For m ≥ 2, every m×3 biquadratic sum of squares is a sum of 2m squares

Real algebraic geometry · sums of squares on projective varieties

Carles Marín

Statements

Theorem 1 3
Lemma 2 7
Proposition 3 9
Corollary 4 11
Lemma 5 12
Lemma 6 14
Corollary 7 15
Question 1 18

At a glance

What the paper contributes

The exact sum-of-squares rank of biquadratic forms of width three: every $m \times 3$ biquadratic sum of squares is a sum of $2m$ squares of bilinear forms, and some need $2m$. This answers the question of Qi, Löfberg and Chen on the exact values of $\mathrm{BSR}(m, 3)$; the cases $m = 2, 3$ were known, $m \geq 4$ is new.

- $\mathrm{BSR}(m, 3) = 2m$ for every $m \geq 2$ (Theorem 1); in particular $\mathrm{BSR}(4, 3) = 8$.
- An integer point where the map of $2m$ squares is a submersion, in every size (Lemma 2).
- A rank inequality for Hankel multipliers without simple tensors: rank at least the support plus one (Lemma 5), replacing the Gorenstein classification of the boundary strategy of Blekherman, Dunbar and Sinn.
- An explicit form needing $2m$ squares (Lemma 6).

🚀 **Start here: Theorem 1**: BSR(m,3) = 2m for every m ≥ 2.

Main results:

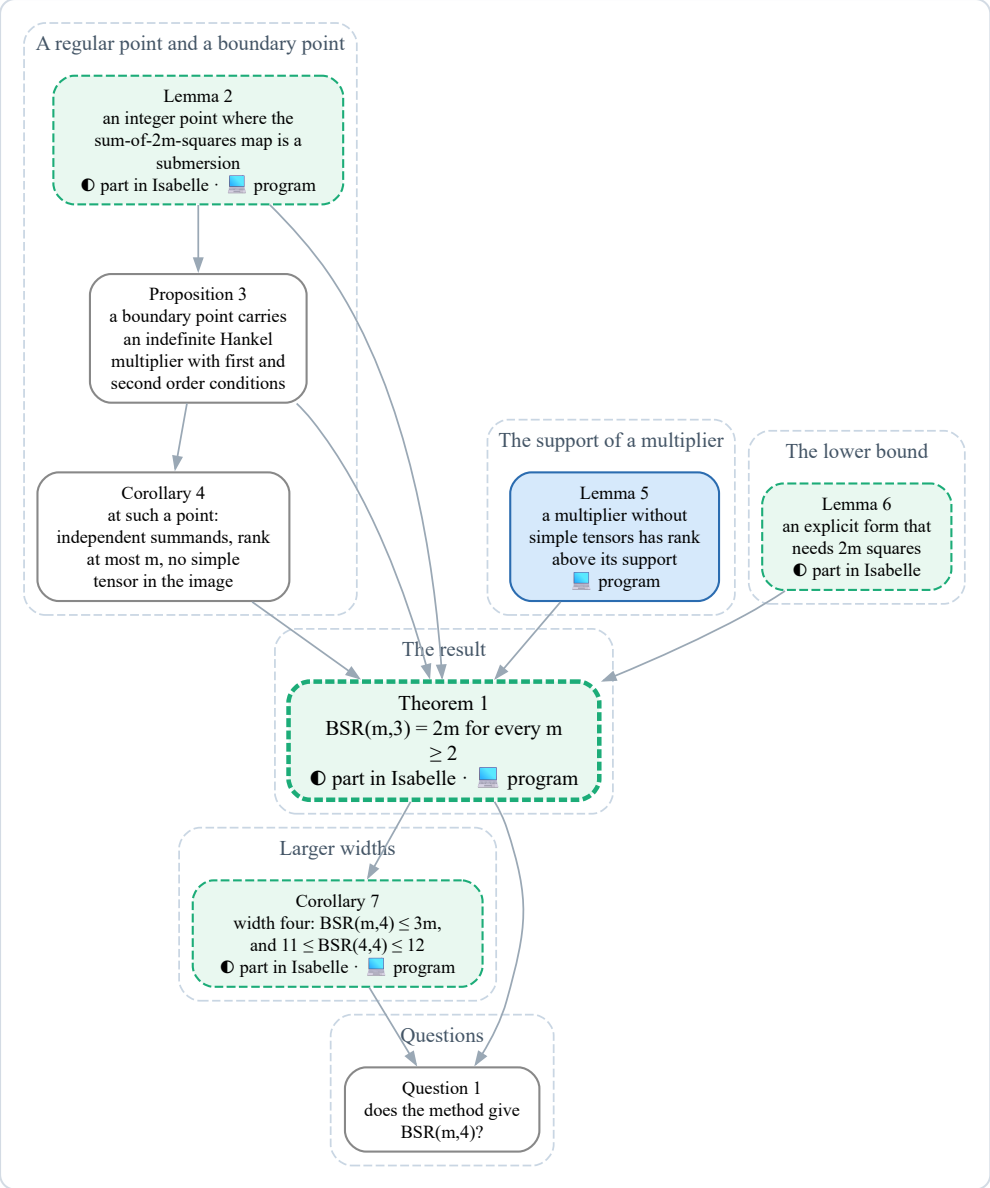
- **Theorem 1**: BSR(m,3) = 2m for every m ≥ 2

8 results | ✅ **0** entirely in Lean | ◐ **4** partly in Lean | 💻 **4** checked by a program | written proof only

(a result can carry several tags)

🗺️ Map of the results (click a box to go to it)

Dependencies in the written proof: an arrow A → B means that this proof of B uses A. Every result rests on its written proof; the tags record what was checked besides, and each statement's “Evidence & scope” block says how far. The colour of a box is its strongest tag, in the order above; the tags in the box give the rest. Boxes with a thick border are the main results; click a box to read the statement.



