hexagon:2611.00176v1

Four-cycle supersaturation and polarity stability at $q^2 + q$ vertices

Zhenyu Sun

Abstract

Let $m_0 = \frac{1}{2}q(q+1)^2 - q$. For every even $q \geq 1600$, we prove that a graph on q^2+q vertices with at least m_0+1 edges contains at least $\lceil (q-1)/12 \rceil$ four-cycles. For powers of two in this range, m_0 is the extremal number and a polarity construction has exactly $q-1$ four-cycles at m_0+1 edges. These bounds determine the order of magnitude of the minimum count on this sequence. We also prove a stability theorem: for every nonnegative integer b and every even $q \geq 10^{18}(b+1)^{16}$, each C_4-free graph on q^2+q vertices with $m_0 - b$ edges is a spanning subgraph of an orthogonal polarity graph with one point deleted. The required degree bounds follow from the edge count. As a consequence, a graph on q^2+q vertices with m_0+1 edges has at least $q-1$ four-cycles if an edge set of size $t \geq 1$ meets every four-cycle and q is even with $q \geq 10^{18}t^{16}$. The counting proof uses a codegree commutator; the stability proof combines degree-deficit concentration with incidence completion and polarity extension.

Keywords. Four-cycle; supersaturation; extremal graph; polarity graph; projective plane; stability.

Contents

1 Introduction 2

2 Preliminaries 5

3 Supersaturation one edge above the bound 9

4 Degree bounds for near-extremal cores 21

5 Polarity stability at a prescribed edge deficit 26

6 Degree deficits at the extremal bound 33

7 Graphs attaining the extremal bound 42

8 Zero full-degree surplus 48

9 Sharp supersaturation from polarity structure 54

10 Further questions 56

A Incidence completion 56

B Polarity extension 60

References 64

1 Introduction

For a graph F, let $\mathrm{ex}(n, F)$ be the largest number of edges in an n-vertex graph containing no copy of F. Write $N_{C_4}(G)$ for the number of unlabelled, not necessarily induced, four-cycles in G, and define

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

The first edge above an extremal bound is a basic case of supersaturation [4]. A conjecture of Erdős and Simonovits predicts $h(n) \geq (1 - o(1))\sqrt{n}$ [5, Conjecture 1]. We determine the order of magnitude of $h(n)$ on the sequence $n = q^2 + q$, where q is a power of two, and establish a quantitative structure theorem for C_4-free graphs just below the corresponding edge bound.

Throughout, put

$$N = q^2 + q + 1, \qquad M = \frac{1}{2}q(q+1)^2, \qquad n = N - 1, \qquad m_0 = M - q. \tag{1.1}$$

The equality $\mathrm{ex}(n, C_4) = m_0$ is known for powers of two [6]. Our statements for general even q use the numerical edge count m_0; the conclusions about $h(n)$ use this known equality.

1.1 Main results

Theorem 1.1 (Supersaturation at the adjacent order)**.** *Let $q \geq 1600$ be even. Every graph G with*

$$|V(G)| = q^2 + q, \qquad e(G) \geq m_0 + 1$$

satisfies

$$12N_{C_4}(G) \geq q - 1.$$

Deleting an absolute point from a classical orthogonal polarity graph and adding an edge between two remaining absolute points gives exactly $q - 1$ four-cycles. This construction and the known extremal number yield the following consequence.

Corollary 1.2. *For every power of two $q \geq 2048$,*

$$\left\lceil \frac{q-1}{12} \right\rceil \leq h(q^2 + q) \leq q - 1.$$

In particular, $h(q^2 + q) = \Theta(q)$ on this sequence.

Our second main result describes the graphs below the same edge bound. A polarity of a projective plane is an incidence-reversing involution between points and lines. Its simple polarity graph joins distinct points x, y when x lies on the polar line of y. A point is absolute if it lies on its polar line, and a polarity of a plane of order q is orthogonal if it has exactly $q + 1$ absolute points.

Theorem 1.3 (Polarity stability)**.** *Let $b \geq 0$ be an integer and let q be even with*

$$q \geq 10^{18}(b+1)^{16}.$$

Every C_4-free graph F with $q^2 + q$ vertices and $m_0 - b$ edges is a spanning subgraph of $P - p$, where P is the simple graph of an orthogonal polarity of a projective plane of order q and p is one of its points. The absolute points of this polarity form a line. More precisely, F is obtained from $P - p$ by deleting b edges when p is absolute, and by deleting $b - 1$ edges when p is nonabsolute.

