FOUR-CYCLES ONE EDGE ABOVE THE EXTREMAL THRESHOLD

WILL HOWE, GRANT HUFFMAN, JENSEN KOHLMEYER, AND LIAM KRUER

ABSTRACT. Let $h(n)$ be the minimum number of four-cycles in an n-vertex graph with $\mathrm{ex}(n, C_4)+1$ edges. We prove that $h(n_j) = o(\sqrt{n_j})$ for a sequence $n_j = q_j^2 + q_j + 2$, where the q_j are powers of four tending to infinity. Consequently, no positive constant gives a universal square-root lower bound one edge above the extremal number. The proof combines a polarity-graph construction with a structural analysis of extremal graphs. Signed deficiencies and a secondary degree-and-triangle minimization produce an open packing of size q. Completing its adjacency matrix reduces the argument to a trace inequality for a symmetric binary matrix whose residual graph is a union of cycles.

hexagon:2611.00751v1

1. INTRODUCTION

For a finite simple graph H, let $N_4(H)$ denote the number of subgraphs isomorphic to C_4. A copy is counted by its four cycle edges. Additional edges among its vertices are allowed. Write

$$h(n) = \min\{N_4(H) : |V(H)| = n,\ e(H) = \mathrm{ex}(n, C_4) + 1\},$$

where $\mathrm{ex}(n, C_4)$ is the maximum number of edges in a C_4-free graph of order n. The square-root supersaturation question, commonly listed as Erdős problem 60, asks whether $h(n) \geq c\sqrt{n}$ for some constant $c > 0$ and all sufficiently large n. The problem belongs to the supersaturation program of Erdős and Simonovits [3, 6].

Theorem 1.1. *There are strictly increasing powers of four q_j and finite simple graphs H_j such that*

$$|V(H_j)| = q_j^2 + q_j + 2, \qquad e(H_j) = \mathrm{ex}(q_j^2 + q_j + 2, C_4) + 1, \qquad \frac{N_4(H_j)}{q_j} \longrightarrow 0.$$

In particular, $\liminf_{n\to\infty} h(n)/\sqrt{n} = 0$.

Polarity graphs supply the classical lower bound for $\mathrm{ex}(q^2+q+1, C_4)$ [1, 2], and Füredi proved the matching upper bound [4]. The vertex-splitting construction used below is present in He, Ma, and Yang's original preprint [5, Section 10]. Ning and Zhai's 2021 preprint records a stronger announced conclusion of He, Ma, and Yang, namely $h(q^2 + q + 2) = 1$ for large powers of four [7, Introduction]. We give a detailed proof of Theorem 1.1 through an extremal-graph and matrix argument.

The proof uses the stability theorem of He, Ma, and Yang [6, Theorem 3.1] to obtain a minimum-degree estimate. All subsequent structural steps are proved below. The main alternatives are a small excess over the polarity edge count and an edge addition that creates few four-cycles. If both alternatives fail, a signed deficiency analysis produces a highly constrained open packing. A matrix completion then forces the excess to be at most three, contradicting its assumed growth.

All asymptotic estimates concern $q \to \infty$ along a specified sequence of powers of four. Implied constants may depend on the fixed constant ϵ introduced below and, where indicated, on fixed auxiliary constants. Estimates are uniform in the vertices under consideration. For a real number x, write $x_+ = \max\{x, 0\}$. For distinct vertices v, w, their *codegree* is $c(v, w) = |N(v) \cap N(w)|$. A graph

Date: November 12, 2026.

2020 *Mathematics Subject Classification.* 05C35, 05C50, 51E15.

Key words and phrases. Four-cycle, extremal number, supersaturation, polarity graph, signed deficiency.

The authors acknowledge assistance from OpenAI's AI systems in developing and checking intermediate arguments and in preparing the mathematical exposition.

is C_4-free precisely when all these codegrees are at most one. An edge is *triangular* if it belongs to a triangle, and *untriangulated* otherwise. Every edge of a C_4-free graph belongs to at most one triangle.

2. Extremizers and degree estimates

Put $n = q^2 + q + 2$ and $E_q = q(q+1)^2/2$. Among the C_4-free graphs on n vertices with $\mathrm{ex}(n, C_4)$ edges, choose G to minimize

$$\Phi(G) = \sum_{v \in V(G)} d(v)^2 + 2N_3(G),$$

where $N_3(G)$ counts triangles. Define

$$a = e(G) - E_q, \qquad \gamma = \min_{vw \notin E(G),\ v \neq w} N_4(G + vw).$$

Adding a nonedge vw creates exactly as many four-cycles as there are length-three paths from v to w. Since $c(v, x) \leq 1$ for $x \neq v$, each neighbor x of w contributes at most one such path. In particular, if r neighbors of w have codegree zero with v, then

$$N_4(G + vw) \leq d(w) - r. \tag{2.1}$$

Lemma 2.1 (Polarity splitting)**.** *For every power of two q and every integer $1 \leq r \leq q$, there is a graph on $q^2 + q + 2$ vertices with $E_q + r + 1$ edges and exactly $\binom{r}{2}$ four-cycles. Consequently, $a \geq 2$.*

Proof. Consider the one-dimensional subspaces of $\mathbb{F}_q^3$, with two distinct points adjacent when their representatives are orthogonal under the standard dot product. There are $q^2 + q + 1$ points. The absolute points lie on the line $x_1 + x_2 + x_3 = 0$. Each of its $q + 1$ points has degree q, and every other point has degree $q + 1$. Thus the graph has E_q edges. Distinct polar lines meet in one point, so distinct graph vertices have at most one common neighbor.

Let u be absolute. Its neighborhood is independent. Indeed, if distinct neighbors x, y of u were adjacent, then x would be orthogonal to both u and y. Since u, y span $u^\perp$, this would give $x \in (u^\perp)^\perp = \langle u \rangle$, a contradiction. Replace u by adjacent vertices u_0, u_1 whose old neighborhoods cover $N(u)$ and overlap in exactly r points. The edge count increases by $r + 1$. Any four-cycle must contain both new vertices. Independence of $N(u)$ prevents them from being consecutive on such a cycle. The cycles are therefore in bijection with the pairs of their r common neighbors. Taking $r = 1$ proves $a \geq 2$. □

We will exclude an infinite sequence of the chosen extremizers for which, for a fixed $0 < \epsilon < 1$,

$$a \geq \epsilon\sqrt{q}, \qquad \gamma \geq \epsilon q. \tag{2.2}$$

In particular, $a \to \infty$ on such a sequence. Until the final section, assume (2.2). Passage to subsequences will always preserve this assumption.

Proposition 2.2. *No infinite sequence of powers of four and chosen extremizers satisfies* (2.2)*.*

Lemma 2.3. *Along a sequence satisfying* (2.2)*,*

$$\max_v d(v) \leq q + 2, \qquad |\{v : d(v) = q + 2\}| = O(1), \qquad \min_v d(v) \geq (1/2 - o(1))q. \tag{2.3}$$

Proof. For the minimum degree, use [6, Theorem 3.1]. For every fixed $c \in (0, 1)$ and sufficiently large even q, a C_4-free graph of order $q^2 + q + 1$ with at least $E_q - cq/2$ edges is contained in a polarity graph of order q. Such a graph has at most E_q edges by Füredi's bound [4], and its diameter is at most two.

We first prove the attachment fact needed here. Suppose that P has diameter at most two and that deleting an edge set D and adjoining one new vertex with neighborhood $T \subseteq V(P)$ produces a C_4-free graph. Then

$$|T| - |D| \leq 2. \tag{2.4}$$

In the retained graph every old vertex has at most one neighbor in T. Construct an auxiliary graph on T. Include the deleted edges with both endpoints in T. For each $w \notin T$ with a retained neighbor in T, join each deleted neighbor of w in T to that retained neighbor. If w has no retained neighbor in T, join its deleted neighbors in T by a tree. The auxiliary graph uses at most $|D|$ edges and has at least $|T| - |D|$ components.

