

Polynomial Identity Testing and Lower Bounds for Sparse Affine Orbits

Vishwas Bhargava*

Akash Kumar Sengupta†

Abstract

Sparsity is a basis-dependent notion: a polynomial with many monomials may become sparse after an invertible affine change of variables. We study polynomial identity testing (PIT) and lower bounds for sparse polynomials under arbitrary invertible affine changes of coordinates, that is, for affine orbits of sparse polynomials. For PIT, the goal is a small explicit set that is not contained in the zero set of any nonzero polynomial that is sparse in some affine coordinate system. For lower bounds, the goal is an explicit polynomial that has many monomials in every affine coordinate system.

Over fields of characteristic zero, we give an explicit hitting set of size $O(n^3 s^2 d)$ for affine orbits of n -variate, degree- d , s -sparse polynomials, improving the quasipolynomial bound $(nd)^{O(\log s)}$ of Medini and Shpilka [CCC 2021]. Over $\mathbb{R}$, we give a hitting set of size $O(n^4 s^5)$, consisting of integer points with $\text{poly}(n, s)$ bits, that works for every degree. We also give explicit polynomials of degree d that have at least $\binom{n-1+d/3}{n-1}$ monomials in every affine coordinate system. This bound grows with the degree even when n is fixed, and gives n^n at degree $3n^2$. The previous best lower bound, $2^{\Omega(n)}$, comes from lower bounds of Nisan and Wigderson [Comput. Complexity 1997] and of Kayal, Saha, and Tavenas [ICALP 2016] for homogeneous depth-three circuits, and does not grow with the degree.

Our hitting set combines one structural lemma with the recent hitting sets of Bhardwaj and Saptharishi for polynomials with small partial derivative spaces, and our lower bound follows from the lemma directly. The lemma says that if f is s -sparse in some affine coordinate system, then after a shift by a generic point $(\beta, \beta^2, \dots, \beta^n)$, the lowest-degree homogeneous component of f has partial derivative space of dimension at most s .

Just as sparse PIT is a building block for PIT of many circuit classes, sparse-orbit PIT is a building block for PIT of many orbit classes. As direct applications, we obtain polynomial-size hitting sets for affine orbits of constant-depth, constant-occur formulas and of circuits over sparse polynomials of bounded transcendence degree, improving the quasipolynomial bounds of Saha and Thankey [ToCT 2024].

*Department of Computer Science and Engineering, Indian Institute of Technology Bombay, India.
vishwas@cse.iitb.ac.in

†Department of Computer Science, Rutgers University, USA. akash.sengupta@rutgers.edu

1 Introduction

Sparsity, the number of monomials of a polynomial, is one of the most basic complexity measures in algebraic complexity, and sparse polynomials are well studied: there are efficient algorithms for identity testing and interpolation [30, 5, 17, 19] and for factoring [29, 8]. However, sparsity is a basis-dependent notion: the number of monomials of a polynomial depends on the choice of coordinates. For example, $(x_1 + \cdots + x_n)^d$ has $\binom{n+d-1}{d}$ monomials, but becomes a single monomial after an invertible linear change of variables. Geometers prefer to work invariantly, which emphasizes the geometry and makes symmetry easier to exploit; as Landsberg puts it, paraphrasing Fulton, “Don’t use coordinates unless someone holds a pickle to your head” [21, §0.3]. Many circuit classes, sparse polynomials being the simplest, are instead defined in a fixed coordinate system. We study sparse polynomials up to invertible affine changes of coordinates, that is, polynomials of the form

$$f(\bar{x}) = p(A\bar{x} + \bar{b}) = \sum_{\bar{\alpha} \in S} c_{\bar{\alpha}} \prod_{i=1}^n L_i(\bar{x})^{\alpha_i}, \quad |S| \leq s, \quad (1.1)$$

where $A \in \text{GL}_n(\mathbb{F})$ and $L_i(\bar{x}) = (A\bar{x} + \bar{b})_i$; here $\text{sp}(p) \leq s$, where $\text{sp}(p)$ is the number of monomials of p . Thus f is computed by a depth-three circuit of formal degree $\deg f$ in which every product gate uses the same n affine forms $L_1, \dots, L_n$. The *sparse-orbit complexity* $\text{osp}(f)$ is the least s for which f has such a representation. We study identity testing and lower bounds for this model.

A *hitting set* for a class of polynomials is a set of points such that every nonzero member of the class is nonzero at some point of the set. Hitting sets for orbits of circuit classes were studied independently and simultaneously by Medini and Shpilka [23] and by Saha and Thankey [26].

To motivate the study of orbit PIT more concretely, we give two reasons, following Saha and Thankey [26]. First, affine projections of polynomials from weak classes, such as constant-width ROABPs, already capture general formulas. In characteristic zero, every affine projection of f lies in the orbit closure of f , so the orbit forms a natural dense subset of its affine projections [26].

Second, a hitting set for an orbit class must not be contained in the zero set of any nonzero polynomial in the class in any affine coordinate system. Thus the same hitting set must work independently of the choice of coordinates. Hitting sets for a class do not automatically transfer to its orbit, since many classes are defined relative to a coordinate system; for example, the orbit of a sparse polynomial contains an explicit polynomial that requires exponential-size ROABPs [26, Appendix E.1].

Sparse polynomials are a basic example of such a coordinate-dependent class. Just as hitting sets for sparse polynomials are a building block for PIT of many circuit classes [4, 1], hitting sets for sparse orbits are a building block for PIT of many orbit classes [26].

Besides PIT and lower bounds, the third basic problem for this model is reconstruction: Baraskar, Dewan, Saha, and Sinha [2] show that deciding whether a sparse polynomial lies in the affine orbit of an s -sparse polynomial, and approximating osp , are NP-hard, which suggests that proper learning of sparse orbits is hard.

1.1 Our results

Unless otherwise stated, we work over a field $\mathbb{F}$ of characteristic zero. As discussed above, we consider invertible affine transformations on the space of polynomials and study the orbits of sparse polynomials under this action. More precisely, for $f \in \mathbb{F}[x_1, \dots, x_n]$ we define

$$\text{orb}(f) = \{f(A\bar{x} + \bar{b}) \mid A \in \text{GL}_n(\mathbb{F}), \bar{b} \in \mathbb{F}^n\}.$$

Similarly, given a set $\mathcal{F} \subseteq \mathbb{F}[x_1, \dots, x_n]$, we define

$$\text{orb}(\mathcal{F}) := \{f(A\bar{x} + \bar{b}) \mid f \in \mathcal{F}, A \in \text{GL}_n(\mathbb{F}), \bar{b} \in \mathbb{F}^n\}.$$

For $d \geq 1$ and $s \geq 1$, let

$$\text{SP}_{n,d,s} = \{f \in \mathbb{F}[x_1, \dots, x_n] \mid \deg(f) \leq d, \text{sp}(f) \leq s\}.$$

We give a polynomial-size hitting set for affine orbits of n -variate, degree- d , s -sparse polynomials.

Theorem 1.2 (Identity testing). *Fix $n, d, s \geq 1$. Let $B = ns^2 + s$ and $G = (G_1, \dots, G_n)$ with $G_i(t) = \sum_{j=1}^B \frac{i^j}{j} t^j$ for $i \in [n]$, and $\bar{a} : \mathbb{F} \rightarrow \mathbb{F}^n$ be given by $\beta \mapsto (\beta, \beta^2, \dots, \beta^n)$. Consider the set*

$$\mathcal{H} := \{\bar{a}(\beta) + t^B G(t) : \beta \in \{0, \dots, n^2\}, t \in \{0, \dots, 2dB\}\}.$$

The following holds:

1. $|\mathcal{H}| \leq (n^2 + 1)(2d(ns^2 + s) + 1)$.
2. $\mathcal{H}$ is a hitting set for $\text{orb}(\text{SP}_{n,d,s})$.
3. The set $\mathcal{H}$ can be constructed using $O(n dB^2 + n^3 dB)$ field operations.

Previously, for individual degree d , Saha and Thankey [26, Theorem 7] and Bhargava and Ghosh [7] gave hitting sets of size $(nd)^{O(d \log s)}$; Medini and Shpilka [23, Corollary 1.33] removed the individual-degree restriction, obtaining $(nd)^{O(\log s)}$. In this paper, we actually show that $\Gamma_{n,s}(\beta, t) = \bar{a}(\beta) + t^B G(t)$ is a hitting-set generator for this class with no degree bound (Corollary 4.3): for all but at most n^2 values of β , substituting it gives a nonzero univariate polynomial in t . The degree bound d in Theorem 1.2 is used only to bound the degree of this univariate by $2dB$, and hence the number of values of t .

Remark 1.3. Over $\mathbb{R}$, using a real-root bound of Koiran, Portier, and Tavenas [20], we can remove the dependence on degree. In particular, in Theorem 4.4 we show that there is an integer hitting set of size $O(n^4 s^5)$ with $O(ns^2 \log(ns))$ -bit coordinates that works for every degree. $\diamond$

Our second result is a lower bound on sparsity in every affine coordinate system that grows with the degree. For $f \in \mathbb{F}[x_1, \dots, x_n]$, we define

$$\text{osp}_{\mathbb{F}}(f) = \min_{A \in \text{GL}_n(\mathbb{F}), \bar{b} \in \mathbb{F}^n} \text{sp}(f(A\bar{x} + \bar{b})).$$

Equivalently, $\text{osp}_{\mathbb{F}}(f)$ is the least sparsity of a polynomial p in a representation $f(\bar{x}) = p(A\bar{x} + \bar{b})$ with A invertible.

