MINIMAL FILLING ARCHITECTURES OF POLYNOMIAL NEURAL NETWORKS: A PROOF FOR THE DAO–RODRIGUEZ TABLE, A CERTIFIED CENSUS AT DEPTH EIGHT, EXACT NEUROVARIETY DIMENSIONS, AND A CORRECTED DEFECTIVITY CRITERION

KOREA SUPERINTELLIGENCE LABS

Abstract. Fix an exponent r and an architecture $d = (d_0, \dots, d_L)$. A polynomial neural network with power activation $z \mapsto z^r$ parametrises a subvariety $V_{d,r}$ — the *neurovariety* — of a space of tuples of forms of degree r^{L-1}; the architecture is *filling* when $V_{d,r}$ is everything and *minimal filling* when in addition no smaller architecture of the same depth and the same input and output widths is filling. Dao and Rodriguez have recently refuted the conjecture of Kileel, Trager and Bruna that every minimal filling architecture has unimodal widths, and list eighteen minimal filling architectures for $d_0 = 2$, $d_L = 1$, $r = 2$, $L \le 7$. They prove that exactly one of the eighteen is minimal filling, and write that finite-field computations "strongly suggest" the rest, a proof needing "substantially more cases". We give that proof. Relative to five dimension bounds, four of them the source's own and the fifth an elementary Grassmann incidence count, we show that all 93 maximal subarchitectures (83 distinct) of the eighteen entries have dimension strictly below the ambient one; with explicit integer parameter points whose Jacobians have a nonzero maximal minor modulo 101, every entry of the table is a minimal filling architecture. The same method proves the source's $r = 3$ counterexample minimal filling, settles its depth-9 counterexample with $d_L = 2$ in seven of eight cases, and determines 71 neurovariety dimensions exactly, of which 63 appear not to have been recorded. At depth 8, where the source reports a count of 82 minimal filling architectures and prints none of them, its published search output flags 111; we certify all 111 filling and prove 107 of them minimal filling, so at least 107 minimal filling architectures of depth 8 exist for $d_0 = 2$, $d_L = 1$, $r = 2$. Seventeen of the 107 have non-unimodal widths, each a new counterexample to the conjecture of Kileel, Trager and Bruna. Of the 605 distinct maximal subarchitectures at depth 8, 533 have their dimension determined exactly; the four flagged architectures not decided are blocked by two named subarchitectures. Both censuses are also shown not to depend on the sharper of two readings of the incidence bound. We also correct a defectivity criterion quoted in the source: as printed it is one step too wide, and $(2,2,3)$, $(2,2,4)$, $(2,2,5)$ satisfy it without being defective. Every statement below is formally verified in Lean 4, with one clause flagged as the exception.

1. Introduction

1.1. **The objects.** Fix $r \in \mathbb{N}$ and a sequence $d = (d_0, \dots, d_L)$ of positive integers. A *polynomial neural network* with power activation is a map

$$p_\theta : \mathbb{R}^{d_0} \to \mathbb{R}^{d_L}, \qquad p_\theta(x) = W_L\, \sigma\, W_{L-1}\, \sigma \cdots \sigma\, W_1 x, \qquad \sigma(z)_j = z_j^r, \tag{1}$$

with parameters $\theta = (W_1, \dots, W_L)$, $W_i \in \mathbb{R}^{d_i \times d_{i-1}}$. The sequence d is the *architecture* and L the *depth*. Each coordinate of p_θ is a homogeneous form of degree $D = r^{L-1}$ in d_0 variables, so (1) defines the *parameter map*

$$\Psi_{d,r} :\ \mathbb{R}^{d_1 \times d_0} \times \cdots \times \mathbb{R}^{d_L \times d_{L-1}} \longrightarrow \mathrm{Sym}_D(\mathbb{R}^{d_0})^{\oplus d_L}, \qquad \theta \mapsto p_\theta, \tag{2}$$

whose codomain, the *ambient space*, has dimension

$$A_r(d) \ :=\ d_L \binom{d_0 + D - 1}{D}, \qquad D = r^{L-1}. \tag{3}$$

The image $\mathcal{F}_{d,r}$ of $\Psi_{d,r}$ is the *functional space*, and its Zariski closure

$$V_{d,r} \ :=\ \overline{\mathcal{F}_{d,r}}$$

is the *neurovariety*, an irreducible algebraic variety. The names are those of [1, Definition 1.3], which calls $\mathcal{F}_{d,r}$ the functional space throughout and nowhere uses the term *neuromanifold* that [4] gives to the same image.

The architecture d is *filling* when $V_{d,r}$ is the whole ambient space. Writing $d' \le d$ for the coordinate-wise order on architectures of the same depth, a *subarchitecture* of d is a $d' < d$; and d is a *minimal filling architecture* (MFA) when d is filling and every subarchitecture $d' = (d_0, d'_1, \dots, d'_{L-1}, d_L) < d$ is not. These are the definitions of [1], which follows [3, 4].

1.2. **The source and the sentence this paper closes.** Kileel, Trager and Bruna [3, Conjecture 12] conjectured that, with r, L, d_0 and d_L fixed, every MFA has *unimodal* widths: there is an i with $d_0 \le \cdots \le d_i$ and $d_i \ge \cdots \ge d_L$. (It is restated as [1, Conjecture 1.7].) Dao and Rodriguez [1] refute it: $(2,3,4,5,4,6,4,1)$ is an MFA at $r = 2$ and is not unimodal, and they exhibit two further counterexamples, one at $r = 3$ and one with $d_L = 2$.

Their Table 2 lists *eighteen* architectures — their complete search output of MFAs for $d_0 = 2$, $d_L = 1$, $r = 2$ and $L \le 7$: one each at depth $2, 3, 4, 5, 6$ and thirteen at depth 7 (the first column of Table 1 below). Of the eighteen they prove exactly one, the counterexample, filling in their Lemma 2.2 and minimal in their Lemma 2.6. Of the rest they write, verbatim:

> "Fix d_0, d_L, r, and L. By Dickson's Lemma, one knows there are finitely many MFAs. Finite field computations over $p = 2^{31} - 1$ strongly suggest that every entry in Table 2 is minimal filling. A proof would follow the same strategy as Lemma 2.6, but with substantially more cases."

The gap between "strongly suggest" and a proof is not rhetorical, and it is not symmetric. A Jacobian of *full* rank at one integer point, read modulo a prime, certifies filling outright (Section 3); a Jacobian of *deficient* rank certifies nothing at all, because rank can drop both off the generic locus and under reduction. Non-filling — which is what minimality asks for — therefore has to be proved by dimension bounds, one architecture at a time. That is what [1] does for six subarchitectures by hand, and what it declines to do for the other eighty-seven.

Depth 8 is treated in [1] by one sentence of its Discussion (Section 4 of `v2`), quoted verbatim as printed:

> "Fix $d_0 = 2$, $d_L = 1$, and $r = 2$. Our search found 13 MFAs of depth $L = 7$ with only one among them being nonunimodal: the architecture discussed in Section 2. For depth $L = 8$, we found 82 MFAs and among them, 20 were nonunimodal"

followed by a footnote naming the repository [2] for "our search and implementation" (the section referred to is the source's own Section 2, on $(2,3,4,5,4,6,4,1)$). No depth-8 architecture, dimension or proof is printed, and two paragraphs later the same Discussion says: "At present, there is no general method for proving completeness of such lists, since we do not know, a priori, how large the hidden widths must be in order to guarantee that all minimal filling architectures have been found." The search output published in [2] — the file `data/processed/2_1_cleaned_architectures.csv`, 12,907 rows, of which 6,639 have depth 8 — flags 111 depth-8 architectures as minimal and filling, 20 of them non-unimodal. The 20 agrees with the sentence; the 82 does not, and we have not been able to reproduce it from the published data (Section 10). The flag itself is provisional by the repository's own notice, which reads: "One should be cautious to check if an architecture is truly minimal. The column "is_minimal" is in reference to the search *thus far*. It is possible for it to actually not be minimal and filling because the search has not uncovered a subarchitecture that is filling yet. One should manually check minimality by computing the dimension of the subarchitectures." The caution is real: of the 605 distinct maximal subarchitectures of the 111 flagged architectures, 601 do not occur in the file at all. Version 2 of this paper does the check the notice asks for, by proof rather than by computation.

1.3. **Results.** Throughout, five upper bounds on neurovariety dimensions are used as *inputs*: they are stated as (D1)–(D5) in Section 4, four are from [1, 3] and the fifth is an elementary incidence count. Every upper-bound statement below is relative to them; every lower bound is relative to the source's

Theorem A.2, which identifies $\dim V_{d,r}$ with the generic rank of $\operatorname{Jac}\Psi_{d,r}$. Nothing here reproves those inputs.

Theorem 1.1 (Theorem A)**.** *Assume* (D1)–(D5) *at $r = 2$. Then for each of the eighteen architectures d of* [1, Table 2] *and each of its maximal subarchitectures s,*

$$\dim V_{s,2} \;<\; A_2(s),$$

so no maximal subarchitecture of a Table-2 entry is filling. The 93 bounds are those of Table 1. Consequently, granting also the filling certificates of Proposition 3.2 and the monotonicity (M) *of Section 2, every entry of* [1, Table 2] *is a minimal filling architecture.*

Theorem 1.2 (Theorem B)**.** *Assume* (D1)–(D5) *at $r = 3$. The architecture $(2, 3, 7, 6, 15, 8, 1)$ at $r = 3$, of ambient dimension 244, is filling — certified by a 244×244 minor of $\operatorname{Jac}\Psi$ equal to 56 modulo 101 at an explicit integer point — and each of its five maximal subarchitectures has dimension at most $84, 240, 225, 235, 233$, all below 244. It is therefore a minimal filling architecture.*

Theorem 1.3 (Theorem C)**.** *Assume* (D1)–(D5) *at $r = 2$, and let d be the depth-9 architecture $(2, 3, 4, 4, 10, 17, 11, 12, 4, 2)$, of ambient dimension 514 and 619 parameters. Then d is filling — certified by a 514×514 minor equal to 54 modulo 101 — and seven of its eight maximal subarchitectures have dimension strictly below 514, of codimension at least $254, \cdot, 176, 5, 17, 34, 11, 124$. The eighth, the one lowering d_2, is not reached: the calculus of Section 4 bounds it only by the ambient dimension.*

Theorem 1.4 (Theorem D)**.** *Assume* (D1)–(D5) *at $r = 2$ and the source's Theorem A.2. Among the 83 distinct maximal subarchitectures of the eighteen entries of* [1, Table 2]*, the certified Jacobian rank at an explicit integer point and the calculus bound agree in 71 cases, determining $\dim V_{s,2}$ exactly there, and never cross. The remaining 12 are listed in Table 2, with gaps between 1 and 3.*

Eleven of the 83 have a dimension on record: the six of [1, Table 1], and five rows of the search data accompanying [2], one of which is the worked example $(2, 1, 1) \mapsto 2$ of [1, Example A.8]. Eight of those eleven are among the 71 determined here, and all eight values agree with ours, computed over a different prime by an independent implementation. We have not found the other 63 in the literature (Section 1.5).

Finally, a correction. [1, Theorem B.6] is quoted from [6] as a criterion for defectivity of a depth-2 neurovariety. As printed it admits architectures that are not defective.

Proposition 1.5 (Proposition E)**.** [1, Theorem B.6] *as printed asserts that if $r = 2$ and $\binom{d_0-1+r}{r} - d_0 < d_1 < \min\{d_2, \binom{d_0-1+r}{r}\}$ — at $r = 2$, that is $\binom{d_0+1}{2} - d_0 < d_1 < \min(d_2, \binom{d_0+1}{2})$ — then $V_{(d_0,d_1,d_2),2}$ is defective. At $d_0 = 2$ the hypothesis reads $d_1 = 2$, $d_2 \geq 3$; but $(2, 2, 3)$, $(2, 2, 4)$ and $(2, 2, 5)$ have Jacobian rank $8, 10, 12$ at explicit integer points, equal to their expected dimensions, hence defect 0. The subspace bound* (D4) *proves defectivity exactly on*

$$\binom{d_0+1}{2} - d_0 + 1 \;<\; d_1 \;<\; \min\Big(d_2, \binom{d_0+1}{2}\Big),$$

one step narrower on the left; the three counterexamples sit on the extra step, and the corrected region still contains $(3, 5, 6)$, the only instance [1] *uses.*

Version 2 carries the same method to depth 8, where the ambient dimension is $A_2(d) = \binom{2+2^7-1}{2^7} = 129$ for every $d = (2, d_1, \ldots, d_7, 1)$. Write $\mathcal{C}_8$ for the 107 architectures of Table 3 and $\mathcal{O}_8$ for the four of Table 4; together they are the 111 architectures that the search output of [2] flags as minimal filling.

Theorem 1.6 (Theorem F)**.** *Assume* (D1)–(D5) *at $r = 2$. Each of the 111 flagged depth-8 architectures carries a filling certificate (Proposition 3.4). For each $d \in \mathcal{C}_8$ and each of its seven maximal subarchitectures s,*

$$\dim V_{s,2} \;<\; 129 \;=\; A_2(s),$$

so, granting also (M)*, every one of the 107 architectures of $\mathcal{C}_8$ is a minimal filling architecture. Consequently there are at least 107 minimal filling architectures of depth 8 with $d_0 = 2$, $d_L = 1$, $r = 2$. The* *four architectures of $\mathcal{O}_8$ are not decided: each has exactly one maximal subarchitecture that the calculus of Section 4 bounds only by* 129*, and only two such subarchitectures occur (Proposition 10.1).*

The number 107 is a lower bound for the census and nothing more is claimed of the census itself: the source does not claim completeness at depth 8, and neither do we (Section 1.4). What Theorem 1.6 settles is that the count 82 printed in [1] is not the number of depth-8 minimal filling architectures; how that sentence came to differ from the artefact it cites we do not know.