For vertices x, y in different components, either x has a retained neighbor in the component of y, or y has one in the component of x. If xy is an edge of P, it survives. Otherwise choose a common neighbor w in P. When $w \notin T$, the construction forces x, y into the same auxiliary component. When $w \in T$, both incident edges cannot survive, and both cannot be deleted. The deleted edge puts w in one endpoint's component, and the retained edge establishes the assertion for the other endpoint.

For each pair of components, this property permits a direction in which every vertex of the source component has a retained neighbor in the target. To justify uniformity, if one source vertex has no neighbor in the other component, every vertex of that other component must have a neighbor in the source. No component can have two outgoing directions, since each of its vertices has at most one retained neighbor in T. A directed triangle is impossible as well. An edge entering its middle component would give its endpoint both the incoming neighbor and a neighbor in the next component. Every orientation of three component pairs has a vertex of outdegree two or is a directed triangle. Hence there are at most two components, proving (2.4).

If $d(v) \le a + cq/2$, then $G - v$ has at least $E_q - cq/2$ edges and is contained in a polarity graph P. Recovering G from P deletes some old edges and attaches v. Since $e(P) \le E_q$, the net edge gain is at least a, contradicting (2.4) once $a \ge 3$. Thus $d(v) > a + cq/2$ for every fixed $c < 1$ and all large q. This proves the minimum-degree assertion.

For the upper bounds, count the two-edge paths centered at $w \ne v$ whose endpoints lie outside $N(v)$, where $d(v) = D_0$. Write $x_w = |N(w) \setminus N(v)| \ge d(w) - 1$. Then $\sum_{w \ne v} x_w \ge q^3 + q^2 + 5 - D_0$, using $a \ge 3$. No unordered endpoint pair occurs twice, so $\sum_{w \ne v} \binom{x_w}{2} \le \binom{n - D_0}{2}$. The integer inequality $\binom{x}{2} \ge \binom{q-1}{2} + (q-1)(x - q + 1)$ gives

$$D_0^2 - (2q^2 + 5)D_0 + 2q^3 + 6q^2 - 8q + 12 \ge 0.$$

The polynomial is negative at both $q + 4$ and $n - 1$ for large q. Its convexity excludes the intervening interval, so $D_0 \le q + 3$.

Set $f_v = q + 1 - d(v)$, $F = \sum_v f_v = 2q + 2 - 2a$, and let Z_0 be the adjacency matrix of zero-codegree pairs of distinct vertices. Counting two-edge paths gives

$$\sum_{v,w} (Z_0)_{vw} = n + 2qF - \sum_v f_v(f_v - 1) = n + (2q + 1)F - \sum_v f_v^2.$$

Since $F \le 2q - 2$ and the left side is nonnegative, this yields $F = O(q)$, $\sum_v f_v^2 = O(q^2)$, and $\|Z_0\|_{\mathrm{F}} = O(q)$. If A is the adjacency matrix, then $A^2 - J = \mathrm{diag}(d - 1) - Z_0$. Consequently, $\|Ax\|^2 = O(q)\|x\|^2$ for $x \perp \mathbf{1}$.

Call a vertex high if $d(v) \ge q + 2$. Its nonreturning two-walks give $(Af)_v \ge qd(v) - (n - 1) \ge q - 1$. The vector $f - (F/n)\mathbf{1}$ is perpendicular to $\mathbf{1}$, and centering changes these coordinates of Af by $O(1)$. The squared norm of its image under A is $O(q^3)$. Each high vertex contributes $\Omega(q^2)$, so there are initially $O(q)$ high vertices. It follows that $F_+ = \sum_v (f_v)_+ = O(q)$.

If v has r high neighbors, their neighborhoods outside v are pairwise disjoint. Each carries positive deficiency weight at least $q - 1 - (f_v)_+$. The minimum-degree bound gives $(f_v)_+ \le (1/2 + o(1))q$, hence $r = O(1)$. Summing positive deficiency incidences at the high vertices now shows that there are $O(1)$ of them and $F_+ \le 2q + O(1)$. A vertex of degree $q + 3$ requires positive neighbor weight at least $2q - 1$. If it is the only high vertex, $F_+ = F + 2 \le 2q - 2$. If another high vertex exists, the two required weights sum to at least $3q - 2$, and their overlap is at most $(1/2 + o(1))q$. Both possibilities contradict the bound on F_+. Every high vertex therefore has degree $q + 2$. □

3. Defect identities and a triangle cover

Write $S = \{v : f_v > 0\}$, $R = \{v : f_v = 0\}$, $H_* = \{v : f_v = -1\}$, and $K = \max_v(f_v)_+$. Then $F_+ = F + |H_*| \le 2q + O(1)$, and $|S| \le F_+$. A vertex of S is called low, a vertex of R regular, and a low of deficiency one is called unit. Put

$$B = A + \operatorname{diag} f, \quad Z = qI + J - B^2, \quad c_f(v) = \sum_{w \in N(v)} f_w, \quad c_f^+(v) = \sum_{w \in N(v)} (f_w)_+.$$

Lemma 3.1. *The matrices B, Z commute and have row sums $q+1, 1$. Write $c(v,w)$ for codegree of distinct vertices and $z(v)$ for the row sum of Z_0. Then*

$$Z_{vv} = f_v - f_v^2, \quad Z_{vw} = 1 - c(v,w) - A_{vw}(f_v + f_w) \ (v \ne w), \quad z(v) = 1 + qf_v + c_f(v). \tag{3.1}$$

Proof. The row sum of B is $d(v) + f_v = q + 1$, so $BJ = JB = (q+1)J$. Thus B commutes with $qI + J - B^2$, whose row sum is $q + n - (q+1)^2 = 1$. The diagonal and off-diagonal entries follow by expanding B^2. Finally, $z(v) = n - 1 - \sum_{w \in N(v)}(d(w) - 1)$, which gives the last identity. □

Let $P_v = \{w : Z_{vw} > 0\}$. Every positive entry in a low row has weight one, and its nonhigh partners are nonadjacent with zero codegree. If $f_l = k \ge 2$, the exact positive row size is

$$|P_l| = (k-1)q + k + c_f(l) + 2t_l, \tag{3.2}$$

where t_l is the number of triangles through l. This follows by summing the negative diagonal $k(k-1)$ and the weights $k - 1 + f_w + c(l,w)$ at neighbors, then adding one. For $k = 1$ the same expression holds with an additive $O(1)$ error from high neighbors. A regular row has positive mass at most $1 + c_f^+(v)$, since negative entries occur only on edges to lows with weight at most their deficiencies. A high row has positive mass $O(q)$ by the negative diagonal and its negative entries on edges to lows.

Two distinct positive-codegree lows v, w satisfy

$$(Y^2)_{vw} + (AYA)_{vw} = 1 + f_v f_w + c_f(v) + c_f(w), \qquad Y = Z_0 + \operatorname{diag} f. \tag{3.3}$$

Indeed $A^2 = qI + J - Y$. Expanding $(A^2)^2 = AA^2A$ at v, w, using $Y_{vw} = 0$ and row sums $1 + (q+1)f_v + c_f(v)$, gives (3.3). The Y^2 entry counts common zero-codegree partners, and $A \operatorname{diag}(f) A \ge -1$ at this pair. Hence the number of common zero-codegree partners is at most $2 + f_v f_w + c_f^+(v) + c_f^+(w)$.