📐 Setting (copied from the paper, §§1–2)

Write $x = (x_1, \dots, x_m)$, $y = (y_1, y_2, y_3)$. A *biquadratic form* is a real form of bidegree $(2,2)$ in (x, y), and $\mathrm{BSR}(m, n)$ is the least r such that every biquadratic sum of squares in $m + n$ variables is a sum of r squares of bilinear forms. Let $A = \mathbb{R}^m$, $B = \mathbb{R}^3$, $V = A \otimes B$ (bilinear forms), R_2 the biquadratic forms, of dimension $N = 3m(m+1)$, $\Phi_k(h_1, \dots, h_k) = h_1^2 + \cdots + h_k^2$, $\Sigma_k = \Phi_k(V^k)$ and Σ the cone of sums of squares. To $\lambda \in R_2^*$ corresponds $H_\lambda(u, v) = \lambda(uv)$, with the Hankel symmetry $H(x \otimes b, x' \otimes c) = H(x \otimes c, x' \otimes b)$.

1 The result

Theorem 1. p. 1

🔁 **Changed in v2. Now:** the trace argument keeps the positive and negative inertia indices Statements that use it: Corollary 7, Question 1.

🔁 **Changed in v2. Before:** the title and statement for every m **Now:** for m ≥ 2 **Why:** for m = 1 the form $x_1^2(y_1^2+y_2^2+y_3^2)$ needs three squares Statements that use it: Corollary 7, Question 1.

> *For every* $m \geq 2$, $\mathrm{BSR}(m, 3) = 2m$.

Isabelle · ours · part **show**

Only a part is in Isabelle: partly: the lower bound 2m (Lemma 6, for the witness with slopes 0, …, m−1), the algebraic core of Lemma 2, and for the case e = 1 the second-variation identity in adapted coordinates and the trace contradiction under symmetry, L ≥ 0, full support, rank ≤ a + 1 and a negative direction (L ≥ 0 is equivalent to the second-order condition for every variation of the family); outside the formal part: the passage from a representation to the summands $x_0 \otimes y_j$ in adapted coordinates, Proposition 3, Lemma 5, the case e ≥ 2 and the assembly of the upper bound

`bsr_width_three_lower` · `Isabelle/BSRWidthThree/BSR_Width3_Lower.thy`, line 352 — the lower bound: if every m × 3 sum of squares is a sum of N squares, then 2m ≤ N (no hypothesis)

```
theorem bsr_width_three_lower:
  assumes "bsr_le m 3 N" shows "2 * m \<le> N"
```

`width_three_trace` · `Isabelle/BSRWidthThree/Width3_Final.thy`, line 196 — the case e = 1 of the induction: no multiplier satisfies the second-order condition, full support, rank at most a + 1 and a negative direction

```
theorem width_three_trace:
  fixes h :: "nat \<Rightarrow> nat \<Rightarrow> nat \
<Rightarrow> nat \<Rightarrow> real"
    and W C :: "nat \<Rightarrow> nat \<times> nat \
<Rightarrow> real"
  assumes hs: "\<forall>p q l l'. h q p l l' = h p q l l'"
and bs: "\<forall>p q l l'. h p q l' l = h p q l l'"
    and L: "\<forall>u. 0 \<le> lform a h u u"
    and supp: "\<And>x. (\<forall>j<3. \<forall>w. hform a
h (blk j x) w = 0) \<Longrightarrow> \<forall>l<a. x l = 0"
    and rank: "\<forall>p<3. \<forall>q<3. \<forall>l<a. \
<forall>l'<a. h p q l l' = (\<Sum>t<Suc a. W t (p, l) * C t
(q, l'))"
    and neg: "hform a h u0 u0 < 0"
  shows False
```

Trust base: ⚠️ not checked. Written proof, read in full and audited by two independent readings; the theorem as a whole is not machine-checked: Isabelle/HOL checks, without sorry, the lower bound 2m (Lemma 6), the algebraic core of Lemma 2 and the case e = 1 (second variation and trace bound); Lemma 5, the case e ≥ 2, Proposition 3 and the assembly are not formalized

Proof: in this note · **Formalized:** Isabelle/HOL, by the author (public formalizations repository) · **Covers:** part: the lower bound 2m (Lemma 6, for the witness with slopes 0, …, m−1), the algebraic core of Lemma 2, and for the case e = 1 the second-variation identity in adapted coordinates and the trace contradiction under symmetry, L ≥ 0, full support, rank ≤ a + 1 and a negative direction (L ≥ 0 is equivalent to the second-order condition for every variation of the family); outside the formal part: the passage from a representation to the summands $x_0 \otimes y_j$ in adapted coordinates, Proposition 3, Lemma 5, the case e ≥ 2 and the assembly of the upper bound · **Cite:** Theorem 1, p. 1, doi:10.5281/zenodo.24051362

🆕 **Classical and new:** the paper: «The statement is new for m ≥ 4» (m = 2 by BPSV and BSSV Theorem 1.4; m = 3 by BSSV Theorem 1.5).

↳ Used in: Corollary 7 (p. 15), 1 (p. 18)

Proof

We prove $\Sigma_{2m} = \Sigma$ by strong induction on $m \geq 2$. Suppose not, and take F, H as in Proposition 3; fix a representation $F = \sum_{i \leq 2m} h_i^2$ and $W = W_h$. By Corollary 4 and Lemma 5, $a + 1 \leq r \leq m$; as $H \neq 0$, $a \geq 1$, so $1 \leq a \leq m - 1$. Let $A = A' \oplus A''$ orthogonally, $e = m - a$.

Case $e \geq 2$. Write $w = w' + w''$ with $w' \in A' \otimes B$, $w'' \in A'' \otimes B$. Since $A'' \otimes B$ lies in the radical of H and $W \subseteq \ker H$, the projection W' of W to $A' \otimes B$ lies in the kernel of $H|_{A' \otimes B}$, which has dimension $3a - r$. An orthogonal change of the $2m$ summands moves all the components in $A' \otimes B$ into at most $s = \dim W' \leq 3a - r$ of them; the other $2m - s$ summands are bilinear forms in the e variables of A'', and their sum of squares is a sum of squares in $e \times 3$ variables. As $2 \leq e < m$, by induction it is a sum of $2e$ squares. Hence F is a sum of at most $3a - r + 2(m - a) = 2m + a - r \leq 2m - 1$ squares, against Corollary 4.