The containing plane need not be Desarguesian. The estimate is uniform in q and b under the displayed condition. It includes every fixed deficit b for all sufficiently large even q, and allows b to grow subject to $b + 1 \leq (q/10^{18})^{1/16}$.

1.2 Relation to earlier work

Füredi proved $\mathrm{ex}(N, C_4) = M$ for powers of two [7].

Firke, Kosek, Nash, and Williford [6, Theorem 1] proved $\mathrm{ex}(N-1, C_4) \le m_0$ for every even q, with equality when q is a power of two. For the equality case and sufficiently large even q, the absolute-point deletion classification was announced in their Theorem 2. Theorem 7.6 gives a complete proof at the edge count m_0, with the explicit sufficient condition $q \ge 10^{15}$.

At order N, He, Ma, and Yang [8, Theorem 1.3] proved that, for even $q \ge 10^{12}$, every graph with $M+1$ edges has at least $q-1$ four-cycles. They also characterized the graphs with fewer than $2q-3$ four-cycles. In particular, $h(N) = q-1$ for $q = 2^k$ with $k \ge 40$. Their subsequent stability theorem [9, Theorem 3.1] states that, for each fixed $0 < c < 1$ and every sufficiently large even q, every C_4-free graph on N vertices with at least $M - cq/2$ edges is contained in a unique polarity graph of order q. Table 1 compares these results with the two principal theorems of the present paper.

Result	Order	Hypothesis	Conclusion
Firke et al. [6, Theorem 1]	$N-1$	C_4-free	$e(G) \le m_0$
He–Ma–Yang [8, Theorem 1.3]	N	$e(G) = M+1$	$N_{C_4}(G) \ge q-1$
He–Ma–Yang [9, Theorem 3.1]	N	C_4-free; $e(G) \ge M - cq/2$	Unique containing polarity graph
Theorem 1.1	$N-1$	$e(G) \ge m_0 + 1$	$12N_{C_4}(G) \ge q-1$
Theorem 1.3	$N-1$	C_4-free; $e(F) = m_0 - b$	$F \subseteq P - p$, with P orthogonal

Table 1: Results at the two consecutive orders. The ranges of q, c, b are those specified in the text and in Theorems 1.1 and 1.3.

The change of order requires a completion argument. Adjoining an isolated vertex to a graph covered by Theorem 1.3 gives

$$|V(F \cup K_1)| = N, \qquad e(F \cup K_1) = M - q - b < M - \frac{c}{2}q \quad (0 < c < 1).$$

Its edge count is therefore below the hypothesis of the existing N-vertex stability theorem. Adding edges at the new vertex instead requires control of the four-cycles they create.

For nearby orders with a larger vertex deficit, Ma and Yang [11, Theorem 1.3] proved $\mathrm{ex}(N-r, C_4) \le M - 0.92rq$ whenever r is sufficiently large and $r \le 0.01q$, with no parity restriction on q. Degree deficits, including negative values at vertices of degree greater than $q+1$, already occur in that work [11, Definition 3.1]. The signed moment and concentration estimates below develop this counting method for the edge deficits $m_0 - e(F)$ at the fixed order $N-1$.

1.3 Sharp counting and structural refinements

For a graph G, let $\tau_4(G)$ be the minimum number of edges whose deletion makes G C_4-free. The stability theorem gives a sharp counting consequence.

Corollary 1.4 (Sharp supersaturation with a small edge transversal)**.** *Let $t \ge 1$ be an integer, let q be even with*

$$q \ge 10^{18} t^{16}, \tag{1.2}$$

and let G have q^2+q vertices and m_0+1 edges. If $\tau_4(G) \le t$, then

$$N_{C_4}(G) \ge q-1.$$

Deleting a transversal of t edges leaves a C_4-free graph at deficit $t-1$. Theorem 1.3 embeds this graph in a one-point deletion of a polarity graph. The single-edge estimates of Proposition 2.3 then give the conclusion; the complete argument appears in Section 9.

