Why three blocks? The limited augmented Zarankiewicz number in fixed width

Extremal combinatorics · Zarankiewicz problems · sums of squares of biquadratic forms

Carles Marín

Statements

Theorem A 8
Theorem B 10
Lemma 4.1 — blocks 12
Lemma 5.1 — vertical pairs, any n 15
Lemma 5.2 — one full row, any n 15
Proposition 6.1 19
Theorem C 20
Corollary 6.2 22
Proposition 6.3 23
Theorem D 25
Theorem E 30
Theorem F 31
Proposition 8.1 32
Question 1 37
Question 2 38
Question 3 38
Question 4 39
Question 5 39
Question 6 39

At a glance

What the paper contributes

The limited augmented Zarankiewicz number $z_L(m,n)$ of Qi, Cui and Xu is a lower bound for the largest sum-of-squares rank of $m \times n$ biquadratic forms. The paper follows one question: with the number n of columns fixed, how many pairs does each new row buy?

- Three columns, exactly: $z_L(m,3) = \lfloor 3m/2 \rfloor + 2$ for every $m \geq 3$ (Theorem A), answering the question of Qi, Löfberg and Chen; proof formalized in Lean. The gap to their weak variant grows like $m/6$.
- Every width: $F_n(m) = z_L(m,n) - z(m,n) = \rho_n m + O_n(1)$ with ρ_n rational, the optimum of finitely many linear programs (Theorem C), and the error is eventually periodic (Theorem F).
- Pairs around a hollow triangle of three column blocks give $\rho_n \geq \lambda_n = n/3 - 1/2 + O(1/n)$; balanced blocks are optimal inside that family (Proposition 6.3), and $\rho_n = \lambda_n$ for $n \leq 6$ by rational dual certificates checked in exact arithmetic (Theorem B). Equality for every n is a conjecture.
- Consequences and companions: a linear gap to the recursive-line, signed and second-order numbers in every width (Corollary 6.2); the slope $\lfloor (n-1)^2/4 \rfloor/(n-1)$ of Lebedev's canonical base for every n (Theorem D); $z_L(m,4)$ for every $m \geq 6$ by certified computation (Theorem E); fourteen new exact values; and a three-dimensional analogue where the naive bound of one quarter already fails in $3 \times 3 \times 5$ (Proposition 8.1).

The admissibility lemmas behind the constructions are formalized in Isabelle/HOL (submitted to the Archive of Formal Proofs). Version 1 on Zenodo: doi:10.5281/zenodo.23778502.

🚀 **Start here: Theorem A**: three columns: z_L(m,3) = ⌊3m/2⌋ + 2 for every m ≥ 3.

Main results:

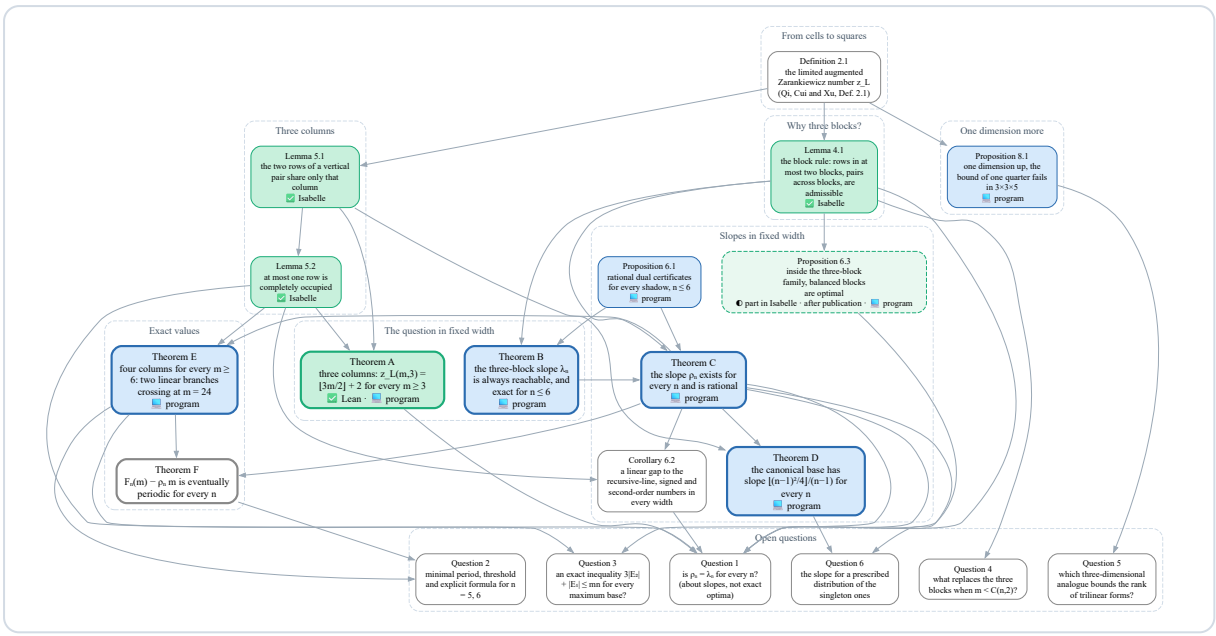
- **Theorem A**: three columns: z_L(m,3) = ⌊3m/2⌋ + 2 for every m ≥ 3
- **Theorem B**: the three-block slope λ_n is always reachable, and exact for n ≤ 6
- **Theorem C**: the slope ρ_n exists for every n and is rational
- **Theorem D**: the canonical base has slope ⌊(n−1)²/4⌋/(n−1) for every n
- **Theorem E**: four columns for every m ≥ 6: two linear branches crossing at m = 24
- **Theorem F**: $F_n(m) - \rho_n$ m is eventually periodic for every n

20 results | ✅ **1** entirely in Lean | ◐ **1** partly in Lean | ✅ **3** in Isabelle | 💻 **8** 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, §§1–3, 5, 6)

Write $z(m,n)$ for the maximum number of ones; when $m \geq \binom{n}{2}$ it is

$$z(m,n) = m + \binom{n}{2},$$

attained by one row for each pair of columns and single ones in all other rows; every extremal matrix has this form. (§1)

Definition 2.1 [28 , Def. 2.1] . Let E_1 be a C_4-free set of cells with $|E_1| = z(m,n)$, and let E_2 be a family of unordered pairs of distinct cells outside E_1, no cell in two pairs; put $O = E_1 \cup \bigcup E_2$. The family is *admissible* if every pair $\{(i,j),(p,q)\}$ satisfies

1. if $i \neq p$ and $j \neq q$, the opposite cells $(i,q),(p,j)$ are not both in O;
2. no cell (k,ℓ) with $k \notin \{i,p\}$, $\ell \notin \{j,q\}$ has $(k,\ell),(k,j),(k,q),(i,\ell),(p,\ell) \in O$.

The *limited augmented Zarankiewicz number* $z_L(m,n)$ is the maximum of $|E_1| + |E_2|$. (§2)

For a pair, let B be the set of rows other than i,p occupied in both columns j,q, and C the set of columns other than j,q occupied in both rows i,p; then (C3) $\iff$ the rectangle $B \times C$ is completely empty. (§2)

If n is fixed and m grows, how many pairs does each new row buy? Put $F_n(m) = z_L(m,n) - z(m,n)$. (§3)

For $m \geq 3$ a maximum base consists of the rows $\{1,2\},\{1,3\},\{2,3\}$ and $m-3$ rows with a single one. For a row r let S_r be its occupied support; the cells of S_r outside E_1 are its *endpoints*, and $2|E_2|$ is the number of endpoints. (§5) “Rows 1–10 are the special rows (two grey cells each) and rows 11–22 singleton rows” (§4, caption of Figure 4).

Each row keeps its original support as a *label* … the *shadow* $\mathcal{H}$, the set of triples of columns contained in some frequent support. (§6.2)

Abstract.

The limited augmented Zarankiewicz number $z_L(m,n)$ of Qi, Cui and Xu bounds the maximum sum-of-squares rank of $m \times n$ biquadratic forms from below: it counts the cells of a maximum C_4-free set in the $m \times n$ grid together with pairs of further cells that avoid two forbidden patterns. We determine $z_L(m,3) = \lfloor 3m/2 \rfloor + 2$ for every $m \geq 3$, answering a question of Qi, Löfberg and Chen, with a proof formalized in Lean; it also shows that the gap to their weak variant grows like $m/6$. For a fixed number n of columns we prove that each new row adds ρ_n pairs up to a bounded error, where ρ_n is a rational number given by finitely many linear programs, and that the error is eventually periodic. Pairs placed around a hollow triangle of three column blocks give $\rho_n \geq \lambda_n = n/3 - 1/2 + O(1/n)$; balanced blocks are optimal within this family, and $\rho_n = \lambda_n$ for $n \leq 6$ by rational dual certificates checked in exact arithmetic; we conjecture equality for all n. Consequently, in every fixed width the original criterion stays linearly below the recursive-line, signed and second-order ones. We also reduce Lebedev's canonical base to a graph parameter whose sharp bound is due to Hof, which gives its slope $\lfloor (n-1)^2/4 \rfloor / (n-1)$ for every n; we find the exact value of $z_L(m,4)$ for every $m \geq 6$ by certified computation, fourteen new exact values by exhaustive computer search (eight of them instances of the three- and four-column formulas), and a three-dimensional analogue in which the naive bound of one quarter fails already in $3 \times 3 \times 5$, the smallest such box. The admissibility lemmas behind the constructions are formalized in Isabelle/HOL.

This paper.

Version 2, doi:10.5281/zenodo.23860725; all versions: doi:10.5281/zenodo.23778502. Changes from version 1: the sharp bound on the graph parameter in Theorem D is due to Hof [18 , 17] , and the theorem's contribution is the reduction to it; the integer program over row types is exact for $n \leq 4$ only, and no value of the table rests on it for $n \geq 5$ (Section 7); eight of the fourteen new values are instances of Theorems A and E; in Question 1, three-block shadows *attain* the optimal value for $n \leq 6$, which is not a classification of the optimal programs. The supplement's type program now refuses $n \geq 5$.

Companion site.

https://karlesmarin.github.io/blueprints/zl-three-blocks/: every statement with its dependencies, its formal proofs and its checks, and labs in the browser that redraw Figures 2–7 for values the reader chooses (Section 10).

1 A question about zeros and ones

How many ones can an $m \times n$ matrix of zeros and ones contain without four of them at the corners of a rectangle? In 1951 Zarankiewicz asked, for square matrices, how many ones

force a $k \times k$ submatrix of ones [24] ; the matrix is one “consisting exclusively of 0’s and 1 ’s”, as Kővári, Sós and Turán put it when they proved the upper bound of order $n^{3/2}$ for 2×2 three years later [14] . Reiman sharpened the bound and observed that finite projective planes attain it [22] , and Guy computed tables of exact values [12 , 13] ; the history of these *degenerate* extremal problems is told by Füredi and Simonovits [9] . When one side is long compared with the other the problem becomes exactly solvable; a classical formula of this kind is due to Čulík [7] (as quoted in [2]), and recent work of Chen, Horsley and Mammoliti determines the values near the triple-system threshold [2 , 3] . For this note one regime matters. Write $z(m,n)$ for the maximum number of ones; when $m \geq \binom{n}{2}$ it is

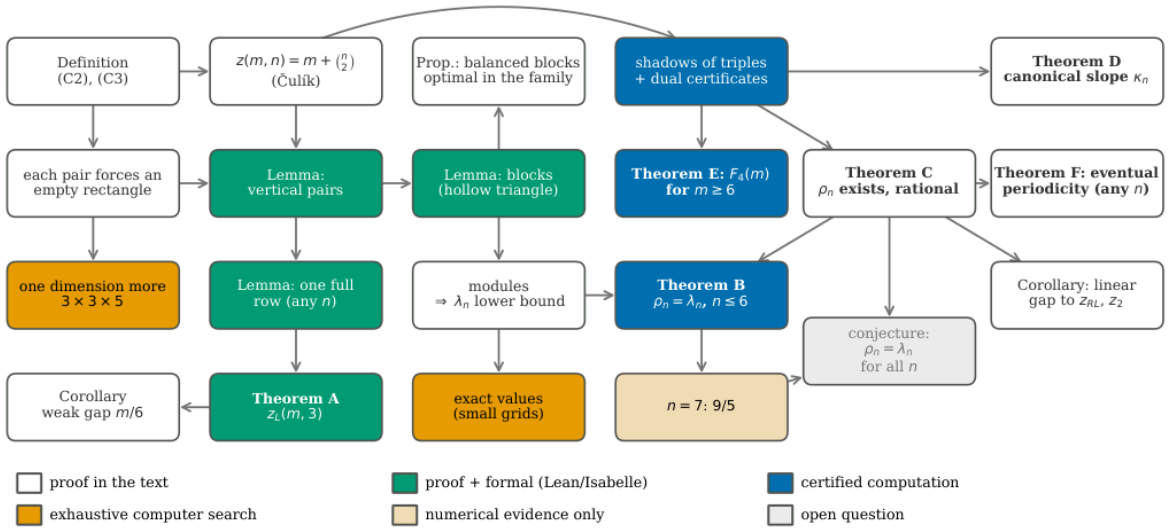
$$z(m,n) = m + \binom{n}{2},$$

attained by one row for each pair of columns and single ones in all other rows; every extremal matrix has this form. Indeed two rows share at most one column, so the $\binom{d_r}{2}$ pairs of columns covered by the rows are distinct, $\sum_r \binom{d_r}{2} \leq \binom{n}{2}$, and $d \leq 1 + \binom{d}{2}$ for every integer $d \geq 0$, with equality only for $d = 1, 2$.

Seventy-five years after Zarankiewicz, the same matrices returned from an unexpected side. In 2026 Qi, Cui and Xu noticed that a C_4-free matrix is a certificate in real algebraic geometry: it forces many squares in a sum of squares of bilinear forms (Section 2). To push the certificate further they let the matrix carry *pairs* of extra cells, and asked how many pairs fit. That is the number $z_L(m,n)$ of this paper. Within a few months a series of papers had computed small values and bounds, and Lebedev had found the density one third when both sides grow (Section 3).

This paper follows one question through to its answer: *if the number of columns is fixed and rows keep coming, how many pairs does each row buy?* For three columns the answer is exactly one half (Theorem A, with a Lean proof). In general a pair must be paid for with empty cells, and the cheapest payment is arranged around a triangle of three blocks of columns whose face stays empty (Section 4). This gives an explicit slope λ_n for every n (Theorem B); certified linear programs show that nothing does better for $n \leq 6$, and the same programs prove that the slope always exists and is rational (Theorem C, Section 6). An exact finite description by profiles of triples then settles four columns for every $m \geq 6$ (Theorem E) and shows that the error term is eventually periodic in every width (Theorem F); for Lebedev’s canonical base the slope is found for every n (Theorem D, a reduction to a known bound for graphical simple games). Exact values and a three-dimensional variant close the paper. Figure 1 shows how the statements depend on each other and what each one rests on.

Figure 1



A map of the paper. Each box is a statement and an arrow means "is used in". The colour says what the statement rests on: a proof in the text; a proof also formalized (Theorem A in Lean; the vertical-pair, full-row and blocks lemmas in Isabelle/HOL, see Formalizations at the end); a computation certified in exact arithmetic; an exhaustive computer search; numerical evidence only; or an open question. Theorems A–F are in bold.

