Black-box identity testing for affine orbits of read-once oblivious algebraic branching programs

Yilin Liu

19 November 2026

Abstract

For every fixed width, we construct a polynomial-size hitting set for read-once oblivious algebraic branching programs under an unknown invertible affine change of variables. A width bound in a single unknown order suffices. The resulting deterministic black-box identity test runs in time polynomial in the number of variables and the total degree over every field with more elements than the degree. The construction preserves the generic ranks of consecutive matrix products in descending order. At fixed dimension and width, the query count is linear in the total degree in characteristic zero or above the individual degree in the hidden coordinates. When the width bound holds in every order, cut ranks give linear dependence up to logarithmic factors even in small characteristic. We also treat multivariate layers with bounded overlap and obtain a linear query bound for bivariate programs of width two.

hexagon:2611.00634v1

1 Introduction

We have black-box access over a field $E \supseteq \mathbb{F}$ to $f(X) = F(AX + c)$, where $A \in \mathrm{GL}_n(\mathbb{F})$ and $c \in \mathbb{F}^n$ are unknown. In the hidden coordinates the polynomial is computed by a width-w read-once oblivious algebraic branching program (ROABP),

$$F(Y) = uA_1(Y_{\pi(1)}) \cdots A_n(Y_{\pi(n)})v,$$

with univariate polynomial matrices of width at most w, a constant row u, a constant column v, and an unknown variable order π. A hitting set is a finite subset of E^n on which every nonzero polynomial in the class has a nonzero value. Write $D \geq 1$ for a total-degree bound and $p = \operatorname{char} \mathbb{F}$. We may use any upper bound $1 \leq d \leq D$ on the individual degree of F, including $d = D$. We count field operations and oracle queries, given distinct field elements for the evaluation grids.

Theorem 1.1 (One unknown order, arbitrary width)**.** *For $n, w, D \geq 1$ and $|E| > D$, the affine orbit of width-w ROABPs in one unknown order admits an explicit hitting set of size*

$$(nD + 1)^{2^{O(w^2)}}.$$

In characteristic zero, or in characteristic $p > d$, there is a hitting set of size

$$(D + 1)(n + 1)^{2^{O(w \log(w+1))}}. \tag{1}$$

Section 3 proves the arbitrary-characteristic bound by two generators. The first uses w scalar parameters. The second uses a moment curve and $2^{O(w^2)}$ interpolation pairs. Section 4 derives (1) from the exponent sets of the top homogeneous component.

Theorem 1.2 (Every variable order)**.** *Suppose F admits width at most w in every variable order, with matrices depending on the order. For $|E| > D$, a hitting set of size $(D+1)(n+1)^{O(w)}$ suffices if $p = 0$, $p > D$, or $p > \max\{d, 2w-2\}$. In positive characteristic let $L = \lceil \log_p(D+1) \rceil$. With no restriction on d the size is*

$$(D + 1)(nL + 1)^{\kappa}, \qquad \kappa = \begin{cases} O(w^2 \log(w + 1)) & \text{uniformly in } p, \\ O_p(w \log(w + 1)) & \text{at fixed } p, \\ O(w \log(w + 1)) & \text{uniformly for } p > 2w - 2. \end{cases} \tag{2}$$

The width in one order bounds the coefficient ranks of its prefix cuts. Bounded width in every order gives a bound on every cut, which in turn bounds the digit weight of a translation-invariant leading coefficient (Section 5). At width two, the singleton and pair cuts alone determine the form of this coefficient.

Theorem 6.1 extends the construction to multivariate layers through the intersections of prefix and suffix spans. For independent blocks of at most b variables the degree exponent is b in characteristic zero or above D, sharp already at width one. In two variables, $15(D+1)$ queries suffice at width two whenever $|E| > D$ (Theorem 7.1).

Related work. Saha and Thankey [12] studied affine orbits of circuit classes including multilinear ROABPs. In characteristic zero or greater than d, Bhargava and Ghosh [1] obtained bounds of $(ndw)^{O(w^2 \log n \min\{w^2, d \log w\})}$ for one unknown order and $(ndw)^{O(\min\{w^2, d \log w\})}$ for every order. For a single unknown order, Theorem 1.1 gives a polynomial bound at each fixed width in every characteristic. Theorem 1.2 separates the dependence on total degree from the dimension. For ROABPs in the original coordinates see Gurjar, Korwar and Saxena [5], Singh and Singh [15], and Shpilka [13]. The rank descent and image-cone arguments build on the source manuscripts [10].

2 Cut ranks, leading coefficients, and interpolation

Lemma 2.1 (Cut ranks)**.** *Let $r(I)$ denote the rank of the coefficient matrix with rows indexed by monomials on I and columns by monomials on I^c. A polynomial admits width at most w in an order precisely when $r(I) \le w$ at every prefix [7, Lemma 8.2]. A cut of a width-w product gives w separated terms, and bases for successive suffix coefficient spaces give the converse. Local linear maps and regular specialization preserve upper bounds on cut ranks, by matrix multiplication and the vanishing of minors.*

Lemma 2.2 (Leading coefficients)**.** *The lexicographically leading amplitude coefficient of $F(Y + \sum_j z_j v_j)$ is invariant along* $\operatorname{span}\{v_j\}$ *and inherits all cut-rank bounds of the translated polynomial in the coordinates under consideration. For a direction a and an integer $M \ge 1$, the lowest nonzero coefficient $J(X)$ in*

$$R(t, X) = t^{Me} f((a + tX)/t^M), \qquad e = \deg f, \tag{3}$$

is invariant along a, inherits the latent cut-rank bounds and has all degrees in one residue class modulo M. Here $\deg_t R \le MD$*. For $M = 1$, J is the leading coefficient in r of $f(ra + X)$.*

Proof. Translating the amplitudes preserves their leading coefficient, which therefore descends to the quotient by the directions. Weights that are successive powers of $D+1$ realize lexicographic order on the amplitude monomials. Lemma 2.1 then gives the cut-rank bounds.

For each nonzero homogeneous part P_k of f, write $J_k = \operatorname{in}_0 P_k(a + X)$ for its lowest nonzero homogeneous component and $j_k = \deg J_k$. Then J_k is the highest radial coefficient of $P_k(ra + X)$, hence invariant along a. In (3) its order is $M(e-k) + j_k$. At the minimum the contributing J_k have distinct degrees in one residue class, so their sum is nonzero. The substitution acts separately on latent variables over $E(t)$, and division by the lowest power followed by specialization gives the rank bound. The terms from P_k have t-degree at most $M(e-k) + k \le MD$. □

Lemma 2.3 (Interpolation in bounded digit degree)**.** *Let* $\mathrm{wt}(j)$ *be the base-p digit sum, with* $\mathrm{wt}(j) = j$ *in characteristic zero. Write* $H_p(n, D, K)$ *for the number of monomials* X^α *with* $|\alpha| \le D$ *and* $\sum_i \mathrm{wt}(\alpha_i) \le K$*. Their span is affine-stable and, over* $|E| > D$*, admits an explicit interpolation set of this size. For* $p > 0$*,* $H_p(n, D, K) \le \binom{nL+K}{K}$ *with* $L = \lceil \log_p(D+1) \rceil$*. For ordinary total degree at most* D *in* m *variables the interpolation set has size* $\Delta_m(D) = \binom{D+m}{m}$*, including* $\Delta_0(D) = 1$*.*

Proof. In positive characteristic choose L prime-field independent v_e from $D+1$ nodes (the smaller spans have size at most D). Order the nodes $x_j = \sum_e j_e v_e$, $0 \le j \le D$, by their digit labels. The first j split into $\mathrm{wt}(j)$ cosets of initial basis spans, by aligned base-p blocks. The vanishing polynomial S of a span is additive, since adjoining v replaces S by $S^p - S(v)^{p-1}S$. Hence the Newton polynomial $N_j(T) = \prod_{h<j}(T - x_h)$ has digit degree at most $\mathrm{wt}(j)$. For the ordinary degree bound, take any $D+1$ distinct nodes in any characteristic.

