Uniform adjacency–degree probes from differential rigidity

Zhipeng Lu

Shenzhen MSU–BIT University and
Guangdong Laboratory of Machine Perception and Intelligent Computing
Shenzhen 518172, Guangdong, China
zhipeng.lu@hotmail.com

Abstract

We construct a single trace invariant that separates all connected trees of order at most N. The invariant is the trace of an explicit self-adjoint polynomial $p_N \in \mathbb{Z}\langle x, y\rangle$, where x, y are noncommuting variables evaluated at the adjacency and degree matrices. Its arithmetic circuit has $O(N^2)$ gates, formal degree at most $3N - 2$, and polynomial bit complexity. The circuit is determined by N alone and operates on matrices of the original input order. A linear lower bound shows that the trace degree has the correct asymptotic scale.

The main ingredient is a fixed spanning family for the joint cyclic module of a real symmetric matrix pair from any prescribed vector. In dimension at most N, when the second matrix has spectral diameter at most $N - 1$, a triangular recurrence gives $N(N + 1)/2$ integer noncommutative polynomials of degree at most $N - 1$, computed by a shared $O(N^2)$ circuit. This replaces the exponentially large family of all short words without increasing the spanning degree. The proof uses a rational differential connection with a small self-adjoint residue: a primitive polynomial exterior vector representing an invariant subspace must have an integer degree smaller than one. For trees, pointed adjacency–degree reconstruction, McKay's distance polynomials, and a factored integer encoding turn this spanning family into the single trace invariant.

1 Introduction

How much spectral information is needed to determine a tree? The adjacency spectrum alone is insufficient: almost all trees have a nonisomorphic cospectral mate [Sch73]. Even the spectra of $A_T - tD_T$ for every real t can agree for nonisomorphic trees [GS25, FXZW26]. McKay showed that the spectra of all adjacency–degree polynomials determine a tree [McK77, Theorem 5.3]. We ask for a quantitative form of this determination: for each order bound N, can one prescribe a single adjacency–degree polynomial whose trace distinguishes all trees of order at most N, with low degree, a small arithmetic circuit, and polynomial bit complexity? Such an expression evaluates by the same sequence of matrix operations on every input and respects vertex relabelling through simultaneous conjugation. Controlling all three resources makes the passage from complete spectral data to a finite algebraic invariant explicit.

McKay's algebraic pruning procedure reconstructs the tree by successive trace tests [McK77, Algorithm 5.4]. These tests select rooted types and the multiplicities of adjacent leaves, and the procedure updates the residual adjacency and degree operators. Its choices therefore depend on the input trace data. The procedure by itself does not specify a single circuit fixed by N with the simultaneous resource bounds sought here. Our construction uses a fixed family of polynomials to determine a representation before applying tree reconstruction; the spanning argument applies to general symmetric matrix pairs and uses no tree pruning.

The pointed reconstruction theorem of Lu and Li continues this spectral approach: the moments $\mathbf{1}^{\mathsf{T}} w(A_T, D_T)\mathbf{1}$, over all words w in the adjacency matrix A_T and degree matrix D_T, already determine the tree [LL26, Theorem 5.4]. The present work makes this description uniformly finite with explicit arithmetic resources. The vectors $w(A_T, D_T)\mathbf{1}$ span its principal module, of dimension at most N. The responses of words of length at most $N-1$ already span this module, but there are $2^N - 1$ such words. Selecting a basis after evaluation reduces their number while making the selection depend on the tree [LL26, Proposition 6.1 and Remark 6.2]. We replace this exponential word family by a fixed quadratic-size polynomial spanning family, preserving the spanning degree and obtaining the following resource bounds.

Throughout, $\mathbb{Z}\langle x, y\rangle$ is the free associative algebra in the two *noncommuting* variables x, y. The involution $p \mapsto p^*$ fixes x, y and reverses each word. Our circuits have the identity and the two input matrices as inputs, and use addition, ordinary matrix multiplication, and multiplication by prescribed integer constants. The involution is compiled by reversing products. Circuit size counts stored instructions, including inputs.

Theorem 1.1 (A uniform linear-degree tree invariant)**.** *For every $N \geq 1$, a deterministic algorithm running in time polynomial in N constructs a self-adjoint polynomial $p_N \in \mathbb{Z}\langle x, y\rangle$, represented by an arithmetic circuit, such that*

$$T \cong T' \quad \Longleftrightarrow \quad \operatorname{tr} p_N(A_T, D_T) = \operatorname{tr} p_N(A_{T'}, D_{T'})$$

for all connected trees T, T' of orders at most N. The circuit has $O(N^2)$ gates and formal degree at most $3N-2$. Its scalar constants have $O(N \log(N+1))$ bits. For every such tree, the bit length of each stored matrix entry and of the final trace is $O(N^5 \log(N+1))$.

Every gate operates on matrices of the original input order using the operations specified above. The circuit and its constants are determined by N alone, and the only scalar readout is the trace of its final output. The simultaneous resource bounds matter. The same reconstruction and integer encoding applied to all $2^N - 1$ short words already gives a single trace polynomial of degree at most $3N-2$: two words contribute at most $2N-2$, an inserted generator at most one, and McKay's

expression for $\mathbf{1}\mathbf{1}^{\mathsf{T}}$ at most $N-1$. The direct encoding of the resulting Gram data, however, has exponentially many digit positions and exponential bit length. The smaller spanning family reduces the number of encoded entries to $O(N^4)$, giving polynomial bit length. Factoring their weights then gives the $O(N^2)$ gate bound.

Tree canonization itself admits logarithmic-space algorithms [Lin92], including a construction of a polynomial canon [ADFK25]. Other degree-based spectral refinements use multivariate spectra with auxiliary rank-one matrices built from degree-class indicators [WZ26]. Our question concerns the resources of a fixed expression in the adjacency and degree matrices alone. Nonisomorphic spiders give a trace-degree lower bound of $2\lfloor (N-3)/3 \rfloor$ for $N \geq 9$ (Section 4), so our invariant has the correct asymptotic degree scale.

The spanning construction applies to general real symmetric pairs. It provides a fixed finite description of the joint cyclic module from any prescribed vector, including when that module is a proper subspace. Together with Gram data, this gives explicit finite invariants for pointed self-adjoint representations (Corollary 3.3), so the algebraic construction extends beyond tree reconstruction. For $H, K \in M_n(\mathbb{R})$ and $v \in \mathbb{R}^n$, define

$$\mathcal{M}(H, K; v) = \operatorname{span}_{\mathbb{R}}\{w(H, K)v : w \in \{x, y\}^*\},$$

including the empty word. For a tree, write $\mathcal{M}_T = \mathcal{M}(A_T, D_T; \mathbf{1})$ for its principal module. Fix $N \geq 1$ and define $U_{i,j} \in \mathbb{Z}\langle x, y\rangle$ by

$$U_{0,0} = 1, \qquad U_{i+1,j} = (y + N^2(j-i))U_{i,j} + N^2 x U_{i,j-1}, \tag{1}$$

with $U_{i,j} = 0$ whenever $j \notin \{0, \ldots, i\}$.

Theorem 1.2 (Uniform spanning probes)**.** *Let $1 \leq n \leq N$, let $H, K \in M_n(\mathbb{R})$ be symmetric, and suppose*

$$\lambda_{\max}(K) - \lambda_{\min}(K) \leq N - 1.$$

Then, for every $v \in \mathbb{R}^n$,

$$\operatorname{span}_{\mathbb{R}}\{U_{i,j}(H, K)v : 0 \leq j \leq i < N\} = \mathcal{M}(H, K; v). \tag{2}$$

The family in (1) is constructed from N alone. It has $N(N+1)/2$ members, shared circuit size $O(N^2)$, and formal degree at most $N-1$. Its scalar constants have $O(\log(N+1))$ bits.

Corollary 1.3 (Rank-sensitive prefix)**.** *Under the hypotheses of Theorem 1.2, if $r = \dim \mathcal{M}(H, K; v)$, then*

$$\operatorname{span}_{\mathbb{R}}\{U_{i,j}(H, K)v : 0 \leq j \leq i < r\} = \mathcal{M}(H, K; v).$$

For $r = 0$ this is the empty span. The construction evaluates the full family indexed by $0 \leq j \leq i < N$ and does not need to determine r.

Either generator may play the bounded role; the ordered statement places it in the second slot. Interchanging the generators gives the corresponding spanning statement for $U_{i,j}(K, H)v$ when $\operatorname{diam}\operatorname{Spec}(H) \leq N - 1$. For trees, we use $H = A_T$, $K = D_T$. The family may be redundant, and the same formulas apply to arbitrary seeds and proper cyclic modules. Its degree $N-1$ is necessary in the worst case (Proposition 2.2). For example, the mixed probes include

$$\begin{aligned} U_{2,1} &= N^2(yx + xy), \\ U_{3,1} &= N^2(y^2x + yxy + xy^2 - N^2yx - 2N^2xy). \end{aligned}$$

The lower-degree coefficients depend on multiplication order. The circuit bound concerns the shared recurrence; its monomial expansion need not have polynomial size.