Case $e = 1$. Then $a = m - 1$ and $r = m = a + 1$, so $W = \ker H$ and $A'' \otimes B \subseteq W$; let x_0 span A''. The coefficient map $\mathbb{R}^{2m} \to W$ of the representation is an isomorphism (Corollary 4); choosing an orthonormal basis of the preimage of $x_0 \otimes B$ and completing it, an orthogonal change of the summands gives $h_i = x_0 \otimes y_i$ $(i = 1, 2, 3)$ for a basis y_1, y_2, y_3 of B. Write H_{jl} for the blocks of $H|_{A' \otimes B}$ in this basis (symmetric, with $H_{jl} = H_{lj}$ by (1)) and $S = H_{11} + H_{22} + H_{33}$. For $v_1, v_2, v_3 \in A'$ let

$$D_i = \sum_{j,p} \varepsilon_{ijp}\, v_p \otimes y_j \quad (i = 1, 2, 3), \qquad D_i = 0 \quad (i > 3),$$

with ε the alternating symbol. Then $\sum_i h_i D_i = \sum_{i,j,p} \varepsilon_{ijp}(x_0 \otimes y_i)(v_p \otimes y_j) = 0$ by antisymmetry in i, j, and $\sum_i \varepsilon_{ijp}\varepsilon_{ilq} = \delta_{jl}\delta_{pq} - \delta_{jq}\delta_{pl}$ gives

$$\sum_i H(D_i, D_i) = \sum_p v_p^\top S v_p - \sum_{p,q} v_p^\top H_{pq} v_q.$$

By Proposition 3(2) this is nonnegative for all v, that is $L := I_3 \otimes S - H|_{A' \otimes B} \succeq 0$ in the block order. The three diagonal blocks of L sum to $3S - S = 2S$, so $S \succeq 0$. If $Sx = 0$, the three nonnegative numbers $L(x \otimes y_j, x \otimes y_j)$ sum to $2x^\top S x = 0$, so $L(x \otimes y_j, \cdot) = 0$ and hence $H(x \otimes y_j, \cdot) = 0$ for all j; by the choice of A', $x = 0$. So $S \succ 0$. The congruence by $\operatorname{diag}(S^{-1/2}, S^{-1/2}, S^{-1/2})$ turns $H|_{A' \otimes B}$ into $\widetilde{H}$ with $\widetilde{H} \preceq I_{3a}$, $\operatorname{tr} \widetilde{H} = \operatorname{tr}(S^{-1/2} S S^{-1/2}) = a$, $\operatorname{rank} \widetilde{H} = a + 1$ and the same positive and negative inertia indices as H, which by Proposition 3(3) has a negative eigenvalue. Then $\widetilde{H}$ has at most a positive eigenvalues, each at most 1, and a negative one, so $\operatorname{tr} \widetilde{H} < a$: a contradiction.

For $m = 2$ only $e = 1$ occurs, so the induction starts without hypothesis.

💡 **In plain words (commentary, not part of the paper)**

1. A biquadratic form in $x \in \mathbb{R}^m$, $y \in \mathbb{R}^3$ that is a sum of squares of bilinear forms can always be written with $2m$ squares, and some forms need all of them. Before this note one knew $2m \le \mathrm{BSR}(m, 3) \le 2m + 1$; the theorem removes the last unit. 2. Why: if some form needed more, a generic form on the edge of the set of sums of $2m$ squares (still inside the cone) would carry a linear functional, the Hankel multiplier, satisfying first and second order conditions. Lemma 5 says such a multiplier has rank larger than its support in the x variables; this either saves a square by induction on m, or leaves one row outside the support, where a second variation built from the alternating symbol contradicts a trace count. 3. New and not new, as the paper says: the statement is new for $m \ge 4$; $m = 2$ follows from [BPSV] and [BSSV, Theorem 1.4], $m = 3$ from [BSSV, Theorem 1.5]; the strategy is that of Blekherman, Dunbar and Sinn [BDS]; the question is asked in [QLC, §9]. 4. The lower bound $2m$ comes from $\mathrm{qp}(\mathbb{P}^{m-1} \times \mathbb{P}^2) = m$ [Mar, Theorem 1] and from the explicit form of Lemma 6.

🛡️ **Evidence & scope**

- **Program:** numerical search: every sampled m×3 form with m = 4, 5, 6 is a sum of 2m squares; the witness F_m is not a sum of 2m − 1 (a control). `supplement.zip: search_width_three.py`
- **Program:** the claims of the two language tabs (N = 60, 78 − 18 = 60, the bound 9 before, F_4 with 8 squares, rank 60 at the regular point),

with two controls that must fail. `lab/comprueba_pestanas.py`

Read this theorem in another mathematical language

Read this theorem in:

Commentary, not part of the paper. Take $m = 4$. Every biquadratic form in $x_1, \dots, x_4, y_1, y_2, y_3$ that is a sum of squares of bilinear forms is a sum of 8 of them. Before, the parallel-line bound gave $4 \cdot 3 - \max(4, 3) + 1 = 9$, so the question was 8 or 9. The form $F_4 = \sum_{i=1}^{4} x_i^2 \left((y_2 - i y_1)^2 + y_3^2 \right)$ is written with its 8 squares and needs all of them (Lemma 6).

Commentary, not part of the paper. The same statement for matrices, $m = 4$: forms correspond to 12×12 symmetric matrices (a 78-dimensional space) modulo the 18-dimensional span of the matrices of the 2×2 minors, and $78 - 18 = 60$ is the dimension of the forms. The theorem says: every positive semidefinite 12×12 matrix can be changed by such a combination of minor matrices into a positive semidefinite matrix of rank at most 8. At the integer point of Lemma 2 the map of 8 squares has full rank 60.

2 Setting

No numbered statement in this section.

It is prose only (no statement, figure or table): read it in the paper (PDF).

3 A regular point and a boundary point

Lemma 2. p. 2

Changed in v2. Before: the ranks checked by a numerical SVD **Now:** checked exactly modulo a prime in the supplement Statements that use it: Theorem 1, Proposition 3, Corollary 4, Corollary 7, Question 1.