The products $\prod_i N_{\alpha_i}(X_i)$ on the indicated indices form a basis by their leading monomials [11]. Evaluation on $(x_{\beta_i})_i$ is triangular with nonzero diagonal, since an entry vanishes unless $\beta_i \ge \alpha_i$ for all i. For ordinary degree the indices are simply $|\alpha| \le D$. Dimension gives optimality over a common coefficient and working field. Counting at most K digit copies among nL positions gives the binomial bound. Under affine substitution each Frobenius power has digit degree at most one, and carries only decrease digit weight. □

Lemma 2.4 (Radial evaluation and field descent)**.** *Set* $Q = n(n-1)+1$ *and* $a(u) = (1, u, \dots, u^{n-1})$*. Among* Q *distinct values of* u*, some* $a(u)$ *is a* good anchor*, where every hidden linear form is nonzero, since their product on the curve has degree at most* $Q-1$*. Rescaling the latent coordinates then turns invariance along the anchor into* $H(Y+t\mathbf{1}) = H(Y)$*.*

For $D < q = |E| < U$*, let* s *be least with* $q^s \ge U$*. In a power basis write an extension point as* $x = \sum_{j=0}^{s-1} \theta^j x_j$*. The restriction* $f(x_0 + \sum_{j=1}^{s-1} t_j x_j)$ *has total degree at most* D*, so Lemma 2.3 uses* $\Delta_{s-1}(D) \le (D+1)^{s-1} \le q^{s-1} < U$ *base-field queries. When* $(s-1)D < q$*, the curve* $\sum_j T^j x_j$ *uses only* $(s-1)D+1$ *nodes. Since* $q^s < U^2$*, enumeration finds a defining irreducible. For coefficients in an overfield take the compositum.*

Suppose a common H*-point set over the anchor field hits the leading coefficient in* r *of* $f(ra+X)$ *at each good anchor. Testing these points with all anchors and* $D+1$ *radial values detects a nonzero restriction. Including descent, the cost is* $Q\delta_Q(D+1)H$*, where* $\delta_Q = 1$ *if* $|E| \ge Q$ *and* $\delta_Q = Q$ *otherwise.*

2.1 Translation in adjacent differences

Put $b_0(r) = r - 1$ and $b_p(r) = \max\{\sum d_e : 0 \le d_e < p,\ \prod(d_e+1) \le r\}$ for $p > 0$, over finite digit vectors. Then $b_p(r) \le r - 1$ and $b_2(r) = \lfloor \log_2 r \rfloor$.

Lemma 2.5 (Translation and ordered cuts)**.** *For an* r*-dimensional polynomial representation* $\rho(t) = \sum_j D_j t^j$ *of the additive group,* $D_j \ne 0$ *implies* $\mathrm{wt}(j) \le b_p(r)$*. If* $H \ne 0$ *and* $H(Y + t\mathbf{1}) = H(Y)$*, its ordered cut ranks* r_i *give a factorization*

$$H(Y) = B_1 \rho_1(Y_1 - Y_2) B_2 \cdots \rho_{n-1}(Y_{n-1} - Y_n) B_n \tag{4}$$

with constant $r_{i-1} \times r_i$ *matrices* B_i*,* $r_0 = r_n = 1$*, and* r_i*-dimensional representations of degree at most* $\deg H$*. The total digit degree is at most* $\sum_{i<n} b_p(r_i)$ *and the weight at coordinate* i *is at most* $b_p(r_{i-1}) + b_p(r_i)$*. For a total-degree or residue-class grading in which* H *is homogeneous, the* B_i *are grade-preserving.*

Proof. Coefficient comparison gives $D_a D_b = \binom{a+b}{a} D_{a+b}$. In characteristic zero $D_j = D_1^j / j!$ with D_1 nilpotent. In positive characteristic $N_e = D_{p^e}$ satisfy $N_e^p = 0$ and $D_j = \prod_e N_e^{j_e} / j_e!$

by Lucas's identity. For $D_j v \neq 0$ take the graded-lexicographically largest α with $N^\alpha v \neq 0$. Multiplying a relation among $N^\beta v$, $\beta \leq \alpha$, by $N^{\alpha-\beta_0}$ for its least index β_0 leaves only that term. Thus $\prod_e(\alpha_e+1) \leq r$ and $\mathrm{wt}(j) \leq |\alpha| \leq b_p(r)$.

Let P_i be a basis row of prefix polynomials obtained from suffix coefficients, $P_0 = (1)$ and $P_n = (H)$. A prefix shift equals the opposite suffix shift on H. Thus $P_i(Y + t\mathbf{1}) = P_i(Y)\rho_i(t)$. Writing $P_i = P_{i-1}A_i(Y_i)$ gives $\rho_{i-1}(t)A_i(y+t) = A_i(y)\rho_i(t)$ and $A_i(y) = \rho_{i-1}(-y)B_i\rho_i(y)$ with $B_i = A_i(0)$. The endpoint representations are trivial, giving (4). Graded bases make each B_i grade-preserving and an order-j row-state transition raises grade by j. Only adjacent factors contribute to a coordinate, and expanding differences cannot increase digit weight. □

3 The arbitrary-width construction

Proposition 3.1 (Two generators)**.** *For $n, w, D \geq 1$ one can compute integers K, M satisfying*

$$K \leq 2^{O(w^2)}, \qquad M \leq C_w n^K (1+\log(d+1))^{2K},$$

where C_w depends only on w. They give the following hitting sets for Theorem 1.1.

1. *Over a working field of size at least $nDM+1$, a polynomial map in w scalar parameters gives a hitting set of size at most*

$$(nDM+1)\Delta_{w-1}(D). \tag{5}$$

   In characteristic zero or greater than $(w-1)d$, one may take $M \leq C_w n^K$.

2. *Over any working field of size at least $nD+1$, a moment curve and K Shpilka–Volkovich pairs give a hitting set of size at most*

$$(nD+1)\Delta_K(D)\min\{(n-1)D+1,\ n(n-1)+1\}^K. \tag{6}$$

Moreover, $K \leq 2^{O(w\log w)}$ in characteristic zero or characteristic greater than $(w-1)d$. In every characteristic the same bound holds when $d \leq w^{cw}$ for a fixed c, with the implicit constant allowed to depend on c.

For $D < |E| < nD+1$, Lemma 2.4 descends (6) at cost below $nD+1$ per query. This proves the first bound of Theorem 1.1. From (5) the all-field count is $(nDM+1)^2\Delta_{w-1}(D)$. For finite $nD < q < nDM+1$, the required extension degree s is $O_w(1)$. If $s-1 \leq n$, the univariate descent costs $O_w(D+1)$ per point. Otherwise the whole n-variable simplex has at most $q^n < nDM+1$ points. At $w=1$ take $K=0$, $M=1$. For $n=1$ use $D+1$ nodes. Given the field nodes, both constructions take polynomial time in n, D for fixed w.

For $1 \leq r < w$ put

$$W_r = \binom{w}{r}, \qquad h_r = (r+2)\prod_{s=r+2}^{w}\left(1+\binom{s-1}{r-1}\right) - 2,$$
$$k_r^{\mathrm{a}} = h_r b_p(W_r^2) + 2b_p(W_r). \tag{7}$$

One may always take $k_r = k_r^{\mathrm{a}}$ and $K = \sum_{r=1}^{w-1} k_r$. Section 3.3 supplies a second bound, and the smaller choice is valid at each rank. Section 3.4 constructs M from these integers.

3.1 Coefficient spaces over truncated rings

Over a field L, position i carries $R_i = L[Y_i]/(Y_i^{e_i+1})$, $e_i \geq 0$. We multiply over the tensor product of these rings. Its ideal of elements with zero constant term is nilpotent, so a matrix with invertible constant part is invertible. The support size of a monomial is the number of variables with positive exponent.

Let $\operatorname{Coeff}(P)$ be the L-span of the rows of the coefficient matrices of P, and put $\operatorname{Coeff}(SP) = \sum_{s\in S} \operatorname{Coeff}(sP)$ for a constant row space S. Multiplication gives $\operatorname{Coeff}(QP) \subseteq \operatorname{Coeff}(P)$. For disjoint positions each coefficient is a single product of local coefficients, so $\operatorname{Coeff}(QP) = \operatorname{Coeff}(\operatorname{Coeff}(Q)P)$ and supports add.

A matrix P has *unit rank* a if its constant rank is a and its $(a+1)$-minors vanish. Given a top left $a \times a$ block U with invertible constant part and top right block V, every row (x, y) in the row module satisfies $y = xU^{-1}V$. The bordered minors are the entries of the Schur complement times $\det U$. Unit rank zero means $P = 0$.

Theorem 3.2. *Let $A_i \in R_i^{W\times W}$, with every nonempty interval of unit rank in $\mathcal{P} = \{0 = p_0 < p_1 < \cdots < p_m = W\}$. For any constant row u, coefficients of support at most*

$$h(\mathcal{P}) = \prod_{j=1}^{m} (p_j - p_{j-1} + 1) - 2 \tag{8}$$

span $\operatorname{Coeff}(uA_1 \cdots A_n)$.

For $a = p_j$ set $G(a) + 1 = \prod_{\ell>j}(p_\ell - p_{\ell-1} + 1)$. If $b > a$ in $\mathcal{P}$, then

$$(b - a + 1)(G(b) + 1) \leq G(a) + 1 \tag{9}$$

by $(1 + x)(1 + y) \geq 1 + x + y$. Suppose $A_i(0) = \operatorname{diag}(I_a, Q_i)$ and let π project onto the first a coordinates. From a constant space S_0, propagate $S_i = \operatorname{Coeff}(S_{i-1}A_i)$. Choose basis lifts B_0 for πS_0 and a basis for $\ker(\pi|_{S_0})$. Give the lifts initial cost zero and the kernel rows initial cost one. A coefficient's cost is its initial row cost plus its support size.

The lemma below proves Theorem 3.2 at $a = 0$ and $S_0 = \operatorname{span}(u)$, leaving support $G(0) - 1 = h(\mathcal{P})$.

Lemma 3.3. *Under the interval hypothesis of Theorem 3.2, if $\pi S_i = \pi S_0$ at every cut, coefficients of cost at most $G(a)$ span S_n.*

Proof. Induct downwards on $a \in \mathcal{P}$. Put $B_i = B_0A_1(0) \cdots A_i(0)$, $H = \pi B_0$, and $T = \operatorname{rowspan} H$. The rows of B_i cost zero and project to H. This proves the case $a = W$. For $a < W$, $G(a) \geq 1$ handles the empty word.