2 From cells to squares

Why should a combinatorialist care about sums of squares? A biquadratic form in $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$ that is nonnegative need not be a sum of squares of bilinear forms: Calderón showed that it is when $m = 2$ or $n = 2$ [1] , and Choi gave a 3×3 form that is not [5] ; see [6] for the wider picture. Among forms that *are* sums of squares one asks for the least number of squares needed, and for its maximum $\mathrm{BSR}(m, n)$ over all such forms, of which only small cases and bounds are known [27] . Write a cell (i, j) as the bilinear monomial $x_i y_j$. A rectangle of four cells must be avoided, because its four squares are only two: $x_i^2 y_j^2 + x_i^2 y_q^2 + x_p^2 y_j^2 + x_p^2 y_q^2 = (x_i y_j + x_p y_q)^2 + (x_i y_q - x_p y_j)^2$. Qi, Cui and Xu [27] show that a C_4-free set of cells gives a sum of squares $\sum x_i^2 y_j^2$ that cannot be written with fewer squares, and that one can do better: a square $(x_i y_j + x_p y_q)^2$ uses two further cells at once, and under two local conditions the enlarged sum is still irreducible. This gives the chain

$$\mathrm{BSR}(m, n) \;\geq\; z_A(m, n) \;\geq\; z_L(m, n) \;\geq\; z(m, n)$$

of lower bounds for the maximum sum-of-squares rank of biquadratic forms; z_A is defined like z_L below but without asking the base E_1 to be maximum.

Definition 2.1. ([28 , Def. 2.1]) p. 3

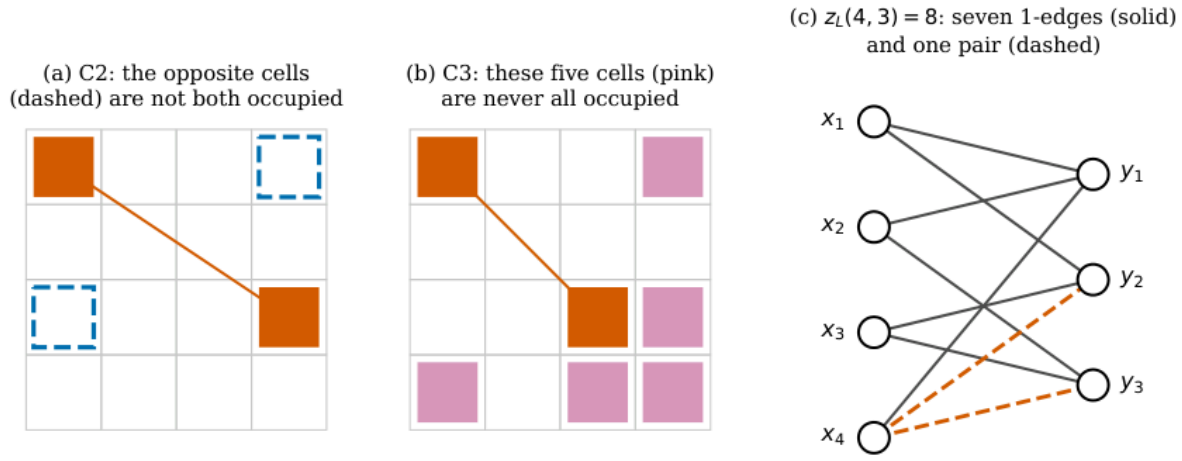
Let E_1 be a C_4-free set of cells with $|E_1| = z(m,n)$, and let E_2 be a family of unordered pairs of distinct cells outside E_1, no cell in two pairs; put $O = E_1 \cup \bigcup E_2$. The family is *admissible* if every pair $\{(i,j),(p,q)\}$ satisfies

1. if $i \neq p$ and $j \neq q$, the opposite cells $(i,q),(p,j)$ are not both in O;
2. no cell (k,ℓ) with $k \notin \{i,p\}$, $\ell \notin \{j,q\}$ has $(k,\ell),(k,j),(k,q),(i,\ell),(p,\ell) \in O$.

The *limited augmented Zarankiewicz number* $z_L(m,n)$ is the maximum of $|E_1| + |E_2|$.

↳ Used in: Lemma 4.1 (p. 12), Lemma 5.1 (p. 15), Proposition 8.1 (p. 32)

Figure 2



The definition at a glance. (a) Condition (C2): for the pair (orange), the two opposite cells (dashed) are not both occupied. (b) Condition (C3): with the pair in orange, the five pink cells, lying in one further row k and one further column ℓ, are never all occupied. (c) The case 4×3 as a bipartite graph, as in [27] : rows x_i, columns y_j; the seven cells of a maximum base (rows $\{1,2\},\{1,3\},\{2,3\},\{1\}$) as solid edges and the single pair, which fills row 4, as two dashed edges; so $z_L(4,3) = 8$.

Why does exactly one pair fit in 4×3*?* The base of Figure 2(c) has rows $\{1,2\},\{1,3\},\{2,3\}$ and $\{1\}$. The pair fills the last row completely; it is admissible because no other row is full, so no witness for (C3) exists. Any second pair would complete another row, which would violate (C3) for the horizontal pair shown: this is the whole idea of the three-column theorem below.

3 The question in fixed width

The series moved fast. The first paper (March 2026) introduced z_A and z_L and computed all values with $m,n \leq 4$ [27] ; the second (April) built lower bounds from complete graphs and settled 5×3, 5×4 and 6×4 [28] ; the third (May) turned the problem into integer programs, with a table up to 7×7 and $z_L(10,5) = 26$ [31] ; a Zenodo supplement of

Howard-Stanley documents $z_L(6,6) = 20$ [25] ; Liu and Song extended the complete-graph constructions to blocks K_{5t} [19] . A weak variant with acyclic dependencies followed in July [29] , and its $m \times 3$ paper computed the original $z_L(m,3)$ for $m \le 13$ by exhaustive search and posed the value for all m as an open problem [30 , § 9] . In parallel, Löfberg and Qi studied a second-order number z_2 and found exact values in three and four columns [20] : fixed width is where exact answers appear first. When both dimensions grow, Lebedev [26] proved that, along sequences with $n \to \infty$ and $m/n^2 \to c \in (0,\infty)$, the density $z_L(m,n)/(mn)$ tends to $1/3$ if and only if $c \ge 1/12$ [26 , Thms. 14–15] , with a stability theorem: near-optimal configurations look like three diagonal blocks left empty [26 , Thm. 5] . His lower bounds come from a construction with three blocks of rows and columns and reciprocal pairs, first with an empty base and then with a maximum one [26 , Thms. 4 and 13] . For a *fixed* number of columns these results leave the linear term in m open.

If n is fixed and m grows, how many pairs does each new row buy? Put $F_n(m) = z_L(m,n) - z(m,n)$.

Theorem A. Lean ✓ p. 4

> *For every $m \ge 3$,*
>
> $$z_L(m,3) = m + 3 + \left\lfloor \frac{m-2}{2} \right\rfloor,$$
>
> *and every maximum base admits at most $\lfloor (m-2)/2 \rfloor$ pairs. The proof is formalized in Lean 4.*

Lean · ours **show**

- ZL.zL_three
- ZL.zL_le
- ZL.card_E2_le
- ZL.no_two_full
- ZL.z_three
- ZL.exists_good
- ZL.exists_odd
- ZL.exists_even

✅ Checked when this page was built: LimitedAugmentedThree.lean compiles with Lean 4.35.0, no errors, no sorry. Axioms — zL_three, zL_le, card_E2_le, no_two_full, z_three, exists_good, exists_odd, exists_even: propext, Classical.choice, Quot.sound.

Reproduce: Lean 4.35.0, Mathlib f0b4c8150aa0; in that Lean project, `lake env lean LimitedAugmentedThree.lean` (SHA-256 of the compiled file: `b3a1168f55b462ae`…; a source page on this site may show it with comments edited, under its own hash).

Trust base: 🔒 kernel only. Lean: 8 results, standard axioms only

Proof: in this note · **Formalized:** Lean 4, by the author (public formalizations repository) · **Covers:** the whole statement · **Cite:** p. 4, doi:10.5281/zenodo.23778502

🆕 **Classical and new:** the paper: «previously known for m ≤ 13» (Card A), posed for all m as an open problem in [QLC, §9]; it also settles the second half of that problem (the gap to the weak number grows like m/6).

↳ Used in: 1 (p. 37)

Proof

A row that is not full has at least one empty cell, so a singleton row carries at most one endpoint and a special row none, except the single full row of Lemma 5.2, which carries at most two. Hence $2|E_2| \le (m-3)+1 = m-2$ for every maximum base. Equality is reached: for m odd pair the singleton rows (all with their one in column 1) two by two as $\{(r,2),(r',3)\}$. For m even take one singleton row t with its one in column 2 and the pair $\{(s,1),(t,3)\}$, where s is the special row $\{2,3\}$, which becomes the full row; split the other $m-4$ singleton rows evenly between ones in columns 1 and 2, and join them by vertical pairs in column 3. For a vertical pair C is empty; for the other pairs $B \times C$ lies in column 1 (odd case) or column 2 (the last pair), where every row of B is empty, and (C2) holds because at least one opposite cell is free; so the families are admissible.

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

Take a $0/1$ matrix with three columns and m rows. First place as many ones as possible with no four of them at the corners of a rectangle: that is $m+3$ ones (one row for each pair of columns, a single one in every other row). Then add *pairs* of extra cells, each pair obeying the two local rules (C2) and (C3). The theorem says that exactly $\lfloor (m-2)/2 \rfloor$ pairs fit, so each new row buys half a pair.

Why: two rows that both fill all three columns always create a forbidden pattern, so at most one row is full (Lemma 5.2); every other row keeps an empty cell, and a row with a single one of the base then carries at most one endpoint of a pair. That caps the endpoints at $m-2$. Explicit families reach the cap, for m odd and for m even.

The paper reports the value as previously known only for $m \le 13$, by exhaustive search [30 , § 9] , where it is posed as an open problem; the whole proof, upper bound for every maximum base and the families, is checked by Lean (module `LimitedAugmentedThree`).

🛡️ **Evidence & scope**

- **Lean:** the definitions of §2 and the whole proof, upper bound for every maximum base and both families; standard axioms only. `Formalizations.Combinatorics.Zarankiewicz.LimitedAugmentedThree (commit 389a576)`
- **Program:** the values for m ≤ 14 by exhaustive search (nauty classes + CP-SAT), equal to the formula. `supplement: zl_nauty.py, zl_cpsat_dlex.py (output: table_zL.json)`
- **Program:** the three language tabs: $F_3(m) = \lfloor (m-2)/2 \rfloor$ for m = 3..7 by an exhaustive search of this blueprint, the closed form for m ≤ 10 000, the base and pair of Figure 2(c),

the rectangle identity in exact arithmetic; each with a control that must fail (29-oct: passes). `lab/comprueba_pestanas.py`

Read this theorem in another mathematical language

Read this theorem in:

Commentary, not part of the paper. An $m \times 3$ matrix of zeros and ones with no four ones at the corners of a rectangle has at most $m + 3$ ones. Start from such a matrix and add *pairs* of further ones, each pair obeying (C2) and (C3). Exactly $\lfloor (m-2)/2 \rfloor$ pairs fit, whatever the starting matrix: $0, 1, 1, 2, 2$ for $m = 3, \ldots, 7$ (each new row buys half a pair), and the total is $z_L(m, 3) = \lfloor 3m/2 \rfloor + 2$.

Commentary, not part of the paper. Rows are vertices $x_1, \ldots, x_m$, columns are y_1, y_2, y_3, a one is an edge $x_i y_j$. The base is a bipartite graph with no 4-cycle, $m + 3$ edges. For $m = 4$ (Figure 2(c)) the base has the seven edges of the rows $\{1,2\}, \{1,3\}, \{2,3\}, \{1\}$, and one pair of new edges, $x_4 y_2$ and $x_4 y_3$, fills x_4; a second full vertex would always create a forbidden pattern. In general at most one row vertex reaches all three columns.

Commentary, not part of the paper. Write a cell as the monomial $x_i y_j$. A rectangle must be avoided because its four squares are only two:
$x_i^2 y_j^2 + x_i^2 y_q^2 + x_p^2 y_j^2 + x_p^2 y_q^2 = (x_i y_j + x_p y_q)^2 + (x_i y_q - x_p y_j)^2$. Through the chain $\mathrm{BSR}(m, n) \geq z_A(m, n) \geq z_L(m, n)$ of Section 2, Theorem A gives a lower bound $\lfloor 3m/2 \rfloor + 2$ for the largest number of squares a sum-of-squares biquadratic form in $x \in \mathbb{R}^m$, $y \in \mathbb{R}^3$ can need.

Theorem B. p. 4

For every $n \geq 3$, $\liminf_{m\to\infty} F_n(m)/m \geq \lambda_n$, where

$$\lambda_n = \begin{cases} \dfrac{2n-3}{6}, & n \equiv 0 \pmod 3, \\ \dfrac{(n-1)(2n-5)}{6(n-2)}, & n \equiv 1 \pmod 3, \\ \dfrac{(n-2)(2n-1)}{6(n-1)}, & n \equiv 2 \pmod 3, \end{cases}$$

and $\lim_{m\to\infty} F_n(m)/m = \lambda_n$ for $3 \leq n \leq 6$: the slopes are $1/2$, $3/4$, $9/8$ and $3/2$.

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library.

Classical and new: lower bound: a fixed-width formulation of Lebedev's construction [Leb, Thms. 4, 13]; the exact slopes 1/2, 3/4, 9/8, 3/2 for n ≤ 6 are the paper's.

↳ Used in: Theorem C (p. 20)

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

With n columns fixed, the number of pairs grows linearly in the number of rows. The theorem gives a slope λ_n, about $n/3 - 1/2$, that is always reachable, and says that for $n = 3, 4, 5, 6$ nothing does better: the slopes are exactly $1/2$, $3/4$, $9/8$ and $3/2$.

Why: the lower bound comes from *modules*, blocks of rows built around a hollow triangle of three column blocks (Lemma 4.1), repeated as often as needed. The upper bound comes from a linear program attached to each "shadow" (the triples of columns that occur in many rows); a rational dual solution of each program is a certificate, and for $n \leq 6$ every certificate has value at most λ_n (Proposition 6.1).

The constructions follow Lebedev's three-block construction [26 , Thms. 4, 13] , adapted to column blocks of any sizes in fixed width; the exact slopes for $n \leq 6$ are the paper's. For $n = 7$ a mixed integer program returns the numerical value 1.8, compatible with $9/5 = \lambda_7$ but not certified.

Evidence & scope

- **Program:** every module rebuilt for n = 3,…,12 and checked by two independent admissibility checkers. `supplement: lambda_modules.json, modulos_lambda.py`
- **Program:** n = 7: a mixed integer program over the shadow returns 1.8 (floating point), compatible with 9/5 = λ_7; not a certificate. `supplement: zl_sombras_mip.py`

In Section 6 we also show that the slope exists for every n: it is a rational number $\rho_n \geq \lambda_n$, the optimum of a finite family of linear programs, and $F_n(m) = \rho_n m + O_n(1)$ (Theorem C).

4 Why three blocks?

What is the cheapest way to pay for pairs with holes? Project the columns onto a few blocks and let every row occupy the union of two blocks. A pair joins an endpoint of a row of type $\{A, C\}$ lying in block A with an endpoint of a row of type $\{B, C\}$ lying in block B. Both opposite cells fall in blocks these rows avoid, and a witness for (C3) would have to meet all three blocks, which no row does (Figure 3). Geometrically the rows are the sides of a triangle whose face is never filled, and each pair is an angle around that hollow face.

The number three has a long history here. In 1941 Turán determined the largest graph on N vertices with no complete subgraph K_{r+1}: the complete r-partite graph with equal parts [23] , the result from which extremal graph theory grew [9] . It reappears as follows. Take a square matrix with empty base and empty diagonal, let the occupied set be the adjacency matrix of a graph G, and pair each edge with its transpose, $\{(i, j), (j, i)\}$. Then (C2) holds automatically, since the opposite cells lie on the diagonal, and a witness for (C3) is exactly a K_4 through the edge ij. So the best such configuration is Turán's complete tripartite graph, and the three blocks are Turán's three parts. (Symmetry alone is not enough: the four-cycle is symmetric and K_4-free, but the pair $\{(1, 2), (3, 4)\}$ breaks (C2).) The lemma below is a

formulation of Lebedev's construction [26 , Thms. 4, 13] in the form that fixed width requires: column blocks only, of any sizes, and rows of mixed types.