Our hard polynomials are powers of a cubic whose singular locus contains the rational normal curve given by $\text{im}(\bar{a}) \subseteq \mathbb{F}^n$, i.e. it vanishes together with its gradient, at every point $\bar{a}(\beta) = (\beta, \beta^2, \dots, \beta^n)$. For $n \geq 4$, we define

$$C_n(\bar{x}) = \sum_{j=3}^{n-1} \det \begin{pmatrix} 1 & x_1 & x_{n+1-j} \\ x_1 & x_2 & x_{n+2-j} \\ x_{j-1} & x_j & x_n \end{pmatrix}. \quad (1.4)$$

Theorem 1.5 (Sparse-orbit lower bound). *For $n \geq 4$ and $k \geq 1$, we have*

1. $\deg C_n^k = 3k$, and
2. $\text{osp}_{\mathbb{F}}(C_n^k) \geq \binom{n-1+k}{n-1} + \binom{n-2+k}{n-1}$.

In terms of the degree $d = 3k$, the bound is at least $\binom{n-1+d/3}{n-1}$, which grows with d even when n is fixed; at $d = 3n^2$ it is at least n^n . The previous best lower bound, $2^{\Omega(n)}$, comes from depth-three lower bounds as follows. The representation (1.1) is a depth-three circuit of formal degree d with at most $\text{sp}(p)$ product gates, and lower bounds for homogeneous depth-three circuits extend to this model. Indeed, keeping the degree- d part of each product gate gives a homogeneous depth-three circuit for $\mathcal{H}_d(f)$ with no more product gates. Nisan and Wigderson [24, Lemmas 3 and 4] give such a bound for explicit polynomials of degree $d \leq \epsilon n$. Kayal, Saha, and Tavenas [18, Theorem 4] give a $2^{\Omega(n)}$ size lower bound for every degree $d \geq n$; since each product gate here has size $O(nd)$, this yields $2^{\Omega(n)}/O(nd)$ monomials. However, neither of these bounds grow with d . Moreover, we note that for $\mathbb{F} = \mathbb{C}$, the polynomial C_n^k has a $\Sigma\Pi\Sigma$ -circuit of size polynomial in n and k (Appendix A). Therefore our Theorem 1.5 is a lower bound for the orbit model only.

The key lemma: partial derivatives of initial form. We prove a structural lemma that bounds the dimension of the space of partial derivatives, which is the key behind both the main results described above. We build on an observation of Forbes [13]: in characteristic zero, after a shift with nonzero coordinates, an s -sparse polynomial has a trailing monomial with at most s monomial divisors. We note that the same argument bounds the partial derivative dimension of the entire *initial form*, i.e. the lowest-degree nonzero homogeneous component.

For $h \in \mathbb{F}[x_1, \dots, x_n]$, let $\dim \partial_{\bar{x} < \infty}(h)$ denote the dimension of the span of the Hasse derivatives of h , i.e. the coefficients of $h(\bar{x} + \bar{z})$ as a polynomial in $\bar{z}$. Note that in characteristic 0, the Hasse derivatives span the same space as the ordinary partial derivatives. For $\bar{a} \in \mathbb{F}^n$, let $\mathcal{H}_{\min, \bar{a}}(f)$ denote the initial form of $f(\bar{x} + \bar{a})$. Our key structural lemma is the following.

Lemma 1.6 (Key lemma). *Let $\mathbb{F}$ be a field of arbitrary characteristic and $n, s \geq 1$. Let $p \in \mathbb{F}[x_1, \dots, x_n]$ be a polynomial of sparsity $\text{sp}(p) \leq s$. Let $0 \neq f = p(A\bar{x} + \bar{b})$, where $A \in \text{GL}_n(\mathbb{F})$ and $\bar{b} \in \mathbb{F}^n$. Then, for all but at most n^2 values of $\beta \in \mathbb{F}$, we have*

$$\dim \partial_{\bar{x} < \infty}(\mathcal{H}_{\min, \bar{a}(\beta)}(f)) \leq s.$$

Write $\bar{u} = A\bar{a}(\beta) + \bar{b}$, so that $f(\bar{x} + \bar{a}(\beta)) = p(A\bar{x} + \bar{u})$. When every coordinate of $\bar{u}$ is nonzero, Lemma 3.2 bounds the derivative dimension of the initial form of $p(\bar{y} + \bar{u})$ by s , and invertible linear substitution preserves this bound. Each coordinate of $\bar{u}$ is a nonzero polynomial in β of degree at most n , excluding at most n^2 values altogether.

We prove Theorem 1.2 by combining this lemma with the hitting sets of Bhardwaj and Saptharishi [6] for polynomials with small partial derivative spaces. By Lemma 1.6 we have that if the initial form of f at $\bar{a}(\beta)$ has partial derivative dimension at least r for more than n^2 values of $\beta \in \mathbb{F}$, then $\text{osp}_{\mathbb{F}}(f) \geq r$. Therefore, to obtain the lower bound in Theorem 1.5, it is enough to lower bound the dimension of partial derivatives of the initial form of C_n^k .

Remark 1.7. *Note that, unlike Forbes's cone-size bound, Lemma 1.6 holds over fields of arbitrary characteristic. For example, in characteristic p , the 2^n -sparse polynomial $\prod_{i=1}^n (y_i^p - 1)$ becomes $\prod_{i=1}^n y_i^p$ after a shift by $\bar{1}$; its trailing monomial has cone size $(p+1)^n > 2^n$, while its derivative dimension is 2^n . The proof lifts a basis of nonzero Hasse derivatives to correlated shifted partials; the comparison of their lowest-degree terms uses only that the shift coordinates are nonzero.*

Our PIT theorem uses the characteristic-zero hitting set of Bhardwaj and Saptharishi [6]. Since our structural lemma holds over every field, an extension of their hitting set to polynomials of

bounded Hasse-derivative dimension in positive characteristic would yield a corresponding sparse-orbit PIT result. $\diamond$

Applications: hitting sets for orbits of other classes. Saha and Thankey [26] reduce PIT for orbits of several classes to sparse-orbit PIT, using the Jacobian technique of [4, 1]; their base case is a quasipolynomial hitting set for sparse orbits. By using the hitting sets generated by $\Gamma_{n,s}$ in the base case of their argument, we obtain the following (Section 6).

Theorem 1.8 (Applications, informal). *There are explicit hitting sets of size polynomial in the number of variables, the size, and the degree for affine orbits of*

1. constant-depth, constant-occur formulas: *constant-depth formulas whose leaves are sparse polynomials and in which each variable occurs in a constant number of leaves* [26, Definition 3];
2. circuits over sparse polynomials of bounded transcendence degree: *polynomials $C(h_1, \dots, h_m)$, for an arbitrary polynomial C and sparse inputs h_i whose transcendence degree is bounded by a constant* [4].

These improve the quasipolynomial bounds of Saha and Thankey [26, Theorem 10 and Remark 1].

1.2 Proof overview

Partial derivatives and shared core of sparse orbit. Given the recent progress on hitting sets for polynomials of bounded partial derivatives [6], one might hope to directly reduce the sparse affine orbits problem to the bounded partial derivatives one. However, there are two immediate obstacles. Firstly, as noted before, orbits of sparse polynomials can contain dense polynomials. Secondly, there is no general bound on partial derivatives in terms of sparsity. In fact, a sparse polynomial can have a large derivative space: $x_1 \cdots x_n$ has dimension 2^n . As in [13, 7], here the key is to consider shifted partial derivatives for an appropriate choice of shift.

Recall that $f(x) = p(A\bar{x} + \bar{b})$, and we may write $f(\bar{x} + \bar{a}) = p(\bar{u} + A\bar{x})$ with hidden shift $\bar{u} = A\bar{a} + \bar{b}$. The correlated shifted partials $(\bar{u} + \bar{y})^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}}$ map each translated monomial $(\bar{u} + \bar{y})^{\bar{\alpha}}$ to a multiple of itself. Therefore the span of the correlated shifted partials of $p(\bar{u} + \bar{y})$ is of dimension at most s . If $\bar{u}$ has nonzero coordinates, the multiplier has nonzero constant term $\bar{u}^{\bar{\beta}}$. Hence we have

$$(\bar{u} + \bar{y})^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} p(\bar{u} + \bar{y}) = \bar{u}^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} h_m + (\text{terms of higher degree}),$$

where h_m is the initial form of $p(\bar{u} + \bar{y})$; so whenever $\partial_{\bar{y}^{\bar{\beta}}} h_m \neq 0$, it is, up to the nonzero scalar $\bar{u}^{\bar{\beta}}$, the lowest-degree component of the correlated shifted partial. Choosing derivative bases order by order and comparing lowest-degree components shows that their lifts are linearly independent. Since the correlated shifted partials span a space of dimension at most s , the derivative dimension of h_m is at most s . For $\bar{a} = \bar{a}(\beta) = (\beta, \beta^2, \dots, \beta^n)$, each coordinate of $\bar{u}$ is a nonzero polynomial of degree at most n in β , so at most n^2 values are excluded.

Identity testing. With B and G as in Theorem 1.2, we use the set

$$\{\bar{a}(\beta) + t^B G(t) : \beta \in S, t \in T\}, \quad |S| = n^2 + 1, \quad |T| = 2dB + 1.$$

Choose a shift $\bar{a} = \bar{a}(\beta)$ satisfying Lemma 1.6 and write $f(\bar{x} + \bar{a}) = h_m + \dots + h_d$ with $h_m \neq 0$. Then $\dim \partial_{\bar{x} < \infty}(h_m) \leq s$; the Bhardwaj–Saptharishi bound then gives $h_m(G(t))$ a nonzero coefficient below degree B . We cannot apply their theorem directly to f , whose derivative space need not be small. A simple rescaling passes from h_m to f :

$$f(\bar{a} + t^B G(t)) = \sum_{j=m}^d t^{Bj} h_j(G(t)).$$

The substitution multiplies $h_j(G(t))$ by t^{Bj} . Thus the chosen coefficient from $h_m(G(t))$ stays below degree $B(m+1)$, while every term from $j > m$ has degree at least $B(m+1)$ and cannot cancel it. The resulting nonzero univariate has degree at most $2dB$ and is hit by T . The algorithm is oblivious to $A, \bar{b}$, and m .