> *For every $m \geq 2$ there are $h \in V^{2m}$ with integer coefficients at which $d\Phi_{2m}$ is onto. In particular Σ_{2m} has nonempty interior.*

Isabelle · ours · part **show**

Only a part is in Isabelle: the linear algebra: a Hankel multiplier vanishing on ker[I P Q] is zero (regular_point_hankel); the duality from there to «dΦ onto» is on paper

`regular_point_algebra` · `Isabelle/BSRWidthThree/Regular_Point.thy`, line 151 — the algebraic core: with P cyclic and Q = DPᵀ, the three block symmetries force M = 0; the passage from a Hankel multiplier vanishing on ker C to these symmetries is in regular_point_hankel

```
theorem regular_point_algebra:
  assumes m: "2 \<le> m"
    and sym: "\<forall>i<m. \<forall>j<m. M j i = M i j"
    and h1: "\<forall>i<m. \<forall>j<m. mmul m M (cycP m)
i j = mmul m (mtr (cycP m)) M i j"
    and h2: "\<forall>i<m. \<forall>j<m. mmul m M (cycQ m)
i j = mmul m (mtr (cycQ m)) M i j"
    and h3: "\<forall>i<m. \<forall>j<m. mmul m (mmul m
(mtr (cycP m)) M) (cycQ m) i j
                        = mmul m (mmul m (mtr (cycQ m)) M)
(cycP m) i j"
  shows "\<forall>i<m. \<forall>j<m. M i j = 0"
```

`regular_point_hankel` · `Isabelle/BSRWidthThree/Regular_Point_Hankel.thy`, line 102 — Lemma 2 in its linear-algebra form: a Hankel multiplier (symmetric blocks, H_jl = H_lj) vanishing on ker[I P Q] is zero

```
theorem regular_point_hankel:
  assumes m: "2 \<le> m"
    and hk: "hankel m h"
    and ker: "\<forall>x. (\<forall>i<m. Cx m x i = 0) \
<longrightarrow> (\<forall>j<3. \<forall>i<m. Hx m h x j i
= 0)"
  shows "\<forall>j<3. \<forall>l<3. \<forall>i<m. \
<forall>i'<m. h j l i i' = 0"
```

Proof: in this note · **Formalized:** Isabelle/HOL, by the author (public formalizations repository) · **Covers:** part: the linear algebra: a Hankel multiplier vanishing on ker[I P Q] is zero (regular_point_hankel); the duality from there to «dΦ onto» is on paper · **Cite:** Lemma 2, p. 2, doi:10.5281/zenodo.24051362

↳ Used in: Theorem 1 (p. 3), Proposition 3 (p. 9)

Proof

Identify V with A^3 by the three columns (y_1, y_2, y_3). Let P be the cyclic permutation matrix of size m, $D = \mathrm{diag}(1, \dots, m)$, $Q = DP^{-1}$ and $C = [\,I \;\; P \;\; Q\,] : A^3 \to A$. Take for h the integer basis $(-Pe_i, e_i, 0)$, $(-Qe_i, 0, e_i)$, $1 \le i \le m$, of $\ker C$, so $W_h = \ker C$. By (2) we must show that a symmetric H with (1) and $\ker C \subseteq \ker H$ vanishes. Such an H is $C^{\top} M C$ with M symmetric, with blocks M, MP, MQ, $P^{\top}MP$, $P^{\top}MQ$, $Q^{\top}MQ$. Symmetry of the off-diagonal blocks gives $MP = P^{\top}M$, $MQ = Q^{\top}M$ and $P^{\top}MQ = Q^{\top}MP$. Then $P^{\top}MQ = P^{\top}Q^{\top}M$ and $Q^{\top}MP = Q^{\top}P^{\top}M$, so $(PQ - QP)^{\top}M = 0$, that is $M(PQ - QP) = 0$. Now

> $PQ - QP = PDP^{-1} - D$ is diagonal with entries $d_{\sigma(i)} - d_i = \pm 1$ or $\pm(m-1)$, hence invertible, and $M = 0$.

In plain words (commentary, not part of the paper)

1. There is an explicit choice of $2m$ bilinear forms with integer coefficients at which small changes of the forms move their sum of squares in every direction: the set of sums of $2m$ squares has interior. 2. Why: a direction that no change reaches would be a symmetric matrix M commuting with P and Q in the sense $MP = P^\top M$, $MQ = Q^\top M$, $P^\top MQ = Q^\top MP$, which forces $M(PQ - QP) = 0$; with P cyclic and $Q = DP^{-1}$ the commutator is diagonal and invertible. 3. This replaces the quaternions used for 4×3 in an earlier version of the argument; it works in every size. 4. The lab computes the rank of the differential for $2 \le m \le 6$ modulo 1000003; the scripts check $2 \le m \le 7$, and the control $Q = P$ loses rank.

Evidence & scope

- **Program:** rank of dΦ equal to N for 2 ≤ m ≤ 7; with Q = P the rank drops (m = 4: 56 < 60). `supplement.zip: check_proof_steps.py`

Proposition 3. p. 3

Changed in v2. Now: the exclusion of a singular point with a real factor is argued through the positive definite matrix M_F(x) Statements that use it: Theorem 1, Corollary 4, Corollary 7, Question 1.

> *Suppose $\Sigma_{2m} \neq \Sigma$. Then there are $F \in \Sigma_{2m}$ in the interior of Σ and $\lambda \in R_2^* \setminus 0$, $H = H_\lambda$, such that for every $h \in V^{2m}$ with $\Phi_{2m}(h) = F$:*
>
> 1. *$W_h \subseteq \ker H$ and $\lambda(F) = 0$;*
> 2. *$\sum_i H(D_i, D_i) \geq 0$ for every $D \in V^{2m}$ with $\sum_i h_i D_i = 0$;*
> 3. *H is indefinite;*
> 4. *the h_i have no common zero $([x], [y]) \in X(\mathbb{C})$.*

Proved by: this note, by the argument of Blekherman–Dunbar–Sinn [BDS, Propositions 2.2 and 3.3] · **Covers:** not formalized · **Cite:** Proposition 3, p. 3, doi:10.5281/zenodo.24051362

Classical and new: the paper: «This is the argument of [BDS, Proposition 2.2 and Proposition 3.3]».