Lemma 3.2 (Triangle cover)**.** *For a low l, let u_l count its regular neighbors on triangular edges and $d_R(l) = |N(l) \cap R|$. Then*

$$u_l(d_R(l) - u_l - 1) \le \sum_{x \in N(l)} c_f^+(x) \le f_l d(l) + F_+. \tag{3.4}$$

Proof. If x, v are regular neighbors of l with lx triangular and lv untriangulated, then $(BZ)_{xv} \ge 1 - f_l$. Its only possible negative entry comes from the unique common neighbor l. The l-term in $(ZB)_{xv}$ is $-f_l$. Thus commutation forces a neighbor of v in P_x, each such positive partner covering at most one v by the codegree bound with l. Consequently $d_R(l) - u_l \le |P_x| \le 1 + c_f^+(x)$ at each such x. Summing over the triangular regular neighbors proves the first inequality. The second bound holds because every low other than l has at most one neighbor in $N(l)$. □

4. Positive-row block estimates

Fix a low l. Put $k = f_l$, $F_1 = k - 1$, $\alpha = 2t_l/q$, $\delta = |N(l) \cap S|$, and $C = c_f^+(l)$. Let $m = |P_l \cap R|$, $b = |P_l \cap S|$, and $m_v = |N(v) \cap P_l \cap R|$, $b_v = |N(v) \cap P_l \cap S|$. Both block sizes vanish on $N(l) \cup \{l\}$, because every nonhigh point of P_l is nonadjacent to l and has codegree zero with l. At any nonhigh center v outside this set, commutation bounds $m_v + b_v$ by $(BZ)_{lv} + \sum_{w:Z_{lw}<0}(-Z_{lw})B_{wv}$, since the column of B is nonnegative. If $v \in P_l$, the diagonal term of B at l contributes k and there is no common neighbor. Otherwise l, v have one common neighbor w. If w is low, its negative-row correction $F_1 + f_w + c(l,w)$ combines with the actual term Z_{wv} to give $F_1 + 1 + c(l,w) - c(w,v) - f_v$. If w is nonlow, its negative-row correction is at most $F_1 + c(l,w)$. This remains valid at $k = 1$, including a high common neighbor. Thus

$$m_v + b_v \le D_v + R_L(v) + R_N(v), \qquad D_v = F_1 + \chi_v + \eta_v, \tag{4.1}$$

where χ_v indicates a positive partner or a triangular common-neighbor edge at l, η_v indicates a low common neighbor, and $D_v = 0$ on $N(l) \cup \{l\}$. The terms $R_L(v)$ and $R_N(v)$ sum $(Z_{wv})_+$ over the low and nonlow neighbors w of l, respectively. There are at most $q^2 + k$ baseline centers, $\sum \chi_v \le |P_l| + 2t_l(q+2)$, and $\sum \eta_v \le q\delta$. Since $\chi_v \eta_v \le \eta_v$,

$$\sum_{v \notin H_*} \binom{D_v}{2} \le q^2 \left(\binom{F_1}{2} + F_1 \alpha \right) + k \binom{F_1}{2} + F_1 |P_l| + O((k+1)q\delta + kq). \tag{4.2}$$

The first and second low moments, and the nonlow first moment, are

$$\begin{aligned} \sum_v R_L(v) &\le qC + O(q(1+\delta)), \\ \sum_{v \notin H_*} \binom{R_L(v)}{2} &= O(\delta^2 + C^2 + \delta(k\delta + F_+)), \\ \sum_v R_N(v) &\le k(q+1-k) + O(q). \end{aligned} \tag{4.3}$$

For the first bound, sum the upper positive mass $1 + (f_w - 1)(q+1) + c_f^+(w) + 2t_w$ over low neighbors w. Use $2t_w \le d(w) \le q$ for low w and $\sum_{w \in N(l) \cap S} c_f^+(w) \le k\delta + F_+$. For the second, two such neighbors share l, so (3.3) bounds their common positive partners by $2 + f_w f_{w'} + c_f^+(w) + c_f^+(w')$. Sum this bound over unordered pairs. For the last bound, regular neighbors w have positive mass at most $1 + c_f^+(w)$, which includes weight k from l, and the excess over $k+1$ sums to at most F_+. The $O(1)$ high neighbors each have mass $O(q)$.

For a nonhigh center, put $s_v = (D_v - b_v)_+$, so $m_v \le s_v + R_L(v) + R_N(v)$. Its pair count satisfies

$$\binom{m_v}{2} \le \binom{D_v}{2} - (D_v - 1)b_v + \binom{b_v}{2} + (k+1)R_L(v) + \binom{R_L(v)}{2} + E_v. \tag{4.4}$$

Indeed the first three terms bound $\binom{s_v}{2}$, even when $b_v > D_v$. Adding R_L increases the pair bound by at most the next two terms since $s_v \le k+1$. Take $E_v = 0$ when $m_v \le s_v + R_L(v)$, and otherwise take $E_v = (m_v + s_v + R_L(v))R_N(v)/2$, by the difference of the binomial coefficients. In particular $E_v \le tR_N(v)$ when $m_v < t$. We can also omit the low-point subtraction by putting $b_v = 0$ in this argument.

The total b_v over nonhigh centers is at least $qb - O(q)$ by the degrees of points of $P_l \cap S$, and $\sum_v \binom{b_v}{2} \le \binom{b}{2} = O(q^2)$. The missing pairs of the m regular points number at most $(m + \sum_{v \in S} f_v m_v)/2 = O(q^2)$ by (3.1). Blocks of different centers meet in at most one point. For blocks of size at least $t > \sqrt{m}$, their total size I is at most

$$I \leq \frac{m}{1 - m/t^2}. \tag{4.5}$$

Indeed, if there are $r \leq I/t$ such blocks and p_i is the number containing point i, then

$$\frac{I^2}{m} \leq \sum_i p_i^2 = I + 2\sum_i \binom{p_i}{2} \leq I + r(r-1) \leq I + \frac{I^2}{t^2}.$$

Rearranging proves the bound.

4.1. **Growing deficiency.**

Lemma 4.1. *Suppose $l = l_q$ is low, $f_l = k \to \infty$, and $|N(l) \cap S| = o(q)$. Then some nonedge incident to l creates $o(q)$ four-cycles on an infinite subsequence.*

Proof. Pass to a subsequence on which k/q converges. First note that

$$\binom{m}{2} = \tfrac{1}{2}(F_1 + \alpha)^2 q^2 + F_1 q(k + C - b) + O(q^2) \tag{4.6}$$

by (3.2). By (4.3), the low moment terms in (4.4) cost at most $kqC + o(kq^2)$ since $C, F_+ = O(q)$. The correction involving b_v subtracts $F_1 qb$ up to $o(kq^2)$ whenever the included centers have at least $qb - o(q^2)$ total b_v. In fact $D_v \geq F_1$ when $b_v > 0$ and the binomial terms in b_v sum to $O(q^2)$.

If $k = o(q)$, choose t with $\sqrt{kq} = o(t)$ and $t = o(q)$. The large blocks have total size at most $m + o(kq)$ and number $o(\sqrt{kq})$ by (4.5). Their total incidences into $P_l \cap S$ are at most $b + o(kq)$ by pairwise intersections. Thus the small nonhigh centers have the required total b_v, and their E_v terms cost $O(tkq) = o(kq^2)$. The extra terms in (4.2) are $o(kq^2)$ since $k, \delta = o(q)$. The missing pairs and all high centers cost $O(q^2)$. Subtract small pair coverage and these losses from (4.6). The C, b terms cancel up to $o(kq^2)$, and $F_1 qk = o(kq^2)$. The large blocks therefore cover at least $(F_1 + \alpha^2)q^2/2 - o(kq^2)$ pairs, with total size at most $(1 + o(1))kq$. Some block has size $q - o(q)$.