Theorem 1.7 (Theorem G)**.** *Assume* (D1)–(D5) *at $r = 2$. Exactly seventeen of the* 107 *architectures of $\mathcal{C}_8$ have non-unimodal width sequences, namely*

$$\begin{array}{llll}(2,3,3,4,7,6,7,6,1), & (2,3,3,5,7,5,8,5,1), & (2,3,3,5,7,6,7,4,1), & (2,3,3,6,6,5,8,8,1),\\ (2,3,3,6,7,6,7,3,1), & (2,3,4,4,7,5,8,8,1), & (2,3,4,4,7,6,7,5,1), & (2,3,4,5,4,7,8,5,1),\\ (2,3,4,5,4,8,8,4,1), & (2,3,4,5,6,5,8,5,1), & (2,3,4,5,7,5,7,5,1), & (2,3,4,5,7,5,8,4,1),\\ (2,3,4,5,7,6,7,3,1), & (2,3,4,5,8,5,7,4,1), & (2,3,4,6,5,8,6,3,1), & (2,3,4,6,6,5,7,7,1),\\ (2,3,4,6,7,5,7,4,1). & & & \end{array}$$

Each is a minimal filling architecture by Theorem 1.6, hence a counterexample to [3, Conjecture 12]*.*

Before this work one non-unimodal minimal filling architecture had a printed proof, namely $(2,3,4,5,4,6,4,1)$, by [1, Theorem 2.1]; Theorem 1.2 proves a second one and Theorem 1.3 a third one in seven of its eight cases. Theorem 1.7 adds seventeen, so the non-unimodal minimal filling architectures in the record now number twenty, nineteen of them with a complete proof (the twentieth, the $d_L = 2$ architecture of Theorem 1.3, is asserted by [1, Example 4.2] and proved here in seven of its eight cases). The source asserts 20 non-unimodal minimal filling architectures at depth 8 and lists none; its search output flags exactly 20, and the three of those 20 that are missing from Theorem 1.7 are three of the four undecided architectures of $\mathcal{O}_8$. Unimodality here is decided exactly as [1, Conjecture 1.7] states it: walk up the non-decreasing initial segment, then down the non-increasing one, and ask whether the walk reaches the end (Section 2).

Theorem 1.8 (Theorem H)**.** *Assume* (D1)–(D5) *at $r = 2$ and the source's Theorem A.2. Among the* 605 *distinct maximal subarchitectures of the* 111 *flagged depth-*8 *architectures, the certified Jacobian rank at an explicit integer point and the calculus bound agree in* 533 *cases, determining* $\dim V_{s,2}$ *exactly there, and never cross. The remaining* 72 *gaps are* 1 *(*24 *cases),* 2 *(*8*),* 3 *(*20*),* 4 *(*11*),* 5 *(*4*),* 6 *(*3*),* 7 *(*1*) and* 8 *(*1*).*

Only four of the 605 occur anywhere in the search output of [2], and [1] prints no depth-8 dimension, so 533 values enter the record here. A wider comparison, made outside the formal development, is the strongest cross-check in this paper: over all 12,907 rows of that search output, at every depth, the calculus bound of Section 4 never falls below the dimension the file records and equals it on 11,779 rows (Section 10.5).

Finally, at depth 8 the subspace bound (D4) is load-bearing — 30 of the 107 certificates need it — and Section 11 re-examines it: as a robustness check, the formal development also proves both the depth-7 and the depth-8 census under a weaker form (D4′) of (D4), and it derives the natural interior-layer generalisation (D6) and measures that it closes none of the open cases.

1.4. **What is not claimed.** Four limits are worth stating before the details.

First, everything above is *relative*. The five dimension inputs (D1)–(D5), the identity of $\dim V_{d,r}$ with the generic Jacobian rank, and the monotonicity $d' \leq d \Rightarrow V_{d',r} \subseteq V_{d,r}$ are cited, not reproved here; they are exactly the facts [1] itself uses. What is proved is that, granting them, the finite arithmetic closes.

Second, we prove that the eighteen listed architectures *are* MFAs. We do not prove that the list is *complete* — that there is no nineteenth MFA for $d_0 = 2$, $d_L = 1$, $r = 2$, $L \leq 7$. Completeness is the subject of [1, Appendix B] and is a statement about infinitely many architectures; it is untouched here.

Third, Theorem 1.3 is partial by design, and the missing case is named. Proposition 1.5 is a correction to a hypothesis *as printed in* [1] and not to [6], which [1] quotes and whose own statement carries the corrected endpoint (Section 8). Nothing in the source's own argument breaks: the corrected region covers its one application.

Fourth, at depth 8 we prove that 107 named architectures are minimal filling architectures and that four more are filling. We do not prove that the four are minimal, we do not prove that the 111 flagged architectures exhaust the search box of [2], and we do not prove anything about architectures outside that box; [1] itself says that no method for completeness at depth 8 is known. The discrepancy between the printed count 82 and the 111 flagged in the published search output is reported as a discrepancy between a sentence and the artefact it cites, not as the refutation of a theorem: the sentence reports what a search found. What is settled is that at least 107 depth-8 minimal filling architectures exist.

1.5. **Related work, and where this claim was looked for.** The dimension theory of neurovarieties is [3] and [4]; the recursive bound (D5) is [1, Lemma 2.4], from [3], and the dimension code of [2] is credited there to the MathRepo page of [4]. Alexander–Hirschowitz [7] enters only at depth 2, where at $r = 2$ it is the classical dimension of a determinantal variety of symmetric matrices. The Segre–Veronese secant computations of [6] and the Alexander–Hirschowitz theorem for neurovarieties of [5] are the two dimension inputs [1, Appendix B] quotes from those two papers; the bound (D4) below supplies both, which matters because [5] was revised on 1 September 2026 with the comment "Thanks to Kevin Dao, we identified a gap in the previous proof of the main non-defectiveness result", so the version [1] cites is the one with the gap.

We searched for a proof of the minimality half, and for the dimension values of Theorem 1.4, on 3 September 2026: the live record and `v2` source of [1]; the whole of [2], whose notebook decides "short" from a Jacobian rank below ambient at a random point — the direction that cannot prove non-filling; the citation graphs of [1] and [3]; arXiv full text, where `"minimal filling architecture"` returns [1] alone, `"filling architecture"` six papers and `"neurovariety"` exactly nine, the whole subject; and a local corpus of 372 arXiv full-text shards, where `"minimal filling architecture"` occurs in two shards only, those holding [1] itself and [3]. We read these as "not found", not as "new". The search data of [2] was pushed three days before this work and already carries minimality flags; a successor of [1] proving Table 2 minimal would supersede Theorems 1.1–1.4 as priority claims, though not as proofs.

The depth-8 searches were repeated on the same day, with the same outcome: the arXiv record of [1] still has two versions; a fresh full-history clone of [2] (seven commits, from 7 May to 30 June 2026; the depth-8 block of the search output — its 6,639 rows with their flags and recorded dimensions — is identical in every one of them, the file having only been moved to `data/processed` on 17 June 2026, the day `v2` was posted) contains no depth-8 proof, no list beyond the flagged rows, and no formal development; OpenAlex lists no work citing [1] and Semantic Scholar exactly one, on rank-deficient critical points; and the three arXiv full-text counts are unchanged. We read this, again, as "not found": a successor of [1] publishing the depth-8 list with proofs would supersede Theorems 1.6–1.8 as priority claims.

Changes from version 1. Version 1 of this paper contained Sections 2–9 in the form in which they stand here: every numbered statement of that version keeps its number and its wording, Tables 1 and 2 are unchanged, and nothing is withdrawn. What is added is the following.

- Depth 8, end to end (Section 10): the 111 candidate architectures read from the search output of [2]; a filling certificate for each (Proposition 3.4); the census of 107 proved minimal filling architectures (Theorem 1.6, Table 3); the seventeen new non-unimodal ones (Theorem 1.7); 533 exact dimensions among 605 (Theorem 1.8); and the four undecided architectures with the two subarchitectures that block them (Proposition 10.1, Table 4).
- The count 82 printed in the Discussion of [1], stated as a discrepancy with the source's own published search output, which flags 111 (Section 10.1).
- The subspace bound re-examined (Section 11): it is needed for 30 of the 107 depth-8 certificates; the weaker form (D4′), under which both censuses are re-proved as a robustness check (Proposition 11.1); and the interior-layer bound (D6), derived and measured to close nothing (Proposition 11.3).

- Section 12 is rewritten to cover the new material, and Appendix A reproduces the formal statement of every machine-checked result. The title and abstract are extended accordingly; Section 1.4 now counts four limits instead of three; a paragraph on unimodality is added to Section 2, Proposition 3.4 to the end of Section 3, and one sentence pointing to Section 11 to the end of Section 6.

2. Preliminaries

We work with the definitions of Section 1. Two conventions and one cited fact are used throughout.

Maximal subarchitectures. For $d = (d_0, \dots, d_L)$ put

$$\mathcal{M}(d) \;:=\; \big\{ (d_0, \dots, d_i - 1, \dots, d_L) \;:\; 1 \le i \le L-1 \big\},$$

the $L-1$ architectures obtained by lowering one hidden width by one. Since A_r depends only on d_0, d_L, r and L, all members of $\mathcal{M}(d)$ have the same ambient dimension as d.

(M) Monotonicity. If $d' \le d$ then $V_{d',r} \subseteq V_{d,r}$ ([1, Lemma 2.3], from [4], together with the induction on one coordinate at a time stated immediately after it). Hence if every $s \in \mathcal{M}(d)$ has $\dim V_{s,r} < A_r(s)$, then every subarchitecture $d' < d$ has $\dim V_{d',r} < A_r(d')$, so d is an MFA as soon as it is also filling. This is the reduction from the definition's infinite-looking quantifier to a finite check, and it is cited, not reproved.

Unimodality. A sequence $d = (d_0, \dots, d_L)$ is *unimodal* when there is an i with $d_0 \le \cdots \le d_i$ and $d_i \ge \cdots \ge d_L$ [1, Conjecture 1.7]. It is decided by a walk: advance while the sequence does not decrease, then advance while it does not increase, and accept when the walk reaches d_L. For example $(2, 3, 4, 5, 4, 6, 4, 1)$ is not unimodal (the walk stops at the 4 before the 6), $(2, 3, 4, 5, 6, 4, 2, 1)$ is, and a sequence with a strict interior valley such as $(2, 2, 1, 2)$ is not.

Expected dimension and defect. Following [1],

$$\operatorname{expdim}(V_{d,r}) \;:=\; \min\Big(A_r(d),\; E(d)\Big), \qquad E(d) := \sum_{i=1}^{L} d_i d_{i-1} - \sum_{i=1}^{L-1} d_i, \tag{4}$$

the second term being the parameter count minus the dimension of the $(\mathbb{C}^*)^{L-1}$ rescaling redundancy of the parametrisation; the *defect* is $\operatorname{expdim}(V_{d,r}) - \dim V_{d,r}$, and d is *defective* when the defect is positive.

3. Certificates for filling

Every entry of $\operatorname{Jac}\Psi_{d,r}$ is a polynomial with integer coefficients in the weights, so at an integer point θ the Jacobian is an integer matrix. The step that turns a finite-field computation into a proof is elementary and one-way.

Lemma 3.1. *Let $J \in \mathbb{Z}^{m \times n}$, let $k \le \min(m, n)$, and let A be the $k \times k$ submatrix of J on some choice of k rows and k columns. If $\det A \ne 0$ in $\mathbb{Z}/N$ for some modulus $N \ge 1$, then $\operatorname{rank}_{\mathbb{Q}} J \ge k$. No primality of N is needed.*

Proof. det commutes with the ring homomorphism $\mathbb{Z} \to \mathbb{Z}/N$, so $\det A = 0$ would force $\det(A \bmod N) = 0$; hence $\det A \ne 0$ in $\mathbb{Z}$, A is invertible over $\mathbb{Q}$, and the rank of a submatrix never exceeds the rank of the matrix. □

The converse fails, and the failure is not academic: reduction can drop the rank. The one-line witness is $\det[11] = 11 \ne 0$ over $\mathbb{Q}$ while $11 = 0$ in $\mathbb{Z}/11$; the same phenomenon on the source's own 65×110 Jacobian is recorded in Section 9.

Combining Lemma 3.1 with [1, Theorem A.2] — the generic rank of $\operatorname{Jac}\Psi_{d,r}$ equals $\dim V_{d,r}$, and the rank at a point is a lower bound for it — we call a triple

$$\big(\theta \in \mathbb{Z}^{\sum_i d_i d_{i-1}},\; p \text{ prime},\; C \text{ a set of } A_r(d) \text{ parameter coordinates}\big) \tag{5}$$

a *filling certificate* for d when the maximal minor of $\operatorname{Jac}\Psi_{d,r}(\theta)$ on the columns C is nonzero modulo p. A filling certificate proves that d is filling; the check verifies that p is prime, that the selected submatrix is square of size exactly $A_r(d)$, and that its determinant is nonzero.

Proposition 3.2. *Each of the eighteen architectures of* [1, Table 2] *carries a filling certificate with* $p = 101$ *and all weight entries in* $\{1, 2, 3\}$*; the minors are the last column of Table 1. So do the five architectures with* $d_0 = 2$ *of* [1, Proposition 2.5]*, namely* $(2, 4, 5, 4)$*,* $(2, 3, 3)$*,* $(2, 3, 4, 4)$*,* $(2, 3, 4, 5, 3)$*,* $(2, 3, 4, 5, 4)$*, at ambient dimensions* $20, 9, 20, 27, 36$ *with minors* $98, 71, 17, 2, 5$.

That these twenty-three architectures are filling is asserted in [1]; what Proposition 3.2 adds is a *reproducible witness*. [1] publishes weight matrices for exactly one architecture (its Appendix C), its search draws random weights per run and stores none, and its Proposition 2.5 is introduced by the sentence "Its proof follows exactly the same computational methods as Lemma 2.2" with no data for any of its seven architectures. The certificates were found by drawing weights uniformly from a small range of positive integers — $\{1, 2, 3\}$ for every architecture above and for that of Theorem 1.2, $\{1, \ldots, 9\}$ for the depth-9 one of Theorem 1.3 — and keeping the first draw of full rank modulo 101; the search is a heuristic, the certificate it returns is a proof. The two architectures of [1, Proposition 2.5] with $d_0 \neq 2$, namely $(3, 6, 4, 1)$ and $(4, 6, 3)$, are outside the binary-form evaluator used here and are not part of the statement.

Proposition 3.3. *The following printed data of* [1] *is reproduced from its definitions by an independent implementation.* (i) *For* $d = (2, 1, 1)$*,* $r = 2$ *at* $\theta = (w_{11}, w_{12}, w_2) = (1, 0, 1)$*,* $\operatorname{Jac}\Psi = \left[\begin{smallmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ *entry by entry, of determinant* 0*, and* $p_\theta = x_1^2$. (ii) *At the weights of* [1, Appendix C] *and its own prime* $p = 101$*,* $\operatorname{Jac}\Psi_{(2,3,4,5,4,6,4,1),2}$ *has a* 65×65 *minor equal to* 75*, hence rank* $65 = A_2(d)$*: its Lemma 2.2.* (iii) *The "Reported Dimension" column* $35, 60, 62, 39, 61, 59$ *of* [1, Table 1] *is the Jacobian rank at integer points for the six maximal subarchitectures of the counterexample.*

Proposition 3.3 is a reproduction, not a new result; it is stated because it is what licenses reading our conventions — the coefficient indexing of a binary form, the ordering of parameter coordinates — as the source's own. As a further cross-check, five of the 83 subarchitectures of Theorem 1.4 occur in the search data of [2], computed there by a different method over a different prime, and all five values agree.

Version 2 adds the depth-8 certificates.