If $q = \operatorname{rank}(Q_1 \cdots Q_n) > 0$, the full word has constant rank $b = a + q \in \mathcal{P}$. In the passive coordinates choose a complement U_0 of the left kernel of the full constant product. Its transported spaces U_i map isomorphically under the suffix to U_n. Pull back a complement of U_n along each suffix to obtain complements of U_i respected by the transitions. Transport bases of U_i to normalize each constant to $\operatorname{diag}(I_b, Q_i')$. These cut basis changes preserve interval unit ranks and supports.

The first-b projections are nested under constant propagation. If the old kernel projects onto β extra dimensions, they start in dimension $\dim T + \beta$ and grow strictly at most $q - \beta$ times. Put $\delta = \mathbf{1}_{\{\beta>0\}}$. Initially one can choose lifts of cost at most δ and kernel rows of cost at most $\delta + 1$. On a stable subword, lifts of cost L_0 and kernel rows of cost at most $L_0 + 1$ yield spanning rows of cost at most $L_0 + G(b)$ by induction and disjointness of supports. One growth factor adds at most one to this common cost. After $t \leq q - \beta$ growth factors, $\delta \leq \beta$ and (9) give the final cost bound

$$\delta + (t + 1)G(b) + t \leq (q + 1)(G(b) + 1) - 1 \leq G(a).$$

If $Q_1 \cdots Q_n = 0$, take c maximal with $Q_c \cdots Q_n = 0$, and write $P = A_c \cdots A_n$ and $B = B_{c-1}$. We claim that $S_n = \mathrm{Coeff}(BP)$. The suffix P has unit rank a and $S_n = \mathrm{Coeff}(S_{c-1}P)$. For $s \in S_{c-1}$, all coefficients of $\pi(sP)$ lie in T. If $T = 0$, the row-module identity gives $sP = 0$. Otherwise write $\pi(BP) = RH$ and $\pi(sP) = \lambda H$. Since $R(0) = I$, the row-module identity yields $sP = \lambda R^{-1}BP$, so $\mathrm{Coeff}(sP) \subseteq \mathrm{Coeff}(BP)$. Conversely every row of B belongs to S_{c-1}.

Restart after c from $\mathrm{Coeff}(BA_c)$ with lifts B_c of cost zero and kernel basis of cost at most one. Every restarted space lies in S_i and contains B_i, hence projects onto T. If $c < n$, the passive suffix product is nonzero by maximality, so the preceding case applies at the same a. For $c = n$ the restart already spans at cost one. In both cases its final span equals $\mathrm{Coeff}(BP) = S_n$. □

3.2 Local minors in the rank descent

For a nonzero polynomial P, write $\mathrm{in}_0 P$ for its lowest nonzero homogeneous component.

Lemma 3.4. *Suppose every interval of $A_1(Y_1) \cdots A_m(Y_m)$ has the same rank over $L(Y)$ as at zero, in $\mathcal{P}$. For constant endpoints u, v, if $P = u \prod_i A_i v \neq 0$, each monomial of $\mathrm{in}_0 P$ has support at most $h(\mathcal{P})$.*

Proof. For a chosen Y^α, reduce modulo $Y_i^{\alpha_i+1}$ at each position. Only divisors of Y^α survive, so minimality leaves a nonzero multiple of Y^α. The intervals have unit rank in the quotient. If the support exceeded $h(\mathcal{P})$, Theorem 3.2 would force v to annihilate every coefficient row, including the chosen one. □

Lemma 3.5. *Let $V(T)$ be a vector of q univariate polynomials over F_0, let $C \in L$ be transcendental over F_0, and let $\Lambda : L^q \to L$ be linear. If $g(Y) = \Lambda(V(C + Y))$ has first nonzero order a, then $\mathrm{wt}(a) \leq b_p(q)$.*

Proof. We use Schmidt's order-sequence criterion [4, Prop. 1.1]. Put $V_j(T) = [Z^j]V(T + Z)$. For $\mu \leq a$ with $\binom{a}{\mu} \neq 0$ in F_0, V_μ lies outside the $F_0(T)$-span of earlier V_j. Indeed $\Lambda(V_j(C + U)) = [Z^j]g(U + Z)$ has order at least $a - j$ in U, with equality at $j = \mu$. Transcendence makes the coefficients of any rational dependence regular at C, so earlier vectors cannot produce that order. Thus there are at most q such indices. There are $a + 1$ in characteristic zero and $\prod_e (d_e + 1)$ for $a = \sum_e d_e p^e$ by expanding $(1 + Z)^a = \prod_e (1 + Z^{p^e})^{d_e}$. □

Fix a monomial Y^α in the lowest homogeneous component of a scalar word. Extracting the coefficient of its prescribed powers outside an interval gives a scalar contraction of that interval with constant endpoints. The retained monomial lies in its lowest homogeneous component by minimality of the original degree. In particular at a single position the first order is α_i. For a width-W matrix at $C_i + Y_i$ with C_i individually transcendental, Lemma 3.5 gives $\mathrm{wt}(\alpha_i) \leq b_p(W^2)$, or $b_p(W)$ for the end row or column of a matrix entry. The contraction coefficients may depend on every center.

Interval tests. Pad the matrices to size $w \times w$ with zero states. Affine invertibility gives the latent scalar total degree at most D. Truncate each factor at degree d. Each scalar coefficient was a single product of local coefficients, so deleting exponents above d preserves the scalar. Use the resulting generic ranks and augment the word by $e_1 u$ and $v e_1^{\mathsf{T}}$, leaving just the scalar in entry $(1, 1)$.

At latent centers $C_i = c_i + \ell_i(G)$, all individually transcendental over $\mathbb{F}$, translation to independent residual variables Y_i preserves generic ranks. Call an interval *preserved* if its rank at $Y = 0$ equals its generic rank. The moment curve $\gamma(t) = (t, t^2, \ldots, t^n)$ preserves

full-rank intervals since each relevant univariate determinant is nonzero at its center. This holds for any transcendental t.

Lemma 3.6. *Suppose every interval of generic rank greater than r is preserved. There are rank-r tests whose preservation suffices for all rank-r intervals. Each test has one possibly empty chunk of generic rank greater than r, with at most one factor on either end. The lowest homogeneous component of every nonzero test minor of order r has digit degree at most k_r^{a} and ordinary degree at most $(h_r+2)rd$.*

Proof. In a target $[l,h]$ of rank r, test its first rank-r prefix $[l,i_0]$. All later prefixes still have rank r. For $i_0 \le i < h$, let $[j,i]$ be the shortest rank-r suffix of $[l,i]$ and test $[j,i+1]$. This has rank r by containment between $[j,i]$ and $[l,h]$. Minimality makes the first test's prefix before i_0 and each later test's interior $[j+1,i]$ empty or of rank greater than r. If $[l,i]$ is preserved, its specialized row space equals that of $[j,i]$ by inclusion and the generic rank bound. Appending the next factor and using the test proves preservation of $[l,i+1]$.

By Cauchy–Binet, each minor is an entry of a word of rth compounds, of width W_r and local degree at most rd. Rank s becomes $\binom{s}{r}$ by a change of bases. All chunk intervals are preserved and have compound ranks in the set

$$\{0, \tbinom{r+1}{r}, \tbinom{r+2}{r}, \dots, \tbinom{w}{r}\}.$$

Its positive gaps after the first are $\binom{s-1}{r-1}$, so (8) gives h_r. For any monomial of lowest degree, extract the exponents outside the chunk and apply Lemma 3.4. The at most h_r selected chunk positions contribute at most $b_p(W_r^2)$ each, and the boundary factors at most $b_p(W_r)$ each. The local degree bound gives the ordinary estimate. □

Substitution $Y = AX = (\ell_i(X))_i$ preserves the nonzero first homogeneous part and its degree. In characteristic $p > 0$, expand each power in base p and apply Frobenius to the linear forms. Adding exponents only decreases digit sums, so digit degree cannot increase. A perturbation detecting these lowest homogeneous components preserves the required minors. Earlier ranks persist if the perturbation can be set to zero. The augmented word has generic rank one for nonzero f, so descending to $r = 1$ suffices. In (7), $W_r \le 2^w$ and $1 + \binom{s-1}{r-1} \le 2^s$ give $K \le 2^{O(w^2)}$.

3.3 Factorial width dependence from image cones

Set $d_w(w) = w^2 - 1$ and $d_w(s) = 2s^2(2\max_{s<t\le w} d_w(t) + 1)$ for $s < w$. We prove the alternative budget

$$k_r^{\mathrm{c}} = \eta_r(b_r+2), \qquad b_r = \max_{r<s\le w} d_w(s),\ \eta_r = \begin{cases} \min\{rd, W_r^2 - 1\}, & p = 0 \text{ or } p > rd, \\ rd, & \text{otherwise.} \end{cases} \tag{10}$$

Take $k_r = \min\{k_r^{\mathrm{a}}, k_r^{\mathrm{c}}\}$ in Proposition 3.1.

