 into $x_{i,1}$ for $i = 1, \ldots, l+1$ and 0 into the remaining variables is the product of Vandermonde matrices in the statement. □

The following theorem formalizes the role ESMVC plays among the degree-$(l+1)$ elements of $\mathrm{Van}[\mathrm{RFE}_l^k]$.

Theorem 28. *Let* $k, l \in \mathbb{N}$ *be parameters, and let* $X_1, \ldots, X_{l+1}$ *be* $l+1$ *disjoint sets of variable indices (of any size). Let* $\mathrm{ESMVC}_l^k(X_1, \ldots, X_{l+1})$ *be the collection of polynomials formed by picking a* $(k+2)$*-subset of each of* $X_1, \ldots, X_{l+1}$ *and instantiating* ESMVC_l^k *with respect to those sets. The linear span of* $\mathrm{ESMVC}_l^k(X_1, \ldots, X_{l+1})$ *equals the set-multilinear polynomials in* $\mathrm{Van}[\mathrm{RFE}_l^k]$ *with variable partition* $X_1 \sqcup \cdots \sqcup X_{l+1}$*.*

Theorem 28 immediately implies that there are no set-multilinear polynomials of degree $l+1$ in $\mathrm{Van}[\mathrm{RFE}_l^k]$ that have at least one partition X_i of size less than $k+2$.

Proving Theorem 28 involves two steps, similar to Theorem 2:

1. Show that any instantiation of ESMVC_l^k is in $\mathrm{Van}[\mathrm{RFE}_l^k]$.
2. Show that, modulo instantiations of ESMVC_l^k, every set-multilinear polynomial with variable partition $X_1, \ldots, X_{l+1}$ takes a particular form such that RFE_l^k hits every nonzero polynomial of that form.

Step 1 is the following claim:

Claim 29. *For every* $k, l \in \mathbb{N}$*, and every choice of* $l+1$ *disjoint sets* $X_1, \ldots, X_{l+1}$ *of* $k+2$ *variable-indices each,* ESMVC_l^k *vanishes at* RFE_l^k*.*

Proof. Let g/h be a seed for RFE_l^k. Let A be the $(l+1) \times (l+1)$ outer matrix defining ESMVC, so that $\mathrm{ESMVC} \doteq \det(A)$. Recall that the columns of A are indexed by $d = l, \ldots, 0$. Let $h \in \mathbb{F}^{l+1}$ be the column vector where the row indexed by d is the coefficient of α^d in $h(\alpha)$. We show that, after substituting $\mathrm{RFE}_l^k(g/h)$, the matrix-vector product $Ah \in \mathbb{F}^{l+1}$ yields the zero vector. It follows that evaluating ESMVC at $\mathrm{RFE}_l^k(g/h)$ vanishes, as it is the determinant of a singular matrix.

Fix $i \in \{1, \ldots, l+1\}$, and focus on the i-th coordinate of Ah. The (i,d) entry of A is a determinant; let $B_{i,d}$ be the inner matrix as in Definition 25. As d varies, only the first column of $B_{i,d}$ changes. Thus, by multilinearity of the determinant, the i-th entry of Ah is itself a determinant. Recalling that the rows of $B_{i,l}, \ldots, B_{i,0}$ are indexed by $j \in X_i$, the j-th row of this determinant is

$$\begin{bmatrix} h(a_j)x_j & a_j^k & \cdots & a_j^0 \end{bmatrix}.$$

After substituting $\mathrm{RFE}_l^k(g/h)$, it becomes

$$\begin{bmatrix} g(a_j) & a_j^k & \cdots & a_j^0 \end{bmatrix}.$$

Since g is a degree-k polynomial, the columns are linearly dependent, so the determinant is zero. □

For step 2, we need a suitable replacement for being (c,t)-cored. The following adaptation of that to the set-multilinear setting suffices.

Definition 30. *Let $X_1 \sqcup \cdots \sqcup X_d \subseteq [n]$ be disjoint sets of variable-indices. For parameters $k, l \geq 0$, a polynomial p that is set-multilinear with respect to the sets $X_1, \ldots, X_d$ is* (c,t)-multi-cored *if there are subsets $C_i \subseteq X_i$ for $i = 1, \ldots, d$, each of size at most c, such that every monomial of p involves at most l variables not indexed by $C_1 \cup \cdots \cup C_d$.*

Claim 31. *Let $k, l \geq 0$ be parameters, and let $X_1 \sqcup \cdots \sqcup X_{l+1} \subseteq [n]$ be disjoint sets of variable-indices. Let $\mathrm{ESMVC}_l^k(X_1, \ldots, X_{l+1})$ be the collection of polynomials formed by picking a $(k+2)$-subset of each of $X_1, \ldots, X_{l+1}$ and instantiating ESMVC_l^k with respect to those sets. Every set-multilinear polynomial with variable partition $X_1, \ldots, X_{l+1}$ equals a $(k+1, l)$-multi-cored polynomial modulo the linear span of $\mathrm{ESMVC}_l^k(X_1, \ldots, X_{l+1})$.*

Claim 31 follows from a monomial elimination argument as in Lemma 9. A formal proof is omitted. From there, Theorem 28 follows from the following claim:

Claim 32. *Let $k, l \geq 0$ be parameters, and let $X_1 \sqcup \cdots \sqcup X_{l+1} \subseteq [n]$ be disjoint sets of variable-indices. Every degree-$(l+1)$ polynomial that is set-multilinear with respect to the partition $X_1, \ldots, X_{l+1}$ and that is $(k+1, l)$-multi-cored is hit by RFE_l^k.*

Proof. Let p satisfy the hypotheses of the claim, and let $C_1, \ldots, C_{l+1}$ be the sets witnessing the $(k+1, l)$-multi-core structure on p, and let $C \doteq C_1 \cup \cdots \cup C_{l+1}$ be their union. Every monomial m factors as $m_0 m_1$ with m_0 supported on variables indexed by C and m_1 supported on variables not indexed by C. Call m_1 the *non-core* of m. Fix $m = m_0 m_1$ to be a monomial in p so that the non-core is maximal under divisibility. The multi-core structure on p implies the non-core has at most $l < l+1$ variables, so m_0 contains at least one variable, x_j. Let i such that $x_j \in X_i$. Let L index the variables in m except for x_j, let $K \doteq C_i \setminus \{j\}$, and set $d^* \in \mathbb{N}^{K \cup L}$ to be the degree pattern with domain $K \cup L$ that matches m. The set-multilinear structure implies that p_{d^*} uses only variables indexed by X_i. By how we chose m and the multi-core structure, it must use variables indexed by C_i. By how we chose K, it must use only x_j. It follows that p_{d^*} is just a nonzero scalar times x_j. The Zoom Lemma applies, and we conclude that $p \notin \mathrm{Van}[\mathrm{RFE}_l^k]$. ☐

7 Read-Once Oblivious Algebraic Branching Programs

In this section we provide some background on ROABPs and establish Theorem 5.

7.1 Background

Algebraic branching programs are a syntactic model for algebraic computation. One forms a directed graph with a designated source and sink. Each edge is labeled by a polynomial that depends on at most one variable among $x_1, \ldots, x_n$. The branching program computes a polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ by summing, over all source-to-sink paths, the product of the labels on the edges of each path.

A special subclass of algebraic branching programs are read-once oblivious algebraic branching programs (ROABPs). In this model, the vertices of the branching program are organized in layers. The layers are totally ordered, and edges exist only from one layer to the next. For each variable, there is at most one consecutive pair of layers between which that variable appears, and for each pair of consecutive layers, there is at most one variable that appears between them. In this way, every source-to-sink path reads each variable at most once (the branching program is *read-once*), and the order in which the variables are read is common to all paths (the branching program is *oblivious*). We can always assume that the number of layers equals one plus the number of variables under consideration.

The number of vertices comprising a layer is called its *width*. The width of an ROABP is the largest width of its layers. The minimum width of an ROABP computing a given polynomial can be characterized in terms of the rank of coefficient matrices constructed as follows.

Definition 33. *Let $U \sqcup V = [n]$ be a partition of the variable indices, and let M_U and M_V be the sets of monomials that are supported on variables indexed by U and V, respectively. For any polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ define the matrix*

$$\mathrm{CMat}_{U,V}(p) \in \mathbb{F}^{M_U \times M_V}$$

by setting the (m_U, m_V) entry to equal the coefficient of $m_U m_V$ in p.

$\mathrm{CMat}_{U,V}(p)$ is formally an infinite matrix, but it has only finitely many nonzero entries. When p has degree at most d, one can just as well truncate $\mathrm{CMat}_{U,V}(p)$ to include only rows and columns indexed by monomials of degree at most d.

Lemma 34 ([Nis91]). *Let $p \in \mathbb{F}[x_1, \ldots, x_n]$ be any polynomial. There is an ROABP of width w computing p in the variable order $x_1, \ldots, x_n$ if and only if, for every $s \in \{0, \ldots, n\}$, with respect to the partition $U = \{1, \ldots, s\}$ and $V = \{s+1, \ldots, n\}$, we have*

$$\mathrm{rank}(\mathrm{CMat}_{U,V}(p)) \leq w.$$

Lemma 34 applies to other variable orders by renaming the variables.