↳ Used in: Theorem 1 (p. 3), Corollary 4 (p. 11)

Proof

> This is the argument of [2 , Proposition 2.2 and Proposition 3.3] ; we give it in our setting. Let Ω be the interior of Σ. The set

> $\Omega \cap \Sigma_{2m}$ contains a nonempty open set (Lemma 2 and the submersion theorem), and $\Omega \setminus \Sigma_{2m}$ is open, and nonempty because $\Sigma \setminus \Sigma_{2m}$ is a nonempty relatively open subset of $\Sigma = \overline{\Omega}$. Fix a semialgebraic stratification of R_2 compatible with Σ_{2m} and Ω and satisfying the frontier condition [1 , § 9] . A generic segment in the convex set Ω joining the two open sets avoids every stratum of dimension at most $N-2$ and meets the strata of dimension $N-1$ transversally; where it leaves Σ_{2m} it crosses a stratum $S \subseteq \partial\Sigma_{2m}$ of dimension $N-1$ with Σ_{2m} on one side and its complement on the other. On a small ball U around the crossing point only S and open strata meet, each component of $U \setminus S$ lies inside or outside Σ_{2m}, and there is a smooth ρ on U with $\rho = 0$ on $S \cap U$, $d\rho \neq 0$ and $\rho \geq 0$ on $\Sigma_{2m} \cap U$.
>
> We choose $F \in S \cap U$ outside the set Z of forms in Ω that are singular at a point of $X(\mathbb{C})$. A form in Ω is positive on $X(\mathbb{R})$, and a singular point $p = ([x],[y])$ has no real factor: if $[x]$ is real, take a real representative $x \neq 0$ and write $F(x,y) = y^{\mathsf{T}} M_F(x)\, y$; since F is positive on $X(\mathbb{R})$, $M_F(x)$ is positive definite, and singularity at p would require $M_F(x)y = 0$, impossible also over $\mathbb{C}$ (symmetrically if $[y]$ is real). So p and $\bar{p}$ differ in both factors. After complex changes of coordinates $p, \bar{p}$ become $(e_1, f_1), (e_2, f_2)$; singularity at (e_i, f_i) forces to vanish, in the chart $x_i = y_i = 1$, the coefficients of $x_i^2 y_i^2$, $x_i x_j y_i^2$ $(j \neq i)$ and $x_i^2 y_i y_l$ $(l \neq i)$, which are $m+2$ conditions, and the two lists are disjoint. For real forms this is $2(m+2)$ real conditions, while p moves in $2(m+1)$ real dimensions; hence $\dim Z \leq N-2 < \dim S$.
>
> Put $\lambda = d\rho_F$. For a preimage h, $\rho \circ \Phi_{2m} \geq 0$ near h and vanishes at h, so it has a local minimum there: the first order condition is $\lambda \circ d\Phi_{2m}(h) = 0$, which is (1) by (2), and $\lambda(F) = \sum_i H(h_i, h_i) = 0$. For D in the kernel of $d\Phi_{2m}(h)$ the second derivative of $\rho \circ \Phi_{2m}$ along D is $d\rho_F(d^2\Phi_{2m}[D,D]) = 2\sum_i H(D_i, D_i)$, which is nonnegative at a minimum: this is (2). If H or $-H$ were positive semidefinite, λ or $-\lambda$ would be nonnegative on Σ and vanish at the interior point F, so $\lambda = 0$; this proves (3). A common complex zero of the h_i is a singular point of F, excluded by $F \notin Z$: this is (4).

💡 **In plain words (commentary, not part of the paper)**

1. If not every sum of squares were a sum of $2m$ squares, the set of sums of $2m$ squares would have an edge inside the cone. At a generic point F of that edge there is a linear functional λ, vanishing on the directions the map can

reach and nonnegative on the second order directions; it is the Hankel multiplier H. 2. Why: the edge is locally a smooth hypersurface with the sums of $2m$ squares on one side; its defining function, composed with the map, has a minimum at every preimage of F, and the first and second derivatives give (1) and (2). A semidefinite H would make $\pm\lambda$ a positive functional vanishing at an interior point, so H is indefinite. 3. This is the argument of Blekherman, Dunbar and Sinn [BDS, Propositions 2.2 and 3.3]; the count that avoids complex base points ($2(m+2)$ real conditions against $2(m+1)$ parameters) is written here for $X = \mathbb{P}^{m-1} \times \mathbb{P}^2$. 4. Not formalized: it uses semialgebraic stratification; in the planned Isabelle formalization it is the explicit hypothesis.

Corollary 4. p. 3

> *In the situation of Proposition 3, for every representation $F = \sum_{i \le 2m} h_i^2$ the h_i are linearly independent, F is not a sum of fewer than $2m$ squares, $r := \operatorname{rank} H \le m$, and $\operatorname{im} H_{\mathbb{C}}$ contains no nonzero simple tensor $x \otimes y$.*

Proof: in this note · **Covers:** not formalized · **Cite:** Corollary 4, p. 3, doi:10.5281/zenodo.24051362

↳ Used in: Theorem 1 (p. 3)

Proof

> This is [2 , Lemma 3.4 and Proposition 3.3(4)] in our setting. If the h_i were dependent, an orthogonal change of the $2m$ summands (which does not change F) would make $h_{2m} = 0$; then $D = (0, \dots, 0, v)$ satisfies $\sum h_i D_i = 0$ for every v, and Proposition 3(2) gives $H(v,v) \ge 0$ for all v, against (3). A representation with fewer squares, completed with zeros, would be dependent. So $\dim W_h = 2m$ and $\operatorname{im} H \subseteq W_h^{\perp}$ has dimension at most $3m - 2m = m$. Finally $\operatorname{im} H_{\mathbb{C}} \subseteq (W_h)^{\perp}_{\mathbb{C}}$, and $x \otimes y \in (W_h)^{\perp}_{\mathbb{C}}$ means $h_i(x,y) = 0$ for all i, excluded by (4).

In plain words (commentary, not part of the paper)

1. At such an edge point every representation by $2m$ squares uses $2m$ independent forms, so the multiplier, which kills them, has rank at most $3m - 2m = m$; and its image contains no simple tensor, because that would be a common zero of the forms. 2. Why the forms are independent: otherwise a rotation would make one of them zero, and changing only that one is allowed by the first order condition, so the second order condition

would make H semidefinite. 3. These are the two facts the induction uses: $r \le m$ and no simple tensor in the image.