Lemma 4.1. Isabelle **(blocks)** p. 5

> *Partition the columns into blocks, give every row of the grid a type T of at most two blocks with its occupied support inside $\bigcup_{x\in T} B_x$, and let every pair join an endpoint of a row of type T in a block $x \in T \setminus T'$ to an endpoint of a row of type T' in a block $y \in T' \setminus T$. Then the pairs are admissible.*

Isabelle · ours **show**

`blocks_admissible` · `Isabelle/AFP/Limited_Augmented_Zarankiewicz/ZL_Labels.thy`, line 89 — the lemma

```
lemma blocks_admissible:
  assumes typ2: "\<And>r. card (typ r) \<le> 2" and fin: "\<And>r. finite
(typ r)"
    and inside: "\<And>r j. j \<in> supp Oc r \<Longrightarrow> blk j \<in>
typ r"
    and rows: "fst c \<noteq> fst d"
    and x: "blk (snd c) \<in> typ (fst c) - typ (fst d)"
    and y: "blk (snd d) \<in> typ (fst d) - typ (fst c)"
  shows "C2 Oc c d \<and> C3 Oc c d"
```

`label_admissible` · `Isabelle/AFP/Limited_Augmented_Zarankiewicz/ZL_Labels.thy`, line 35 — the general label lemma of §6.3, of which it is the special case of labels made of two blocks

```
lemma label_admissible:
  assumes lab: "\<And>r. supp Oc r \<subseteq> lab r"
    and tri: "\<And>k a b e. a \<in> supp Oc k \<Longrightarrow> b \<in>
supp Oc k \<Longrightarrow> e \<in> supp Oc k \<Longrightarrow>
                a \<noteq> b \<Longrightarrow> a \<noteq> e \
<Longrightarrow> b \<noteq> e \<Longrightarrow> {a, b, e} \<in> H"
    and rows: "fst c \<noteq> fst d"
    and vert: "snd c = snd d \<Longrightarrow> lab (fst c) \<inter> lab
(fst d) = {snd c}"
    and diag2: "snd c \<noteq> snd d \<Longrightarrow> snd d \<notin> lab
(fst c) \<or> snd c \<notin> lab (fst d)"
    and diag3: "snd c \<noteq> snd d \<Longrightarrow> \<forall>l \<in> lab
(fst c) \<inter> lab (fst d) - {snd c, snd d}. {snd c, snd d, l} \<notin>
H"
  shows "C2 Oc c d \<and> C3 Oc c d"
```

Trust base: 🔒 kernel only. Isabelle: 2 results, no oracles

Proof: in this note · **Covers:** not formalized · **Cite:** p. 4, doi:10.5281/zenodo.23778502

NEW **Classical and new:** the paper: «a formulation of Lebedev's construction [Leb, Thms. 4, 13] in the form that fixed width requires».

↳ Used in: Theorem B (p. 10), Theorem C (p. 20), Proposition 6.3 (p. 23), 1 (p. 37), 4 (p. 39)

Proof

> The opposite cells lie in block y of the first row and block x of the second, which these rows avoid: (C2). A column common to both rows lies in a block $z \in T \cap T'$ distinct from x, y, so a (C3) witness row would meet three blocks.

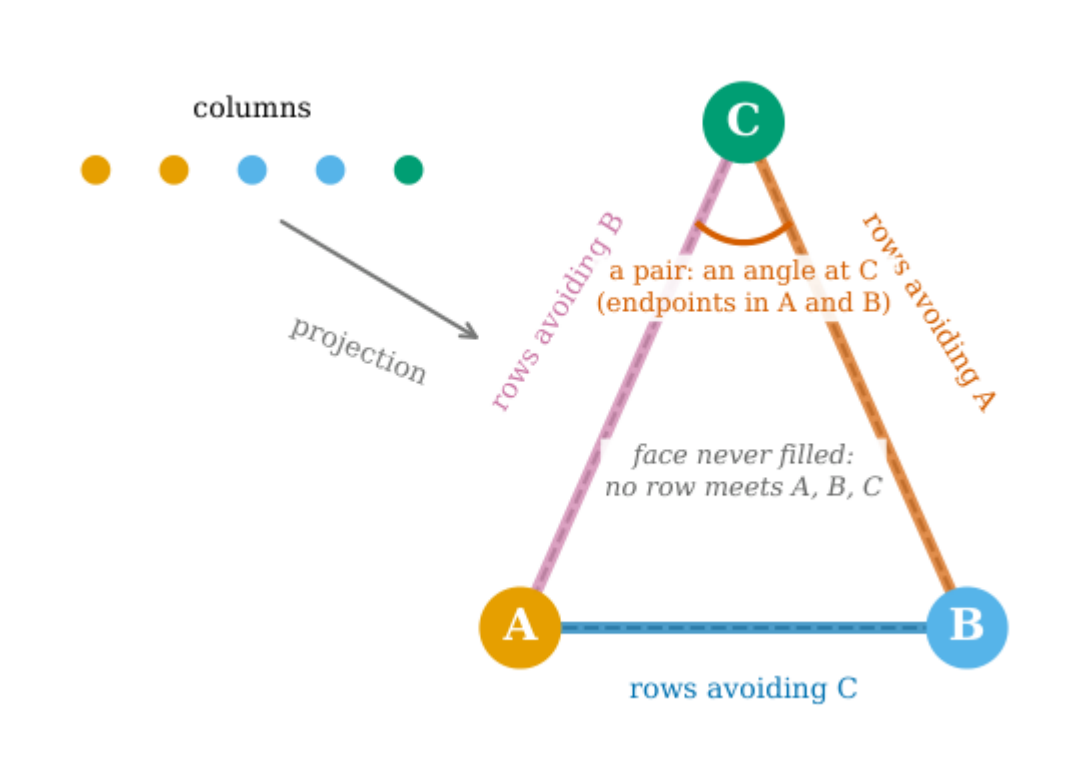
💡 **In plain words (commentary, not part of the paper)**

Split the columns into blocks and let every row live inside one or two blocks. Pair a cell of one row with a cell of another row only if each cell lies in a block the *other* row does not use. Then every such pair is admissible.

Why: the two opposite cells of the pair fall in blocks that their rows avoid, so they are empty, which is (C2); a column shared by the two rows lies in a third block, so a witness row for (C3) would have to touch three blocks, which no row does. With three blocks, the rows are the sides of a triangle and each pair is an angle around its empty face; the paper compares this with Turán's complete tripartite graph.

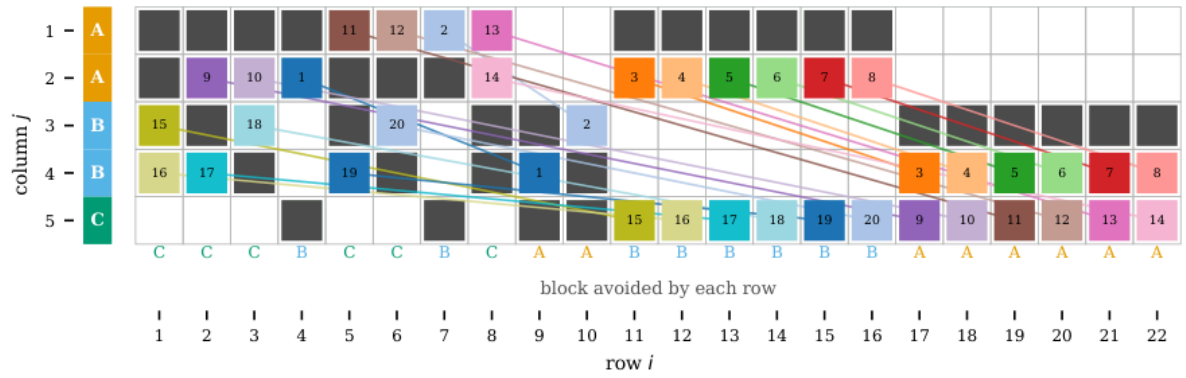
The paper presents the lemma as a formulation of Lebedev's construction [26 , Thms. 4, 13] in the form fixed width requires. It is formalized in Isabelle/HOL (`blocks_admissible`), as a special case of the general label lemma (`label_admissible`).

🖼️ **Figure 3**



The mechanism behind the three-block constructions. Left: the columns are projected onto three blocks A, B, C; a row occupying two blocks is a side of the triangle, a pair is an angle (an endpoint in A on a row of type AC and one in B on a row of type BC, meeting at C), and no row fills the face, which is why (C3) never fires. Right: the same template in fixed width, rows grouped by type and columns by block; each type leaves one block empty, and a pair runs along a channel between two types through their common block (Lemma 4.1).

Figure 4



A three-block configuration on 22×5, drawn transposed (the row index i on the horizontal axis, the column index j on the vertical axis). Blocks $A = \{1, 2\}$, $B = \{3, 4\}$, $C = \{5\}$, the balanced sizes for $n = 5$. Rows 1–10 are the special rows (two grey cells each) and rows 11–22 singleton rows; the letter under each row is the block it avoids. The 32 cells of E_1 are grey and the 20 pairs are numbered in their two cells (16 endpoints lie in special rows, 24 in singleton rows). There are $110 - 32 = 78$ free cells and $20/78 > 1/4$ pairs per free cell; the 40 endpoints occupy more than half of the free cells. Data in `witness_22x5_three_blocks.json`, checked by both admissibility checkers.

Why not four blocks, or two? With two blocks a row occupying both has no hole left; with k blocks of equal size a row occupies only $2/k$ of the columns, and three equal blocks balance occupation against holes. No configuration whatsoever does better than one third asymptotically: Lebedev's density bound [26 , Thm. 3] gives $|E_2| \le mn/3 + O(m + n)$. For fixed n the error term is of the same order as the main term in m, so this bound does not give the slope; for $n \le 6$ the exact slope is λ_n, about one half below $n/3$, and finding it is the subject of Section 6.

How many free cells does a pair use? Each singleton row leaves $n - 1$ free cells. For $n = 3, 4$ the slope is exactly $(n - 1)/4$, as if every pair used four free cells, its two endpoints and two holes. From $n = 5$ on the blocks do better: Figure 4 has 20 pairs on $110 - 32 = 78$ free cells, more than one pair per four free cells, and in the modules of Section 6 the number of pairs per free cell, $\lambda_n/(n - 1)$, tends to one third from below. At most one pair per three free cells is what holds in every case we computed (Question 3).

How to read Figure 4. Every pair is an angle of the triangle, and the three corners are used. Pairs 1–8 join an endpoint in block B of a row avoiding A with an endpoint in block A of a row avoiding B: eight angles at C. Pairs 9–14 join block C of a row avoiding A with block A of a row avoiding C: six angles at B. Pairs 15–20 join block C of a row avoiding B with block B of a row avoiding C: six angles at A. In each pair the two opposite cells fall in blocks that their rows avoid, so (C2) holds, and no row meets all three blocks, so no witness for (C3) exists. The special rows take part too: each avoids the block it does not touch.

5 Three columns

Where do the holes come from when there are only three columns? For $m \geq 3$ a maximum base consists of the rows $\{1,2\},\{1,3\},\{2,3\}$ and $m-3$ rows with a single one. For a row r let S_r be its occupied support; the cells of S_r outside E_1 are its *endpoints*, and $2|E_2|$ is the number of endpoints. Two observations do all the work.

Lemma 5.1. ✅ Isabelle **(vertical pairs, any n)** 📕 p. 6

> *If $m \geq \binom{n}{2}$ and $\{(i,j),(p,j)\} \in E_2$, then $S_i \cap S_p = \{j\}$.*

Isabelle · ours **show**

`vertical_pair` · `Isabelle/AFP/Limited_Augmented_Zarankiewicz/ZL_Labels.thy`, line 189 — the lemma, in a locale whose hypotheses are the properties of a maximum base used in the text

```
lemma vertical_pair:
  assumes "((i, j), (p, j)) \<in> P"
  shows "supp Occ i \<inter> supp Occ p \<inter> {..<n} = {j}"
```

Trust base: 🔒 kernel only. Isabelle: 1 result, no oracles

Proof: in this note · **Formalized:** Isabelle/HOL, by the author (public formalizations repository) · **Covers:** the whole statement · **Cite:** Lemma 5.1, p. 6, doi:10.5281/zenodo.23778502

↳ Used in: Theorem A (p. 8), Lemma 5.2 (p. 15), Theorem D (p. 25)

Proof

> If $\ell \neq j$ lies in both supports, the row k of E_1 through the pair of columns $\{j,\ell\}$ is distinct from i,p (its cell in column j is in E_1, theirs are not) and gives the forbidden cells $(k,\ell),(k,j),(i,\ell),(p,\ell)$ of (C3).

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

A *vertical* pair joins two cells of the same column. When the base is maximum, the two rows of such a pair can share no other column.

Why: if they shared a second column ℓ, the row of the base that covers the pair of columns $\{j,\ell\}$ (a maximum base has one for every pair of columns) would complete the five cells forbidden by (C3).

This is the first of the two observations behind Theorem A; it holds for every n. It is formalized in Isabelle/HOL (`vertical_pair`), inside a locale whose hypotheses are the properties of a maximum base used in the text.

Lemma 5.2. ✅ Isabelle **(one full row, any n)** 📕 p. 7

> *If $n \geq 3$ and $m \geq \binom{n}{2}$, at most one row is completely occupied.*

Isabelle · ours **show**

`at_most_one_full_row` · `Isabelle/AFP/Limited_Augmented_Zarankiewicz/ZL_Labels.thy`, line 328 — the lemma

```
lemma at_most_one_full_row:
  assumes "full F" and "full G"
  shows "F = G"
```

`endpoint_bound` · `Isabelle/AFP/Limited_Augmented_Zarankiewicz/ZL_Counts.thy`, line 64 — its consequence $2F_n(m) \le (n-2)m - C(n,2) + 1$

```
theorem endpoint_bound:
  "card (Occ - E1) + n * (n - 1) div 2 \<le> (n - 2) * m + 1"
```

Trust base: 🔒 kernel only. Isabelle: 2 results, no oracles

Proof: in this note · **Formalized:** Isabelle/HOL, by the author (public formalizations repository) · **Covers:** the whole statement · **Cite:** Lemma 5.2, p. 7, doi:10.5281/zenodo.23778502

↳ Used in: Theorem A (p. 8), Corollary 6.2 (p. 22), Theorem E (p. 30), 3 (p. 38)

Proof

> Let $F \neq G$ be full. No row of the base has more than two cells, so F has an endpoint (F, j). A partner (F, q) in the same row is excluded by G and any third column via (C3). A partner (P, j) in the same column contradicts Lemma 5.1, since S_P contains the E_1-cell of P, outside column j. A partner (P, q) with $P \neq F$, $q \neq j$ forces $(P, j) \notin O$ by (C2), as $(F, q) \in O$; so P is not full, $P \neq G$, an E_1-cell of P lies in a column $\ell \notin \{j, q\}$, and $(G, \ell), (G, j), (G, q), (F, \ell), (P, \ell) \in O$ violates (C3).

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

In the regime $m \geq \binom{n}{2}$, at most one row can be completely occupied.

Why: if two rows F and G were full, F has an endpoint, and its partner can be in the same row, in the same column, or elsewhere. In the same row, G and a third column give a (C3) witness; in the same column, Lemma 5.1 is violated; and elsewhere, (C2) forces an empty cell in the partner's row and then G completes a (C3) witness.

Consequently every other row keeps a hole, which gives the bound $2F_n(m) \leq (n-2)m - \binom{n}{2} + 1$; for $n = 3$ it is the exact bound of Theorem A. Lemma and bound are formalized in Isabelle/HOL (`at_most_one_full_row`, `endpoint_bound`).