We group the monomials in M_U and M_V by their degrees, and order the groups by increasing degree. This induces a block structure on $\mathrm{CMat}_{U,V}(p)$ with one block for every choice of $r, c \in \mathbb{N}$; the (r, c) block is the submatrix with rows indexed by degree-r monomials in M_U and columns indexed by degree-c monomials in M_V. In the case where p is homogeneous, the only nonzero blocks occur for $r + c$ equal to the degree of p. In this case the rank of $\mathrm{CMat}_{U,V}(p)$ is the sum of the ranks of its blocks.

In general, the rank of $\mathrm{CMat}_{U,V}(p)$ is at least the rank of $\mathrm{CMat}_{U,V}(p_\downarrow)$, where $p_\downarrow$ denotes the homogeneous part of p of the lowest degree, $d_\downarrow$. This follows because the submatrix of $\mathrm{CMat}_{U,V}(p)$ consisting of the rows and columns indexed by monomials of degree at most $d_\downarrow$ has a block structure that is triangular with the blocks of $\mathrm{CMat}_{U,V}(p_\downarrow)$ on the hypotenuse. The observation yields the following folklore consequence of Lemma 34.

Proposition 35. *Let $p \in \mathbb{F}[x_1, \ldots, x_n]$ be any nonzero polynomial, and let $p_\downarrow$ be the nonzero homogeneous part of p of least degree. If p can be computed by an ROABP of width w, then so can $p_\downarrow$.*

7.2 Hitting property / lower bound

We now prove Theorem 5—that SV^l, or equivalently RFE_l^{l-1}, hits every polynomial computed by an ROABP of width less than $1+(l/3)$ that contains a monomial of degree at most $l+1$. The novelty lies in the special case where p is homogeneous of degree $l+1$ and multilinear. In the following statement we did not try to optimize the dependence of the width bound on l.

Theorem 36. *Let $l \geq 1$ be an integer. For any nonzero, multilinear, homogeneous polynomial p of degree $l+1$, if p is computable by an ROABP of width less than $(l/3)+1$, then RFE_l^{l-1} hits p.*

Theorem 5 follows from Theorem 36 in a standard way. We provide a proof for completeness.

Proof of Theorem 5 from Theorem 36. Fix p satisfying the hypotheses of Theorem 5. We show that RFE_l^{l-1} hits p; this implies SV^l hits p because RFE_l^{l-1} and SV^l are equivalent up to variable rescaling, and rescaling variables does not affect ROABP width. Let $p_\downarrow$ be the nonzero homogeneous part of p of least degree. We show that RFE_l^{l-1} hits $p_\downarrow$; this implies RFE_l^{l-1} hits p by Proposition 17.

Suppose first that $p_\downarrow$ contains a monomial m^* depending on at most l variables. It is well-known that SV^l hits any such polynomial. Here is an argument based on the Zoom Lemma. Set $K = \varnothing$, let L index the variables appearing in m^*, and set $d^* \in \mathbb{N}^L$ to be the degree pattern with domain L that matches m^*. The homogeneity of $p_\downarrow$ ensures that d^* is (K,L)-extremal in $p_\downarrow$. The coefficient p_{d^*} is a nonzero constant. By the Zoom Lemma, RFE_l^{l-1} hits $p_\downarrow$.

Since $\deg(p_\downarrow) \le \deg(p) \le l+1$, the remaining possibility is that $p_\downarrow$ is multilinear of degree exactly $l+1$. By Proposition 35, $p_\downarrow$ is computable by an ROABP of width less than $1+(l/3)$. That RFE_l^{l-1} hits $p_\downarrow$ then follows from Theorem 36. ☐

In the remainder of this section we establish Theorem 36. Fix a positive integer l, and fix an arbitrary variable order, say $x_1,\dots,x_n$. We show that, for every polynomial p that is nonzero, multilinear, homogeneous of degree $l+1$, and belongs to the vanishing ideal of RFE_l^{l-1}, there exists some $s \in \{0,\dots,n\}$ so that, with respect to the partition $U = \{1,\dots,s\}$, $V = \{s+1,\dots,n\}$, it holds that $\mathrm{rank}(\mathrm{CMat}_{U,V}(p)) \ge (l/3)+1$. Theorem 36 then follows by Lemma 34.

Let $C \doteq \mathrm{CMat}_{U,V}(p)$. As p is homogeneous of degree $l+1$, C is block diagonal, with a block C_d for each $d \in \{0,\dots,l+1\}$ consisting of the rows indexed by monomials of degree d and the columns indexed by monomials of degree $l+1-d$. The block diagonal structure implies $\mathrm{rank}(C) = \sum_{d=0}^{l+1} \mathrm{rank}(C_d)$.

The hypothesis that p belongs to $\mathrm{Van}[\mathrm{RFE}_l^{l-1}]$ induces linear equations on the entries in the blocks C_d. In particular, condition 2 of Theorem 16 stipulates that, for all disjoint subsets $K, L \subseteq [n]$ with $|K| = k = l-1$ and $|L| = l$, $\left(\frac{\partial p}{\partial L}\right)\Big|_{K\leftarrow 0}$ vanishes at the point (20). This condition is linear in the coefficients of $\left(\frac{\partial p}{\partial L}\right)\Big|_{K\leftarrow 0}$, which are entries in the blocks C_d of C. In fact, each of these equations only reads entries from two *adjacent* blocks, i.e., blocks C_d and $C_{d'}$ with $|d-d'| = 1$. This is because L has size l, one less than the degree of p, so the only monomials that contribute to $\left(\frac{\partial p}{\partial L}\right)\Big|_{K\leftarrow 0}$ are those that are one variable x_i times the product of the variables indexed by L. It follows that the corresponding linear equation on C reads only entries that reside in the blocks $C_{|L\cap U|+1}$ (for $i \in U$) and $C_{|L\cap U|}$ (for $i \in V$).

We exploit the structure of these equations and argue that, for an appropriate choice of the partition index s, $\mathrm{rank}(C)$ is high. Our argument looks at the subset of the discrete interval $\{0,\dots,l+1\}$ consisting of those d for which C_d is nonzero.

Ingredients. Our analysis has four ingredients. The first ingredient is the fact that $\mathrm{rank}(C)$ is at least the number of nonzero blocks C_d. This is because a nonzero block has rank at least 1, and $\mathrm{rank}(C)$ is the sum of the ranks of the blocks. This simple observation means we can focus on situations where relatively few of the blocks are nonzero.

The second ingredient establishes an alternative lower bound on $\mathrm{rank}(C)$ in terms of the minimum distance between a nonzero block C_d and either extreme ($d = 0$ or $d = l+1$). Another way to think about this distance is as the maximum t such that every monomial in p depends on at least t variables indexed by U and at least t variables indexed by V.

Lemma 37. *Let $p \in \mathrm{Van}[\mathrm{RFE}_l^{l-1}]$ be nonzero, multilinear, and homogeneous of degree $l+1$, let $U \sqcup V$ be a partition of $[n]$, and let $C \doteq \mathrm{CMat}_{U,V}(p)$. If every monomial in p depends on at least t variables indexed by U and at least t variables indexed by V, then* $\mathrm{rank}(C) \ge t+1$.

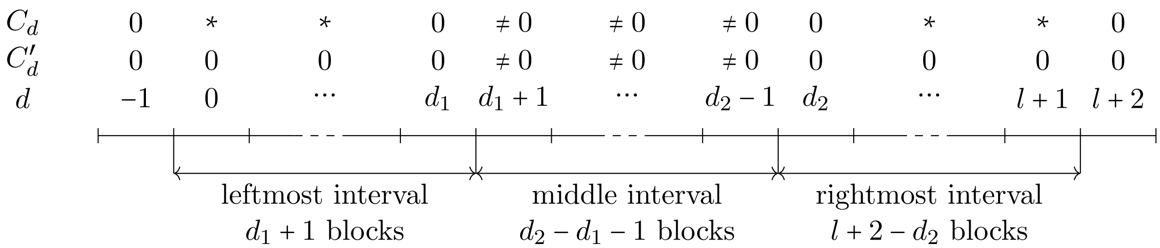


Figure 1: Rank lower bound analysis in terms of the blocks C_d of p and C'_d of p' (Proposition 38)

The proof involves revisiting the equations from condition 2 of Theorem 16 and modifying the underlying instantiations of the Zoom Lemma to obtain a system of linear equations with a simple enough structure that we can analyze.[2]

The remaining ingredients allow us to reduce to situations where either the first or second ingredient applies. The third ingredient lets us fix any two zero blocks and zero out all the blocks that are not between them.

Proposition 38. *Let* $p \in \mathrm{Van}[\mathrm{RFE}_l^{l-1}]$ *be multilinear and homogeneous of degree* $l+1$*. Let* $U \sqcup V$ *be a partition of* $[n]$*, and let* $C \doteq \mathrm{CMat}_{U,V}(p)$*. Suppose that for some* $d_1, d_2 \in \{-1, \ldots, l+2\}$ *with* $d_1 \le d_2$*, we have* $C_{d_1} = 0$ *and* $C_{d_2} = 0$*, where* $C_{-1} \doteq 0$ *and* $C_{l+2} \doteq 0$*. Let* p' *be the polynomial obtained from* p *by zeroing out the blocks* C_d *with* $d < d_1$ *or* $d > d_2$*. Then* p' *belongs to* $\mathrm{Van}[\mathrm{RFE}_l^{l-1}]$*.*

As zeroing out blocks does not increase the rank of C, our lower bound for $\mathrm{rank}(C)$ reduces to the same lower bound for the rank of $\mathrm{CMat}_{U,V}(p')$. This effectively extends the scope of the second ingredient: Alone, the second ingredient requires that *all* nonzero blocks of C be far from the extremes; with the third ingredient, it suffices that there exists a subinterval of nonzero blocks that is surrounded by zero blocks and that is far from the extremes. The proof hinges on the adjacent-block property of the equations from condition 2 of Theorem 16.