Let $\mathcal{C}(M)$ be the Zariski closure, over an algebraic closure, of scalar multiples of the values of M. This cone is irreducible. Restrictions set unselected blocks to zero. Cone equality preserves homogeneous identities and thus the annihilator of the coefficient span of each compound, also over the coefficient field. Cone-preserving restrictions compose on disjoint blocks by density under multiplication.

Cone restrictions for a rank profile. Suppose every interval has its generic rank at zero and the full word has rank $s > 0$. We prove that $d_w(s)$ retained blocks suffice. For $s = w$, normalize the constants to identities. The prefix cones are nested. If a step is stable, fixing

its factor to the identity preserves that prefix cone and hence all later cones by composition. The cones start at the line through I_w in dimension w^2, allowing at most $w^2 - 1$ strict steps.

First consider rank-s factors B_i with $B_i(0) = P = \operatorname{diag}(I_s, 0)$. For constant R of size $a \times w$ put $M_i = RB_1 \cdots B_i$ and $V_i = \mathcal{C}(\pi M_i)$, where π takes the first s columns. These cones are nested since $\pi(HP) = \pi(H)$. If $V_{j-1} = V_j = V_{j+1} = V$, fix $B_j = P$. At any value of B_{j+1} with invertible top left block E and top right block F, every output with first columns Q equals $Q[I_s, E^{-1}F]$ by rank s. The restricted projected cone at that value is VE, a closed irreducible subset of V with the same dimension, so $VE = V$. The graph formula then places each original output at this value in the restricted cone. Such values are dense since $E(0) = I_s$. The restriction keeps $V_j = V_{j-1}$ and preserves the cone of M_{j+1}, hence also every later prefix cone.

Retain strict-growth blocks and the last block of each stable run. When $\pi R \neq 0$, the projected cones start at a line in dimension as, giving at most $2as - 1$ retained blocks. For $B_1 \cdots B_m$, keep B_1 and pass to its function field. Factor $B_1 = U_1[I_s, C_1]$ using its first s columns. Apply the rectangular bound to $R = [I_s, C_1]$ and multiply by U_1. A homogeneous identity on the restriction holds over the function field, hence on the full word as a polynomial identity over the original field. This gives $2s^2$ blocks.

For $s < w$, greedily cut at shortest rank-s prefixes $C_i A_{e_i}$ and absorb the final tail into the last block. Each nonempty C_i and tail, and all their subintervals, have rank greater than s. Induction and composition reduce each selected block to at most $2 \max_{t>s} d_w(t) + 1$ positions. At each internal cut the image of the right constant block complements the kernel of the left, since their product has rank s. At the ends take the first image and a complement of the last kernel. Transport active bases along the maps and choose passive kernel bases to normalize all block constants to P. Selecting at most $2s^2$ such blocks gives the stated bound on $d_w(s)$.

Degree of a test. Local compounds have degree at most rd. If $p = 0$ or $p > rd$, order equals digit weight on this range. Lemma 3.5 bounds the first order of every nonzero linear contraction by $W_r^2 - 1$. Thus the local coefficients of degrees at most η_r span the full coefficient space by duality.

The chunk cone of each test is preserved on at most b_r positions. Expand the retained compounds using their spanning coefficients of degree at most η_r, and do the same at the at most two boundary factors. Disjoint variables make the resulting products actual coefficients of degree at most $\eta_r(b_r + 2)$. They span the coefficient space of the test compound by cone equality, giving (10) for ordinary and digit degree.

The recurrence is decreasing in s, with $d_w(s) + 1 \leq 4s^2(d_w(s + 1) + 1)$. Thus $b_r + 2 \leq 2 \cdot 4^{w-r-1}(w!/r!)^2$. Summing gives $2^{O(w \log w)}$, with a factor at most 4^w from η_r if $p = 0$ or $p > rd$, and at most wd otherwise.

3.4 Construction of the generators

An exponent digit d_e on X_i represents d_e copies of (i, p^e) in positive characteristic. In characteristic zero an exponent a gives a copies of $(i, 1)$. Under $X_i = \sum_j Z_{j,i}$, Frobenius or ordinary expansion assigns each copy to a summand. For a monomial of maximum digit degree h, assign its copies to h distinct indices. Any term with these indices uses exactly one copy at each. Specifying its coordinate and power at each index then recovers the monomial, with coefficient multiplied by the product of digit factorials (ordinary exponent factorials in characteristic zero). This multiplier is nonzero.

A shared parameter. For $w \geq 2$ put $R = \max_{1 \leq r < w} \min\{(h_r + 2)rd, k_r^{\mathrm{c}}\}$. By Lemma 3.6 and Section 3.3, each lowest homogeneous component has ordinary degree at most R. This

bound is $O_w(d)$, and $O_w(1)$ if $p = 0$ or $p > (w-1)d$. For $p > 0$ set $L = \lfloor \log_p R \rfloor$ and $\mathcal{Q} = \{1, p, \dots, p^L\}$. For $p = 0$ set $L = 0$ and $\mathcal{Q} = \{1\}$.

Take the first K primes m with $m - 1 > n(2L+1)$, excluding divisors of $p\prod_{j=1}^{L}(p^j - 1)$ when $L > 0$. At each prime choose nonzero residues $a_{m,i}$ so that

$$(i, q) \longmapsto a_{m,i} q \pmod m \quad \text{is injective on } [n] \times \mathcal{Q}. \tag{11}$$

The exclusions distinguish the powers at each coordinate. The ratios of two powers forbid at most $2L+1$ new residues per previous choice, leaving an admissible residue at each step.

The excluded primes have total logarithm at most $B_0 = (1+L(L+1)/2)\log p$ for $L > 0$, and $B_0 = 0$ otherwise. Set $x = 100(n(2L+1) + K^2 + B_0 + 1)$. The estimates $0.82x \le \vartheta(x) \le 1.2x$ for $x \ge 97$ give $\vartheta(4x) - \vartheta(x) > 2x$ for the prime sum $\vartheta(x) = \sum_{m \le x,\ m \text{ prime}} \log m$ [3]. Since $B_0 + K\log(4x) < 2x$, the interval $(x, 4x]$ contains at least K admissible primes. Their product M satisfies the stated bound, using $L = O(1 + \log R)$ and $B_0 = O((1+\log R)^2)$. For $R = O_w(1)$ we have $x = O_w(n)$.

Partition the chosen primes into sets J_r of size k_r and put $N_r = \prod_{m \in J_r} m$. Define

$$V_{r,i}(s) = \sum_{m \in J_r} s^{(M/m)a_{m,i}}, \qquad 1 \le a_{m,i} < m,$$
$$G_i(s, \epsilon) = s^{iM} + \sum_{r=1}^{w-1} \epsilon_r V_{r,i}(s). \tag{12}$$

Before stage r the earlier exponents are divisible by N_r, so the coordinates lie in $L_0 = E(\epsilon_{>r}, s^{N_r})$. The distinct leading powers s^{iM} dominate all corrections of degree below M, keeping each latent center individually transcendental over $\mathbb{F}$.

Lemma 3.7. *If $0 \ne H \in L_0[X]$ has digit degree at most k_r and ordinary degree at most R, then $H(V_r(s)) \ne 0$.*

Proof. Distinct exponent classes modulo N_r are independent over L_0 by clearing denominators. Assign the h copies of a maximum-digit monomial to distinct primes in J_r. Its exponent has nonzero residue at each chosen m, where only that prime's summand contributes and M/m is a unit. Thus every assignment in this class uses all h primes. By maximality it uses one copy per prime, each determined by (11). These assignments yield the same integer exponent, with the nonzero factorial multiplier above. This includes the constant case $h = 0$. □

The translated minors are over L_0, and field extension preserves ranks. For each required minor with lowest homogeneous component H in observed coordinates, the coefficient of $\epsilon_r^{\deg H}$ after adding $\epsilon_r V_r(s)$ is $H(V_r(s))$. Earlier ranks persist by setting $\epsilon_r = 0$. Rank descent gives $f(G) \ne 0$, with degree at most nDM in s. For fixed s, the map is affine in the entire amplitude vector ϵ, so Lemma 2.3 uses $\Delta_{w-1}(D)$ amplitude nodes. This gives (5), including the moment curve at $w = 1$.

Interpolation over the smaller field. Take distinct nodes $a_1, \dots, a_n \in E$ with degree-$(n-1)$ Lagrange polynomials λ_i, and put $v(y) = (\lambda_1(y), \dots, \lambda_n(y))$. Starting at $\gamma(t)$ add $\sum_{j=1}^{k_r} z_j v(y_j)$ at stage r with fresh pairs [14]. For a lowest homogeneous component H assign the h copies of a maximum-digit monomial to the first h pairs, with powers $q_1, \dots, q_h$. For formal vectors v_j, the coefficient of $\prod_{j=1}^{h} z_j^{q_j}$ is a nonzero linear combination of the products $\prod_{j=1}^{h} v_{j,i_j}^{q_j}$ by the copy argument. Higher homogeneous components have larger amplitude degree. Substituting $v_j = v(y_j)$ keeps the coefficient nonzero, since node tuples recover its individual scalars. A nonzero constant component survives at zero amplitudes. The rank descent leaves a nonzero polynomial. The degrees in t and each y_j are at most nD and

$(n-1)D$, respectively. For fixed t, y the map is affine in the K amplitudes, giving $\Delta_K(D)$ nodes by Lemma 2.3.