Lower bound. By Lemma 1.6, it suffices to make the initial forms of f at the shifts $\bar{a}(\beta)$ have large derivative dimension. Powers of a full-rank quadratic form have maximal derivative dimension, so we seek a cubic C_n whose initial form at $\bar{a}(\beta)$ is a quadratic of rank $n-1$ for all but finitely many β . This needs C_n and $\nabla C_n = (\partial C_n / \partial x_1, \dots, \partial C_n / \partial x_n)$ to vanish at every $\bar{a}(\beta)$, and its Hessian there to have rank $n-1$ for all but finitely many β . Each 3×3 matrix in (1.4) has rank one at every $\bar{a}(\beta)$, which forces both the value and the gradient to vanish. The Hessian annihilates the tangent $\bar{a}'(\beta)$, and a minor nonzero at $\beta = 0$ gives rank exactly $n-1$ for all but finitely many β . With Q_β the quadratic initial form, $\mathcal{H}_{\min, \bar{a}(\beta)}(C_n^k) = Q_\beta^k$, and hence

$$\text{osp}(C_n^k) \geq R_{\max}(n-1, 2k) = \binom{n-1+k}{n-1} + \binom{n-2+k}{n-1}. \quad (1.9)$$

Here $R_{\max}$ is the maximum homogeneous derivative dimension, recorded in Section 5.

1.3 Related work

For general algebraic circuits, both PIT and lower bounds remain wide open. The strongest lower bound for general circuits is still the $\Omega(n \log d)$ bound of Baur and Strassen [3], and for algebraic branching programs it is $\Omega(n^2)$ [10]. For constant-depth circuits, Limaye, Srinivasan, and Tavenas [22] proved the first superpolynomial lower bounds, extended to all fields by Forbes [14], and obtained the first deterministic subexponential-time PIT for this class. Deterministic polynomial-time PIT is known only for restricted classes, sparse polynomials being the most basic.

For sparse polynomials, Klivans and Spielman [19] gave a generator whose seed length grows logarithmically with the degree. Degree-independent deterministic hitting sets go back to Ben-Or and Tiwari [5], and over $\mathbb{R}$ Bläser and Engels [9] gave randomness-efficient tests of unbounded degree. Our degree-independent result over $\mathbb{R}$ (Theorem 4.4) combines our generator with the real-root bound of Koiran, Portier, and Tavenas [20] for sums of products of powers, which does not depend on the exponents.

For orbits of sparse polynomials, Saha and Thankey [26] gave hitting sets of size $(nd)^{O(d \log s)}$ for individual degree d , via commutative ROABPs, and Medini and Shpilka [23] removed the individual-degree restriction, obtaining $(nd)^{O(\log s)}$ by taking directional derivatives and restrictions until the polynomial becomes a product of linear forms. Bhargava and Ghosh [7] observed that a shift by $(t, t^2, \dots, t^n)$ makes every hidden coordinate nonzero, so that Forbes’s low-support bound [13] applies in the hidden coordinates. Our key lemma uses the same shift but keeps the whole initial form. Its derivative dimension, unlike the support or cone size of a monomial, is invariant under invertible

linear substitutions, which is what lets us apply the hitting sets of Bhardwaj and Saptharishi [6]. ROABPs are a natural next target: every algebraic branching program of size s is an affine projection of $\text{IMM}_{s,s}$, which has a linear-size ROABP, affine projections of constant-width ROABPs capture all formulas [26], and affine projections of the determinant give all algebraic branching programs. For orbits of ROABPs, Saha and Thankey [26] gave hitting sets constructible in time $(nd)^{O(d \log w)}$ for width- w commutative ROABPs of individual degree d , and in time $n^{O(w^6 \log n)}$ for multilinear width- w ROABPs. Using low-cone rank concentration, Bhargava and Ghosh [7] improved the dependence on the individual degree, obtaining time $(ndw)^{O(\min\{w^2, d \log w\})}$ for orbits of any-order ROABPs and $(ndw)^{O(w^2 \log n \cdot \min\{w^2, d \log w\})}$ for orbits of ROABPs. To our knowledge, polynomial-size hitting sets for orbits of constant-width ROABPs, and better dependence on the width and the individual degree, remain open.

On the lower-bound side, Nisan and Wigderson [24] and Kayal, Saha, and Tavenas [18] prove $2^{\Omega(n)}$ lower bounds for homogeneous depth-three circuits, and Kayal, Saha, and Tavenas ask for bounds that grow with the degree. Theorem 1.5 grows with the degree, but only for the restricted model (1.1): our hard polynomials have small nonhomogeneous depth-three circuits (Appendix A). Garcia-Marco, Koiran, Pecatte, and Thomassé [15] study the complexity of partial derivative dimension itself, and rank methods such as partial derivatives have inherent limits [11].

2 Notation and preliminaries

Unless otherwise stated, $\mathbb{F}$ has characteristic zero and running times count field operations. Write $[n] = \{1, \dots, n\}$, $\mathbb{N} = \{0, 1, \dots\}$, $\|\bar{\alpha}\|_1 = \sum_i \alpha_i$, $\bar{x}^{\bar{\alpha}} = \prod_i x_i^{\alpha_i}$, and $\bar{\alpha}! = \prod_i \alpha_i!$.

Following Forbes [12], the Hasse derivative $\partial_{\bar{x}\bar{\alpha}} f$ is the coefficient of $\bar{z}^{\bar{\alpha}}$ in $f(\bar{x} + \bar{z})$, and

$$\partial_{\bar{x} < \infty}(f) = \text{span}_{\mathbb{F}}\{\partial_{\bar{x}\bar{\alpha}} f : \bar{\alpha} \in \mathbb{N}^n\}.$$

The subscript $\bar{x} = k$ restricts to $\|\bar{\alpha}\|_1 = k$. Ordinary derivatives differ by the nonzero factors $\bar{\alpha}!$, so they span the same spaces [12]. Write $\dim \partial_{\bar{x} < \infty}(h)$ for the dimension of the Hasse partial derivative space of h .

Let $\mathcal{H}_j(f)$ denote the degree- j homogeneous component. For $f \neq 0$, we call its lowest-degree nonzero homogeneous component its *initial form*, denoted $\mathcal{H}_{\min}(f)$. Thus, if $f = h_m + \dots + h_d$, where $h_j = \mathcal{H}_j(f)$ and $h_m \neq 0$, then

$$\mathcal{H}_{\min}(f) = h_m, \quad \mathcal{H}_{\min, \bar{a}}(f) := \mathcal{H}_{\min}(f(\bar{x} + \bar{a})).$$

In particular, $\mathcal{H}_{\min, \bar{a}}(f) = f(\bar{a})$ if $f(\bar{a}) \neq 0$. We write $\bar{a}(\beta) = (\beta, \beta^2, \dots, \beta^n)$.

For a polynomial p , write $\text{sp}(p)$ for its number of nonzero monomials and define

$$\text{osp}_{\mathbb{F}}(f) = \min_{A \in \text{GL}_n(\mathbb{F}), \bar{b} \in \mathbb{F}^n} \text{sp}(f(A\bar{x} + \bar{b})).$$

Equivalently, $\text{osp}_{\mathbb{F}}(f)$ is the least sparsity of a polynomial p in a representation $f(\bar{x}) = p(A\bar{x} + \bar{b})$ with A invertible. We suppress the field when it is clear. For $\deg f \leq d$,

$$\text{osp}(f) \leq \text{sp}(f) \leq \binom{n+d}{n}. \quad (2.1)$$

The coordinate changes throughout the paper are invertible and preserve the number of variables. Circuit conventions are specified in Appendix A.

A set $H \subseteq \mathbb{F}^n$ is a *hitting set* for a class $\mathcal{C} \subseteq \mathbb{F}[\bar{x}]$ if every nonzero $f \in \mathcal{C}$ satisfies $f(\bar{a}) \neq 0$ for some $\bar{a} \in H$. A polynomial map $\Gamma \in \mathbb{F}[\bar{w}]^n$ is a *hitting-set generator* for $\mathcal{C}$ if $f(\Gamma) \neq 0$ for every

nonzero $f \in \mathcal{C}$. If Γ has q parameters and coordinate degree at most E , then evaluating it on a grid T^q with $|T| > DE$ gives a hitting set for the members of $\mathcal{C}$ of degree at most D .

Invertible linear substitutions preserve derivative dimension: since $\partial_{x_i}[f(A\bar{x})] = \sum_j A_{ji}(\partial_{y_j} f)(A\bar{x})$, the order- k derivatives of $f(A\bar{x})$ lie in the span of $\{(\partial_{\bar{y}^\gamma} f)(A\bar{x}) : \|\bar{\gamma}\|_1 = k\}$, and applying this to A^{-1} gives equality.

Field extension also preserves derivative dimension, since it is the rank of a finite coefficient matrix.

3 Initial forms of translated sparse polynomials

This section proves Lemma 1.6. Following Forbes [13, Definition 5.3], write

$$((\bar{y} + \bar{u}) \circ \partial_{\bar{y}})^{<\infty}(g) = \{(\bar{y} + \bar{u})^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} g : \bar{\beta} \in \mathbb{N}^n\}.$$

The multiplier and the derivative are indexed by the same exponent vector: the correlated shifted partial associated with $\bar{\beta}$ is $(\bar{y} + \bar{u})^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} g$. We therefore call the elements of this family the *correlated shifted partials of g at $\bar{u}$* .