For the proof, set $B = K/N^2$ and introduce an auxiliary variable t. The differential Krylov sequence $q_0 = v$, $q_{i+1} = \nabla q_i$, for $\nabla = \partial_t + B/t + H$, spans an invariant rational subspace. Self-adjointness bounds the degree of a primitive polynomial exterior vector representing this subspace by $N \operatorname{diam}\operatorname{Spec}(B) < 1$. This degree is a nonnegative integer, so the exterior vector is constant. The subspace is therefore constant, and connection invariance makes its constant part invariant under both H and K. It contains v, and every iterate already lies in the rational extension of the joint cyclic module; hence the two constant spaces coincide. Taking coefficients of $N^{2i}t^i q_i$ now gives (1). Only the residue needs a spectral-diameter bound; symmetry of H supplies the inner-product cancellation. Section 2 compares this argument with degree-integrality arguments for small-exponent differential systems [GV15, BG19].

The same self-adjoint argument extends to several symmetric generators. For symmetric $H, A_1, \dots, A_s \in M_n(\mathbb{R})$ with $s \geq 1$, $1 \leq n \leq N$, and $\operatorname{diam}\operatorname{Spec}(A_i) \leq N - 1$, a fixed family of $N + sN(N-1)/2$ integer noncommutative polynomial probes spans the joint cyclic module from every seed, with the same degree bound $N - 1$ and rank-sensitive prefix. A shared circuit uses $O(s^2N^2)$ gates at the original matrix order. There is still no spectral bound on H. We outline the extension after the proof of Theorem 1.2; Appendix D gives its recurrence and proof.

Remark 1.4 (Fixed words versus polynomial probes). For $N \geq 5$, any fixed family of words of length at most $N - 1$ whose responses span $\mathcal{M}(H, K; v)$ for every input in Theorem 1.2 must contain at least 2^{N-1} words. To see this, colour the edges of an N-vertex path arbitrarily with x, y, let H, K be the adjacency matrices of the two colour classes, and take $v = e_1$ at one endpoint. The matrix K is a disjoint union of path adjacency matrices, so $\operatorname{diam}\operatorname{Spec}(K) < 4 \leq N - 1$. Among all words of length at most $N - 1$, exactly one has a nonzero e_N coordinate when applied to e_1: the word following the edge colours from one endpoint to the other, read in the order of matrix action. Its e_N coordinate is one. Thus every spanning word family must contain that word. Varying the colouring forces all 2^{N-1} words of length $N - 1$. In the general symmetric-pair setting, Theorem 1.2 therefore gives an exponential reduction from fixed words to polynomial probes at the same optimal degree.

To obtain the tree invariant, enumerate the family as $U_0, \dots, U_{M-1}$ and put $z_i = U_i(H, K)v$. The three matrices

$$G_{ij} = z_i^{\mathsf{T}} z_j, \qquad G^H_{ij} = z_i^{\mathsf{T}} H z_j, \qquad G^K_{ij} = z_i^{\mathsf{T}} K z_j$$

determine the pointed representation: G determines linear relations and inner products, and G^H, G^K determine the actions of H, K. Hence they determine every pointed moment. For trees, use $(H, K, v) = (A_T, D_T, \mathbf{1})$ and apply pointed reconstruction. McKay's distance polynomials express $\mathbf{1}\mathbf{1}^{\mathsf{T}}$ as a polynomial of degree at most $N - 1$ [McK77, Lemma 5.1 and Corollary 5.2], converting these entries into traces. Bounded integer entries can then be packed into one trace. Factoring this encoding permits $O(N^2)$ gates despite the $O(N^4)$ entries.

Related work. Earlier work on hyperbolic set comparison considered the problem of finding finite, explicit adjacency–degree tests motivated by McKay's theorem [Li+25, Sections D.2 and E.1].

Pointed moments also form a series represented by a weighted automaton. Tzeng's equivalence algorithm and later refinements construct input-dependent reachable bases; polynomial identity testing also gives randomized equivalence tests [Tze92, KMOWW11]. Our self-adjoint setting admits fixed exact polynomial probes. Fixed short word families span the full matrix algebra on generic tuples [KŠ16, JC24], whereas our polynomial family works throughout the stated self-adjoint class. Matrix-algebra length bounds control the span of all short words [Paz84, Pap97, Shi19, Spr26]; the full algebra may have dimension n^2, while the cyclic module used here has dimension at most n.

Trace-generation theorems [Pro76] and trace identities [Raz74] provide the classical invariant-theoretic setting. Trace criteria also classify systems of maps between real inner-product spaces [FHS17, Corollary 1], and finite Gram criteria for unitary similarity were developed by Al'pin and Ikramov [AI03, Proposition 10 and Theorem 5]. Derksen and Makam construct input-dependent pivot-word bases and separating words [DM20, Algorithm 4.8 and Section 5]; Forbes and Shpilka obtain explicit quasipolynomial-size separating families through polynomial identity testing (full version, [FS13, Theorem 1.13]). Our contribution is the fixed low-degree spanning recurrence for self-adjoint pairs and its use to achieve the simultaneous circuit and bit bounds.

Outline. The proof of Theorem 1.1 proceeds from a fixed spanning family to complete finite data and then to a single trace circuit. Section 2 establishes the spanning family for general symmetric pairs: differential rigidity identifies the span of the iterates with the joint cyclic module, and coefficient extraction gives the fixed probes. Section 3 uses their Gram data to recover the pointed representation and hence every principal moment. Principal-moment reconstruction then makes these finite data a complete tree invariant.

The remaining task is to realize this invariant within the stated circuit model. McKay's distance polynomials turn the Gram entries into traces, and uniform integer bounds allow all entries to be encoded injectively in one trace. Factoring the encoding gives the quadratic gate bound; the degrees of its factors and the number of encoded entries give the degree and bit bounds. The circuit evaluates this factored expression directly. Section 4 then shows that linear trace degree is necessary up to a constant factor.

The appendices branch from the spanning theorem. Appendix A uses the probes to construct a single positive semidefinite operator with the prescribed principal module and a cyclic all-ones vector. Appendix B recovers the full adjacency–degree representation in the Frobenius matrix space, allowing reconstruction directly from McKay's theorem at higher resource cost. Appendix C records the computational checks and the scope of formal verification. Appendix D extends the same rigidity argument to several symmetric generators while preserving the degree bound and rank-sensitive prefix.

2 Uniform spanning via differential rigidity

To prove Theorem 1.2, we must show that the fixed recurrence (1) spans the entire joint cyclic module on every admissible input. We first identify the rational span of a differential Krylov sequence from v with $\mathcal{M}(H, K; v) \otimes_{\mathbb{R}} \mathbb{R}(t)$. The rigidity lemma below forces this rational span to be constant and jointly invariant under the two matrices. Taking coefficients of the iterates then gives the prescribed probes. Throughout, t is a commuting auxiliary variable; the polynomials $U_{i,j}$ remain elements of $\mathbb{Z}\langle x, y\rangle$.

For a linear map B on $\mathbb{R}^n$, its additive action on the rth exterior power is

$$B^{[r]}(u_1 \wedge \cdots \wedge u_r) = \sum_{j=1}^{r} u_1 \wedge \cdots \wedge Bu_j \wedge \cdots \wedge u_r.$$

If B is symmetric, then $B^{[r]}$ is symmetric for the induced inner product, and its eigenvalues are sums of r eigenvalues of B with distinct indices. Consequently,

$$\operatorname{diam} \operatorname{Spec}(B^{[r]}) \le r \operatorname{diam} \operatorname{Spec}(B). \tag{3}$$

Lemma 2.1 (Small-residue rigidity)**.** *Let $B, H \in M_n(\mathbb{R})$ be symmetric, and let $V \subseteq \mathbb{R}(t)^n$ be an r-dimensional subspace invariant under the connection*

$$\nabla f = f' + (B/t + H)f.$$

If $r \operatorname{diam}\operatorname{Spec}(B) < 1$, then $V = W \otimes_{\mathbb{R}} \mathbb{R}(t)$ for a real subspace $W \subseteq \mathbb{R}^n$ invariant under both B and H.

Proof. The cases $r = 0$ and $r = n$ are immediate, so assume $0 < r < n$. Take the exterior product of a rational basis of V. Clear its denominators and divide its coordinates by their polynomial gcd to obtain

$$w(t) = w_0 + w_1 t + \cdots + w_\ell t^\ell \in \left(\textstyle\bigwedge^r \mathbb{R}^n\right)[t].$$

This vector is nonzero and its coordinates are coprime. In particular, $w_0 \neq 0$, since otherwise all coordinates would be divisible by t; also $w_\ell \neq 0$ by the choice of ℓ.

Invariance of V makes the rational line spanned by w invariant under the induced connection. Hence there is $c(t) \in \mathbb{R}(t)$ such that

$$tw' + (B^{[r]} + tH^{[r]})w = c(t)w.$$

The left side is polynomial. A polynomial Bezout identity for the coprime coordinates of w therefore shows that $c(t)$ is polynomial. Comparing degrees gives $\deg c \leq 1$, so write $c(t) = \alpha + \beta t$. The constant coefficient and the coefficient of $t^{\ell+1}$ give

$$B^{[r]}w_0 = \alpha w_0, \qquad H^{[r]}w_\ell = \beta w_\ell. \tag{4}$$

If $\ell > 0$, the coefficient of t^ℓ gives

$$\ell w_\ell + B^{[r]}w_\ell + H^{[r]}w_{\ell-1} = \alpha w_\ell + \beta w_{\ell-1}.$$

Take the inner product with w_ℓ. Symmetry of $H^{[r]}$ and (4) cancel the terms containing $w_{\ell-1}$. Since $\alpha \in \operatorname{Spec}(B^{[r]})$, the Rayleigh quotient of $B^{[r]}$ then yields

$$\ell = \alpha - \frac{\langle w_\ell, B^{[r]}w_\ell\rangle}{\|w_\ell\|^2} \leq \operatorname{diam}\operatorname{Spec}(B^{[r]}) \leq r \operatorname{diam}\operatorname{Spec}(B) < 1. \tag{5}$$

This is impossible for the positive integer ℓ. Thus $\ell = 0$.