For the smaller direction grid, keep the inclusion-maximal test intervals with at least two variable factors. Tests with at most one variable are automatic by transcendence. Each other discarded test lies in a retained interval, so its rank is at least the retained rank r and at most its generic rank r. The retained intervals have distinct left endpoints, at most n including the left boundary. For one nonzero minor per retained interval, its selected coefficient has degree at most $n-1$ in each occurring $y_j^{q_j}$. After partial specialization, at most $n-1$ values of the next y_j can kill a surviving coefficient, since $y_j \mapsto y_j^{q_j}$ is the identity or an injective Frobenius power. Thus any grids of size $n(n-1)+1$ let us fix the directions one at a time while preserving all chosen minors. The moment curve keeps later centers transcendental. Interpolate in t and the amplitudes to obtain (6) with the smaller of the two direction grids.

4 Degree bounds for one hidden order

In this section $f(X) = F(AX)$ after absorbing constants. Complete the independent forms to a basis using identity layers. Let r_i be the cut ranks of F in its hidden order. For a finite, nonempty exponent set S, write $b(S)$ for the largest vanishing order at 1 of a nonzero polynomial supported on S.

Proposition 4.1 (Exponent sets and cut ranks)**.** *Assume $|E| > D$.*

1. *The degree-linear bound (1) holds if the exponent sets S_i of the top homogeneous component in the latent variables satisfy $b(S_i) \le w^{cw}$ for a fixed c. This includes characteristic zero and $p > d$. In positive characteristic, distinct residues modulo p, or modulo $p^u \le w^{cw}$, also suffice.*

2. *In positive characteristic put $L = \lceil \log_p(D+1) \rceil$. For a supplied bound M on $\min\{D, \sum_{i<n} b_p(r_i)\}$ the counts are*

$$Q\delta_Q(D+1)H_p(n, D, M), \qquad (D+1)(nL+1)^{2^{O(w\log(w+1))}} \quad (M < p). \tag{13}$$

   In particular the second bound holds whenever $p > \sum_{i<n} b_p(r_i)$, using $M = \min\{D, p-1, (n-1)b_p(w)\}$.

Proof. Write P for the top homogeneous latent component. Scaling all variables and taking the leading term gives P the same cut-rank upper bounds as F. In homogeneous bases of the prefix spaces, each transfer entry is a monomial whose exponent is a difference of basis degrees, so $|S_i| \le w^2$.

In characteristic zero, or when the exponents are distinct modulo p, the determinant of the first $m = |S|$ binomial rows is the Vandermonde determinant on S divided by $\prod_{0\le j<m} j!$. Thus $b(S) = m-1$. In positive characteristic group the exponents of S by residue, writing $S_r = \{(j-r)/p : j \in S,\ j \equiv r \bmod p\}$. With empty classes omitted and $b(S) = 0$ for a singleton, the exact recursion is

$$b(S) = ph + \#\{r : b(S_r) = h\} - 1, \qquad h = \max_r b(S_r). \tag{14}$$

At $1+Y$ the polynomial is $\sum_r (1+Y)^r P_r(1+Y^p)$. The independent polynomials $(1+Y)^r$ separate the residue classes in each block of p consecutive orders. Vanishing below ph forces each nonzero P_r to attain order h. In the next block the leading scalars of precisely those classes with $b(S_r) = h$ vary freely. The same Vandermonde calculation gives (14). Iteration

bounds $b(S)$ by $p^u - 1$ for distinct residues modulo p^u. The substitution $T \mapsto aT$ gives the same result at any $a \neq 0$.

Choose a good anchor a and put $J(X) = \mathrm{in}_0 P(Aa + AX)$. To bound an exponent of a monomial of J in the latent variables, extract its prescribed powers in the other variables. The exponent is the first surviving order of a shifted univariate polynomial supported on S_i. We may bound all $b(S_i)$ by the known value $B = w^2 - 1$ if $p = 0$ or $p > d$, and by $B = \lfloor w^{cw} \rfloor$ under the exponent-set promise. Thus J has latent width at most w, latent individual degree at most B and total degree at most $T = \min(D, nB)$.

Use (6) with total degree T and individual degree at most $\min(B, T)$ to find h with $J(h) \neq 0$, taking $h = 0$ when $T = 0$. Write $e = \deg f$ and $j = \deg J$. In $[r^{e-j}]f(ra + th)$ the top homogeneous part contributes $t^j J(h)$, while each lower homogeneous part contributes a smaller power of t. Hence $T + 1$ choices for t and $D + 1$ for r suffice. Work over a field supplying $U = \max\{Q, nT + 1, D + 1\}$ nodes and descend the final query points together by Lemma 2.4. If $|E| < U$, then $U = \max\{Q, nT + 1\}$ since $|E| > D$. All factors other than $D + 1$ are therefore bounded by $(n + 1)^{2^{O_c(w \log(w+1))}}$.

For the rank-profile bounds, use the entire highest radial coefficient $C(X)$ of $f(ra + X)$. After latent rescaling it is invariant under common translation. Lemmas 2.5, 2.3 and 2.4 give the first count from its digit budget M. For $M < p$, give each factor ρ_i of (4) its own digit variables $Z_{i,e}$, $0 \leq e < L$. Replacing t^{p^e} in the divided-power expansion by $Z_{i,e}$ factors the lift into univariate matrices $\sum_{j<p} N_{i,e}^j Z_{i,e}^j / j!$ of the original width, where $N_{i,e} = [t^{p^e}]\rho_i(t)$. Passing to independent difference coordinates preserves the digit bound by affine stability. In the lift each digit variable of an edge has exponent below p, so distinct monomials encode distinct exponent vectors. The lifted polynomial therefore has ordinary degree at most M.

Complete the difference forms to a frame $l_i(X) = \sum_j c_{ij} X_j$ with an identity layer. For the prime-field basis v_a in Lemma 2.3, set $X_j = \sum_a v_a z_{ja}$ with $z_{ja} \in \mathbb{F}_p$. The powers of the forms become $\sum_{j,a} c_{ij}^{p^e} v_a^{p^e} z_{ja}$. To see that these nL linear forms are independent, observe that the Moore matrix $(v_a^{p^e})_{e,a}$ is invertible by the root bound for an additive polynomial of degree at most p^{L-1} on a span of size p^L, and each coefficient matrix $(c_{ij}^{p^e})$ is invertible. The lift agrees with C on the prime grid and is nonzero by interpolation on that span. Apply (1) at degree $M < p$ in nL variables. The construction can take its nodes in a finite extension of $\mathbb{F}_p$ in a compositum with the coefficient field. Descend its points to $\mathbb{F}_p$ by power-basis interpolation at degree M. We may cap M at $(n - 1)b_p(w)$, so these grid and descent thresholds are polynomial in nL, w (here the individual degree is below p). Basis summation returns physical directions. The outer $D + 1$ radial nodes and any anchor descent give the second bound. □

The ordinary rank-profile estimate is sharp on radial coefficients. In characteristic zero, $\prod_{i<n}(Y_{i+1} - Y_i)^{w-1}$ at $\mathbf{1}$ attains total degree $(n - 1)(w - 1)$ and, for $n \geq 3$, individual degree $2w - 2$. Each cut crosses just one factor. In characteristic p, $(Y_2 - Y_1)^{p^e}$ already has ordinary vanishing order p^e at width two.

5 Digit degree across all variable orders

Write $\bar{r}(H) = 2^{-n} \sum_{I \subseteq [n]} r(I)$ for the mean cut rank and $\deg_p H$ for the largest total digit weight of a monomial. For an invariant H, shifting I is equivalent to shifting I^c oppositely, so every cut coefficient space is translation-stable.

Lemma 5.1 (Ordinary degree from all cuts)**.** *If $H \neq 0$ is homogeneous of degree m, invariant under common translation, and $p = 0$ or $p > m$, then*

$$m \leq 2(\bar{r}(H) - 1) \leq 2(1 - 2^{1-n})(w - 1) \qquad \text{if all } r(I) \leq w. \tag{15}$$

Proof. Lemma 2.5 bounds $\deg_t H(Y + t\mathbf{1}_I)$ by $r(I) - 1$. By homogeneity its multiplicity at $\mathbf{1}_I$ is at least $m - r(I) + 1$. The multiplicity Schwartz–Zippel bound gives mean multiplicity at most $m/2$ on $\{0,1\}^n$ [2]. Use $r(\varnothing) = r([n]) = 1$ for the last bound. □

The three-variable Vandermonde attains equality in characteristic zero, with degree and nontrivial cut ranks all equal to three.

Lemma 5.2 (Permutation support)**.** *Let T be a nonzero tensor supported on permutations of $[a]$, with all cut ranks at most w. Suppose each supported permutation has, for every k-set of positions, a distinct supported permutation agreeing outside that set. If $2 \le k \le a$, then*

$$2^{(a-k+1)/2} \le \bar{r}(T) \le w. \tag{16}$$

Proof. Build a binary tree of depth $a - k + 1$. At each internal node at least k positions remain unqueried, so two compatible permutations differ at some unqueried position. Branch on two distinct values there. Choose one compatible permutation at each leaf.