The degree control used in the stability proof is also quantitative. If F is C_4-free with n vertices and $m_0 - b$ edges, then, for even q,

$$q \geq 10^9(b+1)^8 \quad \Longrightarrow \quad \Delta(F) \leq q+1$$

(Theorem 4.7). A preliminary result for $q \geq 2048$ and $0 \leq b \leq q/16$ gives $\Delta(F) \leq q+2$, at most one vertex of degree $q+2$, and minimum degree at least $q-b-2$.

At $b=0$, a separate two-vertex concentration argument improves the sufficient threshold to $q \geq 10^{15}$ and yields a unique completion by the missing absolute point (Theorem 7.6). This classification also gives a sharp bound throughout a linear range of edge modifications. If G has n vertices and m_0+1 edges and F is a C_4-free graph on the same vertices with m_0 edges, then, for even $q \geq 10^{15}$,

$$b = |E(F) \setminus E(G)| \leq q-2 \quad \Longrightarrow \quad N_{C_4}(G) \geq (b+1)(q-1-b) \geq q-1. \tag{1.3}$$

The proof counts cycles containing exactly one edge outside F. For a fixed new edge, their supporting three-edge paths in F are pairwise edge-disjoint, so each deleted edge removes at most one of these cycles.

A further refinement concerns the full-degree surplus. With $\delta_v = q+1-d_F(v)$, define

$$\eta(F) = \sum_{v : d_F(v) = q+1} \left(\sum_{w \in N_F(v)} \delta_w - 2 \right). \tag{1.4}$$

For even q, its summands are nonnegative in a C_4-free graph on n vertices. Under $\Delta(F) \leq q+1$, zero surplus at deficit b forces an absolute-point deletion at $b=0$ or a nucleus deletion at $b=1$, once $q \geq 8192(b+1)^4$. The nucleus is the nonabsolute point whose polar line is the absolute-point line. At $b=1$, the sufficient threshold is $q \geq 12$ under the same degree bound. Section 8 proves these refinements, including uniqueness of the one-point completions.

1.4 Proofs and organization

The proof of Theorem 1.1 is combinatorial. The codegree matrix differs from $qI+J$ by the degree deficits, the uncovered-pair matrix, and a nonnegative excess matrix controlled by the number of four-cycles. Commuting this identity with the adjacency matrix constrains the unique uncovered partners of full-degree vertices. Neighbourhood counts then treat all four possible maximum-degree configurations. In the case of a unique vertex of degree $q+2$, an injective pairing of vertices with no uncovered partner sharpens the exceptional-set estimate and permits cancellation of the excess codegree at the high vertex.

For Theorem 1.3, the edge count first forces the required degree bounds. The total excess deficit is then either $O(b+1)$ or at least $q-O(b+1)$. The latter alternative supplies sufficiently many full-degree vertices directly. In the former, first and second moments concentrate local weight $q-O(b+1)$ in $2b+3$ vertices. Deleting these vertices leaves enough complete neighbourhoods to recover a projective plane. The incidence-completion theorem eliminates the uncovered point pairs and embeds the resulting incidence system in a plane. A symmetric-incidence argument extends the resulting assignment to a polarity. Finally, a count of missing edges restores every deleted vertex and every edge between them. The uniform estimates needed in this construction are stated explicitly in Section 5 and Appendix A.

Section 2 gives the counting identities, polarity construction, and elementary geometric bounds. Section 3 proves Theorem 1.1. Sections 4 and 5 establish the automatic degree

bounds and prove Theorem 1.3. Sections 6, 7, and 8 develop the extremal and zero-surplus refinements. Section 9 proves the sharp counting consequences. The remaining sharp-count question is stated in Section 10. Appendices A and B contain the incidence-completion and polarity-extension results used in the structural proofs.

2 Preliminaries

All graphs are finite, simple, and undirected. We write $N_G(v)$ and $d_G(v)$ for the neighbourhood and degree of v, omitting G when the ambient graph is fixed. For disjoint vertex sets X, Y, write $e(X,Y)$ for the number of edges with one endpoint in each set. The parameters N, M, n, m_0 are as in (1.1).