The ingredients thus far suffice provided there exists a nonzero block far from the extremes: Such a block belongs to some subinterval of nonzero blocks that is surrounded by zero blocks, say C_{d_1} to the left and C_{d_2} to the right, and the subinterval either is large and therefore has many nonzero blocks such that the first ingredient applies, or else it is small and therefore stays far from the extremes such that the combination of the second and third ingredients applies. See Figure 1 for an illustration. The fourth and final ingredient lets us ensure there is a nonzero block far from the extremes by setting the partition index s appropriately. In fact, it lets us guarantee a zero-to-nonzero transition at a position of our choosing.

Proposition 39. *For every* $d \in \{-1, \ldots, l\}$*, there is* $s \in \{0, \ldots, n\}$ *such that* $C_d = 0$ *and* $C_{d+1} \neq 0$ *with respect to the partition* $U = \{1, \ldots, s\}$*,* $V = \{s+1, \ldots, n\}$*, where* $C_{-1} \doteq 0$*.*

Combining ingredients. Let us find out what lower bound on $\mathrm{rank}(C)$ the prior ingredients give us as a function of the position $d = d_1$ in the interval where we have a guaranteed zero-to-nonzero transition as in Proposition 39. Starting from position d_1, keep increasing the position index until we hit the next zero block, say at position d_2, where we use $C_{l+2} \doteq 0$ as a sentinel. See Figure 1.

1. By the first ingredient, since the middle interval consists of nonzero blocks only, $\mathrm{rank}(C) \ge d_2 - d_1 - 1$.

[2]This is the setting where we exploit the possibility of the sets K and L in the Zoom Lemma to overlap.

2. By the combination of the second and the third ingredient, we have that $\mathrm{rank}(C) \geq t+1$ where $t = \min(d_1+1, l+2-d_2)$ is the minimum length of the leftmost and rightmost intervals. Indeed, let p' be the polynomial obtained from p by zeroing out the blocks C_d with $d < d_1$ or $d > d_2$. By Proposition 38 $p' \in \mathrm{Van}[\mathrm{RFE}_l^{l-1}]$. The polynomial p' is nonzero as it contains the original block C_{d+1}, which is nonzero. It is homogeneous of degree $l+1$ and multilinear as all of its monomials also occur in the homogeneous multilinear polynomial p of degree $l+1$. By construction, every monomial in p' contains at least d_1+1 variables indexed by U, and at least $l+2-d_2$ variables indexed by V. As such, p' satisfies the conditions of Lemma 37 with $t = \min(d_1+1, l+2-d_2)$. It follows that $\mathrm{rank}(C) \geq \mathrm{rank}(\mathrm{CMat}_{U,V}(p')) \geq t+1$.

If the rightmost interval has length at least the leftmost interval ($l+2-d_2 \geq d_1+1$), then item 2 yields $\mathrm{rank}(C) \geq d_1+2$. Otherwise, the rightmost interval is strictly shorter than the leftmost interval ($d_1+1 > l+2-d_2$); this implies that the middle interval has length at least $l-2d_1+1$, which by item 1 yields $\mathrm{rank}(C) \geq l-2d_1+1$. In any case, the bound $\mathrm{rank}(C) \geq \min(d_1+2, l-2d_1+1)$ holds. Taking $d_1 = \lfloor \frac{l-1}{3} \rfloor$ optimizes this expression, achieving $\mathrm{rank}(C) \geq \lfloor \frac{l-1}{3} \rfloor + 2 \geq (l/3)+1$. This completes the proof of Theorem 36 modulo the proofs of ingredients two through four.

Proofs. We conclude by proving ingredients two through four. We start with the one that requires the least specificity (ingredient 4, Proposition 39), then do ingredient 3 (Proposition 38), and end with the one that involves the most structure (ingredient 2, Lemma 37).

Proof of Proposition 39. When $s=0$, C_0 contains all entries. As s increases by 1, some entries move from their current block $C_{d'}$ to the next block $C_{d'+1}$. Finally, when $s=n$, C_{l+1} contains all entries. It follows that every nonzero entry moves from C_d to C_{d+1} at some time. If we stop increasing s right after the last nonzero entry of C moves out of C_d, we have $C_d = 0$ and $C_{d+1} \neq 0$. $\square$

Proof of Proposition 38. It suffices to show that whenever p satisfies the two conditions in Theorem 16, then so does p'. Condition 1 holds for p' as p' either is zero or else has the same degree as p. Regarding condition 2, as mentioned, the condition is equivalent to a system of homogeneous linear equations on $C' \doteq \mathrm{CMat}_{U,V}(p')$, each involving only an adjacent pair of blocks in C'. Those that involve only blocks C'_d with $d \leq d_1$ are met as the equations are homogeneous and the involved blocks are all zero. The same holds for the equations that involve only blocks C'_d with $d \geq d_2$. The remaining equations involve only blocks C'_d with $d \in \{d_1, \ldots, d_2\}$, on which p and p' agree. As the equations hold for C, they also hold for C'. $\square$

It remains to argue Lemma 37. Our proof makes use of linear equations that are closely related to those given by Theorem 16, which in turn come from the Zoom Lemma. We revisit the application of the Zoom Lemma so as to obtain a simpler coefficient matrix—ultimately a Cauchy matrix—that enables a deeper analysis. To facilitate the discussion, we utilize the following notation. As p is multilinear, we only need to consider rows indexed by monomials of the form $\prod_{i\in I} x_i$ for $I \subseteq U$ and columns indexed by monomials of the form $\prod_{j\in J} x_j$ for $J \subseteq V$. This allows us to index rows by subsets $I \subseteq U$ and columns by subsets $J \subseteq V$. For $I \subseteq U$ and $J \subseteq V$ we denote by $C(I,J)$ the corresponding entry of C. The following proposition describes the linear equations we use.

Proposition 40. *Let $p \in \mathrm{Van}[\mathrm{RFE}_l^{l-1}]$ be multilinear, and homogeneous of degree $l+1$, let $U \sqcup V$ be a partition of $[n]$, and let $C \doteq \mathrm{CMat}_{U,V}(p)$. For every $I \subseteq U$ and $J \subseteq V$ with $|I|+|J| = l$, and for every $i^* \in I \cup J$,*

$$\sum_{i\in U\setminus I} \frac{C(\{i\}\cup I, J)}{a_i - a_{i^*}} + \sum_{i\in V\setminus J} \frac{C(I, \{i\}\cup J)}{a_i - a_{i^*}} = 0. \tag{22}$$

Proof. Set $L \doteq I \cup J$ and $K \doteq L \smallsetminus \{i^*\}$, and note that $K \subseteq L$. Let $d^* \in \mathbb{N}^L$ be the all-1 degree pattern with domain L, and let $m^* \doteq \prod_{i \in L} x_i$ be the monomial supported on L that matches d^*. As p is multilinear, it follows that d^* is (K, L)-extremal in p. Since p is in $\mathrm{Van}[\mathrm{RFE}_l^{l-1}]$, the contrapositive of the Zoom Lemma tells us that the coefficient p_{d^*} of p vanishes at the point (13) for each $z \in \mathbb{F}$. We take $z = 1$.

The multilinear monomials m of degree $l+1$ that match d^* have the form $m = x_i \cdot m^*$, where $i \in [n] \smallsetminus L$. Thus, we can write the coefficient p_{d^*} as

$$p_{d^*} = \sum_{i \in U \smallsetminus I} C(\{i\} \cup I, J) \cdot x_i + \sum_{i \in V \smallsetminus J} C(I, \{i\} \cup J) \cdot x_i. \tag{23}$$

For each $i \in [n] \smallsetminus L$, (13) with $z = 1$ substitutes $1/(a_i - a_{i^*})$ into x_i. Plugging this into (23) yields (22). $\square$

Proof of Lemma 37. The proof goes by induction on t. The base case is $t = 0$, where the lemma holds because the rank of a nonzero matrix is always at least 1. For the inductive step, where $t \geq 1$, we zoom in on the contributions of the monomials that contain a particular variable. More precisely, for $i^* \in [n]$, let p_{i^*} denote the unique polynomial such that $p = p_{i^*} x_{i^*} + r$ for some polynomial r that does not depend on x_{i^*}. (In our terminology of degree patterns, the polynomial p_{i^*} is the coefficient of p corresponding to the degree pattern d with domain $\{i^*\}$ and $d_{i^*} = 1$.) Consider any $i^* \in [n]$ such that p_{i^*} is nonzero. As p is multilinear and homogeneous of degree $l+1$, p_{i^*} is multilinear and homogeneous of degree l. As every monomial in p depends on at least t variables indexed by U and at least t variables indexed by V, every monomial in p_{i^*} depends on at least $t-1$ variables indexed by U and at least $t-1$ variables indexed by V. In a moment, we argue that for every $i^* \in [n]$, $p_{i^*} \in \mathrm{Van}[\mathrm{RFE}_{l-1}^{l-2}]$. Then we will show the following:

Claim 41. *There exists $i^* \in [n]$ such that $p_{i^*} \neq 0$ and*

$$\mathrm{rank}(\mathrm{CMat}_{U,V}(p)) \ \geq \ \mathrm{rank}(\mathrm{CMat}_{U,V}(p_{i^*})) + 1. \tag{24}$$

Given an i^* as in Claim 41, we conclude by induction that

$$\mathrm{rank}(\mathrm{CMat}_{U,V}(p)) \ \geq \ \mathrm{rank}(\mathrm{CMat}_{U,V}(p_{i^*})) + 1 \ \geq \ (t-1) + 1 + 1 \ = \ t + 1.$$

To see that p_{i^*} belongs to the vanishing ideal of RFE_{l-1}^{l-2}, we use Theorem 16. Condition 1 of Theorem 16 is satisfied by p_{i^*} since it it is satisfied by p, and all of k, l, n, and the degree of p_{i^*} are one less. Given K and L as in condition 2 of Theorem 16, we have

$$\left(\frac{\partial p_{i^*}}{\partial L}\right)\bigg|_{K \leftarrow 0} = p_{d^*} \tag{25}$$

where d^* is the degree pattern with domain $K \cup L \cup \{i^*\}$ that has $d_j^* = 1$ for $j \in L \cup \{i^*\}$ and $d_j^* = 0$ for $j \in K$. Since $p \in \mathrm{Van}[\mathrm{RFE}_l^{l-1}]$, the contrapositive of the Zoom Lemma applied to p with $K' = K \cup \{i^*\}$, $L' = L \cup \{i^*\}$, d^* and $z = 1$, says that (25) vanishes at (20). So $p_{i^*} \in \mathrm{Van}[\mathrm{RFE}_{l-1}^{l-2}]$ by Theorem 16. This concludes the proof of Lemma 37 modulo the proof of Claim 41. $\square$

Proof of Claim 41. Let $U' \subseteq U$ be the indices of variables x_i such that p depends on x_i, and similarly define $V' \subseteq V$. We first consider the possibility that (24) fails for every $i^* \in V'$. We show that this can only happen when $|V'| < |U'|$. A symmetric argument shows that if (24) fails for all $i^* \in U'$, then it must be that $|U'| < |V'|$. As both inequalities cannot simultaneously occur, this guarantees the existence of the desired i^*.

Suppose that (24) fails for each $i^* \in V'$. Observe that the column of $\mathrm{CMat}_{U,V}(p_{i^*})$ corresponding to a monomial m equals the column of $\mathrm{CMat}_{U,V}(x_{i^*}p_{i^*})$ corresponding to the monomial $x_{i^*}m$; all other columns of $\mathrm{CMat}_{U,V}(x_{i^*}p_{i^*})$ are zero. The matrix $\mathrm{CMat}_{U,V}(x_{i^*}p_{i^*})$ can also be formed from $\mathrm{CMat}_{U,V}(p)$ by zeroing out all the columns indexed by subsets that do not contain i^* (corresponding to multilinear monomials not involving x_{i^*}). The failure of (24) for i^* implies that $\mathrm{CMat}_{U,V}(p_{i^*})$ has the same rank as $\mathrm{CMat}_{U,V}(p)$, which is to say that the columns of $\mathrm{CMat}_{U,V}(p)$ indexed by subsets that contain i^* span *all* the columns of $\mathrm{CMat}_{U,V}(p)$. Going block by block, this implies that for every block C_d of $C = \mathrm{CMat}_{U,V}(p)$, the columns within C_d that are indexed by subsets containing i^* span all the columns of C_d. This goes for every $i^* \in V'$, as we are assuming that (24) fails for all of them.

Let d be minimal such that $C_d \neq 0$, i.e., such that p has a monomial depending on exactly d variables indexed by U. We have $d \geq t \geq 1$ and $C_{d-1} = 0$. The entries of C_d appear in the linear equations (22) given in Proposition 40, either with entries from C_{d-1} or from C_{d+1}. Since C_{d-1} is zero, the equations involving C_{d-1} and C_d simplify to equations on C_d only. Namely, for every $I \subseteq U$ with $|I| = d-1$, every $J \subseteq V$ with $|J| = l-(d-1)$, and every $i^* \in I \cup J$, equation (22) simplifies to

$$\sum_{i \in U \setminus I} \frac{C_d(\{i\} \cup I, J)}{a_i - a_{i^*}} = 0. \tag{26}$$

For any fixed $i \in U \setminus U'$, all entries of the form $C_d(\{i\} \cup I, J)$ are zero. Thus, we can restrict the range of i in (26) from $U \setminus I$ to $U' \setminus I$:

$$\sum_{i \in U' \setminus I} \frac{C_d(\{i\} \cup I, J)}{a_i - a_{i^*}} = 0. \tag{27}$$

Since $C_d \neq 0$, there is at least one fixed I for which not all entries of the form $C_d(\{i\} \cup I, J)$ are zero as i and J vary. Let I^* be such an I, and let C_d^* denote the submatrix of C_d that consists of all entries of the form $C_d(\{i\} \cup I^*, J)$ as i and J vary. For every $J \subseteq V$ with $|J| = l-(d-1)$ and every $i^* \in I^* \cup J$, we have

$$\sum_{i \in U' \setminus I^*} \frac{C_d^*(\{i\} \cup I^*, J)}{a_i - a_{i^*}} = 0. \tag{28}$$

For each $i^* \in V'$, consider the equations (28) where J ranges over all subsets of V of size $|J| = l-(d-1)$ that contain i^*. Observe that the coefficients $\frac{1}{a_i - a_{i^*}}$ in (28) are independent of the choice of J. We argued that the columns of C_d indexed by subsets J that contain i^* span all columns of C_d. The same holds for C_d^*, as C_d^* is obtained from C_d by removing rows. It follows that (28) holds for *every* subset J of V of size $l-(d-1)$ (not just the ones containing i^*).

In particular, consider any one nonzero column of C_d^*. The column represents a nontrivial solution to the homogeneous system (28) of $|V'|$ linear equations (one for each choice of $i^* \in V'$) in $|U' \setminus I^*|$ unknowns (one for each $i \in U' \setminus I^*$). The coefficient matrix $[\frac{1}{a_i - a_{i^*}}]$ is a Cauchy matrix, which is well-known to have full rank. In order for there to be a nontrivial solution, the number of equations must be strictly less than the number of unknowns. In other words, we have $|V'| < |U' \setminus I^*| \leq |U'|$, as desired. □

8 Alternating Algebra Representation

In this section we present in greater detail the alternating algebra-based representation of polynomials suited to studying the vanishing ideal of RFE.

Per Proposition 17, RFE acts separately on the homogeneous parts of any polynomial, so we focus on homogeneous polynomials. We use d as the parameter for degree. Nonzero polynomials with $d \le l$ are automatically outside the ideal, leaving the case of $d = l + 1$ as the simplest nontrivial case. Subsection 8.1 expands the informal discussion in the introduction, describing the representation and characterization for this case where moreover $l = 1$ and $k = 0$. With that in hand, Subsection 8.2 provides a brief introduction to alternating algebra suited to our purpose, and then Subsection 8.3 formalizes the discussion in Subsection 8.1 and extends it to the case of general k, l, and d.

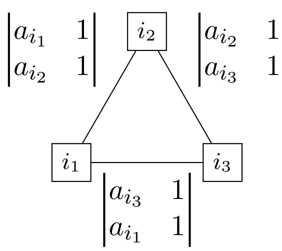
8.1 Basic case

Let $\{x_1, \dots, x_n\}$ be a set of variables. For the purposes of this subsection, we fix the parameters $k = 0$, $l = 1$, and $d = 2$. That is to say, we are studying which degree-2 polynomials belong to the vanishing ideal for RFE^0_1. When $d = l + 1$, polynomials that are not multilinear are automatically outside the ideal by an application of the Zoom Lemma (cf. the proof of Theorem 5 from Theorem 36), so we focus on multilinear polynomials.

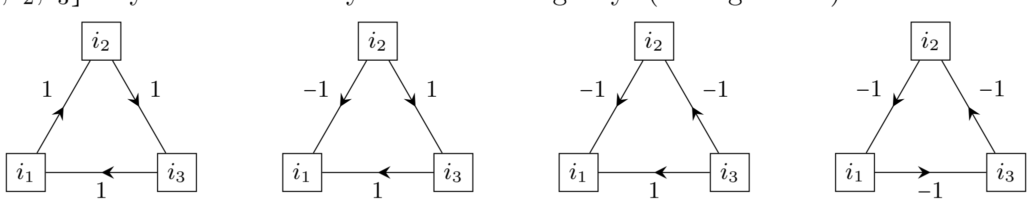
In Theorem 2, we proved that the polynomials $\mathrm{EVC}^0_1[i_1, i_2, i_3]$ as i_1, i_2, i_3 range over $[n]$ generate $\mathrm{Van}[\mathrm{RFE}^0_1]$. As these generators are all degree-2 polynomials, a degree-2 polynomial is in the ideal if and only if it is a linear combination of instantiations of EVC^0_1. Consider the generator when expanded as a linear combination of monomials:

$$\mathrm{EVC}^0_1[i_1, i_2, i_3] = \begin{vmatrix} a_{i_1} & 1 \\ a_{i_2} & 1 \end{vmatrix} x_{i_1} x_{i_2} + \begin{vmatrix} a_{i_3} & 1 \\ a_{i_1} & 1 \end{vmatrix} x_{i_3} x_{i_1} + \begin{vmatrix} a_{i_2} & 1 \\ a_{i_3} & 1 \end{vmatrix} x_{i_2} x_{i_3}.$$