The constant exterior vector w_0 determines a constant subspace: explicitly, V consists of the rational vectors u satisfying $u \wedge w_0 = 0$. This equation has constant real coefficients, so its solution space is $W \otimes_{\mathbb{R}} \mathbb{R}(t)$ for a real subspace $W \subseteq \mathbb{R}^n$. For $u \in W$, invariance gives $(B/t+H)u \in W \otimes_{\mathbb{R}} \mathbb{R}(t)$. Multiply by t and pass to $\mathbb{R}^n/W$. The constant and linear coefficients of $Bu+tHu$ vanish separately, proving $Bu, Hu \in W$. □

The strict bound in the lemma is necessary: for $B = \operatorname{diag}(0, -1)$ and $H = 0$, the nonconstant rational line spanned by $(1, t)^{\mathsf{T}}$ is horizontal for $\partial_t + B/t$. Its primitive polynomial degree and the spectral diameter of B both equal one.

We apply the lemma to the rational span of the differential iterates from v. The goal is to identify this span with $\mathcal{M}(H, K; v) \otimes_{\mathbb{R}} \mathbb{R}(t)$; the number of iterates required will then equal $\dim_{\mathbb{R}} \mathcal{M}(H, K; v)$.

Proof of Theorem 1.2 and Corollary 1.3. If $v = 0$, both sides of the spanning identities are the zero subspace. Assume $v \neq 0$, put $B = K/N^2$, and define

$$q_0 = v, \qquad q_{i+1} = \nabla q_i, \qquad \nabla = \partial_t + B/t + H.$$

Let $s \geq 1$ be the first index for which q_s belongs to the $\mathbb{R}(t)$-span of $q_0, \dots, q_{s-1}$. Such an index exists with $s \leq n$. Write $q_s = \sum_{i=0}^{s-1} a_i(t) q_i$. The Leibniz rule gives

$$q_{s+1} = \sum_{i=0}^{s-1} a_i'(t) q_i + \sum_{i=0}^{s-1} a_i(t) q_{i+1}.$$

Every term on the right lies in the same span. More generally, that span is invariant under ∇, because $\nabla q_i = q_{i+1}$ and differentiating a rational coefficient preserves $\mathbb{R}(t)$. Thus all later iterates lie in it, and

$$V := \operatorname{span}_{\mathbb{R}(t)}\{q_i : i \geq 0\} = \operatorname{span}_{\mathbb{R}(t)}\{q_0, \dots, q_{s-1}\}, \qquad \dim_{\mathbb{R}(t)} V = s.$$

Let $\mathcal{M} = \mathcal{M}(H, K; v)$. It is invariant under H and K, so the recurrence for q_i gives $V \subseteq \mathcal{M} \otimes_{\mathbb{R}} \mathbb{R}(t)$. Also,

$$s \operatorname{diam} \operatorname{Spec}(B) \leq \frac{N(N-1)}{N^2} < 1.$$

Lemma 2.1 therefore gives $V = W \otimes_{\mathbb{R}} \mathbb{R}(t)$ with W invariant under H, K and containing v. The definition of $\mathcal{M}$ implies $\mathcal{M} \subseteq W$, and hence

$$V = \mathcal{M} \otimes_{\mathbb{R}} \mathbb{R}(t). \tag{6}$$

In particular, $s = r := \dim_{\mathbb{R}} \mathcal{M}$. This is where the spectral-diameter bound on K enters: K/N^2 is the residue in the connection, while symmetry of H supplies the cancellation in (5).

To extract fixed polynomial expressions from these iterates, put $f_i = t^i q_i$. Their recurrence is

$$f_0 = v, \qquad f_{i+1} = t f_i' + (B - iI + tH) f_i.$$

Induction shows that f_i is a polynomial vector of degree at most i. Write

$$f_i(t) = N^{-2i} \sum_{j=0}^{i} c_{i,j} t^j.$$

Comparison of coefficients gives

$$c_{0,0} = v, \qquad c_{i+1,j} = \big(K + N^2(j-i)I\big) c_{i,j} + N^2 H c_{i,j-1},$$

with coefficients outside the indicated range taken to be zero. This is exactly (1), so $c_{i,j} = U_{i,j}(H, K)v$.

Let S be the real span of the coefficients of $f_0, \dots, f_{r-1}$. The coefficient recurrence gives $S \subseteq \mathcal{M}$. Conversely, $q_i = t^{-i} f_i \in S \otimes_{\mathbb{R}} \mathbb{R}(t)$ for $i < r$, and these q_i span V. Equation (6) therefore gives $\mathcal{M} \otimes_{\mathbb{R}} \mathbb{R}(t) \subseteq S \otimes_{\mathbb{R}} \mathbb{R}(t)$. Taking dimensions yields $S = \mathcal{M}$. Consequently,

$$\operatorname{span}_{\mathbb{R}}\{U_{i,j}(H, K)v : 0 \leq j \leq i < r\} = \mathcal{M}.$$

Since $r \leq N$, this proves both spanning assertions. The circuit evaluates the full family with $0 \leq j \leq i < N$, independently of r.

It remains to bound the resources. There are $\sum_{i=0}^{N-1}(i+1) = N(N+1)/2$ polynomials, and each new entry of (1) uses a constant number of gates from the preceding entries. The shared circuit therefore has $O(N^2)$ gates. Induction in (1) gives $\deg U_{i,j} \leq i$, so the formal degree is at most $N-1$. The prescribed integer scalars have absolute value at most N^3, and thus $O(\log(N+1))$ bits. All choices in this recurrence depend only on N. $\square$

Several symmetric generators. The mechanism extends by placing the additional generators at distinct real poles. For $\nabla = \partial_t + H + \sum_{i=1}^{s} B_i/(t-a_i)$, the primitive exterior-vector argument bounds its degree ℓ by $\ell \le r \sum_i \operatorname{diam} \operatorname{Spec}(B_i)$. Each pole supplies an eigenvalue of $B_i^{[r]}$, and the highest two coefficients supply a sum of Rayleigh quotients; the self-adjointness of H gives the same cancellation as above. Thus $r \sum_i \operatorname{diam} \operatorname{Spec}(B_i) < 1$ again forces constancy. Taking $a_i = i-1$ and $B_i = A_i/(sN^2)$ meets this condition uniformly. Clearing denominators in the differential iterates then yields a fixed integer recurrence, whose first r rows already span the joint cyclic module. The polynomial coefficients in row k still have noncommutative degree at most k. Appendix D provides the details, including the shared circuit bound. This extension uses the same self-adjoint rigidity argument without a block enlargement.

The differential-module framework is standard [vPS03, Chapter 2]. Classical cyclic-vector constructions, such as Katz's algorithm [Kat87], select a vector whose differential iterates form a basis. Here the seed is prescribed: Lemma 2.1 identifies its differential cyclic module with the joint invariant module of the two matrices. The use of degree integrality also has precedents in constant-reduction arguments. Gontsov and Vyugin bound a nonpositive subbundle degree strictly above -1 and hence force it to vanish [GV15, Lemma 2]. Barkatou and Gontsov give a Liouvillian-solvability criterion using differences of the real parts of formal exponents [BG19, Theorem 2.3]. Their criterion also requires distinct leading eigenvalues at irregular singularities and exponent differences outside $\mathbb{Q} \setminus \mathbb{Z}$ at Fuchsian singularities. The latter condition already fails for our horizontal tree system $f' = -(D_T/(N^2 t) + A_T)f$ whenever T has distinct vertex degrees: the residue then has nonzero rational eigenvalue differences of absolute value less than one. Lemma 2.1 instead gives the required conclusion directly for every rational invariant subspace of the stated self-adjoint connection, including when H has repeated eigenvalues.

The degree bound on the spanning family cannot be improved in general. This concerns the degree needed to generate the cyclic module; the trace degree needed to distinguish trees is considered in Section 4.

Proposition 2.2 (The probe degree is necessary)**.** *For every $N \ge 1$, there is a symmetric pair satisfying Theorem 1.2 for which every polynomial family spanning the cyclic module has a member of degree at least $N-1$.*

Proof. Take $K = 0$, $H = \operatorname{diag}(\lambda_1, \dots, \lambda_N)$ with distinct real λ_i, and $v = \mathbf{1}$. The Vandermonde matrix shows that $v, Hv, \dots, H^{N-1}v$ are independent. Every polynomial of total degree at most d evaluated at $(H, 0)$ sends v into $\operatorname{span}\{v, Hv, \dots, H^d v\}$, whose dimension is at most $d+1$. Spanning this N-dimensional module therefore requires $d \ge N-1$. □

3 A single trace invariant for trees

The fixed spanning family lets us replace all principal moments by three finite Gram matrices recording the inner product and both generator actions on the principal module. We first identify these complete data, then give their trace encoding and its circuit factorization. The proof of Theorem 1.1 verifies the encoding and its resource bounds together. The graph-theoretic input is the following reconstruction theorem.