Every other row keeps an empty cell, so a special row carries at most $n - 3$ endpoints, a singleton row at most $n - 2$, and the full row one more. Hence, for every $n \geq 3$ and $m \geq \binom{n}{2}$,

$$2F_n(m) \;\leq\; (n-2)\,m - \binom{n}{2} + 1.$$

For $n = 3$ this is the exact bound of Theorem A; for larger n the slope $(n-2)/2$ is weaker than $\lambda_n \approx n/3$, and the real work is in Section 6. Lemmas 5.1 and 5.2 and this bound are formalized in Isabelle/HOL (see Formalizations at the end). The definitions of Section 2 and this proof are formalized in Lean 4 with Mathlib, with no axioms beyond the standard ones; the module is

`Formalizations.Combinatorics.Zarankiewicz.LimitedAugmentedThree`

in the public repository https://github.com/karlesmarin/formalizations (commit `389a576`, Lean `v4.35.0`, Mathlib `f0b4c81`; main theorem `ZL.zL_three`, checked with `lake build`). It proves the upper bound for every maximum base, not only for one, and builds the families above.

Theorem A also settles the second half of the open problem of [30] , which asks whether the small gap they observed between the weak number $z_{wL}(m,3)$ and $z_L(m,3)$ persists. They prove $z_{wL}(m,3) = m + 3 + \lfloor (2m-4)/3 \rfloor$ for $m \geq 16$ [30 , Thm. 26] , so

$$z_{wL}(m,3) - z_L(m,3) = \left\lfloor \frac{2m-4}{3} \right\rfloor - \left\lfloor \frac{m-2}{2} \right\rfloor = \frac{m}{6} + O(1) \qquad (m \geq 16) :$$

the gap grows linearly. The weak conditions allow slope $2/3$ per row, the original ones only $1/2$.

With this, five of the numbers studied in the series are known exactly in three columns: the original z_L, the weak z_{wL} of [29 , 30] , and the recursive-line, signed and second-order numbers z_{RL}, z_{SL}, z_2 of Löfberg and Qi [20] (relaxed versions of the irreducibility test behind z_L). They prove $z_2(m,3) = z_{RL}(m,3) = 2m$ for $m \geq 3$ [20 , Prop. 4.13] , and their chain $z_2 \geq z_{SL} \geq z_{RL} \geq z_{wL} \geq z$ [20 , Prop. 7.4] then gives $z_{SL}(m,3) = 2m$ as well; it sits above $z_L \leq z_{wL}$ [29] . (z_A and BSR in three columns are not determined here.) Measured from $z(m,3) = m + 3$, each new row buys

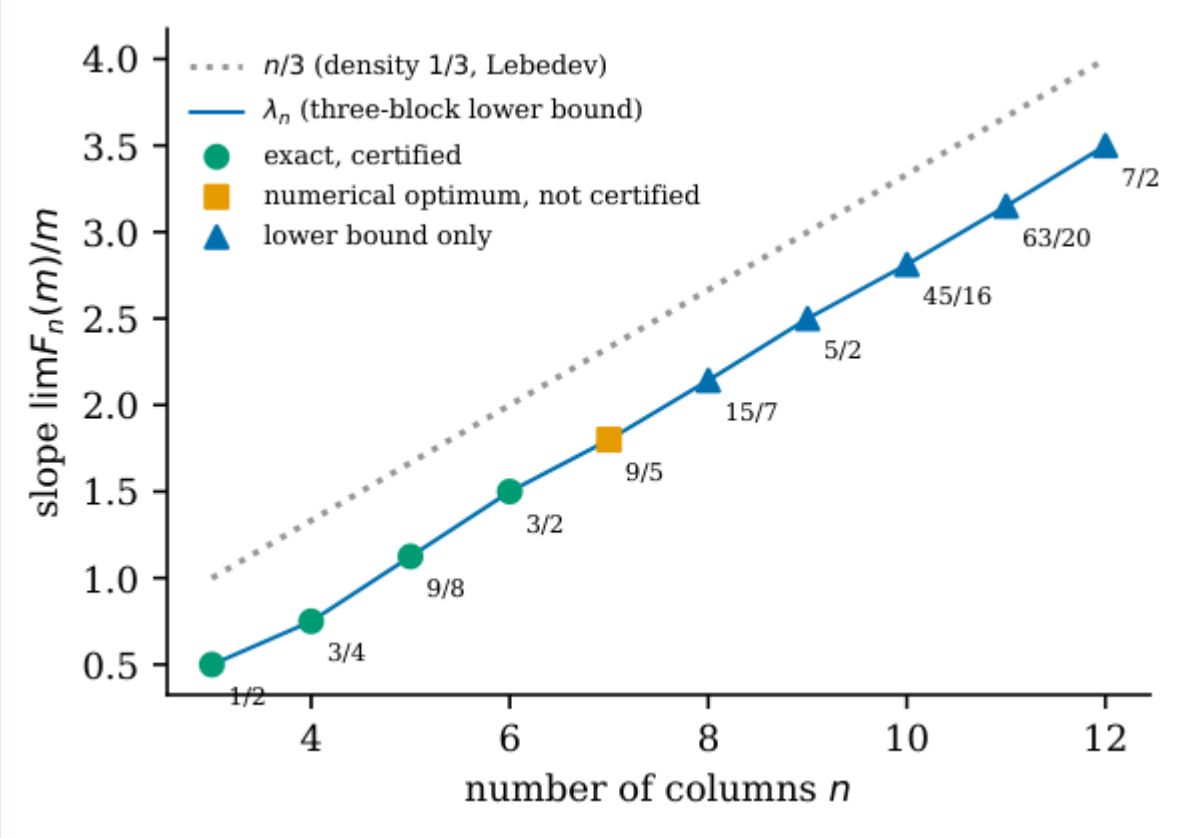
$\frac{1}{2}$ pair for z_L, $\frac{2}{3}$ for z_{wL} $(m \geq 16)$, 1 for $z_{RL} = z_{SL} = z_2$, and the three formulas $\lfloor 3m/2 \rfloor + 2$, $m + 3 + \lfloor (2m-4)/3 \rfloor$ and $2m$ are consistent with the published gap $z_2(m,3) - z_{wL}(m,3) = \lfloor (m-3)/3 \rfloor$ for $m \geq 16$ [20 , Cor. 4.14] , which we rechecked.

Card A — three columns. $z_L(m,3) = m + 3 + \lfloor (m-2)/2 \rfloor$ for all $m \geq 3$; previously known for $m \leq 13$.

6 Slopes in fixed width

How many pairs does each new row buy when n is fixed? Lemma 5.1 already shows the price of a pair. The answer comes in two halves.

Figure 5



The slope $\rho_n = \lim F_n(m)/m$ for $n = 3, \dots, 12$. Green: $\rho_n = \lambda_n$ certified by rational duals ($1/2, 3/4, 9/8, 3/2$ for $n \le 6$); orange: the numerical candidate 1.8 for ρ_7 from a mixed integer program, compatible with $9/5$ but not certified; blue: the lower bound λ_n of the three-block modules (Theorem B), with ρ_n unknown ($n = 8$ was inconclusive). The solid line is λ_n and the dotted line the density $n/3$ of Lebedev's regime where both dimensions grow.

6.1 Lower bound: modules around a hollow triangle

Number the blocks $1, 2, 3$ and let a row of type $\{a, b\}$ occupy blocks a and b completely, apart from its single cell of E_1. For block sizes (B, c, c) a module consists of $2c - 1$ rows of type $\{1, 2\}$ and $2c - 1$ rows of type $\{1, 3\}$ with their cell of E_1 in block 1, $B - 1$ rows of type $\{2, 3\}$ with it in block 2 and $B - 1$ with it in block 3. For sizes (B, B, c) it consists of c rows of type $\{1, 2\}$ with the cell in block 1, c with it in block 2, $2B - 1$ rows of type $\{1, 3\}$ with it in block 1 and $2B - 1$ of type $\{2, 3\}$ with it in block 2. Matching the endpoints along the three channels of Lemma 4.1 (any bijection within a channel) gives

$$(B, c, c) : \ (2c - 1)(2B + c - 2) \text{ pairs on } 2(2c + B - 2) \text{ rows},$$

$$(B, B, c) : \ (2B - 1)(B + 2c - 1) \text{ pairs on } 2(2B + c - 1) \text{ rows}$$

The special rows and any further singleton rows carry no endpoints and meet at most two blocks. With the module (B, c, c) of sizes (b, b, b) for $n = 3b$ and $(b + 1, b, b)$ for $n = 3b + 1$, and the module (B, B, c) of sizes $(b + 1, b + 1, b)$ for $n = 3b + 2$, the ratio of pairs to rows is exactly λ_n; repeating the module gives $F_n(m) \ge \lambda_n m - O(1)$. Every module was rebuilt for $n = 3, \dots, 12$ and checked by two independent admissibility checkers (`lambda_modules.json`).

6.2 Upper bound: shadows of triples

Discard every pair meeting a special row or a row whose occupied support occurs fewer than three times; there are at most $(n-2)\binom{n}{2} + 2(n-1)(2^n-1)$ of them, a constant. Each row keeps its original support as a *label*, even if discarding removes one of its endpoints; the inequalities below are capacities, so unused cells are allowed. The remaining pairs join singleton rows with *frequent* labels, and they obey:

- no pair lies inside one row (a second row with the same support gives a (C3) witness);
- a vertical pair in column j joins supports S, T with $S \cap T = \{j\}$ (Lemma 5.1);
- rows with the same support are never paired (by (C2), or by the previous rule for a vertical pair);
- a pair from (S, j) to (T, q), $j \neq q$, needs $q \notin S$ or $j \notin T$ (C2), and no $\ell \in (S \cap T) \setminus \{j, q\}$ with $\{j, q, \ell\}$ in the *shadow* $\mathcal{H}$, the set of triples of columns contained in some frequent support: such a triple lies in at least three rows, one of them external to the pair (C3).

In particular a full row carries no remaining pair: partners in its row or column are excluded by the first two rules, any other partner (T, q) would need $j \notin T$, and then the E_1-cell of T gives a forbidden ℓ. For each shadow this defines a linear program: variables $x_{S,a} \geq 0$, the proportion of rows with label S and cell of E_1 in column a (over supports all of whose triples lie in $\mathcal{H}$), with $\sum x_{S,a} \leq 1$, capacities $\sum_{a \neq j} x_{S,a}$ at the port (S, j), and flows on the permitted pairs of ports. Every actual configuration gives a feasible point, so the value of the program bounds $\limsup F_n(m)/m$. A dual solution, weights $y_{S,j} \geq 0$ with $y_{S,j} + y_{T,q} \geq 1$ on permitted pairs and a budget $\Lambda \geq \sum_{j \in S \setminus \{a\}} y_{S,j}$ for all (S, a), certifies the value $\leq \Lambda$: the flow is at most $\sum_{S,a} x_{S,a} \sum_{j \in S \setminus \{a\}} y_{S,j} \leq \Lambda$. The maximum over all shadows bounds the slope.

What does a certificate look like? For $n = 4$ there are five shadows up to symmetry, and only one, $\{123, 124\}$, reaches $3/4$: it is the three-block shadow with blocks $\{1, 2\}, \{3\}, \{4\}$. Its dual puts weight $1/4$ on the ports of the support $\{1, 2\}$, weight $3/4$ on the ports of the other two-column supports, and on the supports 123 and 124 weight $1/4$ on columns $1, 2$ and $1/2$ on the third column. Every permitted pair of ports carries total weight at least 1, and every row, whatever column holds its one, carries at most $3/4$; summing over rows gives at most $3m/4 + O(1)$ pairs. The other four shadows have certificates of value $1/2$ or $2/3$.

Proposition 6.1. p. 9

> *For $n = 3, 4, 5, 6$ the shadows form $2,\ 5,\ 34$ and 2136 orbits under column permutations, and each orbit has a rational dual certificate of value at most λ_n; an independent checker in exact arithmetic verifies $33,\ 501,\ 14\,316$ and $2\,892\,596$ inequalities.*

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL

Light/Metamath): not found in any formal library; certified by rational duals checked in exact arithmetic, not formalized.

↳ Used in: Theorem B (p. 10), Theorem C (p. 20)

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

For $n = 3, 4, 5, 6$ the shadows fall into 2, 5, 34 and 2136 classes under permuting the columns. For each class there is a dual solution with rational weights whose value is at most λ_n, and an independent program in exact arithmetic checks all of them: 33, 501, 14 316 and 2 892 596 inequalities.

What a certificate is: weights on the ports (a support and one of its columns) such that every permitted pair of ports carries weight at least 1 and every row carries at most Λ; summing over rows bounds the number of pairs by $\Lambda m + O(1)$. For $n = 4$ the paper writes the one certificate that reaches $3/4$ explicitly.

These certificates are the upper half of Theorem B. All numbers are the paper's; the certificates and the checker are the files `shadow_certificate_n3..6.json` and `cert_sombras_check.py`.

🛡️ **Evidence & scope**

- **Program:** 2, 5, 34 and 2136 orbits; the independent checker verifies 33, 501, 14 316 and 2 892 596 inequalities in exact arithmetic. `supplement: shadow_certificate_n3..6.json, zl_cert_sombras.py, cert_sombras_check.py`

Together the two halves prove Theorem B. For $n = 7$ the 2^{35} shadows are too many to list; solving the shadow as $0/1$ variables in one mixed integer program returns the numerical optimum 1.8, compatible with $9/5 = \lambda_7$; a decimal output does not certify the rational value, and we report it as evidence only.

6.3 The slope always exists

The linear programs do more than bound the slope: their optimum is attained.

Theorem C. p. 9

> *For every $n \geq 3$ let ρ_n be the largest value of the shadow programs over all shadows $\mathcal{H}$. Then ρ_n is rational and*
>
> $$F_n(m) = \rho_n\, m + O_n(1) \qquad (m \to \infty).$$
>
> *Moreover $\rho_n \geq \lambda_n$, with equality for $n \leq 6$.*

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library.

↳ Used in: Corollary 6.2 (p. 22), Theorem D (p. 25), Theorem E (p. 30), Theorem F (p. 31), 1 (p. 37), 3 (p. 38), 6 (p. 39)

Proof

> The upper bound $F_n(m) \le \rho_n m + O_n(1)$ is the reduction above. For the lower bound take an optimal shadow and a rational optimal solution of its program, with $\sum x_{S,a} = 1$, and multiply by a common denominator D: this gives integers $X_{S,a}$ (rows with label S and E_1-cell in column a) and Y_e (pairs on each permitted pair e of ports). Build these rows as singleton rows; at the port (S, j) choose as many rows of label S with $a \neq j$ as the port needs, which the capacities allow, put an endpoint in column j of each, and join the endpoints along the pairs e. A row's real support is contained in its label. Two ports of the same label are never permitted, so the two cells of a pair lie in different rows. (C2) holds because supports only shrink. A vertical pair has $S \cap T = \{j\}$, so no column for a (C3) witness. For a diagonal pair, a witness row would contain three columns j, q, ℓ with $\ell \in (S \cap T) \setminus \{j, q\}$; the special rows are built with their two base cells only, so the witness is a labelled row, whose label puts $\{j, q, \ell\}$ in $\mathcal{H}$, which the permitted pairs exclude. The same argument applies to any number of copies of this module together with the $\binom{n}{2}$ special rows and filler singleton rows, so $F_n(m) \ge \rho_n D \lfloor (m - \binom{n}{2})/D \rfloor$. There are finitely many programs, all with rational data, so ρ_n is rational. The modules of Section 6.1 give $\rho_n \ge \lambda_n$, and Proposition 6.1 gives equality for $n \le 6$.

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

For every number of columns the slope exists: $F_n(m) = \rho_n m + O_n(1)$, where ρ_n is the best value of finitely many linear programs with rational data, so a rational number, and at least λ_n .

Why: the upper bound is the reduction to the shadow programs (Theorem B's proof). For the lower bound, an optimal rational solution is scaled to integers and built row by row: each row gets the support its label allows and endpoints at its ports, and the pairs follow the flows. The general label lemma shows that every pair built this way is admissible. Repeating this module, plus the special rows, gives $\rho_n m - O(1)$ pairs.

The paper checks the realization by machine for $n = 3, 4, 5, 6$ (configurations 7×3, 14×4, 26×5, 31×6 with exactly ρ_n pairs per labelled row). So whether $\rho_n = \lambda_n$ for every n (Question 1) is a question about one explicit finite family of linear programs.

