hexagon:2611.00206v2

THE L'VOV–KAPLANSKY CONJECTURE FOR MATRICES OF SUFFICIENTLY LARGE ORDER

QIYUE TANG

Abstract. For each positive degree, the L'vov–Kaplansky conjecture holds over arbitrary fields in all matrix sizes beyond a threshold depending only on the degree. In these sizes the image of a nonzero multilinear polynomial is the traceless subspace if every cyclic coefficient sum vanishes, and the full matrix algebra otherwise.

1. Introduction

For a field K and an integer $n \geq 1$, write $M_n(K)$ for the algebra of n-by-n matrices and $\mathfrak{sl}_n(K)$ for the kernel of the usual matrix trace, the sum of the diagonal entries. Write I_n for the identity matrix, E_{ij} for the matrix unit in row i and column j, and $\overline{K}$ for an algebraic closure of K. We use the full trace pairing $(P, A) \mapsto \operatorname{tr}(PA)$ on $M_n(K)$ for orthogonal complements.

The commutator $AB - BA$ has trace zero. Shoda [18, Satz 3] proved that every traceless matrix over a field of characteristic zero is a commutator. Albert and Muckenhoupt [1] extended this to arbitrary fields. Thus the image of $aX_1X_2 + bX_2X_1$ is $\mathfrak{sl}_n(K)$ if $a + b = 0$ and $a \neq 0$, since the polynomial is then a nonzero multiple of the commutator. If $a+b \neq 0$, setting $X_2 = I_n$ gives the full algebra.

An ordinary homogeneous multilinear polynomial of degree d has the form

$$f(X_1, \ldots, X_d) = \sum_{\sigma \in S_d} c_\sigma X_{\sigma(1)} \cdots X_{\sigma(d)}, \qquad c_\sigma \in K.$$

Here S_d is the set of permutations of $\{1, \ldots, d\}$. The image $f(M_n(K)^d)$ consists of the individual matrices $f(A_1, \ldots, A_d)$ with $A_i \in M_n(K)$. Multilinearity gives closure under scalar multiplication by scaling one input. In Problem 1.98 of the fourth edition of the *Dniester Notebook*, I. V. L'vov asked whether this image is a vector space [6, Problem 1.98]. The question is also associated with Kaplansky [10, Introduction]. In his 1956 Shelter Island report Kaplansky asked for polynomials whose values on matrices of prescribed order are scalar, with at least one nonzero value [12]. Formanek and Razmyslov constructed such polynomials [9, 17]. Polynomial identities give the zero image, and a multilinear central polynomial has image KI_n by scaling one input. The standard identity of Amitsur and Levitzki [2] illustrates how the zero-image case depends on the matrix order.

Kanel-Belov, Malev, and Rowen proved the image-linearity assertion for 2-by-2 matrices under the hypothesis that each nonconstant polynomial over K of degree at most $2d$ has a root in K [10, Theorem 2]. Their work in order three distinguishes actual images, Zariski-dense subsets of linear spaces, and possible exceptional spectra [11]. For trilinear polynomials, Dykema and Klep [7] proved image linearity over $\mathbb{C}$ in even orders and in orders below 17. Their proof freezes inputs in a clock-and-shift basis and solves linear equations in the remaining input. Vitas [24, Theorem B] proved image linearity in all orders for degree three over algebraically

Date: November 8, 2026.

2020 *Mathematics Subject Classification.* Primary 16R10. Secondary 15A24.

Keywords. L'vov–Kaplansky conjecture, multilinear polynomials, matrix algebras.

closed fields of characteristic zero. In degree four, Fagundes, de Mello, and dos Santos [8] corrected the proof of traceless containment in order at least three over algebraically closed fields of characteristic zero. They also proposed a degree-dependent threshold for traceless containment in the Mesyan conjecture. Recent preprints classify images of ordinary polynomials in degrees four through six over algebraically closed fields of characteristic zero [21, 22, 16]. Miah [15, Theorem 1.1] treats the alternating standard polynomials in all orders and degrees over those fields.

For arbitrary degree, Vitas [23, Theorem 1.1] proved that over an infinite field every traceless matrix of a prescribed order is a value of a given nonzero multilinear polynomial after embedding the target into a larger matrix algebra by adjoining a zero block. We study the image in the target's own matrix order and ask for a threshold uniform both in the coefficients and in the field.

Two words are cyclically equivalent if one is obtained from the other by cyclic rotation of its letters. We call f formally traceless if the sum of its coefficients in each cyclic class is zero, and formally traceful otherwise. For instance, $X_1X_2 - X_2X_1$ is formally traceless over every field. We set

$$V_n = \begin{cases} \mathfrak{sl}_n(K), & f \text{ formally traceless,} \\ M_n(K), & f \text{ formally traceful.} \end{cases}$$

We write $V_n(E)$ for the space given by the same rule after extending scalars to E.

Theorem 1.1. *For each $d \geq 2$ there is an integer $B(d)$ such that for every field K, every nonzero ordinary homogeneous multilinear polynomial f of degree d over K, and every $n \geq B(d)$,*

$$f(\mathfrak{sl}_n(K)^d) = f(M_n(K)^d) = V_n.$$

For the zero polynomial the image is zero. In degree one a nonzero polynomial cX_1 has image $M_n(K)$ on unrestricted inputs.

Cyclic invariance of trace gives the upper containment in V_n. For $n \geq d$, assign the letters of a word with nonzero cyclic coefficient sum to the successive units on the cycle $1 \to 2 \to \cdots \to d \to 1$. A product has nonzero trace precisely when its edges follow the cycle from one of its vertices, so its word belongs to the chosen cyclic class and its trace is one. A directed path using $E_{12}, E_{23}, \ldots, E_{d,d+1}$ isolates one chosen nonzero coefficient when $n \geq d+1$. In each construction the inputs are off-diagonal matrix units.

For the proof, fix K, $d \geq 2$, and a nonzero f as in the theorem. Fix the free variable $Y = X_d$. At order $s \geq 1$, freeze the other variables at $B \in \mathfrak{sl}_s(K)^{d-1}$ and define $L_B(y) = f(B_1, \ldots, B_{d-1}, y)$ on $\mathfrak{sl}_s(K)$ and $N_B = (\operatorname{im} L_B)^{\perp}$. Write $K^{\times} = K \setminus \{0\}$. The trace pairing is nondegenerate since $\operatorname{tr}(PE_{ij}) = P_{ji}$.

For a rank-one matrix uv, with u a column and v a row, its tangent space in the rank-one variety consists of $uw + zv$ for arbitrary rows w and columns z. We call $uv \in W \subset M_s(K)$ transverse to the linear space W if this tangent space and W span $M_s(K)$. Since $\operatorname{tr}(P(uw + zv)) = wPu + vPz$, transversality says that $P = 0$ is the only element of $W^{\perp}$ with $Pu = 0$ and $vP = 0$. Write $[M_s(K), R] = \{AR - RA : A \in M_s(K)\}$. At a rank-one square-zero value R, the sum $\operatorname{im} L_B + [M_s(K), R]$ combines the variations in the original free slot and the variations by simultaneous conjugation.

We will use the conductor formula for coprime positive integers a, b. Each integer $n \geq (a-1)(b-1)$ is a sum of nonnegative multiples of a, b. To see this, take the coefficient of b in $\{0, \ldots, a-1\}$ to match n modulo a. The remaining multiple of a is at least $1 - a$ and hence is nonnegative.

Contents

1. Introduction 1
2. Comparison factors and cyclic matrix slices 3
3. Transverse rank-one values 9
4. Nonscalar values over infinite fields 11
5. Nonscalar values over finite fields 17
6. Scalar values in positive characteristic 22
7. Scalar values in characteristic zero 23
8. Uniformity of the scalar bound 32
References 33

2. Comparison factors and cyclic matrix slices

Label the letters in their original order. For a word w in these letters, let $\nu(w)_{ij}$ for $i < j$ be one if i follows j in w, and zero otherwise. Write c_w for its coefficient and call w a support word when $c_w \neq 0$. Put

$$Q(z) = \sum_{w : c_w \neq 0} c_w z^{\nu(w)}, \qquad z^\nu = \prod_{i<j} z_{ij}^{\nu_{ij}}.$$

We work in the Laurent polynomial ring over K in the pair variables z_{ij}. A comparison factor is a divisor $1 - z^\xi$ of Q, where ξ is a nonzero primitive integer vector, meaning that the greatest common divisor of its coordinates is one. We identify ξ and $-\xi$ in this definition. Its graph has all d letters as vertices and edges ij where $\xi_{ij} \neq 0$.

A module in a graph is a vertex set such that every outside vertex is adjacent either to all of the set or to none of it. It is strong if every other module meeting it either contains it or is contained in it. Strong modules form an inclusion tree. The children of a node are its maximal proper strong submodules. In the quotient at a node, two children are adjacent when all pairs between them are edges.

Lemma 2.1. *The comparison factors have disjoint edge sets. For each support word, simultaneously reversing its comparisons on any selected collection of these edge sets gives another support word and multiplies its coefficient by* (-1) *per selected factor.*

The union of the strong-module trees of all the factor graphs, with all singletons and the full set adjoined, is a tree of blocks consecutive in every support word. In this tree each factor quotient at a node is empty, complete, or prime, where prime means it has no proper nonsingleton module. At a node with a nonempty quotient, every comparison between two children is an affine-linear function over $\mathbb{Q}$ *of the comparisons with any fixed child.*

Proof. A primitive vector extends to a basis of the integer exponent lattice. The associated Laurent monomial change takes $1 - z^\xi$ to $1 - z_{12}$, whose quotient ring is a Laurent polynomial domain. Hence the comparison factors are prime. Distinct directions up to sign give nonassociate primes, so their product divides Q. The width in one variable is the largest exponent minus the smallest exponent. Widths add under multiplication because the coefficient ring in the other variables is an integral domain. Each width of Q is at most one. Thus each comparison factor occurs once, its nonzero exponents are 1 or -1, and the edge sets are disjoint. The remaining factor has width zero in all these coordinates. Expansion gives distinct terms for the choices of factor states, with the coefficient multiplied by -1 for each flipped state.

For each factor graph, any support word and its flipped word represent the graph by their discrepancies. In a permutation graph represented by two orders, its strong modules are consecutive in both orders. The canonical quotient graphs are empty, complete, or prime [14, Section 2.3 and Proposition 4]. A set consecutive in the two representing orders is a module, since each outside letter lies on one side of the whole set in each order. Thus every strong module of any factor graph is a module of every factor graph. Two such strong modules cannot overlap without containment.

Fix one factor graph and a node C of the union tree. If C is a canonical node, any grouping of its canonical children in the union tree gives a module in the canonical quotient. Such a proper grouping of two or more children is impossible in a prime quotient. Otherwise let P be the smallest canonical node containing C. Then C meets at least two children of P and contains each child that it meets by strongness of that child. This proper union of children gives a nontrivial module in the quotient at P, which is therefore empty or complete. Its type is unchanged on the corresponding children of C.

At a complete quotient, a factor flip reverses all pairs and a single orientation bit determines their comparisons. At a prime quotient, the graph and its complement each have two transitive orientations, interchanged by reversal [14, Sections 2.3 and 5.1]. Every child is incident with an edge of each, since both graphs are connected. The comparisons with that child therefore determine both orientation bits, and each pair comparison is affine-linear in the bit for its own graph. □

2.1. **Balanced shifts.** Use the path from $\{Y\}$ to the root of the tree in Lemma 2.1. At a node with a nonempty quotient in some factor graph, take each sibling of the child containing Y as a part of a partition of the frozen letters. At a node where all factor quotients are empty, take the union of the siblings as one part. Each frozen letter occurs in exactly one part. Write H for the space of rational vectors $b = (b_i)_{i<d}$ satisfying $\sum_{i\in A} b_i = 0$ for every part A.

For a pair of support words, let ϵ_i be the difference between their indicators of X_i following Y. Let D_{ij} be the difference between their indicators of X_i following X_j, with $D_{ii} = 0$. Thus D is an alternating matrix. The cross-part comparisons at different levels are determined by the position relative to Y of the outside part. At a common level with several parts they are determined affinely by the comparisons with the child containing Y, by Lemma 2.1. We use the following two properties.

(1) For any linear combination of support differences with $\epsilon = 0$, D is block diagonal in the parts. If also $Db = 0$ for all $b \in H$, each alternating diagonal block has rank at most one and hence is zero.
(2) If a support difference has ϵ constant on each part, its cross-part blocks in D have constant entries. Thus Db on H uses only the within-part blocks.

For the difference belonging to a comparison factor, ϵ is constant on every part, and is zero on a part formed at a node where all quotients are empty. Indeed a child is a module, and its side relative to Y is uniform in both words.

Lemma 2.2. *One can choose integral b of sum zero and integral vectors $a^{(0)}, a^{(1)}$ with all coordinates nonzero as follows. Fix a support word as reference. For $j = 0, 1$, put $P_j(q, x, y) = \sum_w c_w q^{\alpha_w} x^{\beta_w} y^{\gamma_w}$, summed over support words, with exponent triple*

$$(-a^{\mathsf{T}} Db,\ \epsilon \cdot b,\ -\epsilon \cdot a),$$

where $a = a^{(j)}$ and (D, ϵ) is the difference from the reference word. Both specializations keep all support exponents distinct. Call a root-one binomial divisor $1 - q^{\alpha} x^{\beta} y^{\gamma}$ nonvertical if $(\beta, \gamma) \neq (0, 0)$. It gives the line $\alpha + \beta u + \gamma v = 0$ in the rational (u, v)-plane. Each such line

has the form $v = \rho$. A nonzero intercept for one specialization differs from the intercepts of the other. It also differs from C/A for every one-dimensional torus coset in the other zero locus parametrized by $(q_0 z^A, x_0 z^B, y_0 z^C)$ with z nonzero and $A \neq 0$, lying in no binomial-divisor hypersurface.

Proof. For a generic rational pair (a, b) with $b \in H$, distinct support exponents cannot collide. If $\epsilon \neq 0$, the last coordinate is generically nonzero, and if $\epsilon = 0$, fact (1) gives nonvanishing of the first coordinate. A binomial divisor $1 - z^\xi$ in a Laurent ring amounts to cancellation of the coefficient sums in each coset of the exponent lattice modulo $\mathbb{Z}\xi$. Hence a specialized root-one binomial gives a partition of the support into blocks of coefficient sum zero, with all differences within blocks parallel after specialization. There are finitely many partitions. We work on the rational open set where a required rank-one identity for any such partition holds only if it holds identically in a, b.

Suppose a difference in such a partition has $\epsilon \cdot b$ not identically zero. The last-two-coordinate minor implies, for every other difference (E, η),

$$(\epsilon \cdot a)(\eta \cdot b) = (\eta \cdot a)(\epsilon \cdot b).$$

Since a is unrestricted, η is proportional to ϵ. A difference with $\eta = 0$ would have $Eb = 0$, contrary to (1). Along a fixed nonzero direction among the binary indicator vectors, a block therefore has exactly two words. Orienting each difference with the same ϵ, the remaining minors imply that their D's agree on H, and (1) implies that they agree as matrices. The pairs then give a comparison factor whose ϵ does not annihilate H, contrary to the preceding observation.

It follows that for a nonvertical factor every ϵ is constant on each part, and at least one is nonzero. A difference between distinct words in the same block cannot have $\epsilon = 0$, since parallelism with a nonvertical difference would force $Db = 0$ on H, contrary to (1). If all ϵ's are proportional, the same pair argument gives a comparison factor. Otherwise the first and third coordinates give, after canceling relatively prime linear forms in a,

$$Db = \lambda(b)\epsilon$$

for every difference in the partition, with the same linear function λ. To see that the relation applies to all differences, compare each one with a nonproportional difference from the chosen pair. Fact (2) shows that on any part where $\epsilon \neq 0$, λ can depend only on that part's b-coordinates. A nonzero λ would force every difference to use the same single part. Then all the ϵ's would be proportional. Thus the new lines are $v = 0$.