We may represent it graphically by making a vertex for each variable, an undirected edge for each monomial, and assigning to each edge a weight equal to the coefficient of that monomial:



Observe that the coefficient of $x_{i_1} x_{i_2}$ has no dependence on a_{i_3}. In particular, as i_3 varies, the coefficient of $x_{i_1} x_{i_2}$ in $\mathrm{EVC}^0_1[i_1, i_2, i_3]$ does not change. In any other instantiation of EVC^0_1 involving both i_1 and i_2, the coefficient is either the same, or else differs by a sign, according to whether i_1 or i_2 precedes the other in the determinant. Similar holds with respect to all other monomials. This suggests to modify the graphical representation by rescaling the weights on edges. To capture the signs, we use oriented edges. More precisely, for each edge $\{i_1, i_2\}$, we consider either of its two orientations, say $i_1 \to i_2$, and then divide its coefficient by $\begin{vmatrix} a_{i_1} & 1 \\ a_{i_2} & 1 \end{vmatrix}$. Note that considering the opposite orientation coincides with flipping the sign of the scaling factor. With these changes, $\mathrm{EVC}^0_1[i_1, i_2, i_3]$ may be drawn in any of the following ways (among others).



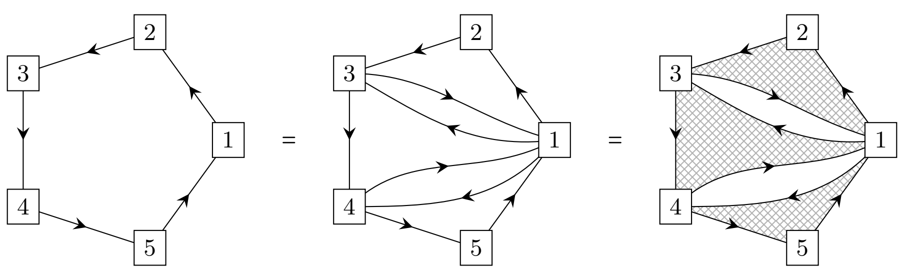
While different choices of edge orientations lead to different illustrations, any one illustration can be transformed into any other by considering edges in opposite orientations as needed, and flipping the sign of each associated coefficient. By identifying each edge in one orientation with the negative of itself in the opposite orientation, we can view all the illustrations as renditions of the same underlying object.

In general, we can represent any degree-2 homogeneous multilinear polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ in a similar way: create a vertex i for each variable x_i, and create an oriented edge for each monomial. For each edge $i_1 \to i_2$, set its coefficient to be the coefficient of $x_{i_1} x_{i_2}$ in p divided by $\begin{vmatrix} a_{i_1} & 1 \\ a_{i_2} & 1 \end{vmatrix}$. The representation determines the polynomial: simply undo the scaling on each edge, and read off a linear combination of monomials. Note moreover that this graphical representation is linear in the polynomial: adding or rescaling polynomials coincides with adding or rescaling coefficients on the edges.

Observe that, in every graphical representation of $\mathrm{EVC}_1^0[i_1, i_2, i_3]$, at every vertex, the sum of the coefficients on edges oriented out of that vertex equals the sum of the coefficients on edges oriented in to that vertex. Indeed, we can interpret $\mathrm{EVC}_1^0[i_1, i_2, i_3]$ as a *circulation* in which one unit of flow travels around a simple 3-cycle $i_1 \to i_2 \to i_3 \to i_1$. The coefficient on an oriented edge $i_1 \to i_2$ measures how much flow is traveling in the direction $i_1 \to i_2$, with negatives representing flow in the opposite direction. That the sum of coefficients on outgoing edges equals the sum of coefficients on incoming edges means that this circulation satisfies a *conservation law* at every vertex: the total flow in equals the total flow out.

Since every degree-2 polynomial in $\mathrm{Van}[\mathrm{RFE}_1^0]$ is a linear combinations of instantiations of EVC_1^0, each is represented by a circulation *that also satisfies the conservation law.* Thus conservation is a necessary condition for membership in $\mathrm{Van}[\mathrm{RFE}_1^0]$. Not every polynomial satisfies this condition: consider, for example, any lone monomial, or a sum of variable-disjoint monomials.

Indeed, conservation is *sufficient* for ideal membership as well. It is folklore that every circulation that satisfies conservation can be decomposed as a linear combination of unit circulations around simple cycles. Unit circulations around simple cycles can be decomposed as a sum of unit circulations on 3-cycles; this is depicted for a 5-cycle below, where each edge indicates unit flow:



The basis of the first equality above is that a unit flow $i_1 \to i_2$ cancels with a unit flow $i_2 \to i_1$.

In summary, a homogeneous degree-2 multilinear polynomial is in $\mathrm{Van}[\mathrm{RFE}_1^0]$ *if and only if*, when represented as a circulation, it satisfies the conservation law. This is the representation and ideal membership characterization for such polynomials in the special case $k = 0$, $l = 1$, and $d = 2$.

8.2 Alternating algebra

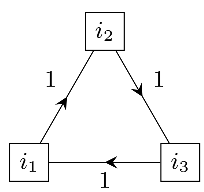
In order to generalize Subsection 8.1, we need to be able to discuss higher-dimensional analogues of "flow" and "circulation", as well as appropriately-generalized notions of "conservation". Suited to this purpose is the language of *alternating algebra*. Alternating algebra was introduced in the 1800s by Hermann Grassmann [Gra44; GK00] and is the formalism underlying differential geometry and its applications to physics. We give only a brief introduction to alternating algebra here, tailored heavily toward our purposes.

For each variable x_i, we create for it with a fresh vertex, labeled i. The alternating algebra provides a multiplication, denoted $\wedge$, that can be thought of as a constructor to make *oriented simplices* out of these vertices. For example, the $\wedge$-product of i_1 with i_2, written $i_1 \wedge i_2$, encodes the simplex with vertices i_1 and i_2. When $i_1 = i_2$, $i_1 \wedge i_2$ is defined to be zero. $\wedge$-multiplication is associative. Rather than being commutative, the $\wedge$-product is *anticommutative* in the sense that $i_1 \wedge i_2 = -i_2 \wedge i_1$. In this way the order of the vertices in the product encodes an orientation. There are only ever two orientations: in a larger product such as $i_1 \wedge i_2 \wedge i_3$, we have

$$\begin{aligned} & i_1 \wedge i_2 \wedge i_3 &=& \ -i_1 \wedge i_3 \wedge i_2 \\ = \ & i_3 \wedge i_1 \wedge i_2 &=& \ -i_3 \wedge i_2 \wedge i_1 \\ = \ & i_2 \wedge i_3 \wedge i_1 &=& \ -i_2 \wedge i_1 \wedge i_3. \end{aligned}$$

In general, permuting the vertices in a $\wedge$-product by an even permutation has no effect, while permuting by an odd permutation flips the sign. Any $\wedge$-product that uses the same vertex more than once is zero. The alternating algebra consists of all formal linear combinations of $\wedge$-products of vertices formed in the preceding way. This includes as a distinct simplex the empty product, denoted 1, which is an identity for $\wedge$. The $\wedge$-product distributes over addition.

To connect this with Subsection 8.1, recall the graphical depiction of $\mathrm{EVC}_1^0[i_1, i_2, i_3]$:



Adopting the convention that an arrow $i_1 \to i_2$ is $i_1 \wedge i_2$ (and so an arrow $i_2 \to i_1$ is $i_2 \wedge i_1 = -i_1 \wedge i_2$), we can alternatively express the above as

$$i_1 \wedge i_2 \ + \ i_2 \wedge i_3 \ + \ i_3 \wedge i_1.$$

In general, the graphical representation of a degree-2 multilinear polynomial is some linear combination of 2-vertex oriented simplices. When we go to higher-degree polynomials, we will use oriented simplices with more vertices.

To express conservation, we use *boundary maps*. Denoted by ∂_w where $w : [n] \to \mathbb{F}$ is any function, they are linear maps that send each simplex to a linear combination of its boundary faces (and the empty simplex to zero) according to the following formula:

$$\partial_w(i_1 \wedge \cdots \wedge i_r) = \sum_{j=1}^{r} (-1)^{1+j} w(i_j) \ i_1 \wedge \cdots \wedge i_{j-1} \wedge i_{j+1} \wedge \cdots \wedge i_r.$$

The map ∂_w extends linearly to all elements of the alternating algebra. In the simplest case, w is the constant-1 function. In this case, the boundary of some 2-vertex simplex is given by

$$\partial_w(i_1 \wedge i_2) = i_2 - i_1.$$

In particular, $i_1 \wedge i_2$ contributes -1 toward i_1 and $+1$ toward i_2. This coincides with the contribution of the edge $i_1 \to i_2$ toward the net flow in to the vertices i_1 and i_2. In exactly this way, conservation is identified with having a *vanishing boundary*. Different choices of w give rise to different boundary operators.