For a uniform random subset $B \subseteq [a]$, let $N_v(B)$ count the distinct preimages $\pi^{-1}(B)$ at leaves below node v. At a branch the two families are disjoint if exactly one of its two values belongs to B, an event of probability $1/2$. Thus N_v is at least the geometric mean of the two child counts, and at least twice that mean on this event. Taking expected base-two logarithms and inducting from the leaves gives $\mathbb{E} \log_2 N_{\text{root}} \ge (a - k + 1)/2$.

For fixed I, distinct images $\pi(I)$ index disjoint nonzero blocks of the cut matrix. Counting pairs $(I, \pi(I))$ therefore gives $\mathbb{E}N_{\text{root}} \le 2^{-a} \sum_I r_T(I) = \bar{r}(T)$. Jensen's inequality gives (16). □

Lemma 5.3 (Digit degree from layer capacities)**.** *Let $H \neq 0$ be invariant under common translation in characteristic $p > 0$, with every cut rank at most $w \ge 2$. Put $b = b_p(w)$ and $k = \lfloor \log_2 w \rfloor + 1$. In a maximum-digit-degree monomial, chosen with lexicographically largest digit counts from the highest layer down, let a be the maximum matching size between coordinates and occupied layers. If each layer contains at most C copies, then*

$$\deg_p H \le a \max\{b, C\}, \qquad a \le k - 1 + \lfloor 2 \log_2 \bar{r}(H) \rfloor. \tag{17}$$

Put $j = \lfloor \log_p w \rfloor$, $c = \lfloor w/p^j \rfloor$, and $a_0 = (c+1)p^j - 1 \le 2w - 1$. One may always use

$$C = \min\{n, a_0 - 1\} \min\{p - 1, b\}. \tag{18}$$

If $p > 2w - 2$, one may also use $C = 2w - 3$.

Proof. By Lemma 2.5, a singleton shift followed by setting the shifted variable to zero bounds every exponent's digit weight by b, and every subset shift has zero coefficient at orders of weight greater than b. Inclusion–exclusion on shifts by subsets of S retains precisely the Taylor increments with support S.

Write Y^α for the chosen monomial. Packing digits toward $p - 1$ minimizes the product of one plus each digit at fixed digit sum, so $b = (p-1)j + c - 1$ and $\mathrm{wt}(a_0) = b + 1$. If a power $q = p^e$ occupied a_0 coordinates S, use inclusion–exclusion on S at order $a_0 q$ and extract $Y^{\alpha - q\mathbf{1}_S}$. Lucas' identity and maximal digit degree force exactly one translated copy per coordinate in S. The descending-layer tie excludes powers above q, so their sum forces all copies to equal q. The unique source Y^α contributes nontrivially, contradicting the shift-order bound. This proves (18).

For the matching bound, encode each power $Y_i^{p^e}$ by $Z_{i,e}$ and use the truncated algebra $Z_{i,e}^p = T_e^p = 0$. In the physical shift $Y_i \mapsto Y_i + t$, Frobenius sends $Z_{i,e}$ to $Z_{i,e} + t^{p^e}$. A carry in t lowers the combined digit degree. The terms of highest digit degree are therefore invariant

under the truncated shifts $Z_{i,e} \mapsto Z_{i,e} + T_e$. That identity respects the combined layer counts in Z and T. Keep the lexicographically largest layer-count component G, which contains the encoding of Y^α. Relabeling monomials separately by coordinate and taking leading components by digit degree and successive layer degrees preserve every original cut-rank bound, with the variables $Z_{i,e}$ grouped by i.

At each cut the coefficient space of G is multigraded by layer counts. A truncated shift on one side equals the opposite shift on the other by simultaneous invariance, so this space is stable under its side's shifts. Write D_δ for the coefficient operator of T^δ. The shifts compose in the truncated algebra, giving $D_{\delta-\beta}D_\beta = \binom{\delta}{\beta}D_\delta$ for $\beta \le \delta < p$ coordinatewise. If $D_\delta v \neq 0$ for a homogeneous v, then the $D_\beta v$ are nonzero of distinct multidegrees. Thus $\prod_e(1+\delta_e) \le w$, and the coefficient of every squarefree T-monomial of degree k vanishes at every subset shift.

For a maximum matching with coordinates S, differentiate G by the Hasse multiorder that leaves one copy of each matched layer in its matched coordinate of Y^α. Set variables outside S to zero and retain terms using all coordinates of S. Exactly one copy of each matched layer remains, giving a nonzero permutation tensor T. Each cut of T has rank at most that of any extension of the cut to the omitted coordinates, since all operations are local. Averaging gives $\bar{r}(T) \le \bar{r}(H)$. For a supported assignment and a k-set $J \subseteq S$, work before deleting terms missing coordinates. Apply inclusion–exclusion over shifts of subsets of J at the squarefree multiorder given by its layers on J. The Hasse derivative is coefficient extraction after an independent truncated shift, so it commutes with these shifts. Evaluation outside S commutes with them as well. Setting the groups in J to zero and extracting the assignment off J sums exactly the coefficients of permutations agreeing off J, since inclusion–exclusion enforces support on all of J. The sum vanishes, so a distinct alternative exists. For $a \ge k$ Lemma 5.2 bounds a as claimed. For $a < k$ the bound follows from $\bar{r}(H) \ge 1$. Kőnig's theorem gives a vertex cover of size a, accounting for all digit copies at cost at most b per coordinate and C per layer.

For $p > 2w-2$, suppose a layer has at least $m = 2w-2$ copies. Differentiate G to retain m of them. The resulting degree-m polynomial remains invariant under truncated shifts. Project onto the selected layer by setting the others to zero. The selected monomial survives, and all cut ranks remain at most w. The shift in this layer is now a formal identity since $m < p$, contradicting (15). □

Proof of Theorem 1.2. Choose a good anchor a by Lemma 2.4. The highest coefficient of $f(ra+X)$ becomes an invariant H after rescaling the latent coordinates, and inherits all cut-rank upper bounds by Lemma 2.2. For $w = 1$ the coefficient is constant by Lemma 2.5, so at each anchor the direction set is $\{0\}$. If $p = 0$ or $p > D$, apply Lemma 5.1 to the highest homogeneous component. If $p > \max\{d, 2w-2\}$, all latent exponents of H lie below p, so the sharper layer capacity bounds ordinary degree. Thus the ordinary interpolation grid of degree at most $2w-2$ supplies the directions.

For $w \ge 2$ put $u = \min\{n, L, \lfloor 2\log_2 w \rfloor + \lfloor \log_2 w \rfloor\}$. By (17) the digit budget is $K = \min\{D, u\max(b, C)\}$, using the smaller available capacity. Here $b \le w-1$ and $C \le (2w-2)\min\{p-1, w-1\}$, with $C \le 2w-3$ when $p > 2w-2$.

Affine substitution preserves the digit budget, so Lemmas 2.3 and 2.4 give $Q\delta_Q(D+1)H_p(n, D, K)$ queries. A padded list of digit positions gives $H_p(n, D, K) \le (nL+1)^K$, proving (2). The ordinary grid has at most $(n+1)^{2w-2}$ nodes. □

Proposition 5.4 (Singleton and pair cuts at width two)**.** *Let $H(Y + t\mathbf{1}) = H(Y)$ as a polynomial identity. Its singleton and pair cut ranks are at most two if and only if*

$$H(Y) = c + \sum_i A_i(Y_i), \qquad A_i(U+V) = A_i(U) + A_i(V), \qquad \sum_i A_i(T) = 0.$$

Consequently the affine orbit of an input with these cut ranks admits a hitting set of size $Q\delta_Q(D+1)(1+nL)$*, with* Q, δ_Q *as in Lemma 2.4. Take* $L = 1$ *in characteristic zero and* $L = \lceil \log_p(D+1) \rceil$ *otherwise.*

Proof. The coefficient space on either side of a cut is stable under simultaneous translation of its variables. An active singleton space of dimension at most two contains 1, by taking the highest translation coefficient of a nonconstant member. Its other basis vector A_i, normalized by $A_i(0) = 0$, is additive, since every positive-order coefficient in $A_i(Y_i + t)$ has smaller Y_i-degree and is therefore constant. Thus H is multilinear in the separate variables $z_i = A_i(Y_i)$.

If two such variables occur together, some coefficient on their side of the cut is $V = az_iz_j + bz_i + cz_j + d$ with $a \neq 0$. Under simultaneous translation the coefficient of $t^{\deg A_i + \deg A_j}$ in V is a nonzero constant. Another positive-order coefficient has a nonzero linear part, since that part of the increment is $a(A_j(t)z_i + A_i(t)z_j)$. Together with V, these coefficients give three independent vectors in the pair space. Hence H is a sum of additive polynomials, and invariance gives the displayed identity. Conversely an additive sum has rank at most two across every cut.

Apply this description to the highest radial coefficient at a good anchor. Its digit degree is at most one, also after the observed affine substitution. Lemmas 2.4 and 2.3 give the stated count since $H_p(n, D, 1) = 1 + nL$. □

The invariant $(Y_1 - Y_2)(Y_3 - Y_4)$ has singleton ranks two and rank four at $\{1, 3\} \mid \{2, 4\}$.