For a comparison factor, a line can have a nonzero intercept independent of a only if its ϵ is supported on one part A and the restriction of D to A sends the sum-zero hyperplane into the span of the constant vector $\mathbf{1}_A$. This is equivalent to $D_{ij} = h_j - h_i$ on A for some rational h_i, because the alternating form vanishes on two vectors from that hyperplane. In fact taking $e_i - e_{i_0}$ and $e_j - e_{i_0}$ for a fixed $i_0 \in A$ gives $h_i = D_{i_0,i}$. Orient the factor with respect to one base word and list A in its order in that word. Up to the overall sign, the upper-triangular entries belong to $\{0, 1\}$. A nonzero such block therefore comes from a two-level h, with all cross-level pairs flipped. The set A is a child at a node with nonempty quotient and the two levels are consecutive blocks in every support word, since their cross pairs have one common flip. Split A into those two parts. The factor contains every edge between Y and A. The remaining Laurent factor has width zero in those coordinates, so each support difference has constant Y-comparison on A. The new cross-part comparisons are affine-linear in this same factor state. Consequently (1) and (2) persist. Carry out this refinement for every factor with one-part ϵ and a nonzero block of the displayed form, even if its intercept already varies

with a. These refinements use disjoint parts because the factors have disjoint Y-edges. After refinement a nonzero intercept cannot be independent of a: a factor that meets two parts cannot give a nonzero λ, and a one-part factor whose within-part block remains cannot give a nonzero constant intercept. The preceding partition proof applies to the refined H.

For the remaining avoidance, write each exponent-difference triple as (α, β, γ). There are finitely many possible v-slopes obtained by solving $\alpha + \beta u + \gamma v = \alpha' + \beta' u + \gamma' v = 0$ with two difference rows and nonzero coefficient determinant in u, v. These include every slope in the assertion. Indeed on a one-dimensional torus coset with direction (A, B, C), $A \neq 0$, partition the monomials by their common weight along this direction. The differences within the blocks have rank at most two and rank at least one, since a singleton cannot cancel. If the rank is one, the primitive integer direction of those differences gives a binomial divisor with the coset's constant, by cancellation in the Laurent quotient. If the rank is two, solving the two equations gives $v = C/A$. Choose generic rational b so that each moving intercept remains nonconstant in a. Choose $a^{(0)}, a^{(1)}$ independently to avoid equality of each moving intercept with zero, the intercepts at the other endpoint, and the finite list of slopes at the other endpoint, in both directions. Each excluded equality is a proper rational equation because the intercept is nonconstant in an independent tuple. Clearing denominators by a common scale preserves the identities and inequalities: first coordinates scale quadratically, second and third linearly, and intercepts and slopes linearly. □

2.2. **Mixed rectangles.** At size s invertible in K, over an algebraic closure choose a primitive s-th root ζ. With indices in $\mathbb{Z}/s\mathbb{Z}$, define the clock and shift by $Ue_i = \zeta^i e_i$ and $Ve_i = e_{i+1}$ for the standard column basis e_i. The s^2 matrices $U^u V^v$ for $u, v \in \mathbb{Z}/s\mathbb{Z}$ form a basis of M_s, and those other than I_s form a basis of $\mathfrak{sl}_s$. The matrices of shift v are supported on the cyclic diagonal $\{(i, j) : i - j = v \text{ in } \mathbb{Z}/s\mathbb{Z}\}$. For fixed v, the vectors of entries of $U^u V^v$ on its cyclic diagonal form a Vandermonde basis on the s distinct roots of unity. The trace of U^u is $\sum_i \zeta^{ui}$, which vanishes for $u \neq 0$ modulo s. Also $VU = \zeta^{-1} UV$.

For frozen tuples B of order s and C of order t, put $\operatorname{diag}(B_i, C_i)$ in each frozen slot. An input in the upper right s-by-t block of Y gives an output in that same block. Denote the resulting endomorphism of the rectangular matrix space by $T_{s,t}(B, C)$, with determinant $D_{s,t}(B, C)$. Equivalently a word contributes its prefix, evaluated at B, acting by left multiplication, and its suffix, evaluated at C, acting by right multiplication. Empty prefixes and suffixes act by the identity.

Lemma 2.3. *If $s, t \geq d + 1$ are invertible in K, then $D_{s,t}$ is a nonzero polynomial on the product of the full traceless frozen spaces. Its total degree is at most $(d - 1)st$.*

Proof. Group the words by the subset S of frozen letters preceding Y, and let T_S be the operator contributed by one group containing a nonzero coefficient. The matrix-unit path witness for that word uses separate paths in the left and right spaces and one rectangle unit joining them. It isolates that word, so T_S is nonzero on traceless inputs. Over $\overline{K}$, expand in the nonidentity clock-shift bases on the two spaces. Some basis tuple gives a nonzero operator. All the left products in that tuple have the same clock-shift degree and all the right products have the same degree. Thus the operator is a nonzero scalar times $Y \mapsto PYQ'$ for invertible clock-shift matrices P, Q'. This proves $\det T_S \neq 0$ as a polynomial.

Scale the left variables indexed by S and the right variables indexed by its complement by z. The coefficient of the highest power z^{d-1} in the mixed operator is T_S, since each other group loses one power per letter whose side changes. The leading determinant coefficient is $\det T_S$. The degree bound follows from the $(d-1)$-linearity of each operator entry in the frozen slots. □

Lemma 2.4. *Over $K = \mathbb{F}_q$, let $B^{[Q]}$ mean entrywise Q-th power for $Q = q^j$, $j \geq 1$. If $s \geq d+1$ is prime to the characteristic, both $D_{s,s}(B, B^{[Q]})$ and $D_{s,s}(B^{[Q]}, B)$ are nonzero polynomials, of degree at most $(d-1)s^2(1+Q)$.*

Proof. Take S as in Lemma 2.3 and replace B_i by zA_i for $i \in S$ and by $z^{-1}A_i$ for $i \notin S$. In the first operator the group with left set S has uniquely lowest z-degree. Changing the side of any letter increases this degree by $Q-1$. Its determinant is nonzero: it uses the independent variable sets A_i for $i \in S$ and $A_i^{[Q]}$ for $i \notin S$, and substitution of powers of independent coordinates preserves nonvanishing. In the reversed operator exchange the signs of the z-exponents. The asserted degree follows by substitution from Lemma 2.3. □

2.3. **Prime-power slices.**

Lemma 2.5. *Let K be algebraically closed, let ℓ be a prime invertible in K, and let X be a closed subvariety of a split algebraic torus $(K^\times)^r$. The closure of the points of X whose coordinates are roots of unity of ℓ-power order is a finite union of translates of subtori by such points.*

Proof. Choose a finitely generated field of definition K_0 and let k be the algebraic closure of its prime field inside K_0. This is a number field or a finite field. Since k is perfect and algebraically closed in K_0, the extension is regular and the cyclotomic Galois image is the same over k and K_0. For a number field it has finite index in $\mathbb{Z}_\ell^\times$. For $k = \mathbb{F}_q$ it is the closure of the group generated by q. Take a positive integer u with $q^u \in 1+\ell\mathbb{Z}_\ell$, requiring $q^u \in 1+4\mathbb{Z}_2$ for $\ell = 2$. If j is the finite ℓ-adic valuation of q^u-1, the binomial expansion gives $v_\ell(q^{u\ell^i}-1) = j+i$ for $i \geq 0$. Therefore the order of q^u modulo ℓ^{j+i} is ℓ^i, the order of the whole subgroup $1+\ell^j\mathbb{Z}_\ell$ modulo that modulus. This proves openness in positive characteristic as well. Choose an integer $a > 1$ in this open group with a invertible in K. For example take $1+\ell^j$ in characteristic zero or $1+p\ell^j$ in characteristic $p > 0$, with j large enough. Coordinatewise a-th powering acts on the indicated torsion points as a Galois automorphism fixing the equations of X, so preserves their closure.

The power map is finite and permutes the irreducible components of this closure. Replace a by a power fixing a given component and quotient the torus by the identity subtorus of its translation stabilizer. The component is the inverse image of a closed irreducible subvariety X' in the quotient, since the quotient map is open and the component is a union of fibers. The stabilizer of X' is finite. Compactify the quotient torus by a product of projective lines. On the closure of X', the degree of the map given by the a^j-th power is $a^{j \dim X'}$. Indeed the map pulls back $\mathcal{O}(1,\dots,1)$ to its a^j-th tensor power, so the projection formula gives the degree after intersecting $\dim X'$ times.

In the torus the ambient power map is finite étale. Its preimage of X' is the union of the translates of X' by its kernel elements. Distinct translates meet in proper closed subsets. After removing their intersections over a dense open of X', restriction to the original component is therefore finite étale, and its degree counts the translates of a generic point by the kernel elements stabilizing X'. This bounds the degree by the order of the finite stabilizer for every j, forcing $\dim X' = 0$. Restoring the stabilizing subtorus gives a coset. Density of the original torsion points on each component gives the stated translating point. □

Lemma 2.6. *There is an integer $M(d)$ such that for every K, f as in Theorem 1.1, there are coprime integers r, t between $\max(3, d+1)$ and $M(d)$, invertible in K, with the following property for each $s = r, t$. For some shifts $b_i \bmod s$ of total zero, let $B_i = D_i V^{b_i}$, where the entries of D_i are independent indeterminates except for the constraint $\operatorname{tr} D_i = 0$ when $b_i = 0$.*

Over their rational parameter field,

$$N_B \subseteq \mathcal{D} + E,$$

where $\mathcal{D}$ is the diagonal space, E lies in at most c nonzero cyclic diagonals for some integer $c \geq 0$, $\dim E \leq c$, and every nonzero matrix of E over the algebraic closure has rank at least s/c if $c > 0$. We can take $s \geq 4096c^6$ when $c > 0$. For $c = 0$, take $E = 0$.

Proof. Use Lemma 2.2 over $\overline{K}$. At the two freezes $B_i = U^{a_i^{(j)}} V^{b_i}$ the original slot sends each $U^u V^v$ to a fixed basis direction times a common nonzero scalar times $P_j(\zeta, \zeta^u, \zeta^v)$. The commutation relation gives the exponent differences in Lemma 2.2. The output shift remains v since $\sum b_i = 0$. For sufficiently large s, each frozen matrix is traceless by the nonzero $a_i^{(j)}$. Restricting the free variable to traceless matrices deletes only the identity input, in shift zero.

Choose a prime ℓ different from the characteristic, avoiding the prime divisors of the orders of the nontrivial torsion constants of binomial divisors of the two P_j. Also exclude the prime divisors of the denominators of all rational line intercepts and of the numerators of their nonzero differences, including differences from zero. At $s = \ell^k$, Lemma 2.5 describes the zeros of $P_j(\zeta, \zeta^u, \zeta^v)$. A two-dimensional torsion coset corresponds to a binomial divisor, which gives a root-one line by the choice of ℓ. A coset with constant first coordinate disappears for large k. A one-dimensional coset with first direction coordinate $A \neq 0$ contributes at most $|A|$ points at a given ζ, and zero-dimensional cosets disappear.

An exceptional one-dimensional coset with direction (A, B, C) at one endpoint cannot meet the grade $v = h/j$ of a nonzero line at the other endpoint for large k. Here h/j is in lowest terms and denotes its residue modulo ℓ^k. On that grade $y^j = q^h$. If the coset is parametrized by $(q_0 z^A, x_0 z^B, y_0 z^C)$, eliminating z from $q = \zeta$ and $y^j = q^h$ gives $\zeta^{jC - hA} = q_0^{jC} y_0^{-jA}$. The right side has fixed order, whereas the left side has order tending to infinity since $jC - hA \neq 0$ by Lemma 2.2. A grade is bad for an endpoint if its restriction of L_B fails to be invertible, with the degree-zero grade allowed separately because of the trace restriction. Only grade zero can therefore be a common bad grade containing a whole line.

Freeze each slot in the span of its two endpoint matrices. The original linear map preserves the cyclic grading, and a grade that is invertible at either endpoint remains generically invertible. Take the limit of the generic normal space in the Grassmannian along the line from the first endpoint to a generic tuple. This limit exists by properness of the Grassmannian over the discrete valuation ring at the endpoint. The orthogonality equations are closed, so the limit lies in the first endpoint normal. The pairing between opposite grades is nondegenerate and all other grades pair to zero. Consequently outside grade zero the limit is supported only in the negatives of the common bad grades. It lies there in at most c clock-shift basis directions, where c bounds the number of zeros from the first endpoint's one-dimensional cosets, independently of k. Grade zero is the diagonal space.

A nonzero combination of at most c clock-shift matrices has rank at least s/c. Conjugation by the clock-shift basis multiplies the nonzero terms by distinct characters of $(\mathbb{Z}/s)^2$. Independence of the characters means that if there are $c' \leq c$ nonzero terms, some c' such conjugates have linearly independent coefficient vectors. One can express an individual term, which is invertible, as a linear combination of these conjugates. Rank subadditivity gives $s \leq c' \operatorname{rank} A$ for the original combination A. Choose k large enough for $s \geq 4096c^6$. Repeat with a different prime ℓ to obtain a coprime pair.

Fix s, the shifts, c, and at most c permitted nonzero normal grades from the limit. Since the weighted maps are graded, the off-diagonal part of a normal space is a direct summand. On the generic rank stratum of the two-endpoint family its Grassmann limit at the first endpoint is off-diagonal and contained in the indicated at most c basis directions. Avoidance of projective

matrices of rank $< s/c$ is open on the Grassmannian, so the generic tuple satisfies the rank condition. In the full weighted affine parameter space the support restriction and dimension bound on the off-diagonal normal are open rank conditions. The low-rank incidence is closed in the product with the fixed projective space of off-diagonal matrices, and its projection is closed. Thus the required conditions hold at a generic point of the full weighted space too.

Take an independent coefficient coordinate c_w for every permutation word, including coordinates that specialize to zero. In $\operatorname{Spec}\mathbb{Z}[c_w]$, remove the common zero locus of all coefficients. For each proposed pair of cores, impose invertibility of both orders in the field, and the open conditions just described in the two weighted affine bundles. The images in the coefficient scheme are open since affine bundle projections are open. The pointwise construction over every algebraically closed field proves that these images cover every scheme point, by passage to an algebraic closure of its residue field. The scheme is Noetherian and hence quasi-compact. A finite subcover gives $M(d)$ as the largest order occurring. Over the rational parameter field of a K-valued coefficient tuple in one member of the cover, the nonempty weighted opens hold at the generic points, giving the claimed result over K. □

3. Transverse rank-one values

Write $\mathcal{R}_s$ for the variety of nonzero rank-one matrices in M_s. By the trace pairing, $R = uv$ is transverse to W if and only if the only $P \in W^\perp$ with $Pu = 0$ and $vP = 0$ is zero.

Lemma 3.1. *Let F be infinite, let $\mathcal{D} \subset M_s(F)$ be the diagonal subspace, and let E be a subspace of off-diagonal matrices. For an integer $c \geq 0$, suppose $\dim E \leq c$, the union of the supports of E has at most c entries in each row and column, and every nonzero element of $E \otimes_F \overline{F}$ has rank at least s/c if $c > 0$. Suppose $s \geq 4096c^6$ if $c > 0$ and $s \geq 2$ otherwise. Then $(\mathcal{D} + E)^\perp$ has a basis of F-rational rank-one matrices transverse to it.*

Proof. First suppose $E \neq 0$. Work over $\overline{F}$ to choose a partition of the coordinate indices into sets S, T. In each support pattern of a nonzero element of $E \otimes_F \overline{F}$, choose positions (i, j) with mutually disjoint row and column indices, also disjoint across the two index sets, and with all cross entries zero. This can be done greedily. Upon choosing (i, j), delete positions whose row lies in the support of column j or equals j, or whose column lies in the support of row i or equals i. At most $2c(c + 1)$ positions are deleted at each step. There are at least s/c nonzero entries, so at least $s/(4c^3)$ positions are chosen. Every subset of the chosen positions forms a diagonal square submatrix after independently ordering its rows and columns.