4 The support of a multiplier

Lemma 5. p. 4

Changed in v2. Now: the example with m = 5 is described as attaining equality in Lemma 5 under its hypotheses Statements that use it: Theorem 1, Corollary 7, Question 1.

> *Let $H \neq 0$ be symmetric with (1) and support of dimension a in the first factor. If $\operatorname{im} H_{\mathbb{C}}$ contains no nonzero simple tensor, then* $\operatorname{rank} H \geq a + 1$.

Proof: in this note · **Covers:** not formalized · **Cite:** Lemma 5, p. 4, doi:10.5281/zenodo.24051362

↳ Used in: Theorem 1 (p. 3)

Proof

Restrict to $A' \otimes B$, complexify and write $H(u,v) = J(Tu, Tv)$ with $T\colon A'_{\mathbb{C}} \otimes B_{\mathbb{C}} \to E$ onto, $\dim E = r = \operatorname{rank} H$, and J a nondegenerate symmetric bilinear form on E. Put $T_b x = T(x \otimes b)$. Trivial radical means $\bigcap_b \ker T_b = 0$, and (1) reads

$$J(T_b x, T_c v) = J(T_c x, T_b v). \tag{1}$$

A line gives a simple tensor. If $x \neq 0$ and $T_b x = \beta(b) e$ for a linear form $\beta \neq 0$ and $e \neq 0$, choose c with $\beta(c) = 1$; by (3), $J(e, T_b v) = \beta(b)\, J(e, T_c v)$ for all $v \in A'$, $b \in B$. The left side, as a linear form in $v \otimes b$, lies in $\operatorname{im} H_{\mathbb{C}}$ (T is onto, J nondegenerate) and is nonzero; the right side is the simple tensor $\beta \otimes J(e, T_c \cdot)$. So it suffices to find such an x when $r \le a$.

Assume $r \le a$. Choose b with $s = \operatorname{rank} T_b$ maximal, $U = \operatorname{im} T_b$, $N_0 = \ker T_b$. For every c,

$$T_c(N_0) \subseteq U \cap U^{\perp_J}. \tag{2}$$

> Orthogonality follows from (3) with $x \in N_0$. For $T_c(N_0) \subseteq U$, write T_b as an identity block of size s and zeros; the coefficient of t in an $(s+1)$-minor of $T_b + tT_c$ bordering that block is an entry of $T_c\colon N_0 \to E/U$, and maximality of s forces it to vanish.
>
> If $s = r<a$, then $U = E$, $U \cap U^{\perp_J} = 0$, and (4) gives a common kernel $N_0 \neq 0$, a contradiction. If $s = r = a$, T_b is invertible; complete b to a basis b, c, d. The operators $T_b^{-1}T_c$, $T_b^{-1}T_d$ are self-adjoint for $J(T_b\,\cdot, T_b\,\cdot)$ by (3), and (3) for the pair c, d makes their product self-adjoint as well, so they commute; a common eigenvector $x \neq 0$ gives $T_c x, T_d x \in \mathbb{C}\,T_b x$, a line. If $s<r$, put $R = U \cap U^{\perp_J}$; then $\dim N_0 = a - s \geq r - s \geq \dim R$. If $R = 0$ there is again a common kernel. Otherwise $T_c|_{N_0}, T_d|_{N_0}\colon N_0 \to R$ form a pencil: if $\dim N_0 > \dim R$ every member has a kernel, and if the dimensions agree the binary form $\det(\sigma T_c + \tau T_d)|_{N_0}$ is zero or of positive degree, so it vanishes at some $[\sigma : \tau] \in \mathbb{P}^1(\mathbb{C})$. Either way there is $x \in N_0 \setminus 0$ with $T_b x = 0$ and $T_c x, T_d x$ dependent, not both zero; they span a line. In all cases the first paragraph gives a contradiction.

💡 **In plain words (commentary, not part of the paper)**

1. A Hankel multiplier whose image avoids the simple tensors cannot be small: its rank is at least one more than the number a of x directions it actually uses. 2. Why: if $r \leq a$, linear algebra over $\mathbb{C}$ produces a vector x whose three images $T_y x$ lie on one line, and the Hankel symmetry turns that line into a simple tensor in the image. The three cases are: a common kernel; commuting self-adjoint operators with a common eigenvector; or a square pencil that is singular at some point of $\mathbb{P}^1(\mathbb{C})$. 3. With Corollary 4 this gives $a + 1 \leq r \leq m$, the inequality that drives the induction; in [BDS] this place is taken by the classification of Gorenstein ideals. 4. A numerical search over Hankel matrices of rank at most four in size 4×3 found no counterexample, and found the case rank four, support three, no simple tensor, so the last case of the induction is not empty.

🛡️ **Evidence & scope**

- **Program:** search for counterexamples among Hankel matrices of rank ≤ 4 (4×3): none; the case rank 4, support 3, no simple tensor occurs. `supplement.zip: search_support_lemma_4x3.py`

5 Induction and second variation

No numbered statement in this section.

It is prose only (no statement, figure or table): read it in the paper, p. 4.