Lemma 3.1 (Dimension of correlated shifted partials (Forbes [13, Corollary 5.6])). *Over any field, for every $p \in \mathbb{F}[\bar{y}]$ and $\bar{u} \in \mathbb{F}^n$,*

$$\dim \text{span}_{\mathbb{F}} \left(((\bar{y} + \bar{u}) \circ \partial_{\bar{y}})^{<\infty}(p(\bar{u} + \bar{y})) \right) \leq \text{sp}(p).$$

Proof. For a translated monomial, the Hasse derivative formula gives

$$(\bar{u} + \bar{y})^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} (\bar{u} + \bar{y})^{\bar{\alpha}} = \binom{\bar{\alpha}}{\bar{\beta}} (\bar{u} + \bar{y})^{\bar{\alpha}}.$$

Thus its correlated derivative space is one-dimensional, as in [13, Lemma 5.5]. Writing $p = \sum_{\bar{\alpha} \in S} c_{\bar{\alpha}} \bar{y}^{\bar{\alpha}}$ and using linearity gives

$$\text{span}_{\mathbb{F}} \left(((\bar{y} + \bar{u}) \circ \partial_{\bar{y}})^{<\infty}(p(\bar{u} + \bar{y})) \right) \subseteq \sum_{\bar{\alpha} \in S} \text{span}_{\mathbb{F}} \{(\bar{u} + \bar{y})^{\bar{\alpha}}\}.$$

The dimension is therefore at most $|S|$. □

Lemma 3.2 (Derivative dimension of the initial form). *Over any field, let $0 \neq p \in \mathbb{F}[\bar{y}]$ and $\bar{u} \in (\mathbb{F}^\times)^n$. Write*

$$g(\bar{y}) = p(\bar{u} + \bar{y}) = h_m + h_{m+1} + \cdots + h_d, \quad h_i = \mathcal{H}_i(g), \quad h_m \neq 0.$$

Then

$$\dim \boldsymbol{\partial}_{\bar{y}^{<\infty}}(h_m) \leq \text{sp}(p).$$

Proof. Put $s = \text{sp}(p)$ and $V = \text{span}_{\mathbb{F}} \left(((\bar{y} + \bar{u}) \circ \partial_{\bar{y}})^{<\infty}(g) \right)$. By Lemma 3.1, $\dim V \leq s$. Suppose, for a contradiction, that $\dim \boldsymbol{\partial}_{\bar{y}^{<\infty}}(h_m) > s$.

For each order $0 \leq k \leq m$, let

$$W_k = \text{span}_{\mathbb{F}} \{ \partial_{\bar{y}^{\bar{\beta}}} h_m : \|\bar{\beta}\|_1 = k \}, \quad r_k = \dim W_k.$$

The spaces W_k have distinct homogeneous degrees $m-k$, so $\sum_{k=0}^m r_k = \dim \boldsymbol{\partial}_{\bar{y} < \infty}(h_m) > s$. Choose multiindices $\bar{\beta}_{k,1}, \dots, \bar{\beta}_{k,r_k}$ of order k such that

$$\partial_{\bar{y}^{\bar{\beta}_{k,1}}} h_m, \dots, \partial_{\bar{y}^{\bar{\beta}_{k,r_k}}} h_m$$

form a basis of W_k . Define their lifts by

$$F_{k,j} = (\bar{u} + \bar{y})^{\bar{\beta}_{k,j}} \partial_{\bar{y}^{\bar{\beta}_{k,j}}} g \in V.$$

The multiplier has nonzero constant term $\bar{u}^{\bar{\beta}_{k,j}}$. Consequently $F_{k,j}$ has no terms of degree below $m-k$, and

$$\mathcal{H}_{m-k}(F_{k,j}) = \underbrace{\bar{u}^{\bar{\beta}_{k,j}}}_{\neq 0} \partial_{\bar{y}^{\bar{\beta}_{k,j}}} h_m.$$

There are $\sum_k r_k > s$ lifts in the space V of dimension at most s . They therefore satisfy a nontrivial relation

$$\sum_{k=0}^m \sum_{j=1}^{r_k} c_{k,j} F_{k,j} = 0.$$

Let K be the largest order for which some $c_{K,j} \neq 0$. Taking the homogeneous component of degree $m-K$ gives

$$\sum_{j=1}^{r_K} c_{K,j} \bar{u}^{\bar{\beta}_{K,j}} \partial_{\bar{y}^{\bar{\beta}_{K,j}}} h_m = 0.$$

All terms with $k < K$ start in higher degrees and hence disappear. But the displayed derivatives form a basis of W_K , and every $\bar{u}^{\bar{\beta}_{K,j}}$ is nonzero. Thus all $c_{K,j}$ must be zero, a contradiction. It follows that $\dim \boldsymbol{\partial}_{\bar{y} < \infty}(h_m) \leq s$. $\square$

Proof of Lemma 1.6. For $\bar{a} = \bar{a}(\beta)$, put $\bar{u} = A\bar{a} + \bar{b}$. Each coordinate $b_i + \sum_{j=1}^n A_{ij}\beta^j$ is a nonzero polynomial of degree at most n , since row i of A is nonzero. Thus at most n^2 parameters make any coordinate of $\bar{u}$ zero. At every other parameter, Lemma 3.2 gives $\dim \boldsymbol{\partial}_{\bar{y} < \infty}(\mathcal{H}_{\min}(p(\bar{u} + \bar{y}))) \leq \text{sp}(p)$. The initial form of $f(\bar{a} + \bar{x}) = p(\bar{u} + A\bar{x})$ is $\mathcal{H}_{\min}(p(\bar{u} + \bar{y}))|_{\bar{y}=A\bar{x}}$. Invariance under invertible linear substitution proves the claim. $\square$

The nonzero-coordinate hypothesis is essential: at zero, the single monomial $y_1 \cdots y_n$ is its own initial form and has derivative dimension 2^n .

4 A hitting set for sparse affine orbits

By Lemma 1.6, for some β among any $n^2 + 1$ values, the initial form of f at $\bar{a}(\beta)$ has derivative dimension at most s . We apply the following Bhardwaj–Saptharishi multiplicity bound at that point.

Theorem 4.1 (Bhardwaj–Saptharishi [6]). *Let $r \geq 1$ and $B = nr^2 + r$. Set*

$$G_i(t) = \sum_{j=1}^B \frac{i^j t^j}{j}, \quad i \in [n].$$

If $0 \neq h \in \mathbb{F}[x_1, \dots, x_n]$ and $\dim \boldsymbol{\partial}_{\bar{x} < \infty}(h) \leq r$, then the polynomial $h(G_1(t), \dots, G_n(t))$ has a nonzero coefficient in some degree below B .

The cited theorem first proves the assertion for the formal series $-\log(1-it)$. Truncating these series through degree B preserves all coefficients in degrees below B .

Proof of Theorem 1.2. Put $B = ns^2 + s$, let G be as in Theorem 4.1, and choose sets $S, T \subseteq \mathbb{F}$ with $|S| = n^2 + 1$ and $|T| = 2dB + 1$. We claim that

$$\mathcal{H}_{n,d,s} = \{\bar{a}(\beta) + t^B G(t) : \beta \in S, t \in T\} \quad (4.2)$$

is a hitting set. Let $0 \neq f = p(A\bar{x} + \bar{b})$ be in the promised class. By Lemma 1.6, some $\beta_0 \in S$ gives a shift $\bar{a} = \bar{a}(\beta_0)$ at which the expansion

$$f(\bar{a} + \bar{x}) = h_m(\bar{x}) + \cdots + h_d(\bar{x}), \quad h_m \neq 0,$$

has $\dim \boldsymbol{\theta}_{\bar{x} < \infty}(h_m) \leq s$, where $h_j = \mathcal{H}_j(f(\bar{a} + \bar{x}))$. Since each h_j is homogeneous of degree j , we have

$$f(\bar{a} + t^B G(t)) = \sum_{j=m}^d t^{Bj} h_j(G(t)).$$

By Theorem 4.1, $h_m(G(t))$ has a nonzero coefficient in some degree $e < B$. The $j = m$ summand therefore contributes a nonzero coefficient in degree $Bm + e < B(m+1)$. Every summand with $j > m$ has all its terms in degrees at least $B(m+1)$, so this coefficient cannot cancel. The resulting univariate is nonzero and has degree at most $2dB$, since every substituted coordinate has degree at most $2B$. Some $t \in T$ gives a nonzero evaluation.

Taking $S = \{0, \dots, n^2\}$ and $T = \{0, \dots, 2dB\}$ gives the required explicit construction. The coordinates are rational numbers of polynomial bit length in n, d, s , and their construction uses polynomially many field operations. $\square$

4.1 A degree-free generator

The proof of Theorem 1.2 gives the following generator form.

Corollary 4.3 (Sparse-affine generator). *For $n, s \geq 1$, put $B_s = ns^2 + s$ and define*

$$\Gamma_{n,s,i}(\beta, t) = \beta^i + t^{B_s} \sum_{j=1}^{B_s} \frac{i^j t^j}{j}, \quad i \in [n],$$

and $\Gamma_{n,s} = (\Gamma_{n,s,1}, \dots, \Gamma_{n,s,n})$. For every nonzero $p \in \mathbb{F}[\bar{x}]$ with $\text{sp}(p) \leq s$, every $A \in \text{GL}_n(\mathbb{F})$, and every $\bar{b} \in \mathbb{F}^n$, $p(A\Gamma_{n,s}(\beta, t) + \bar{b}) \neq 0$. There is no degree restriction on p .

Proof. With $f = p(A\bar{x} + \bar{b})$ and G as in Theorem 4.1 with $r = s$, we have $p(A\Gamma_{n,s}(\beta, t) + \bar{b}) = f(\bar{a}(\beta) + t^{B_s} G(t))$. The proof of Theorem 1.2 shows that this is a nonzero polynomial in t for all but at most n^2 values of β ; it uses $\deg p \leq d$ only to bound the degree in t . $\square$

Over $\mathbb{R}$, the degree-free generator combines with a real-root bound to give a hitting set that works for every degree.