There are at most $(cs + 1)^c$ support patterns. Indeed, the entries at the at most cs possible positions are nonzero linear forms on E. At a nonzero vector, a basis among the vanishing forms determines all the vanishing forms and has size at most $c - 1$. Put each coordinate in S or T independently with probability $1/2$. For any of the chosen sets of m positions, the number Z retained in $T \times S$ is a sum of m independent Bernoulli variables of mean $1/4$. With $\mu = m/4$, Markov's inequality applied to 2^{-Z} gives

$$\Pr(Z < \mu/2) \leq 2^{\mu/2}(1 - 1/8)^m \leq \exp(-\mu/8).$$

The union bound is less than one since $c \log(cs+1) < s/(128c^3)$ for $s \geq 4096c^6$. At $s = 4096c^6$ the left side is at most $10c^2$ and the right side is $32c^3$. The difference increases thereafter. Thus there is a cut S, T for which every nonzero element restricts to a matrix of rank at least $s/(32c^3)$ on $T \times S$. This also gives rank greater than $2c - 1$ on the rectangle of each cut obtained by moving at most two indices, and on its overlap with $T \times S$, since at most two rows or columns are lost.

Fix one of these cuts. Let $E_1, \dots, E_e$ be an F-basis of E, and let $B_j = (E_j)_{T,S}$ be the restrictions for this cut. The preceding rank bound implies independence of these restrictions. Take u supported on S and v supported on T, subject to $vB_ju = 0$ for $1 \le j \le e$, with the row and column in this expression restricted to their respective supports. Over $\overline{F}$, the locus of u for which B_ju are dependent has dimension at most $|S| - R + e - 1$, where R is the minimum rank of a nonzero combination of the B_j. Indeed one can choose a projective combination in dimension $e-1$, after which its kernel has dimension at most $|S| - R$. Above this bad locus the common zero set has dimension at most $|S| + |T| - R + e - 1 < |S| + |T| - e$. Above the complementary open the zero set is an irreducible vector bundle of dimension $|S| + |T| - e$. Every component of the full zero set has dimension at least this dimension, so each meets that same open. The zero set is therefore irreducible and generically smooth. The defining equations form a complete intersection, so its coordinate ring has no embedded associated primes. Generic reducedness now implies that their ideal is prime. Its part of bidegree $(1,1)$ is the span of the defining bilinear forms, because the polynomial ring is bigraded by the degrees in u and v.

We can take every supported coordinate nonzero. For u this excludes coordinate hyperplanes. For v we need its kernel fiber to lie in no coordinate hyperplane. For a fixed coordinate vector in $\overline{F}^T$, the locus of u where that vector belongs to the span of the B_ju has dimension at most $|S| - R + e < |S|$, by parametrizing a nonzero linear combination and solving the corresponding inhomogeneous equation. On a chart of the remaining open, the vector bundle is given by solving an invertible minor in the equations for v. Its F-points are dense, since the chart is an open subset of affine space over the infinite field F. The resulting matrices uv therefore span the entire subspace of matrices supported on $S \times T$ and orthogonal to E, by the bidegree assertion.

These matrices are transverse. If $P \in \mathcal{D} + E$ satisfies $Pu = vP = 0$, the T-coordinates of Pu give a linear combination of the independent B_ju. Its vanishing implies that the E-component of P is zero. The nonzero coordinates of u on S and of v on T then kill the diagonal component of P.

Finally a matrix orthogonal to all the displayed rectangular subspaces restricts to a linear combination of the E_j on each dual rectangle $T \times S$. The rank bounds on overlaps force the coefficients for the nearby cuts to coincide with those for the initial cut. By moving at most two indices, every off-diagonal position appears in one of these dual rectangles. The common annihilator is therefore $\mathcal{D} + E$. If $E = 0$, use all cuts into two nonempty sets. Rank-one matrices with all supported coordinates nonzero span each rectangular space, are transverse by the diagonal calculation above, and their dual rectangles cover the off-diagonal positions. In both cases a basis can be selected from the transverse matrices. □

Lemma 3.2. *For either core order s of Lemma 2.6, let B_0 be the generic weighted frozen tuple in that lemma. Extend its purely transcendental parameter field by two independent generic full traceless frozen tuples C_1, C_2, and call the resulting field F. There is a geometrically integral curve over F, of degree at most $(4(d-1)s^2+2)^{2s^2}$ in an affine embedding, with an F-rational smooth point above $z = 0$ and a map to the z-line étale at this point. On a nonempty open it carries $y \in \mathfrak{sl}_s$ and a rank-one value $L_{B_0+z^2C_1+z^4C_2}(y)$ transverse to the image of that operator. If f is formally traceful, this value has nonzero trace.*

Proof. Let $A(t)$ be the matrix of the original linear operator with frozen tuple $B_0 + tC_1 + t^2C_2$, in fixed bases, and let ℓ be its rank over $F(t)$. Each entry has degree at most $2(d-1)$ in t. Choose ℓ independent columns. Among their nonzero ℓ-row minors choose one with minimal order of vanishing at zero, on rows indexed by J. Dividing all minors by the common power

of t gives a denominator $D(t)$ with $D(0) \neq 0$ and polynomial numerators of degree at most $2(d-1)\ell$. In the corresponding chart for the column space $W(t)$, the rows on J are an identity matrix. By Cramer's rule the other rows have entries $P_{ij}(t)/D(t)$, with the same degree bound. Each column of $A(t)$ is the chart matrix times its coordinates on J, so its value at zero lies in the limiting plane $W(0)$. In particular,

$$(\mathcal{D} + E)^{\perp} \subseteq \operatorname{im} L_{B_0} \subseteq W(0),$$

where $\mathcal{D}, E$ are the diagonal space and the sparse normal envelope of the core.

Let $R_1, \dots, R_k$ be the basis in Lemma 3.1 for the leftmost space. Each is transverse to $W(0)$. The intersection $T_{R_1}\mathcal{R}_s \cap W(0)$ maps onto the quotient of $W(0)$ by $(\mathcal{D}+E)^{\perp}$, because the latter space and $T_{R_1}\mathcal{R}_s$ together span $M_s(F)$. Choose tangent vectors T_j lifting a basis of the quotient.

Replace t by z^2. The equations for R to lie in the plane are

$$D(z^2)x_i - \sum_{j \in J} P_{ij}(z^2)x_j = 0 \qquad (i \notin J),$$

where x_i is the coordinate of R on row i of the chosen vectorization. On a nonzero-entry chart $R = uv$, normalize the corresponding entry of u to one. There are $2s-1$ chart coordinates, and each equation has degree at most $4(d-1)s^2 + 2$. At any of the chosen basis points over zero the Jacobian in those coordinates has rank $s^2 - \ell$, by transversality. Keep a corresponding set of pivot coordinates and prescribe the others as affine linear functions of z. This gives an étale curve near the point. On the pivot open the curve is smooth near the marked point. The reduced irreducible component through this point is geometrically integral by the smooth F-point criterion [19]. There are at most $2s-1$ equations including the coordinate slices. The affine and projective closure degrees are bounded by $(4(d-1)s^2+2)^{2s^2}$ by cumulative Bézout [13, Theorem 1.1].

For the points through R_i, choose constant free coordinates. The resulting formal lifts over $F[[z]]$ are $R_i + O(z^2)$. For each T_j choose the free-coordinate derivatives at R_1 induced by T_j. Since the plane coordinates change in order at least two, this gives lifts $R_1 + zT_j + O(z^2)$. Subtracting the constant-coordinate lift through R_1 from the latter lifts gives a family which, after dividing these differences by z, reduces to a basis of $W(0)$. Thus the lifts span $W(z^2)$ over $F((z))$. In the traceful case this plane is traceful, since the frozen tuple is generic over $K(z)$ and a directed matrix-unit cycle detects a nonzero cyclic coefficient sum on the traceless source. At least one lift therefore has nonzero trace. Choose its curve, or any of these curves in the formally traceless case.

On the open where $z \neq 0$ and the original operator has rank ℓ, a nonzero minor gives a rational preimage y in the traceless original slot. Remove its poles and the zeros of a transversality minor, and in the traceful case remove the trace-zero locus. The formal lift shows this is a nonempty open. The curve function field embeds into $F((z))$ by the marked smooth branch. The formal transversality minor has nonzero constant term in the Grassmann chart, so the asserted conditions hold at the generic point of the curve. □

4. Nonscalar values over infinite fields

We first record how a curve with a marked smooth rational point supplies compatible tuples. Let F be infinite and let C be a geometrically integral curve of degree at most D in an affine embedding, with a marked smooth F-point p. Its normal projective completion has arithmetic genus at most $(D-1)(D-2)/2$. In fact a general projection of the projective closure to a plane is birational and has image of degree at most D. For a fixed secant line, meeting

that line imposes codimension two on the space of projection centers. Distinct point pairs vary in dimension two, so a general center identifies at most finitely many pairs. Tangent lines at smooth points vary in dimension one, so a general center avoids them and the map is separable. This yields birationality, and the assertion is immediate if the span already has dimension at most two. Let $\nu : \widetilde{C} \to C_{\mathrm{pl}}$ be the normalization of the plane image. The quotient of $\nu_*\mathcal{O}_{\widetilde{C}}$ by $\mathcal{O}_{C_{\mathrm{pl}}}$ has finite support. Its first cohomology is zero, so the induced map on first cohomology is surjective. The arithmetic genus $h^1(\widetilde{C}, \mathcal{O}_{\widetilde{C}})$ is therefore at most $(D'-1)(D'-2)/2 \le (D-1)(D-2)/2$, where $D' \le D$ is the actual degree of the plane image. The divisor p is Cartier of degree one. On this geometrically integral completion, $\mathcal{O}(ep)$, the line bundle of the divisor ep, is very ample whenever $e \ge \max(2, (D-1)(D-2)+1)$ [20].

For either one curve or two curves with a good open on each, suppose each ordered-pair condition to be imposed is the nonvanishing of a nonzero rational function on a product of curves. Only distinct pairs are considered on a single curve. Choose sections of $\mathcal{O}(ep)$ on the completions for the prescribed degrees above. A double zero in a good smooth open imposes two independent conditions on a section. So does vanishing on two distinct points of a single curve, while a pair on different curves imposes one condition on each section. The points at which a double zero can occur range over a curve. The pairs at which a nonzero rational function vanishes lie in a set of dimension at most one. In either case the incidence has dimension at most $1 + (h-2) = h-1$, with h the dimension of the section space or the sum of the two section-space dimensions. Its image has proper closure. Avoiding the finitely many boundary points also excludes only proper closed subsets. Hence sections can be chosen over F whose zero schemes are finite étale algebras of the prescribed degrees and satisfy all pair conditions. One can at the same time exclude any finitely many additional zeros or poles of nonzero rational functions on either curve.

We use this with F a purely transcendental extension of K. The finite étale algebras, solution data and nonvanishing conditions specialize to K when K is infinite. Indeed choose a basis of each finite algebra over F and spread its structure constants, the coordinates of the matrices, and the inverse of its trace discriminant to a nonempty affine open of the transcendental parameters. For the pair conditions, after splitting a finite étale algebra A, the coordinates of $A \otimes_F A$ are indexed by ordered pairs of embeddings. The idempotent equal to one on distinct pairs and zero on equal pairs is invariant under Galois, and cuts out the direct factor over F for the distinct pairs. For pairs between two algebras use their tensor product. Spread the factor idempotents and the inverses of the elements that must remain units. The algebra identities and degree persist on the resulting open, which has a K-point by infinitude.

Proposition 4.1. *There exists an integer N depending only on d such that over any infinite K, for every $n \ge N$, there are a traceless frozen tuple B and a traceless input y in the fixed slot Y with $R = L_B(y)$ nonzero of rank one and $R^2 = 0$, such that*

$$\operatorname{im} L_B + [M_n(K), R] = V_n.$$

Proof. Choose the two coprime orders r, t of Lemma 2.6 and their curves from Lemma 3.2, using independent parameter sets. For two independent curve parameters z, w on one curve, the coefficient determinant for the two full frozen tuples as functions of C_1, C_2 is $z^2w^2(w^2 - z^2)$. It is nonzero, so the two frozen tuples are independent generic tuples over the field generated by the weighted parameters and z, w. By Lemma 2.3 every ordered mixed determinant is a nonzero function on the product of the two copies of the curve. This follows by dominance of the product map to the (z, w)-plane. The same holds between different curves by independence of their parameters. Apply the preceding curve construction to obtain compatible finite étale algebras A_r, A_t of any prescribed degrees e_r, e_t above a common bound depending only on d.

The bound follows from the core size bound and the curve degree estimate. If the common degree bound is E_0, every $n \geq (r+t)E_0 + (r-1)(t-1)$ has the form $re_r + te_t$ with $e_r, e_t \geq E_0$ by the conductor formula. This lower bound is bounded in terms of d.

View the core matrices as endomorphisms of $A_r^r \oplus A_t^t$ over K. Over a common separable splitting field the frozen matrices are block diagonal, with one block per embedding of the two algebras. In the traceful case the local values uv have trace vu a unit in their respective algebras. On each algebra A choose a unit λ with $\mathrm{Tr}_{A/K}(\lambda vu) = 0$, where $\mathrm{Tr}_{A/K}$ is the ordinary linear trace of multiplication on A. Over a splitting field the trace equation is a hyperplane with every coefficient nonzero, and since $\dim_K A \geq 2$ it is contained in no coordinate hyperplane. Thus its unit open has a K-point. Scale y and v by λ. In the formally traceless case no scaling is needed.

Concatenate the local columns to form a K-column and use the sum of the local trace rows $x \mapsto \mathrm{Tr}_{A/K}(vx)$ as a K-row. Their product is a nonzero rank-one matrix R, with trace zero by construction, so $R^2 = 0$. At every embedding its diagonal block is the chosen local value, since the trace is the sum of the embeddings. Each off-diagonal rectangle is attained by a unique original Y-block by the mixed determinant conditions. Uniqueness gives descent over K. The resulting frozen inputs and y are traceless by the trace formula for restriction of scalars and because off-diagonal rectangles contribute zero trace.

The image of the original Y-map after splitting contains all off-diagonal rectangles and the individual local traceless-slot images on the diagonal. A normal P to this image and to $[M_n, R]$ is block diagonal and commutes with R. At each block it satisfies $P_i u_i = \mu u_i$ and $v_i P_i = \mu v_i$ for a single scalar μ, by the equality of the two rank-one products. If f is formally traceless, $P_i - \mu I$ belongs to the local normal space, and transversality implies $P_i = \mu I$. If f is traceful, $\mathrm{tr}(P_i u_i v_i) = \mu v_i u_i = 0$, with $v_i u_i \neq 0$, so $\mu = 0$. Transversality gives $P_i = 0$. These two annihilators give precisely the asserted equality. □

We call a frozen tuple B of order n compatible with an order u if the two mixed rectangle determinants between B and an independent traceless tuple of order u are nonzero polynomials in that independent tuple. This includes the two orientations of the rectangle.

Proposition 4.2. *Fix $r_0 \geq 1$. There is an integer depending only on d, r_0 above which every $T \in V_n$ with $1 \leq \mathrm{rank}(T - \lambda I_n) \leq r_0$ for some $\lambda \in K$ is an original traceless value over any infinite K. Its frozen tuple can also be made compatible with any prescribed finite list of orders $u \geq d+1$ invertible in K.*

Proof. Choose two distinct primes a, b, different from the characteristic and greater than the bound in Proposition 4.1 and than $\max(d+1, 2r_0+1)$. They are bounded in terms of d, r_0, since it suffices to consider the first three primes beyond these bounds. At either order $s = a, b$, conjugate a source from Proposition 4.1 so its value is E_{12}. Denote its tuple by (B_0, y_0).