As we generalize parameters, it will become important to iterate boundary operators. The first fact here is that taking the same boundary multiple times always vanishes. That is, for any w, $\partial_w \circ \partial_w = 0$. The second is that for any w, w' and $\beta, \beta' \in \mathbb{F}$, $\partial_{\beta w + \beta' w'} = \beta \partial_w + \beta' \partial_{w'}$, which is to say that the boundary operators themselves are linear in w. It follows from these that, for any w, w', $\partial_w \circ \partial_{w'} = -\partial_{w'} \circ \partial_w$. This means that the boundary operators themselves behave like an alternating algebra, with $\circ$ as the multiplication rather than $\wedge$. For any $w_1, \ldots, w_k$, write $\omega = w_1 \wedge \cdots \wedge w_k$, and define $\partial_\omega = \partial_{w_k} \circ \cdots \circ \partial_{w_1}$. That is, $w_1 \wedge \cdots \wedge w_k$ means apply ∂_{w_1}, then ∂_{w_2}, and so on, up to ∂_{w_k}. We extend this by linearity to any linear combination of such constructions. The result is well-defined.

8.3 General case

With the formalism of alternating algebra in hand, we turn now to formalizing and generalizing the representation that we introduced in Subsection 8.1, henceforth the *simplicial representation*. The parameters $k, l \in \mathbb{N}$ may be arbitrary; however, we will continue to restrict to polynomials of degree $d = l + 1$. As before, since $d = l + 1$, polynomials of degree d that are not multilinear are automatically outside $\mathrm{Van}[\mathrm{RFE}_l^k]$, so we also restrict our attention to multilinear polynomials.

Let $\{x_1, \ldots, x_n\}$ be the set of variables, and correspond each variable x_i with a fresh vertex, labeled i. Degree-$(l+1)$ multilinear polynomials are represented by linear combinations of $(l+1)$-vertex simplices, and we will use simplices of other dimensions to study membership in $\mathrm{Van}[\mathrm{RFE}_l^k]$. Together, all live in the following spaces:

Definition 42 (space of oriented simplices). *For each $t \in \mathbb{N}$, we let*

$$\Sigma^t \doteq \mathrm{span}(i_1 \wedge \cdots \wedge i_t : i_1, \ldots, i_t \in [n])$$

denote the space of linear combinations of t-vertex oriented simplices.

To emphasize, the t in Σ^t counts the number of vertices in the simplices; this is one more than the usual notion of dimension of a simplex.

The representation makes use of certain Vandermonde determinants. We abbreviate them using the following notation:

$$|i_1 \wedge \cdots \wedge i_t|_a \doteq \begin{vmatrix} a_{i_1}^{t-1} & \cdots & 1 \\ \vdots & & \vdots \\ a_{i_t}^{t-1} & \cdots & 1 \end{vmatrix}.$$

Note that the determinants depend on the abscissas a_i that underlie RFE.

Algebraically, the representation of a polynomial in this representation can be concisely understood as follows. For distinct $i_1, \ldots, i_{l+1} \in [n]$, the monomial $x_{i_1} \cdots x_{i_{l+1}}$ is represented as the following element of Σ^{l+1}:

$$\frac{i_1 \wedge \cdots \wedge i_{l+1}}{|i_1 \wedge \cdots \wedge i_{l+1}|_a}.$$

Note that the above is indeed symmetric in the order of the variable indices: exchanging any two indices causes both the numerator and denominator to change signs, to a net effect of no change. Formally, we define the following "decoder map" that maps a simplicial representation to the polynomial it represents:

Definition 43 (representation). *Let $\rho : \Sigma^{l+1} \to \mathbb{F}[x_1, \ldots, x_n]$ be the linear map extending*

$$i_1 \wedge \cdots \wedge i_{l+1} \mapsto |i_1 \wedge \cdots \wedge i_{l+1}|_a \cdot x_{i_1} \cdots x_{i_{l+1}}.$$

ρ is a vector space isomorphism between Σ^{l+1} and the space of homogeneous degree-$(l+1)$ multilinear polynomials.

Reasoning about membership in Van[RFE] makes use of boundary operators $\bigoplus_{t=0}^{n} \Sigma^t \to \bigoplus_{t=0}^{n} \Sigma^t$. Boundary operators are parametrized by functions $w : [n] \to \mathbb{F}$. For our purposes, it is useful to view these functions as *univariate polynomials* in the sense that a univariate polynomial $w \in \mathbb{F}[\alpha]$ determines the function $i \mapsto w(a_i)$.

Definition 44 (boundary operator). *For any univariate polynomial $w \in \mathbb{F}[\alpha]$, the boundary operator with weight function w, $\partial_w : \bigoplus_{t=0}^{n} \Sigma^t \to \bigoplus_{t=0}^{n} \Sigma^t$, is defined to be the linear map extending*

$$i_1 \wedge \cdots \wedge i_t \;\mapsto\; \sum_{j=1}^{t} (-1)^{1+j} w(a_{i_j})\,(i_1 \wedge \cdots \wedge i_{j-1} \wedge i_{j+1} \wedge \cdots \wedge i_t).$$

For each $t \geq 1$, $\partial_w(\Sigma^t) \subseteq \Sigma^{t-1}$, while $\partial_w(\Sigma^0) = \{0\}$.

One may restrict attention to w with degree less than n, since any two polynomials that are the same as functions $\{a_i : i \in [n]\} \to \mathbb{F}$ determine the same boundary operator. In general, we will be interested in the boundaries that are weighted by low-degree polynomials. When w has degree δ or less, we say that ∂_w is a degree-δ boundary.

As discussed in Subsection 8.2, boundary operators under composition behave like an alternating algebra. In this way, the expression ∂_ω for some $\omega = w_1 \wedge \cdots \wedge w_r$ is defined to be $\partial_{w_r} \circ \cdots \circ \partial_{w_1}$.

We can now describe the simplicial representation of EVC:

Lemma 45. *For any $k, l \in \mathbb{N}$ and $i_1, \ldots, i_{k+l+2} \in [n]$,*

$$\mathrm{EVC}_l^k[i_1, \ldots, i_{k+l+2}] \;=\; (-1)^{(k+1)(l+1)} \cdot \rho\left(\, \partial_{\alpha^k \wedge \cdots \wedge \alpha^0} (i_1 \wedge \cdots \wedge i_{k+l+2}) \,\right). \tag{29}$$

That is, EVC_l^k is the polynomial formed (up to sign) from a given $(k+l+2)$-vertex simplex by iteratively applying to it the $k+1$ boundaries weighted by $\alpha^0, \alpha^1, \ldots, \alpha^k$. The sign factor in (29) can be removed if one rearranges (4) so that the first $l+1$ columns come after the last $k+1$ columns.

Every degree-$(l+1)$ polynomial in $\mathrm{Van}[\mathrm{RFE}_l^k]$ is a linear combination of instantiations of EVC_l^k, which is to say that it is in the image of Σ^{k+l+2} through $\rho \circ \partial_{\alpha^k \wedge \cdots \wedge \alpha^0}$. Equivalently, for any degree-$(l+1)$ homogeneous multilinear polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$, p belongs to $\mathrm{Van}[\mathrm{RFE}_l^k]$ if and only if $\rho^{-1}(p)$ is in the set $\partial_{\alpha^k \wedge \cdots \wedge \alpha^0}(\Sigma^{l+k+2})$. The following relationship is an instance of a general phenomenon in alternating algebra:

$$\mathrm{Im}\left(\partial_{\alpha^k \wedge \cdots \wedge \alpha^0}\right) \;=\; \bigcap_{r=0}^{k} \mathrm{Ker}\left(\partial_{\alpha^r}\right).$$

This leads to the following characterization of the degree-$(l+1)$ elements of $\mathrm{Van}[\mathrm{RFE}_l^k]$:

Theorem 46. *Let $k, l \in \mathbb{N}$. For any homogeneous multilinear polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ of degree $l+1$, $p(\mathrm{RFE}_l^k) = 0$ if and only if*

$$\partial_w(\, \rho^{-1}(p)\,) = 0$$

for every $w \in \mathbb{F}[\alpha]$ of degree at most k.

In other words, for any degree-$(l+1)$ homogeneous multilinear polynomial, it vanishes at RFE_l^k if and only if in the simplicial representation it satisfies conservation with respect to all degree-k boundaries. This is the representation and ideal membership characterization for such polynomials for general k and l in the special case of $d = l+1$.

Theorem 46 is ultimately a reformulation of Theorem 4 (or, more precisely, Theorem 16) in the specific case of homogeneous, multilinear polynomials of degree $l+1$. For sets K and L as in Theorem 16, let $w(\alpha) \doteq \prod_{y \in K}(\alpha - a_y)$ and let $i_1, \ldots, i_l$ enumerate L. The coefficient of $i_1 \wedge \cdots \wedge i_l$ in $\partial_w(\rho^{-1}(p))$, multiplied by $|i_1 \wedge \cdots \wedge i_l|_a$, precisely equals the evaluation of $\left(\frac{\partial p}{\partial K}\right)\Big|_{L \leftarrow 0}$ at (20).