Proposition 3.4. *Each of the* 111 *depth-*8 *architectures* $(2, d_1, \ldots, d_7, 1)$ *that the search output of* [2] *flags as minimal filling carries a filling certificate with* $p = 101$ *and all weight entries in* $\{1, 2, 3\}$*: a* 129×129 *minor of* $\operatorname{Jac}\Psi_{d,2}(\theta)$ *at an explicit integer point* θ*, nonzero modulo* 101*. The* 111 *minors are the second column of Tables 3 and 4. Euler's identity of Proposition 9.2(i) holds layer by layer, at degree* 128*, for four of the* 111 *certificates in the formal development (the first three and the last), and for all* 111 *in the independent implementation.*

That these 111 architectures are filling is what the search output asserts: it records a computed dimension equal to the ambient dimension for each of them, obtained from the backpropagation routine of [4] at random weights over GF(100003), and it stores no weights. By the source's own Corollary A.6 such a run is evidence; the certificates are the proof, reproducible from the integer points printed in the formal development. They were found as before, by drawing weights from $\{1, 2, 3\}$ and keeping the first draw of full rank modulo 101; the first draw succeeded for 75 of the 111 architectures and the worst case needed 84 draws. The Jacobians at depth 8 are $129 \times \sum_i d_i d_{i-1}$ integer matrices with entries computed modulo 101, the parameter counts ranging from 168 to 211.

4. The dimension calculus and its soundness

4.1. **The five inputs.** Let r be fixed and let $d \mapsto \dim V_{d,r}$ be the dimension function. We use exactly the following five upper bounds.

- (D1) *Ambient.* $\dim V_{d,r} \leq A_r(d)$, with A_r as in (3) [1, Definition 1.3].
- (D2) *Expected dimension.* $\dim V_{d,r} \leq E(d)$ with E as in (4): the source's naive parameter-count bound.
- (D3) *Depth two, $r = 2$.* $\dim V_{(d_0,d_1,1),2} = d_0 d_1 - \binom{d_1}{2}$ for $d_1 \leq d_0$ and $\binom{d_0+1}{2}$ otherwise — [1, Theorem B.4], quoting [7]; at $r = 2$ this is the classical dimension of the variety of symmetric $d_0 \times d_0$ matrices of rank at most d_1.

(D4) *Subspace bound.* With $N = \binom{d_0 + r^{L-1} - 1}{r^{L-1}}$ and $m = d_{L-1} \le N$,

$$\dim V_{d,r} \;\le\; m(N-m) + d_L\, m.$$

(D5) *Recursive split.* For $1 \le k \le L-1$,

$$\dim V_{(d_0,\dots,d_L),r} \;\le\; \dim V_{(d_0,\dots,d_k),r} + \dim V_{(d_k,\dots,d_L),r} - d_k$$

— [1, Lemma 2.4], which is [3, Proposition 17].

Of these, (D1), (D2), (D3), (D5) are the source's own and are used here as cited. Only (D4) needs a word. The d_L output forms are linear combinations of the d_{L-1} forms $\sigma(u_{L-1})_1, \dots, \sigma(u_{L-1})_{d_{L-1}}$, so they all lie in a single subspace of dimension at most $m = d_{L-1}$ of the N-dimensional space $\mathrm{Sym}_{r^{L-1}}(\mathbb{R}^{d_0})$. Hence $\mathcal{F}_{d,r}$ is contained in the image of the projection of the incidence variety $\{(W,q) : W \in \mathrm{Gr}(m,N),\ q \in W^{d_L}\}$, whose dimension is $\dim \mathrm{Gr}(m,N) + d_L m = m(N-m) + d_L m$. This is a Grassmann secant count and is not new; the same construction appears as the intermediate Grassmann condition [5, (4.7)–(4.8)] in the same setting, and the depth-2 case is the count in Remark 4.4 of [6] (Section 8). What it buys here is self-containedness: it supplies the two dimension inputs [1, Appendix B] quotes from [6] and [5], namely $\dim V_{(3,5,6),2} \le 5 \cdot 1 + 6 \cdot 5 = 35$ (its Theorem B.6) and $\dim V_{(2,d_1,4,5),2} \le 4 \cdot 1 + 5 \cdot 4 = 24$ for $d_1 \ge 3$ (its citation of [5]).

4.2. **The calculus.** Define, for r fixed,

$$B_r(d) \;:=\; \min\Big(A_r(d),\ E(d),\ \mathrm{AH}_2(d),\ S_r(d)\Big), \tag{6}$$

$$U_r^{(0)}(d) \;:=\; B_r(d),$$

$$U_r^{(f+1)}(d) \;:=\; \min\Big(B_r(d),\ \min_{1\le k\le L-1}\big[\,U_r^{(f)}(d_0,\dots,d_k) + U_r^{(f)}(d_k,\dots,d_L) - d_k\,\big]\Big), \tag{7}$$

where AH_2 and S_r are the right-hand sides of (D3) and (D4) and are omitted from the minimum when their hypotheses do not apply. So $U_r^{(f)}$ folds the four non-recursive bounds over every contiguous sub-architecture reached by at most f nested applications of the split (D5).

Proposition 4.1. *Let $\delta : \{architectures\} \to \mathbb{N}$ satisfy* (D1)–(D5)*. Then $\delta(d) \le U_r^{(f)}(d)$ for every f and every d.*

Proof. Induction on f. At $f = 0$ the four non-recursive hypotheses give $\delta(d) \le B_r(d)$ directly. For the step, each term of the inner minimum in (7) is bounded below by the corresponding instance of (D5) after replacing the two summands by their inductive bounds; a lower bound survives a fold of minima; and the split term is guarded so that it can never fall below $B_r(d)$ by an underflow. □

Proposition 4.1 is bookkeeping, but it is the bookkeeping that makes the finite arithmetic of the next section into a statement about neurovarieties. Three facts about the hypotheses are recorded: they are consistent ($\delta \equiv 0$ satisfies all five), the conclusion is not vacuous ($U_2^{(2)}(2,2,4,5,4,6,4,1) = 35$ against an ambient 65), and U cannot be lowered in general, since on three subarchitectures of the counterexample the bound is attained (Section 9).

In the natural numbers the subtraction in (D5) and (7) is truncated. For the true dimension function the right-hand side of (D5) is nonnegative, so the truncated and ordinary forms agree there; in (7) the split term is discarded in favour of $B_r(d)$ whenever d_k exceeds the sum, so truncation can never lower U.

5. Minimality: the whole table

Proof of Theorem 1.1. The eighteen architectures of [1, Table 2] have depths $2, 3, 4, 5, 6$ and 7, so they carry $1 + 2 + 3 + 4 + 5 + 13 \cdot 6 = 93$ maximal subarchitectures, 83 of them distinct. For each, $U_2^{(2)}$ is evaluated and compared with the ambient dimension; the values are those of Table 1, and every one is strictly smaller. By Proposition 4.1, $\dim V_{s,2} \le U_2^{(2)}(s) < A_2(s)$, so s is not filling. With (M) every subarchitecture of a Table-2 entry is then non-filling, and with Proposition 3.2 each entry is filling, so each is an MFA. □

L	architecture d	$A_2(d)$	bounds $U_2^{(2)}(s)$, $s \in \mathcal{M}(d)$	minor
2	$(2,2,1)$	3	2	74
3	$(2,2,2,1)$	5	2, 3	45
4	$(2,3,3,2,1)$	9	7, 7, 5	38
5	$(2,3,3,3,2,1)$	17	11, 9, 11, 9	4
6	$(2,3,3,4,4,2,1)$	33	19, 13, 24, 26, 17	5
7	$(2,3,4,5,4,6,4,1)$	65	35, 61, 63, 39, 61, 62	89
7	$(2,3,4,5,6,4,2,1)$	65	35, 62, 59, 62, 53, 33	26
7	$(2,3,3,4,5,6,3,1)$	65	35, 21, 56, 56, 61, 59	14
7	$(2,3,4,5,5,5,2,1)$	65	35, 60, 60, 52, 62, 33	19
7	$(2,3,3,5,6,4,4,1)$	65	35, 21, 60, 57, 54, 64	90
7	$(2,3,3,5,7,4,2,1)$	65	35, 21, 57, 62, 53, 33	18
7	$(2,3,3,4,6,5,2,1)$	65	35, 21, 42, 58, 57, 33	76
7	$(2,3,4,5,5,4,4,1)$	65	35, 57, 57, 54, 54, 64	48
7	$(2,3,4,4,5,5,4,1)$	65	35, 63, 51, 57, 57, 63	14
7	$(2,3,3,6,6,4,3,1)$	65	35, 21, 64, 56, 54, 63	75
7	$(2,3,3,5,5,5,4,1)$	65	35, 21, 63, 55, 57, 63	52
7	$(2,3,4,6,5,4,3,1)$	65	35, 56, 64, 53, 54, 63	65
7	$(2,3,5,5,5,4,3,1)$	65	35, 64, 56, 53, 54, 63	16

TABLE 1. The eighteen architectures of [1, Table 2]. Column 3 is the ambient dimension (3); column 4 lists $U_2^{(2)}(s)$ for the maximal subarchitectures s in the order $i = 1, \dots, L-1$, every value strictly below column 3 (Theorem 1.1); column 5 is the maximal minor modulo 101 of the filling certificate (Proposition 3.2).

The computation is exhaustive over a finite, explicitly given set: 93 evaluations of (7) at $f = 2$, each unfolding into at most 73 evaluations of B_2 (at depth 7; fewer below), and 93 comparisons. Nothing is sampled and nothing is random. Two things about the fuel are worth stating because they are what makes the sweep both correct and cheap. At $f = 1$ the sweep fails — one level of the recursive bound is not enough. At $f = 2$ the recursion has already reached its fixed point. The formal development checks that at $f = 4$ on the counterexample's subarchitectures (Section 9); outside it we have checked that $f = 2$ and $f = 12$ agree on all 83 subarchitectures, on the $r = 3$ counterexample of Theorem 1.2 and on the depth-9 one of Theorem 1.3. Nothing below depends on that wider check: the bounds of Table 1 are the $f = 2$ values, and $U_2^{(f)}$ is an upper bound at every f.

The subspace bound (D4) is not decoration: five of the eighteen entries fail without it.

Remark 5.1. Exactly two of the 83 subarchitectures are left at 65, the ambient dimension, when the minimum in (6) is taken over (D1), (D2), (D3) only, the recursion (D5) still being applied:

$$(2,3,4,5,5,4,3,1), \quad \text{shared by } (2,3,4,5,5,4,4,1),\ (2,3,4,6,5,4,3,1),\ (2,3,5,5,5,4,3,1);$$
$$(2,3,3,5,6,4,3,1), \quad \text{shared by } (2,3,3,5,6,4,4,1) \text{ and } (2,3,3,6,6,4,3,1).$$

With (D4) both are bounded by 64 — two of the entries of Table 1, and so part of the verified sweep — and all five architectures are certified. The value 65 without (D4) is a computation outside the formal development; what is verified there is the value 64 with it, and the strict improvement (D4) makes at the depth-2 and depth-3 instances [1, Appendix B] had to cite (Section 9).

Finally, the source's own minimality lemma is recovered and slightly improved.

Proposition 5.2. *For the counterexample* $d = (2,3,4,5,4,6,4,1)$*, the six maximal subarchitectures are exactly those of* [1, Table 1]*, and* $U_2^{(2)}$ *gives* $35, 61, 63, 39, 61, 62$ *against the bounds* $\le 53, \le 61, \le 63, = 39, \le 62, \le 62$ *proved in* [1, Lemma 2.6] *— an improvement on the first and the fifth. On* s_1*,* s_4*,* s_5 *the bound equals the dimension* [1, Table 1] *reports, namely* $35, 39, 61$.

Proposition 5.2 is the source's own argument run to exhaustion rather than by hand, with (D2) and (D3) added; the two better constants follow immediately and no new ingredient is involved. Its last sentence is what Theorem 1.4 generalises: where the upper bound meets a certified rank, the dimension is pinned.

6. The two further counterexamples

Proof of Theorem 1.2. The ambient dimension is $1 \cdot \binom{2+3^5-1}{3^5} = \binom{244}{243} = 244$. A filling certificate at $p = 101$ with weights in $\{1, 2, 3\}$ gives a 244×244 minor equal to 56, so by Lemma 3.1 and [1, Theorem A.2] the architecture is filling. Its five maximal subarchitectures evaluate to $U_3^{(2)} = 84, 240, 225, 235, 233$, all below 244; Proposition 4.1 and (M) finish it. □

[1, Example 4.3] states only "We successfully verified it is a minimal filling architecture using the same techniques as above", and prints the dimensions $84, 240, 223, 233, 231$ for the five subarchitectures without a verification; our bounds $84, 240, 225, 235, 233$ are proved upper bounds and agree on the first two. Neither this architecture nor any of its maximal subarchitectures occurs in the $r = 3$ search data of [2].

Proof of Theorem 1.3. The ambient dimension is $2\binom{2+2^8-1}{2^8} = 2 \cdot 257 = 514$ and the parameter count is 619. A filling certificate at $p = 101$ gives a 514×514 minor equal to 54. Of the eight maximal subarchitectures, seven satisfy $U_2^{(2)} < 514$, with $514 - U_2^{(2)}$ — a lower bound for the codimension — equal to $254, 176, 5, 17, 34, 11, 124$ in their order; these match seven of the eight codimensions $254, 17, 176, 5, 17, 34, 11, 124$ [1, Example 4.2] reports. For $(2, 3, 3, 4, 10, 17, 11, 12, 4, 2)$ the calculus returns 514 and proves nothing. □

The one gap is a mathematical gap, not a computational one: the sweep takes seconds, and the missing ingredient is a bound the present five inputs do not give. The natural candidate is a subspace bound applied at an interior layer rather than only at the last, which would replace (D4) by a family of Grassmann terms indexed by the cut. [1, Example 4.2] reports 497 for this subarchitecture, so a bound of 513 would suffice. That candidate is carried out in Section 11, and it does not suffice: with the interior-layer bound (D6) added to the calculus, this subarchitecture is still bounded only by 514 (Proposition 11.3).

7. Exact neurovariety dimensions

Proof of Theorem 1.4. For each of the 83 distinct maximal subarchitectures s, an explicit integer point θ_s is given and $\operatorname{rank}(\operatorname{Jac}\Psi_{s,2}(\theta_s) \bmod 101)$ computed; by Lemma 3.1 and [1, Theorem A.2] this is a lower bound for $\dim V_{s,2}$, and by Proposition 4.1 $U_2^{(2)}(s)$ is an upper bound. The two are compared in all 83 cases: the lower bound never exceeds the upper one, and they coincide in 71. □

Every one of the 71 equalities is a proved statement, not a measurement: a kernel-checked nonzero minor below, and Proposition 4.1 above. The overlap with the record was measured rather than assumed. Six of the 83 appear in [1, Table 1] and five in the search data of [2], with $(2, 1, 1)$ in the latter and in the worked example of [1, Example A.8]; of these eleven architectures, eight have their dimension determined here and all eight values agree with the reported ones — over a different prime, by an independent implementation. Three of the six values [1, Table 1] reports fall in Table 2, which is to say that they remain reported and unproved.