🛡️ **Evidence & scope**

- **Program:** the realization of the optimal shadow program for n = 3,4,5,6: configurations 7×3, 14×4, 26×5, 31×6 with exactly ρ_n pairs per labelled row, admissible under both checkers. `supplement: zl_rho_realiza.py`

The admissibility step of this proof is a general *label lemma*: if every support lies in its label, any three occupied columns of a row form a triple of $\mathcal{H}$, and the pair is compatible with the labels, then (C2) and (C3) hold. It is formalized in Isabelle/HOL together with Lemma 4.1, which is the special case of labels made of two blocks (theory `ZL_Labels`, no `sorry`). We also checked the realization step by machine: for $n = 3, 4, 5, 6$ the program of the optimal

shadow, built as above with two copies, gives configurations of sizes 7×3, 14×4, 26×5 and 31×6 with exactly ρ_n pairs per labelled row $(1/2,\ 3/4,\ 9/8,\ 3/2)$, admissible under both of our checkers (`zl_rho_realiza.py`). So Question 1 is a question about one explicit finite family of linear programs.

How far is the original criterion from the relaxed ones? Lebedev has shown that the recursive-line, signed and second-order numbers all satisfy $z_X(m,n) = \frac{n+1}{2}m + O_n(1)$ for fixed n [15 , Thm. 1.2] . Theorem C and the bound $\rho_n \le (n-2)/2$ that follows from Lemma 5.2 give at once:

Corollary 6.2. p. 10

> *For every fixed $n \ge 3$ and $X \in \{RL, SL, 2\}$,*
>
> $$z_X(m,n) - z_L(m,n) = \Big(\frac{n-1}{2} - \rho_n\Big)m + O_n(1), \qquad \frac{n-1}{2} - \rho_n \ge \frac{1}{2}.$$

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library.

NEW **Classical and new:** combines Lebedev's asymptotics of the relaxed numbers [LebII, Thm. 1.2] with Theorem C.

↳ Used in: 1 (p. 37)

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

Qi and coauthors also studied relaxed versions of the irreducibility test behind z_L: the recursive-line, signed and second-order numbers. Lebedev proved that all three grow like $\frac{n+1}{2}m$ in fixed width [15 , Thm. 1.2] . Together with Theorem C this says that the original number stays linearly below them in every width: the gap has slope $\frac{n-1}{2} - \rho_n$, at least $\frac{1}{2}$.

Why at least $\frac{1}{2}$: Lemma 5.2 gives $\rho_n \le (n-2)/2$. For $n = 3, 4, 5, 6$ the gap slopes are $1/2$, $3/4$, $7/8$ and 1, from the certified ρ_n.

The relaxed numbers' asymptotics are Lebedev's; the corollary is the paper's combination of it with Theorem C.

So the separation is linear in every fixed width; for $n = 3, 4, 5, 6$ the gap slope is $1/2$, $3/4$, $7/8$ and 1. (In three and four columns the relaxed gain per row, $(n-1)/2$, happens to be exactly $2\rho_n$; from $n = 5$ on it is not.)

What is optimal inside the three-block family? Partition the columns into blocks A, B, C of sizes $x \ge y \ge z \ge 1$, $x + y + z = n$. An *ordinary module* is a finite set of R rows, each with exactly one base cell (its anchor), and P disjoint pairs of further cells, obeying the rule of Lemma 4.1: each row has a type of one or two blocks containing its anchor and its occupied support, with partial supports allowed. Let $M(x,y,z)$ be the supremum of P/R over ordinary modules. The maximum-base setting is recovered by adding the $\binom{n}{2}$ special rows bare, which

costs a bounded number of rows; so M is a slope, not a ratio $F_n(m)/m$ (for instance $M(1,1,1) = 1/2$, while $F_3(m)/m < 1/2$ for every m).

Proposition 6.3. p. 10

If $z(x-1) \geq y(y-1)$, then $M = \frac{(s-1)((x-1)s+yz)}{s(n-2)}$ with $s = y+z$; otherwise $M = \frac{(u+1)(zu+(x-1)(y-1))}{u(n-1)}$ with $u = x+y-2$. In the first case

$$\frac{n}{3} - \frac{1}{2} - M = \frac{\eta(2\eta-1)}{6(n-2)} + \frac{(s-1)(y-z)^2}{4s(n-2)}, \qquad \eta = x - \frac{y+z}{2},$$

and in the second case

$$\frac{n}{3} - \frac{1}{2} - M = \frac{\eta(2\eta-1)}{6(n-1)} + \frac{(u+1)(x-y)^2}{4u(n-1)}, \qquad \eta = \frac{x+y}{2} - z.$$

(On the boundary $z(x-1) = y(y-1)$ the two formulas agree wherever both are defined; for $x = y = z = 1$ only the first is defined, and it gives $M = 1/2$.) Consequently $\max_{x+y+z=n} M = \lambda_n$, attained only by the sizes (k,k,k), $(k+1,k,k)$ and $(k+1,k+1,k)$ according as $n = 3k$, $3k+1$, $3k+2$.

Isabelle · ours · part **show**

Only a part is in Isabelle: the two identities and the non-negativity of the deficit, after publication; that M is the value of the family (construction and weights) and the maximizers are on paper

`identity_case1` · `Isabelle/ZLThreeBlockFamily/ZL_Three_Block_Family.thy`, line 34 — the closed form of n/3 − 1/2 − M in the first case

```
theorem identity_case1:
  fixes x y z :: real
  assumes "x+y+z-2 \<noteq> 0" and "y+z \<noteq> 0"
  shows "(x+y+z)/3 - 1/2 - M1 x y z =
     eta1 x y z * (2 * eta1 x y z - 1) / (6 * (x+y+z-2))
     + (y+z-1) * (y-z)^2 / (4 * (y+z) * (x+y+z-2))"
```

`identity_case2` · `Isabelle/ZLThreeBlockFamily/ZL_Three_Block_Family.thy`, line 43 — the closed form in the second case

```
theorem identity_case2:
  fixes x y z :: real
  assumes "x+y+z-1 \<noteq> 0" and "x+y-2 \<noteq> 0"
  shows "(x+y+z)/3 - 1/2 - M2 x y z =
     eta2 x y z * (2 * eta2 x y z - 1) / (6 * (x+y+z-1))
     + (x+y-2+1) * (x-y)^2 / (4 * (x+y-2) * (x+y+z-1))"
```

`deficit_case1` · `Isabelle/ZLThreeBlockFamily/ZL_Three_Block_Family.thy`, line 89 — the deficit is non-negative, first case (x ≥ y ≥ z ≥ 1)

```
theorem deficit_case1:
  fixes x y z :: nat
  assumes "x \<ge> y" and "y \<ge> z" and "z \<ge> 1"
```

```
  shows "(real x + real y + real z)/3 - 1/2 - M1 (real x) (real y) (real z)
\<ge> 0"
```

`deficit_case2` · `Isabelle/ZLThreeBlockFamily/ZL_Three_Block_Family.thy`, line 112 — the deficit is non-negative, second case (also y ≥ 2)

```
theorem deficit_case2:
  fixes x y z :: nat
  assumes "x \<ge> y" and "y \<ge> z" and "z \<ge> 1" and "y \<ge> 2"
  shows "(real x + real y + real z)/3 - 1/2 - M2 (real x) (real y) (real z)
\<ge> 0"
```

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library; the closest is Lebedev's Lean companion of arXiv:2609.16555 (CleanHolePairCounting.lean, reciprocal_block_charge), a charging lemma for reciprocal pairs between three classes inside his density argument, not this optimization over the block sizes.

Trust base: 🔒 kernel only — **only part of the statement is formal** (see the box). Isabelle: 4 results, no oracles

Proof: in this note · **Covers:** not formalized · **Cite:** Proposition 6.1, p. 9, doi:10.5281/zenodo.23778502

↳ Used in: 1 (p. 37)

Proof

Lower bound: modules in which every row fills its two blocks and pairs are matched along the three channels of Lemma 4.1. In the first case take $z(s-1)$ rows of type AB and $y(s-1)$ of type AC anchored in A, and $(x-1)y$, $(x-1)z$ rows of type BC anchored in B, C; the channels then carry $yz(s-1)$, $(x-1)z(s-1)$ and $(x-1)y(s-1)$ pairs, so $R=s(n-2)$ and $P=(s-1)((x-1)s+yz)$. In the second case take $z(y-1)$ and $z(x-1)$ rows of type AB anchored in A and B, $(y-1)(u+1)$ of type AC anchored in A, and $(x-1)(u+1)$ of type BC anchored in B; the channels carry $z(x-1)(u+1)$, $z(y-1)(u+1)$ and $(x-1)(y-1)(u+1)$ pairs, so $R=u(n-1)$ and $P=(u+1)(zu+(x-1)(y-1))$. Upper bound: weights on endpoint cells by row type and block, $(AB,A)\mapsto p$, $(AB,B)\mapsto q$, $(AC,A)\mapsto r$, $(AC,C)\mapsto 1-q$, $(BC,B)\mapsto 1-r$, $(BC,C)\mapsto 1-p$, with $p=r=\frac{s(s-1)-yz}{s(n-2)}$, $q=\frac{z}{s}$ in the first case and $p=q=\frac{zu+(x-1)(y-1)}{u(n-1)}$, $r=\frac{y-1}{u}$ in the second, and on one-block rows $w_A=\max(p,r)$, $w_B=\max(q,1-r)$, $w_C=\max(1-p,1-q)$. Every pair of the block rule has weight at least 1; a row of any type and anchor carries at most M (the regime condition is exactly the inequality that makes the two non-active row costs at most M); and leaving cells unused only lowers a row's weight. Summing gives $P\le MR$. The identities are algebra, and η is a non-negative half-integer, so both terms are non-negative. For $n=3k$ they vanish only at $x=y=z$; for $n=3k+1$ the least deficit $1/(6(n-2))$ is attained only at $(k+1,k,k)$, in the first case, and for

> $n = 3k + 2$ the least deficit $1/(6(n-1))$ only at $(k+1, k+1, k)$, in the second; the other case is excluded there by its regime condition.

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

Inside the three-block family, how should the n columns be split? The proposition computes the best ratio of pairs to rows $M(x, y, z)$ for every split $x \geq y \geq z$ (over ordinary modules: rows with one base cell each, before the special rows are added) and shows it is largest exactly for the balanced splits (k, k, k), $(k+1, k, k)$, $(k+1, k+1, k)$, where it equals λ_n.

Why: explicit modules give the lower bound and explicit dual weights on the endpoint cells give the matching upper bound, in two regimes. Two identities write $\frac{n}{3} - \frac{1}{2} - M$ as a sum of two non-negative terms, one measuring how unequal the blocks are. This also explains the shape of λ_n: the $-\frac{1}{2}$ is the price of the base cell reserved in each row, and δ_n the price of blocks of unequal integer sizes.

The paper checks the identities symbolically and the weights and modules in exact arithmetic for every partition with $n \leq 60$. It does not settle Question 1: a general configuration need not follow the block rule.

Evidence & scope

- **Program:** the two deficit identities symbolically; weights, row budgets, modules and the maximum λ_n in exact arithmetic for every partition with n ≤ 60 (0 failures). `supplement: tres_bloques_identidades.py, tres_bloques_dual_check.py`

Theorem D. p. 11

> *For every $n \geq 3$ the largest number of pairs on the canonical base is $\kappa_n m + O_n(1)$, where*
>
> $$\kappa_n = \frac{\lfloor (n-1)^2/4 \rfloor}{n-1},$$
>
> *the number of edges of a balanced complete bipartite graph on the $n-1$ free columns divided by $n-1$.*

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library.

Proof: in this note · **Covers:** not formalized · **Cite:** p. 11, doi:10.5281/zenodo.23778502

NEW **Classical and new:** new: the reduction from the canonical base to the graph parameter γ, with a construction attaining the slope. Known: the sharp bound γ(G) ≤ ⌊k²/4⌋/k for graphs (F. Hof, Weight distribution in matching games, University of Twente, 2016, Thm. 7.2; recorded in Hof–Kern–Kurz–Paulusma, SAGT 2018, footnote 7), the graphical case of the Freixas–Kurz

conjecture; for odd n the upper bound also follows from the counting in the proof of Lebedev, Thm. 9. Version 1 on Zenodo did not say this; version 2 does (1-nov-2026).

↳ Used in: 6 (p. 39)

Proof

Let column 1 be the anchor, V the set of the $k = n - 1$ free columns, and, for an ordinary row r of an arbitrary admissible configuration on the canonical base, $L_r = S_r \setminus \{1\}$ its *label*. A label is *frequent* if at least three rows carry it; at most $2 \cdot 2^k$ rows carry a rare label. Discard the pairs meeting a special row or a row with a rare label, at most $(n-2)\binom{n}{2} + 2(n-1)2^k$ of them. Let $\{(r,j),(r',q)\}$ be a remaining pair; its cells avoid column 1, which holds the base cells. It is not horizontal: if $r = r'$, a third row r'' with label L_r gives the (C3) witness $(r'',1),(r'',j),(r'',q),(r,1)$. It is not vertical: both rows contain column 1, against Lemma 5.1. So $r \neq r'$, $j \neq q$, and no frequent label contains both j and q, since among its three rows one, $r'' \notin \{r, r'\}$, would give the witness $(r'',1),(r'',j),(r'',q),(r,1),(r',1)$. Let G be the graph on V whose edges are the pairs $\{j,q\}$ used by remaining pairs: every frequent label is an independent set of G. For any $p \in Q(G) = \{p \in \mathbb{R}^V_{\geq 0} : p_u + p_v \geq 1 \text{ for } uv \in E(G)\}$, charging each pair $p_j + p_q \geq 1$ and each row the weights of its endpoints, which lie in its label, gives

$$F \;\leq\; \sum_r p(L_r) + O_n(1) \;\leq\; \gamma(G)\, m + O_n(1), \qquad \gamma(G) = \min_{p \in Q(G)} \max_{I \text{ independen}}$$

the critical threshold of the graphical simple game of G [16] .

It remains to bound $\gamma(G)$. The sharp bound $\gamma(G) \leq \lfloor k^2/4 \rfloor / k$ for every graph on $k \geq 4$ vertices is known: it is [18 , Thm. 7.2] , and it is recorded for odd k in [17 , footnote 7] , whose Theorem 1 gives $k/4$. For completeness we include a short proof, following the minimal-norm argument of [16 , Thm. 3] ; it also covers $k = 2, 3$. Let p^* be the point of minimal Euclidean norm of the closed convex set $Q(G)$, so $\langle p^*, q - p^* \rangle \geq 0$ for every $q \in Q(G)$. For an independent set I the indicator of $V \setminus I$ lies in $Q(G)$, which gives

$$p^*(I) \;\leq\; \sum_v p^*_v - \|p^*\|^2 \;=\; \frac{k}{4} - \left\| p^* - \tfrac{1}{2}\mathbf{1} \right\|^2 .$$

For even k this is the claim. Let k be odd, and call an edge *tight* if $p^*_u + p^*_v = 1$. A vertex with $p^*_v > 0$ on no tight edge could be lowered, so the vertices isolated in the tight graph have $p^*_v = 0$ and contribute $\frac{1}{4}$ to $\|p^* - \frac{1}{2}\mathbf{1}\|^2$. On a component of the tight graph with at least two vertices p^* alternates between t and $1 - t$ along tight edges. If the component is not bipartite, $t = \frac{1}{2}$. If it is bipartite with parts of sizes a, b, moving t to $t + \varepsilon$ on one part and $1 - t - \varepsilon$ on the other keeps p^* in $Q(G)$ for small ε of the admissible sign, so minimality forces $0 < t < 1$ and $at = b(1-t)$, that is $t = b/(a+b)$; the component contributes

> $(a-b)^2/(4(a+b))$, which is at least $1/(4k)$ when $a \neq b$. Hence either $p^* = \frac{1}{2}\mathbf{1}$, and then $|I| \leq k/2$, so $|I| \leq \frac{k-1}{2}$ and $p^*(I) \leq \frac{k-1}{4}$; or $\|p^* - \frac{1}{2}\mathbf{1}\|^2 \geq \frac{1}{4k}$ and $p^*(I) \leq \frac{k}{4} - \frac{1}{4k}$. In both cases $\gamma(G) \leq \frac{k^2-1}{4k} = \lfloor k^2/4 \rfloor / k$.
>
> For the lower bound split $V = X \cup Y$ with $|X| = \lceil k/2 \rceil$, $|Y| = \lfloor k/2 \rfloor$. A module has $|Y|$ singleton rows occupying X and $|X|$ singleton rows occupying Y, all anchored in column 1, and joins the $|X||Y|$ endpoints of each kind by any bijection: $\lfloor k^2/4 \rfloor$ pairs on k rows. A pair joins (r, x) with (r', y), $x \in X$, $y \in Y$; the opposite cell (r, y) is empty, which gives (C2); the two rows share only column 1, and no row contains 1, x and y (special rows have two cells), which gives (C3). Repeating the module and adding the special rows and filler rows gives $\kappa_n m - O_n(1)$ pairs.