Theorem 46 extends to higher degrees in the following way. Let $d > l$, and let $\mathbb{F}[x_1, \ldots, x_n]_{=d}$ be the space of degree-d homogeneous polynomials. For each monomial $x_{i_1} \cdots x_{i_d} \in \mathbb{F}[x_1, \ldots, x_n]_{=d}$ with distinct $i_1, \ldots, i_d$, we represent it as $i_1 \wedge \cdots \wedge i_d / |i_1 \wedge \cdots \wedge i_d|_a$. This means extending our decoder map ρ to Σ^d:

$$\begin{array}{rccl} \rho \ : & \Sigma^d & \to & \mathbb{F}[x_1, \ldots, x_n]_{=d} \\ & i_1 \wedge \cdots \wedge i_d & \mapsto & |i_1 \wedge \cdots \wedge i_d|_a \cdot x_{i_1} \cdots x_{i_d} \end{array} . \tag{30}$$

ρ is a vector space isomorphism between Σ^d and the multilinear subspace of $\mathbb{F}[x_1, \ldots, x_n]_{=d}$.

The following theorem characterizes $\mathrm{Van}[\mathrm{RFE}_l^k]$ within this representation. It is likewise ultimately a reformulation of Theorem 4 (or Theorem 16 to be precise).

Theorem 47. *Let $k, l \in \mathbb{N}$ and $\Delta \geq 1$. For any degree-$(l+\Delta)$ homogeneous multilinear polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$, $p(\mathrm{RFE}_l^k) = 0$ if and only if*

$$\partial_{w_1 \wedge \cdots \wedge w_\Delta}(\ \rho^{-1}(p)\ ) = 0$$

for every $w_1, \ldots, w_\Delta \in \mathbb{F}[\alpha]$ of degree at most $k + \Delta - 1$.

Theorems 46 and 47 do well for understanding the multilinear elements of the vanishing ideal. For non-multilinear elements, one may do the following. Let $\widehat{\Sigma}^t$ be Σ^t except that coefficients may be arbitrary polynomials in $\mathbb{F}[x_1, \ldots, x_n]$ rather than just scalars in $\mathbb{F}$. The decoder map ρ and boundary maps ∂_w carry over to $\widehat{\Sigma}^t$ directly, though now ρ is no longer injective. The following variation of Theorem 46 characterizes ideal membership for arbitrary polynomials.

Theorem 48. *Let $k, l \in \mathbb{N}$. For any polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$, $p(\mathrm{RFE}_l^k) = 0$ if and only if there exists $\eta \in \widehat{\Sigma}^{l+1}$ with $\rho(\eta) = p$ such that, for every $w \in \mathbb{F}[\alpha]$ of degree at most k,*

$$\partial_w(\eta) = 0.$$

While Theorem 48 applies to a broader class of polynomials, it has the drawback that representing polynomials with $\widehat{\Sigma}^{l+1}$ is too redundant. Specifically, whenever p has a representation in $\widehat{\Sigma}^{l+1}$, there are many $\eta \in \widehat{\Sigma}^{l+1}$ that represent p, and most of them do *not* satisfy the boundary conditions, even when p belongs to the vanishing ideal. This weakens the utility of the characterization. Theorems 46 and 47 yield straightforward tests: given p, form the unique η with $\rho(\eta) = p$, and then check whether the boundary conditions hold for η. Theorem 48, on the other hand, leaves η comparatively underspecified.

Acknowledgements

We are grateful to Hervé Fournier and Arpita Korwar for their presentation at WACT'18 in Paris [FK18]. We are indebted to Gautam Prakriya for helpful discussions and detailed feedback. We also thank Amir Shpilka for helpful comments and suggestions, as well as Ilya Volkovich and the anonymous reviewers of the conference version.

A RFE as a Hitting Set Generator

In Definition 1, we defined RFE as a set of substitutions formed by varying the seed f over certain rational functions with coefficients in $\mathbb{F}$. Meanwhile, our analyses proceed by parametrizing f by scalars, abstracting the scalar parameters as fresh formal variables, and calculating in the field of rational functions in those variables. The approaches are equivalent over large enough fields, however, and the flexibility to choose is a source of convenience. Here are some natural parametrizations of f:

Coefficients. Select scalars $g_0, \ldots, g_k, h_0, \ldots, h_l \in \mathbb{F}$ and set

$$f(\alpha) = \frac{g_k\alpha^k + g_{k-1}\alpha^{k-1} + \cdots + g_1\alpha + g_0}{h_l\alpha^l + h_{l-1}\alpha^{l-1} + \cdots + h_1\alpha + h_0},$$

ignoring choices of $h_0, \ldots, h_l$ for which the denominator vanishes at some abscissa.

Evaluations. Fix two collections, $B = \{b_1, \ldots, b_{k+1}\}$ and $C = \{c_1, \ldots, c_{l+1}\}$, each of distinct scalars from $\mathbb{F}$. Then select scalars $g_1, \ldots, g_{k+1}$ and $h_1, \ldots, h_{l+1}$ and set

$$f(\alpha) = \frac{g(\alpha)}{h(\alpha)}$$

where g is the unique degree-k polynomial with $g(b_1) = g_1,\ g(b_2) = g_2,\ \ldots,\ g(b_{k+1}) = g_{k+1}$, and h is defined similarly with respect to C. Choices of $h_1, \ldots, h_{l+1}$ that lead h to vanish at some abscissa are ignored.

Note that an explicit formula for g and h in terms of the parameters can be obtained using the Lagrange interpolants with respect to B and C.

Roots. Select scalars $z, s_1, \ldots, s_{k'}, t_1, \ldots, t_{l'} \in \mathbb{F}$ for some $k' \leq k$ and $l' \leq l$ and set

$$f(\alpha) = z \cdot \frac{(\alpha - s_1) \cdot \cdots \cdot (\alpha - s_{k'})}{(\alpha - t_1) \cdot \cdots \cdot (\alpha - t_{l'})},$$

where $\{t_1, \ldots, t_{l'}\}$ is disjoint from the set of abscissas.

In fact, it is no loss of power to restrict to $k' = k$ and $l' = l$.

Hybrids are of course possible, too. For example, Proposition 51 uses the evaluations parametrization for the numerator and roots parametrization for the denominator.

The following lemma justifies that, for any polynomial p, as long as $\mathbb{F}$ is large enough, $p(\text{RFE})$ vanishes with respect to a particular parametrization of RFE if and only if it vanishes with respect to RFE as defined in Definition 1. The lemma is an extension of the well-known analogous result for polynomials [Ore22; DL78; Zip79; Sch80].

Lemma 49. *Let $\mathbb{F}$ be field, and $f = g/h \in \mathbb{F}(\tau_1, \ldots, \tau_l)$ be a rational function in l variables with $\deg(g) \leq d$ and $\deg(h) \leq d$. Let $S \subseteq \mathbb{F}$ be finite. Then the probability that f vanishes or is undefined when each τ_i is substituted by a uniformly random element of S is at most $2d/|S|$.*

In particular, if $\mathbb{F}$ is infinite, then, for all polynomials p, all the above parametrizations and Definition 1 are equivalent for the purposes of hitting p; when p is fixed, the equivalence holds provided $|\mathbb{F}| \geq \text{poly}(n, \deg(p))$. Quantitative bounds on the number of substitutions to perform when testing whether RFE hits p in the blackbox algorithm likewise follow from Lemma 49. As is customary in the context of blackbox derandomization of PIT, if $\mathbb{F}$ is not large enough, then one works instead over a sufficiently large extension of $\mathbb{F}$.

B Equivalence between RFE and SV

The Shpilka-Volkovich generator can be defined as follows in the format of our definition of RFE.

Definition 50 (SV Generator). *The* Shpilka–Volkovich (SV) Generator *for polynomials in the variables $x_1, \ldots, x_n$ is parametrized by the following data:*

- *For each $i \in [n]$, a distinct $a_i \in \mathbb{F}$.*
- *A positive integer, l.*

The generator takes as seed l pairs of scalars $(y_1, z_1), \ldots, (y_l, z_l)$ and substitutes

$$x_i \leftarrow \sum_{t=1}^{l} \left(z_t \cdot \prod_{j \in [n] \setminus \{i\}} \frac{y_t - a_j}{a_i - a_t} \right).$$

We abbreviate the generator to SV^l or just SV.

Shpilka and Volkovich designed the generator SV^l so that any selection of l of the variables could remain independent while the others were forced to zero. This can be viewed as an algebraic version of l-wise independence. SV^1 was realized with two seed variables, y and z, using Lagrange interpolation. The fresh variable y enables selecting one of the original variables x_i, namely by setting $y = a_i$. The selected variable x_i is then set to z, while the other variables are set to zero. For larger l, SV^l is the sum of l independent copies of SV^1.

We now formally state and argue the close relationship between SV^l and RFE_l^{l-1} that we sketched in Section 1.

Proposition 51. *Let $\{x_1, \ldots, x_n\}$ be a set of variables and $l \geq 1$. There is an invertible diagonal transformation $A : \mathbb{F}^n \to \mathbb{F}^n$ such that, for any polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$, $p(\mathrm{SV}^l) = 0$ if and only if $(p \circ A)(\mathrm{RFE}_l^{l-1}) = 0$.*

In particular, the vanishing ideals of RFE_l^{l-1} and of SV^l are the same up to the rescaling of Proposition 51.