For any fixed nonscalar $S \in V_s$, change basis over K so $S_{12} = c \neq 0$. Such a basis exists by choosing x with x, Sx independent. For $D(z) = \mathrm{diag}(1, z^2, z, \ldots, z)$ set $C(z) = c^{-1} z^2 D(z) S D(z)^{-1}$. This has polynomial entries of degree at most four and $C(0) = E_{12}$. Over the purely transcendental field F of two independent full traceless tuples C_1, C_2, put $B(z) = B_0 + zC_1 + z^2 C_2$ and consider the equations

$$gh = I_s, \qquad f(B(z), y) = gC(z)h, \qquad y \in \mathfrak{sl}_s.$$

At $z = 0$ the point $(y, g, h) = (y_0, I_s, I_s)$ is rational. Differentiating the inverse equation gives $\delta h = -\delta g$. The relative value derivative is $L_{B_0}(\delta y) - [\delta g, E_{12}]$, with image $\mathrm{im}\, L_{B_0} + [M_s, E_{12}] = V_s$. In the formally traceless case one diagonal equation follows from the others using $gh = I_s$ and the trace. The remaining equations therefore have independent relative differentials. Fix

the coordinates outside a pivot minor at their values at the marked point. The resulting curve is étale over the z-line near zero, with a geometrically integral component through that point by the smooth F-point criterion [19]. Its degree is at most $\max(6, 2d-1)^{2s^2}$ by cumulative Bézout [13, Theorem 1.1]. There are at most $2s^2$ equations besides linear coordinate slices. The source term has degree at most $2d-1$ and the conjugated target term has degree at most six. On $z \neq 0$ replace y by $cz^{-2}y$ and g by $gD(z)$. This gives original-slot solutions to the fixed orbit equation $f(B, y) = gSg^{-1}$ on a smooth nonempty open of the curve.

The coefficient determinant for $B(z), B(w)$ as functions of C_1, C_2 is $zw(w-z)$. The mixed rectangle determinants are consequently nonzero on the product of two copies of the curve, and on products for different core orders with independent parameters. For a future order u in the statement, expand each mixed determinant as a polynomial in the coordinates of the future tuple and select a nonzero coefficient polynomial in the current tuple. Lemma 2.3 gives such a coefficient in both orientations. Its pullback along $B(z)$ is nonzero by genericity, so its zeros can be removed. The curve construction preceding Proposition 4.1 now gives compatible fixed orbit solutions over finite étale algebras of any two prescribed degrees L, M above a bound depending only on d, r_0. The degree estimate is uniform in the local targets. Removing finitely many points leaves the prescribed-degree threshold unchanged.

We specify the local targets before choosing these algebras. Let $r = \operatorname{rank}(T - \lambda I_n)$, and factor $T - \lambda I_n = U_0 V_0$ with $U_0 : K^r \to K^n$ injective and $V_0 : K^n \to K^r$ surjective. Put $Q = V_0 U_0$. For fixed n, r, this determines the similarity class of $U_0 V_0$ up to similarity of Q. Indeed, in a basis beginning with the columns of U_0, the matrix is $\left(\begin{smallmatrix} Q & R \\ 0 & 0 \end{smallmatrix}\right)$, with the map induced by R from K^{n-r} onto $\operatorname{coker} Q$ surjective. Any two such surjections are related by an automorphism of K^{n-r}. Their remaining difference has image in $\operatorname{im} Q$ and hence has the form QX for a linear map $X : K^{n-r} \to K^r$. Conjugating with $\left(\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right)$, with the inverse on the left, adds QX to the upper right block.

Write $n = aL + bM$ with L, M above the common degree bound and $L \neq 0$ in K. This is possible uniformly for all large n: choose L in its residue class modulo b within b of $n/(2a)$ and put $M = (n - aL)/b$. If $L = 0$ in K, replace (L, M) by $(L + b, M - a)$. Since a, b are bounded and are units, both degrees still grow uniformly and the new L is a unit.

In the formally traceless case set

$$H_a = -a\lambda E_{11} + \frac{Q + n\lambda E_{11}}{L}, \qquad H_b = -b\lambda E_{11} \quad \text{in } M_r(K).$$

Then $LH_a + MH_b = Q$ and $\operatorname{tr} H_s = -s\lambda$ for $s = a, b$, since $\operatorname{tr} Q = -n\lambda$. In the traceful case use $H_a = Q/L$, $H_b = 0$. For $s = a, b$ take $U_s : K^r \to K^s$ to be the inclusion in the first r coordinates and $V_s' : K^s \to K^r$ to have blocks $(H_s, I_r, 0)$. The two fixed local targets $S_s = \lambda I_s + U_s V_s'$ belong to V_s and are nonscalar because of their off-diagonal identity block.

Apply the orbit-curve construction to S_a, S_b with these L, M and let A_a, A_b be the resulting algebras. On $A_a^a \oplus A_b^b$ use the maps

$$U(x) = (U_a x, U_b x), \qquad V(v_a, v_b) = \operatorname{Tr}_{A_a/K}(V_a' v_a) + \operatorname{Tr}_{A_b/K}(V_b' v_b),$$

where the traces act coordinatewise. The map U is injective. The identity block and nonzero trace functional make V surjective. Since the traces of 1 in the two algebras are L, M, respectively, $VU = LH_a + MH_b = Q$. It follows that $\lambda I_n + UV$ is similar to T over K. Its diagonal embedding blocks after splitting the algebras are precisely S_a, S_b at the corresponding embeddings, since each trace row splits into its component rows.

Conjugate by the selected algebra-linear g's so these blocks are the actual local values. Freeze the other original slots to the selected algebra-linear tuples. Solve each remaining

rectangle in the original Y-slot using the invertible mixed operators. The solutions descend by uniqueness and all original inputs are traceless as in Proposition 4.1. Conjugating back realizes T. The nonzero future coefficient conditions ensure compatibility of the final frozen tuple. After splitting, each mixed determinant with a future tuple factors as the product of its local determinants, which are all nonzero polynomials in that tuple. Simultaneous conjugation of the frozen tuple changes the rectangle operator by invertible changes of bases and preserves this property. □

For the remaining targets we use the geometric scalar rank

$$\rho(A) = \min_{\lambda \in \overline{K}} \operatorname{rank}(A - \lambda I).$$

For an order-m matrix A, write $i(A) = m - \rho(A)$. This is the largest eigenspace dimension over $\overline{K}$, equivalently the number of nonconstant invariant factors in rational canonical form. Each invariant factor contributes one to the eigenspace dimension at each of its roots, and a root of the smallest factor is a root of them all.

Lemma 4.3. *Let K be infinite and $\rho(A) \geq 2s$. Then A is similar over K to a matrix with a zero leading diagonal block of order s. The remaining diagonal block D satisfies $i(D) \leq \max(i(A) - s, s)$. In particular if $i(A) \geq 2s$, then $\rho(D) \geq \rho(A)$.*

Proof. We construct U with $\dim U = \dim AU = s$ and $U \cap AU = 0$. If $u_1, Au_1, \ldots, u_j, Au_j$ are independent with $j < s$, let H be their span. If no further pair is independent modulo H, then Av modulo H is in the line of v for every $v \notin H$. Comparing v and $v + h$ for $h \in H$ shows $AH \subseteq H$, since the quotient has dimension at least two. The induced map on the quotient preserves every line and is scalar. This would give $\rho(A) \leq 2j < 2s$. Thus the required U exists, and its defining rank condition is open in the Grassmannian.

For every geometric eigenvalue λ, let $k_\lambda = \dim \ker(A - \lambda I)$. Impose in addition the open transversality conditions

$$\dim\big(U \cap \operatorname{im}(A - \lambda I)\big) = \max(0, s - k_\lambda).$$

There are only finitely many of them, and each is nonempty over $\overline{K}$. Their intersection with the preceding open has a K-point, since a Grassmannian affine cell has Zariski dense K-points even after extension of scalars. Choose a complement W to U containing AU, which makes the first compression zero. For a geometric subspace $H_\lambda = (A - \lambda I)^{-1}(U)$, a generic such W satisfies

$$\dim(W \cap H_\lambda) = \max\big(\dim H_\lambda - s,\ \dim(H_\lambda \cap AU)\big).$$

Indeed, quotient by AU. A complement is the kernel of a projection onto the image of U. Its restriction to the image of H_λ is fixed on the intersection with the image of U, and on a complement to this intersection its values can be prescribed arbitrarily. Write h for the dimension of the image of H_λ in the quotient and α for the dimension of its intersection with the image of U. On this intersection the projection is the identity. The generic rank on the whole image is therefore $\alpha + \min(s - \alpha, h - \alpha) = \min(s, h)$. Adding back $H_\lambda \cap AU$ gives the displayed formula. The finitely many resulting rank conditions hold simultaneously for a K-complement since the space of these projections is affine and K is infinite.

For old eigenvalues, $\dim H_\lambda = \max(k_\lambda, s)$, so the formula gives $\dim \ker(D - \lambda I) \leq \max(k_\lambda - s, s)$. For any other λ, $A - \lambda I$ is invertible and $\dim H_\lambda = s$, giving the same overall bound. Taking the maximum proves the claim. □

Proposition 4.4. *There is an integer depending only on d above which every nonscalar $T \in V_n$ over an infinite field K is a value of f on traceless inputs.*

Proof. Among $d+1, d+2, 2d+3$ choose two orders invertible in K. The three orders are pairwise coprime. Their pairwise two-generator conductors are at most $S = 2(d+1)^2$, so $S, S+1$ are sums of the chosen orders. Mixed determinant nonvanishing holds between these larger orders. To see this, restrict each frozen tuple to be block diagonal with the chosen small orders. The operator on its rectangle is the direct sum of the small rectangle operators, and its determinant is their product, a nonzero polynomial by Lemma 2.3. The same argument applies when one side is fixed and is compatible with the two small orders.

Fix an order $h \geq 2S$ above the bound in Proposition 4.2 for $r_0 = 1$. If f is traceful, apply that proposition to E_{11} at order h, requiring compatibility with the two small orders. It gives a fixed frozen tuple compatible with $S, S+1$. Scaling its free input gives τE_{11} for any $\tau \in K$, retaining the same frozen tuple.

For a target of sufficiently large $\rho(T)$, repeatedly apply Lemma 4.3 at $s = S$ while the remainder has $i \geq 2S$. Its scalar rank does not decrease, its trace is preserved, and the process leaves zero diagonal blocks of order S and a remainder A of order $m \geq \rho(T)$ with $k = i(A) < 2S$. In rational canonical form, replace each companion by a Fiedler companion [5, Theorem 2.3]. The needed form is the product of its odd-indexed factors followed by its even-indexed factors. For a monic polynomial $x^\ell + a_1 x^{\ell-1} + \cdots + a_\ell$, factor $j < \ell$ is the identity with the square on $j, j+1$ replaced by $\left(\begin{smallmatrix} -a_j & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and the last factor is $\operatorname{diag}(1, \ldots, 1, -a_\ell)$. This product is pentadiagonal and has only the first diagonal entry possibly nonzero. Its similarity to the companion holds over the coefficient field. Thus the only possibly nonzero scalar diagonal entries of A are its k companion heads, and the undirected graph in which distinct coordinates i, j are adjacent when A_{ij} or A_{ji} is nonzero has maximum degree at most four.

To justify the companion product over K, denote its factors by $F_1, \ldots, F_\ell$. Nonadjacent factors commute, and F_j is invertible for $j < \ell$. In a product order with an adjacent pair of indices out of order, take the smallest j for which F_{j+1} precedes F_j. Both neighbors of F_j precede it (only F_{j+1} when $j = 1$). Commute F_j to the end and move it to the front by conjugation. This corrects j and can introduce at most $j-1$ as an out-of-order pair. The sum of the out-of-order indices decreases strictly. The increasing product, by multiplication of the displayed factors, has first row $(-a_1, \ldots, -a_\ell)$, unit subdiagonal, and other entries zero. The odd and even products each have bandwidth one. Their product has bandwidth two. For each index beyond the first, one product has diagonal entry zero, and an off-diagonal pair cannot contribute to the diagonal because its two entries would have to come from opposite parities. This also proves the stated diagonal pattern.

By similarity on just the heads, with one additional coordinate if needed, we concentrate the trace τ at one head and make the other scalar diagonal entries zero. In fact the head compression is diagonal, and if scalar nonzero, adjoining a nonhead makes it nonscalar, since the additional diagonal entry is zero. A nonhead exists because $m \geq \rho(T) > 2S > k$. A zero head compression already has the desired diagonal. For a nonscalar compression, the prescribed-diagonal argument works over any field. To prescribe the first entry γ, choose x, Ax independent and begin a basis with $x, Ax - \gamma x$. Its first column is $(\gamma, 1, 0, \ldots)^t$. In order at least three, conjugating by $P = I + (1 - A_{23})E_{13}$ in the convention $P^{-1}AP$ preserves the first entry and changes its $(2,3)$-entry by $(1 - A_{23})A_{21} = 1 - A_{23}$, making the trailing compression nonscalar. The first diagonal entry stays fixed because $A_{31} = 0$. Induct on the order, with the last diagonal entry forced by the trace. Here the trace prescription is $(\tau, 0, \ldots, 0)$.

The similarity used at most $2S$ coordinates. Every mixed vertex among these can have at most $2S - 1$ neighbors inside and at most $8S$ outside, since each old vertex had at most four neighbors. An outside vertex has at most four old outside neighbors and at most $2S$ new

neighbors. Hence the graph now has maximum degree at most $10S-1$. If $\tau \neq 0$, greedily choose h independent vertices including the one carrying the trace, which is possible when $m \geq 10Sh$. This yields a distinguished diagonal block τE_{11} of order h. Remove its vertices. If $\tau = 0$, keep the whole graph and use no distinguished block.

A graph of maximum degree Δ with sufficiently many vertices can be partitioned into independent sets of sizes $S, S+1$. If the number of vertices is $m' \geq 30S^2$ and $\Delta \leq 10S-1$, first partition into those sizes, using $S\lfloor m'/S \rfloor$ and distributing the remainder by increasing individual sizes by one. Choose a partition minimizing the number of edges internal to a part. If v is incident to such an edge in a part J, choose w outside J, with no neighbor in $J \setminus \{v\}$, and whose part contains no neighbor of v. The excluded vertices are at most $S+1+S\Delta+(S+1)\Delta < 30S^2$. Swapping v, w removes at least one internal edge and creates none, a contradiction. Applying this to the remaining graph yields zero diagonal blocks of orders $S, S+1$. In particular, every step is valid for $\rho(T) \geq \max(10Sh, h+30S^2)$.

On each zero diagonal block use an independent traceless frozen tuple and set its diagonal free input to zero. On the distinguished block use the fixed compatible source, scaled in its free slot by τ. All ordered mixed determinants are nonzero polynomials by the first paragraph and the choice of the distinguished tuple. Choose the independent tuples over K so that all these determinants are nonzero. In each off-diagonal rectangle the original free slot can then be solved uniquely for the corresponding target block. Since all frozen matrices are block diagonal, no rectangle influences any other block of the value. The resulting inputs are all traceless and realize a K-conjugate of T.

Finally let C be the preceding degree-dependent bound on geometric scalar rank. For $n > 2C$ and $\rho(T) < C$, the number of nonconstant invariant factors exceeds $n/2$. Their degrees sum to n, so the smallest has degree one. Its root λ lies in K and realizes $\rho(T)$. If T is nonscalar, $1 \leq \operatorname{rank}(T-\lambda I_n) < C$, and Proposition 4.2 applies with $r_0 = C-1$ for all sufficiently large n. Combining the two bounds gives the proposition. □

5. Nonscalar values over finite fields

For a tuple B of frozen inputs in order s, write $L_B(y) = f(B,y)$ for the linear map on the designated slot Y, restricted to $\mathfrak{sl}_s$. Over a field E we call a tuple (B,y) *strong* if all inputs are traceless, its value $R = L_B(y)$ is a nonzero rank-one nilpotent, and

$$\operatorname{im} L_B + [M_s(E), R] = V_s(E). \tag{1}$$

We use the following two point estimates. If X is a geometrically integral closed affine variety of dimension $j > 0$ and degree δ over a field of size Q, and $Q > 2(j+1)\delta^2$, then

$$\left|\#X(\mathbb{F}_Q) - Q^j\right| \leq (\delta-1)(\delta-2)Q^{j-1/2} + 5\delta^{13/3}Q^{j-1}.$$

If a polynomial G of degree g is nonzero on X, then

$$\#(X \cap \{G=0\})(\mathbb{F}_Q) \leq \delta g Q^{j-1}.$$

The first estimate is [4, Theorem 7.1]. In the second, the sum of the degrees of the geometric components of $X \cap \{G=0\}$ is at most δg, and each has dimension at most $j-1$. The estimate follows from [13, Theorems 1.1 and 2.1]. We use the same degree inequality for components of intersections of several equations.