Theorem 4.4 (Degree-independent hitting set over $\mathbb{R}$). *Let $N = n^2 + 1$, $B = ns^2 + s$, $M = B!$, and $L = \lfloor \frac{2}{3}s^3nB + 4snB + s \rfloor + 1$. The set*

$$\mathcal{H} = \{M\Gamma_{n,s}(\beta, t) : \beta \in [N], t \in [L]\} \subseteq \mathbb{Z}_{>0}^n$$

has size $O(n^4 s^5)$, its coordinates have $O(ns^2 \log(ns))$ bits, and it hits every nonzero $p(A\bar{x} + \bar{b})$ with $p \in \mathbb{R}[\bar{y}]$, $\text{sp}(p) \leq s$, $A \in \text{GL}_n(\mathbb{R})$, and $\bar{b} \in \mathbb{R}^n$, of arbitrary degree.

Proof. Let $0 \neq f = p(A\bar{x} + \bar{b})$, so $f(M\bar{x}) = p(MA\bar{x} + \bar{b})$. As in the proof of Lemma 1.6, some $\beta_0 \in [N]$ makes every coordinate of $MA\bar{a}(\beta_0) + \bar{b}$ nonzero, and the proof of Theorem 1.2 shows that $F(t) = f(M\Gamma_{n,s}(\beta_0, t))$ is nonzero. With $q_i(t) = (MA\Gamma_{n,s}(\beta_0, t) + \bar{b})_i$, of degree at most $2B$, $F = \sum_{\bar{\alpha}} c_{\bar{\alpha}} \prod_i q_i^{\alpha_i}$ has at most s terms. By [20, Theorem 13], F has fewer than L real roots, so $F(t) \neq 0$ for some $t \in [L]$. Since $j \mid B!$ for $j \leq B$, all coordinates are integers; the size and bit bounds are direct. $\square$

For sparse polynomials themselves, degree-independent hitting sets follow from the Ben-Or–Tiware argument: the points $(p_1^k, \dots, p_n^k)$, $0 \leq k < s$, for the first n primes p_i , suffice [5]; see also [9]. Theorem 4.4 is a basis-independent version of these degree-independent hitting sets over $\mathbb{R}$.

4.2 Extension

The same construction handles sparse polynomials whose coefficients have small derivative spaces. We enlarge the correlated space in the proof of Lemma 1.6 and reuse its lifting argument.

Corollary 4.5 (Structured coefficients). *Let $\mathbb{F}$ have characteristic zero, $0 \leq k \leq n$, and $r \geq 1$. Partition the hidden variables as $\bar{y} \in \mathbb{F}^k$ and $\bar{z} \in \mathbb{F}^{n-k}$. Suppose*

$$p(\bar{y}, \bar{z}) = \sum_{\bar{\alpha} \in S} \bar{y}^{\bar{\alpha}} q_{\bar{\alpha}}(\bar{z}), \quad \sum_{\bar{\alpha} \in S} \dim \boldsymbol{\theta}_{\bar{z} < \infty}(q_{\bar{\alpha}}) \leq r, \quad \deg p \leq d.$$

Put $B = nr^2 + r$ and $G = (G_1, \dots, G_n)$ with $G_i(t) = \sum_{j=1}^B i^j t^j / j$ for $i \in [n]$. Choose $S_0, T \subseteq \mathbb{F}$ of sizes $nk + 1$ and $2dB + 1$. Then

$$\{\bar{a}(\beta) + t^B G(t) : \beta \in S_0, t \in T\}, \quad \bar{a}(\beta) = (\beta, \dots, \beta^n),$$

hits every nonzero polynomial $p(A\bar{x} + \bar{b})$ of this form, for $A \in \text{GL}_n(\mathbb{F})$ and $\bar{b} \in \mathbb{F}^n$. Its size is at most $(nk + 1)(2d(nr^2 + r) + 1)$.

Proof. Assume $p \neq 0$. At a hidden shift $(\bar{u}, \bar{v})$ with every $u_i \neq 0$, put $g = p(\bar{u} + \bar{y}, \bar{v} + \bar{z})$ and let h_m be its lowest nonzero homogeneous component, of degree m . Use the space

$$V = \sum_{\bar{\alpha} \in S} (\bar{u} + \bar{y})^{\bar{\alpha}} \boldsymbol{\theta}_{\bar{z} < \infty}(q_{\bar{\alpha}}(\bar{v} + \bar{z})).$$

It has dimension at most r and contains all mixed lifts $(\bar{u} + \bar{y})^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} \partial_{\bar{z}^{\bar{\gamma}}} g$. For $j = \|\bar{\beta}\|_1 + \|\bar{\gamma}\|_1 \leq m$, their lowest possible component is $\bar{u}^{\bar{\beta}} \partial_{\bar{y}^{\bar{\beta}}} \partial_{\bar{z}^{\bar{\gamma}}} h_m$ in degree $m - j$. The basis-lifting argument in the proof of Lemma 3.2 therefore gives $\dim \boldsymbol{\theta}_{(\bar{y}, \bar{z}) < \infty}(h_m) \leq r$. Only k hidden coordinates must be nonzero, so the same root count excludes at most nk values of β . Apply the preceding hitting-set proof with r in place of s and $nk + 1$ shifts. $\square$

The cases $k = n$ and $k = 0$ recover, respectively, sparse affine orbits and polynomials with small partial derivative spaces.

5 The sparse-orbit lower bound

By Lemma 1.6, if the initial form of $f(\bar{x} + \bar{a}(\beta))$ has derivative dimension at least r outside a finite set of parameters, then $\text{osp}(f) \geq r$. For the cubic C_n in (1.4), we will compute this derivative

dimension using the following formula. For $N, m \geq 1$, define

$$R_{\max}(N, m) = \sum_{j=0}^m \binom{N-1 + \min(j, m-j)}{\min(j, m-j)}.$$

We recall standard facts about quadratic forms. A quadratic form in N variables can be written uniquely as $Q(\bar{z}) = \bar{z}^T M \bar{z}$ with M symmetric. Its *rank* is $\text{rank } M$, and it is *nondegenerate* if $\text{rank } M = N$. Its Hessian (the matrix of ordinary second derivatives) is $2M$, so quadratic rank and Hessian rank agree. Diagonalizing M shows that, over an algebraically closed field of characteristic zero, Q is linearly equivalent to $z_1^2 + \cdots + z_{\text{rank } M}^2$. Equality for quadratic powers in the following lemma follows from Reznick's maximal-rank theorem [25, Theorem 8.15], restated and proved in [16, Theorem 2.2 and Remark 2.3].

Lemma 5.1 (Homogeneous derivative dimension). *Every degree- m homogeneous polynomial h in N variables satisfies $\dim \boldsymbol{\partial}_{\bar{z} < \infty}(h) \leq R_{\max}(N, m)$. If Q is a nondegenerate quadratic form in N variables and $k \geq 1$, then*

$$\dim \boldsymbol{\partial}_{\bar{z} < \infty}(Q^k) = R_{\max}(N, 2k) = \binom{N+k}{N} + \binom{N+k-1}{N}. \quad (5.2)$$

The cited result is stated over $\mathbb{C}$. The full derivative ranks of $(\sum_i z_i^2)^k$ are witnessed by nonzero rational minors, which remain nonzero over every characteristic-zero field. Every nondegenerate Q is linearly equivalent to $\sum_i z_i^2$ over an algebraic closure; field extension and invertible linear substitution preserve derivative dimension, giving the formulation above.

Proof of Theorem 1.5. At $\bar{x} = \bar{a}(\beta)$, the j th matrix in (1.4) is

$$\begin{pmatrix} 1 \\ \beta \\ \beta^{j-1} \end{pmatrix} (1 \quad \beta \quad \beta^{n+1-j}).$$

It has rank one. Its determinant and all its first derivatives vanish, because its 2×2 cofactors vanish. Thus C_n and its gradient vanish at $\bar{a}(\beta)$ for every β .

Let $\mathcal{A}(\beta)$ be the Hessian of C_n at $\bar{a}(\beta)$. Since C_n has degree three, Taylor expansion at $\bar{a}(\beta)$ gives

$$C_n(\bar{a}(\beta) + \bar{x}) = \underbrace{C_n(\bar{a}(\beta))}_{=0} + \underbrace{\nabla C_n(\bar{a}(\beta)) \cdot \bar{x}}_{=0} + \underbrace{\frac{1}{2} \bar{x}^T \mathcal{A}(\beta) \bar{x}}_{=: Q_\beta(\bar{x})} + \mathcal{H}_3(C_n)(\bar{x}).$$

The constant and linear terms vanish because C_n and ∇C_n vanish at $\bar{a}(\beta)$, and the cubic term is the top-degree part of C_n , since the third derivatives of a cubic are constant.

Thus $\text{rank } Q_\beta = \text{rank } \mathcal{A}(\beta)$, and whenever $Q_\beta \neq 0$ it is the initial form, since the constant and linear components vanish; then the initial form of C_n^k at $\bar{a}(\beta)$ is Q_β^k , a power of a quadratic. To apply Lemma 5.1 in $n-1$ variables, we show that this rank is $n-1$ outside a finite set.