8. A correction to a quoted defectivity criterion

Proof of Proposition 1.5. At $r = 2$ write $N = \binom{d_0+1}{2}$. For $d_0 = 2$, $N = 3$ and the printed hypothesis $N - d_0 < d_1 < \min(d_2, N)$ reads $1 < d_1 < \min(d_2, 3)$, i.e. $d_1 = 2$ and $d_2 \geq 3$. For $(2, 2, 3)$, $(2, 2, 4)$, $(2, 2, 5)$ the expected dimensions (4) are $\min(9, 8) = 8$, $\min(12, 10) = 10$, $\min(15, 12) = 12$, and explicit integer points give Jacobian ranks $8, 10, 12$ modulo 101. A rank is a lower bound for the dimension and the expected dimension is an upper bound, so the dimension equals the expected dimension and the defect is 0 in all three cases.

	subarchitecture s	rank	$U_2^{(2)}(s)$	gap
1	$(2,3,3,3,4,2,1)$	23	24	1
2	$(2,3,3,3,5,6,3,1)$	54	56	2
3	$(2,3,3,3,6,5,2,1)$	41	42	1
4	$(2,3,3,4,4,6,3,1)$	53	56	3
5	$(2,3,3,4,5,6,2,1)$	58	59	1
6	$(2,3,3,5,4,6,4,1)$	60	61	1
7	$(2,3,3,5,6,3,4,1)$	51	54	3
8	$(2,3,3,5,7,3,2,1)$	50	53	3
9	$(2,3,3,6,6,3,3,1)$	51	54	3
10	$(2,3,4,4,4,6,4,1)$	62	63	1
11	$(2,3,4,5,4,5,2,1)$	51	52	1
12	$(2,3,4,5,4,6,3,1)$	59	62	3

TABLE 2. The 12 of the 83 distinct maximal subarchitectures where the certified rank and the calculus bound do not meet; the other 71 dimensions are determined exactly (Theorem 1.4). Entries 6, 10 and 12 are three of the six maximal subarchitectures of the counterexample $(2,3,4,5,4,6,4,1)$; [1, Table 1] reports dimensions for them, and those values remain unproved.

For the corrected region, at depth 2 the bound (D4) reads $S = d_1(N-d_1)+d_2d_1$. Then $S < E(d) = d_0d_1 + d_1d_2 - d_1$ if and only if $N - d_1 < d_0 - 1$, i.e. $d_1 > N - d_0 + 1$; and $S < A_2(d) = d_2N$ if and only if $d_1 < d_2$ (given $d_1 < N$). So (D4) certifies $\dim V_{d,2} < \operatorname{expdim}(V_{d,2})$ exactly on $N - d_0 + 1 < d_1 < \min(d_2, N)$. At $d_0 = 2$ that region is empty, consistently with the three counterexamples; at $(3,5,6)$ it gives $S = 35 < 36 = \operatorname{expdim}$. $\square$

The three counterexamples sit exactly on the step $d_1 = N - d_0 + 1$ by which the printed hypothesis exceeds the corrected one. The same failure occurs at $d_0 = 3$, where the printed criterion admits $(3,4,5)$ whose generic Jacobian rank is 28, its expected dimension; that computation was performed outside the formal development, whose evaluator is specialised to $d_0 = 2$, and is reported here only as a remark.

We stress what the proposition is about: the criterion *as printed in* [1], and not the source [6] from which [1] quotes it. Indeed [6] states the correct endpoint. A depth-2 network at $r = 2$ has output $\left(\sum_{j\le d_1}(W_2)_{ij}\ell_j^2\right)_{i\le d_2}$ with ℓ_j linear in d_0 variables, i.e. $\sum_j c_j \otimes \ell_j^2$ with $c_j \in \mathbb{R}^{d_2}$: so $V_{(d_0,d_1,d_2),2}$ is the affine cone over the s-th secant variety $\sigma_s(X_{m,n})$ of the Segre–Veronese $X_{m,n} = \mathbb{P}^m \times \mathbb{P}^n$ embedded by $\mathcal{O}(1,2)$, where

$$m = d_2 - 1, \qquad n = d_0 - 1, \qquad s = d_1, \qquad d = 2,$$

and under this dictionary the source's expdim and ambient dimension are the affine forms $s(m+n+1)$ and $(m+1)\binom{n+d}{d}$ of the ones in [6]. Remark 4.4 of [6] records, for $(m,n;1,d)$ unbalanced — and credits to earlier work — that $\sigma_s(X_{m,n})$ fails to have the expected dimension if and only if

$$\binom{n+d}{d} - n \;<\; s \;<\; \min\Big\{m+1,\ \binom{n+d}{d}\Big\},$$

which is exactly the corrected region above: $\binom{n+2}{2} - n = \binom{d_0+1}{2} - d_0 + 1$. What [1] prints has $-\,d_0$ in place of $-\,n = -(d_0-1)$, one step too wide, which is the projective-to-affine slip the translation invites. The bound [6] uses there, $s\big(\binom{n+d}{d} + m + 1 - s\big)$, is in the same dictionary $d_1(N-d_1) + d_2d_1$ — the depth-2 case of (D4). Nothing in the source's own argument depends on the difference: $(3,5,6)$, its only application of the criterion, lies in the corrected region.

9. Controls

A calculus of upper bounds and a family of certificates can both fail silently, so the following negative controls were run. They are not results; they are the reasons to believe the results.

Proposition 9.1. (i) *No entry of* [1, Table 2] *is certified non-filling by the calculus — as it must not be, since all eighteen carry filling certificates.* (ii) *The constant-width* $(2,2,2,2,2,2,2,1)$ *is certified non-filling,* $U_2^{(2)} = 13 < 65$*, and so is the source's own* $(2,1,1)$*,* $U_2^{(2)} = 2 < 3$*.* (iii) $(2,4,4,5,4,6,4,1)$ *is* not *certified minimal, correctly, since its maximal subarchitecture* $(2,3,4,5,4,6,4,1)$ *is itself filling.* (iv) $f = 1$ *does not suffice for the sweep while* $f = 2$ *and* $f = 4$ *agree on the counterexample's subarchitectures.* (v) *The subspace bound is strictly stronger than (D1)–(D2) at the two instances* [1, Appendix B] *takes from the literature —* $S_2(2,3,4,5) = 24 < 25$ *and* $S_2(3,5,6) = 35 < 36$ *— and* $U_2^{(2)}(2,3,4,5,5,4,3,1) = 64$*, the value Remark 5.1 turns on.*

Proposition 9.2. (i) *Euler's identity* $\sum_{a,b}(W_i)_{ab}\,\partial\Psi/\partial(W_i)_{ab} = r^{L-i}\Psi$ *holds layer by layer for the Table-2 certificates, the Proposition-2.5 certificates, the source's own point, and the* $r = 3$ *counterexample (not, in the formal development, for the depth-*9 *certificate of Theorem 1.3).* (ii) *The source's own integer point read modulo* 11 *or modulo* 29 *instead of* 101 *has Jacobian rank* $60 < 65$*, although the architecture is filling.* (iii) $(2,4,4,5,4,6,4,1)$ *and its maximal subarchitecture* $(2,3,4,5,4,6,4,1)$ *are both filling.* (iv) $(2,2,2,2,2,2,2,1)$ *has Jacobian rank* 13 *against an ambient* 65*.*

Part (i) of Proposition 9.2 is the one that ties the Jacobian code to the network evaluation: a wrong product rule, a mis-indexed layer or a dropped term each break the identity. Part (ii) is the reason the minimality half needs a calculus at all, and it is the exact phenomenon behind the source's word "suggest": deficiency modulo a prime is not evidence. Part (iii) of Proposition 9.1 is the corresponding statement on the other side: a filling certificate alone says nothing about minimality.

One control fired during this work and is worth recording. An earlier version of the calculus used [1, Theorem B.6] as printed in place of (D4). The soundness control "no subarchitecture has certified rank above its calculus bound" then failed on three of the 83, a strict contradiction between a proved lower bound and a claimed upper bound; that is how Proposition 1.5 was found. After the replacement the control passes on all 83 (Theorem 1.4).

10. Depth eight

Throughout this section $d_0 = 2$, $d_L = 1$, $r = 2$ and $L = 8$, so every architecture and every maximal subarchitecture has ambient dimension $\binom{129}{128} = 129$.

10.1. **The candidate set and the printed count.** The 111 architectures were read from the file `data/processed/2_1_cleaned_architectures.csv` of [2]: of its 12,907 rows, 6,639 have depth 8, all distinct, and exactly 111 of those carry both flags `is_minimal` and `is_full_dimension`; 20 of the 111 are not unimodal, and 91 are. Every one of the 111 has recorded dimension equal to 129 and recorded defect 0. The same file flags 13 architectures at depth 7, the thirteen of Table 1.

The sentence of [1] quoted in Section 1 counts 82. We looked for a reading of the published data that gives 82 and found none: restricting the hidden widths to $[1,6]$, $[1,7]$, $[1,8]$, $[1,9]$, $[1,10]$ leaves 1, 37, 91, 108, 111 of the flagged architectures; no threshold on the parameter count, above or below, leaves 82; the unimodal ones number 91 and the non-unimodal 20; and all 111 have defect 0. We also note that the file's depth-8 block is the same in every version of it the repository has held: all seven commits, from 7 May 2026 on, carry the same 6,639 rows with the same flags and recorded dimensions, and the same 111. We therefore report the 82 as a discrepancy between the Discussion of [1] and the artefact its footnote cites, and draw no conclusion about which of the two the authors intend. The 20 non-unimodal, in the same sentence, matches the file exactly.

10.2. **The fuel ladder.** At depth 7 the sweep evaluated $U_2^{(2)}$ on every subarchitecture. At depth 8 there are 777 subarchitectures, 605 distinct, and the recursion in (7) is more expensive by a factor of about three per level, so the sweep is organised as a ladder: a subarchitecture s is certified non-filling as soon as one of $U_2^{(0)}(s)$, $U_2^{(1)}(s)$, $U_2^{(2)}(s)$ is below 129, the cheaper rung first. Since each $U_2^{(f)}$ is an upper bound by Proposition 4.1, any rung that fires proves $\dim V_{s,2} < 129$; the formal development proves this once, as a lemma over the three rungs, and everything below uses it. All three rungs are used: $(2,3,3,3,6,8,6,3,1)$ is settled by $U_2^{(0)} = 127$, $(2,2,3,4,4,7,8,7,1)$ needs one level ($U_2^{(0)} = 129$,

$U_2^{(1)} = 67$) and $(2,3,3,3,4,8,8,6,1)$ needs two (129, 129, 125). Outside the formal development we counted where the 605 are decided: 54 at the base rung, 402 at the first level, 147 at the second, and the remaining 2 — the blocking pair of Proposition 10.1 — at none.

10.3. **The census.**

Proof of Theorem 1.6. The 111 certificates of Proposition 3.4 prove every flagged architecture filling, by Lemma 3.1 and [1, Theorem A.2]. For each of the 107 architectures of Table 3, each of its seven maximal subarchitectures is run through the ladder of Section 10.2 and every one fires; by Proposition 4.1, $\dim V_{s,2} < 129$ for all 777 of them (605 distinct). With (M), every subarchitecture of a listed architecture is non-filling, and with the certificate each listed architecture is an MFA. For the four architectures of Table 4 the ladder fails on exactly one subarchitecture each, and the sweep reports which (Proposition 10.1). □

The computation is exhaustive over an explicitly given finite set — 777 ladder evaluations, none sampled — and is checked by the kernel of the proof assistant alone, in eleven blocks of at most ten architectures (Section 12). Which of the 111 flagged architectures land in $\mathcal{C}_8$ is an output of the sweep, not an input: the census is what survives, and the four of $\mathcal{O}_8$ are what does not.

10.4. **The seventeen counterexamples.**

Proof of Theorem 1.7. Filtering $\mathcal{C}_8$ by the unimodality walk of Section 2 returns exactly the seventeen sequences listed, and each of them is a minimal filling architecture by Theorem 1.6. The walk itself is checked against five test cases: the source's counterexample $(2,3,4,5,4,6,4,1)$ is rejected, its neighbour $(2,3,4,5,6,4,2,1)$ in Table 1 and $(2,2,1)$ are accepted, and $(1,3,2,4)$, $(2,2,1,2)$ are rejected. □

The source's file `counterexamples.csv` lists the same 20 flagged non-unimodal depth-8 architectures without a verification of minimality; the three of them not in Theorem 1.7 are the three non-unimodal entries of Table 4.

10.5. **Exact dimensions at depth eight.**

Proof of Theorem 1.8. As in the proof of Theorem 1.4: for each of the 605 distinct maximal subarchitectures s an explicit integer point θ_s is given, $\operatorname{rank}(\operatorname{Jac}\Psi_{s,2}(\theta_s) \bmod 101)$ is a lower bound for $\dim V_{s,2}$ by Lemma 3.1 and [1, Theorem A.2], and $U_2^{(2)}(s)$ is an upper bound by Proposition 4.1. The pairs are computed in four blocks of 152, 152, 152 and 149; the lower bound never exceeds the upper one in any block, and the numbers of equalities are 152, 130, 122 and 129, which sum to 533. □

The points θ_s were chosen by a search for a good witness, not by a change of method: up to 30 random integer points per subarchitecture, drawn from $\{1,2,3\}$, then $\{1,\dots,5\}$, then $\{1,\dots,9\}$, keeping the point of largest rank. A first pass with six points from $\{1,2,3\}$ gave 416 equalities; a second pass over the 189 subarchitectures then short of their bound raised 167 ranks and lifted the count to 533. Every rank in the theorem is certified at the point that achieves it, and the search that found the point is not part of the proof. The 72 gaps are of the same kind as the twelve of Table 2: an upper bound that is not attained by any integer point tried, or a neurovariety whose dimension the five inputs do not pin.