5.1. **Strong tuples in two coprime orders.**

Lemma 5.1. *There are constants A_0 and Q_0, depending only on d, with the following property. Over every finite field K, for each nonzero f as in the main theorem, there are distinct primes*

$a, b \le A_0$, each at least $\max(3, d+1)$ *and different from* $\operatorname{char} K$*, such that both orders admit strong tuples over every finite extension E/K of cardinality at least Q_0. Both tuples use the same designated slot Y.*

Proof. Take the two coprime core orders from Lemma 2.6, and let M be their degree-only upper bound. For either core order s, consider the affine incidence

$$f(B, y) = uv, \qquad B_i, y \in \mathfrak{sl}_s, \quad u \text{ a column of length } s, \quad v \text{ a row of length } s.$$

By Lemma 3.2, over $F((z))$, for a purely transcendental extension F/K, there is a point with $uv \neq 0$ for which

$$\operatorname{im} L_B + \{u\, v' + u'\, v : \ u' \in F((z))^{s\times 1}, \ v' \in F((z))^{1\times s}\} = M_s(F((z))).$$

If f is formally traceful, the point also has $vu \neq 0$. The displayed condition is exactly surjectivity of the relative Jacobian of the incidence over the affine space of frozen tuples. Thus the projection is smooth at the point. The Laurent series field $F((z))$ contains no algebraic extension of K: an element algebraic over K generates a finite field, and a nonzero element of a finite field in $F((z))$ has valuation zero and is determined by its residue. Indeed an element of $1 + zF[[z]]$ of finite order prime to $\operatorname{char} K$ is 1, by comparing its first nonzero term with that of its power. The residue is algebraic over K in F, hence lies in K. As K is perfect, $F((z))/K$ is regular. Over the compositum of $F((z))$ and $\overline{K}$, the smooth point lies on a unique geometric irreducible component. Regularity allows each Galois automorphism of $\overline{K}/K$ to act on this compositum fixing $F((z))$. It fixes the point, hence fixes the component. Its defining ideal descends to K, giving the required geometrically integral component. The smooth projection has open image near the point, so that component dominates the full frozen space. The Jacobian condition and, when required, $vu \neq 0$ hold on a nonempty open of the component. Its affine degree is at most d^{s^2}, since the incidence has s^2 equations of degree at most d.

Let A_0 be the largest of the first three primes greater than $\max(M^2, d+1)$. Two of those primes differ from $\operatorname{char} K$. Denote them by a, b. Each is a sum of at least two core orders, because the two core orders are coprime and their numerical semigroup contains all integers at least the product of their predecessors. For either order a or b, write the chosen sum as $\sum_{i=1}^{\ell} s_i$ and take the product of the corresponding incidence components. The product is geometrically integral, and its map to the independent frozen spaces is dominant. Hence all ordered mixed rectangle determinants from Lemma 2.3 can be required nonzero at once. Require also $u_i v_i \neq 0$ and the relative Jacobian condition on each block. If f is formally traceful, restrict further to $v_i u_i \neq 0$ and introduce nonzero scalars λ_i satisfying

$$\sum_{i=1}^{\ell} \lambda_i v_i u_i = 0.$$

Over the locus where each $v_i u_i$ is nonzero this equation describes a vector bundle of rank $\ell - 1$. The condition that all λ_i are nonzero gives an open meeting every geometric fibre, because $\ell \ge 2$. Thus the resulting variety is still geometrically integral and dominant over the independent frozen spaces. If f is formally traceless use $\lambda_i = 1$.

The closure of the variety has dimension and degree bounded in terms of d, A_0. Indeed the product of the component degrees is at most $d^{\sum_i s_i^2}$, and the additional equation has degree 3. The required open is obtained by removing one hypersurface of bounded degree. One can take the product of one nonzero relative Jacobian minor for each block, one nonzero coordinate of each u_i, v_i, all mixed determinants, and, if f is formally traceful, the factors λ_i and $v_i u_i$. These minors and determinants have bounded sizes and degrees since $\sum_i s_i \le A_0$.

The point estimates therefore give an E-point on the required open for every $|E| \geq Q_0$, with Q_0 depending only on d.

At such a point, concatenate the columns u_i and the rows $\lambda_i v_i$ to obtain u and v. Freeze block diagonal tuples with blocks B_i. Take the diagonal blocks of the original free input to be $\lambda_i y_i$, and solve for its off-diagonal blocks using the invertible mixed rectangle operators, so the value is $R = uv$. This value is rank one and has trace zero. All inputs are traceless.

For (1), use the trace pairing. A matrix annihilating $\operatorname{im} L_B$ and $[M_s(E), R]$, where $s = \sum_i s_i$, is block diagonal by mixed invertibility and commutes with uv. Thus its blocks P_i satisfy $P_i u_i = \mu u_i$ and $v_i P_i = \mu v_i$ with one common μ. If f is formally traceless, $P_i - \mu I_{s_i}$ annihilates both the local image and the tangent space at $u_i v_i$, so the relative Jacobian condition gives $P_i = \mu I_{s_i}$. If f is formally traceful, annihilation of $u_i v_i$ gives $\mu v_i u_i = 0$, hence $\mu = 0$. The same Jacobian condition now gives $P_i = 0$. The annihilator is therefore the scalar line or zero, respectively, which proves (1). □

5.2. **Incidence with a conjugacy class.**

Lemma 5.2. *Suppose order s admits a strong tuple over a finite field E. For every nonscalar $S \in V_s(E)$, the incidence*

$$gh = I_s, \qquad f(B, y) = gSh, \qquad B_i, y \in \mathfrak{sl}_s$$

has a geometrically integral E-component dominating the frozen space. Its degree is at most $(2d)^{s^2}$.

Proof. Conjugate the strong tuple so that its value is E_{12}, and write B_0 for its frozen tuple. The map

$$\mathrm{GL}_s \times \mathfrak{sl}_s \longrightarrow V_s, \qquad (g, y) \longmapsto g^{-1} L_{B_0}(y) g$$

is smooth at the resulting source by (1). Choose a basis over E in which $S_{12} = c \neq 0$. Such a basis exists because some vector x has $Sx \notin Ex$. Take $D(z) = \operatorname{diag}(1, z^2, z, \ldots, z)$. Then $c^{-1} z^2 D(z) S D(z)^{-1}$ extends polynomially to $z = 0$ with value E_{12}. Smoothness lifts this path to an $E[[z]]$-point through the source. Over $E((z))$, absorb the invertible scalar into the linear input y and the conjugating $D(z)$ into g. This gives an incidence point over $E((z))$, where the relative differential in y, g still surjects onto V_s. When $V_s = \mathfrak{sl}_s(E)$, the trace relation removes one value equation on the locus $gh = I_s$. Thus the remaining equations satisfy the relative Jacobian criterion over the frozen space. Regularity of $E((z))/E$ and uniqueness of the geometric component through this point give the asserted E-component, as in the preceding proof. Its projection is dominant by smoothness. The affine model uses s^2 inverse equations of degree 2 and at most s^2 value equations of degree at most d, giving the stated degree bound. □

5.3. **Prescribing the traces of the diagonal blocks.** Let $A = \prod_i E_i$ be a finite product of finite separable field extensions E_i/K. For an A-module $W = \bigoplus_i E_i^{s_i}$, the A-linear part of a K-endomorphism T of W is obtained by retaining its diagonal embedding blocks after passing to a separable splitting field of A. The projection descends to K, since its block description is Galois invariant. Its diagonal factor in $M_{s_i}(E_i)$ has ordinary E_i-matrix trace. Collect these traces into $h_T \in A$. If $\rho(\alpha)$ denotes multiplication by $\alpha \in A$ on W, then

$$\operatorname{Tr}_{A/K}(\alpha h_T) = \operatorname{tr}_K(T \rho(\alpha)). \tag{2}$$

Here $\operatorname{Tr}_{A/K}$ is the sum of the field traces of the factors. The identity follows after splitting A, and its pairing in α is nondegenerate.

Lemma 5.3. *Suppose $A = \prod_i E_i$ is generated by one element over K and $s_i \geq 3$ for every i. A nonscalar K-endomorphism T on a space of dimension $\sum_i s_i[E_i : K]$ can be represented on $\bigoplus_i E_i^{s_i}$ so that each diagonal factor of its A-linear part is nonscalar and $h_T = h$, for any $h \in A$ with $\operatorname{Tr}_{A/K}(h) = \operatorname{tr}_K(T)$.*

Proof. Put $e = \dim_K A$. First, for every $1 \leq k < e$ there is a k-dimensional subspace H of A whose multiplicative stabilizer $\{\alpha : \alpha H \subseteq H\}$ is $K1$. Write $A = K[\theta]$ with power basis $1, \theta, \ldots, \theta^{e-1}$. For $k \leq e/2$, take the first k basis vectors. A stabilizing α belongs to H, and if its highest power of θ has positive exponent, multiplication by θ^{k-1} gives highest exponent between k and $2k - 2 < e$. This contradicts stability. For $k > e/2$, use the orthogonal complement of such a subspace of dimension $e - k$ under the trace pairing. Commutativity gives the same stabilizer for a subspace and its orthogonal complement.

Choose an e-dimensional subspace U containing a vector x_0 but excluding Tx_0, with $Tx_0 \notin Kx_0$, and a complement W' to U. The lower block $C : U \to W'$ of T is nonzero. We make $U = A$ and give W' the required complementary A-module structure, with at least two regular summands. If $k = \operatorname{rank} C < e$, identify $\ker C$ with a subspace H of dimension $e - k$ as above. For a basis $w_1, \ldots, w_k$ of $\operatorname{im} C$, write

$$C(x) = \sum_{j=1}^{k} w_j \operatorname{Tr}_{A/K}(\alpha_j x).$$

The α_j form a basis of $H^\perp$. They generate the unit ideal of A: otherwise they vanish on a field factor, that factor is contained in H, and its nontrivial idempotent stabilizes H. In the one-factor case every nonzero α_j is a unit. Choose $x_j \in A$ with $\sum_j x_j \alpha_j = 1$, and linearly independent $z_j \in A$. Identify w_j with $(x_j, z_j, 0)$ in the first two regular summands and the remaining complement. These vectors are independent, so the identification extends to all of W'. After splitting, the first coordinate of the A-linear part on an embedding σ is multiplication by $\sum_j \sigma(x_j)\sigma(\alpha_j)$, since the trace in the formula for $C(x)$ is the sum of its embedding coordinates and only the diagonal coordinate is kept. This sum is one. Moreover any $\alpha \in A$ intertwining C preserves H, hence lies in $K1$.

If $\operatorname{rank} C = e$, identify $C(x)$ with $(x, \operatorname{Tr}_{A/K}(x)1, 0)$ in the same decomposition of W'. This is again an identification on its image and extends to all of W'. The first coordinate is the identity. An intertwining α must satisfy

$$\operatorname{Tr}_{A/K}(\alpha x)1 = \alpha \operatorname{Tr}_{A/K}(x) \qquad (x \in A).$$

Taking an x of trace 1 gives $\alpha \in K1$. In both cases the identity block in the lower A-linear part makes every diagonal factor of the A-linear part of T nonscalar.

Finally conjugate T by $\left(\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right)$, with $X : W' \to U$, taking the inverse on the left. The lower block C stays fixed, and the two diagonal blocks change by $-XC$ and CX. If $\rho_U, \rho_{W'}$ denote the two multiplication actions, the change in the right side of (2) is

$$\operatorname{tr}_U(X(\rho_{W'}(\alpha)C - C\rho_U(\alpha))).$$

The rectangular trace pairing is nondegenerate. The annihilator in A of these changes is precisely the space $K1$ of intertwiners. Their image in the dual of A is therefore the full hyperplane of functionals vanishing at 1. The desired change in (2) lies in this hyperplane by the hypothesis on h, so some X makes $h_T = h$. The lower identity blocks still ensure nonscalarity. □

5.4. **Frobenius avoidance and descent.**

Proposition 5.4. *There is a bound depending only on d such that, for every finite field K, every nonzero f as in the main theorem, and every n at least that bound, each nonscalar $T \in V_n(K)$ is a value on traceless inputs over K.*

Proof. Write $K = \mathbb{F}_q$. Take a, b from Lemma 5.1. For all sufficiently large n write $n = aL + bM$ with positive unequal integers L, M, each at least $n/(4A_0)$, where A_0 is the bound in that lemma. Choose L in the required residue class modulo b within b of $n/(2a)$, and set $M = (n - aL)/b$. If they are equal, replace them by $L + b, M - a$. The errors from $n/(2a)$ and $n/(2b)$ are bounded in terms of A_0. Put $E_L = \mathbb{F}_{q^L}$, $E_M = \mathbb{F}_{q^M}$ and $A = E_L \times E_M$. The algebra A is generated by one element over K, since field generators in the two factors have irreducible minimal polynomials of distinct degrees. Apply Lemma 5.3 to T with multiplicities a, b, choosing $h = 0$ when f is formally traceless. In the other case take any h of the required total trace, which exists by nondegeneracy. The two diagonal factors $S_a \in M_a(E_L)$ and $S_b \in M_b(E_M)$ of the A-linear part are nonscalar and belong to $V_a(E_L), V_b(E_M)$, respectively.

For each order $s \in \{a, b\}$, let e be its extension degree L or M. Use the determinants $D_{s,t}$ of Lemma 2.3 and the entrywise powers $B^{[q^j]}$ of Lemma 2.4. By Lemma 2.4 the polynomial

$$\Phi_{s,e}(B) = \prod_{j=1}^{\lfloor e/2 \rfloor} D_{s,s}(B, B^{[q^j]}) D_{s,s}(B^{[q^j]}, B)$$

is nonzero. Its degree is at most

$$2(d-1)s^2 \left(\lfloor e/2 \rfloor + \sum_{j=1}^{\lfloor e/2 \rfloor} q^j \right) \leq 6(d-1)s^2 q^{e/2}.$$

The inequality uses $\sum_{j=1}^{k} q^j \leq 2q^k$ and $k \leq q^k$. Nonvanishing gives both ordered mixed inverses between every pair of distinct embedding blocks of the same extension. In fact, apply Frobenius to reduce an ordered pair to one starting at the identity embedding, and use its reverse if the difference exceeds $e/2$.

For the first choice of frozen tuple B of order a, also require that a nonzero coefficient of each of $D_{a,b}(B, C)$ and $D_{b,a}(C, B)$, viewed as polynomials in the independent order-b tuple C, remains nonzero. Such coefficients exist by Lemma 2.3 and have bounded degrees. For the second tuple C, after choosing B, require nonvanishing of

$$\prod_{j=0}^{L-1} D_{a,b}(B^{[q^j]}, C) D_{b,a}(C, B^{[q^j]}).$$

This is a nonzero K-polynomial of degree at most $2L(d-1)ab$. The coefficient conditions ensure each factor is nonzero, and Frobenius permutes the factors. Nonvanishing therefore gives all ordered mixed inverses between the conjugates of the first and second tuples.