If $k/q \to \kappa > 0$, then $\kappa \leq 1/2$ by (2.3). Suppose all nonedges from l give at least ϵq cycles on a subsequence. Each nonempty block has a center nonadjacent to l, and its points have codegree zero with l. Thus $m_v \leq (1 - \epsilon)q + 2$. Applying (4.4) to all nonhigh centers, the nonlow extra terms cost at most

$$\tfrac{1}{2}\kappa(1 - \kappa)(1 + \kappa - \epsilon)q^3 + o(q^3)$$

by $s_v \leq k + 1$, $R_L(v) \leq \delta = o(q)$, and (4.3). Compare (4.6) with (4.2) and (4.4). The b, C terms cancel at scale q^3. Also $F_1 qk - F_1|P_l| = O(q^2)$ by (3.2). After the baseline, the remaining demand is

$$\tfrac{1}{2}(F_1 + \alpha^2)q^2 - k\binom{F_1}{2} + o(q^3) = \tfrac{1}{2}\kappa(1 - \kappa^2)q^3 + o(q^3),$$

including the $O(q^2)$ cost from high centers and missing pairs. This exceeds the nonlow extras by $\epsilon\kappa(1 - \kappa)q^3/2 + o(q^3)$, impossible. This handles every positive limiting ratio k/q. In the sublinear case, the block of size $q - o(q)$ gives a nonedge from l creating $o(q)$ four-cycles because the center has degree at most $q + 2$. Every new cycle corresponds to a length-three path between the endpoints, and the block points contribute none.

□

4.2. **Bounded local deficiency.**

Lemma 4.2. *Fix K_0. There are constants $b_0 > 0$ and $\theta_* > 0$, depending on K_0, ϵ, such that the following holds for all sufficiently large q. If a low vertex l satisfies*

$$f_l \le K_0, \qquad |N(l) \cap S| \le \theta q, \qquad c_f^+(l) \le K_0 \theta q, \qquad 0 < \theta \le \theta_*,$$

and either $f_l \ge 2$ or $u_l \ge q/2$, then l has at least $b_0 q$ zero-codegree partners in S.

Proof. Choose $b_0 > 0$ small in terms of K_0, ϵ, and then choose θ sufficiently small in terms of these constants. Under the stated hypotheses, (3.4) gives $u_l = O(1)$ or $u_l = d_R(l) - O(1)$, distinguished by $u_l < q/2$ for $\theta < 1/8$. Thus $\alpha^2 - \alpha \ge -\theta - o(1)$ and $F_1 + \alpha \ge 1 - \theta - o(1)$.

If there are at most $b_0 q$ zero-codegree low partners, then $m = (F_1 + \alpha)q + O_{K_0}((\theta + b_0)q + 1)$ by (3.2) and its unit version. Use (4.4) with $b_v = 0$, and threshold $t = q^{3/4}$. By (4.2) and (4.3), the small blocks cover at most $(\binom{F_1}{2} + F_1\alpha)q^2 + O_{K_0}(\theta q^2) + o(q^2)$ pairs. The low moments cost $O_{K_0}(\theta q^2 + q)$, the nonlow extras cost $O(tq) = o(q^2)$, and the small high blocks cost $O(t^2) = o(q^2)$. The weighted missing-pair bound above (4.5) has small-center term at most $tF_+ = o(q^2)$. Subtract small coverage and missing pairs from $\binom{m}{2}$. The remaining inequality is

$$\sum_{m_v \ge t} m_v(m_v + (f_v)_+) \ge (F_1 + \alpha^2)q^2 - O_{K_0}((\theta + b_0)q^2) - o(q^2).$$

The large blocks have total size at most $m + o(q)$ by (4.5). For a nonempty block, (2.2) implies $m_v + (f_v)_+ \le (1-\epsilon)q + 2$. The upper bound on the weighted sum is therefore $(1-\epsilon)(F_1+\alpha)q^2 + O_{K_0}((\theta + b_0)q^2) + o(q^2)$. This is impossible for sufficiently small b_0, θ, since $\epsilon(F_1+\alpha) + \alpha^2 - \alpha \ge \epsilon(1-\theta) - \theta - o(1)$. Only the bounds on k, δ, C, F_+ entered this proof. Deficiencies elsewhere may be unbounded. □

5. Light vertices and propagation

Pass to a subsequence on which K is bounded or $K \to \infty$. In the bounded case fix $K_0 \ge K$. In the growing case the preceding growing-deficiency result and (2.2) give constants $K_0, \theta_0 > 0$ such that every low of deficiency larger than K_0 has at least $\theta_0 q$ low neighbors. Otherwise we could select increasing q and lows with deficiencies tending to infinity and induced low degrees $o(q)$.

Take $\theta > 0$ small enough for the bounded local estimate with K_0, and in the growing case take $\theta < \theta_0$. Let $U = \{v \in S : |N(v) \cap S| \ge \theta q\}$, the low hubs. There are $O_\theta(1)$ hubs because

$$\sum_v \binom{|N(v) \cap S|}{2} \le \binom{|S|}{2}. \tag{5.1}$$

Set $W = U$ in the bounded case, and $W = U \cup (S \cap N(U))$ in the growing case. In the bounded case every vertex sees $O(1)$ members of W. In the growing case every vertex outside U sees $O(1)$ members of W because it has at most one common neighbor with each member of U.

For a low in $S \setminus W$, $k \le K_0$, $\delta < \theta q$, and $C \le K_0 \theta q$. Call such a low light if it is unit and has $u_l < q/2$, and active otherwise. Denote the light set by L. Equation (3.4) gives $u_l = O(1)$ for lights. Each active low has at least $b_0 q$ zero-codegree low partners by the bounded local estimate. Any two active neighbors of a graph vertex have positive codegree and by (3.3) have at most $2 + K_0^2 + 2K_0\theta q$ common zero-codegree low partners. Fix an integer $r > 4/b_0$, and make θ small enough that $\binom{r}{2} 2K_0\theta < 1$. The union of the zero-codegree low neighborhoods of r active neighbors would exceed $3q - o(q) > |S|$. Therefore every graph vertex sees $O(1)$ active lows. Call the union of the active lows and W nonlight. Every graph vertex sees $O(1)$ nonlight lows in the bounded case. In the growing case this holds at every vertex outside U.

Lemma 5.1. *The set L is independent. Every light vertex has $O(1)$ low neighbors, triangles, positive partners, and zero-codegree low partners. The set J_* of regular vertices on triangular edges to lights*

has bounded size. Every regular row has bounded positive support in the bounded-deficiency case. In the growing-deficiency case the same holds for regular vertices with no hub neighbor.

Proof. We first prove independence. Suppose $l, w \in L$ are adjacent and v is an untriangulated regular neighbor of l. Then w, v are nonadjacent with unique common neighbor l. The only possible negative term in $(BZ)_{wv}$ is at l, where $Z_{lv} = 0$. The l-term of $(ZB)_{wv}$ is negative. Thus v has a neighbor in P_w. Each such positive partner covers at most one of these v by its codegree with l. There are $(1-\theta)q - O(1)$ choices of v, while $|P_w| \le (K_0+1)\theta q + O(1)$ from the negative entries to bounded-deficiency low neighbors and the $O(1)$ triangular regular neighbors. Take $\theta < 1/(4(K_0+1))$. This contradiction proves independence. Each light now has $O(1)$ low neighbors and triangles. Its low neighbor deficiencies are bounded, since a light has no hub neighbor in the growing case. Hence its positive row has bounded size and it has $O(1)$ zero-codegree low partners.