Two comparisons with the record were made, both outside the formal development. First, four of the 605 occur in the search output of [2]; with the dimension recorded there, our certified rank and our calculus bound, they are

s	recorded	rank	$U_2^{(2)}(s)$
$(2,3,4,7,7,6,5,2,1)$	126	126	126
$(2,3,3,4,7,9,4,2,1)$	128	128	128
$(2,3,4,5,5,7,8,2,1)$	122	122	123
$(2,3,3,4,5,9,6,2,1)$	124	124	124

so all four recorded values agree with our ranks, and three of the four dimensions are among the 533 determined exactly. Second, and more widely: for every one of the 12,907 rows of that file — every depth

architecture d	minor	architecture d	minor	architecture d	minor
$(2,3,3,4,4,7,8,7,1)$	15	$(2,3,3,5,7,8,5,2,1)$	89	$(2,3,4,5,6,7,7,2,1)$	92
$(2,3,3,4,4,8,8,6,1)$	36	$(2,3,3,5,7,9,4,2,1)$	71	$(2,3,4,5,6,8,5,3,1)$	36
$(2,3,3,4,5,6,8,8,1)$	75	$(2,3,3,5,8,6,6,3,1)$	37	$(2,3,4,5,6,8,6,2,1)$	48
$(2,3,3,4,5,7,8,4,1)$	37	$(2,3,3,5,8,7,5,2,1)$	9	$(2,3,4,5,6,9,4,3,1)$	84
$(2,3,3,4,5,8,7,4,1)$	3	$(2,3,3,5,8,8,4,2,1)$	15	$(2,3,4,5,6,10,4,2,1)$	34
$(2,3,3,4,5,8,8,2,1)$	70	$(2,3,3,6,6,5,8,8,1)^\dagger$	93	$(2,3,4,5,7,5,7,5,1)^\dagger$	10
$(2,3,3,4,5,9,6,4,1)$	47	$(2,3,3,6,6,6,7,7,1)$	83	$(2,3,4,5,7,5,8,4,1)^\dagger$	99
$(2,3,3,4,5,9,7,2,1)$	96	$(2,3,3,6,6,8,6,2,1)$	46	$(2,3,4,5,7,6,6,4,1)$	54
$(2,3,3,4,6,6,8,5,1)$	6	$(2,3,3,6,7,6,7,3,1)^\dagger$	82	$(2,3,4,5,7,6,7,3,1)^\dagger$	19
$(2,3,3,4,6,7,7,4,1)$	84	$(2,3,3,6,7,7,6,2,1)$	96	$(2,3,4,5,7,7,5,3,1)$	53
$(2,3,3,4,6,7,8,3,1)$	77	$(2,3,3,6,8,7,4,4,1)$	30	$(2,3,4,5,7,7,6,2,1)$	36
$(2,3,3,4,6,8,6,3,1)$	68	$(2,3,4,4,4,7,8,6,1)$	4	$(2,3,4,5,7,8,4,2,1)$	25
$(2,3,3,4,6,8,7,2,1)$	53	$(2,3,4,4,4,8,8,5,1)$	35	$(2,3,4,5,8,5,7,4,1)^\dagger$	97
$(2,3,3,4,6,9,5,3,1)$	43	$(2,3,4,4,5,6,8,7,1)$	29	$(2,3,4,5,8,6,5,3,1)$	55
$(2,3,3,4,6,9,6,2,1)$	90	$(2,3,4,4,5,7,7,7,1)$	12	$(2,3,4,5,8,6,6,2,1)$	73
$(2,3,3,4,6,10,5,2,1)$	43	$(2,3,4,4,5,8,7,3,1)$	100	$(2,3,4,5,8,7,4,2,1)$	28
$(2,3,3,4,7,6,7,6,1)^\dagger$	67	$(2,3,4,4,5,9,6,3,1)$	13	$(2,3,4,5,9,6,5,2,1)$	99
$(2,3,3,4,7,7,6,4,1)$	61	$(2,3,4,4,6,6,8,4,1)$	46	$(2,3,4,6,5,8,6,3,1)^\dagger$	46
$(2,3,3,4,7,7,7,2,1)$	7	$(2,3,4,4,6,7,7,3,1)$	31	$(2,3,4,6,6,5,7,7,1)^\dagger$	84
$(2,3,3,4,7,8,5,3,1)$	50	$(2,3,4,4,7,5,8,8,1)^\dagger$	45	$(2,3,4,6,6,6,6,6,1)$	25
$(2,3,3,4,7,8,6,2,1)$	68	$(2,3,4,4,7,6,7,5,1)^\dagger$	13	$(2,3,4,6,6,6,7,3,1)$	59
$(2,3,3,4,7,9,4,3,1)$	84	$(2,3,4,4,7,7,6,3,1)$	84	$(2,3,4,6,6,7,6,2,1)$	73
$(2,3,3,4,7,9,5,2,1)$	56	$(2,3,4,4,7,8,5,2,1)$	3	$(2,3,4,6,6,8,5,2,1)$	70
$(2,3,3,4,7,10,4,2,1)$	69	$(2,3,4,4,7,9,4,2,1)$	69	$(2,3,4,6,6,9,4,2,1)$	30
$(2,3,3,4,8,8,5,2,1)$	96	$(2,3,4,5,4,7,8,5,1)^\dagger$	25	$(2,3,4,6,7,5,7,4,1)^\dagger$	23
$(2,3,3,4,8,9,4,2,1)$	55	$(2,3,4,5,4,8,8,4,1)^\dagger$	87	$(2,3,4,6,7,6,5,5,1)$	78
$(2,3,3,5,5,6,8,7,1)$	22	$(2,3,4,5,5,6,8,4,1)$	31	$(2,3,4,6,7,6,6,2,1)$	63
$(2,3,3,5,5,7,7,7,1)$	16	$(2,3,4,5,5,7,7,4,1)$	15	$(2,3,4,6,7,7,4,4,1)$	75
$(2,3,3,5,5,8,7,3,1)$	62	$(2,3,4,5,5,7,8,3,1)$	15	$(2,3,4,6,7,7,5,2,1)$	99
$(2,3,3,5,5,9,6,3,1)$	79	$(2,3,4,5,5,8,6,4,1)$	61	$(2,3,4,6,8,6,5,2,1)$	78
$(2,3,3,5,6,6,8,4,1)$	40	$(2,3,4,5,5,8,7,2,1)$	2	$(2,3,4,7,7,6,5,3,1)$	33
$(2,3,3,5,6,7,7,3,1)$	5	$(2,3,4,5,5,9,6,2,1)$	4	$(2,3,4,7,7,7,4,2,1)$	98
$(2,3,3,5,6,9,5,2,1)$	81	$(2,3,4,5,6,5,8,5,1)^\dagger$	17	$(2,3,5,5,5,8,6,3,1)$	43
$(2,3,3,5,7,5,8,5,1)^\dagger$	90	$(2,3,4,5,6,6,7,4,1)$	11	$(2,3,5,5,7,6,6,3,1)$	42
$(2,3,3,5,7,6,7,4,1)^\dagger$	85	$(2,3,4,5,6,6,8,3,1)$	28	$(2,3,5,5,7,7,5,2,1)$	47
$(2,3,3,5,7,7,6,3,1)$	79	$(2,3,4,5,6,7,6,3,1)$	26		

TABLE 3. The census $\mathcal{C}_8$: the 107 depth-8 architectures with $d_0 = 2$, $d_L = 1$, $r = 2$ proved to be minimal filling architectures (Theorem 1.6), in the order of the formal development, each with the 129×129 minor modulo 101 of its filling certificate (Proposition 3.4). A dagger marks the seventeen non-unimodal entries of Theorem 1.7.

architecture d	minor	blocking $s \in \mathcal{M}(d)$	$U_2^{(2)}(s)$	rank
$(2,3,3,4,7,6,8,4,1)^\dagger$	23	$(2,3,3,4,7,6,8,3,1)$	129	127
$(2,3,3,5,7,6,8,3,1)^\dagger$	75	$(2,3,3,4,7,6,8,3,1)$	129	127
$(2,3,3,5,9,7,4,3,1)$	63	$(2,3,3,5,9,7,4,2,1)$	129	128
$(2,3,4,4,7,6,8,3,1)^\dagger$	77	$(2,3,3,4,7,6,8,3,1)$	129	127

TABLE 4. The four flagged depth-8 architectures $\mathcal{O}_8$ that Theorem 1.6 does not decide: each is filling (column 2 is the minor of its certificate), and each has exactly one maximal subarchitecture s that the calculus leaves at the ambient dimension 129; only two such s occur. The last column is the certified Jacobian rank of s at an explicit integer point, a lower bound for $\dim V_{s,2}$ and not an upper one (Proposition 10.1). A dagger marks the non-unimodal entries.

from 2 to 8, filling or not — we evaluated $U_2^{(2)}$ and compared it with the dimension the file records from its own random-point computation over GF(100003). The bound is never below the recorded dimension, and equals it on 11,779 rows. Two implementations, two primes, two methods, and no disagreement anywhere; a single row the other way would have refuted one of (D1)–(D5) or the code. The comparison was re-run independently by the author of this version with the same three numbers.

10.6. **The gap, named, and the depth-eight controls.**

Proposition 10.1. (i) *The four architectures of Table 4 are not certified minimal, and for each the ladder of Section 10.2 fails on exactly one maximal subarchitecture:*

$$(2,3,3,4,7,6,8,3,1) \quad \textit{for } (2,3,3,4,7,6,8,4,1),\ (2,3,3,5,7,6,8,3,1),\ (2,3,4,4,7,6,8,3,1);$$
$$(2,3,3,5,9,7,4,2,1) \quad \textit{for } (2,3,3,5,9,7,4,3,1).$$

(ii) *The ladder is sound: for any δ satisfying* (D1)–(D5)*, a rung that fires on s gives $\delta(s) < A_r(s)$.* (iii) *The constant-width* $(2,2,2,2,2,2,2,2,1)$ *is certified non-filling with $U_2^{(2)} = 15 < 129$, and the width-one tower* $(2,1,1,1,1,1,1,1,1)$ *with $U_2^{(2)} = 2$.* (iv) $(2,4,3,4,4,7,8,7,1)$ *is* not *certified minimal, correctly, since* $(2,3,3,4,4,7,8,7,1) \in \mathcal{C}_8$ *is one of its maximal subarchitectures.* (v) *The three rung witnesses of Section 10.2 hold as stated.* (vi) $U_2^{(1)}(2,2,3,4,4,7,8,7,1) = 67$ *against an ambient* 129*, and* $|\mathcal{C}_8| = 107$*.* (vii) $\delta \equiv 0$ *satisfies* (D1)–(D3)*,* (D5) *and the weaker form* (D4′) *of Section 11, so the census theorem has no content without the filling certificates.*

Part (i) names the gap. The two blocking subarchitectures have certified Jacobian ranks 127 and 128 at explicit integer points (Theorem 1.8), two and one short of the ambient dimension, at every point tried — strong evidence that neither is filling, and no proof, for the reason of Proposition 9.2(ii): a deficient rank modulo a prime is not an upper bound. More fuel does not help either; outside the formal development, $U_2^{(f)}$ is 129 on both for $f = 1, \ldots, 6$. Closing them needs a bound of a different kind, one that sees defectivity rather than a parameter count; the only such input in the calculus is (D3), which lives at depth 2 (see also Proposition 11.3). If both are non-filling, as the ranks suggest, all four architectures of Table 4 are minimal filling architectures and the depth-8 census within the flagged set is 111 with 20 non-unimodal, which is the source's second number exactly. Parts (iii)–(vii) are the depth-8 instances of the controls of Section 9.

11. The subspace bound revisited

11.1. **It is load-bearing at depth eight.** Remark 5.1 showed that five of the eighteen depth-7 entries need (D4). At depth 8 the dependence is much heavier.

Proposition 11.1. (i) *With the minimum in* (6) *taken over* (D1)*,* (D2)*,* (D3) *only, the recursion* (D5) *still being applied through the ladder,* 30 *of the* 107 *architectures of $\mathcal{C}_8$ lose their certificate:* 0, 1, 0, 4, 6, 2, 1, 3, 5, 2, 6 *in the eleven blocks of Table 3. For instance* $(2,3,3,4,7,5,8,8,1)$ *and* $(2,3,3,5,8,6,6,2,1)$ *are left at* 129 *without* (D4) *and bounded by* 127 *and* 126 *with it.* (ii) *Let* (D4′) *be the bound*

$$\dim V_{d,r} \ \le\ S'_r(d) \ :=\ \max_{1\le t\le m}\ t(N-t) + d_L\, t, \qquad m = d_{L-1} \le N,$$

and let $U_r'^{(f)}$ be the calculus (6)–(7) *with S'_r in place of S_r. Then $\delta(d) \le U_r'^{(f)}(d)$ for every f and d whenever δ satisfies* (D1)–(D3)*,* (D4′)*,* (D5)*; and under these hypotheses both censuses hold: every maximal subarchitecture of every entry of* [1, Table 2]*, and of every architecture of $\mathcal{C}_8$, has dimension strictly below its ambient dimension.* (iii) $S_2(2,5,8,8,4) = 40$ *while* $S'_2(2,5,8,8,4) = 42$*; and $S_2 = S'_2$ at* $(3,5,6)$ *(*35*), at* $(2,3,4,5)$ *(*24*), and* $U_2^{(2)} = U_2'^{(2)} = 64$ *at* $(2,3,4,5,5,4,3,1)$*.*

Proof. (i) and (iii) are evaluations. For (ii), the proof of Proposition 4.1 goes through verbatim with S'_r in place of S_r, since it uses (D4) only through the inequality $\delta(d) \le S_r(d)$ at the base rung; the censuses are the sweeps of Theorem 1.1 and Theorem 1.6 re-run with U', the ladder of Section 10.2 included, and every evaluation lands strictly below the ambient dimension again. ☐

Outside the formal development we recorded two more numbers: the 30 lost certificates are blocked by 17 distinct maximal subarchitectures, and U' returns *the same value* as U on every one of the 605 depth-8 and 93 depth-7 subarchitectures, so the two calculi do not merely agree on the verdicts but on the bounds.

Remark 11.2. Where (D4′) comes from, and why the censuses were re-run under it. The incidence count of Section 4 can be stratified by the exact dimension t of the span of $\sigma(u_{L-1})_1, \ldots, \sigma(u_{L-1})_m$: the stratum of span dimension t lies in the projection of $\{(W, q) : W \in \mathrm{Gr}(t, N),\ q \in W^{d_L}\}$, of dimension $f(t) := t(N - t) + d_L t$, and the maximum of f over $t \le m$ is (D4′). Since $f(t) = t(N + d_L - t)$ peaks at $t = (N + d_L)/2$, the value $f(m)$ of (D4) is that maximum only when $2m \le N + d_L + 1$, and this side condition does fail: at $(2, 5, 8, 8, 4)$ one has $N = 9$, $d_L = 4$, $m = 8$ and $f(8) = 40 < 42 = f(6)$, and outside the formal development we found $S_2 < S_2'$ on 18 of the 7,933 architectures the depth-8 calculus visits. The stratified reading therefore justifies only (D4′). It is not, however, the reading of Section 4, which does not stratify: a span of dimension $t \le m \le N$ lies in some subspace of dimension exactly m, so the whole functional space lies in the projection of the single incidence variety with $W \in \mathrm{Gr}(m, N)$; that projection is closed, because $\mathrm{Gr}(m, N)$ is complete, and has dimension at most $f(m)$, so the closure $V_{d,r}$ lies in it as well. Hence (D4) is justified as it stands, and (D4′), which is never smaller than (D4), is a *weaker* valid bound, not a repair of the statement. What Proposition 11.1(ii) adds is robustness: neither census depends on the sharper reading, and both survive under the weakest hypothesis that either derivation supports. Nothing in version 1 is withdrawn.

11.2. **The interior-layer bound, and that it does not help.** Section 6 named the obvious candidate for the one subarchitecture of Theorem 1.3 that the calculus does not reach: a subspace bound applied at an interior layer. Here is that bound.