At zero, the quadratic component is

$$Q_0 = (n-3)x_2x_n - \sum_{j=3}^{n-1} x_jx_{n+2-j}. \quad (5.3)$$

Its Hessian on the variables $x_2, \dots, x_n$ is antidiagonal, with endpoint entries $n-3$ and all interior entries -2 . It is invertible. Hence the minor of $\mathcal{A}(\beta)$ on rows and columns $2, \dots, n$ is a polynomial

in β that is nonzero at $\beta = 0$, so it vanishes for only finitely many β . On the other hand, the identity $\nabla C_n(\bar{a}(\beta)) = \bar{0}$ holds in $\mathbb{F}[\beta]^n$; differentiating each coordinate with respect to β by the chain rule gives

$$\mathcal{A}(\beta)\bar{a}'(\beta) = \bar{0}, \quad \bar{a}'(\beta) = (1, 2\beta, \dots, n\beta^{n-1}) \neq \bar{0}.$$

Thus $\mathcal{A}(\beta)$ has rank exactly $n - 1$ outside that finite set. At each such parameter the initial form of $C_n(\bar{x} + \bar{a}(\beta))$ is a quadratic form Q_β of rank $n - 1$. For nonzero polynomials u, v , the product of their lowest nonzero components is nonzero, so $\mathcal{H}_{\min}(uv) = \mathcal{H}_{\min}(u)\mathcal{H}_{\min}(v)$. Consequently

$$\mathcal{H}_{\min, \bar{a}(\beta)}(C_n^k) = Q_\beta^k.$$

By Lemma 5.1, its derivative dimension is $R_{\max}(n - 1, 2k)$: a rank- $(n - 1)$ quadratic becomes nondegenerate in $n - 1$ variables after an invertible linear change. Lemma 1.6 therefore proves the lower bound.

Each determinant has degree at most three. The coefficient of $x_1^2 x_n$ in each summand is -1 , and no other term contributes to this monomial. Hence its coefficient in C_n is $-(n - 3) \neq 0$. Therefore $\deg C_n = 3$ and $\deg C_n^k = 3k$. $\square$

The n^n specialization. Put $N = n - 1$ and $k = n^2$. The arithmetic–geometric mean inequality gives $N! \leq ((N + 1)/2)^N = (n/2)^N$, so

$$\text{osp}(C_n^{n^2}) \geq \binom{N + n^2}{N} \geq \frac{n^{2N}}{N!} \geq (2n)^{n-1} \geq n^n.$$

Why degree three suffices. We need shifts at which $\mathcal{H}_0$ and $\mathcal{H}_1$ of $C_n(\bar{x} + \bar{a}(\beta))$ vanish, that is, C_n and ∇C_n vanish at $\bar{a}(\beta)$ for every β . No nonzero polynomial of degree at most two does this: each first derivative is affine-linear and vanishes at $\bar{a}(\beta)$ for every β , hence is zero, so the polynomial is a constant vanishing at these points. A 3×3 determinant vanishes with its gradient at rank-one matrices, which gives the required behaviour in degree three.

Dependence on the degree. For a degree budget $D \geq 12$, taking $k = \lfloor D/3 \rfloor \geq D/4$ gives an explicit polynomial of degree at most D with $\text{osp} \geq \binom{n-1+k}{n-1} \geq (D/(4n))^{n-1}$. For $D = n^c$ with $c > 1$ this is $n^{(c-1+o(1))n}$, matching the exponent of the trivial upper bound $\binom{n+D}{n} \leq (e(1+D/n))^n$.

Remark 5.4. Let $M = \binom{n+D}{n}$, with $n \geq 2$ and $D \geq 3$. In a fixed coordinate system, a polynomial of degree at most D can have all M monomials. Orbit sparsity can be smaller, since any affine coordinate system may be used. For example, diagonalizing the quadratic homogeneous component by an invertible linear change reduces its $\binom{n+1}{2}$ possible monomials to at most n squares. Hence, over any field of characteristic zero,

$$\max_{\deg f \leq D} \text{osp}(f) \leq M - \binom{n}{2}.$$

The saving cannot be much larger. For simplicity, we argue over finite fields: counting sparse polynomials and affine changes of coordinates, absorbing diagonal rescalings into the polynomial, shows that for every sufficiently large odd prime q ,

$$\max_{\deg f \leq D} \text{osp}_{\mathbb{F}_q}(f) \geq M - n^2.$$

We leave as an open problem to construct explicit polynomials whose orbit sparsity matches these bounds. $\diamond$

6 Applications to orbit PIT

Sparse PIT is a building block for PIT of many circuit classes. Sparse-orbit PIT plays a similar role for affine orbits. We illustrate this through two applications: polynomial combinations of sparse inputs of bounded transcendence degree, and constant-depth formulas in which each variable occurs in only constantly many sparse leaves.

Both results in this section use the algebraic-independence method of Beecken, Mittmann, and Saxena [4] and Agrawal, Saha, Saptharishi, and Saxena [1], which Saha and Thankey [26] adapted to affine orbits. The work of Agrawal et al. unified and extended polynomial-time black-box PIT for several central models, including bounded-top-fan-in depth-three circuits and constant-depth constant-read multilinear formulas. In each case the only new ingredient is the generator of Corollary 4.3, which replaces the quasipolynomial sparse-orbit generator at the base.

Since $\Gamma_{n,s}$ has two parameters, of degree at most n and $2B_s$, the grid rule of Section 2 turns each generator below into an explicit hitting set.

6.1 Faithful maps

Let $g = (g_1, \dots, g_m)$ with $g_i \in \mathbb{F}[\bar{x}]$, $\bar{x} = (x_1, \dots, x_n)$, and let $J_{\bar{x}}(g) = (\partial g_i / \partial x_j)_{i,j}$. Write $\text{trdeg}_{\mathbb{F}}(g)$ for the largest size of an algebraically independent subset of $\{g_1, \dots, g_m\}$. A polynomial map $\Phi \in \mathbb{F}[\bar{w}]^n$, in some parameter variables $\bar{w}$, is *faithful* to g if $\text{trdeg}_{\mathbb{F}}(g(\Phi)) = \text{trdeg}_{\mathbb{F}}(g)$. Faithfulness implies that substitution $\bar{x} \mapsto \Phi$ is injective on $\mathbb{F}[g_1, \dots, g_m]$; consequently it preserves every nonzero polynomial combination of the g_i [4, Theorem 14 and Corollary 15].

Lemma 6.1 ([1, Lemma 2.7], see also [4]). *Let $\text{trdeg}_{\mathbb{F}}(g) \leq r$. Let $\Psi \in \mathbb{F}[\bar{z}]^n$ be a polynomial map in parameter variables $\bar{z}$, and write $J_{\bar{x}}(g)(\Psi)$ for the Jacobian evaluated at $\bar{x} = \Psi(\bar{z})$. Suppose*

$$\text{rank}_{\mathbb{F}(\bar{z})} J_{\bar{x}}(g)(\Psi) = \text{rank}_{\mathbb{F}(\bar{x})} J_{\bar{x}}(g).$$

Using fresh variables $\tau, v_1, \dots, v_r$, define

$$\Phi_i(\bar{z}, \tau, \bar{v}) = \Psi_i(\bar{z}) + \sum_{j=1}^r v_j \tau^{ij}, \quad i \in [n].$$

Then Φ is faithful to g . It adds $r+1$ parameters and has coordinate degree at most $\max\{\deg \Psi, nr+1\}$.

We discuss briefly the idea for Lemma 6.1. We want $\text{trdeg}_{\mathbb{F}}(g(\Phi)) = \text{trdeg}_{\mathbb{F}}(g)$. By the Jacobian criterion, it suffices to preserve the corresponding Jacobian rank. The chain rule gives

$$J_{\bar{v}}(g(\Phi))|_{\bar{v}=0} = J_{\bar{x}}(g)(\Psi) V(\tau), \quad V(\tau)_{ij} = \tau^{ij}.$$

The added terms produce the Vandermonde factor $V(\tau)$, which preserves the rank of $J_{\bar{x}}(g)(\Psi)$. This gives the required faithfulness; see [1, Lemma 2.7] for a detailed proof. In applications $g_i(\bar{x}) = h_i(A\bar{x} + \bar{b})$. The chain rule gives

$$J_{\bar{x}}(g) = J_{\bar{y}}(h)(A\bar{x} + \bar{b})A. \tag{6.2}$$

Therefore it suffices to preserve a nonzero maximal minor of the *hidden-coordinate* Jacobian after the affine substitution. A generator that hits every such nonzero minor does this without identifying the witnessing minor or the matrix A .

6.2 Inputs of bounded transcendence degree

We begin with polynomial combinations of sparse inputs of bounded transcendence degree, the model studied by Beecken, Mittmann, and Saxena [4]. Saha and Thankey [26] observed that their method also gives a quasipolynomial hitting set for its affine orbits. We obtain a polynomial-size hitting set for fixed transcendence degree. The argument serves as a warm-up for the constant-depth, constant-occur formulas considered next.

Theorem 6.3. *Let $h_1, \dots, h_m \in \mathbb{F}[\bar{y}]$ satisfy*

$$\text{sp}(h_i) \leq s, \quad \text{trdeg}_{\mathbb{F}}(h_1, \dots, h_m) \leq r, \quad r \geq 1.$$

Let C be any polynomial and

$$f(\bar{x}) = C(h_1(A\bar{x} + \bar{b}), \dots, h_m(A\bar{x} + \bar{b})), \quad \deg f \leq D,$$

where $A \in \text{GL}_n(\mathbb{F})$. Set $\sigma = r!s^r$ and $B = n\sigma^2 + \sigma$. An explicit hitting set for this class has size at most

$$(nD + 1)(2BD + 1)(nrD + 1)(D + 1)^r. \quad (6.4)$$

In particular, the size is polynomial for fixed r .

Proof. Let $\rho = \text{trdeg}(h) \leq r$. Each $\rho \times \rho$ minor of $J_{\bar{y}}(h)$ has sparsity at most $\rho!s^\rho \leq \sigma$. Consequently $\Psi = \Gamma_{n,\sigma}$ preserves every nonzero maximal minor after substitution by $A\bar{x} + \bar{b}$. By (6.2) and Lemma 6.1, the map

$$\Phi_i = \Gamma_{n,\sigma,i}(\beta, t) + \sum_{j=1}^r v_j \tau^{ij}$$

is faithful to $(h_i(A\bar{x} + \bar{b}))_{i \in [m]}$ and preserves nonzero f . The degrees of $f(\Phi)$ in β, t, τ , and each v_j are at most $nD, 2BD, nrD$, and D , respectively. Evaluating on the corresponding Cartesian grid gives (6.4). $\square$

The restriction is on the joint transcendence degree of the inputs, not on the number m of inputs or on the depth of C . Only the degree of the resulting f enters the grid bound.