Let $J_* \subseteq R$ be the regular vertices on triangular edges to lights. The local cover in (3.4) gives $c_f^+(v) \ge q - O(1)$ for $v \in J_*$, because a light has $q - O(1)$ regular neighbors and $O(1)$ triangular edges. In the bounded case each such v gets only $O(1)$ weight from nonlight neighbors. In the growing case, at most $O(1)$ vertices see two hubs by the codegree bound. Every other $v \in J_*$ gets weight at most $(1/2 + o(1))q$ from its hub neighbor and $O(1)$ weight from nonhub nonlight neighbors. Thus all but $O(1)$ vertices of J_* see $\Omega(q)$ lights in either case. By (5.1), $|J_*| = O(1)$.

At a regular vertex x, the number of triangular edges to lights is $O(1)$. The third vertex in such a triangle is regular and in J_*, high, or a nonlight low neighbor of x, each with bounded count. Every edge lies in at most one triangle. Thus every regular row has $|P_x| = O(1)$ in the bounded case, and so does every regular row with no hub neighbor in the growing case. Its negative mass through lights occurs only on the bounded number of triangular light edges, and its negative mass through the other low neighbors has bounded count and weights.

□

Lemma 5.2 (Propagation)**.** *We have*

$$|P_l \cap (R \cup L)| = O(1) \tag{5.2}$$

for all lows l in the bounded case and all $l \notin U$ in the growing case.

Proof. Put $Y_l = P_l \cap (R \cup L) \setminus J_*$. Each $y \in Y_l$ is nonadjacent to l and has codegree zero, so y is nonadjacent to all $x \in N(l)$. Consider the $q - O(1)$ neighbors $x \in N(l) \cap (R \cup L)$ with no neighbor in J_*, and in the growing case no neighbor in U. These exclusions cost $O(1)$ by the codegree bound because l is distinct from J_* and, when required, from all hubs. All selected rows satisfy $|P_x| \le B_0$ for a constant B_0.

For each $y \in Y_l$, $(BZ)_{xy} \ge 1$ except for $O(1)$ selected x. The l-term equals one and a loop at x is nonnegative since xy is absent. If y is light, a negative term through a regular neighbor of x requires that neighbor in J_*, and the remaining negative terms can only come from one of the $O(1)$ low neighbors of y. If y is regular, a light neighbor cannot give a negative term because $y \notin J_*$, and only one of its $O(1)$ nonlight low neighbors can give such a term. In either case the potential negative neighbors are distinct from l and each sees at most one $x \in N(l)$. High neighbors give no negative entries to regular or light y.

Choose x_0 for which at most half the points of Y_l are exceptional, possible by the per-point exception count. Every other y needs a B-neighbor in P_{x_0} by commutation. If $|Y_l|$ exceeds a fixed constant, some $z \in P_{x_0}$ is graph-adjacent to more than $4B_0$ of these points, allowing for one loop. All but $O(1)$ selected x have at most half of those points exceptional. Any $w \ne z$ can cover at most two points of $Y_l \cap N(z)$ in B, by the codegree bound and one possible loop. Hence $z \in P_x$ for all but $O(1)$ selected x. Since $Z_{x_0 l} \le 0$, $z \ne l$. A positive entry at x, z gives zero codegree, except possibly when x is regular and adjacent to a high z. The latter exception costs at most one $x \in N(l)$. Thus lz is a nonedge and all but $O(1)$ neighbors of l have codegree zero with z. This gives $O(1)$ new cycles, contrary to (2.2). The claim follows. □

6. Bounded maximum deficiency

Lemma 6.1. *The maximum positive deficiency K remains bounded along a sequence satisfying (2.2).*

Proof. Suppose otherwise and pass to a subsequence with $K \to \infty$. Take u with $f_u = k = K \to \infty$, so $u \in U$. Its regular positive row outside J_*, denoted Y_u, has size $kq + O(q)$ by (3.2) and $|S| = O(q)$. Every point of Y_u is nonadjacent to all $x \in N(u)$. Suppose $x \in N(u) \setminus U$ is regular or low with $|P_x| \le k + O(1)$. Then $(BZ)_{xy} \ge 1$ for all but $O(q)$ points $y \in Y_u$. The u-term equals one. A negative term requires a nonlight low neighbor of x adjacent to y, since $y \notin J_*$ and a loop at x gives no negative entry. There are $O(1)$ such neighbors and each accounts for at most q points. Commutation gives a graph neighbor of each remaining y in P_x. Some $z \in P_x$ therefore has at least $(kq - O(q))/(k + O(1)) = q - o(q)$ neighbors in Y_u. This forces uz to be a nonedge and gives $o(q)$ new cycles since every point of Y_u has codegree zero with u. Hence such x cannot occur along (2.2).

Every regular neighbor of u with no second hub neighbor would have $|P_x| \le k + O(1)$. Its negative mass at u is at most k, and its other negative contributions through nonlight lows and triangular light edges sum to $O(1)$. There are only $O(1)$ neighbors of u with a second hub neighbor and only $O(1)$ highs. Thus $|N(u) \cap S| = q + 1 - k - O(1) = \Omega(q)$ by (2.3).

Choose a low $x \in N(u) \setminus U$ with no second hub neighbor and with $O(1)$ low neighbors. Such a choice exists because the hubs and second-hub neighbors cost $O(1)$, and the sum of induced low degrees over $N(u) \cap S$ is at most $|S| - 1 + |N(u) \cap S|$. Each low other than u contributes at most one incidence to that sum by the codegree bound. Now $f_x \le K_0$, $c_f^+(x) = k + O(1)$, and $d_R(x) = q - O(1)$. If x is unit and has $O(1)$ triangular regular edges, then $|P_x| \le k + O(1)$, already excluded. If it is unit and has a larger number, (3.4) forces $q - O(1)$ triangular regular edges. These give negative mass $q - O(1)$ in addition to at least k from its edge to u. If it is nonunit, (3.2) gives the same lower bound $|P_x| \ge k + q - O(1)$, using $c_f(x) \ge k - |H_*|$.

Every other low in $N(u)$ shares u with x, so cannot be a positive partner of x. The diagonal is nonpositive. There are $O(1)$ regular positive partners of x by (5.2) and $O(1)$ high partners. Hence

$$k + q - O(1) \le |P_x| \le |S| - |N(u) \cap S| + O(1).$$

This gives $|S| \ge 2q - O(1)$, which contradicts $F_+ \ge |S| + k - 1$ and $F_+ \le 2q + O(1)$. □

7. A large open packing

Now $K = O(1)$. Let M be the zero-codegree graph on S. We sharpen (3.3) to show that any nonadjacent pair in M has $O(1)$ common neighbors in M.

At any low v, only $O(1)$ edges to light neighbors are triangular. As for regular vertices, the third vertex of such a triangle is in J_*, high, or a nonlight low neighbor of v. For distinct positive-codegree lows v, t, all but $O(1)$ lights $x \in N(v)$ have positive codegree with t. The possible exceptions are $x = t$, a common graph neighbor of v, t, or a positive light partner of t. The latter number is $O(1)$ by (5.2). For each remaining x, let $y \in N(t)$ be its unique common neighbor with t. Then $y \notin L$, and xy is untriangulated for all but $O(1)$ choices of x. If it is triangular, y is in J_*, high, or one of the $O(1)$ nonlight low neighbors of t. Each such $y \ne v$ is adjacent to at most one $x \in N(v)$, and the case $y = v$ costs only the bounded number of triangular light edges at v.