6 Transfer for dependent forms and multivariate layers

Consider the more general model

$$f(X) = uA_1(L_1X + c_1) \cdots A_m(L_mX + c_m)v \tag{19}$$

with matrices of width at most w. Their arguments can be tuples, with repeated or dependent forms. Let U_i be the row span of L_i, $P_i = \sum_{j \leq i} U_j$, $S_i = \sum_{j \geq i} U_j$, and $P_0 = S_{m+1} = 0$. On the effective space $W = (\sum_i U_i)^*$, of dimension ρ, put

$$\lambda_i = \dim(P_i \cap S_{i+1}), \quad \gamma_i = \dim(P_i \cap S_i), \quad \sigma = \max_i \{\lambda_i, \gamma_i - 1, 0\}. \tag{20}$$

The matroid cut connectivity is λ_i [8]. For univariate layers its maximum equals σ.

Theorem 6.1 (Transfer to independent forms)**.** *If* $\sigma \leq s \leq n$*, model (19) has a polynomial-size hitting set over every field* E *with* $|E| > D$ *for fixed* s, w*, independently of the number of layers and their local degrees. In characteristic zero or* $p > D$*, and in any characteristic at* $w = 1$*, its size is at most*

$$(D+1)^{\min(n, s+1)}(n+1)^{2^{O(w \log(w+1))}}. \tag{21}$$

Over $|E| \geq nD + 1$ *the reduction preserves width and uses an independent-form set* $\mathcal{H}_{w,n,D}$ *from Proposition 3.1, with query count*

$$|\mathcal{H}_{w,n,D}|\Delta_s(D) \min\{(n-1)D + 1, (2n+1)s(n-1) + 1\}. \tag{22}$$

For independent blocks of size at most b, $\lambda_i = 0$ and $\sigma \leq b - 1$. A single width-one layer contains all degree-D polynomials in b variables, which require at least $\binom{D+b}{b}$ evaluations over a common coefficient and working field. This proves sharpness of the degree exponent for fixed b, and for fixed s by taking $b = \min(n, s+1)$. In a univariate word $\lambda_i \leq m - \rho$. If the forms are drawn from a basis, λ_i counts the basis directions used on both sides.

Proof. Write $f = F \circ \pi$ with π onto W and $\deg F = \deg f$. Set $\mathcal{A}_i = \operatorname{ann}(S_{i+1})$, $\mathcal{B}_i = \operatorname{ann}(P_i)$. The first flag increases and the second decreases. The codimensions of $\mathcal{A}_i + \mathcal{B}_i$ and $\mathcal{A}_{i-1} + \mathcal{B}_i$ are λ_i and γ_i, respectively. The interleaved sequence of pairs changes at most 2ρ times.

At $s = 0$ put $C = 0$ and $Q_0 = 1$. For $s > 0$ choose $\omega \in E^*$ of order at least n and set $C(t, z) = \sum_{j=1}^{s} z_j a(\omega^{j-1} t)$ for the curve a of Lemma 2.4. Testing the first $n - 1$ powers excludes at most $\binom{n}{2}$ candidates. If needed, enumerate E^*, which is then finite and cyclic. For each flag pair (A, B) take $h = \min(s, \operatorname{codim}(A + B))$ independent forms on $W/(A + B)$. Their pullbacks evaluated on $a(t)$ give independent polynomials of degrees at most $n-1$. Row reduction gives distinct leading degrees e_i. The first h folded directions have a minor of degree at most $h(n - 1)$ with leading coefficient proportional to the nonzero Vandermonde in ω^{e_i} [6, Lemma 12]. The product of the selected minors has degree below $Q_0 = (2n + 1)s(n - 1) + 1$.

Work first over $E(t)$. Quotient by the span of the projected directions, writing primes for images in W'. The minors give $\mathcal{A}'_i + \mathcal{B}'_i = W'$ and $\operatorname{codim}_{W'}(\mathcal{A}'_{i-1} + \mathcal{B}'_i) \le 1$. At $s = 0$ both follow from (20). Refine the ascending flag one dimension at a time. At each step add a vector in the smallest descending space whose intersection grows. It accounts for the growth in every larger space, giving a basis adapted to both flags. Group the vectors by their first $\mathcal{A}'_i$. Since $\mathcal{A}'_i + \mathcal{B}'_i = W'$, every later group lies in $\mathcal{B}'_i$. Group i has precisely $\operatorname{codim}(\mathcal{A}'_{i-1} + \mathcal{B}'_i) \le 1$ vectors outside $\mathcal{B}'_i$. Put the exceptional vector first when it exists. At every cut after a vector of group i, the prefix thus lifts to $\mathcal{A}_i$ and the suffix to $\mathcal{B}_i$.

The leading amplitude coefficient of $F(Y + \pi C(t, z))$ descends to W' by Lemma 2.2. With these lifts at a given cut, layers through i ignore the suffix, and later layers ignore the prefix. The word splits into w terms across the cut, and taking the leading coefficient preserves its rank bound. By Lemma 2.1 the quotient polynomial has width w in this basis. Pullback to the observed coordinates uses independent forms, completed by identity layers. Some $x \in \mathcal{H}_{w,n,D}$ therefore makes $f(x + C(t, z))$ nonzero over $E(t)$, since the construction works with grid nodes in E and coefficients in $E(t)$.

The same argument works at any t where the selected minors survive, so Q_0 nodes suffice. Alternatively, for the x just obtained the degree in t is at most $(n - 1)D$, giving the other grid in (22). At a surviving value of t the restriction is affine in z before applying f, and $\Delta_s(D)$ nodes suffice by Lemma 2.3.

For (21) use numerical t and (1). Apart from $D + 1$, the field thresholds in its proof are polynomial in n, w here since $b(S_i) \le w^2 - 1$. At width one all independent univariate factors survive on a good radial line in any characteristic. Lemma 2.4 descends both constructions, and the full grid suffices if $s + 1 \ge n$. For unrestricted characteristic descend (22) from a field supplying $nD + 1$ nodes. □

7 Width two in hidden coordinates

Theorem 7.1 (Width two in one order)**.** *In two variables,* $15(D + 1)$ *queries suffice over every field* E *with* $|E| > D$*. In* n *variables put* $Q = n(n - 1) + 1$*. For* $|E| > D$ *the counts are* $Q^2(D + 1)$ *for affine translates of homogeneous inputs and* $Q^2\Delta_2(D)$ *for arbitrary inputs, with an extra factor of at most* Q *when* $|E| < Q$*.*

For arbitrary inputs there is also a set of size $1 + Q^2 N_M$ *over* $|E| \ge \max\{Q, N_M + 1\}$*, where* $N_M = M(D - 1) + 2$ *and* $M = O((1 + \log(D + 1))^2)$ *uniformly in the characteristic.*

Lemma 7.2 (Factorization at width two)**.** *Let* $H \neq 0$ *of degree at most* D *be invariant under common translation and have width at most two in one order. Suppose* H *is homogeneous, or homogeneous modulo* M *with the signed powers* $\pm p^e$ *(*$p^e \le D$*) distinct and nonzero modulo* M*. In characteristic zero use* ± 1*. Up to a scalar over the algebraic closure,* H *is a product of at most* $n - 1$ *powers of linear forms on disjoint intervals of adjacent differences. The powers are* p^e *in positive characteristic and one in characteristic zero.*

Proof. Use graded bases in (4). Each constant bridge is nonzero. Split every rank-one bridge through a line in its grade. Between splits the remaining bridges are invertible and identify at most two grades. A nonconstant transition has order p^e (one in characteristic zero), by $b_p(2) = 1$. Its grade difference is nonzero and uniquely determines its order and orientation. Each path in a factor therefore uses at most one such transition. The endpoint grades make the factor constant or a sum of powers of the same order in separate differences, hence a power of a linear form. The nonconstant factors occupy disjoint nonempty intervals. □

Proof of the n-variable bounds. Choose Q anchors on the curve a of Lemma 2.4. Choose another Q points on $h(s) = (s, \dots, s^n)$. These hit every product of powers of at most $n-1$ nonzero affine linear forms, since the product of their distinct underlying forms restricts to a nonzero polynomial of degree at most $Q-1$.

For the homogeneous top part P of degree e, the lowest homogeneous component $J = \operatorname{in}_0 P(a+X)$ has the factorization of Lemma 7.2 after latent rescaling. Write $j = \deg J$. For a translate $f(X) = P(X-b)$, some h has $J(h-b) \neq 0$, and $[r^{e-j}]f(ra+h) = J(h-b)$. For arbitrary f, some h has $[r^{e-j}t^j]f(ra+th) = J(h) \neq 0$. Interpolation uses $D+1$ radial nodes in the first case and $\Delta_2(D)$ nodes in the second. Lemma 2.4 handles the anchor threshold.