Both choices can be made on the respective orbit components from Lemma 5.2. The strong tuples needed for the lemma exist once both extension sizes exceed Q_0. On each component the avoidance polynomial remains nonzero by dominance over the frozen space. The component has positive dimension bounded in terms of d, A_0 and degree at most $(2d)^{A_0^2}$. If $Q = q^e$ is the size of the current extension, the number of admissible points divided by Q^j, where j and δ are the component's dimension and degree, is at least

$$1 - \frac{(\delta - 1)(\delta - 2) + 6\delta(d-1)s^2}{\sqrt{Q}} - \frac{5\delta^{13/3} + \delta C(1 + L)}{Q}.$$

Here C depends only on d, A_0 and bounds the degree of the cross conditions divided by $1+L$, including the two coefficient conditions on the first tuple. Since $L, M \geq n/(4A_0)$, $L \leq n$, and $q \geq 2$, the required size conditions hold and the losses are less than 1 for all sufficiently large n, uniformly in q, f, T.

The resulting local values are conjugates of S_a, S_b over E_L, E_M, respectively. Conjugate the entire K-linear target by the direct sum of these two local conjugating maps, which is a K-linear automorphism. Freeze the $d-1$ original inputs as the corresponding $E_L \times E_M$-linear matrices. Over a finite common splitting field, their blocks are the Frobenius conjugates of the two chosen tuples. Use the conjugates of the two chosen local Y-inputs for the diagonal embedding blocks of Y. For each ordered pair of distinct blocks, solve the invertible mixed rectangle equation with right hand side the prescribed target block. Each solution is unique, so Frobenius equivariance of the equations makes the collection of solutions equivariant. Thus Y descends to K. All inputs are traceless: the diagonal blocks have trace zero and the off-diagonal blocks contribute no trace. Original multilinearity in Y gives the prescribed value in every block. Undoing the target's changes of basis realizes T over K. □

6. Scalar values in positive characteristic

Lemma 6.1. *Suppose* $\operatorname{char} K = p > 0$ *and write* $n = p^a m$ *with* $p \nmid m$. *Assume* K *contains a primitive* ℓ*-th root of unity for each prime* $\ell \mid m$. *If* I_n *belongs to the linear span of* $f(\mathfrak{sl}_n(K)^d)$, *then* I_n *belongs to* $f(\mathfrak{sl}_n(K)^d)$.

Proof. Index the matrix units by the abelian group $G = P \times A$, where $P = (\mathbb{Z}/p\mathbb{Z})^a$ and A is the product, over the primes $\ell \mid m$, of $v_\ell(m)$ copies of $\mathbb{Z}/\ell\mathbb{Z}$. Here $v_\ell(m)$ is the exponent of ℓ in m. Give E_{xy} degree $y - x \in G$. The degree-zero matrices are diagonal. This is an algebra grading, and $\mathfrak{sl}_n(K)$ is the direct sum of its homogeneous pieces. Conjugation by a translation of G permutes matrix indices and preserves each piece. Its action on the diagonal is the regular permutation action.

Project onto degree zero and then onto the A-invariant diagonal matrices by averaging over A. The composite fixes I_n, so some value of f has nonzero projection. Multilinearity gives such a value on homogeneous inputs of total degree zero. Because the exponent of A splits over K and is prime to p, each input space decomposes into K-valued character spaces. Restricting to a suitable tuple of characters with trivial product gives a nonzero multilinear map with A-invariant diagonal values. Its input spaces are P-stable.

The commuting translations of P act unipotently on each such space. They have a common triangular basis over K: intersect fixed spaces successively to find a common fixed vector and use induction on the quotient. In these bases choose a tuple with nonzero value and minimal sum of indices. Translating every basis vector changes it by lower-index terms. All resulting terms of smaller total index vanish, so the chosen value is P-fixed by equivariance. Being also A-fixed and diagonal, it is a nonzero scalar multiple of I_n. Scale one original input to obtain I_n. □

The image span commutes with extension of scalars because it is spanned by values on input basis tuples. Over $\overline{K}$, the span of the conjugates of a nonscalar matrix contains $\mathfrak{sl}_n$. To see this, first choose a basis in which the matrix has an off-diagonal nonzero entry, using a vector x such that x, Ax are independent for the matrix A in question. Conjugating by $\operatorname{diag}(t^{2^1}, \ldots, t^{2^n})$ multiplies the ij entry by $t^{2^i - 2^j}$. The nonzero differences $2^i - 2^j$ for $i \neq j$ are distinct, as seen from their sign and 2-adic valuation. The corresponding Laurent monomials and the constant function are linearly independent on the infinite field $\overline{K}$. Thus their evaluation vectors span the full coefficient space, and linear combinations of the conjugates isolate the matrix unit of

a nonzero off-diagonal entry. Permutations give all off-diagonal units. Conjugation of E_{ij} by $I_n + E_{ji}$ gives $E_{ij} + E_{jj} - E_{ii} - E_{ji}$, so the span also contains every diagonal difference.

A nonzero scalar value itself gives I_n in the image span. In its absence, if the image is nonzero and $p \mid n$, the nonscalar argument gives $\mathfrak{sl}_n$ and hence I_n. If there is a value of nonzero trace, adjoining that value to $\mathfrak{sl}_n$ gives M_n. Membership in the span descends by linear algebra.

Proposition 6.2. *In positive characteristic, there is a bound depending only on d beyond which every scalar in V_n is attained on traceless inputs if f is nonzero and formally traceless. If f is formally traceful, the same holds in every fixed positive characteristic, with a bound depending only on d and that characteristic.*

Proof. Direct sums add the orders of identity values. We apply the conductor formula to coprime orders and to orders whose greatest common divisor is p, after dividing by p. A restriction of scalars from a finite extension L/K of degree e sends an identity value of order s to one of order es. Indeed each matrix entry acts by multiplication on L in a K-basis. This substitution preserves products and sums, and the trace is the sum of the field traces of the original diagonal entries. In particular it preserves tracelessness.

First assume f is formally traceless. A nonzero scalar can occur in V_n only if $p \mid n$. Let a be the least p-power at least $d+1$. The directed path witness and Lemma 6.1 give I_a. For odd p, take b to be the least power of two at least $(d+1)/p$, allowing $b = 1$. The same lemma gives I_{pb} since the order is at least $d+1$ and the required roots are already in K. The reduced orders $a/p, b$ are coprime. If $p \geq d+1$, $a/p = 1$. Otherwise $a/p < d+1$ and $pb < 2(d+1)$, so the bound $p(a/p-1)(b-1)$ is less than $2(d+1)^2$.

If $p = 2$, let w be the least power of seven at least $(d+1)/2$. In $L = K\mathbb{F}_8$, formed inside an algebraic closure, the seventh roots split, so I_{2w} is attained. The extension degree is 1 or 3 since $\mathbb{F}_8/\mathbb{F}_2$ is cyclic of degree 3 and $\mathrm{Gal}(L/K)$ injects into that cyclic group. Restriction and direct sum yield I_{6w} over K. The reduced orders $a/2, 3w$ are coprime. Since $a/2 < d+1$ and $w < 7(d+1)/2$, the resulting bound for even orders is at most $21(d+1)^2$.

Now let f be formally traceful. For odd p, the cycle witness and Lemma 6.1 give I_a and I_b, with a the least p-power and b the least two-power at least $d+1$. They are coprime. For $p = 2$, take a the least two-power and w the least seven-power at least $d+1$. The same extension argument gives the coprime orders $a, 3w$. In either case this supplies all large identity orders. Scaling one input gives every scalar multiple, and zero is attained by a zero input. □

7. Scalar values in characteristic zero

Proposition 7.1. *Let K have characteristic zero and let f be formally traceful. Then for every sufficiently large n, the identity I_n is a value of f on $\mathfrak{sl}_n(K)^d$.*

7.1. Disconnected factor graphs.

We use the following cyclotomic fact for coefficients in an arbitrary characteristic-zero field.

Lemma 7.2. *Let K have characteristic zero and let $c_1, \dots, c_t \in K$ be nonzero. For all but at most $t-1$ primes $\ell > t$, the following implication holds for every $e_1, \dots, e_t \in \mathbb{Z}/\ell\mathbb{Z}$ and every primitive ℓ-th root ζ in an algebraic closure of K: if $\sum_i c_i \zeta^{e_i} = 0$, then $\sum_{i:e_i=r} c_i = 0$ for every $r \in \mathbb{Z}/\ell\mathbb{Z}$.*

Proof. For each exceptional prime choose a witnessing vector $(\zeta^{e_1}, \dots, \zeta^{e_t})$. All these vectors are in the hyperplane annihilated by $c = (c_1, \dots, c_t)$. If one lies in the span of vectors from distinct earlier primes, the same holds for all its cyclotomic conjugates. Indeed, the span

condition is a rank condition over the compositum of the cyclotomic fields, and the automorphisms of the new prime cyclotomic field extend fixing the previous fields. This follows from multiplicativity of Euler's totient function on coprime conductors. If there are $h \le t$ distinct exponents in the new vector, the conjugates span their h class indicator vectors. To see this, use representatives $0 \le r < \ell$. A linear relation among the h columns $(\zeta^{ar})_{1\le a<\ell}$ gives a polynomial of degree at most $\ell - 1$ with fewer than ℓ possible terms that vanishes at all the nontrivial ℓ-th roots. It is a multiple of $1 + X + \cdots + X^{\ell-1}$ and must be zero. The class indicator vectors would therefore all be annihilated by c, contradicting the witness. Thus the witnessing vectors are independent and there are at most $t - 1$. □

Lemma 7.3. *If every comparison factor has disconnected graph, Proposition 7.1 holds.*

Proof. Use the common tree of Lemma 2.1. Every connected component of a factor graph belongs to its canonical tree and hence lies inside a root child of the common tree. Thus each factor quotient at the root is empty. Let C be the root child containing Y, and put $I = C \setminus \{Y\}$ and $B = \{1, \dots, d\} \setminus C$, identifying each letter with its index. If there are no factors, the root children are singletons. In particular $B \neq \varnothing$. Follow the path from Y to C. At each node where some factor has a nonempty quotient, use each sibling of the child containing Y as one part. At each other node use the union of these siblings as one part. The resulting parts partition I. In $\mathbb{Q}^{d-1}$, with coordinates indexed by the letters other than Y, let H be the subspace of b defined by $\sum_{i\in A} b_i = 0$ for each such part A. The coordinates b_i for $i \in B$ are free.

For a difference of support words $w - u$, let ϵ_i be the difference of the indicators that i follows Y. Let D be the comparison matrix defined in Section 2, with D_{ij} recording the difference of the indicators that i follows j. Rotate each word so that Y is last and delete Y. The comparison difference of the resulting words records precedence and gives the alternating matrix

$$R = -D + \epsilon \mathbf{1}^{\mathsf{T}} - \mathbf{1}\epsilon^{\mathsf{T}}, \qquad R = \begin{pmatrix} R_{II} & \epsilon_I \mathbf{1}_B^{\mathsf{T}} \\ -\mathbf{1}_B \epsilon_I^{\mathsf{T}} & R_{BB} \end{pmatrix}.$$

Here $\mathbf{1}$ denotes the all-ones vector of the indicated size. The block formula holds because C is consecutive in each word. It follows that $R = 0$ exactly when w, u are cyclic rotations.

Two consequences of the common block structure will be used. First, for a pair differing by a single factor flip, $\epsilon^{\mathsf{T}} b = 0$ on H. The flip changes no comparison across root children. Its Y-comparison is constant on each sibling at a nonempty quotient, and zero on a union at an empty quotient. Second, for any rational linear combination of word differences with $\epsilon_I = 0$, D_{II} has no entries across the parts of I. At distinct levels the outer sibling compares with the entire inner block as it compares with Y. At a nonempty quotient, comparisons depend on one orientation for a complete quotient, or on the two orientations of a prime graph and its complement. Each orientation is detected by a comparison with the child containing Y. Thus the cross-part comparison functions are affine linear functions of the Y-comparisons in I.

For any rational linear combination of word differences, if $Rb = 0$ on H, varying b_B gives $\epsilon_I = 0$ and $R_{BB} = 0$. Now $R_{II} = -D_{II}$ is block diagonal on the parts of I, and each alternating block kills the sum-zero hyperplane on its part. Its rank is at most one and hence zero. Thus $R = 0$. If in addition $\epsilon^{\mathsf{T}} b = 0$ on H, then $\epsilon = 0$, since b_B is free.

For an unrestricted vector $a \in \mathbb{Q}^{d-1}$ and $b \in H$, assign to each word difference the pair of polynomials

$$A(a, b) = a^{\mathsf{T}} R b, \qquad L(b) = \epsilon^{\mathsf{T}} b.$$

Partition the support words into blocks whose coefficients sum to zero, and choose one anchor word in each block. Write A_j, L_j for the differences from the anchors to the other words. We

claim that some nonzero polynomial in a, b precludes all equations $A_j + sL_j = 0$ simultaneously, for any s.

If all $L_j = 0$, each $A_j \neq 0$. Otherwise $R_j = \epsilon_j = 0$ by the preceding argument, which would identify distinct words. If some minor $A_jL_k - A_kL_j$ is nonzero, it supplies the desired polynomial. Assume all minors are identically zero. When two slopes are independent, unique factorization applied to $(R_jb)L_k = (R_kb)L_j$ gives $R_jb = vL_j$ for all j, with a common rational vector v. Write $e_j = \epsilon_{j,I}^{\mathsf{T}} b_I$. The B-coordinates give

$$R_{j,BB}b_B - e_j\mathbf{1}_B = v_B(e_j + \epsilon_{j,B}^{\mathsf{T}} b_B).$$

If $e_j \neq 0$, comparison in b_I gives $v_B = -\mathbf{1}_B$. Comparison in b_B and the zero diagonal of $R_{j,BB}$ then give $\epsilon_{j,B} = 0$. The I-coordinates, with b_B free, force $\epsilon_{j,I} = 0$, a contradiction. Thus all $e_j = 0$. Each $R_{j,BB} = v_B\epsilon_{j,B}^{\mathsf{T}}$ has rank at most one and is alternating, so it is zero. Some $\epsilon_{j,B} \neq 0$, whence $v_B = 0$. The I-coordinates give $\epsilon_{j,I}\mathbf{1}_B^{\mathsf{T}} = v_I\epsilon_{j,B}^{\mathsf{T}}$. If $v_I \neq 0$, all $\epsilon_{j,B}$ are multiples of $\mathbf{1}_B$, contrary to independence of two slopes. Hence $v = 0$ and every $R_j = 0$. Each balanced block would lie in a cyclic class. This would make f formally traceless.

It remains to consider proportional slopes with one $L_0 \neq 0$. Write $L_j = q_jL_0$ with $q_j \in \mathbb{Q}$. The vanishing minors and the faithfulness statement for R, ϵ give $(R_j, \epsilon_j) = q_j(R_0, \epsilon_0)$. Distinct words force $q_j \neq 0$, and their indicator differences force $q_j = \pm 1$. Within a block, the indicator vectors lie on a line in a Boolean cube, which contains at most two Boolean points. Repeated indicator vectors would also give repeated comparison vectors. Each block therefore consists of two opposite coefficients, with the same full comparison difference up to sign. The symbol is divisible by $1 - z^v$, where v is that full difference and z^v denotes the Laurent monomial in the comparison coordinates. The vector v is primitive because its nonzero entries are ± 1. But this factor has slope $L_0 \neq 0$ on H, contradicting the first consequence of the common tree. This proves the claim about the balanced partitions.

Multiply one nonzero witness for each balanced partition by $(\sum_i b_i)\prod_i a_i$, with all indices i here different from Y. The product is nonzero on $\mathbb{Q}^{d-1} \times H$: the sum uses the free coordinates indexed by B. Use integral coordinates on H by eliminating one coordinate in each part. For all sufficiently large primes ℓ the product is a nonzero polynomial modulo ℓ of total degree less than ℓ. The univariate root bound, applied successively in each coordinate, gives an $\mathbb{F}_\ell$-point where the product is nonzero. In particular $\sum_i b_i \neq 0$ and $a_i \neq 0$. Take ℓ also outside the exceptional set of Lemma 7.2.