Theorem 3.1 (Principal-moment reconstruction; [LL26, Theorem 5.4])**.** *If two trees T, T' satisfy*

$$\mathbf{1}^{\mathsf{T}} w(A_T, D_T)\mathbf{1} = \mathbf{1}^{\mathsf{T}} w(A_{T'}, D_{T'})\mathbf{1} \quad \text{for every word } w \in \{x, y\}^*,$$

then $T \cong T'$.

To recover every pointed moment, it suffices to recover the inner product, the seed, and the action of both generators on its cyclic module. The correspondence is standard; for the all-ones seed, see [GRS25, Lemma 3.4]. Finite word-Gram reconstruction also underlies the unitary-similarity criteria of Al'pin and Ikramov [AI03, Proposition 10 and Theorem 5]. The following pointed real-symmetric form shows that a prescribed spanning family and its three Gram matrices recover precisely this information.

Lemma 3.2 (Recovering a pointed representation)**.** *Let H, K be real symmetric matrices, let v be a vector, and suppose the vectors $z_i = U_i(H, K)v$, $0 \le i < M$, span $\mathcal{M}(H, K; v)$, where $U_0 = 1$. The three matrices G, G^H, G^K with entries*

$$G_{ij} = z_i^{\mathsf{T}} z_j, \qquad G^H_{ij} = z_i^{\mathsf{T}} H z_j, \qquad G^K_{ij} = z_i^{\mathsf{T}} K z_j \tag{7}$$

determine this pointed representation up to a seed-preserving isometry.

Proof. Suppose two pointed pairs give the same three matrices. Define $\sum_i c_i z_i \mapsto \sum_i c_i z'_i$. Equality of G makes this map well defined and isometric, and the spanning hypothesis makes it onto. Equality of G^H implies that the map intertwines the restrictions of H and H': pair the desired identity with each spanning vector z'_i. Equality of G^K gives the same conclusion for K, K'. Finally, $z_0 = v$ and $z'_0 = v'$, so the seed is preserved. □

Corollary 3.3 (Finite invariants for pointed representations)**.** *For real symmetric pairs of dimension at most N satisfying the hypothesis of Theorem 1.2, the matrices in (7) formed from its fixed polynomials are complete invariants of the pointed cyclic representation. Their entries are polynomial functions of (H, K, v), of degree at most $2N - 1$ in the matrix entries and degree two in the seed. Equality of the three matrices is equivalent to a seed-preserving isometry intertwining both restricted generators.*

Proof. One direction is Lemma 3.2. For the other, an intertwining isometry preserves every vector $U_i(H, K)v$ and hence all three matrices. The degree bound follows from $\deg U_i \le N - 1$. □

From Gram data to a trace expression. For trees, the resulting isometry preserves every principal moment, so Theorem 3.1 gives a graph isomorphism. To express each of these integer matrix entries as a trace, we use McKay's distance polynomials [McK77, Lemma 5.1 and Corollary 5.2]:

$$\begin{aligned} &P_0 = 1, \qquad P_1 = x, \qquad P_2 = x^2 - y, \\ &P_k = xP_{k-1} - (y-1)P_{k-2} \quad (k \ge 3), \qquad J_N = \sum_{k=0}^{N-1} P_k. \end{aligned} \tag{8}$$

All these expressions belong to $\mathbb{Z}\langle x, y\rangle$. On a tree T, let $F_k(u, v)$ be the indicator of $\operatorname{dist}(u, v) = k$. Then $F_0 = I$ and $F_1 = A_T$. The diagonal of A_T^2 is D_T, and two distinct vertices have one common neighbour exactly when their distance is two; hence $F_2 = A_T^2 - D_T$. For $k \ge 3$, count neighbours w of u with $\operatorname{dist}(w, v) = k - 1$. The unique path from u to v gives one such neighbour when $\operatorname{dist}(u, v) = k$, and $\deg(u) - 1$ when $\operatorname{dist}(u, v) = k - 2$; otherwise the count is zero. Thus $A_T F_{k-1} = F_k + (D_T - I)F_{k-2}$, proving $P_k(A_T, D_T) = F_k$ by induction. Every distance is at most $n - 1 \le N - 1$, so

$$J_N(A_T, D_T) = \mathbf{1}\mathbf{1}^{\mathsf{T}}. \tag{9}$$

In particular, $\operatorname{tr}(J_N(A_T, D_T)X) = \mathbf{1}^{\mathsf{T}} X \mathbf{1}$ for every matrix X. The expression J_N has degree at most $N - 1$; we keep its formal adjoint J_N^* in the construction below.

It remains to encode the entries of G, G^A, G^D in a single trace and to compute that trace with $O(N^2)$ gates. Enumerate the $M = N(N+1)/2$ polynomials in (1) in lexicographic order of the index pair, as $U_0, \dots, U_{M-1}$, with $U_0 = 1$. Set

$$X_\nu = U_i^* S U_j, \qquad S \in \{1, x, y\}, \quad 0 \le i, j < M,$$

where $\nu = sM^2 + iM + j$ and $s = 0, 1, 2$ corresponds to $S = 1, x, y$, respectively. There are $Q = 3M^2$ entries to encode. Define

$$L = 3N^3, \qquad R = L^{N-1}, \qquad H_N = N^2 R^2, \qquad b_N = 2H_N + 1. \tag{10}$$

For $N = 1$ take $p_1 = 1$; there is only one nonempty tree of order at most one. For $N \ge 2$, define

$$p_N = \sum_{\nu=0}^{Q-1} b_N^\nu (J_N X_\nu + X_\nu^* J_N^*). \tag{11}$$

This expression is formally self-adjoint. The bounds in (10) will ensure that its trace uniquely encodes all Q integer entries.

The sum in (11) has $O(N^4)$ terms. The exponent $sM^2 + iM + j$ separates the generator index from the two probe indices, allowing the circuit to share two probe combinations and one generator combination. Put

$$Y_N = \sum_{j=0}^{M-1} b_N^j U_j, \qquad Z_N = \sum_{i=0}^{M-1} b_N^{iM} U_i,$$
$$W_N = 1 + b_N^{M^2} x + b_N^{2M^2} y.$$

Expanding the product and using $W_N^* = W_N$ gives

$$p_N = J_N Z_N^* W_N Y_N + Y_N^* W_N Z_N J_N^*. \tag{12}$$

The coefficient of $J_N U_i^* S U_j$ is exactly $b_N^{sM^2+iM+j}$, as in (11). The circuit evaluates (12) directly from the original-order probe matrices and J_N; the Gram entries are recovered only from the final trace in the correctness argument.

The only large prescribed scalar is b_N. Starting from $B = b_N I$, at most M multiplications construct $B^M = b_N^M I$, and at most another M construct $(B^M)^M = b_N^{M^2} I$. Squaring the latter gives $b_N^{2M^2} I$. Horner's rule with bases B and B^M constructs Y_N and Z_N in $O(M)$ further gates. Thus these large scalar powers are computed by the circuit.

Correctness and resource bounds. The construction is now fixed. Its correctness rests on the completeness of the Gram data and the injectivity of their bounded-integer encoding; the factorization then supplies the circuit bounds.

Proof of Theorem 1.1. The case $N = 1$ follows from the identity output. Assume $N \ge 2$. For every permitted tree, A_T, D_T are symmetric and $\mathrm{Spec}(D_T) \subseteq [0, N-1]$. Theorem 1.2 therefore applies. Lemma 3.2 and Theorem 3.1 show that G, G^A, G^D together form a complete tree invariant.

Their entries are integers. Since the induced row-sum norms of A_T, D_T are at most $N - 1$, induction in (1) gives

$$\|U_{i,j}(A_T, D_T)\mathbf{1}\|_\infty \le L^i.$$

The same bound holds for the induced matrix norms of $U_{i,j}$. Consequently, for $S = I, A_T, D_T$,

$$|\mathbf{1}^\mathsf{T} U_i^* S U_j \mathbf{1}| \le N^2 R^2 = H_N.$$

Equation (9) now gives

$$\operatorname{tr} p_N(A_T, D_T) = 2 \sum_{\nu=0}^{Q-1} b_N^{\nu} h_\nu(T), \qquad h_\nu(T) = \mathbf{1}^{\mathsf{T}} X_\nu(A_T, D_T)\mathbf{1}.$$

This integer encoding is injective for entries in $[-H_N, H_N]$. Indeed, at the least index where two sequences differ, the nonzero difference has absolute value at most $2H_N < b_N$ and cannot be divisible by b_N. Equality of final traces is therefore equivalent to equality of all entries of G, G^A, G^D, and hence to tree isomorphism. In particular, $h_0(T) = n$ records the order.

For the gate count, the shared probe circuit uses $O(N^2)$ gates, the distance recurrence constructs J_N in $O(N)$ gates, and the factored expression (12) uses $O(M) = O(N^2)$ further gates. Reversing multiplication gates computes formal adjoints at a constant-factor increase in size. Every operation is prescribed by N.

For the degree bound, scalar-power registers have degree zero, Y_N, Z_N have degree at most $N-1$, and W_N has degree at most one. Together with $\deg J_N \le N-1$, this gives degree at most $3N-2$. All scalar coefficients are integers, and (12) is self-adjoint.