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

In Lebedev's *canonical* base all the single ones of the base sit in column 1. The theorem gives its exact slope for every n: $\kappa_n = \lfloor (n-1)^2/4 \rfloor / (n-1)$, the number of edges of a balanced complete bipartite graph on the $n-1$ free columns, per column. For $n = 3, 4, 5, 6$ this is $1/2$, $2/3$, 1, $6/5$, against $\rho_n = 1/2, 3/4, 9/8, 3/2$.

Why: since every ordinary row contains column 1, every surviving pair is diagonal and its two columns are never together in a frequent support. The columns then form a graph whose independent sets are the supports, and the slope is the critical threshold of the graphical simple game of that graph [16] . The minimal-norm argument of Hof et al. plus a parity step bounds it by $\lfloor k^2/4 \rfloor / k$; the balanced complete bipartite graph attains it.

The values for $n \leq 6$ are also certified by rational duals and realizations.

Attribution (corrected 29-oct-2026; the Zenodo text v1 does not say it). The sharp bound $\gamma(G) \leq \lfloor k^2/4 \rfloor / k$ for graphs is known: it is Theorem 7.2 of F. Hof, *Weight distribution in matching games* (University of Twente, 2016), and it is recorded for odd k in footnote 7 of Hof, Kern, Kurz and Paulusma (SAGT 2018). It is the graphical case of the Freixas–Kurz conjecture [10] . What is new here is the reduction from the canonical base to γ and the construction attaining the slope; the proof in the paper is a self-contained one by the minimal-norm method. For odd n the upper bound also follows from the counting in the proof of Lebedev's Theorem 9 [26] .

🛡️ **Evidence & scope**

- **Program:** n ≤ 6: rational dual certificates for 2, 8, 90 and 10 688 orbits (independent exact checker) and realizations anchored in one column; n ≤ 5 also by exact linear programming in Sage. `supplement: cert_canonica_3..6.json, zl_cert_canonica.py, cert_canonica_check.py, zl_rho_realiza.py -canonica, zl_rho_exacto.sage`

The formula agrees with the certified values for $n \leq 6$. The contribution of Theorem D is the reduction from augmentations of the canonical base to the graph parameter γ, with a construction attaining the resulting slope; the bound on γ is the known graphical case of a conjecture of Freixas and Kurz [10] . For odd n the upper bound $\kappa_n = (n-1)/4$ also

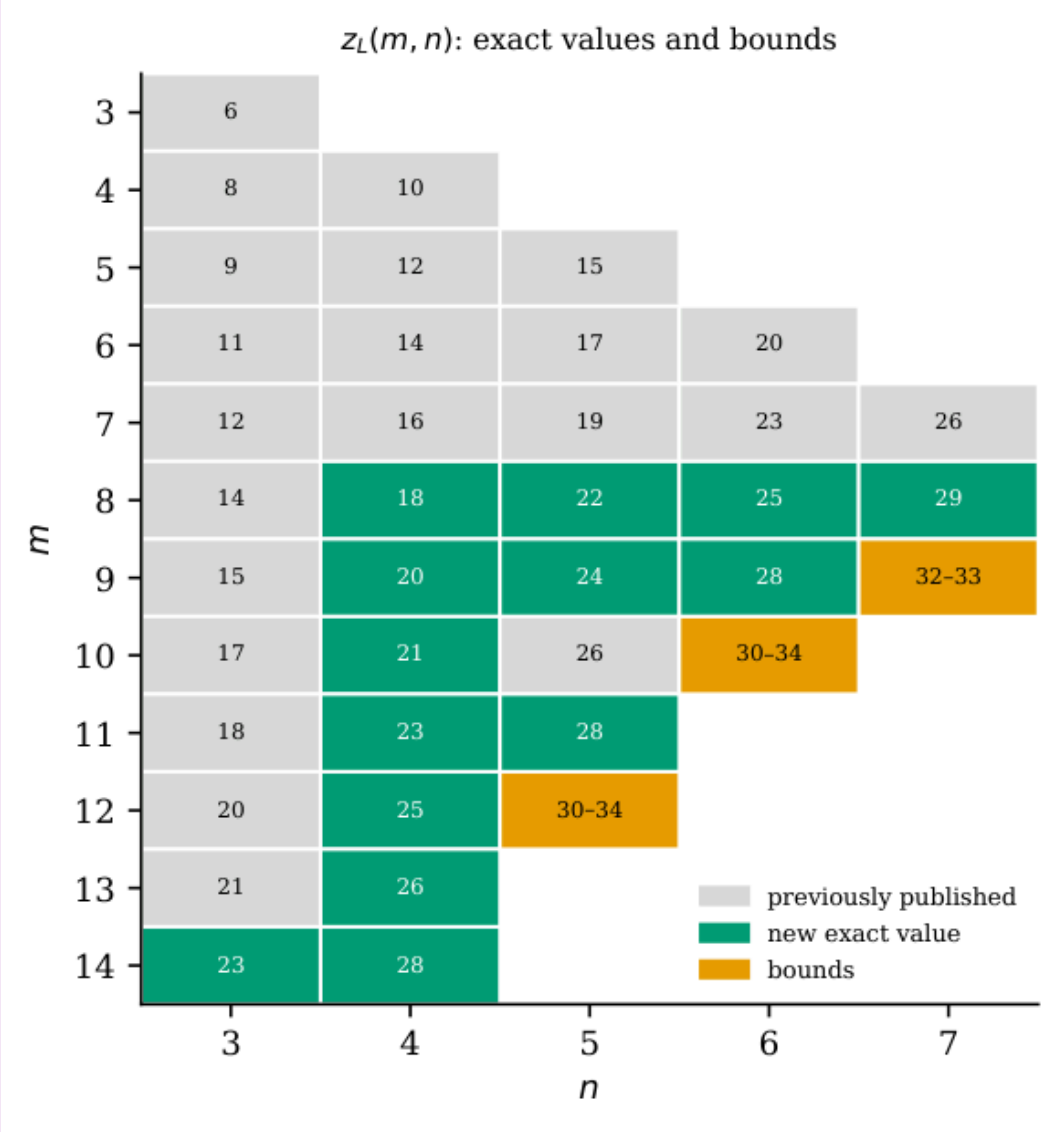
follows from the counting in the proof of [26 , Thm. 9] , which shows that the pairs on the canonical base avoiding the special rows number at most $(S+2)(n-1)/4$, where S is the number of singleton rows.

Card B — slopes in fixed width. For every n each new row buys ρ_n pairs, a rational number given by finitely many linear programs, up to a bounded error;
$\rho_n \geq \lambda_n = n/3 - 1/2 + O(1/n)$, with equality for $n \leq 6$.

7 Exact values

How far can one compute? Exact values have been part of the Zarankiewicz problem from the start: Guy's tables of the 1960s [12 , 13] were computed largely by hand, and their extension near the triple-system threshold is recent [3] . For small grids we enumerate the maximum bases up to isomorphism with nauty [21] and, for each, maximize $|E_2|$ by constraint programming. For $m \geq \binom{n}{2}$ the base has only special and singleton rows, so an integer program over the *numbers* of rows of each type, with integer flows between ports, gives $F_4(m)$ exactly: this reduces 14×4 from two hours to seconds and reaches $m = 40$. (From $n = 5$ on a row can carry several horizontal pairs, and counts by type no longer determine the rows; the exact formulation is then the profile system of the appendix, in which each type records its matching. Version 1 of this paper presented the type program for every n; its scope is $n \leq 4$, and no value of the table rests on it for $n \geq 5$.) Figure 6 collects the values; fourteen are new, of which the eight in three and four columns are instances of Theorems A and E, and six (8×5, 9×5, 11×5, 8×6, 9×6, 8×7) are not covered by any formula.

Figure 6



Left: the 39 values of $z_L(m,n)$ for $3 \le m \le 14$, $3 \le n \le 7$ that we computed, coloured by origin: 22 previously published (all reproduced), 14 new exact values, and 3 entries with lower and upper bounds (9×7, 10×6, 12×5). Right: $F_4(m) = z_L(m,4) - (m+6)$ for $6 \le m \le 40$ (staircase, from the type reduction) against the two floors $\lfloor (2m-3)/3 \rfloor$ and $\lfloor (3m-12)/4 \rfloor$; the lines cross at $m = 24$, and the second branch has the slope $3/4 = \rho_4$. Data in `table_zL.json` and `F4_exact.json`.

Can four columns be settled for every $m \ge 6$? Yes, by a finite description that is exact. For a configuration in the regime $m \ge \binom{n}{2}$ record, for every triple U of columns, its *profile* $p_U = \min(2, c_U)$, where c_U is the number of rows, special rows included, whose occupied support (base cells and endpoints) contains U. Then admissibility depends only on the profile and on the supports of the two rows of each pair: a vertical pair in column j is admissible iff $S \cap T = \{j\}$; a horizontal pair $\{j,q\}$ in a row of support S iff $p_{\{j,q,\ell\}} = 1$ for every $\ell \in S \setminus \{j,q\}$; a diagonal pair from (S,j) to (T,q) iff (C2) holds and, for every $\ell \in (S \cap T) \setminus \{j,q\}$, the triple $U = \{j,q,\ell\}$ lies in fewer than $1 + [U \subseteq S] + [U \subseteq T]$ rows (so that no row outside the pair contains it). Indeed a (C3) witness for a vertical pair is any further row occupied in j and in a common column ℓ, and the special row of $\{j,\ell\}$ is always one; for a horizontal or diagonal pair it is a further row whose support contains j, q and ℓ, which is counted in c_U whatever its base; and (C2) forbids both rows of a diagonal pair to

contain U. For a fixed profile, the numbers of rows of each type and of pairs between each two ports satisfy a finite system of linear equations and inequalities, written out in Appendix A together with the realization of its integer solutions.

Theorem E. p. 14

> *For every $m \geq 7$,*
>
> $$F_4(m) = \max\left\{\left\lfloor \frac{2m-3}{3} \right\rfloor, \left\lfloor \frac{3m-12}{4} \right\rfloor\right\},$$
>
> *and $F_4(6) = 2$. In particular $z_L(m,4) = m + 6 + F_4(m)$. The first floor is at least the second for $m \leq 27$ and the second, of slope $3/4 = \lambda_4$, is at least the first for $m \geq 22$; the lines cross at $m = 24$.*

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library; computer-assisted (certified rational duals), not formalized.

Proof: in this note · **Covers:** not formalized · **Cite:** p. 14, doi:10.5281/zenodo.23778502

↳ Used in: Theorem F (p. 31), 1 (p. 37), 2 (p. 38)

Proof

Proof (computer-assisted). The four triples give $3^4 = 81$ profiles, 15 up to permuting the columns; a displayed nondecreasing tuple such as $(0,0,2,2)$ denotes an orbit, its positions not tied to a fixed order of the triples (for $(0,0,2,2)$ we rename the columns so that the triples of profile 2 are 123 and 124). For each, the linear relaxation of the system above, with the valid inequality "at most one full row" (Lemma 5.2), has a rational dual certificate of $4f - 3m \leq -12$ for the profile $(0,0,2,2)$ and of $3f - 2m \leq -3$ for the other fourteen, where f is the number of pairs. The duals were computed in exact arithmetic (Sage) and checked by an independent program that rebuilds the types, ports and permitted pairs from scratch and verifies every column (29 certificates: the 15 orbits for $4f - 3m$, of which only $(0,0,2,2)$ is used, and 14 for $3f - 2m$; `cert_perfiles_check.py`). This gives the upper bound for every $m \geq 6$. For the lower bound, two explicit families, built from a seed by repeating a module of three rows and two pairs (slope $2/3$, from $m = 7$) or of four rows and three pairs (slope $3/4$, from $m = 10$), attain the two expressions; we checked them with both of our admissibility checkers for every $m \leq 200$, and repeating the module adds rows whose supports contain only triples already of profile 2, so the profile stays the same in each residue class and, by the exact criterion, the repeated families stay admissible for every m. The case $m = 6$ is in the table.

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

Four columns are settled for every $m \geq 6$: $F_4(m)$ is the larger of $\lfloor (2m-3)/3 \rfloor$ and $\lfloor (3m-12)/4 \rfloor$ (and $F_4(6) = 2$). Two straight lines, of slopes $2/3$ and $3/4$, cross at $m = 24$; after that the slope is $\rho_4 = 3/4$.

Why: record for each triple of columns whether it lies in 0, 1 or at least 2 rows (its *profile*). Admissibility depends only on the profile and the supports, so each of the 15 profile classes gives a finite linear system. Rational dual certificates, computed in Sage and checked by an independent program (29 certificates), give the upper bound; two explicit periodic families, checked for every $m \leq 200$ and with a profile that does not change when a module is added, give the lower bound.

Computer-assisted; the paper lists it as certified. The first branch has the slope $2/3$ of the canonical base (Theorem D).

Evidence & scope

- **Program:** 15 profile classes; 29 rational dual certificates (Sage, exact) verified column by column by an independent program that rebuilds types, ports and permitted pairs. `supplement: zl_perfiles.py, zl_perfiles_dual.sage, perfiles_n4/, cert_perfiles_check.py`
- **Program:** the two periodic families admissible under both checkers for every m ≤ 200, with constant profile per residue class. `supplement: f4_familias_check.py`
- **Program:** F_4(m) for 6 ≤ m ≤ 40 by the exact integer program over row types. `supplement: F4_exact.json, zl_tipos.py`

Theorem F. p. 14

> *For every $n \geq 3$ there are integers $P_n \geq 1$ and M_n such that $F_n(m + P_n) = F_n(m) + \rho_n P_n$ for all $m \geq M_n$. If $\rho_n = a/b$ in lowest terms, every such P_n is a multiple of b.*

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL Light/Metamath): not found in any formal library.

Proof: in this note · **Covers:** not formalized · **Cite:** p. 14, doi:10.5281/zenodo.23778502

↳ Used in: 2 (p. 38)

Proof

> By Appendix A, the set of pairs (m, f) such that some admissible configuration with a maximum base on an $m \times n$ grid, $m \geq \binom{n}{2}$, has f pairs is a finite union, over the profiles, of projections of sets of non-negative integer solutions of linear systems with constant coefficients; so it is definable in Presburger arithmetic, and so is the graph of its maximum F_n (add that no realizable $g > f$ exists). By Theorem C, $D(m) = bF_n(m) - am$ is an integer function taking finitely many

values. Each set $\{m : D(m) = k\}$ is definable in Presburger arithmetic, hence semilinear [11] , hence eventually periodic as a subset of $\mathbb{N}$. A common period P and threshold for these finitely many sets give $D(m + P) = D(m)$, which is the claim with $P_n = P$; since F_n is an integer, $\rho_n P_n$ must be one, so $b \mid P_n$.

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

For every n, after some point $F_n(m)$ repeats with a fixed period: $F_n(m + P_n) = F_n(m) + \rho_n P_n$. If $\rho_n = a/b$ in lowest terms, the period is a multiple of b.