(D6). Fix $1 \le k \le L - 1$ and let $N_k = \binom{d_0 + r^k - 1}{r^k} = \dim \mathrm{Sym}_{r^k}(\mathbb{R}^{d_0})$, the dimension of the space in which the layer-k activations $\sigma(u_k)_1, \ldots, \sigma(u_k)_{d_k}$ live. If the tuple $v = (\sigma(u_k)_j)_j$ is replaced by another basis Av of the same span, the change is absorbed by $W_{k+1} \mapsto W_{k+1}A^{-1}$, so the output of the layers above k depends on v only through $\mathrm{span}(v) \in \mathrm{Gr}(m, N_k)$ with $m \le d_k$; and for a fixed span, substituting fixed forms for the inputs of the top network $(m, d_{k+1}, \ldots, d_L)$ is a linear map on coefficient vectors, so the fibre has dimension at most $\dim V_{(m,d_{k+1},\ldots,d_L),r}$. Hence

$$\dim V_{(d_0,\ldots,d_L),r} \ \le \ \max_{1 \le m \le \min(d_k, N_k)} \left[\, m(N_k - m) + \dim V_{(m,d_{k+1},\ldots,d_L),r} \,\right].$$

At $k = L - 1$ the top network is a single linear layer, $\dim V_{(m,d_L),r} = d_L m$, and (D6) is exactly (D4′); so (D6) generalises the incidence bound, and it is elementary in the same way. Write $U_6^{(f)}$ for the calculus (7) with, at each fuel level, the minimum also taken over (D6) at every interior layer k, the inner dimension evaluated by the calculus at the previous level.

Proposition 11.3. (i) *At the last layer (D6) evaluates to (D4′): the interior bound at $k = 2$ for $(2, 3, 4, 5)$ is* 24 *and at $k = 1$ for $(3, 5, 6)$ is* 35*.* (ii) *(D6) does not close the depth-*8 *gap of Proposition 10.1:*

$$U_6^{(1)}(2, 3, 3, 4, 7, 6, 8, 3, 1) \ = \ U_6^{(1)}(2, 3, 3, 5, 9, 7, 4, 2, 1) \ = \ 129,$$

the ambient dimension. (iii) *Nor the depth-*9 *gap of Theorem 1.3: $U_6^{(1)}(2, 3, 3, 4, 10, 17, 11, 12, 4, 2) = 514$, again the ambient dimension.* (iv) *(D6) is not the trivial bound: at $k = 1$ for $(2, 2, 2, 2, 1)$ it gives* 7 *against an ambient* 9*; and for $(2, 3, 3, 4, 7, 6, 8, 3, 1)$ one has $N_1 = 3$ and $N_5 = 33$.*

Outside the formal development, $U_6^{(2)}$ agrees with $U_2^{(2)}$ on every one of the 605 distinct depth-8 subarchitectures, so at that fuel adding (D6) at every interior layer changes no value of the calculus at all; evaluating the inner dimension in (D6) by the full calculus $U_2^{(2)}$ instead of the previous level lowers four of the 605 values, by at most six, and moves no verdict. The reason is visible in part (iv): $N_k = 2^k + 1$ grows with k, so at the middle layers the Grassmann term $m(N_k - m)$ alone exceeds the ambient dimension ($N_5 = 33$ and $m = 8$ give $200 > 129$), while at $k = 1, 2$ the maximum over m is attained where the top network's own bound is its parameter count, again above the ambient dimension. A remark on the statement of (D6): it is stated with the maximum over m, which is what the formal development evaluates; by the enlargement step of Remark 11.2 the single value at $m = \min(d_k, N_k)$ is also a valid upper bound, never larger than the maximum, and outside the formal development we checked that this sharper form, with the inner dimension evaluated by $U_2^{(2)}$, leaves the two depth-8 blocking subarchitectures and the depth-9 one at 129, 129 and 514 as well. The negative result is recorded so that it need not be re-derived: closing the three open subarchitectures of Proposition 10.1 and Theorem 1.3 needs an input that sees defectivity at depth 3 or 4, not another parameter count.

12. Verification

Every statement above — Lemma 3.1, Propositions 3.2, 3.3, 3.4, 4.1, 5.2, 1.5, 9.1, 9.2, 10.1, 11.1, 11.3 and Theorems 1.1–1.4 and 1.6–1.8 — has been formally verified in Lean 4 [8] against *Mathlib* [9], in twenty-four modules: the eight of version 1 (the evaluator and the calculus, the certificate data, the subarchitecture data, the minimality sweep, the filling certificates, the dimension sweep, the soundness proof, the rank lemma) and sixteen for depth 8 (the two architecture lists and the ladder, the 111 certificates, the sweep in eleven blocks, the certificate checks, the certificate reports and the Euler control, four blocks of subarchitecture points and four of rank-and-bound sweeps, the weaker bound (D4′) with both censuses under it, the necessity of (D4) and the interior bound (D6), and the soundness bridge for the ladder and for (D4′)). Appendix A reproduces the statement of every machine-checked result. One clause is the exception and is flagged as such: of Proposition 1.5, the three architectures that refute the printed criterion are machine-checked, while the identification of the region on which (D4) certifies defectivity is the two-line computation displayed in Section 8, done by hand. There is no `sorry` anywhere and no axiom is declared by hand.

The split between the two halves is deliberate, and it is the same at depth 8 as at depth 7. The *minimality* half — the sweeps of Theorems 1.1 and 1.6, the bounds of Theorems 1.2 and 1.3, the seventeen of Theorem 1.7, the soundness Proposition 4.1 and its counterparts for the ladder and for (D4′), the whole of Section 11, and every calculus control — is checked by the kernel alone: a print of its axiom set returns nothing beyond `propext` and `Quot.sound`, and the list of seventeen and its count return no axiom at all. The *filling* half and the rank sweeps of Theorems 1.4 and 1.8 — finite integer arithmetic on Jacobians up to 514×514 — use compiled evaluation, and each such theorem carries the corresponding evaluation axiom (one per theorem, four for the depth-8 Euler control; the exact-count and no-violation statements inherit the axiom of the rank sweep they are computed from and add none); the statements about $\mathbb{Z}$ and $\mathbb{Q}$ matrices in Lemma 3.1 use *Mathlib* and `Classical.choice`. The claims that rest on exhaustive finite computation are precisely: the 93 calculus evaluations of Theorem 1.1 and the 777 ladder evaluations of Theorem 1.6; the five and eight of Theorems 1.2 and 1.3; the 166 evaluations of Theorem 1.4 and the 1,210 (605 ranks and 605 bounds) of Theorem 1.8; the 23 certificate checks of Proposition 3.2, the two of Theorems 1.2 and 1.3 and the 111 of Proposition 3.4; the three rank computations of Proposition 1.5; the 777 evaluations without (D4), the $777 + 93$ with (D4′) and the handful of (D6) evaluations of Section 11. All of these ranges are given explicitly above; none is sampled. The only searches that are heuristic are the searches for the certificates and for the rank witnesses — random integer weights, best draw kept — and their output is checked, not trusted. The evaluator is independent of the code accompanying [1]: binary forms are represented as coefficient arrays and derivatives are taken by forward-mode dual numbers, where [2] uses the backpropagation routine of [4] over a different prime; the depth-8 integer points were produced by a second, faster implementation of the same evaluator, cross-checked entry by entry against the first on eight architectures, two of them of depth 8, and both agree with the source's printed $(2, 1, 1)$ Jacobian; the five depth-7 and four depth-8 values that overlap with the record agree. Two limitations are structural: the evaluator is specialised to $d_0 = 2$, so the two architectures of [1, Proposition 2.5] with $d_0 \neq 2$ and the $d_0 = 3$ remark of Section 8 lie outside it; and the step from a full-rank Jacobian to "$V_{d,r}$ is filling" is [1, Theorem A.2], cited and not formalised, since it needs generic smoothness. The statements in the text labelled as made outside the formal development support no theorem; they are the 12,907-row comparison and the four-row overlap of Section 10.5, the counts 54/402/147 of Section 10.2, the fuel range $f \leq 6$ of Section 10, and in Section 11 the 17 blocking subarchitectures, the 18 of 7,933, the two statements that (D4′) and U_6 change no value of the calculus, and the two measurements of the sharper forms of (D6).

The version-1 development compiles in about 0.8 CPU-hours and the depth-8 layer in about one more, the kernel sweep of Theorem 1.6 taking 149 seconds in eleven blocks where a single block over all 107 needed 435 seconds and 11.3 GB, and the 111 certificate checks under a minute once the Mathlib-free evaluator is compiled to a shared library; these timings are measurements outside the formal development and support no claim above.

The source of the development is currently held in a private repository: repository reference to be added at submission.

On the reference list. Every theorem, lemma, proposition, table, example and appendix number attributed to [1] above is the numbering printed in its **v2**, read off the compiled e-print. That numbering is version-sensitive, and so is the sentence quoted in Section 1: `v1` carries no appendices at all, and there the sentence reads "The results of highly reliable finite field computations for $p = 2^{31} - 1$ suggests that every entry in Table 2 is minimal filling. The proof is similar to the proof of Lemma 2.6, but with many more cases."

Appendix A. Formal statements

For reference, the statement of every machine-checked result of the formal development, grouped by the result of the text it supports; the modules are named in Section 12, and the line `-- Name` before each statement is the module it lives in. Throughout, an architecture is a `List N` $(d_0, \ldots, d_L)$; `ambDim r d` is $A_r(d)$, `expDim d` is $E(d)$, `ahBound`, `subspaceBound` and `subspaceBoundSafe` are the right-hand sides of (D3), (D4) and (D4′) as `Option N` (`none` when the hypothesis does not apply), `ub r f d` is $U_r^{(f)}(d)$, `ubSafe` is U', `ub6` is U_6, `ubNoSub` is the calculus without (D4), `interiorBound r rec d k` is the right-hand side of (D6) at layer k with the inner dimension evaluated by `rec`, `layerAmb r d k` is N_k, `notFilling r f d` decides $U_r^{(f)}(d) < A_r(d)$, `notFillingFast`, `notFillingSafe` and `notFillingNoSub` are the ladders of Section 10.2 for the three calculi, `minimalityCertified`, `minimalFast`, `minimalSafe` and `minimalNoSub` run them over `maximalSubs d` $= \mathcal{M}(d)$, `isUnimodal` is the walk of Section 2, `DimBounds r dim` and `DimBoundsSafe r dim` are the conjunctions (D1)–(D5) and (D1)–(D3), (D4′), (D5) for a function `dim`, and `IsFilling r dim d` is `dim d = ambDim r d`. A certificate `c` carries an architecture, the exponent, the prime, the integer weights and the selected columns; `certReport c` is the triple (rows, columns, minor mod p) of the selected submatrix, `fillingCertified c` the check of Section 3, `jacFull c` and `jacSub c` the whole Jacobian and its selected submatrix modulo p, `rankMod` and `detMod` rank and determinant over $\mathbb{F}_p$, `netEval c` the network's output, and `eulerCheck c` the identity of Proposition 9.2(i). The named lists of architectures and of certificates are the data described in the text: `table2` and `table2Archs` are Table 1 (certificates and architectures), `prop24` the five of Proposition 3.2, `table1Subs` the six of Proposition 5.2, `subPoints` the 83 of Theorem 1.4, `depth8Certs` the 111 of Proposition 3.4, `d8chunk01` to `d8chunk11` the eleven blocks of Table 3 with `depth8Chunks` their list and `depth8MFAs` their concatenation, `depth8Open` Table 4, and `subPoints8_1` to `subPoints8_4` the four blocks of Theorem 1.8; `paperCert` is the source's Appendix-C point (`paperCert11`, `paperCert29` the same point read modulo 11 and 29), `cSmall` the source's $(2, 1, 1)$ example, `cR3` and `cDL2` the certificates of Theorems 1.2 and 1.3, `cTooLarge` and `cTooSmall` those of Proposition 9.2, `a10_3` to `a10_5` the three points of Proposition 1.5, and `d8c000` to `d8c110` the 111 depth-8 certificates in the order of Tables 3 and 4. `dimReport` and `dimReport8_1` to `dimReport8_4` pair each subarchitecture's certified rank with its calculus bound. The statements are reproduced unchanged except that the source's Unicode is transliterated into ASCII: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ as `N`, `Z`, `Q`, $\forall$ as `forall`, $\in$ as `in`, $\to$ as `->`, $\leq$ as `<=`, $\neq$ as `<>`, $\times$ as `x`, and $\wedge, \neg$ as `and`, `not`; a literal list of more than a few hundred characters is shown by its first three entries and its length, as `[a, b, c, ... (n entries in all) ...]`. Proofs are omitted.