For the weighted construction take $M = 3$ if $p = 0$ or $p > D$. Otherwise set $l = \lfloor \log_p D \rfloor$ and choose an odd prime $M \nmid p \prod_{j=1}^{2l}(p^j - 1)$. All signed powers are then distinct nonzero residues. The excluded primes have total logarithm $O(1 + l^2 \log p)$, so the prime estimate in Section 3.4 gives $M = O((1 + \log(D+1))^2)$. Some h hits the factored leading coefficient in (3). Query the origin once. If $f(0) = 0$, a term of degree $k \geq 1$ contributes t-degree at most $M(e-k) + k \leq N_M - 1$. Evaluating at N_M nonzero scales per pair (a, h) suffices. □

7.1 Width two in every order

Corollary 7.3. *If the latent width is at most two in every order, $(2n+1)(nD+1)$ queries suffice for $|E| \geq nD+1$. For $|E| \geq \max\{Q, 2D+1\}$ the same construction with quadratic curves gives $Q(2n+1)(2D+1)$ queries. Over a common coefficient and working field, sums of univariates give the lower bound $nD+1$ by dimension.*

Proof. Work over the algebraic closure and split off the rank-one modes from the coefficient tensor. With at most two remaining modes use ordinary matrix rank. Otherwise their singleton ranks are two, and the first pair spans a two-dimensional pencil of 2×2 matrices. Indeed, if the pair cut had rank one, the tensor would be $A_{12} \otimes B$ with A_{12} of rank two. Every remaining mode of B has rank two, so a cut containing mode 1 and one such mode has rank four. An entirely singular pencil has a common left or right factor, which would give singleton rank one. Over $\mathbb{C}$, the classification also follows from the secant equations of Landsberg and Manivel [9, Theorem 5.1].

Two distinct determinant roots give the form $e_{00} \otimes U + e_{11} \otimes V$. Cuts separating the first pair are block diagonal, so U, V are products. At a double root normalize an invertible member to I. For the singular member N, the double root of $\det(N + zI)$ gives $\operatorname{tr} N = 0$, so N is rank-one nilpotent. A change of bases gives $e_{00} \otimes U + (e_{01} + e_{10}) \otimes V$. The same cuts have block matrix $\left(\begin{smallmatrix} U & V \\ V & 0 \end{smallmatrix}\right)$ of rank at least $2 \operatorname{rank} V$, so V is a product. Write n' for the number of remaining modes and choose bases with $V = e_0^{\otimes(n'-2)}$. A minor bordering the two off-diagonal pivots forces any entry of U with deviations on both sides of the cut to vanish. Thus U is a sum of products with at most one factor replaced.

Apart from pure products, vanishing on a curve with nonconstant latent coordinates $L_i(s)$ therefore gives $\sum_i R_i(L_i) = \lambda$ or $\prod_i R_i(L_i) = \lambda \neq 0$ for univariate rational functions, at least

one nonconstant. In positive characteristic remove common Frobenius powers until some $R_i' \neq 0$ (over a perfect field $\ker(d/dT) = K(T^p)$). Differentiation gives

$$\sum_i S_i(L_i)L_i' = 0, \qquad S_i = R_i' \text{ or } R_i'/R_i, \tag{23}$$

with some $S_i \neq 0$ in every characteristic. Pure products already survive on such a curve.

For $n \geq 2$ use $\Gamma = (s, s^2, \ldots, s^n)$ and its perturbations $\Gamma + s^k e_j$, $k = 0, 1$. Write L_i for a latent coordinate on Γ and ℓ_i for its linear part in the observed variables. Each L_i is nonconstant. A degree-one L_i requires ℓ_i proportional to X_1, so there is at most one. In that case use $j \geq 2$, on which the other linear forms remain independent.

For $q_i = \deg L_i \geq 2$ and $S_i \neq 0$, let $c_i T^{d_i}$ be the leading Laurent term of S_i and A_i the leading coefficient of L_i. Put $h = \max q_i d_i$ over these indices if any exist. Subtract (23) on Γ from the perturbed identity. The coefficient at s^{h+k-1} is

$$(h + k) \sum_{q_i d_i = h} c_i A_i^{d_i} \ell_i(e_j).$$

Indeed the first variation in $\delta = s^k \ell_i(e_j)$ is $(S_i(L_i)\delta)'$. Each remaining term contains at least two perturbation factors δ or δ'. Its Laurent degree is at most $q_i d_i + 2k - q_i - 1 < q_i d_i + k - 1$, since $k < q_i$. One of $k = 0, 1$ makes $h + k$ nonzero, contradicting independence. If only a degree-one term survives, (23) already fails on Γ. The $2n + 1$ curves have degree at most n, giving the first count. For the second take $\Gamma = s^2 b$ for each of the Q anchors. At a good b every $q_i = 2$, so the same argument applies with $2D + 1$ nodes per curve. For $n = 1$ use univariate interpolation. □

Two variables. The two orders have the same width in two variables, as their coefficient matrices are transposes. The moment curve family gives $5(2D + 1)$ queries for $|E| \geq 2D + 1$. For $D < q = |E| \leq 2D$, test $D + 1$ points per line. Absorb the translations into the univariates so that T_1, T_2 are linear. A rank-two polynomial is $A_1(T_1)B_1(T_2) + A_2(T_1)B_2(T_2)$ with independent pairs. On a line where both latent coordinates vary, vanishing implies $R_1(T_1) = R_2(T_2)$ with $R_1 = -A_1/A_2$, $R_2 = B_2/B_1$ nonconstant of degree at most D. A product already survives there. At most two directions are bad. If the test vanishes on a good origin line $T_2 = mT_1$ and a distinct parallel $T_2 = mT_1 + c$, then c is a nonzero period of R_2. The period group has at most D elements. Indeed the numerator of $R_2(Z) - R_2(T)$ over $\overline{E}(T)$ has degree at most D in Z and has a distinct root $T + c$ for each period.

When $d_E = [E : \mathbb{F}_p] \leq 4$, take three origin lines. For each, add its translates by the d_E prime-field basis multiples of a transverse vector. A good line gives periods spanning a group of size $q > D$. This uses $3(1 + d_E) \leq 15$ lines.

For $d_E \geq 5$, choose a generator θ of $E/\mathbb{F}_p$ and three disjoint quartets of directions of cross-ratio θ. Map $(\infty, 0, 1, \theta)$ to a finite quartet S and take disjoint translates. Each previous quartet excludes at most $|S - S| \leq 13$ shifts, and $q \geq 32$. Use all twelve origin lines and add one distinct parallel for each of three of them. At least one parallel pair is good, giving a nonzero period c, and some quartet is wholly good. Fix a slope m_0 in that quartet. For each of its slopes m, the origin identities give $R_2(mT) = R_2(m_0 T)$. The period group is additive and stable under multiplication by the ratios m/m_0, hence by the $\mathbb{F}_p$-algebra they generate. This finite subalgebra of E is a field. The change to latent coordinates and the common scaling preserve the cross-ratio, so the field contains θ and equals E. The period group contains cE, of size $q > D$. This proves the bivariate bound.

Acknowledgment. The author used OpenAI models to formulate results, develop and review proofs, draft the manuscript, and implement finite checks. The source manuscripts [10] were prepared with GPT-6 Astra Pro.

References

[1] V. Bhargava and S. Ghosh. Improved hitting set for orbit of ROABPs. *Computational Complexity* 31 (2022), article 15.

[2] Z. Dvir, S. Kopparty, S. Saraf and M. Sudan. Extensions to the method of multiplicities, with applications to Kakeya sets and mergers. *SIAM Journal on Computing* 42 (2013), 2305–2328.

[3] M. Eberl. Concrete bounds for Chebyshev's prime counting functions. *Archive of Formal Proofs* (2024).

[4] A. Garcia and R. F. Lax. Weierstrass weight of Gorenstein singularities with one or two branches. arXiv:alg-geom/9208003 (1992).

[5] R. Gurjar, A. Korwar and N. Saxena. Identity testing for constant-width, and any-order, read-once oblivious arithmetic branching programs. *Theory of Computing* 13(2) (2017), 1–21.

[6] V. Guruswami and S. Kopparty. Explicit subspace designs. *Combinatorica* 36 (2016), 161–185.

[7] I. Hu, D. van Melkebeek and A. Morgan. Polynomial identity testing via evaluation of rational functions. *Theory of Computing* 20(1) (2024), 1–70.

[8] N. Kashyap. Matroid pathwidth and code trellis complexity. *SIAM Journal on Discrete Mathematics* 22 (2008), 256–272.

[9] J. M. Landsberg and L. Manivel. On the ideals of secant varieties of Segre varieties. *Foundations of Computational Mathematics* 4 (2004), 397–422.

[10] Yilin Liu and GPT-6 Astra Pro. Affine-orbit polynomial identity testing. *Algorithm Frontier*, source snapshot `77140c2` (2026).

[11] R. D. Neidinger. Multivariate polynomial interpolation in Newton forms. *SIAM Review* 61 (2019), 361–381.

[12] C. Saha and B. Thankey. Hitting sets for orbits of circuit classes and polynomial families. *ACM Transactions on Computation Theory* 16(3) (2024), article 14.

[13] A. Shpilka. Polynomial-size hitting sets for multilinear any-order read-once algebraic branching programs and for set-multilinear depth-3 circuits with constant-size variable sets. ECCC TR26-314 (2026).

[14] A. Shpilka and I. Volkovich. On reconstruction and testing of read-once formulas. *Theory of Computing* 10(18) (2014), 465–514.

[15] S. Singh and V. Singh. An algebraic rigidity framework for order-oblivious deterministic black-box PIT of ROABPs. arXiv:2602.13449v1 (2026).