The ordered pairs (x, y) just obtained contribute $|N(v) \cap L| - O(1)$ to $(AZ_0A)_{vt}$. Interchanging v, t supplies disjoint ordered contributions with second endpoint light and first endpoint nonlight. Since $c_f(v) = |N(v) \cap L| + O(1)$, this cancels the two c_f terms of (3.3) up to $O(1)$. The common-neighbor term $(Y^2)_{vt}$ is therefore $O(1)$.

Lemma 7.1. *There is a nonlight low for all sufficiently large q.*

Proof. If all lows were light, they would be unit and independent. A high v would see at least $q - 1$ low neighbors by (3.1), with at most three nonlow neighbors. The graph on $N(v)$ is a matching, so

at least $q-4$ of its low neighbors would have zero codegree with v, contradicting $z(v) \le 3$. With no highs, Z has zero diagonal and only $O(1)$ negative entries. Indeed each triangle through a light uses an edge between two regular vertices of J_*. On the rational subspace $\mathbf{1}^\perp$, the trace of B is the odd integer $|S| - (q+1)$. The trace on each rational $Z = 1$ and $Z = -1$ eigenspace within this subspace is zero, since $B^2 = qI - Z$ there and $q \mp 1$ is nonsquare for large dyadic q. Thus the orthogonal projection onto the remaining Z-eigenspaces is nonzero and commutes with B. Its support has bounded size, since a nonnegative row of Z has exactly one positive entry of weight one and two such rows joined together form an isolated edge carrying only eigenvalues $1, -1$. Only the bounded negative endpoints and their bounded positive support can belong to other components. If an outside vertex has a graph neighbor in the projection support, it has at least two, since its adjacency column projects to zero by commutation. Counting such cross edges by twice the number of pairs of support vertices gives $O(1)$ edges by four-cycle-freeness. The support's internal edges also number $O(1)$, contradicting minimum degree q. This proves that a nonlight exists. The trace assertion used here is Lemma 9.1 below. Since q is a square and $q \ge 4$, both $q-1$ and $q+1$ lie strictly between successive squares. □

Lemma 7.2. *There is a maximal set $T \subseteq S$ of pairwise zero-codegree vertices, with $p = |T| = q - O(1) \le q$, consisting entirely of nonlight vertices. Every outside low has $O(1)$ zero-codegree partners in T.*

Proof. Every nonlight has linearly many positive partners. For a hub, this follows from its negative mass through low neighbors. A nonunit low uses (3.2), and a heavy unit low has at least $q/2$ triangular regular incident edges. By (5.2), all but $O(1)$ of these partners are nonlight lows, and each pair has codegree zero. By Lemma 7.1, one such vertex exists. Its partners give $\Omega(q)$ nonlight vertices, each with $\Omega(q)$ partners. Hence $e(M) = \Omega(q^2)$. The bounded common-neighbor estimate gives only $O(q^2)$ induced two-paths in M. Since $|S| = O(q)$ and $e(M) = \Omega(q^2)$, some vertex has a linear-size M-neighborhood with $O(q)$ missing edges. To see the last selection explicitly, the number of induced two-paths is $O(q^2)$, while the total degree in M is $\Omega(q^2)$. Vertices of sufficiently small degree account for only a fixed small fraction of this total. Among the remaining vertices, one has $\Omega(q)$ neighbors and $O(q)$ missing pairs within its neighborhood. The complement on that neighborhood has bounded average degree and therefore an independent set of size $\Omega(q)$. For completeness, the bound $\alpha(H) \ge |V(H)|/(\overline{d}(H)+1)$ follows by placing vertices in a random order and selecting those that precede all their neighbors, followed by Cauchy–Schwarz. Hence M has a clique of size $\Omega(q)$. Extend it to a maximal clique T, of size p. It is an open packing, meaning its graph neighborhoods are pairwise disjoint. Each light has M-degree $O(1)$, so T consists of nonlight lows. For any outside low, maximality gives a packing vertex with which it has positive codegree. Its zero-codegree partners in T are common M-neighbors of that pair, so number $O(1)$.

For any vertex v, replacing its neighborhood by a set $U \subseteq T \setminus \{v\}$ preserves four-cycle-freeness. A new four-cycle would require two vertices of U with a common neighbor distinct from v, contradicting the packing property. Edge extremality gives $p \le q$. Replacing the neighborhood of a packing vertex by the remaining packing points preserves four-cycle-freeness and gains an edge if $p \ge q+2$. If $p = q+1$, the inside exchange forces every packing low to have degree q, and the outside exchange, replacing an outside low's neighborhood by T, excludes any outside low. A high would then have at most one unit low neighbor, contradicting (3.1). With no highs the total deficiency $p = q+1$ would be odd, whereas $F = 2q + 2 - 2a$ is even.

Put $d = q + 2 - p$. Let F_T be the binary matrix recording zero-codegree partners of T outside T, A_T the rows of A indexed by T, and D_T the diagonal of their degrees. Write $g = (f_t - 1 : t \in T)$, $b_T = (c_f(t) : t \in T)$, and use J for the all-ones matrix of the relevant size. The packing Gram identity is

$$F_T F_T^\top + A_T Z_0 A_T^\top + \operatorname{diag}(b_T) + D_T^2 - qD_T = dJ + gg^\top + \mathbf{1}b_T^\top + b_T\mathbf{1}^\top. \tag{7.1}$$

To obtain it, the T-rows of A^2 consist of D_T inside T and the all-ones matrix minus F_T outside T. Their outside zero counts are $d\mathbf{1} + qg + b_T$ by (3.1). Expand the T-block of $(A^2)^2$ using these rows and equate it to the T-block of $qA^2 + (A\mathbf{1})(A\mathbf{1})^\top - AYA$, from $A^2 = qI + J - Y$. The diagonal contribution from $A\,\mathrm{diag}(f)A$ is b_T because the neighborhoods of T are disjoint. The degrees on T are $q\mathbf{1} - g$, giving (7.1).

Summing (7.1) gives $dp^2 = O(q^2)$. In detail the low outside columns have bounded size by maximality. A regular outside column has size at most $1 + c_f^+(v)$ by (3.1), and

$$\sum_v c_f^+(v)^2 \le q \sum_{w\in S} f_w^2 + F_+^2 = O(q^2)$$

by the pair codegree bound. The constant number of high columns also contribute $O(q^2)$ squared size. The sum of $A_T Z_0 A_T^\top$ is at most $\sum_v z(v) = O(q^2)$. The diagonal terms on the left sum to at most F_+, since $D_T \le qI$ and packing neighborhoods are disjoint. The last two terms on the right sum to at least $-2p|H_*|$, and the $gg^\top$ term is nonnegative. Since $p = \Omega(q)$, we have $2 \le d = O(1)$. □

8. The full packing and its completion

Proposition 8.1. *The extremizer has a set T of q vertices such that*

1. *the neighborhoods of the vertices in T are pairwise disjoint, and T is independent,*
2. *every low vertex has degree q, and there are no high vertices,*
3. *$X = S \setminus T$ is independent and no vertex of X belongs to a triangle.*

Adjoining one looped vertex adjacent precisely to T, and adding a loop at each vertex of X, gives a symmetric binary matrix C of order $q^2 + q + 3$, with row sum $q + 1$, off-diagonal entries of C^2 at most one, and $\operatorname{tr} C = q + 3 - 2a$.