Lemma 3.1 (the rank step), with its two controls.

```
-- Rank
theorem control_drop :
    ((!![(11 : Z)]).map (Int.cast : Z -> Q)).det <> 0 and
      ((!![(11 : Z)]).map (Int.cast : Z -> ZMod 11)).det = 0

-- Rank
```

```
theorem control_nonvacuous :
    ((!![(3 : Z)]).map (Int.cast : Z -> ZMod 5)).det <> 0

-- Rank
theorem det_ne_zero_of_map_det_ne_zero {k : N} (n : N) (A : Matrix (Fin k) (Fin k) Z)
    (h : (A.map (Int.cast : Z -> ZMod n)).det <> 0) : A.det <> 0

-- Rank
theorem rank_ge_of_minor_ne_zero_mod {m n k : N} (N : N) (J : Matrix (Fin m) (Fin n) Z)
    (ri : Fin k -> Fin m) (ci : Fin k -> Fin n)
    (h : ((J.submatrix ri ci).map (Int.cast : Z -> ZMod N)).det <> 0) :
    k <= (J.map (Int.cast : Z -> Q)).rank
```

Proposition 3.3 (the source's printed data reproduced).

```
-- Filling
theorem paper_lemma22 :
    certReport paperCert = (65, 65, 75) and fillingCertified paperCert = true

-- Filling
theorem source_example_jacobian :
    jacSub cSmall = #[#[2, 0, 1], #[0, 2, 0], #[0, 0, 0]] and
      netEval cSmall = #[#[1, 0, 0]] and detMod 101 (jacSub cSmall) = 0

-- Filling
theorem table1_reported_dimensions :
    table1Subs.map (fun c => rankMod c.p (jacFull c)) = #[35, 60, 62, 39, 61, 59]
```

Proposition 3.2 (the $18 + 5$ filling certificates).

```
-- Filling
theorem prop24_filling :
    prop24.all fillingCertified = true and
      (prop24.map certReport).toList =
        [(20, 20, 98), (9, 9, 71), (20, 20, 17), (27, 27, 2), (36, 36, 5)]

-- Filling
theorem table2_filling : table2.all fillingCertified = true

-- Filling
theorem table2_reports :
    (table2.map certReport).toList =
      [(3, 3, 74), (5, 5, 45), (9, 9, 38), (17, 17, 4), (33, 33, 5),
       (65, 65, 89), (65, 65, 26), (65, 65, 14), (65, 65, 19), (65, 65, 90),
       (65, 65, 18), (65, 65, 76), (65, 65, 48), (65, 65, 14), (65, 65, 75),
       (65, 65, 52), (65, 65, 65), (65, 65, 16)]
```

Theorem 1.1 (the minimality half of Table 1, and the census as a theorem).

```
-- Minimal
theorem table2_ambient :
    table2Archs.map (fun d => ambDim 2 d) =
      [3, 5, 9, 17, 33, 65, 65, 65, 65, 65, 65, 65, 65, 65, 65, 65, 65, 65]

-- Minimal
theorem table2_minimality :
    table2Archs.all (fun d => minimalityCertified 2 2 d) = true

-- Minimal
theorem table2_sub_bounds :
    table2Archs.map (fun d => (maximalSubs d).map (fun s => ub 2 2 s)) =
      [[2], [2, 3], [7, 7, 5], [11, 9, 11, 9], [19, 13, 24, 26, 17],
       [35, 61, 63, 39, 61, 62], [35, 62, 59, 62, 53, 33], [35, 21, 56, 56, 61, 59],
       [35, 60, 60, 52, 62, 33], [35, 21, 60, 57, 54, 64], [35, 21, 57, 62, 53, 33],
       [35, 21, 42, 58, 57, 33], [35, 57, 57, 54, 54, 64], [35, 63, 51, 57, 57, 63],
       [35, 21, 64, 56, 54, 63], [35, 21, 63, 55, 57, 63], [35, 56, 64, 53, 54, 63],
```

```
      [35, 64, 56, 53, 54, 63]]

-- Bounds
theorem table2_minimal_filling {dim : List N -> N} (h : DimBounds 2 dim)
    (hfill : forall d in table2Archs, IsFilling 2 dim d) :
    forall d in table2Archs, IsFilling 2 dim d and forall s in maximalSubs d,
    not IsFilling 2 dim s

-- Bounds
theorem table2_subs_not_filling {dim : List N -> N} (h : DimBounds 2 dim) :
    forall d in table2Archs, forall s in maximalSubs d, dim s < ambDim 2 s
```

Proposition 5.2 (the source's minimality lemma recovered).

```
-- Minimal
theorem lemma25_bounds :
    (maximalSubs [2,3,4,5,4,6,4,1]).map (fun s => ub 2 2 s) = [35, 61, 63, 39, 61, 62]

-- Minimal
theorem lemma25_subs :
    maximalSubs [2,3,4,5,4,6,4,1] =
      [[2,2,4,5,4,6,4,1], [2,3,3,5,4,6,4,1], [2,3,4,4,4,6,4,1],
       [2,3,4,5,3,6,4,1], [2,3,4,5,4,5,4,1], [2,3,4,5,4,6,3,1]]
```

Theorem 1.2 (the $r = 3$ counterexample).

```
-- Minimal
theorem r3_minimality :
    minimalityCertified 3 2 [2,3,7,6,15,8,1] = true and
      (maximalSubs [2,3,7,6,15,8,1]).map (fun s => ub 3 2 s) = [84, 240, 225, 235,
      233] and
      ambDim 3 [2,3,7,6,15,8,1] = 244

-- Filling
theorem r3_filling : fillingCertified cR3 = true

-- Bounds
theorem r3_subs_not_filling {dim : List N -> N} (h : DimBounds 3 dim) :
    forall s in maximalSubs [2,3,7,6,15,8,1], dim s < ambDim 3 s
```

Theorem 1.3 (the $d_L = 2$ counterexample, seven of eight).

```
-- Minimal
theorem dL2_minimality_partial :
    (maximalSubs [2,3,4,4,10,17,11,12,4,2]).map (fun s => ambDim 2 [2,3,4,4,10,17,11,12,4,2] - ub 2 2 s)
      = [254, 0, 176, 5, 17, 34, 11, 124] and
    ((maximalSubs [2,3,4,4,10,17,11,12,4,2]).filter (fun s => notFilling 2 2 s)).length = 7 and
    minimalityCertified 2 2 [2,3,4,4,10,17,11,12,4,2] = false

-- Filling
theorem dL2_filling : fillingCertified cDL2 = true
```

Proposition 9.1 (controls on the calculus).

```
-- Minimal
theorem control_fuel_converged :
    (maximalSubs [2,3,4,5,4,6,4,1]).map (fun s => ub 2 4 s)
      = (maximalSubs [2,3,4,5,4,6,4,1]).map (fun s => ub 2 2 s)

-- Minimal
theorem control_fuel_needed :
    table2Archs.all (fun d => minimalityCertified 2 1 d) = false

-- Minimal
theorem control_no_table2_entry_refuted :
    (table2Archs.filter (fun d => notFilling 2 2 d)).length = 0
```

```
-- Minimal
theorem control_subspace_bound_needed :
    ub 2 2 [2,3,4,5,5,4,3,1] = 64 and subspaceBound 2 [2,3,4,5] = some 24 and
      subspaceBound 2 [3,5,6] = some 35 and min (ambDim 2 [2,3,4,5]) (expDim [2,3,4,5]) = 25 and
      min (ambDim 2 [3,5,6]) (expDim [3,5,6]) = 36

-- Minimal
theorem control_too_large :
    minimalityCertified 2 2 [2,4,4,5,4,6,4,1] = false and
      [2,3,4,5,4,6,4,1] in maximalSubs [2,4,4,5,4,6,4,1]

-- Minimal
theorem control_too_small :
    notFilling 2 2 [2,2,2,2,2,2,2,1] = true and ub 2 2 [2,2,2,2,2,2,2,1] = 13 and
      notFilling 2 2 [2,1,1] = true and ub 2 2 [2,1,1] = 2 and ambDim 2 [2,1,1] = 3
```

Proposition 9.2 (controls on the certificates).

```
-- Filling
theorem control_euler :
    table2.all eulerCheck = true and prop24.all eulerCheck = true and
      eulerCheck paperCert = true and eulerCheck cR3 = true

-- Filling
theorem control_too_large_filling :
    fillingCertified cTooLarge = true and fillingCertified paperCert = true

-- Filling
theorem control_too_small_rank :
    rankMod cTooSmall.p (jacFull cTooSmall) = 13 and certReport cSmall = (3, 3, 0)

-- Filling
theorem control_wrong_prime :
    rankMod paperCert11.p (jacFull paperCert11) = 60 and
      rankMod paperCert29.p (jacFull paperCert29) = 60 and
      rankMod paperCert.p (jacFull paperCert) = 65
```

Proposition 4.1 (soundness of the calculus), with its controls.

```
-- Bounds
theorem control_consistent (r : N) : DimBounds r (fun _ => 0)

-- Bounds
theorem control_gap_nontrivial :
    ub 2 2 [2,2,4,5,4,6,4,1] = 35 and ambDim 2 [2,2,4,5,4,6,4,1] = 65

-- Bounds
theorem control_sharp :
    ub 2 2 [2,2,4,5,4,6,4,1] = 35 and ub 2 2 [2,3,4,5,3,6,4,1] = 39 and
      ub 2 2 [2,3,4,5,4,5,4,1] = 61

-- Bounds
theorem dim_le_baseBound {r : N} {dim : List N -> N} (h : DimBounds r dim) (d : List N) :
    dim d <= baseBound r d

-- Bounds
theorem dim_le_ub {r : N} {dim : List N -> N} (h : DimBounds r dim) :
    forall (fuel : N) (d : List N), dim d <= ub r fuel d
```

Theorem 1.4 (71 exact dimensions at depth ≤ 7).

```
-- Dims
theorem control_no_violation :
    (dimReport.toList.filter (fun q => q.2 < q.1)).length = 0

-- Dims
```

```
theorem dimReport_value :
    dimReport.toList =
      [(2, 2), (2, 2), (3, 3), ... (83 entries in all) ...]

-- Dims
theorem exact_dimension_count :
    (dimReport.toList.filter (fun q => q.1 = q.2)).length = 71

-- Dims
theorem subPoint_archs :
    (subPoints.map (fun c => c.arch.toList)).toList =
      [[2, 1, 1], [2, 1, 2, 1], [2, 2, 1, 1], ... (83 entries in all) ...]
```

Proposition 1.5 (the corrected defectivity criterion).

```
-- Dims
theorem thmA10_offByOne :
    (#[a10_3, a10_4, a10_5].map
      (fun c => (rankMod c.p (jacFull c),
      min (ambDim 2 c.arch.toList) (expDim c.arch.toList),
                 c.arch.toList))).toList
      = [(8, 8, [2,2,3]), (10, 10, [2,2,4]), (12, 12, [2,2,5])]

-- Dims
theorem thmA10_repaired :
    subspaceBound 2 [3,5,6] = some 35 and min (ambDim 2 [3,5,6]) (expDim [3,5,6]) = 36 and
      subspaceBound 2 [2,2,3] = some 8 and min (ambDim 2 [2,2,3]) (expDim [2,2,3]) = 8
```

Proposition 3.4 (the 111 depth-8 certificates and the Euler control).

```
-- Depth8Filling
theorem depth8_archs :
    (depth8Certs.map (fun c => c.arch.toList)).toList =
      [[2,3,3,4,4,7,8,7,1], [2,3,3,4,4,8,8,6,1], [2,3,3,4,5,6,8,8,1],
      ... (111 entries in all) ...]

-- Depth8Filling
theorem depth8_filling : depth8Certs.all fillingCertified = true

-- Depth8Reports
theorem control_euler_depth8 :
    eulerCheck d8c000 = true and eulerCheck d8c001 = true and eulerCheck d8c002 = true and
      eulerCheck d8c110 = true

-- Depth8Reports
theorem depth8_reports :
    (depth8Certs.map certReport).toList =
      [(129, 129, 15), (129, 129, 36), (129, 129, 75), ... (111 entries in all) ...]
```

Theorem 1.6 (the depth-8 census: the eleven blocks, their gluing, and the census as a theorem).

```
-- Depth8Min
theorem d8min01 : d8chunk01.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min02 : d8chunk02.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min03 : d8chunk03.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min04 : d8chunk04.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min05 : d8chunk05.all (fun d => minimalFast 2 d) = true
```

```
-- Depth8Min
theorem d8min06 : d8chunk06.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min07 : d8chunk07.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min08 : d8chunk08.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min09 : d8chunk09.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min10 : d8chunk10.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem d8min11 : d8chunk11.all (fun d => minimalFast 2 d) = true

-- Depth8Min
theorem depth8_ambient :
    depth8MFAs.length = 107 and depth8MFAs.all (fun d => decide (ambDim 2 d = 129)) = true

-- Depth8Bounds
theorem depth8_all_min : forall d in depth8MFAs, minimalFast 2 d = true

-- Depth8Bounds
theorem depth8_chunks_min :
    forall c in depth8Chunks, c.all (fun d => minimalFast 2 d) = true

-- Depth8Bounds
theorem depth8_minimal_filling {dim : List N -> N} (h : DimBounds 2 dim)
    (hfill : forall d in depth8MFAs, IsFilling 2 dim d) :
    forall d in depth8MFAs, IsFilling 2 dim d and forall s in maximalSubs d,
    not IsFilling 2 dim s

-- Depth8Bounds
theorem depth8_subs_not_filling {dim : List N -> N} (h : DimBounds 2 dim) :
    forall d in depth8MFAs, forall s in maximalSubs d, dim s < ambDim 2 s
```

Theorem 1.7 (the seventeen non-unimodal entries).

```
-- Depth8Min
theorem depth8_nonunimodal :
    depth8MFAs.filter (fun d => ! isUnimodal d) =
      [[2,3,3,4,7,6,7,6,1], [2,3,3,5,7,5,8,5,1], [2,3,3,5,7,6,7,4,1],
       [2,3,3,6,6,5,8,8,1], [2,3,3,6,7,6,7,3,1], [2,3,4,4,7,5,8,8,1],
       [2,3,4,4,7,6,7,5,1], [2,3,4,5,4,7,8,5,1], [2,3,4,5,4,8,8,4,1],
       [2,3,4,5,6,5,8,5,1], [2,3,4,5,7,5,7,5,1], [2,3,4,5,7,5,8,4,1],
       [2,3,4,5,7,6,7,3,1], [2,3,4,5,8,5,7,4,1], [2,3,4,6,5,8,6,3,1],
       [2,3,4,6,6,5,7,7,1], [2,3,4,6,7,5,7,4,1]]

-- Depth8Min
theorem depth8_nonunimodal_count :
    (depth8MFAs.filter (fun d => ! isUnimodal d)).length = 17 and
      isUnimodal [2,3,4,5,4,6,4,1] = false and isUnimodal [2,3,4,5,6,4,2,1] = true and
      isUnimodal [2,2,1] = true and isUnimodal [1,3,2,4] = false and
      isUnimodal [2,2,1,2] = false

-- Depth8Bounds
theorem depth8_nonunimodal_counterexamples {dim : List N -> N} (h : DimBounds 2 dim)
    (hfill : forall d in depth8MFAs, IsFilling 2 dim d) :
    forall d in depth8MFAs.filter (fun d => ! isUnimodal d),
      isUnimodal d = false and IsFilling 2 dim d and
        forall s in maximalSubs d, not IsFilling 2 dim s
```