6 The lower bound

Lemma 6. p. 5

> *The form* $F_m = \sum_{i=1}^{m} x_i^2\big((y_2 - iy_1)^2 + y_3^2\big)$ *is a sum of* $2m$ *squares of bilinear forms and of no fewer.*

Isabelle · ours · part **show**

Only a part is in Isabelle: the witness with slopes 0, …, m−1 (F_wit): its 2m squares and the lower bound 2m; the paper's F_m uses slopes 1, …, m, and the change $y_2 \mapsto y_2 - y_1$ that carries one to the other (with the conservation of the SOS length) is not formalized

`wit3_sos` · `Isabelle/BSRWidthThree/BSR_Width3_Lower.thy`, line 93 — F_m is the sum of its 2m displayed squares

```
theorem wit3_sos: "sos_form m 3 (2 * m) wit3 x y = F_wit m
x y"
```

`wit3_length` · `Isabelle/BSRWidthThree/BSR_Width3_Lower.thy`, line 241 — every representation of F_m has at least 2m squares (Bessel inequality on 2m orthonormal vectors)

```
theorem wit3_length:
  assumes rep: "\<forall>x y. sos_form m 3 s c x y = F_wit
m x y"
  shows "2 * m \<le> s"
```

Proof: in this note · **Formalized:** Isabelle/HOL, by the author (public formalizations repository) · **Covers:** part: the witness with slopes 0, …, m−1 (F_wit): its 2m squares and the lower bound 2m; the paper's F_m uses slopes 1, …, m, and the change $y_2 \mapsto y_2 - y_1$ that carries one to the other (with the conservation of the SOS length) is not formalized · **Cite:** Lemma 6, p. 5, doi:10.5281/zenodo.24051362

↳ Used in: Theorem 1 (p. 3)

Proof

> The displayed expression has $2m$ squares. Write any representation as

$$F_m = \Big\| \sum_i x_i M_i y \Big\|^2, \qquad M_i \in \mathbb{R}^{s\times 3}.$$

The coefficient of x_i^2 gives $M_i^\top M_i$ of rank two with kernel spanned by $b_i = (1, i, 0)$. For $i \neq j$ the coefficient of $x_i x_j$ vanishes, so $M_i^\top M_j$ is antisymmetric; it kills b_j on the right and b_i on the left, hence, being antisymmetric, b_i and b_j on the right. An antisymmetric 3×3 matrix with two independent vectors in its kernel is zero. So the images of the M_i are m planes, pairwise orthogonal in $\mathbb{R}^s$, and $s \geq 2m$.

7 Larger widths

Corollary 7. p. 5

🔄 **Changed in v2. Now:** new in version 2 (Section 7): BSR(m,4) ≤ 3m for m ≥ 4, so 11 ≤ BSR(4,4) ≤ 12; the bound mn − max(m,n) is new only in width four (Figure 1) Statements that use it: Question 1.

> *For $m \geq 4$, $\mathrm{BSR}(m, 4) \leq 3m$. In particular $11 \leq \mathrm{BSR}(4, 4) \leq 12$.*

Isabelle · ours · part **show**

Only a part is in Isabelle: the implication only: Theorem 1 enters as a hypothesis (bsr_width_four, bsr_upper_three), and Theorem 1 is itself only partly formalized; the comparison identity is fully formalized

`bsr_width_four` · `Isabelle/BSRWidthThree/Width_Four.thy`, line 89 — Corollary 7 from Theorem 1 (taken as hypothesis): BSR(m,3) ≤ 2m implies BSR(m,4) ≤ 3m

```
corollary bsr_width_four:
  assumes Theorem_1: "bsr_le m 3 (2 * m)"
  shows "bsr_le m 4 (3 * m)"
```

`bsr_upper_three` · `Isabelle/BSRWidthThree/Width_Four.thy`, line 96 — the general bound mn − max(m,n) for m, n ≥ 3, from Theorem 1 for every number of rows

```
theorem bsr_upper_three:
  assumes Theorem_1: "\<And>p. 3 \<le> p \<Longrightarrow>
bsr_le p 3 (2 * p)"
    and m: "3 \<le> m" and n: "3 \<le> n"
  shows "bsr_le m n (m * n - max m n)"
```

`bssv_wins_from_five` · `Isabelle/BSRWidthThree/Width_Four.thy`, line 129 — the identity of the comparison and its sign: for min(m,n) ≥ 5 the bound of [BSSV, Theorem 2.1] is already at most mn − max(m,n)

```
lemma bssv_wins_from_five:
  fixes M s :: int
  assumes s: "5 \<le> s" and M: "3 \<le> M"
  shows "M * (M + 1) * s * (s + 1) < 2 * (M * (s - 1) + 1)
* (M * (s - 1) + 2)"
```

`width_four_wins` · `Isabelle/BSRWidthThree/Width_Four.thy`, line 147 — in width four [BSSV, Theorem 2.1] allows 3m + 1: Corollary 7 is better

```
lemma width_four_wins:
  fixes M :: int
  assumes M: "2 \<le> M"
  shows "2 * (3 * M + 1) * (3 * M + 2) < M * (M + 1) * 4 *
5"
```

Proof: in this note · **Formalized:** Isabelle/HOL, by the author (public formalizations repository) · **Covers:** the implication only (Theorem 1 enters as a hypothesis (bsr_width_four, bsr_upper_three), and Theorem 1 is itself only partly formalized; the comparison identity is fully formalized) · **Cite:** Corollary 7, p. 5, doi:10.5281/zenodo.24051362

🆕 **Classical and new:** the paper: the previous upper bound in width four was the covering bound 3m + 1 (for m = 4 also [BSSV, Theorem 2.1], 13); for n ≥ 5 the general bound mn − max(m,n) is not better than [BSSV, Theorem 2.1]. The lower bound 11 is [QC, Theorem 3.1].

↳ Used in: 1 (p. 18)

Proof

> We prove more generally $\mathrm{BSR}(m,n) \le mn - \max(m,n)$ for $m,n \ge 3$; for $n = 4 \le m$ this is $3m$. This is the argument of [7 , Theorem 9] with Theorem 1 in place of width two. Let $m \ge n$ and $f = \sum_{k \le r} b_k^2$; let $v_{ij} \in \mathbb{R}^r$ be the coefficient vector of $x_i y_j$ in $(b_1, \dots, b_r)$. The vectors v_{ij} with $j \ge 4$ span a space W of dimension at most $m(n-3)$. After an orthogonal change of the r coordinates, which does not change f, W lies in the first $\dim W$ of them, and the remaining coordinates give bilinear forms in x and y_1, y_2, y_3 only. Their sum of squares is an $m \times 3$ biquadratic sum of squares, hence a sum of $2m$ squares by Theorem 1. So f is a sum of $m(n-3) + 2m = mn - m$ squares. For $m<n$ exchange the roles of x and y. The lower bound $\mathrm{BSR}(4,4) \ge 11$ is [8 , Theorem 3.1] .