Why: the exact profile system (Appendix A) describes all realizable pairs (m, f) by linear equations and inequalities over the integers, so F_n is definable in Presburger arithmetic. By Theorem C, $bF_n(m) - am$ takes finitely many values; each set where it takes one value is Presburger-definable, hence semilinear [11] , hence eventually periodic.

The proof gives neither the minimal period nor the threshold; for $n = 4$ Theorem E makes them explicit (period 4, threshold 22). For $n = 5, 6$ they are open (Question 2).

Theorem E is the case $n = 4$ made explicit, with minimal period 4 and threshold 22.

Card C — exact values. Fourteen new exact values (widths 3 to 7, $m \le 14$), bounds for three more entries, and $z_L(m, 4)$ for every $m \ge 6$ (Theorem E), with two linear branches of slopes $2/3$ and $3/4$.

8 One dimension more

If pairs live around a hollow triangle, do they live around a hollow tetrahedron one dimension up? The Zarankiewicz problem itself was lifted to higher dimensions early: Erdős showed in 1964 that an $n \times n \times n$ array of zeros and ones with no box of ones has $O(n^{3-1/4})$ ones [8] . Replace the matrix by an $m \times n \times p$ array, rectangles by boxes $\{a, a'\} \times \{b, b'\} \times \{c, c'\}$, and the base by a maximum box-free set. Two cells of O are adjacent when they are opposite corners of a box whose eight cells lie in O, and the two cells of a pair must be at distance at least three in this graph: the distance form of (C2)–(C3) from Section 2. In the plane every computed case has at most one pair per three free cells, and three blocks are a triangle; a hollow tetrahedron of four blocks suggests at most one pair per four free cells here, $4|E_2| \le mnp - |E_1|$. We take all sides at least three, and the guess fails already for $3 \times 3 \times 5$.

Proposition 8.1. p. 15

> *In $3 \times 3 \times 5$ a maximum box-free base has 34 cells, and there is an admissible family of 3 pairs, so that $4|E_2| = 12 > 11$ free cells.*

Lean · none found **show**

Searched 29-oct-2026 with existe-ya (our Lean, Mathlib, Loogle, LeanSearch, Isabelle/AFP, GitHub, the clones atlas-lean, TauCeti, openai-math, and Rocq/HOL

Light/Metamath): not found in any formal library; checked by computer, not formalized.

Proof: in this note · **Covers:** not formalized · **Cite:** Proposition 8.1, p. 15, doi:10.5281/zenodo.23778502

↳ Used in: 5 (p. 39)

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

One dimension up, matrices become $m \times n \times p$ arrays and rectangles become boxes. In the plane, every computed case has at most one pair per three free cells, matching the three blocks of a triangle; a hollow tetrahedron of four blocks suggests at most one pair per four free cells in 3D. The proposition shows that this guess already fails in $3 \times 3 \times 5$: a maximum box-free base has 34 cells, leaving 11 free cells, and 3 pairs fit, so $4|E_2| = 12 > 11$.

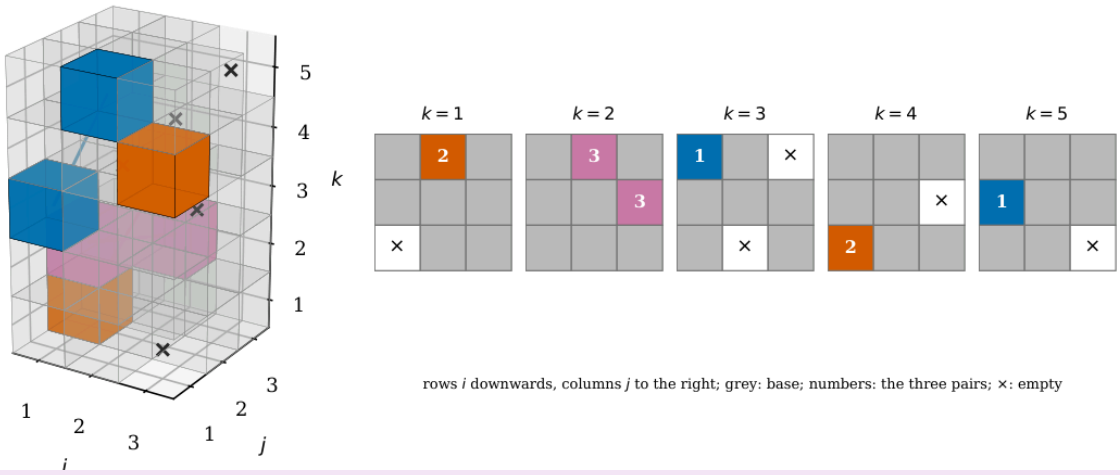
Why the base is maximum: a 3×3 slice with a ones contains at least $2a - 12$ full rectangles, and two slices of a box-free array never share one. Exhaustive search over the maximum bases of $3 \times 3 \times 3$ and $3 \times 3 \times 4$ finds no counterexample there, so $3 \times 3 \times 5$ is the smallest by volume with all sides at least three.

Whether this variant means anything for sums of squares of trilinear forms is left open in the paper.

🛡️ Evidence & scope

- **Program:** the witness: base box-free and of size 34, the three pairs at distance at least 3 in the box graph. `supplement: witness_3x3x5_tensor.json, figures/datos.py (tensor_3x3x5)`
- **Program:** the counting lemma on all 512 slices; the maximum box-free bases of 3×3×p up to permuting the slices (9, 162, 1413 for p = 3, 4, 5); the largest admissible family (0 and 2 pairs for p = 3, 4), so 3×3×5 is the smallest by volume. `supplement: tensor_busqueda.py (output: outputs/tensor_busqueda.txt)`

Figure 7



The $3 \times 3 \times 5$ configuration of Proposition 8.1. Left: in perspective. Right: the five 3×3 slices $k = 1, \dots, 5$, where every cell can be checked: the 34 base cells grey, the two cells of each of the 3 pairs with its number, and the 5 empty cells $\times$. The 11 free cells are 6 paired and 5 empty, so $4|E_2| = 12 > 11$. Data in `witness_3x3x5_tensor.json`.

The base is maximum because a 3×3 slice with a ones contains at least $2a - 12$ full rectangles and two slices of a box-free array never share one, so $2|E_1| - 60 \leq 9$. The configuration is shown in Figure 7 and listed in `witness_3x3x5_tensor.json`; we checked it with our own box-graph checker. It is the smallest example by volume with all sides at least three: the same counting shows that every maximum base of $3 \times 3 \times p$, $p \leq 5$, consists of slices with 6, 7 or 8 ones and pairwise disjoint sets of full rectangles, which leaves 9, 162 and 1413 bases up to permuting slices for $p = 3, 4, 5$, and an exhaustive search over them gives at most 0 pairs in $3 \times 3 \times 3$ and 2 pairs in $3 \times 3 \times 4$, where $4|E_2|$ only equals the number of free cells. Whether this variant has a meaning for sums of squares of trilinear forms is open.

Card D — one dimension more. The hollow-tetrahedron guess of one quarter fails already in $3 \times 3 \times 5$, the smallest volume where it can fail with all sides at least three.

9 What each statement rests on

Table 1 lists every statement with its level of guarantee, its basis and the file where it can be checked. The levels, coloured as in Figure 1, are: *formal* (proof checked by Lean or Isabelle/HOL), *certified* (finite certificate checked in exact arithmetic by an independent program), *computer* (exhaustive search or witness checked by our programs), *proof* (proof in the text), *evidence* (numerical only), and *partial* or *open*.

3whitelvH!7

Table 1: What each statement rests on.

Level	Statement	Basis	Where to check

lvF!28**formal**	Label, blocks, vertical, one full row	proofs in the text; Isabelle/HOL, from the properties of a maximum base (Formalizations)	theory `ZL_Labels`
black!18black lvF!28**formal**	Distance form of (C2)–(C3); endpoint bound	proofs in the text; Isabelle/HOL	theory `ZL_Counts`
black!18black lvF!28**formal**	Theorem A	proof in §5; Lean	module `LimitedAugmentedThree`
black!18black lvP!28**proof**	Weak gap $m/6 + O(1)$	Theorem A and the cited formula [30 , Thm. 26] ; the floor arithmetic is formal	theory `ZL_Counts` (`weak_gap_arith`)
black!18black lvP!28**proof**	Three-column ladder $\frac{1}{2}, \frac{2}{3}, 1$	Theorem A, [30] , [20]	`cruce_escalera_m3.py`
black!18black lvP!28**proof**	λ_n lower bound, all n	explicit modules; Lemma 4.1 (formal)	`lambda_modules.json`, two check ($n \leq 12$)
black!18black lvC!28**certified**	Slopes exact, $n \leq 6$	shadow reduction + rational dual certificates	`shadow_certificate_n3..6.js` independent checker
black!18black lvE!28**evidence**	Slope $9/5$ for $n = 7$	mixed integer program, floating point	not certified
black!18black lvP!28**proof**	Slope ρ_n exists, rational	Theorem C: reduction + realization	`zl_rho_realiza.py` ($n \leq 6$, two checkers)
black!18black lvP!28**proof**	Linear gap to z_{RL}, z_{SL}, z_2	Corollary 6.2: Theorem C, Lemma 5.2, [15]	proof in the text
black!18black lvP!28**proof**	Balanced blocks optimal	explicit modules and dual weights;	`tres_bloques_dual_check.py`, `tres_bloques_identidades.py`

	in the family (Prop. 6.3)	checked exactly for $n \leq 60$	
black!18black lvC!28**certified**	Canonical slopes, $n \leq 6$	rational dual certificates + realizations	`cert_canonica_3..6.json`, independent checker
black!18black lvP!28**proof**	Canonical slope, all n (Theorem D)	graph reduction; bound on γ from [18] , proved in the text; agrees with the certified $n \leq 6$	proof in the text
black!18black lvS!28**computer**	New exact values	nauty classes + CP-SAT; type reduction	`table_zL.json`
black!18black lvC!28**certified**	$F_4(m)$, every $m \geq 6$ (Theorem E)	profile system + rational duals; periodic families	`perfiles_n4/`, `cert_perfiles_check.py`; two checkers ($m \leq 200$)
black!18black lvS!28**computer**	Proposition 8.1	explicit witness, maximality by counting	`witness_3x3x5_tensor.json`, b graph checker
black!18black lvS!28**computer**	$3 \times 3 \times 5$ smallest by volume	exhaustive search over the maximum bases of $3 \times 3 \times 3$, $3 \times 3 \times 4$	`tensor_busqueda.py`
black!18black lvP!28**proof**	Exact profile system (Appendix A)	both directions proved in the text	`zl_perfiles.py`
black!18black lvP!28**proof**	Eventual periodicity, every n (Theorem F)	Theorem C, Appendix A, Presburger arithmetic [11]	proof in the text
black!18black			

Colour: red = no / fails.

10 The companion site

This paper has a companion website, its *blueprint*, at https://karlesmarin.github.io/blueprints/zl-three-blocks/. It shows the paper as a whole: what it proves, how its results depend on each other, and what exactly has been checked. There one finds:

- *The statements.* Every numbered statement, copied verbatim with its number (6 theorems, 3 propositions, 3 lemmas, 1 corollary, 1 definition and 6 questions), each with its proof, where it has one, a click away, and with a commentary in plain words.
- *The dependencies.* For each statement, what it uses and where it is used, as a list and as a graph.
- *The map.* One page with the main results, the tools they rest on and the fields they feed.
- *What was checked, and how.* Besides the written proof, the site says for each statement what else was checked: Theorem A in Lean 4 (8 declarations) and 3 lemmas in Isabelle/HOL (5 theorems), linked by name to the public repository karlesmarin/formalizations; 8 statements corroborated by programs, each named with what it computes.
- *Labs.* Next to Definition 2.1, Lemma 4.1, Proposition 6.3, Theorem E and Proposition 8.1, a form runs the computations of this paper in the reader's browser (Python, nothing to install) for values the reader chooses, checks every configuration against (C2) and (C3), and redraws Figures 2–7, in 3D where it helps; a die chooses random values. The captions of the figures and some phrases of the text link to these labs with their values already set.
- *The paper itself.* Its PDF, in English and in Spanish, opened at the right page from every statement.

The site is a reading aid, not a part of the proofs: the statements and proofs are those of this paper, whose text prevails, and a formal or computational check is reported for exactly the part it covers.

11 Open questions

Question 1. p. 17

Is $\rho_n = \lambda_n$ for every n? Equivalently (Theorem C), do three-block modules give $F_n(m) - O_n(1)$ pairs as $m \to \infty$? For $n \le 6$ three-block shadows attain the optimal value (certified); other sets of triples can attain it too (for $n = 3$, $\mathcal{H} = \{123\}$ does, with one vertical pair per two rows), so this is not a classification of the optimal programs. For $n = 7$ the numerical candidate 1.8 equals λ_7. The question is about slopes, not exact optima: no exactly optimal configuration obeys

the block rule of Lemma 4.1 when $n = 3$ and m is even. Indeed, then $2F_3(m) = m - 2$, while without a full row the special rows carry no endpoints and the other rows at most one each, so $2F_3(m) \le m - 3$; hence every optimum has a full row carrying an endpoint, and under the rule a full row carries none, since its type contains every block. Optimal configurations are also not all close to the rule: the even construction of Theorem A uses $(m - 4)/2$ vertical pairs, which the rule never allows, and the optimal four-column families of Theorem E repeat a module with a pair whose opposite corner is occupied, so in both cases a linear number of rows must be removed to reach the rule. Separate questions are the stability of the *supports* or of the shadow (are the occupied supports of almost every row unions of two of three blocks?) and the classification of all optimal pairings. A positive answer to the first question would fix the gap slope of Corollary 6.2 at $g_n = n/6 + \delta_n$; the relaxed criteria would then buy $(n - 1)/(2\lambda_n)$ times as many pairs per row as the original one, which is 2 for $n = 3, 4$, $16/9$ for $n = 5$, $5/3$ for $n = 6$, and tends to $3/2$ only as $n \to \infty$.

Cite: Question 1, p. 17, doi:10.5281/zenodo.23778502

↳ Not used by another result of the paper.

Corrected in v2 (1-nov-2026). Version 1 said that for $n \le 6$ the optimal shadows are three-block ones. The certificates prove less: three-block shadows *attain* the optimal value. Other sets of triples can attain it too: for $n = 3$ the program of $\mathcal{H} = \{123\}$ reaches $1/2$ with $x_{13,1} = x_{23,2} = 1/2$ and one vertical pair in column 3, while the three-block shadow in three columns is empty. So the certificates do not classify the optimal programs.

Question 2. p. 18

Theorem F says that $F_n(m) - \rho_n m$ is eventually periodic for every n, and Theorem E gives period 4 and threshold 22 for $n = 4$. What are the minimal period and threshold, and the explicit formula, for $n = 5, 6$? The denominators force the period to be a multiple of 8 and 2; the profile certificates of Section 7 could settle them, as they did for $n = 4$.

Cite: Question 2, p. 18, doi:10.5281/zenodo.23778502

↳ Not used by another result of the paper.

Question 3. p. 18

Is there an exact inequality $3|E_2| + |E_1| \le mn$ for every maximum base, without the error terms of the density bound? It fails if the base is not maximum, and it holds in every case we computed. For $n = 3, 4$ and $m \ge \binom{n}{2}$ it follows from the bound $2F_n(m) \le (n - 2)m - \binom{n}{2} + 1$ of Section 5; from $n = 5$ on that bound is too weak (it allows a violation from $m = 11$ for $n = 5$). Asymptotically the inequality says $\rho_n \le (n - 1)/3$, which holds for $n \le 6$ by the certified slopes (

$9/8<4/3$, $3/2<5/3$) and would hold for every n if $\rho_n = \lambda_n$, since $\lambda_n<(n-1)/3$; the exact inequality for every m is open from $n = 5$ on.

Cite: Question 3, p. 18, doi:10.5281/zenodo.23778502

↳ Not used by another result of the paper.

Question 4. p. 18

What replaces the three blocks when $m<\binom{n}{2}$, where maximum bases are linear hypergraphs with rows of several sizes rather than pairs of columns, for instance when the base is a projective plane?