Over an algebraic closure of K, let $Ue_j = \zeta^j e_j$ and $Ve_j = e_{j+1}$, where $j \in \mathbb{Z}/\ell\mathbb{Z}$ and ζ is a primitive ℓ-th root of unity. Freeze the input at each letter $i \neq Y$ to $U^{a_i}V^{b_i}$. In the free slot use

$$U^{-\sum_i a_i + s}V^{-\sum_i b_i}, \qquad s \in \mathbb{Z}/\ell\mathbb{Z}.$$

All word products are scalar multiples of U^s. Since $UV = \zeta VU$, the difference of the phases, written as exponents of ζ, is $A(a, b) + sL(b)$. Indeed with free exponents u, v, it is $-a^{\mathsf{T}}Db + u\epsilon^{\mathsf{T}}b - v\epsilon^{\mathsf{T}}a$, and substitution gives the formula. If a multiplier of U^s were zero, Lemma 7.2 would partition the words by equal phases into balanced blocks. The chosen point excludes this. The free linear map from $\{D_0V^{-\sum_i b_i} : D_0 \text{ diagonal}\}$ onto the diagonal space is therefore invertible. The displayed free inputs form a basis because $(\zeta^{sj})_{s,j}$ is invertible.

To realize the inverse over K, keep V and replace each frozen U^{a_i} by a diagonal matrix of parameters. Impose trace zero on this diagonal matrix when $b_i = 0$. These constraints contain the constructed specialization, since then $a_i \neq 0$ and $\sum_{j=0}^{\ell-1} \zeta^{a_ij} = 0$. For $b_i \neq 0$ the corresponding weighted shift is already traceless. The determinant of the free linear map is nonzero as a polynomial on the parameter space, and there is a K-point where it is nonzero

because K is infinite. Solving for I_ℓ gives a traceless free input since $\sum_i b_i \neq 0$. This works at two distinct primes, say ℓ and ℓ'. Direct sums give I_n for every $n \geq (\ell-1)(\ell'-1)$ by the conductor formula for two coprime integers. □

7.2. **Weighted shifts and comparison factors.** We will substitute one matrix for the product on each consecutive block of original letters. For this purpose a block is treated as a single letter, and its degree is the number of original letters in it. Fix an integer step $c_i \neq 0$ for each of s such letters. On the standard basis $e_0, \dots, e_{n-1}$ use weighted shifts $X_i e_j = x_{i,j} e_{j+c_i}$ if both indices are in range, and zero otherwise, with all available weights $x_{i,j} \in K^\times$. Each X_i is similar to $J_n^{|c_i|}$, where J_n is a nilpotent Jordan block. This follows by normalizing the weights along the chains of the shift. For an order v of the effective letters let $r_i(v)$ be the sum of the steps strictly after i in v. If the prefix sums, including the empty prefix, have minimum m_v, maximum M_v, and total h, the surviving starting indices are

$$[M_v - h, n + m_v - h) \cap \mathbb{Z}.$$

The entry from index j to $j+h$ is $\prod_i x_{i,j+r_i(v)}$ there. The formula follows by applying the letters right to left and testing whether each visited index lies in $\{0, \dots, n-1\}$.

For any list of orders and any n, let E have one row for each surviving monomial, indexed by (v, j), and one column for each available shift edge. Its entry records whether that edge occurs in the monomial. Let $M(T) = (T^{r_i(v)})_{i,v}$, with T an indeterminate. A row relation of E over any field, with coefficients $a_{v,j}$ extended by zero to every integer j, is equivalent to

$$M(T)\big(\textstyle\sum_j a_{v,j} T^j\big)_v = 0.$$

Indeed the coefficient at T^k in row i records the relation on edge (i, k), and a surviving monomial uses only available edges. If $M(T)$ has full column rank over $\mathbb{Q}(T)$ and $\mathbb{F}_p(T)$ for every prime p, then E has full row rank over all these coefficient fields, so its integer Smith invariants are all one. More generally, if the rank of E over $\mathbb{Q}$ and over every $\mathbb{F}_p$ is the same, the nonzero Smith invariants are one. In this case one can prescribe arbitrary nonzero values of the monomials subject to the multiplicative relations given by the integer left kernel of E. This is because the unimodular row and column operations in Smith reduction correspond to invertible monomial changes of coordinates on powers of $K^\times$. After these changes the image consists of arbitrary entries in the nonzero Smith positions and ones elsewhere.

To lift an effective letter on a common block to original inputs, take positive integers b_i indexed by the original letters in that block, with total $h \geq 1$. Given $q_i \in K^\times$ put $D(q_i) = \operatorname{diag}(1, q_i, \dots, q_i^{n-1})$ and evaluate the original letter i at $D(q_i) J_n^{b_i}$, using $J_n e_j = e_{j-1}$ with $J_n e_0 = 0$. For any word v on the block, let $\beta_i(v) = \sum_{j \text{ preceding } i \text{ in } v} b_j$. Then the product is

$$\Big(\prod_i q_i^{\beta_i(v)}\Big) D\Big(\prod_i q_i\Big) J_n^h.$$

For $n > h$, simultaneous conjugation of the original matrices can replace the common right factor by any prescribed conjugate of J_n^h.

Suppose the selected global comparison factor flips internal comparisons in this block. Let v_0 be the difference of the β vectors across a flip, independent of the base word. The vectors $(\beta_i(v) + b_i/2)_i$ all have the same squared norm for the inner product $\langle x, y\rangle = \sum_i b_i x_i y_i$. Indeed their coordinates are midpoints of consecutive intervals partitioning $[0, h]$, and the squared norm is $\int_0^h t^2\, dt - \sum_i b_i^3/12$. For each base word the equality of norms before and after the flip gives $2\langle \beta + b/2, v_0\rangle + \langle v_0, v_0\rangle = 0$. Subtracting this equality for two base words shows that their β-difference is orthogonal to v_0. The β-vector determines the order because

its coordinates strictly increase along the word. Consequently any two distinct base words have β-difference outside $\mathbb{Q}v_0$, unless $v_0 = 0$, in which case distinctness alone suffices. Take independent parameters $u_j \in K^\times$ for the letters of the block. If $v_0 \neq 0$, set

$$q_i = \prod_j u_j^{T_{ij}}, \qquad T = \Big(\sum_i b_i (v_0)_i^2\Big) I - \operatorname{diag}(b_i) v_0 v_0^{\mathsf{T}}.$$

For $v_0 = 0$ use $q_i = u_i$. In the nonzero case $v_0^{\mathsf{T}} T = 0$, and for each difference $\Delta\beta$ of distinct base words orthogonality gives

$$(\Delta\beta)^{\mathsf{T}} T = \Big(\sum_i b_i (v_0)_i^2\Big) (\Delta\beta)^{\mathsf{T}} \neq 0.$$

Hence the paired block products agree, and the coefficient functions for distinct base words on a block are distinct Laurent monomials. The prescribed matrix for the block can be held fixed by conjugating all its inputs together. These inputs remain nilpotent.

7.3. **Connected factor graphs.** Choose a connected factor graph of f and its canonical module tree. In each construction below, the matrices assigned to the blocks have the Jordan types required by the preceding construction. To see how the substitution recovers f, choose the set $\mathcal{W}$ of support words in one state of the selected factor and write $f = \sum_{w \in \mathcal{W}} c_w (w - \widetilde{w})$, where $\widetilde{w}$ denotes the paired word. For each exposed block C, let Z_C be its prescribed effective matrix and let $\mu_C(w)$ be the Laurent monomial from the construction, with $\mu_C(w) = \mu_C(\widetilde{w})$. If σ is the order of the blocks in w, write Z_σ for their ordered product and $\widetilde{\sigma}$ for the order in $\widetilde{w}$. The actual value is

$$\sum_\sigma \Big(\sum_{\substack{w \in \mathcal{W} \\ w \text{ has block order } \sigma}} c_w \prod_C \mu_C(w) \Big) (Z_\sigma - Z_{\widetilde{\sigma}}).$$

For a fixed σ, distinct base words differ internally in a block. The block parameters are disjoint, so the bracketed coefficient is a nonzero Laurent polynomial whenever the corresponding set of base words is nonempty. This permits a specialization with a nonzero coefficient, since K is infinite. In the constructions using products of powers of a single nilpotent, one first prescribes the matrix of each block as the appropriate power and then uses the same substitution.

For a complete root with at least three children, the effective pair is $1 \cdots s$ and its reversal. Choose steps $(c_1, \dots, c_s) = (d, \dots, d, -(s-1)d)$. The prefix ranges of the orders are $[0, (s-1)d]$ and $[-(s-1)d, 0]$. The two columns of $M(T)$ are independent over every field because their suffix differences in rows $1, 2$ are $d, 3d$. Thus the surviving monomials may be prescribed independently. When $n \geq 2(s-1)d$ the two supports cover all diagonal entries. Where both products survive, choose two nonzero entries with difference one, and where only one survives choose its entry to be 1 or -1 as required. The paired difference is I_n. The nonzero residual coefficient obtained by the original-input lift can be normalized by scaling one original input.

At a prime root let $w = 1 \cdots s$, let π be the paired order, and use the four orders

$$w, \quad \pi, \quad \operatorname{rev}(\pi), \quad \operatorname{rev}(w),$$

where rev reverses an order. The base orders $w, \operatorname{rev}(\pi)$ have the same orientation of the selected factor graph and opposite orientations of its complement. Their partners are $\pi, \operatorname{rev}(w)$, respectively. The paired suffix differences agree because $r_i(\operatorname{rev}(v)) = h - c_i - r_i(v)$ for total step h. By Lemma 2.1 the normalized permutation π is simple, meaning that no proper consecutive interval of positions of length at least two has consecutive values. The connected discrepancy graph implies that no proper prefix of w and π has the same set of letters. In the hyperplane $\sum_i c_i = 0$ there is an open cone on which all proper w-prefix sums are positive and all proper

π-prefix sums are negative. For example $c_i = \mathrm{pos}_\pi(i) - i$ belongs to it, where $\mathrm{pos}_\pi(i)$ is the position of letter i. For a w-prefix of length j, the sum of the positions in π is greater than $1 + \cdots + j$, since equality would give a common prefix. Interchanging the two orders proves the negative inequality for the π-prefix.

Suppose $s \geq 5$. Form the s-by-four matrix with (i, v)-entry $\prod_{j \text{ after } i \text{ in } v} x_j$ in the Laurent ring $\mathbb{Z}[x_1^{\pm 1}, \ldots, x_s^{\pm 1}]/(\prod_i x_i - 1)$. It has a four-row minor with relatively prime integer coefficients. By the Schmerl–Trotter theorem for permutations [3], every simple permutation of length at least five contains a simple pattern of length five or six. In fact a simple permutation that is not a parallel alternation can be shortened by one, and a simple parallel alternation of length at least eight can be shortened by two. Setting the variables outside such a pattern to one specializes the selected rows to the corresponding minor for the pattern. The length-five and length-six checks are finite. In the pseudocode below, a coefficient table has exponent tuples as keys and integer coefficients as values. It computes the checks by signed determinant expansion, using the last variable to enforce the product-one relation.

```
for s in (5, 6):
    good = total = 0
    for p in permutations(1, ..., s):
        if any proper interval of length >= 2 has consecutive values:
            continue
        total += 1
        words = ((1, ..., s), p, reverse(p), (s, ..., 1))
        for row in combinations((1, ..., s), 4):
            C = empty coefficient table, default 0
            for perm in permutations(0, 1, 2, 3):
                e = zero vector of length s - 1
                for k in (0, 1, 2, 3):
                    A = letters after row[k] in words[perm[k]]
                    for j in (1, ..., s-1):
                        e[j-1] += indicator(j in A) - indicator(s in A)
                C[tuple(e)] += sign(perm)
            if gcd of the entries of C is 1:
                good += 1
                break
    assert (total, good) == ((6, 6) if s == 5 else (46, 46))
```

Here sign is the parity sign of the determinant permutation and the gcd of an all-zero table is zero. Relatively prime coefficients at the specialized minor imply relatively prime coefficients before specialization. Choose rational steps in the opposite-prefix cone avoiding all $c_i = 0$ and equalities between the exponents of distinct terms of this minor under $x_i = T^{c_i}$. The excluded hyperplanes are proper on $\sum_i c_i = 0$ because the terms are distinct modulo the product-one relation. Clear denominators and scale so $|c_i|$ meets each degree. The resulting minor is nonzero over every field since its coefficients remain relatively prime.

Let the maximum w-prefix and minus the minimum π-prefix be A and B. The four support intervals in order are $[A, n), [0, n - B), [B, n), [0, n - A)$. For $n \geq A + B$, the rank-four monomial solve therefore permits simultaneous paired differences $\alpha I_n, \beta I_n$ for any $\alpha, \beta \in K^\times$. On overlaps one chooses nonzero entries with the prescribed difference, as above. The two residual coefficients after lifting can be made a nonzero pair, and then α, β can be chosen so their contraction is nonzero.

For $s = 4$ the two simple orders are 2413 and 3142. It suffices to treat 2413, since exchanging the graph with its complement exchanges the roles of the two available pairs. More explicitly the ordered pairs for 3142 are $(1234, 3142), (2413, 4321)$. Interchanging the labels by $1 \mapsto$

$2,\ 2 \mapsto 4,\ 3 \mapsto 1,\ 4 \mapsto 3$ takes the 2413 pairs to these pairs in the opposite order. Choose an integer H at least the four child degrees. For $n \geq 3H+4$ let $\varepsilon \in \{0,1\}$ have the parity of $n-H-2$ and take steps

$$(a,b,c,-a-1) = \left(\frac{n-H-2-\varepsilon}{2}, -H-\varepsilon, H+1+\varepsilon, -\frac{n-H-\varepsilon}{2}\right).$$

The four supports are $[a+1,n), [0,a+1), [a+c,n), [0,a+c)$. In the order of these four words, the suffix-exponent matrix is

$$\begin{pmatrix} -a & c & -a-c & 0 \\ -a-b & -b & 0 & a \\ -a-b-c & 0 & -c & a+b \\ 0 & a+c & b & a+b+c \end{pmatrix}.$$

Their suffix matrix has kernel generated over the Laurent ring by

$$(T^{2a+b+c}, -T^{a+b}, -T^{a+c}, 1)^{\mathsf{T}}.$$

Indeed substitution gives a kernel vector. The first three rows and columns have determinant

$$-T^{-2(a+b+c)}(T^{b+c}-1)(T^{a+b+c}-1),$$

which is nonzero over every field. The last coordinate is one. Subtracting the four kernel exponents from the support intervals gives lower endpoints $-a, -a-b, 0, 0$ and upper endpoints $c, c, a+1, a+c$, using $b+c=1$ and $n = 2a+c+1$. Their intersection is $[0,c)$. A supported row relation is thus the kernel vector multiplied by a polynomial supported on the integers from 0 to $c-1$. This describes the integer left kernel and gives the same rank in every characteristic. Assign constants $\alpha, -\alpha, \beta, -\beta$ to the surviving monomials of the four words respectively. Each translated relation holds multiplicatively, since $\alpha(-\beta) = (-\alpha)\beta$. The Smith argument realizes the assignment and gives the two desired scalar channels simultaneously.

It remains to consider a complete root with two children, denoted A, B. If neither child has an internal comparison flip, the paired differences are commutators of words on A and B and f is formally traceless. Thus one child, say A, has a nontrivial internal restriction. Since the children of a canonical complete root are the connected components of its complement graph, the canonical root of A is prime or empty. In the latter case retain one nonsingleton connected component C of the graph on A. The other connected components $D_1, \ldots, D_k$, with $k \geq 1$, occur in the orders allowed by the support of f, and the selected factor preserves their order relative to C. The retained component C has complete or prime root. We treat the three possibilities in turn.

7.3.1. *A prime child of a binary root.* Let $w, \pi, \mathrm{rev}(\pi), \mathrm{rev}(w)$ be the four orders on the canonical children of A. The full orders are $wB, B\pi, \mathrm{rev}(\pi)B, B\,\mathrm{rev}(w)$. The first and third orders are paired with the second and fourth respectively. We will also use the following minor in the case of a retained prime component.