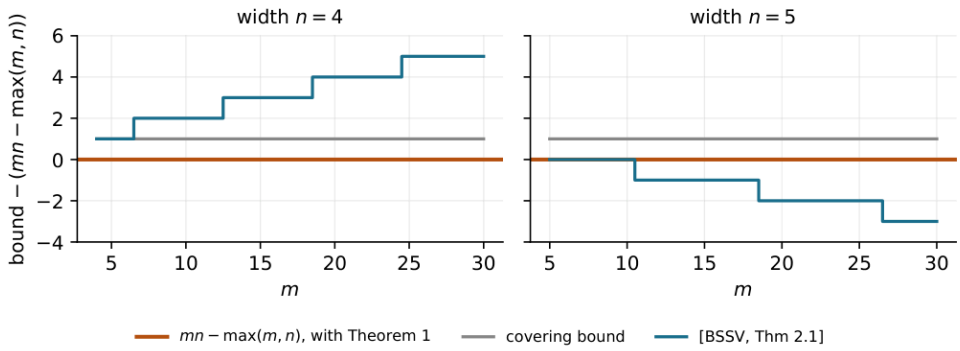
In plain words (commentary, not part of the paper)

1. Every $m \times 4$ biquadratic sum of squares is a sum of at most $3m$ squares; for 4×4 this is 12, and Qi and Cui showed that 11 are sometimes needed, so $11 \leq \mathrm{BSR}(4,4) \leq 12$. 2. Why: in any representation, rotate the squares so that the coefficients of y_4 live in at most m of them. The other squares do not see y_4: they form an $m \times 3$ sum of squares, which Theorem 1 rewrites with $2m$ squares. Total $m + 2m = 3m$. 3. The same rotation gives $mn - \max(m,n)$ for all $m, n \geq 3$. It is new only when $\min(m,n) = 4$: for $\min(m,n) = 3$ it is Theorem 1 itself, and for $\min(m,n) \geq 5$ the bound of [BSSV, Theorem 2.1] is as good or better (for 6×6: 29 against 30; but for 4×5 ours gives 15 against 16). The lab draws the three bounds. 4. The rotation is the one of our earlier note [Mar, Theorem 9], which used width two and gave $3m + 1$.

Evidence & scope

- **Program:** where mn − max(m,n) beats the covering bound and [BSSV, Theorem 2.1] for $3 \leq n \leq 7$, $m \leq 40$ (only n = 3, where it is Theorem 1, and n = 4); controls 13 for 4×4 and 29 for 6×6. `supplement.zip: compare_upper_bounds.py`

Figure 1



Upper bounds for $\mathrm{BSR}(m,n)$ minus $mn - \max(m,n)$, the bound obtained with Theorem 1: the covering bound of [4 , Corollary 2.15] and the bound of [4 , Theorem 2.1] . In width four both lie above for every $m \geq 4$ (Corollary 7); in width five, the BSSV bound equals $mn - \max(m,n)$ for $5 \leq m \leq 10$ and is lower for every $m \geq 11$.

8 Checks and formal status

No numbered statement in this section.

It is prose only (no statement, figure or table): read it in the paper (PDF).

The companion site

No numbered statement in this section.

It is prose only (no statement, figure or table): read it in the paper, p. 7.

Questions

Question 1. p. 7

> 🔄 **Changed in v2. Now:** the closing question bounds the rank of a multiplier whose kernel contains k linearly independent summands by 4m − k

Does the method give $\mathrm{BSR}(m,4)$? For $m \geq 4$, our earlier note and Corollary 7 give $\lfloor 8m/3 \rfloor \leq \mathrm{BSR}(m,4) \leq 3m$. In width four, if the kernel of the multiplier H contains k linearly independent summands, its rank is at most $4m - k$, and the second variation would have to come from a structure replacing the alternating symbol on $\mathbb{R}^3$.

Cite: Question 1, p. 7, doi:10.5281/zenodo.24051362

↳ Not used by another result of the paper.

💡 **In plain words (commentary, not part of the paper)**

1. In width four the value is open: $\lfloor 8m/3 \rfloor \leq \mathrm{BSR}(m,4) \leq 3m$ for $m \geq 4$, and $11 \leq \mathrm{BSR}(4,4) \leq 12$. 2. The method of this note would need two things in width four: a rank inequality for the multiplier, and a second variation. In width three the second variation came from the alternating symbol on $\mathbb{R}^3$, which has no analogue of the same kind on $\mathbb{R}^4$.

Acknowledgement

No numbered statement in this section.

It is prose only (no statement, figure or table): read it in the paper (PDF).

References

1

J. Bochnak, M. Coste, M.-F. Roy, Real Algebraic Geometry, Springer, 1998. 🔎 find it

2

G. Blekherman, A. Dunbar, R. Sinn, Pythagoras numbers for ternary forms, arXiv:2410.17123v2.

3

G. Blekherman, D. Plaumann, R. Sinn, C. Vinzant, Low-rank sum-of-squares representations on varieties of minimal degree, Int. Math. Res. Not. IMRN 2019, no. 1, 33–54. 🔎 find it

4

G. Blekherman, R. Sinn, G. G. Smith, M. Velasco, Sums of squares and quadratic persistence on real projective varieties, J. Eur. Math. Soc. 24 (2022), 925–965; statement numbers as in arXiv:1902.02754v2.

5

G. Blekherman, R. Sinn, M. Velasco, S. Zhang, Spurious local minima in nonconvex sum-of-squares optimization, arXiv:2411.02208.

6

J. Löfberg, L. Qi, Second order Zarankiewicz number, arXiv:2608.30555v3.

7

C. Marín, Quadratic persistence of Segre varieties and a linear gap between biquadratic SOS rank and Zarankiewicz bounds, Zenodo, doi:10.5281/zenodo.24036320, 2026.

8

L. Qi, C. Cui, The second-order augmented Zarankiewicz number, arXiv:2610.04448v1.

9

L. Qi, J. Löfberg, Y. Chen, Exact and asymptotic values for weak limited augmented Zarankiewicz numbers in the $m \times 3$ case, arXiv:2608.06050v3.

10

Y. Xu, C. Cui, L. Qi, Narrowing the gap: SOS ranks of 4×3 biquadratic forms and a lower bound of 8, arXiv:2602.21570.