Theorem 1.8 (533 **exact dimensions at depth** 8**, in four blocks).**

```
-- Depth8Dims1
theorem control8_1_no_violation :
    (dimReport8_1.toList.filter (fun q => q.2 < q.1)).length = 0

-- Depth8Dims1
theorem dimReport8_1_value :
    dimReport8_1.toList =
      [(67, 67), (67, 67), (67, 67), ... (152 entries in all) ...]

-- Depth8Dims1
theorem exact8_1_count :
    (dimReport8_1.toList.filter (fun q => q.1 = q.2)).length = 152

-- Depth8Dims1
theorem subPoints8_1_archs :
    (subPoints8_1.map (fun c => c.arch.toList)).toList =
      [[2,2,3,4,4,7,8,7,1], [2,2,3,4,4,8,8,6,1], [2,2,3,4,5,6,8,8,1],
      ... (152 entries in all) ...]

-- Depth8Dims2
theorem control8_2_no_violation :
    (dimReport8_2.toList.filter (fun q => q.2 < q.1)).length = 0

-- Depth8Dims2
theorem dimReport8_2_value :
    dimReport8_2.toList =
      [(37, 37), (37, 37), (37, 37), ... (152 entries in all) ...]

-- Depth8Dims2
theorem exact8_2_count :
    (dimReport8_2.toList.filter (fun q => q.1 = q.2)).length = 130

-- Depth8Dims2
theorem subPoints8_2_archs :
    (subPoints8_2.map (fun c => c.arch.toList)).toList =
      [[2,3,2,5,8,7,5,2,1], [2,3,2,5,8,8,4,2,1], [2,3,2,5,9,7,4,3,1],
      ... (152 entries in all) ...]

-- Depth8Dims3
theorem control8_3_no_violation :
    (dimReport8_3.toList.filter (fun q => q.2 < q.1)).length = 0

-- Depth8Dims3
theorem dimReport8_3_value :
    dimReport8_3.toList =
      [(121, 121), (65, 65), (122, 122), ... (152 entries in all) ...]

-- Depth8Dims3
theorem exact8_3_count :
    (dimReport8_3.toList.filter (fun q => q.1 = q.2)).length = 122

-- Depth8Dims3
theorem subPoints8_3_archs :
    (subPoints8_3.map (fun c => c.arch.toList)).toList =
      [[2,3,3,5,6,9,4,3,1], [2,3,3,5,6,9,5,1,1], [2,3,3,5,6,10,4,2,1],
      ... (152 entries in all) ...]

-- Depth8Dims4
theorem control8_4_no_violation :
    (dimReport8_4.toList.filter (fun q => q.2 < q.1)).length = 0
```

```
-- Depth8Dims4
theorem dimReport8_4_value :
    dimReport8_4.toList =
      [(101, 102), (117, 117), (125, 125), ... (149 entries in all) ...]

-- Depth8Dims4
theorem exact8_4_count :
    (dimReport8_4.toList.filter (fun q => q.1 = q.2)).length = 129

-- Depth8Dims4
theorem subPoints8_4_archs :
    (subPoints8_4.map (fun c => c.arch.toList)).toList =
      [[2,3,4,5,4,8,7,2,1], [2,3,4,5,4,8,7,4,1], [2,3,4,5,4,8,8,3,1],
      ... (149 entries in all) ...]
```

Proposition 10.1 (the ladder is sound; the gap; the depth-8 controls).

```
-- Depth8Min
theorem control_depth8_fuel :
    ub 2 0 [2,3,3,3,6,8,6,3,1] = 127 and
      notFilling 2 0 [2,2,3,4,4,7,8,7,1] = false and ub 2 1 [2,2,3,4,4,7,8,7,1] = 67 and
      notFilling 2 1 [2,3,3,3,4,8,8,6,1] = false and ub 2 2 [2,3,3,3,4,8,8,6,1] = 125 and
      notFillingFast 2 [2,3,3,3,4,8,8,6,1] = true

-- Depth8Min
theorem control_depth8_too_large :
    minimalFast 2 [2,4,3,4,4,7,8,7,1] = false and
      [2,3,3,4,4,7,8,7,1] in maximalSubs [2,4,3,4,4,7,8,7,1] and
      [2,3,3,4,4,7,8,7,1] in d8chunk01

-- Depth8Min
theorem control_depth8_too_small :
    notFillingFast 2 [2,2,2,2,2,2,2,2,1] = true and ub 2 2 [2,2,2,2,2,2,2,2,1] = 15 and
      ambDim 2 [2,2,2,2,2,2,2,2,1] = 129 and notFillingFast 2 [2,1,1,1,1,1,1,1,1] = true and
      ub 2 2 [2,1,1,1,1,1,1,1,1] = 2

-- Depth8Min
theorem depth8_open_blocked :
    depth8Open.all (fun d => minimalFast 2 d) = false and
      depth8Open.map (fun d => (maximalSubs d).filter (fun s => ! notFillingFast 2 s)) =
        [[[2,3,3,4,7,6,8,3,1]], [[2,3,3,4,7,6,8,3,1]], [[2,3,3,5,9,7,4,2,1]],
         [[2,3,3,4,7,6,8,3,1]]]

-- Depth8Bounds
theorem control_depth8_gap :
    ub 2 1 [2,2,3,4,4,7,8,7,1] = 67 and ambDim 2 [2,2,3,4,4,7,8,7,1] = 129 and
      depth8MFAs.length = 107

-- Depth8Bounds
theorem control_needs_filling (r : N) : DimBoundsSafe r (fun _ => 0)

-- Depth8Bounds
theorem dim_lt_of_notFillingFast {r : N} {dim : List N -> N} (h : DimBounds r dim)
    {d : List N} (hd : notFillingFast r d = true) : dim d < ambDim r d
```

Proposition 11.1 (necessity of (D4) in eleven blocks; (D4′), its soundness, and both censuses under it).

```
-- Depth8Safe
theorem control_safe_agrees :
    subspaceBoundSafe 2 [3,5,6] = some 35 and subspaceBound 2 [3,5,6] = some 35 and
      subspaceBoundSafe 2 [2,3,4,5] = some 24 and subspaceBound 2 [2,3,4,5] = some 24 and
      ubSafe 2 2 [2,3,4,5,5,4,3,1] = 64 and ub 2 2 [2,3,4,5,5,4,3,1] = 64

-- Depth8Safe
theorem control_safe_differs :
```

```
    subspaceBound 2 [2,5,8,8,4] = some 40 and subspaceBoundSafe 2 [2,5,8,8,4] = some 42

-- Depth8Safe
theorem d8safe01 : d8chunk01.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe02 : d8chunk02.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe03 : d8chunk03.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe04 : d8chunk04.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe05 : d8chunk05.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe06 : d8chunk06.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe07 : d8chunk07.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe08 : d8chunk08.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe09 : d8chunk09.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe10 : d8chunk10.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem d8safe11 : d8chunk11.all (fun d => minimalSafe 2 d) = true

-- Depth8Safe
theorem table2_minimality_safe : table2Archs.all (fun d => minimalSafe 2 d) = true

-- Depth8Subspace
theorem control_nosub_witness :
    notFillingNoSub 2 [2,3,3,4,7,5,8,8,1] = false and ub 2 2 [2,3,3,4,7,5,8,8,1] = 127 and
      notFillingNoSub 2 [2,3,3,5,8,6,6,2,1] = false and ub 2 2 [2,3,3,5,8,6,6,2,1] = 126 and
      ambDim 2 [2,3,3,4,7,5,8,8,1] = 129

-- Depth8Subspace
theorem d8nosub01 : (d8chunk01.filter (fun d => ! minimalNoSub 2 d)).length = 0

-- Depth8Subspace
theorem d8nosub02 : (d8chunk02.filter (fun d => ! minimalNoSub 2 d)).length = 1

-- Depth8Subspace
theorem d8nosub03 : (d8chunk03.filter (fun d => ! minimalNoSub 2 d)).length = 0

-- Depth8Subspace
theorem d8nosub04 : (d8chunk04.filter (fun d => ! minimalNoSub 2 d)).length = 4

-- Depth8Subspace
theorem d8nosub05 : (d8chunk05.filter (fun d => ! minimalNoSub 2 d)).length = 6

-- Depth8Subspace
theorem d8nosub06 : (d8chunk06.filter (fun d => ! minimalNoSub 2 d)).length = 2
```

```
-- Depth8Subspace
theorem d8nosub07 : (d8chunk07.filter (fun d => ! minimalNoSub 2 d)).length = 1

-- Depth8Subspace
theorem d8nosub08 : (d8chunk08.filter (fun d => ! minimalNoSub 2 d)).length = 3

-- Depth8Subspace
theorem d8nosub09 : (d8chunk09.filter (fun d => ! minimalNoSub 2 d)).length = 5

-- Depth8Subspace
theorem d8nosub10 : (d8chunk10.filter (fun d => ! minimalNoSub 2 d)).length = 2

-- Depth8Subspace
theorem d8nosub11 : (d8chunk11.filter (fun d => ! minimalNoSub 2 d)).length = 6

-- Depth8Bounds
theorem depth8_chunks_safe :
    forall c in depth8Chunks, c.all (fun d => minimalSafe 2 d) = true

-- Depth8Bounds
theorem depth8_minimal_filling_safe {dim : List N -> N} (h : DimBoundsSafe 2 dim)
    (hfill : forall d in depth8MFAs, IsFilling 2 dim d) :
    forall d in depth8MFAs, IsFilling 2 dim d and forall s in maximalSubs d,
    not IsFilling 2 dim s

-- Depth8Bounds
theorem dim_le_baseBoundSafe {r : N} {dim : List N -> N} (h : DimBoundsSafe r dim)
    (d : List N) : dim d <= baseBoundSafe r d

-- Depth8Bounds
theorem dim_le_ubSafe {r : N} {dim : List N -> N} (h : DimBoundsSafe r dim) :
    forall (fuel : N) (d : List N), dim d <= ubSafe r fuel d

-- Depth8Bounds
theorem dim_lt_of_notFillingSafe {r : N} {dim : List N -> N} (h : DimBoundsSafe r dim)
    {d : List N} (hd : notFillingSafe r d = true) : dim d < ambDim r d

-- Depth8Bounds
theorem table2_subs_not_filling_safe {dim : List N -> N} (h : DimBoundsSafe 2 dim) :
    forall d in table2Archs, forall s in maximalSubs d, dim s < ambDim 2 s
```

Proposition 11.3 (the interior bound (D6)).

```
-- Depth8Subspace
theorem control_interior_nontrivial :
    interiorBound 2 (ub 2 2) [2,2,2,2,1] 1 = 7 and ambDim 2 [2,2,2,2,1] = 9 and
      layerAmb 2 [2,3,3,4,7,6,8,3,1] 1 = 3 and layerAmb 2 [2,3,3,4,7,6,8,3,1] 5 = 33

-- Depth8Subspace
theorem interior_does_not_close :
    ub6 2 1 [2,3,3,4,7,6,8,3,1] = 129 and ub6 2 1 [2,3,3,5,9,7,4,2,1] = 129 and
      ambDim 2 [2,3,3,4,7,6,8,3,1] = 129

-- Depth8Subspace
theorem interior_does_not_close_dL2 :
    ub6 2 1 [2,3,3,4,10,17,11,12,4,2] = 514 and
      ambDim 2 [2,3,3,4,10,17,11,12,4,2] = 514

-- Depth8Subspace
theorem interior_is_subspace_at_last :
    interiorBound 2 (ub 2 2) [2,3,4,5] 2 = 24 and interiorBound 2 (ub 2 2) [3,5,6] 1 = 35
```

References

[1] K. Dao and J. I. Rodriguez, *Minimal Filling Architectures of Polynomial Neural Networks: Counterexamples, Frontier Search, and Defects*, arXiv:2605.09609v2 [cs.LG], cross-listed math.AG; v1 submitted 10 May 2026, v2 17 June 2026. The source of the problem and of everything attributed above: Definitions 1.1, 1.3, 1.5, 1.6 and Conjecture 1.7, Table 2 and its eighteen architectures, Lemma 2.2, Lemma 2.3, Lemma 2.4, Proposition 2.5, Lemma 2.6, Table 1, Examples 4.2 and 4.3, Theorem A.2, Corollary A.6, Example A.8, Theorems B.4 and B.6, Appendix B, the Appendix-C weights, and the Discussion sentence quoted in Section 1. arXiv record checked live on 3 September 2026: two versions, the two authors above, primary category cs.LG, journal-reference and comment fields both empty. Presented at the ICML 2026 workshop *CoLoRAI — the 2nd Workshop on Connecting Low-rank Representations in AI* (workshop programme at `icml.cc`, checked 3 September 2026; recorded the same way in the bibliography of [5]). The "Proceedings of the 43rd International Conference on Machine Learning" line on the first page of the e-print is the boilerplate of the `icml2026` class file, not a statement about the venue.

[2] K. Dao, *MFA_PNNs*, `github.com/daokevin06/MFA_PNNs`, the search harness accompanying [1]; created 8 May 2026, last push 31 August 2026 (GitHub API `created_at` and `pushed_at`), read on 3 September 2026. Jupyter and Sage; no formal development. Cited for its search data (`data/processed`, dimension values over GF(100003)), for its minimality notebook, which decides "short" from a Jacobian rank below the ambient dimension at a random point, and for the attribution of its dimension routine to the MathRepo page of [4]. For version 2 also cited for the depth-8 block of `2_1_cleaned_architectures.csv` (6,639 of 12,907 rows; 111 rows flagged `is_minimal` and `is_full_dimension`), for `counterexamples.csv`, for `is_unimodal.py`, and for the notice in `notebooks/process_csv_files.ipynb` quoted in Section 1; all re-read from a fresh clone on 3 September 2026.

[3] J. Kileel, M. Trager and J. Bruna, *On the Expressive Power of Deep Polynomial Neural Networks*, in: Advances in Neural Information Processing Systems 32 (NeurIPS 2019), 10310–10319; arXiv:1905.12207 [cs.LG] (the arXiv record carries no journal reference; the venue and pages are those of the NeurIPS proceedings entry, checked 3 September 2026). Source of the unimodality conjecture (its Conjecture 12) that [1] refutes, and of the recursive dimension bound (D5), which is its Proposition 17.

[4] K. Kubjas, J. Li and M. Wiesmann, *Geometry of Polynomial Neural Networks*, arXiv:2402.00949 [math.AG]; Algebraic Statistics **15** (2024), no. 2, 295–328, `doi:10.2140/astat.2024.15.295` (arXiv abstract page and Crossref, both checked 3 September 2026). Cited for the geometry of neurovarieties, for the monotonicity (M), for the term *neuromanifold* it gives the functional space, and as the origin of the dimension routine used by [2].

[5] A. Massarenti and M. Mella, *The Alexander–Hirschowitz theorem for neurovarieties*, arXiv:2511.19703 [math.AG]; v1 24 November 2025, v2 1 September 2026, checked live 3 September 2026. Cited by [1] for the bound $\dim V_{(2,d_1,4,5),2} \leq 24$; its v2 comment field reads "Thanks to Kevin Dao, we identified a gap in the previous proof of the main non-defectiveness result. We replaced the transversality argument with a direct analysis of the differential…". Its intermediate Grassmann condition (its (4.7)–(4.8) in `v2`) is the construction behind (D4).

[6] H. Abo and M. C. Brambilla, *Secant varieties of Segre–Veronese varieties $\mathbb{P}^m \times \mathbb{P}^n$ embedded by $\mathcal{O}(1,2)$*, Experimental Mathematics **18** (2009), no. 3, 369–384, `doi:10.1080/10586458.2009.10129051`; arXiv:0809.4837 [math.AG]. Bibliographic data from Crossref and the arXiv abstract page, both checked 3 September 2026. Quoted by [1] as its Theorem B.6. Read here in its **v3** of 6 February 2012, which the arXiv record labels the published version: its Remark 4.4 and the inequality (4.1) there are what Section 8 compares with the criterion as printed in [1], and the numbering is version-sensitive.

[7] J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Algebraic Geom. **4** (1995), no. 2, 201–222 (Zbl 0829.14002; checked against zbMATH on 3 September 2026). Cited by [1] as its Theorem B.4; used here only through (D3), where at $r = 2$ it reduces to the dimension of a determinantal variety of symmetric matrices.

[8] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science **12699**, Springer, 2021, 625–635, `doi:10.1007/978-3-030-79876-5_37`.

[9] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, 367–381, `doi:10.1145/3372885.3373824`.