Proof. The hole count. Put $X = S\setminus T$, $F_X = \sum_{x\in X} f_x$, and $\eta = |X \cap N(T)|$, where $N(T) = \bigcup_{t\in T} N(t)$. Let h count the nonlow vertices outside $N(T)$, called holes. The deficiency sum gives at least $F_X + 2a + 2 - 2d - |H_*|$ unit rows of T. Each such row t has $d + c_f(t) = |N(t) \cap L| + O(1)$ outside zero-codegree partners. All but $O(1)$ of its light neighbors are untriangulated and hence among these partners. Each hole has a graph neighbor in the outside zero partners of every unit row. Indeed it has degree at least $q + 1$ and no neighbor in T, and at most q of its neighbors can have positive codegree with that unit row, since each of the row's q neighbors can contribute at most one.

Each light ℓ meets $p - j_\ell$ row neighborhoods in one point each, where $j_\ell = O(1)$ is its zero-partner count in T. Thus ℓ has $q - p + j_\ell = O(1)$ graph neighbors outside $N(T)$. It can serve as an adjacent light partner of at most one unit row. Therefore the adjacent lights cover $O(\eta)$ total holes over unit rows.

If $h \to \infty$ and $\eta = O(h)$, at most $O(1)$ unit rows have at least half the holes covered by their adjacent lights. Each other unit row has only $O(1)$ remaining partners and one of these must meet a fixed positive fraction of the holes. Only $O(1)$ vertices can meet that many holes, by the pair count on the hole set. Each such low occurs in only $O(1)$ rows. Each nonlow occurs in at most $1 + c_f^+(v)$ rows by (3.1). Each of the $O(1)$ nonlow connectors meets T at most once, and distinct connectors share at most one low neighbor, of bounded weight. Their total weighted low incidence is therefore at most $F_X + O(1)$. So there are at most $F_X + O(1)$ unit rows, contradicting their lower bound and (2.2).

The size of the packing. For $d \ge 3$, write $\tau = \sum_{t\in T}(f_t - 1)$ and $\sigma = \sum_{t\in T} |N(t) \cap S|$. Counting nonlow covered points gives

$$h = (d-1)q + 2 - |S| + \tau + \sigma \ge (d-3)q + 2a + \eta - |H_*|.$$

This gives $h \to \infty$ and $\eta = O(h)$, so $d = 2$ and $p = q$.

Edge extremality now forces $f_x = 1$ on X and $f_t \leq 2$ on T by the outside and inside neighborhood replacements. Put $k_T = 2e(G[T])$. The packing induces a matching, and $\sigma = k_T + \eta$. Since $F_+ = q + \tau + |X|$, we have

$$h = 2a + 2\tau + k_T + \eta - q - |H_*|. \tag{8.1}$$

For any paired row t, a nonlow vertex in $N(t)$ has at most one neighbor in each row, and in the mate's row that neighbor is t itself. Its only neighbor in T is t. Thus at most q of its neighbors lie in $T \cup N(T)$, and it needs a neighbor outside. Each exterior point meets $N(t)$ at most once. The number of exterior points is $2 + \tau + k_T$, since $|N(T)| = q^2 - \tau$ and $|T \cap N(T)| = k_T$. If $x_t = |N(t) \cap X|$, a paired row has at least $q - 2 - x_t$ nonlow vertices. Therefore summing over the k_T paired rows gives

$$k_T(q - 4 - \tau - k_T) \leq |X|. \tag{8.2}$$

If $k_T = \Omega(q)$ along a subsequence, (8.2) gives $q - \tau - k_T \leq O(1)$ since $|X| = O(q)$. Then (8.1) gives $h \geq 2a + \tau + \eta - O(1)$, which the hole argument excludes. Hence take $k_T = o(q)$.

The secondary minimization. For $x \in X$, replace $N(x)$ by T. The change in degree-square sum is $2(c_f(x) - \tau - |N(x) \cap T|)$ and the change in triangles is $k_T/2 - t_x$. To see the first identity, the number of net gain vertices equals the number of net loss vertices, both $q - |N(x) \cap T|$. For a gain at v, the square changes by $2(q + 1 - f_v) + 1$, and for a loss by $-2(q + 1 - f_v) + 1$. The sum of deficiencies at the loss vertices minus the gain vertices is $c_f(x) - q - \tau$. For the triangle count, all new triangles through x correspond to edges of $G[T]$. The exchange preserves the edge count, so minimality of the objective gives

$$\tau + t_x \leq c_f(x) - |N(x) \cap T| + k_T/2. \tag{8.3}$$

For a deep packing point t with $f_t = 2$, replace $N(t)$ by $T \setminus \{t\}$. The same square calculation now gives change $2(1 + c_f(t) - \tau - d_T(t))$, with $d_T(t) = |N(t) \cap T|$. The triangle change is $k_T/2 - d_T(t) - t_t$. Hence

$$\tau + t_t \leq 1 + c_f(t) + k_T/2 - 2d_T(t). \tag{8.4}$$

If $\tau > 0$, select a deep packing point with $c_f^+(t) \leq F_+/\tau$, since the neighborhoods of deep packing points are disjoint. Equation (8.4) implies $\tau = o(q)$ as $k_T = o(q)$. Equation (8.2) now excludes $k_T \geq 2$, since $|X| \leq q + O(1)$ and $q - 4 - \tau - k_T = q - o(q)$. Thus T is independent.

Every outside nonhub low x is unit and satisfies $t_x \leq c_f(x) \leq K\theta q$ by (8.3). An active unit low has at least $q/2$ triangular regular incident edges and hence at least $q/4$ triangles. The choice $\theta < 1/(4(K_0 + 1))$ therefore makes each outside nonhub light. Any outside hub would have linear positive mass, but could have only $O(1)$ positive partners among regulars or lights by (5.2), highs and outside nonlight hubs by bounded counts, and T by the outside zero-partner bound. Thus $X \subseteq L$, so X is independent.

Any deep $t \in T$ now has $|P_t| \leq q + O(1)$ by (5.2), because the only nonlight lows lie in T. Its exact row size $q + 2 + c_f(t) + 2t_t$ gives $c_f^+(t) + 2t_t = O(1)$. Thus t has $q - O(1)$ untriangulated regular neighbors v. If t were adjacent to $x \in X$, then for each such v the only possible negative term of $(BZ)_{xv}$ is the t-term -1. The t-term in $(ZB)_{xv}$ is at most -2. This forces a positive partner of x adjacent to v, and each such partner covers at most one v. That contradicts $|P_x| = O(1)$ for lights. Therefore each $x \in X$ has low neighbors only among unit packing points. Equation (8.3), with $k_T = 0$, implies $\tau = t_x = 0$ and no high neighbor at x if X is nonempty.

In either case all low neighbors of a high would lie in T, at most one of weight at most two, contradicting (3.1). If X is empty and $\tau > 0$, the inside exchange (8.4) now gives $\tau = 1$, which is impossible by parity of $q + \tau = F$. Thus every packing low is unit and there are no highs.

The matrix completion. Add a new looped vertex adjacent precisely to T, and put a loop on each vertex of X. The resulting symmetric binary matrix C has order q^2+q+3 and row sum $q+1$. All its off-diagonal Gram entries are at most one. The new vertex covers the pairs in T, which originally had zero codegree. Each old loop adds one to a pair containing its looped endpoint only when the pair is an edge, and those edges have zero codegree because the looped endpoint is in no triangles. Independence of X ensures that no old pair receives two loop contributions. For the new vertex and any old v, the Gram entry is $1_{v\in T}+|N(v)\cap T|\le 1$ by the packing and independence of T. The number of loops is $1+|X|=q+3-2a$. □

9. A trace bound for the completion

Lemma 9.1 (Rational square-root trace)**.** *Let V be a finite-dimensional vector space over $\mathbb{Q}$. If an endomorphism T satisfies $T^2=bI$ for a positive rational number b that is not a rational square, then* $\operatorname{tr} T=0$.