For the bit bounds, $\log_2 b_N = O(N \log(N+1))$. The reversed-product recurrence gives the same matrix norm bound R for U_i^*. Scalar-power registers have exponent at most $2M^2$, and the partial Horner sums for Y_N, Z_N have exponents at most $M-1$ and $M(M-1)$, respectively. On trees, $\|J_N\|_\infty \le N$, and intermediate products in the distance recurrence have polynomially bounded entries. Every register is consequently bounded by polynomial factors in N times $R^2 b_N^{3M^2}$. Since $M = O(N^2)$, all stored integer entries and the final trace have $O(N^5 \log(N+1))$ bits. These bounds also give polynomial time for constructing and evaluating the circuit in the bit model. □

4 A linear trace-degree lower bound

We now prove that the linear degree in Theorem 1.1 is necessary up to a constant factor. The obstruction comes from two spiders whose short degree-weighted closed walks have identical counts. This is a bound on polynomial degree; gate counts and bit complexity are separate resources.

Proposition 4.1. *For $m \ge 2$, let T_m and T'_m be the spiders with arm lengths*

$$(m, m, m+2) \quad \text{and} \quad (m, m+1, m+1),$$

respectively. Both have $3m+3$ vertices and maximum degree three. Every word $w \in \{x, y\}^*$ containing fewer than $2m$ occurrences of x satisfies $\operatorname{tr} w(A_{T_m}, D_{T_m}) = \operatorname{tr} w(A_{T'_m}, D_{T'_m})$. Consequently, any polynomial in $\mathbb{R}\langle x, y\rangle$ whose trace separates all trees of order at most N has degree at least $2\lfloor (N-3)/3 \rfloor$ for $N \ge 9$.

Proof. A trace word with s adjacency factors is a sum over closed walks of length s, with products of vertex degrees attached at prescribed positions. On a tree, such a walk traverses every edge of its image at least twice. Its image therefore has at most $s/2 < m$ edges.

Call the center and the three leaves special. The image cannot contain two special vertices. If it contains the center, its degree-labelled embeddings are the same in both spiders, since every arm has length at least m. If it contains a leaf but not the center, its image is a short path and the three leaf neighbourhoods agree. If it contains no special vertex, it is a path of length $\ell < m$ inside the degree-two portion of an arm. An arm of length L_a has $L_a - 1 - \ell$ such undirected placements, so their total is $\sum_a L_a - 3(\ell+1)$, equal for the two spiders. Orientations, choices of starting position, and repeated-walk patterns contribute the same multiplicities to each placement.

Hence every degree-labelled walk count, and therefore every trace word in question, agrees. The case $s = 0$ follows directly from the common degree multiset.

A polynomial of total degree less than $2m$ is a linear combination of such words. The unique degree-three center has two leaves at distance m in T_m and one in T'_m, so the spiders are nonisomorphic. Taking $m = \lfloor (N-3)/3 \rfloor$ proves the stated bound. □

Together with Theorem 1.1, this gives degree $\Theta(N)$ for a single-trace separator. The constants in the upper and lower bounds need not agree.

A A fixed cyclic representative

We give a further construction from the fixed spanning family: a polynomial $C_N \in \mathbb{Z}\langle x, y\rangle$ that is a sum of Hermitian squares and whose value on each tree has image $\mathcal{M}_T$ and is cyclic from $\mathbf{1}$ on that module. This construction is independent of the trace encoding in Theorem 1.1. In vector formulas in this appendix, polynomial expressions are evaluated at (A_T, D_T).

A.1 Fixed rank condensation

The probes span $\mathcal{M}_T$, but a fixed initial segment need not be a basis. We need fixed linear combinations whose first r responses are independent whenever $r = \dim \mathcal{M}_T$. This prefix property will make the orthogonal residuals vanish exactly after the rank, without choosing pivots from the input.

Assume $N \geq 2$. Let $U_0 = 1, U_1, \ldots, U_{M-1}$ be the probes. With L, R as in (10), set

$$H_0 = (M-1)R\,2^{(N-2)(M-2)}, \qquad H_1 = N!H_0^{N-1}, \qquad \tau = H_1 + 2.$$

Define $Z_0 = 1$ and

$$Z_{j+1} = \sum_{i=1}^{M-1} (2^j \tau)^{i-1} U_i, \qquad 0 \leq j \leq N-2. \tag{13}$$

These sums again use Horner's rule. Their degree is at most $N-1$, their shared size is polynomial, and their coefficients and evaluated entries have polynomial bit length. More explicitly, $\log \tau = O(N^4)$ and the evaluated vector entries have $O(N^6)$ bits.

Lemma A.1 (A basis in the first positions)**.** *On every tree of order at most N, if $r = \dim \mathcal{M}_T$, then*

$$z_0 = \mathbf{1},\ z_1 = Z_1\mathbf{1}, \ldots, z_{r-1} = Z_{r-1}\mathbf{1}$$

are independent. All N vectors z_i belong to $\mathcal{M}_T$.

Proof. If $r = 1$, the assertion follows from $z_0 = \mathbf{1} \neq 0$. Assume $r \geq 2$. Work first with a commuting variable t in place of τ, and pass to the quotient of $\mathcal{M}_T$ by $\mathbb{R}\mathbf{1}$. The coefficient vectors of

$$F(t) = \sum_{i=1}^{M-1} U_i\mathbf{1}\, t^{i-1}$$

span this quotient. Choose linear coordinates on the quotient, and by constant row operations put the resulting $r-1$ independent coordinate polynomials in a basis of distinct degrees $d_1, \ldots, d_{r-1}$. The leading coefficient of the determinant of their evaluations at

$$t, 2t, \ldots, 2^{r-2}t$$

is the product of their nonzero leading coefficients times

$$\det\left[(2^{d_i})^j\right]_{1\le i\le r-1,\,0\le j\le r-2}\neq 0.$$

Independence in the quotient gives full column rank over $\mathbb{R}(t)$ in the original vertex coordinates. Thus some $r\times r$ coordinate minor of $[\mathbf{1},F(t),F(2t),\dots,F(2^{r-2}t)]$ is a nonzero integer polynomial.

Each polynomial entry outside the first column has coefficient ℓ_1 norm at most H_0. Every such minor has coefficient ℓ_1 norm at most $N!H_0^{N-1}=H_1$. A nonzero integer polynomial with this bound has no real root greater than H_1+1: its leading coefficient has absolute value at least one and dominates the lower coefficients there. Therefore the same minor is nonzero at $t=\tau$. □

The evaluations $f(t)\mapsto(f(t),f(2t),\dots)$ use the folded-Wronskian rank-preservation construction of Forbes and Shpilka (full version, [FS12, Theorem 4.1 and Corollary 4.3]), building on the method of Gabizon and Raz (earlier ECCC version, [GR08, Lemma 6.1]). See also its use in subspace designs [GK16] and the full version of [FG15, Construction 2.7]. This construction gives a nonzero minor polynomial. The coefficient bound above then supplies a single integer specialization τ valid for all the bounded integer inputs considered here.

A.2 Fraction-free orthogonal residuals

The prefix property permits one orthogonalization formula for every rank. Multiplying each residual by the preceding Gram determinant removes division and keeps the expression polynomial. For $0\le k<N$, let $K_k=[z_0\ \cdots\ z_{k-1}]$ and $G_k=K_k^{\mathsf{T}}K_k$. Define

$$u_0=\mathbf{1},\qquad u_k=(\det G_k)z_k-K_k\operatorname{adj}(G_k)K_k^{\mathsf{T}}z_k\quad(k\ge 1).\tag{14}$$

For $k<r$, u_k is a nonzero multiple of the orthogonal residual of z_k against its predecessors. Hence $u_0,\dots,u_{r-1}$ are mutually orthogonal and nonzero. For $k=r<N$, the matrix G_r is invertible and $z_r\in\operatorname{im}K_r$, so $u_r=0$. For $k>r$ they also vanish: then $\det G_k=0$ and

$$K_k\operatorname{adj}(G_k)=0.$$

Indeed $G_k\operatorname{adj}(G_k)=0$, and $\ker(K_k^{\mathsf{T}}K_k)=\ker K_k$.

To make the first vector cyclic, we couple consecutive residuals through outer products of their adjacent sums.

Set

$$C^{\mathrm{aux}}=u_0u_0^{\mathsf{T}}+\sum_{k=0}^{N-2}(u_k+u_{k+1})(u_k+u_{k+1})^{\mathsf{T}}.\tag{15}$$

In the orthonormal basis $u_i/\|u_i\|$, its restriction to $\mathcal{M}_T$ is tridiagonal with strictly positive off-diagonal entries. Powers of this matrix applied to the first basis vector successively have nonzero coordinates in positions $1,2,\dots,r$. Thus it is cyclic from $u_0=\mathbf{1}$. Moreover, u_0 together with the adjacent sums u_k+u_{k+1} spans $\mathcal{M}_T$. The kernel of their positive-semidefinite outer-product sum is exactly $\mathcal{M}_T^{\perp}$, so its image is $\mathcal{M}_T$.

The vector construction gives the required image and cyclicity. To realize these properties by a polynomial in x,y, we must express the Gram determinants using matrices of the original input order. On a tree of order n, multiples of J_N carry the scalar expressions through $J_N^2=nJ_N$; the resulting positive rescalings of the residuals preserve both properties. This yields the following circuit construction.