For the internal words on a prime quotient, consider their four suffix columns in the variables x_i, now without imposing $\prod_i x_i = 1$. The quotient contains a copy of 2413 or 3142 by the simple-permutation theorem. For a simple parallel alternation, two entries from each of its alternating halves can be retained to obtain such a pattern. Select the four rows of this pattern and set other variables to one. For 2413 the specialized matrix is

$$\begin{pmatrix} x_2x_3x_4 & x_3 & x_2x_4 & 1 \\ x_3x_4 & x_1x_3x_4 & 1 & x_1 \\ x_4 & 1 & x_1x_2x_4 & x_1x_2 \\ 1 & x_1x_3 & x_2 & x_1x_2x_3 \end{pmatrix}.$$

Its determinant is

$$(h-1)\left(h^2 - h\sum_i x_i + \sum_{i<j<l} x_i x_j x_l - 1\right), \qquad h = x_1x_2x_3x_4,$$

where the sums run over the four pattern letters. For 3142 the two middle columns exchange roles. The term h^3 has coefficient one, so the unspecialized minor has relatively prime coefficients.

In the full orders of this binary case, the internal rows of the first and third columns acquire the factor x_B, since B comes last in those words. After eliminating $x_B = (\prod_i x_i)^{-1}$ the same minor is multiplied by a Laurent monomial and remains primitive. Choose positive integer steps on the children of A such that all monomials of this minor have distinct exponents under the one-variable substitution. This is possible by avoiding finitely many proper rational hyperplanes in the positive cone, then clearing denominators and scaling to meet the child degree bounds. Write L for their total and use step $-L$ on B, increasing the scale if needed so L also meets the degree of B. The four supports are $[L, n), [0, n-L), [L, n), [0, n-L)$. For $n \geq 2L$ the rank-four monomial solve gives simultaneous paired differences $\alpha I_n, \beta I_n$, just as at a prime global root. The two coefficients after the block lift are not identically zero, so a choice of parameters and of $\alpha, \beta \in K^\times$ gives I_n after scaling.

7.3.2. *A retained complete component.* Let U denote the effective letter on the first child of C. Use powers of one common nilpotent for all remaining children of C, so their product in either direction is the same effective matrix V. In this construction U, V denote matrices as well as their effective letters in the displayed orders. Their steps are $u, -\Delta$, respectively. The powers realizing V can have any positive exponents with sum Δ that meet the degrees of the corresponding children. Indeed a weighted shift of step $-\Delta$ is a conjugate of J_n^Δ, so one uses powers of the same conjugate of J_n. The other letters D_i have positive steps d_i at least their degrees and B has step $-P$, where

$$D = \sum_{i=1}^{k} d_i, \qquad p = u - \Delta > 0, \qquad \Delta \geq D, \qquad P = p + D, \qquad n = 2u + D.$$

For an order σ of $C, D_1, \ldots, D_k$, let S be the sum of the sibling steps before C, and write $R = D - S$. The paired full words have the orders

$$W_\sigma = (D_{\text{before}})UV(D_{\text{after}})B, \qquad P_\sigma = B(D_{\text{before}})VU(D_{\text{after}}),$$

where the products in parentheses are the ordered lists of siblings on either side of C. The first prefix range is $[0, u+S]$, because $u+S \geq p+D$. The second range is $[-(u+R), 0]$. Their diagonal supports are therefore $[u+S, n)$ and $[0, u+S)$.

Use unit edge weights on the siblings and on B. On U use $a(j) = (-1)^{\lfloor j/p \rfloor}$ for $0 \leq j < u+D$. On V use weights $v(j+\Delta) = a(j)^{-1}$ and $v(j+u) = -a(j)^{-1}$ for the same range of j. The assignments cover all starts of V and agree on the overlap by $a(j+p) = -a(j)$. A surviving first word at index z uses the U, V starts $j, j+\Delta$ with $j = z - u - S$. A surviving second word uses $j, j+u$ with $j = z + R$. Thus the surviving values are 1 and -1, respectively, for every σ.

To separate the residual coefficient directions, multiply each effective edge at start z by q_i^z for its letter i. The multiplier on a whole word of net step zero is $(\prod_i q_i)^z \prod_i q_i^{r_i(w)}$ at index z. Moving B from last to first increases every suffix by P, including the suffix of B, which changes from 0 to P. Swapping UV to VU then adds Δ to the suffix of U and u to the suffix

of V. With $g = \gcd(u, \Delta)$, choose

$$q_U = t^{u/g}, \qquad q_V = t^{-\Delta/g}, \qquad q_{D_i} = z_i, \qquad q_B = (q_U q_V \prod_i z_i)^{-1} \quad (t, z_i \in K^\times).$$

The paired multipliers agree and are independent of z. For a difference of suffix vectors of two first words, its B-coordinate is zero and its U, V-coordinates agree. Vanishing of the Laurent exponent differences in z_i, t forces all the D_i-coordinates to vanish and forces the common U, V difference to vanish, since $u - \Delta > 0$. The suffix positions of these positive-length components determine their order. Hence the paired scalar values have distinct Laurent coefficients for distinct σ. Choose the inner block parameters so that at least one residual coefficient is nonzero. These outer Laurent monomials then give a nonzero scalar for some t, z_i.

For each large n, take D equal to the sum of the sibling degrees or one more, choosing the parity of n. This is achieved by increasing one sibling step if necessary. Take Δ at least this sum plus one and at least the sum of the degrees of the collapsed children, independently of n, and put $u = (n - D)/2$. For all sufficiently large n, the steps satisfy the above conditions and the required lower bounds for U, B. The twisted V still has its prescribed Jordan type and hence can still be realized as the product of the collapsed children.

7.3.3. *A retained prime component.* On the children of C, choose steps c_i with positive total $p = \sum_i c_i$, positive proper w-prefix sums and negative proper π-prefix sums. Write A for the largest w-prefix sum and $-\delta$ for the smallest π-prefix sum. The four internal prefix ranges in order are

$$[0, A], \quad [-\delta, p], \quad [0, p + \delta], \quad [p - A, p].$$

The admissible steps form a nonempty open cone in the full step space by perturbing the earlier sum-zero cone to small positive total. Choose a rational point avoiding all $c_i = 0$ and the collision hyperplanes for the primitive minor without the product-one relation, proved in the prime-child case. The internal four-column suffix matrix then has full rank in every characteristic. Require also that the two vectors

$$v = r(\pi) - r(w), \qquad v' = r(\mathrm{rev}(\pi)) - r(w)$$

be linearly independent over $\mathbb{Q}$, where these suffix vectors are internal to C. They are the images of the step vector under two alternating comparison matrices with disjoint nonempty supports. Each matrix has rank at least two. The two images cannot be dependent for every step vector. Indeed, if linear maps F, G always have dependent images on each vector and $\operatorname{rank} F \geq 2$, take two vectors with independent F-images and apply dependence to their sum. This forces the same proportionality factor on both vectors, and repeating with a vector in the kernel of F forces its G-image to be zero. Thus G would be a scalar multiple of F. The independence condition therefore excludes only a proper algebraic set. Clear denominators to obtain a fixed integer choice of steps satisfying all these conditions.

As in the complete component case, give positive steps d_i to the siblings, put $D = \sum_i d_i$, $P = p + D$, and give B step $-P$. Require $A - p \geq D$ and $\delta \geq D$, and set $n = P + A + \delta$. With S the total of the steps before C, the first channel consists of the words $(D_{\mathrm{before}})w(D_{\mathrm{after}})B$ and $B(D_{\mathrm{before}})\pi(D_{\mathrm{after}})$. The first has prefix range $[0, A + S]$ and the second has range $[-(P + \delta - S), 0]$. The second channel uses $\mathrm{rev}(\pi)$ and $\mathrm{rev}(w)$ in place of w and π, with ranges $[0, p + \delta + S]$ and $[-(P + A - p - S), 0]$. These ranges follow by inserting the internal ranges between the positive sibling steps and the step $-P$. In each pair the supports are disjoint and cover the diagonal. The separating index is $A + S$ for the first channel and $p + \delta + S$ for the second.

By the rank-four monomial solve prescribe all surviving internal monomials of the four orders to be $1, -1, 1, -1$, respectively. This solve applies to the net step p since the incidence

uses the starting indices of the monomials. Give unit weights to the siblings and B. The paired differences are then I_n for both channels and every order of the siblings. Apply the same edge twist as above. Moving B from last to first adds P to every suffix, including its own. The internal change of order then adds v on C and zero elsewhere. Impose $\prod_i q_i = 1$ and $\prod_{i \text{ in } C} q_i^{v_i} = 1$. Explicitly parameterize q_i on C by the monomial exponent matrix $(v^{\mathsf{T}}v)I - vv^{\mathsf{T}}$, choose the siblings' parameters freely, and choose q_B to make the product one. The annihilator of this monomial map on rational exponent vectors is generated by $\mathbf{1}$ and the extension of v by zero. The difference of two base suffix vectors has B-coordinate zero. If all its parameter exponents vanish, the siblings have the same suffix positions. The sibling steps are positive, so sorting their suffix sums determines their relative order. After subtracting the steps of siblings to the right and the step of B, each remaining suffix contribution is p if C lies to its right and zero otherwise. Since $p > 0$, this determines the position of C. The only remaining difference could be between the two internal base states of C. That difference is v' up to sign, and it does not belong to $\mathbb{Q}v$. Thus the base scalar multipliers are distinct Laurent monomials. Their coefficients, after lifting the inner blocks, can be made a nonzero array, and evaluation of this Laurent polynomial followed by scaling gives I_n.

To see that every large n occurs, let $Q_0 = p + A + \delta$ for the fixed integer steps before scaling, and let D_0 be the sum of the sibling degrees. Take $k_0 = \lfloor (n - D_0)/Q_0 \rfloor$, multiply all internal steps by k_0, and enlarge one sibling so that $D = n - k_0 Q_0$. Then $D_0 \le D < D_0 + Q_0$. For n sufficiently large the scaled dominance margins $A - p, \delta$ exceed D. The absolute steps also meet all degree bounds. Scaling preserves both the monomial distinctions and the independence of the two suffix differences. The equality $n = P + A + \delta$ now holds with the scaled values.

Proof of Proposition 7.1. If every comparison factor has disconnected graph, apply Lemma 7.3. Otherwise the connected-factor constructions apply to a connected factor graph. At a complete root of arity two, tracefulness ensures the nontrivial action on a child used above. All the connected-factor constructions produce nilpotent original inputs, which are traceless. □

8. Uniformity of the scalar bound

Proposition 8.1. *For each $d \ge 2$ there is a bound depending only on d beyond which every formally traceful degree-d polynomial over every field takes the identity on traceless inputs.*

Proof. Fix d and consider the $d!$ coefficients as constants in the language of fields. Formal tracefulness is a finite disjunction of inequalities between cyclic coefficient sums and zero. For an integer $m \ge 2$, let P_m be the first-order formula asserting the existence of traceless m-by-m inputs with value I_m. This formula is a finite system of polynomial equations in their entries.

Propositions 6.2 and 7.1 imply that in every field with a formally traceful coefficient tuple, some adjacent pair P_m, P_{m+1} holds. If no finite list of adjacent pairs sufficed, the field axioms, the tracefulness condition, and all the sentences $\neg(P_m \wedge P_{m+1})$ would be finitely satisfiable. First-order compactness would give a field in which they all hold, a contradiction. Thus there is an integer M depending only on d such that some pair with $2 \le m \le M$ holds for every such field and tuple. Direct sums in the two coprime orders $m, m+1$ give all orders at least $M(M-1)$. □

Proof of Theorem 1.1. Take the maximum of the degree-dependent thresholds for nonscalar targets in Propositions 4.4 and 5.4 and the two scalar thresholds of Propositions 6.2 and 8.1. Every matrix in V_n is attained on traceless inputs. All values on unrestricted inputs lie in V_n by cyclicity of trace in the formally traceless case, and by definition of V_n otherwise. The two inclusions give the desired equalities. □

References

[1] A. A. Albert and B. Muckenhoupt, *On matrices of trace zero*, Michigan Math. J. **4** (1957), 1–3.
[2] S. A. Amitsur and J. Levitzki, *Minimal identities for algebras*, Proc. Amer. Math. Soc. **1** (1950), 449–463.
[3] R. Brignall and V. Vatter, *A simple proof of a theorem of Schmerl and Trotter for permutations*, J. Combin. **6** (2015), 47–54.
[4] A. Cafure and G. Matera, *Improved explicit estimates on the number of solutions of equations over a finite field*, Finite Fields Appl. **12** (2006), 155–185.
[5] M. Fiedler, *A note on companion matrices*, Linear Algebra Appl. **372** (2003), 325–331.
[6] V. T. Filippov, V. K. Kharchenko, and I. P. Shestakov (eds.), *Dniester Notebook: Unsolved Problems in the Theory of Rings and Modules*, 4th ed., Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 1993 (in Russian).
[7] K. Dykema and I. Klep, *Instances of the Kaplansky–Lvov multilinear conjecture for polynomials of degree three*, Linear Algebra Appl. **508** (2016), 272–288.
[8] P. S. Fagundes, T. C. de Mello, and P. H. da Silva dos Santos, *The Mesyan conjecture: a restatement and a correction*, arXiv:2111.13698 (2021).
[9] E. Formanek, *Central polynomials for matrix rings*, J. Algebra **23** (1972), 129–132.
[10] A. Kanel-Belov, S. Malev, and L. Rowen, *The images of non-commutative polynomials evaluated on* 2×2 *matrices*, Proc. Amer. Math. Soc. **140** (2012), 465–478.
[11] A. Kanel-Belov, S. Malev, and L. Rowen, *The images of multilinear polynomials evaluated on* 3×3 *matrices*, Proc. Amer. Math. Soc. **144** (2016), 7–19.
[12] I. Kaplansky, *Problems in the theory of rings*, in *Report of a Conference on Linear Algebras*, National Academy of Sciences–National Research Council, Washington, DC, 1957.
[13] G. Lachaud and R. Rolland, *On the number of points of algebraic sets over finite fields*, J. Pure Appl. Algebra **219** (2015), 5117–5136.
[14] M. Münch, I. Rutter, and P. Stumpf, *Partial and simultaneous transitive orientations via modular decompositions*, Algorithmica **86** (2024), 1263–1292.
[15] Y. T. Miah, *The L'vov–Kaplansky conjecture for standard polynomials*, Hexagon preprint 2611.00184v1 (2026).
[16] S. Pan, *The L'vov–Kaplansky conjecture for multilinear polynomials of degrees four to six*, Hexagon preprint 2611.00173v1 (2026).
[17] Yu. P. Razmyslov, *On a problem of Kaplansky*, Math. USSR-Izv. **7** (1973), 479–496.
[18] K. Shoda, *Einige Sätze über Matrizen*, Jpn. J. Math. **13** (1936), 361–365.
[19] The Stacks Project Authors, *The Stacks Project*, Lemma 33.25.10, Tag 0CDW.
[20] The Stacks Project Authors, *The Stacks Project*, Lemma 53.22.5, Tag 0H2V.
[21] Q. Tang, *The L'vov–Kaplansky conjecture for polynomials of degree four*, Hexagon preprint 2611.00119v1 (2026).
[22] Q. Tang, *The L'vov–Kaplansky conjecture in degrees five and six*, aiXiv preprint 2611.00692 (2026).
[23] D. Vitas, *Images of multilinear polynomials in the algebra of finitary matrices contain trace zero matrices*, Linear Algebra Appl. **626** (2021), 221–233.
[24] D. Vitas, *The L'vov–Kaplansky conjecture for polynomials of degree three*, Linear Algebra Appl. (2026), doi:10.1016/j.laa.2025.12.016.

Zhili College, Tsinghua University, Beijing 100084, China
Email address tangqy24@mails.tsinghua.edu.cn