6.3 Constant-depth, constant-occur formulas

An *occur- k formula* is a rooted tree whose leaves are sparse polynomials and whose internal gates compute linear combinations or products of powers $\prod_j q_j^{e_j}$, with positive integer exponents e_j . Each variable appears in at most k leaves. Size is the sum of the sizes of the depth-two leaf circuits and the edge weights: an edge entering a sum gate has weight one, and a power edge has weight equal to its exponent. Depth is the depth of the underlying tree plus two, accounting for the sparse leaves. These are the conventions of [26, Definition 3]. For fixed depth, gate degrees are polynomially bounded in size. Neither multilinearity nor bounded individual degree is assumed.

Consider first the depth-four case

$$f = \sum_i T_i, \quad T_i = \prod_j q_{ij}^{e_{ij}},$$

where the q_{ij} are sparse. Agrawal et al. [1] first use a shift-and-subtract reduction to bound the top fan-in in terms of k . Lemma 6.1 then reduces the construction of a faithful map for the

summands to preserving a nonzero maximal Jacobian minor. The occur bound limits how many sparse factors involve the variables indexing this minor. Extracting common powers from its rows leaves a determinant of polynomial sparsity for fixed k . Thus sparse PIT suffices to preserve the minor.

For higher depth, the same argument reduces faithfulness for a tuple of gate derivatives to faithfulness for bounded tuples of derivatives of gates one level below [1, Lemma 4.4]. Iterating reaches the sparse leaves. For fixed Δ, k , all tuple sizes are bounded by a constant $R = R(\Delta, k)$.

Saha and Thankey [26] adapt this argument to affine orbits. They replace the shift-and-subtract reduction by a derivative reduction; adding G_1^{SV} [28] then transfers a generator for the derivatives back to the original polynomial. The chain rule reduces the remaining requirements to preserving the same hidden-coordinate Jacobian minors after the affine substitution. At the leaves, this requires a generator for sparse orbits.

For fixed Δ, k , the quasipolynomial bound enters only through that base generator. After preprocessing, the leaf requirements have sparsity at most $\sigma = R!S^R$, where $S = \text{poly}_{\Delta, k}(s)$. We replace the sparse-orbit generator used in [26] by $\Gamma_{n, \sigma}$. The recursive lifts add $O(\Delta R)$ parameters and terms of degree at most $nR + 1$, giving a generator with constantly many parameters and polynomial coordinate degree.

Theorem 6.5. *Fix positive integers Δ and k . The affine orbits of degree-at-most- D polynomials computed by size- s , depth- Δ , occur- k formulas have an explicit hitting set of size $\text{poly}_{\Delta, k}(n, s, D)$.*

We omit the proof, which is the reduction of [26, Appendix D] with $\Gamma_{n, \sigma}$ as the base generator; equation (6.2) handles translations. This improves the quasipolynomial bound of [26, Theorem 10].

Acknowledgments

We thank Chandan Saha for discussions on the lower bound and for pointing out the current bottleneck in lower bounds for homogeneous depth-three circuits, and Sumanta Ghosh for numerous discussions on the orbit PIT problem.

AI disclosure

The authors' original hard family was the product $P_{n, k} = \prod_{i=1}^{n-1} (x_{i+1} - x_1 x_i)^k$ of Appendix B, whose lower bound grows more slowly with the degree. The cubic C_n in (1.4), which gives the stronger bound of Theorem 1.5, was found with the help of an AI system (GPT-6 ASTRA). AI assistants were also used for L^AT_EX formatting and language editing. AI assistants also helped us prepare this write-up quickly; a lightly revised version may follow in the next few weeks. The authors checked all statements and proofs and take full responsibility for the content.

References

- [1] Manindra Agrawal, Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. Jacobian hits circuits: Hitting sets, lower bounds for depth- D occur- k formulas and depth-3 transcendence degree- k circuits. *SIAM Journal on Computing*, 45(4):1533–1562, 2016. <https://doi.org/10.1137/130910725>. [Cited on pages 2, 5, 14, 15, and 16.]
- [2] Omkar Baraskar, Agrim Dewan, Chandan Saha, and Pulkit Sinha. NP-hardness of testing equivalence to sparse polynomials and to constant-support polynomials. In *ICALP*, volume 297 of *LIPIcs*, 2024. Article 16. [Cited on page 2.]

- [3] Walter Baur and Volker Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22(3):317–330, 1983. [Cited on page 6.]
- [4] Malte Beecken, Johannes Mittmann, and Nitin Saxena. Algebraic independence and blackbox identity testing. *Information and Computation*, 222:2–19, 2013. <https://doi.org/10.1016/j.ic.2012.10.004>. Theorem numbers refer to the full version: <https://arxiv.org/pdf/1102.2789>. [Cited on pages 2, 5, 14, and 15.]
- [5] Michael Ben-Or and Prason Tiwari. A deterministic algorithm for sparse multivariate polynomial interpolation. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing (STOC)*, pages 301–309, 1988. <https://doi.org/10.1145/62212.62241>. [Cited on pages 2, 6, and 11.]
- [6] Shubham Bhardwaj and Ramprasad Saptharishi. Hitting sets for polynomials with small partial derivative spaces. ECCC TR26-217. Also arXiv:2609.34436, 2026. <https://arxiv.org/abs/2609.34436>. [Cited on pages 4, 5, 7, and 9.]
- [7] Vishwas Bhargava and Sumanta Ghosh. Improved hitting set for orbit of ROABPs. In *APPROX/RANDOM*, volume 207 of *LIPIcs*, pages 30:1–30:23, 2021. [Cited on pages 3, 5, 6, and 7.]
- [8] Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. Deterministic factorization of sparse polynomials with bounded individual degree. *Journal of the ACM*, 67(2), 2020. [doi:10.1145/3365667](https://doi.org/10.1145/3365667). [Cited on page 2.]
- [9] Markus Bläser and Christian Engels. Randomness efficient testing of sparse black box identities of unbounded degree over the reals. In *STACS*, volume 9 of *LIPIcs*, pages 555–566, 2011. <https://doi.org/10.4230/LIPIcs.STACS.2011.555>. [Cited on pages 6 and 11.]
- [10] Prerona Chatterjee, Mrinal Kumar, Adrian She, and Ben Lee Volk. A quadratic lower bound for algebraic branching programs. In *35th Computational Complexity Conference (CCC)*, volume 169 of *LIPIcs*, pages 2:1–2:21, 2020. <https://doi.org/10.4230/LIPIcs.CCC.2020.2>. [Cited on page 6.]
- [11] Klim Efremenko, Ankit Garg, Rafael Oliveira, and Avi Wigderson. Barriers for rank methods in arithmetic complexity. In *9th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 94 of *LIPIcs*, pages 1:1–1:19, 2018. <https://doi.org/10.4230/LIPIcs.ITCS.2018.1>. [Cited on page 7.]
- [12] Michael A. Forbes. *Polynomial Identity Testing of Read-Once Oblivious Algebraic Branching Programs*. PhD thesis, Massachusetts Institute of Technology, 2014. [Cited on pages 7, 20, and 21.]
- [13] Michael A. Forbes. Deterministic divisibility testing via shifted partial derivatives. In *FOCS*, pages 451–465, 2015. Full version (April 16, 2015): <https://www.ias.edu/sites/default/files/math/csdm/14-15/Forbes2015.pdf>. [Cited on pages 4, 5, 6, and 8.]
- [14] Michael A. Forbes. Low-depth algebraic circuit lower bounds over any field. In *39th Computational Complexity Conference (CCC)*, volume 300 of *LIPIcs*, pages 31:1–31:16, 2024. <https://doi.org/10.4230/LIPIcs.CCC.2024.31>. [Cited on page 6.]

- [15] Ignacio Garcia-Marco, Pascal Koiran, Timothée Pecatte, and Stéphan Thomassé. On the complexity of partial derivatives. In *STACS*, volume 66 of *LIPIcs*, pages 37:1–37:13, 2017. <https://doi.org/10.4230/LIPIcs.STACS.2017.37>. [Cited on page 7.]
- [16] Fulvio Gesmundo and Joseph M. Landsberg. Explicit polynomial sequences with maximal spaces of partial derivatives and a question of K. Mulmuley. *Theory of Computing*, 15(3):1–24, 2019. <https://doi.org/10.4086/toc.2019.v015a003>. [Cited on page 12.]
- [17] Dima Yu. Grigoriev, Marek Karpinski, and Michael F. Singer. Fast parallel algorithms for sparse multivariate polynomial interpolation over finite fields. *SIAM Journal on Computing*, 19(6):1059–1063, 1990. doi:10.1137/0219073. [Cited on page 2.]
- [18] Neeraj Kayal, Chandan Saha, and Sébastien Tavenas. An almost cubic lower bound for depth three arithmetic circuits. In *ICALP*, volume 55 of *LIPIcs*, pages 33:1–33:15, 2016. Full version: https://www.csa.iisc.ac.in/~chandan/research/d3_cubic_lb.pdf. [Cited on pages 4, 7, and 20.]
- [19] Adam R. Klivans and Daniel A. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 216–223, 2001. doi:10.1145/380752.380801. [Cited on pages 2 and 6.]
- [20] Pascal Koiran, Natacha Portier, and Sébastien Tavenas. A Wronskian approach to the real τ -conjecture. *Journal of Symbolic Computation*, 68:195–214, 2015. <https://doi.org/10.1016/j.jsc.2014.09.036>. Theorem numbers refer to the full version, <https://arxiv.org/abs/1205.1015>. [Cited on pages 3, 6, and 11.]
- [21] J. M. Landsberg. *Tensors: Geometry and Applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012. [Cited on page 2.]
- [22] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. Superpolynomial lower bounds against low-depth algebraic circuits. In *62nd IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 804–814, 2021. <https://doi.org/10.1109/FOCS52979.2021.00083>. [Cited on page 6.]
- [23] Dori Medini and Amir Shpilka. Hitting sets and reconstruction for dense orbits in VP_e and $\Sigma\Pi\Sigma$ circuits. In *CCC*, volume 200 of *LIPIcs*, pages 19:1–19:27, 2021. [Cited on pages 2, 3, and 6.]
- [24] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6:217–234, 1997. [Cited on pages 4 and 7.]
- [25] Bruce Reznick. *Sums of Even Powers of Real Linear Forms*, volume 96 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 1992. No. 463. <https://doi.org/10.1090/memo/0463>. [Cited on page 12.]
- [26] Chandan Saha and Bhargav Thankey. Hitting sets for orbits of circuit classes and polynomial families. *ACM Transactions on Computation Theory*, 16(3):14:1–14:50, 2024. <https://doi.org/10.1145/3665800>. Theorem numbers refer to the full version, ECCC TR21-015: <https://eccc.weizmann.ac.il/report/2021/015>. [Cited on pages 2, 3, 5, 6, 7, 14, 15, and 16.]
- [27] Nitin Saxena. Diagonal circuit identity testing and lower bounds. In *ICALP*, volume 5125 of *LNCS*, pages 60–71, 2008. [Cited on page 20.]