Theorem A.2 (Uniform cyclic principal matrix)**.** *There is a deterministic polynomial-time construction, from N alone, of a polynomial $C_N \in \mathbb{Z}\langle x, y\rangle$, represented by an integer matrix circuit and formally a sum of Hermitian squares, with degree at most $2(4N-3)(N-1)$ for $N \geq 2$ and polynomial size and bit complexity such that, on every tree of order at most N,*

$$C_N \succeq 0, \qquad \operatorname{im} C_N = \mathcal{M}_T, \qquad \operatorname{span}\{\mathbf{1}, C_N\mathbf{1}, \ldots, C_N^{N-1}\mathbf{1}\} = \mathcal{M}_T.$$

All its gates act on the original matrix space. Formula (14) requires no determinant, inverse, rank, or trace oracle.

Proof. For $N = 1$ return $C_1 = I$. Assume $N \geq 2$, and abbreviate J_N to J. Compute original-order matrix registers

$$B_{ij} = JZ_i^* Z_j J.$$

On a tree of order n, these satisfy $B_{ij} = g_{ij}J$, where $g_{ij} = z_i^{\mathsf{T}} z_j$, and $J^2 = nJ$. A homogeneous scalar polynomial f of positive degree q in the Gram entries can therefore be compiled so that

$$\widehat{f}(B) = n^{q-1} f(g) J. \tag{16}$$

Indeed, addition of equal degrees preserves this identity, while multiplication of positive degrees p, q gives the factor n^{p+q-1} by $J^2 = nJ$. Degree-zero constants act by prescribed integer scaling; they are never represented by a putative multiplicative identity J.

Division-free determinant circuits are classical [Ber84, Sol02, Bir11]. We use the homogeneous form of Bird's recurrence [Bir11], also stated in [Prz25, Theorem 1.1]: for a $k \times k$ scalar matrix G, put $F_1 = G$ and $F_{q+1} = \mu(F_q)G$, where $\mu(F)$ keeps the strictly upper-triangular entries, has zero entries below the diagonal, and has diagonal entries $-\sum_{\ell>i} F_{\ell\ell}$. Then $(-1)^{k-1}(F_k)_{00} = \det G$. Each stage has homogeneous degree q; shared ordinary matrix multiplication implements the scalar recurrence with $O(k^4)$ arithmetic gates. Thus (16) applies directly, with all integer coefficients implemented as scaling gates. The auxiliary $k \times k$ arrays index original-order matrix registers. Let $b_k = (g_{ik})_{i<k}$ and let $G_k^{(i)}$ be G_k with column i replaced by b_k. The degree-k determinant circuits give

$$\widehat{\Delta}_k = n^{k-1} \det(G_k) J, \qquad \widehat{\Delta}_{k,i} = n^{k-1} \det(G_k^{(i)}) J.$$

Define $V_0 = J$ and, for $1 \leq k < N$,

$$V_k = Z_k \widehat{\Delta}_k - \sum_{i<k} Z_i \widehat{\Delta}_{k,i}.$$

The polynomial column-replacement identity $\det(G_k^{(i)}) = (\operatorname{adj}(G_k) b_k)_i$ holds also when G_k is singular. Thus on trees $V_k = n^{k-1} u_k \mathbf{1}^{\mathsf{T}}$. With $t_0 = \mathbf{1}$ and $t_k = n^{k-1} u_k$ for $k \geq 1$, set

$$C_N = V_0 V_0^* + \sum_{k=0}^{N-2} (V_k + V_{k+1})(V_k + V_{k+1})^*. \tag{17}$$

Its value is $n[t_0 t_0^{\mathsf{T}} + \sum_k (t_k + t_{k+1})(t_k + t_{k+1})^{\mathsf{T}}]$. The t_i have the same orthogonal nonzero prefix and zero tail as the u_i. The Jacobi and image arguments following (15) apply with these positive rescalings, proving the three conclusions. Formula (17) is formally self-adjoint and evaluates to a positive semidefinite matrix on every real symmetric pair; the image and cyclicity claims use the spanning and independence properties proved above.

Since $\deg Z_i, \deg J \leq N-1$, we have

$$\deg B_{ij} \leq 4(N-1), \quad \deg \widehat{\Delta}_k, \deg \widehat{\Delta}_{k,i} \leq 4k(N-1),$$

and $\deg V_k \le (4k+1)(N-1)$. Equation (17) consequently has degree at most $2(4N-3)(N-1)$. There are polynomially many determinant circuits of order at most N. Their auxiliary arrays consist of original-order matrix registers.

For the bit bound, let c_q be the largest coefficient ℓ_1 norm of an entry of F_q, viewed as a polynomial in the scalar matrix entries. The stated Bird recurrence gives

$$c_1 = 1, \qquad c_{q+1} \le k^2 c_q, \qquad c_q \le k^{2(q-1)}.$$

Indeed, an entry of $\mu(F_q)$ is an entry of F_q or a signed sum of at most k such entries, and matrix multiplication sums at most k products. If the scalar inputs have absolute value at most $A \ge 1$, then $|(F_q)_{ij}| \le k^{2(q-1)} A^q$, with the same type of bound for partial sums in each stage. Thus their bit lengths are $O(q(\log(k+1) + \log(A+1)))$. Here the Gram entries have $O(N^6)$ bits and $q \le k \le N$. The homogeneous lift adds powers of n of exponent at most $N-1$; the subsequent products and sums in (17) preserve a polynomial bound on every register's bit length. □

B A full-representation route using only McKay

There is also a polynomial-resource construction using McKay's full adjacency–degree reconstruction theorem as its graph-theoretic input. It has degree $O(N^2)$ and $O(N^4)$ gates. The written argument below uses the classical reconstruction theorem directly; Appendix C specifies the different reconstruction dependency of the formal development.

To invoke McKay's theorem directly, we recover all word traces. These are pointed moments for left multiplication on the Frobenius space with seed I, so the same finite Gram reconstruction applies with dimension bound N^2. Fix a tree T of order $n \le N$ and write $A = A_T$, $D = D_T$. Equip $M_n(\mathbb{R})$ with the Frobenius inner product and seed I. Left multiplication by A and D gives symmetric operators in dimension $n^2 \le N^2$; the latter has spectral diameter at most $N-1$. Theorem 1.2 with dimension bound N^2, and hence recurrence scale N^4, gives $O(N^4)$ polynomials of degree at most N^2-1 spanning $\langle A, D\rangle$.

The three matrices of Frobenius inner products, with entries

$$\operatorname{tr}(U_i^* U_j), \quad \operatorname{tr}(U_i^* A U_j), \quad \operatorname{tr}(U_i^* D U_j)$$

determine the pointed left-regular representation with seed I. Their $(0,0)$ Gram entry is $\operatorname{tr} I = n$, so equality first gives the same order. It also implies equality of $\operatorname{tr} w(A,D)$ for every word w. For every fixed matrix polynomial q, equality also holds for $\operatorname{tr}(q^k)$ for every k. Newton identities therefore give the same characteristic polynomial for q on the two inputs. McKay's theorem [McK77, Theorem 5.3] then distinguishes trees.

These Gram entries are already traces, so their integer encoding requires no additional factor J_N. The factoring used in (12) applies directly. If $M' = O(N^4)$ is the number of probes and b exceeds twice a uniform bound on their integer entries, put $Y = \sum_j b^j U_j$, $Z = \sum_i b^{iM'} U_i$, and $W = 1 + b^{M'^2} x + b^{2M'^2} y$. Then

$$p_N^{\mathrm{full}} = Z^* W Y + Y^* W Z$$

encodes all entries of the three matrices using $O(M') = O(N^4)$ gates and degree at most $2N^2 - 1$. All evaluations still use matrices of the original order; the Frobenius-space enlargement is only in the spanning proof. The resulting polynomial belongs to $\mathbb{Z}\langle x, y\rangle$. The recurrence and integer encoding give polynomial bit complexity by the same norm estimates as in Section 3, with the larger dimension bound. The output expression uses only the polynomials spanning the full algebra and their formal adjoints.

C Computational verification and proof audit

The accompanying artifact at https://github.com/Lzp88/AD contains the matrix-circuit emitters, finite checks, Lean proof sources, and audit records. The emitter `quadratic_ad_circuit.py` takes only N as input. Its exported $N = 5$ example has 407 stored instructions, including input instructions, and output formal degree 13; for $N = 1$ it returns the identity. The instruction set consists of identity, adjacency, degree, zero, addition, multiplication, and integer scaling. Formal adjoints are compiled by reversing products.

The record `current_checks.json` covers circuit syntax for $N = 1, \dots, 15$, all 14 nonisomorphic nonempty trees of order at most six with the common bound $N = 6$, three boundary cases with $N = 1, 2$, and 42 seeded integer symmetric-pair cases. The checks compare the probe span with an independently generated word span or rooted-tree orbit count, evaluate the distance polynomials, and decode the integer trace. Integer trace fingerprints are recorded for the tree and boundary cases. All rank comparisons use exact arithmetic over the rationals.

The record `export_checks.json` compares the Python instruction arrays and output references with those exported by the Lean definitions for $N = 1, \dots, 15$. It also evaluates the division-free PSD factory for $N = 1, \dots, 4$. The tree tests comprise 11 cases (N, T) involving five distinct trees: they compare the output with a separately computed determinant-based Gram-residual formula and check its image, cyclicity, and positive semidefiniteness exactly. Three fixed 2×2 symmetric pairs are each tested at the four values of N, giving 12 additional cases that check symmetry and positive semidefiniteness. These finite comparisons do not prove equality of the Python and Lean emitters for arbitrary N.