Proof. After extending scalars to $\mathbb{R}$, the minimal polynomial divides $(x-\sqrt{b})(x+\sqrt{b})$ and has distinct roots. Thus T is diagonalizable with eigenvalues $\sqrt{b}$ and $-\sqrt{b}$. If their multiplicities are r,s, then $\operatorname{tr} T=(r-s)\sqrt{b}$. The trace is rational, whereas $\sqrt{b}$ is irrational, forcing $r=s$. □

Lemma 9.2 (Cycle transitions)**.** *Let D be a symmetric binary matrix indexed cyclically by $\mathbb{Z}/s\mathbb{Z}$, where $s\ge 3$, and suppose D commutes with the adjacency matrix of the s-cycle. Its row sums are constant, say r. Put $u_i=D_{i,i+1}$. The number of indices with $u_{i-1}\ne u_i$ is at most $4r$.*

Proof. If $R=D\mathbf{1}$, commutation gives $AR=2R$, where A is the cycle adjacency matrix. The equation $R_{i-1}+R_{i+1}=2R_i$ makes all cyclic successive differences equal, and hence zero. Thus every row sum is r. The assertion is immediate if the bits u_i are constant. Otherwise, we claim that

$$D_{i,i+2j+1}=u_{i+j}\qquad (j\ge 0). \tag{9.1}$$

The case $j=0$ is the definition. For the inductive step, compare the $(i,i+2j)$ entries of DA and AD. If $v_i=D_{i,i+2j+1}$, the previous case gives

$$v_i-v_{i-1}=u_{i+j}-u_{i+j-1}.$$

The arrays (v_i) and (u_{i+j}) differ by a constant. Both are binary, and the latter is nonconstant, so this constant is zero. This proves (9.1).

Fix a row i and let j run through $\mathbb{Z}/s\mathbb{Z}$. The map $j\mapsto i+2j+1$ has fibers of size at most two, so

$$\sum_j u_j=\sum_j D_{i,i+2j+1}\le 2r.$$

Each transition is incident to a bit equal to one, and each such bit is incident to at most two transitions. The transition count is therefore at most $2\sum_j u_j\le 4r$. □

Proposition 9.3. *Let $q>2$ be even, with $q-2$ nonsquare. If C is a symmetric binary matrix of order q^2+q+3, every row sum is $q+1$, and $(C^2)_{ij}\le 1$ for $i\ne j$, then* $\operatorname{tr} C\ge q-3$.

Proof. Write $N=q^2+q+3$, $\ell=\operatorname{tr} C$, and

$$Z'=qI+J-C^2.$$

The matrix Z' is symmetric and binary, has zero diagonal, and has row sum two. It is the adjacency matrix of a disjoint union of cycles of lengths at least three. It commutes with C.

For each vertex i, let $m_i=\sum_j C_{ij}C_{jj}$, the number of incident looped vertices, counting i itself when it is looped. In the diagonal entry of C^3, terms with distinct intermediate indices cancel modulo two in symmetric pairs. The remaining terms give $(C^3)_{ii}\equiv m_i \pmod 2$. Since

$$C^3=qC+(q+1)J-CZ',$$

we obtain $(CZ')_{ii}\equiv 1-m_i \pmod 2$.

On a residual cycle, write $u_i = C_{i,i+1}$ in cyclic order. Then $(CZ')_{ii} = u_{i-1} + u_i$. If $m_i = 0$, the two bits differ. In all cases,

$$1 \leq m_i + \mathbf{1}_{\{u_{i-1} \neq u_i\}}. \tag{9.2}$$

Each diagonal cycle block of C commutes with its cycle adjacency matrix. Lemma 9.2 bounds its number of transitions by four times its internal row sum r.

It remains to sum these internal row sums. Let E be the rational orthogonal projection onto the subspace of vectors constant on each residual cycle. On a cycle of length s, its diagonal block is J_s/s. Because this is the 2-eigenspace of Z', the matrices C and E commute. Put $P_0 = J/N$. Then $E - P_0$ is a rational orthogonal projection, and

$$C^2(E - P_0) = (q-2)(E - P_0).$$

Apply Lemma 9.1 on its image. The constant vector contributes $q+1$, so

$$\sum_{\text{cycles}} r = \operatorname{tr}(CE) = q + 1.$$

Also $\sum_i m_i = (q+1)\ell$, by the column sums of C. Summing (9.2) therefore gives

$$q^2 + q + 3 \leq (q+1)\ell + 4(q+1).$$

Hence $\ell \geq q - 4 + 3/(q+1)$, and integrality yields $\ell \geq q - 3$. □

10. Conclusion of the proof

Proof of Proposition 2.2. Lemma 6.1 bounds the maximum deficiency. Proposition 8.1 supplies a matrix C with $\operatorname{tr} C = q + 3 - 2a$. Since q is a power of four, $q - 2 \equiv 2 \pmod 4$ is nonsquare. Proposition 9.3 gives

$$q + 3 - 2a \geq q - 3,$$

so $a \leq 3$. This contradicts $a \geq \epsilon\sqrt{q}$ as $q \to \infty$. □

Proof of Theorem 1.1. For each power of four q, choose the extremizer used above and its parameters a_q, γ_q. Proposition 2.2 implies

$$\min\left\{\frac{a_q}{\sqrt{q}}, \frac{\gamma_q}{q}\right\} \longrightarrow 0.$$

Indeed, failure would give a fixed positive ϵ and an infinite sequence satisfying (2.2). On an infinite subsequence one of these two ratios tends to zero. If $a_q/\sqrt{q} \to 0$, then $1 \leq a_q \leq q$ for all sufficiently large members of the subsequence. Apply Lemma 2.1 with $r = a_q$. The resulting graph has $E_q + a_q + 1 = \operatorname{ex}(q^2 + q + 2, C_4) + 1$ edges and $\binom{a_q}{2} = o(q)$ four-cycles. If $\gamma_q/q \to 0$, add a nonedge attaining γ_q to the chosen extremizer. This again gives the required edge count and $o(q)$ four-cycles. Since $\sqrt{q^2 + q + 2}/q \to 1$, either alternative proves the theorem. □

References

[1] W. G. Brown, *On graphs that do not contain a Thomsen graph*, Canad. Math. Bull. **9** (1966), 281–285. doi:10.4153/CMB-1966-036-2.

[2] P. Erdős, A. Rényi, and V. T. Sós, *On a problem of graph theory*, Studia Sci. Math. Hungar. **1** (1966), 215–235.

[3] P. Erdős and M. Simonovits, *Cube-supersaturated graphs and related problems*, in *Progress in Graph Theory*, Academic Press, 1984, 203–218.

[4] Z. Füredi, *On the number of edges of quadrilateral-free graphs*, J. Combin. Theory Ser. B **68** (1996), 1–6. doi:10.1006/jctb.1996.0052.

[5] J. He, J. Ma, and T. Yang, *Stability and supersaturation of* 4*-cycles*, preprint, first version, 2019. arXiv:1912.00986v1.

[6] J. He, J. Ma, and T. Yang, *Some exact results on* 4*-cycles: stability and supersaturation*, CSIAM Trans. Appl. Math. **4** (2023), no. 1, 74–128. doi:10.4208/csiam-am.SO-2021-0006.

[7] B. Ning and M. Zhai, *Counting substructures and eigenvalues II: quadrilaterals*, preprint, first version, 2021. arXiv:2112.15279v1.

Email address, Will Howe: howew@purdue.edu

Email address, Grant Huffman: huffma97@purdue.edu

Email address, Jensen Kohlmeyer: jenwin@purdue.edu

Email address, Liam Kruer: kruerl@purdue.edu