Cite: Question 4, p. 18, doi:10.5281/zenodo.23778502

↳ Not used by another result of the paper.

Question 5. p. 18

Which three-dimensional analogue, if any, bounds the sum-of-squares rank of trilinear forms, and what plays the role of the hollow triangle there?

Cite: Question 5, p. 18, doi:10.5281/zenodo.23778502

↳ Not used by another result of the paper.

Question 6. p. 18

The canonical base (all singleton ones in one column) has slope $\lfloor (n-1)^2/4 \rfloor/(n-1)$ (Theorem D), and the best bases reach ρ_n. What is the slope for a prescribed distribution of the singleton ones over the columns? Adding these frequencies to the shadow programs gives a parametric family of linear programs that contains both the canonical and the optimal distributions; is the slope a concave function of the distribution?

Cite: Question 6, p. 18, doi:10.5281/zenodo.23778502

↳ Not used by another result of the paper.

A The exact profile system

Fix $n \geq 3$, $m \geq \binom{n}{2}$ and a profile $p \in \{0,1,2\}^{\binom{[n]}{3}}$. A *type* is a triple $\tau = (B, S, M)$ with $B \subseteq S \subseteq [n]$, $|B| \in \{1,2\}$, no triple of profile 0 inside S, and M a partial matching of $S \setminus B$ (the horizontal pairs of the row) such that $p_{\{j,q,\ell\}} = 1$ for every $\{j,q\} \in M$ and $\ell \in S \setminus \{j,q\}$

. Its *ports* are the columns $J_\tau = S \setminus (B \cup V(M))$: cells that must be paired with another row. Two ports (τ, j), (τ', q) of different types with supports S, S' are *permitted* if $j = q$ and $S \cap S' = \{j\}$, or $j \neq q$, not both $q \in S$ and $j \in S'$, and every $U = \{j, q, \ell\}$ with $\ell \in (S \cap S') \setminus \{j, q\}$ has $p_U < 1 + [U \subseteq S] + [U \subseteq S']$.

Variables: $x_\tau \in \mathbb{Z}_{\geq 0}$ (rows of type τ) and $y_e \in \mathbb{Z}_{\geq 0}$ (pairs on each permitted pair e of ports). Constraints:

$$\sum_\tau x_\tau = m, \qquad \sum_{\tau:\, B_\tau = B} x_\tau = 1 \quad \left(B \in \tbinom{[n]}{2}\right),$$

$$c_U = \sum_{\tau:\, U \subseteq S_\tau} x_\tau : \quad c_U = 0,\ c_U = 1,\ c_U \geq 2 \ \text{ according as } p_U = 0, 1, 2,$$

$$\sum_{e \ni (\tau, j)} y_e = x_\tau \qquad (\tau,\ j \in J_\tau).$$

The objective is $f = \sum_\tau |M_\tau| x_\tau + \sum_e y_e$. The port equations are equalities: every cell of $S \setminus B$ announced in a support is used by exactly one pair, so the realized supports are exactly the announced ones and the counts c_U are the true ones. (In Section 6 capacities sufficed, because supports there are frozen labels and real supports may be smaller.)

From configurations to solutions. Recording for every row its base, its occupied support and its horizontal pairs, and counting the other pairs between ports, gives a solution for the profile of the configuration: the constraints are counts, and every pair is permitted by the exact criterion of Section 7.

From solutions to configurations. Given an integer solution, create x_τ rows of each type, with base B_τ and the horizontal pairs M_τ. For each port (τ, j), the equation lets us assign the pairs of the incident edges to distinct rows of type τ, using column j of each; along each edge, match the two lists. The two cells of a pair lie in rows of different types, hence in different rows. Every non-base cell of every support is used exactly once, so the supports and the counts c_U are as prescribed, the profile is p, and the exact criterion makes every pair admissible. The rows with $|B| = 2$ give one row per pair of columns and the others one cell each, so the base is maximum. Hence the set of pairs (m, f) realized by admissible configurations is the union over the profiles of the projections of the solution sets, as used in Theorems E and F; for Theorem E the linear relaxation is strengthened by the valid inequality "at most one full row" (Lemma 5.2).

Formalizations

Lean 4. Theorem A with the definitions of Section 2: module `Formalizations.Combinatorics.Zarankiewicz.LimitedAugmentedThree` of the repository https://github.com/karlesmarin/formalizations, commit `389a576`, Lean `v4.35.0`, Mathlib `f0b4c81`; standard axioms only, checked by the repository's continuous integration.

Isabelle/HOL. Directory `Isabelle/AFP/Limited_Augmented_Zarankiewicz` of the same repository, commit `5754a1f`, built with Isabelle2026 without `quick_and_dirty`; its current status is on the companion site. Theory `ZL_Labels`: the label lemma (`label_admissible`), Lemma 4.1 (`blocks_admissible`), Lemma 5.1 (`vertical_pair`) and Lemma 5.2 (`at_most_one_full_row`). Theory `ZL_Counts`: the distance form of (C2)–(C3) (`C2_iff_not_adj`, `C3_iff_no_common_neighbour`), the bound $2F_n(m) \le (n-2)m - \binom{n}{2} + 1$ (`endpoint_bound`) and the floor arithmetic of the weak gap (`weak_gap_arith`). The two lemmas on maximum bases and the bound are proved inside a locale whose hypotheses are the properties of a maximum base used in the text: every pair of columns is covered by a base row, every occupied row has one or two base cells, all cells lie in the grid, and $|E_1| = m + \binom{n}{2}$.

Data and code

The supplementary archive deposited with this paper on Zenodo (`zl-three-blocks-supplement.zip`, also browsable on the companion site of Section 10) contains, with the complete outputs of the checks: the witnesses (`data/`, JSON, indices from 1); the shadow certificates for $n \le 6$ (`shadow_certificate_n3..6.json`) with the generator and the independent checker (`zl_cert_sombras.py`, `cert_sombras_check.py`); the canonical certificates (`cert_canonica_3..6.json`, `zl_cert_canonica.py`, `cert_canonica_check.py`); the profile system for four columns (`zl_perfiles.py`, `zl_perfiles_dual.sage`, the 29 certificates and `cert_perfiles_check.py`) and the check of the periodic families (`f4_familias_check.py`); the realizations (`zl_rho_realiza.py`), the module builder, the type reduction, the exact linear programs in Sage, the three-block check (`tres_bloques_dual_check.py`), the exhaustive search of Section 8 (`tensor_busqueda.py`), the two admissibility checkers, and the figure scripts, which re-check every witness before drawing it. The exact values of the table come from exhaustive search (nauty and CP-SAT reaching proven optimality), not from independently checked proof certificates.

Use of AI tools

AI assistants were used as research assistants throughout, for computation, writing and review. The author checked every argument as described in the text and takes full responsibility for the content; results that are only computed or only proposed are marked as such in Table 1.

📊 Tables

Every table of the paper, in its order: click one to unfold it, with where the paper cites it, how its numbers were checked, and its data as CSV. The typeset table is one click away, on its page of the PDF.

Table 1. What each statement rests on.

Level	Statement	Basis	Where to check
lvFformal	Label, blocks, vertical, one full row	proofs in the text; Isabelle/HOL, from the properties of a maximum base (Formalizations)	theory `ZL_Labels`
black!18black lvFformal	Distance form of (C2)–(C3); endpoint bound	proofs in the text; Isabelle/HOL	theory `ZL_Counts`
black!18black lvFformal	Theorem A	proof in §5; Lean	module `LimitedAugmentedThree`
black!18black lvPproof	Weak gap $m/6 + O(1)$	Theorem A and the cited formula [Thm. 26]QLC; the floor arithmetic is formal	theory `ZL_Counts` (`weak_gap_arith`)
black!18black lvPproof	Three-column ladder $\frac{1}{2}, \frac{2}{3}, 1$	Theorem A, [QLC], [LQ]	`cruce_escalera_m3.py`
black!18black lvPproof	λ_n lower bound, all n	explicit modules; Lemma 4.1 (formal)	`lambda_modules.json`, two checkers ($n \leq 12$)
black!18black lvCcertified	Slopes exact, $n \leq 6$	shadow reduction + rational dual certificates	`shadow_certificate_n3..6.json`, independent checker
black!18black lvEevidence	Slope $9/5$ for $n = 7$	mixed integer program, floating point	not certified
black!18black lvPproof	Slope ρ_n exists, rational	Theorem C: reduction + realization	`zl_rho_realiza.py` ($n \leq 6$, two checkers)
black!18black lvPproof	Linear gap to z_{RL}, z_{SL}, z_2	Corollary 6.2: Theorem C,	proof in the text

		Lemma 5.2, [LebII]	
black!18black lvPproof	Balanced blocks optimal in the family (Prop. 6.3)	explicit modules and dual weights; checked exactly for $n \leq 60$	`tres_bloques_dual_check.py`, `tres_bloques_identidades.py`
black!18black lvCcertified	Canonical slopes, $n \leq 6$	rational dual certificates + realizations	`cert_canonica_3..6.json`, independent checker
black!18black lvPproof	Canonical slope, all n (Theorem D)	graph reduction; bound on γ from [HofMSc], proved in the text; agrees with the certified $n \leq 6$	proof in the text
black!18black lvScomputer	New exact values	nauty classes + CP-SAT; type reduction	`table_zL.json`
black!18black lvCcertified	$F_4(m)$, every $m \geq 6$ (Theorem E)	profile system + rational duals; periodic families	`perfiles_n4/`, `cert_perfiles_check.py`; two checkers ($m \leq 200$)
black!18black lvScomputer	Proposition 8.1	explicit witness, maximality by counting	`witness_3x3x5_tensor.json`, box-graph checker
black!18black lvScomputer	$3 \times 3 \times 5$ smallest by volume	exhaustive search over the maximum bases of $3 \times 3 \times 3$, $3 \times 3 \times 4$	`tensor_busqueda.py`
black!18black lvPproof	Exact profile system (Appendix A)	both directions proved in the text	`zl_perfiles.py`
black!18black lvPproof	Eventual periodicity, every n (Theorem F)	Theorem C, Appendix A, Presburger arithmetic [GinsburgSpanier]	proof in the text
black!18black			

Colour: red = no / fails.

p. 16 in the PDF ⬇ CSV cited in §9 What each statement rests on cited in §A The exact profile system

References

1

A. P. Calderón, A note on biquadratic forms, Linear Algebra Appl. 7 (1973), 175–177, doi:10.1016/0024-3795(73)90052-9.

2

G. Chen, D. Horsley, A. Mammoliti, Exact values for some unbalanced Zarankiewicz numbers, J. Graph Theory 106 (2024), 81–109, doi:10.1002/jgt.23068.

3

G. Chen, D. Horsley, A. Mammoliti, Zarankiewicz numbers near the triple system threshold, J. Combin. Des. 32 (2024), 556–576, doi:10.1002/jcd.21948.

4

Z. Chen, Y. Chen, Recursive-line Zarankiewicz numbers with four columns, arXiv:2609.11093v2 (2026).

5

M.-D. Choi, Positive semidefinite biquadratic forms, Linear Algebra Appl. 12 (1975), 95–100, doi:10.1016/0024-3795(75)90058-0.

6

M.-D. Choi, T.-Y. Lam, B. Reznick, Sums of squares of real polynomials, in: K-Theory and Algebraic Geometry (Santa Barbara, 1992), Proc. Sympos. Pure Math. 58, Part 2, Amer. Math. Soc., 1995, 103–126, doi:10.1090/pspum/058.2/1327293.

7

K. Čulík, Teilweise Lösung eines verallgemeinerten Problems von K. Zarankiewicz, Ann. Polon. Math. 3 (1956), 165–168, doi:10.4064/ap-3-1-165-168.

8

P. Erdős, On extremal problems of graphs and generalized graphs, Israel J. Math. 2 (1964), 183–190, doi:10.1007/BF02759942.

9

Z. Füredi, M. Simonovits, The history of degenerate (bipartite) extremal graph problems, in: Erdős Centennial, Bolyai Soc. Math. Stud. 25, Springer, 2013, 169–264, doi:10.1007/978-3-642-39286-3_7.

10

J. Freixas, S. Kurz, On α-roughly weighted games, Internat. J. Game Theory 43 (2014), 659–692, doi:10.1007/s00182-013-0402-x.

11

S. Ginsburg, E. H. Spanier, Semigroups, Presburger formulas, and languages, Pacific J. Math. 16 (1966), 285–296, doi:10.2140/pjm.1966.16.285.

12

R. K. Guy, A problem of Zarankiewicz, in: Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, 119–150. 🔎 find it

13

R. K. Guy, A many-facetted problem of Zarankiewicz, in: The Many Facets of Graph Theory, Lecture Notes in Math. 110, Springer, 1969, 129–148, doi:10.1007/BFb0060112.

14

T. Kővári, V. T. Sós, P. Turán, On a problem of K. Zarankiewicz, Colloq. Math. 3 (1954), 50–57, doi:10.4064/cm-3-1-50-57.

15

N. Lebedev, Asymptotic equivalence and exact values for second-order Zarankiewicz numbers, arXiv:2610.05442v1 (2026).

16

F. Hof, W. Kern, S. Kurz, K. Pashkovich, D. Paulusma, Simple games versus weighted voting games: bounding the critical threshold value, Soc. Choice Welf. 54 (2020), 609–621, doi:10.1007/s00355-019-01221-6.

17

F. Hof, W. Kern, S. Kurz, D. Paulusma, Simple games versus weighted voting games, in: Algorithmic Game Theory (SAGT 2018), Lecture Notes in Comput. Sci. 11059, Springer, 2018, 69–81, doi:10.1007/978-3-319-99660-8_7.

18

F. Hof, Weight distribution in matching games, individual research assignment, master's programme Science Education and Communication, University of Twente, March 2016, https://essay.utwente.nl/69350/.

19

H. Liu, Y. Song, New bounds for limited Zarankiewicz numbers from K_{5t} blocks, arXiv:2609.07227v1 (2026).

20

J. Löfberg, L. Qi, Second order Zarankiewicz number, arXiv:2608.30555v3 (2026).

21

B. D. McKay, A. Piperno, Practical graph isomorphism, II, J. Symbolic Comput. 60 (2014), 94–112, doi:10.1016/j.jsc.2013.09.003.

22

I. Reiman, Über ein Problem von K. Zarankiewicz, Acta Math. Acad. Sci. Hungar. 9 (1958), 269–273, doi:10.1007/BF02020254.

23 P. Turán, Eine Extremalaufgabe aus der Graphentheorie, Mat. Fiz. Lapok 48 (1941), 436–452 (in Hungarian). 🔎 find it

24 K. Zarankiewicz, Problem P 101, Colloq. Math. 2 (1951), 301. 🔎 find it

25 J. Howard-Stanley, Supplement for “The exact limited augmented Zarankiewicz number $z_L(6,6)$”, Zenodo (2026), doi:10.5281/zenodo.21816535.

26 N. Lebedev, Density and separation for augmented Zarankiewicz numbers, arXiv:2609.16555v1 (2026).

27 L. Qi, C. Cui, Y. Xu, Biquadratic SOS rank and augmented Zarankiewicz number, Mathematics 14 (2026), no. 9, 1552, doi:10.3390/math14091552; arXiv:2603.04912.

28 L. Qi, C. Cui, Y. Xu, A general lower bound for the limited augmented Zarankiewicz number based upon complete graphs, arXiv:2604.04111v5 (2026); published as: New lower bounds for the limited augmented Zarankiewicz number based on complete graphs, Symmetry 18 (2026), no. 7, 1076, doi:10.3390/sym18071076.

29 L. Qi, C. Cui, Y. Xu, A weak condition for limited augmented Zarankiewicz numbers, arXiv:2607.07071v4 (2026).

30 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 (2026).

31 Y. Xu, G. Yu, L. Qi, Optimization models and computational bounds for limited augmented Zarankiewicz numbers in the incidence-graph family of complete graphs, arXiv:2605.29658v2 (2026).