The Lean project formalizes the rational differential spanning argument, the prescribed matrix-circuit compiler, and the resource bounds. It also contains a complete formal proof of the principal-moment reconstruction theorem [LL26, Theorem 5.4], in the `PrincipalMomentAudit` namespace (33 source files), and the compiled classification theorem depends on this formal proof. A source-to-statement map identifies the formal declarations and summarizes their hypotheses and scope. These include arbitrary prescribed seeds and proper cyclic modules, the tree distance identity, and the compiled single-trace classification theorem, including $N = 1$. The resource results bound every stored instruction's formal degree and every integer register's bit length on graph inputs of order at most N. The PSD results cover the complete fixed determinant and sum-of-squares construction, and the degree lower bound uses nonisomorphic spider trees and arbitrary real polynomial coefficients. The elementary fixed-word lower bound in Remark 1.4 is proved in the text and is not part of this formalization.

For Appendix B, the project also proves Frobenius-space spanning, the three Gram-matrix identities, and the packed invariant with its resource bounds. Its tree-reconstruction step converts word traces into principal moments using the distance polynomial and then applies [LL26, Theorem 5.4]. The written reconstruction argument in that section instead uses McKay's theorem and Newton identities. The source-to-statement file records these different proof dependencies.

The original formalization comprises 189 modules with a pinned Lean 4.33.0 toolchain and dependencies. The original audit is preserved with separate audited-source and distributed-source hashes, since public identifiers were subsequently standardized. A fresh rebuild of all 189 distributed modules on 25 October 2026 compiled without errors or warnings, and its declaration scan checked 4,704 declarations and 4,639 bodies. The declaration scan permits only `propext`, `Classical.choice`, and `Quot.sound` as transitive axioms. All source hashes remained unchanged, and the 19 fresh circuit export records agree with the original exports. The fresh report and declaration and export logs are in `lean/provenance/rebuild_20261025/`. Running `python -B run_audit.py` in the `lean` directory performs another fresh build and writes its records to `rebuild_logs/`; the `--verify-only`

mode checks package integrity without compiling.

Integrated audit. The public artifact also includes the 25 modules for Appendix D and a common audit entry point. A fresh build compiles all 215 modules, replays every declaration of each module through the Lean kernel with `leanchecker`, and scans the combined declaration environment with the same axiom allowlist. Before compiling, a source scan admits only three elaboration options and confines metaprogramming to the audit and export commands. The recorded audit also replayed the complete import closure, including the Lean core library and mathlib, from an empty environment with `leanchecker --fresh`. The extension audit covers multiple-pole rigidity, the integer recurrence and its evaluation, actual rank-sensitive spanning, coefficient count and word degree, shared circuit resources, and exact specialization to the two-generator recurrence. Its 15 endpoints have explicit type checks; the source-to-statement map records their hypotheses. The integrated report and logs are in `lean/provenance/integrated_audit/`. The additional 43 exact finite tests supplement the universal proofs. No uniform bit bound on evaluated entries of arbitrary real inputs is claimed.

D Several symmetric generators

We give the direct extension of Theorem 1.2. Here $\mathbb{Z}\langle x_0, \dots, x_s\rangle$ denotes the free associative algebra in $s+1$ noncommuting variables, and t commutes with every generator. For $s, N \geq 1$, set

$$\kappa = sN^2, \qquad a_i = i - 1, \qquad D(t) = \prod_{i=1}^{s}(t - a_i), \qquad Q_i(t) = \prod_{j \neq i}(t - a_j).$$

Define $F_k \in \mathbb{Z}\langle x_0, \dots, x_s\rangle[t]$ by

$$\begin{aligned} F_0 &= 1, \\ F_{k+1} &= \kappa D F_k' + \left(\kappa D x_0 + \sum_{i=1}^{s} Q_i x_i - k\kappa D'\right) F_k. \end{aligned} \tag{18}$$

Write $F_k = \sum_{j=0}^{sk} P_{k,j} t^j$, including zero coefficients.

Corollary D.1 (Fixed probes for several symmetric generators)**.** *Let $s, N \geq 1$ and $1 \leq n \leq N$. Suppose that $H, A_1, \dots, A_s \in M_n(\mathbb{R})$ are symmetric and $\operatorname{diam}\operatorname{Spec}(A_i) \leq N - 1$ for every i. For every $v \in \mathbb{R}^n$, let*

$$\mathcal{M} = \operatorname{span}_{\mathbb{R}}\{w(H, A_1, \dots, A_s)v : w \in \{x_0, \dots, x_s\}^*\}, \qquad r = \dim \mathcal{M}.$$

Then

$$\operatorname{span}_{\mathbb{R}}\{P_{k,j}(H, A_1, \dots, A_s)v : 0 \leq k < r,\ 0 \leq j \leq sk\} = \mathcal{M}.$$

In particular, the full family with $k < N$ spans $\mathcal{M}$ and has $N + sN(N-1)/2$ indexed members. It is constructed from (s, N) alone, has noncommutative degree at most $N - 1$, and admits a shared division-free circuit of $O(s^2N^2)$ gates on matrices of the original order. Its prescribed integer constants have $O(s\log(s+1) + \log(N+1))$ bits. No spectral bound is required for H. For $r = 0$, the prefix is the empty family.

Proof. We first extend Lemma 2.1. Let the a_i be distinct real numbers and let $H, B_1, \dots, B_s$ be symmetric. If a rational subspace V of rank r is invariant under

$$\nabla = \partial_t + H + \sum_{i=1}^{s} \frac{B_i}{t - a_i}, \qquad r \sum_{i=1}^{s} \operatorname{diam} \operatorname{Spec}(B_i) < 1,$$

then it is the scalar extension of a real subspace invariant separately under $H, B_1, \dots, B_s$. The rank-zero case is immediate. Otherwise choose a primitive polynomial exterior vector $w(t)$ representing V, of degree ℓ with nonzero leading coefficient w_ℓ. The induced connection satisfies

$$w' + \left(H^{[r]} + \sum_i \frac{B_i^{[r]}}{t - a_i} \right) w = c(t) w.$$

A Bézout combination of the coordinates shows that Dc is polynomial; degree comparison gives $\deg(Dc) \le s$. Consequently $c = \beta + \sum_i \alpha_i/(t - a_i)$. Primitivity gives $w(a_i) \neq 0$ at every pole, and taking residues yields $B_i^{[r]} w(a_i) = \alpha_i w(a_i)$.

Suppose $\ell > 0$. Comparing the two highest terms at infinity gives

$$H^{[r]} w_\ell = \beta w_\ell,$$
$$\ell w_\ell + H^{[r]} w_{\ell-1} + \sum_i B_i^{[r]} w_\ell = \beta w_{\ell-1} + \left(\sum_i \alpha_i \right) w_\ell.$$

Equivalently, these are the highest two coefficient identities after multiplication by D, with its next-to-leading coefficient cancelled using the first identity. Taking the inner product of the second identity with w_ℓ cancels the $H^{[r]}$ terms and gives

$$\ell = \sum_i \left(\alpha_i - \frac{\langle w_\ell, B_i^{[r]} w_\ell \rangle}{\|w_\ell\|^2} \right) \le \sum_i \operatorname{diam} \operatorname{Spec}(B_i^{[r]}) \le r \sum_i \operatorname{diam} \operatorname{Spec}(B_i) < 1.$$

Thus $\ell = 0$. The constant exterior line defines a real subspace W with $V = W \otimes_{\mathbb{R}} \mathbb{R}(t)$. Projecting ∇u to $\mathbb{R}^n/W$ for $u \in W$ and comparing its residues and constant part proves separate invariance.

Apply this result with $B_i = A_i/\kappa$. For every $r \le N$,

$$r \sum_i \operatorname{diam} \operatorname{Spec}(B_i) \le \frac{N-1}{N} < 1.$$

Let $q_0 = v$ and $q_{k+1} = \nabla q_k$. As in the proof of Theorem 1.2, the first dependence is stable by the Leibniz rule. Its rational span lies in $\mathcal{M} \otimes_{\mathbb{R}} \mathbb{R}(t)$; rigidity makes it constant and separately invariant, and it contains v. Minimality of $\mathcal{M}$ therefore identifies this span with $\mathcal{M} \otimes_{\mathbb{R}} \mathbb{R}(t)$. The first r iterates span it. Direct differentiation shows that $f_k = \kappa^k D^k q_k$ obeys (18) evaluated on v. Induction gives $\deg_t F_k \le sk$ and $\deg P_{k,j} \le k$. Every coefficient response lies in $\mathcal{M}$. Conversely, the first r iterates belong to the scalar extension of the span of those coefficients; faithfulness of scalar extension proves the asserted real spanning equality. The full indexed family contains $\sum_{k=0}^{N-1} (sk + 1) = N + sN(N-1)/2$ coefficients.