Proof of Proposition 51. Let $\widehat{\mathbb{F}}$ be the field of rational functions in indeterminates $\upsilon_1, \ldots, \upsilon_l, \zeta_1, \ldots, \zeta_l$ over $\mathbb{F}$. A polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$ has $p(\mathrm{SV}^l) = 0$ if and only if p vanishes at the point

$$\left(\sum_{t=1}^{l} \zeta_t \prod_{j \in [n] \setminus \{i\}} \frac{\upsilon_t - a_j}{a_i - a_j} \; : \; i \in [n] \right) \in \widehat{\mathbb{F}}^n. \tag{31}$$

Set $A : \mathbb{F}^n \to \mathbb{F}^n$ to be the diagonal linear transformation that divides the coordinate for x_i by $\prod_{j \in [n] \setminus \{i\}} (a_i - a_j)$. It is invertible. Applying A^{-1} to (31) yields the point

$$\left(\sum_{t=1}^{l} \zeta_t \prod_{j \in [n] \setminus \{i\}} (\upsilon_t - a_j) \; : \; i \in [n] \right) = \left(\sum_{t=1}^{l} \left(\zeta_t \prod_{j \in [n]} (\upsilon_t - a_j) \right) \frac{1}{\upsilon_t - a_i} \; : \; i \in [n] \right). \tag{32}$$

p vanishes at (31) if and only if $p \circ A$ vanishes at (32). Now let $\widehat{\mathbb{F}}'$ be the field of rational functions in indeterminates $\tau_1, \ldots, \tau_l, \sigma_1, \ldots, \sigma_l$ over $\mathbb{F}$. After the invertible change of variables

$$\zeta_t \leftarrow \frac{1}{\prod_{j \in [n]} (\tau_t - a_j)} \cdot \frac{-\sigma_t}{\prod_{s \neq t} (\tau_t - \tau_s)} \quad \text{and} \quad \upsilon_t \leftarrow \tau_t$$

(32) becomes

$$\left(\sum_{t=1}^{l} \frac{\sigma_t}{(\prod_{s\neq t} \tau_t - \tau_s)} \frac{1}{a_i - \tau_t} \ : \ i \in [n]\right) = \left(\frac{\sum_{t=1}^{l} \sigma_t \prod_{s\neq t} \frac{a_i - \tau_s}{\tau_t - \tau_s}}{\prod_{t=1}^{l} a_i - \tau_t} \ : \ i \in [n]\right) \in \widehat{\mathbb{F}}'^n. \tag{33}$$

Since the change of variables is invertible, $p \circ A$ vanishes at (32) if and only if it vanishes at (33).

Now, viewing $\sigma_1, \ldots, \sigma_l, \tau_1, \ldots, \tau_l$ as seed variables, observe that the right-hand side of (33) is $\mathrm{RFE}_l^{l-1}(g/h)$ where g is parametrized by evaluations ($g(\tau_t) = \sigma_t$) and h is parametrized by roots $(\tau_1, \ldots, \tau_l)$. (See appendix A for a discussion on parametrizations of RFE.) It follows that $p \circ A$ vanishes at (33) if and only if $(p \circ A)(\mathrm{RFE}_l^{l-1}) = 0$. The proposition follows. □

References

[AFS+18] Matthew Anderson, Michael A. Forbes, Ramprasad Saptharishi, Amir Shpilka, and Ben Lee Volk. Identity testing and lower bounds for read-k oblivious algebraic branching programs. *ACM Transactions on Computation Theory*, 10(1):1–30, 2018.

[AGK+15] Manindra Agrawal, Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Hitting-sets for ROABP and sum of set-multilinear circuits. *SIAM Journal on Computing*, 44(3):669–697, 2015.

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

[ASS13] Manindra Agrawal, Chandan Saha, and Nitin Saxena. Quasi-polynomial hitting-set for set-depth-Δ formulas. In *ACM Symposium on Theory of Computing (STOC)*, pages 321–330, 2013.

[AvMV15] Matthew Anderson, Dieter van Melkebeek, and Ilya Volkovich. Deterministic polynomial identity tests for multilinear bounded-read formulae. *Computational Complexity*, 24(4):695–776, 2015.

[BG21] Vishwas Bhargava and Sumanta Ghosh. Improved hitting set for orbit of ROABPs. In *International Conference on Randomization and Computation (RANDOM)*, 30:1–30:23, 2021.

[CLO13] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer Science & Business Media, 2013.

[DL78] Richard A. Demillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978.

[FK18] Hervé Fournier and Arpita Korwar. Limitations of the Shpilka–Volkovich generator. Workshop on Algebraic Complexity Theory (WACT), Paris, 2018.

[For15] Michael A. Forbes. Deterministic divisibility testing via shifted partial derivatives. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 451–465, 2015.

[FS13] Michael A. Forbes and Amir Shpilka. Quasipolynomial-time identity testing of non-commutative and read-once oblivious algebraic branching programs. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 243–252, 2013.

[FSS14] Michael A. Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *ACM Symposium on Theory of Computing (STOC)*, pages 867–875, 2014.

[FSV17] Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct hitting sets and barriers to proving algebraic circuits lower bounds. In *ACM Symposium on Theory of Computing (STOC)*, pages 653–664, 2017.

[FSV18] Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct hitting sets and barriers to proving algebraic circuits lower bounds, 2018. arXiv: `1701.05328 [cs.CC]`. Full version of [FSV17].

[GG20] Zeyu Guo and Rohit Gurjar. Improved explicit hitting-sets for ROABPs. In *International Conference on Randomization and Computation (RANDOM)*, 2020.

[GK00] Hermann Grassmann and Lloyd C Kannenberg. *Extension theory*. American Mathematical Society, 2000. Translation of [Gra44].

[GKS+17] Rohit Gurjar, Arpita Korwar, Nitin Saxena, and Thomas Thierauf. Deterministic identity testing for sum of read-once oblivious arithmetic branching programs. *Computational Complexity*, 26(4):835–880, 2017.

[GKS17] Rohit Gurjar, Arpita Korwar, and Nitin Saxena. Identity testing for constant-width, and any-order, read-once oblivious arithmetic branching programs. *Theory of Computing*, 13(2):1–21, 2017.

[Gra44] Hermann Grassmann. *Die lineale Ausdehnungslehre ein neuer Zweig der Mathematik: dargestellt und durch Anwendungen auf die übrigen Zweige der Mathematik, wie auch auf die Statik, Mechanik, die Lehre vom Magnetismus und die Krystallonomie erläutert*. O. Wigand, 1844.

[HS80] Joos Heintz and Claus-Peter Schnorr. Testing polynomials which are easy to compute. In *ACM Symposium on Theory of Computing (STOC)*, pages 262–272, 1980.

[IW97] Russell Impagliazzo and Avi Wigderson. P=BPP if E requires exponential circuits: derandomizing the XOR lemma. In *ACM Symposium on Theory of Computing (STOC)*, pages 220–229, 1997.

[JQS09] Maurice Jansen, Youming Qiao, and Jayalal Sarma. Deterministic identity testing of read-once algebraic branching programs, 2009. arXiv: `0912.2565 [cs.CC]`.

[JQS10] Maurice Jansen, Youming Qiao, and Jayalal Sarma M. N. Deterministic black-box identity testing π-ordered algebraic branching programs. In *IARCS Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 296–307, 2010.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1–2):1–46, 2004.

[KM96] Klaus Kühnle and Ernst W. Mayr. Exponential space computation of Gröbner bases. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 63–71, 1996.

[KMS+13] Zohar S. Karnin, Partha Mukhopadhyay, Amir Shpilka, and Ilya Volkovich. Deterministic identity testing of depth-4 multilinear circuits with bounded top fan-in. *SIAM Journal on Computing*, 42(6):2114–2131, 2013.

[Kor21] Arpita Korwar. Personal communication, 2021.

[KS17] Mrinal Kumar and Shubhangi Saraf. Arithmetic circuits with locally low algebraic rank. *Theory of Computing*, 13(1):1–33, 2017.

[May97] Ernst W. Mayr. Some complexity results for polynomial ideals. *Journal of Complexity*, 13(3):303–325, 1997.

[MS21] Dori Medini and Amir Shpilka. Hitting sets and reconstruction for dense orbits in VP_e and $\Sigma\Pi\Sigma$ circuits. In *Computational Complexity Conference (CCC)*, 19:1–19:27, 2021.

[MV18] Daniel Minahan and Ilya Volkovich. Complete derandomization of identity testing and reconstruction of read-once formulas. *ACM Transactions on Computation Theory (TOCT)*, 10(3):1–11, 2018.

[Nis91] Noam Nisan. Lower bounds for non-commutative computation. In *ACM Symposium on Theory of Computing (STOC)*, pages 410–418, 1991.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994.

[Ore22] Øystein Ore. Über höhere Kongruenzen. *Norsk Mat. Forenings Skrifter*, 1(7), 1922.

[RS05] Ran Raz and Amir Shpilka. Deterministic polynomial identity testing in non-commutative models. *Computational Complexity*, 14(1):1–19, 2005.

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

[ST21] Chandan Saha and Bhargav Thankey. Hitting sets for orbits of circuit classes and polynomial families. In *International Conference on Randomization and Computation (RANDOM)*, 50:1–50:26, 2021.

[SV15] Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. *Computational Complexity*, 24(3):477–532, 2015.

[SY10] Amir Shpilka and Amir Yehudayoff. *Arithmetic circuits: A survey of recent results and open questions*. Now Publishers Inc, 2010.

[vMM22] Dieter van Melkebeek and Andrew Morgan. Polynomial identity testing via evaluation of rational functions. In *Innovations in Theoretical Computer Science Conference (ITCS)*, volume 215 of *LIPIcs*, 119:1–119:24, 2022.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *International symposium on symbolic and algebraic manipulation*, pages 216–226, 1979.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il