- [28] Amir Shpilka and Ilya Volkovich. Read-once polynomial identity testing. *Computational Complexity*, 24(3):477–532, 2015. doi:10.1007/s00037-015-0105-8. [Cited on page 16.]
- [29] Joachim von zur Gathen and Erich Kaltofen. Factoring sparse multivariate polynomials. *Journal of Computer and System Sciences*, 31(2):265–287, 1985. doi:10.1016/0022-0000(85)90044-3. [Cited on page 2.]
- [30] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation (EUROSAM '79)*, volume 72 of *Lecture Notes in Computer Science*, pages 216–226. Springer, 1979. doi:10.1007/3-540-09519-5_73. [Cited on page 2.]

A Small depth-three circuits for the hard family

Depth-three circuit conventions. As in [18], a depth-three circuit computes $\sum_{i=1}^s \prod_{j=1}^{d_i} \ell_{i,j}$, where the $\ell_{i,j}$ are affine-linear polynomials. Its *top fan-in* is s ; its *multiplicative complexity* is $\sum_i d_i$; and its size is its number of edges. Scalar coefficients can be placed outside the products. The product-gate formal degree is $\max_i d_i$, counting nonconstant factors with multiplicity. These parameters are different: our explicit upper bounds first bound top fan-in and formal degree, which together give an edge-size bound $O(ns \max_i d_i)$.

The circuit construction uses interpolation and the standard duality idea for powers of sums [27]. We give the details because both its top fan-in and its formal degree matter.

Lemma A.1 (Powers of quadratic polynomials). *Let q be a polynomial of degree at most two in m variables over $\mathbb{C}$. For $k \geq 1$, the polynomial q^k has a depth-three circuit with top fan-in at most $(m+1)k+1$ and product-gate formal degree at most $2(m+1)k$.*

Proof. Homogenize q using one new variable and diagonalize the resulting quadratic form. Setting the new variable to one gives $q = \sum_{i=1}^r \ell_i^2$ with affine forms ℓ_i and $r \leq m+1$. The zero polynomial needs no gates. Set

$$E_k(z) = \sum_{j=0}^k \frac{z^j}{j!}.$$

The multinomial identity gives

$$q^k = k! [t^k] \prod_{i=1}^r E_k(t\ell_i^2). \quad (\text{A.2})$$

The product on the right has degree at most rk in t . Interpolation [12] expresses its t^k coefficient as a linear combination of $rk+1$ evaluations at distinct constants. At each evaluation, $E_k(t\ell_i^2)$ is a univariate polynomial of degree at most $2k$ in ℓ_i , and hence factors over $\mathbb{C}$ into affine forms. Every evaluation is therefore a product of at most $2rk$ affine forms. This proves both bounds. $\square$

Proof of the circuit upper bound. Expanding (1.4) gives

$$\begin{aligned} C_n &= (n-3)(x_2 - x_1^2)x_n - \sum_{j=3}^{n-1} x_j x_{n+2-j} \\ &\quad + x_1 \sum_{j=3}^{n-1} (x_{j-1}x_{n+2-j} + x_j x_{n+1-j}) - x_2 \sum_{j=3}^{n-1} x_{j-1}x_{n+1-j}. \end{aligned}$$

Thus $\deg_{x_1} C_n \leq 2$, $\deg_{x_2} C_n \leq 3$, and fixing x_1, x_2 leaves a quadratic in the $n-2$ remaining variables. Choose distinct $a_0, \dots, a_{2k}$ and $b_0, \dots, b_{3k}$. Let $U_a(x_1)$ and $V_b(x_2)$ be their Lagrange basis polynomials. Bivariate interpolation gives

$$C_n(\bar{x})^k = \sum_{a,b} U_a(x_1) V_b(x_2) C_n(a, b, x_3, \dots, x_n)^k.$$

Each $U_a V_b$ is a scalar times a product of $5k$ affine forms. Apply Lemma A.1 with $m = n-2$ to each remaining power. The number of product gates is at most $(2k+1)(3k+1)((n-1)k+1)$, and each has formal degree at most $5k+2(n-1)k = (2n+3)k$. These bounds give $O(n^3 k^4)$ edges. For $k = n^2$, the top fan-in is $O(n^7)$, the formal degree is $O(n^3)$, and the number of edges is $O(n^{11})$. $\square$

Sparse orbits versus depth-three circuits. The lower bound applies to affine transforms of sparse $\Sigma\Pi$ representations in the same n variables. The upper bound allows different affine forms at different product gates. Among polynomial-degree families, polynomial-sparsity affine orbits therefore form a proper subclass of polynomial-size depth-three circuits over $\mathbb{C}$.

B An exact bound with no formal-degree increase

The following construction uses Lemma 1.6. Its role is to give an exact sparse-orbit complexity together with a depth-three circuit of formal degree equal to the polynomial's degree.

Theorem B.1. *In characteristic zero, let $n \geq 2$, $k \geq 1$, and*

$$P_{n,k} = \prod_{i=1}^{n-1} (x_{i+1} - x_1 x_i)^k, \quad d = 2k(n-1).$$

Then $\deg P_{n,k} = d$ and $\text{osp}(P_{n,k}) = (k+1)^{n-1}$. Moreover, $P_{n,k}$ has a depth-three circuit over $\mathbb{F}$ with top fan-in at most $(k+1)(k(n-2)+1)$ and product-gate formal degree at most d .

Proof. Write $q_i = x_{i+1} - x_1 x_i$, so $q_1 = x_2 - x_1^2$. At $\bar{a}(\beta)$,

$$q_i(\bar{a}(\beta) + \bar{x}) = \ell_i(\bar{x}) - x_1 x_i, \quad \ell_i = x_{i+1} - \beta x_i - \beta^i x_1.$$

The forms $x_1, \ell_1, \dots, \ell_{n-1}$ are linearly independent: their coefficient matrix is triangular with diagonal entries one. Therefore the initial form of $P_{n,k}$ is linearly equivalent to $z_2^k \cdots z_n^k$. Its Hasse derivatives span all $(k+1)^{n-1}$ monomial divisors, since the relevant binomial coefficients are nonzero; see also [12, Lemmas 8.4.3–8.4.7]. Lemma 1.6 proves the lower bound.

For the upper bound, expand the $n-1$ powers. A choice $(j_1, \dots, j_{n-1}) \in \{0, \dots, k\}^{n-1}$ has nonzero coefficient $(-1)^{\sum j_i} \prod_i \binom{k}{j_i}$. Its exponent of x_n is $k - j_{n-1}$; the exponent of x_{i+1} for $1 \leq i \leq n-2$ is $k - j_i + j_{i+1}$. Reading these exponents backwards recovers the choice uniquely. Thus no two terms collide and $\text{sp}(P_{n,k}) = (k+1)^{n-1}$. The highest-degree product of the nonzero quadratic parts proves $\deg P_{n,k} = 2k(n-1)$.

For the circuit, write $P_{n,k} = (x_2 - x_1^2)^k H$, where

$$H = \prod_{i=2}^{n-1} (x_{i+1} - x_1 x_i)^k, \quad t = k(n-2).$$

Interpolate H as a degree- t polynomial in x_1 :

$$H = \sum_{j=0}^t W_j(x_1) \prod_{i=2}^{n-1} (x_{i+1} - c_j x_i)^k.$$

Each summand is a product of $2t$ affine forms. Expand $(x_2 - x_1^2)^k$ into its $k+1$ monomials $(-1)^r \binom{k}{r} x_1^{2r} x_2^{k-r}$. The resulting circuit has $(k+1)(t+1)$ product gates, each of formal degree $2t + k + r \leq 2t + 2k = d$. When $n = 2$, take $H = 1$, $t = 0$. $\square$

Proposition B.2 (Comparison at a fixed degree budget). *Let $n \geq 4$, $D \geq 2(n-1)$, $k = \lfloor D/(2(n-1)) \rfloor$, and $K = \lfloor D/3 \rfloor$. The certified cubic bound is at least the exact product bound:*

$$\binom{n-1+K}{n-1} + \binom{n-2+K}{n-1} \geq (k+1)^{n-1}.$$

Proof. Put $N = n-1$ and $T = Nk \geq 3$. Then $K \geq \lfloor 2T/3 \rfloor$, so $2K \geq T$. The function $R_{\max}(N, m)$ is nondecreasing in m , directly from its defining sum. The degree- T monomial $z_1^k \cdots z_N^k$ has derivative dimension $(k+1)^N$; hence

$$(k+1)^N \leq R_{\max}(N, T) \leq R_{\max}(N, 2K).$$

Apply (5.2). □