For the circuit bound, put $G = DF_k$ and $T = DF_k'$. Multiplication by D can be performed successively through its s monic linear factors. From the shared G, synthetic division by $t - a_i$ computes $Q_i F_k$ using only additions and multiplication by the fixed integer a_i. Since $D' F_k = G' - T$, the next row can be computed as

$$F_{k+1} = \kappa(k+1) T + \kappa x_0 G + \sum_i x_i (Q_i F_k) - k\kappa G'.$$

Scalar polynomials commute with the noncommuting generators, so this is exactly (18). The factor products and all s synthetic divisions require $O(s^2(k+1))$ gates in row k; the derivative and generator operations obey the same bound. Summing over $k < N$ gives $O(s^2N^2)$. This factor implementation needs only pole values, derivative indices and integer multiples of κ, which satisfy the stated constant-bit bound. All generator multiplications are on the left and all matrix registers retain the input order. For $s = 1$, $D = t$ and (18) gives exactly (1), with $x_0 = x$ and $x_1 = y$. □

The bound $N - 1$ remains necessary: set every $A_i = 0$ and use the diagonal H and all-ones seed from Proposition 2.2. The case of a single generator ($s = 0$) is the ordinary Krylov family $1, H, \ldots, H^{N-1}$ and is separate from the formula $\kappa = sN^2$.

AI disclosure

ChatGPT and OpenAI Codex helped develop parts of the supporting theory, including the formulation and refinement of intermediate results, and assisted with checking the proofs of most theoretical results, including results in Sections 2–4 and Appendices A, B, and D. They also assisted with the literature review, drafting and editing, and the development and verification of the Lean formalizations described in Appendix C. Claude Code assisted with re-checking the Lean audit through an independent rebuild, and with adding the kernel replay to the audit runner. The author takes full responsibility for the content of this paper.

References

[AI03] Yu. A. Al'pin and Kh. D. Ikramov, On the unitary similarity of matrix families, *Mathematical Notes* **74** (2003), no. 5–6, 772–782. Publisher's version; Original Russian version and metadata.

[ADFK25] V. Arvind, Samir Datta, Salman Faris, and Asif Khan, Revisiting tree canonization using polynomials, in *SOSA 2025*, 465–472. Publisher's version; arXiv:2408.10338.

[BG19] Moulay A. Barkatou and Renat R. Gontsov, Linear differential systems with small coefficients: Various types of solvability and their verification, *SIGMA* **15** (2019), article 058, 15 pp. Publisher's version; arXiv:1901.09951.

[Ber84] Stuart J. Berkowitz, On computing the determinant in small parallel time using a small number of processors, *Information Processing Letters* **18** (1984), no. 3, 147–150. Publisher's version.

[Bir11] Richard S. Bird, A simple division-free algorithm for computing determinants, *Information Processing Letters* **111** (2011), no. 21–22, 1072–1074. Publisher's version.

[DM20] Harm Derksen and Visu Makam, Algorithms for orbit closure separation for invariants and semi-invariants of matrices, *Algebra & Number Theory* **14** (2020), no. 10, 2791–2813. Publisher's version; arXiv:1801.02043.

[FXZW26] Yi-Zheng Fan, Ruo-Jie Xing, Yi-Liu Zhang, and Wei Wang, Degree-similar graphs and cospectral graphs, *Electronic Journal of Combinatorics* **33** (2026), no. 2, article P2.62. Publisher's version; arXiv:2509.01520.

[FG15] Michael A. Forbes and Venkatesan Guruswami, Dimension expanders via rank condensers, in *APPROX/RANDOM 2015*, LIPIcs 40 (2015), 800–814. Publisher's version; full version: arXiv:1411.7455.

[FS12] Michael A. Forbes and Amir Shpilka, On identity testing of tensors, low-rank recovery and compressed sensing, in *STOC 2012*, 163–172. Publisher's version; full version: arXiv:1111.0663.

[FS13] Michael A. Forbes and Amir Shpilka, Explicit Noether normalization for simultaneous conjugation via polynomial identity testing, in *APPROX/RANDOM 2013*, LNCS 8096, 527–542. Publisher's version; full version: arXiv:1303.0084.

[FHS17] Vyacheslav Futorny, Roger A. Horn, and Vladimir V. Sergeichuk, Specht's criterion for systems of linear mappings, *Linear Algebra and its Applications* **519** (2017), 278–295. Publisher's version; arXiv:1701.08826.

[GR08] Ariel Gabizon and Ran Raz, Deterministic extractors for affine sources over large fields, *Combinatorica* **28** (2008), no. 4, 415–440. Publisher's version; Earlier ECCC version.

[GS25] Chris Godsil and Wanting Sun, Degree-similar graphs, *Electronic Journal of Combinatorics* **32** (2025), no. 4, article P4.12. Publisher's version; arXiv:2407.11328.

[GV15] Renat R. Gontsov and Ilya V. Vyugin, Solvability of linear differential systems with small exponents in the Liouvillian sense, *Arnold Mathematical Journal* **1** (2015), no. 4, 445–471. Publisher's version; earlier preprint: arXiv:1312.2518.

[GRS25] Martin Grohe, Gaurav Rattan, and Tim Seppelt, Homomorphism tensors and linear equations, *Advances in Combinatorics* (2025), Paper 4, 50 pp. Publisher's version; arXiv:2111.11313.

[GK16] Venkatesan Guruswami and Swastik Kopparty, Explicit subspace designs, *Combinatorica* **36** (2016), no. 2, 161–185. Publisher's version.

[JC24] Yifan Jia and Angela Capel, A generic quantum Wielandt's inequality, *Quantum* **8** (2024), article 1331. Publisher's version.

[Kat87] Nicholas M. Katz, A simple algorithm for cyclic vectors, *American Journal of Mathematics* **109** (1987), no. 1, 65–70. Publisher's version; Author's copy.

[KMOWW11] Stefan Kiefer, Andrzej S. Murawski, Joël Ouaknine, Björn Wachter, and James Worrell, Language equivalence for probabilistic automata, in *CAV 2011*, LNCS 6806, 526–540. Publisher's version; Author's version.

[KŠ16] Igor Klep and Špela Špenko, Sweeping words and the length of a generic vector subspace of $M_n(\mathbb{F})$, *Journal of Combinatorial Theory, Series A* **143** (2016), 56–65. Publisher's version; arXiv:1506.03460.

[Li+25] Pengxiang Li, Wei Wu, Zhi Gao, Xiaomeng Fan, Peilin Yu, Yuwei Wu, Zhipeng Lu, Yunde Jia, and Mehrtash Harandi, A Set-to-Set Distance Measure in Hyperbolic Space, preprint, 2025, arXiv:2506.18529, version 1.

[Lin92] Steven Lindell, A logspace algorithm for tree canonization (extended abstract), in *STOC 1992*, 400–404. Publisher's version.

[LL26] Zhipeng Lu and Pengxiang Li, Adjacency-degree algebras and spectral determination of graphs, preprint, 2026, arXiv:2607.21494, version 1.

[McK77] Brendan D. McKay, On the spectral characterisation of trees, *Ars Combinatoria* **3** (1977), 219–232. Author's scan: https://users.cecs.anu.edu.au/~bdm/papers/SpectralTrees.pdf.

[Pap97] Christopher J. Pappacena, An upper bound for the length of a finite-dimensional algebra, *Journal of Algebra* **197** (1997), no. 2, 535–545. Publisher's version.

[Paz84] Azaria Paz, An application of the Cayley–Hamilton theorem to matrix polynomials in several variables, *Linear and Multilinear Algebra* **15** (1984), no. 2, 161–170. Publisher's version.

[Pro76] Claudio Procesi, The invariant theory of $n \times n$ matrices, *Advances in Mathematics* **19** (1976), no. 3, 306–381. Publisher's version.

[Prz25] Adam J. Przeździecki, A simple division-free algorithm for computing Pfaffians, *Information Processing Letters* **189** (2025), article 106550. Publisher's version; arXiv:2302.12081.

[Raz74] Yuri P. Razmyslov, Trace identities of full matrix algebras over a field of characteristic zero, *Mathematics of the USSR-Izvestiya* **8** (1974), no. 4, 727–760. Publisher's version.

[Sch73] Allen J. Schwenk, Almost all trees are cospectral, in Frank Harary (ed.), *New Directions in the Theory of Graphs*, Academic Press, New York, 1973, 275–307.

[Shi19] Yaroslav Shitov, An improved bound for the lengths of matrix algebras, *Algebra & Number Theory* **13** (2019), no. 6, 1501–1507. Publisher's version; arXiv:1807.09310.

[Sol02] Michael Soltys, Berkowitz's algorithm and clow sequences, *Electronic Journal of Linear Algebra* **9** (2002), 42–54. Publisher's version.

[Spr26] Florian Ito Sprung, A log-log saving for matrix-algebra length and terseness, preprint, 2026, arXiv:2607.16679, version 2.

[Tze92] Wen-Guey Tzeng, A polynomial-time algorithm for the equivalence of probabilistic automata, *SIAM Journal on Computing* **21** (1992), no. 2, 216–227. Publisher's version.

[vPS03] Marius van der Put and Michael F. Singer, *Galois Theory of Linear Differential Equations*, Grundlehren der mathematischen Wissenschaften, vol. 328, Springer-Verlag, Berlin–Heidelberg, 2003. Publisher's version.

[WZ26] Wei Wang and Da Zhao, Graph isomorphism and multivariate graph spectrum, *Advances in Applied Mathematics* **173** (2026), Part B, article 102994. Publisher's version; arXiv:2412.20016.