2.1 Codegrees and degree deficits

For a vertex v and a set $X \subseteq V(G)$, write $d_X(v) = |N(v) \cap X|$, and put $d(v) = |N(v)|$. For the supersaturation estimates, put

$$n = q^2 + q, \qquad m_0 = \frac{1}{2}q(q+1)^2 - q, \qquad m_+ = m_0 + 1, \qquad Q = N_{C_4}(G).$$

Unless stated otherwise, q is an integer with $q \geq 2$. For distinct vertices u, v, let

$$c_{uv} = |N(u) \cap N(v)|.$$

Writing $x_+ = \max\{x, 0\}$, define

$$\begin{aligned} \delta_v &= q + 1 - d(v), \qquad a_v = \sum_{w \in N(v)} \delta_w, \\ \varepsilon_v &= \sum_{w \neq v} (c_{vw} - 1)_+, \quad u_v = |\{w \neq v : c_{vw} = 0\}|. \end{aligned} \tag{2.1}$$

The deficits δ_v and their neighbourhood sums a_v are allowed to be negative. A pair of distinct vertices is *uncovered* if its codegree is zero. We consistently write

$$H = \{v : d(v) = q+1\}, \qquad L = \{v : d(v) \leq q\}.$$

Lemma 2.1. *Let G have $q^2 + q$ vertices, and let Q_v be the number of four-cycles containing v. Then*

$$Q_v = \sum_{w \neq v} \binom{c_{vw}}{2}, \qquad 2Q = \sum_{\{v,w\} \subseteq V(G)} \binom{c_{vw}}{2}, \tag{2.2}$$

$$\varepsilon_v \leq Q_v \leq Q, \qquad \sum_v \varepsilon_v \leq 4Q, \tag{2.3}$$

$$u_v = -1 + q\delta_v + a_v + \varepsilon_v. \tag{2.4}$$

If $e(G) = m_+$, then

$$\sum_v \delta_v = 2q - 2. \tag{2.5}$$

Let A be the adjacency matrix, let J be the all-ones matrix, and define symmetric matrices U, R with zero diagonal by

$$U_{uv} = \mathbf{1}_{\{c_{uv}=0\}}, \qquad R_{uv} = (c_{uv} - 1)_+ \qquad (u \neq v).$$

Then

$$A^2 = qI + J - U - \mathrm{diag}(\delta_v) + R, \tag{2.6}$$

and, with $[S,T] = ST - TS$,

$$[A, U]_{uv} = (d(u) - d(v))(1 - A_{uv}) + [A, R]_{uv}. \tag{2.7}$$

Proof. A four-cycle containing v is determined by its vertex opposite v and its two common neighbours with v. This proves the first identity in (2.2); counting the two opposite pairs of every four-cycle proves the second. The inequalities in (2.3) follow from $(r-1)_+ \le \binom{r}{2}$ for nonnegative integers r.

Counting two-edge paths starting at v gives

$$n - 1 - u_v + \varepsilon_v = \sum_{w \neq v} c_{vw} = \sum_{w \in N(v)} (d(w) - 1) = q\, d(v) - a_v.$$

Substituting $n = q^2 + q$ and $d(v) = q + 1 - \delta_v$ proves (2.4). The degree sum yields (2.5) when $e(G) = m_+$.

The off-diagonal entries of (2.6) express $c_{uv} = 1 - U_{uv} + R_{uv}$, and its diagonal entries express $d(v) = q + 1 - \delta_v$. Commuting this identity with A, and using

$$[A, J]_{uv} = d(u) - d(v), \qquad [A, \operatorname{diag}(\delta_v)]_{uv} = (d(u) - d(v))A_{uv},$$

gives (2.7). □

For even q, call a vertex v *good* if

$$v \in H, \qquad \varepsilon_v = 0, \qquad a_v = 2,$$

and denote the set of good vertices by Z_0.

Lemma 2.2. *Suppose that q is even and $|V(G)| = q^2 + q$. Every vertex $v \in H$ with $\varepsilon_v = 0$ satisfies $a_v \ge 2$. Every good vertex v lies in no four-cycle and has a unique uncovered partner $f(v)$; moreover, $vf(v) \in E(G)$. For any two good vertices u, v,*

$$A_{u,f(v)} = A_{f(u),v}. \tag{2.8}$$

If u, w are good and $f(u) = w$, then $f(w) = u$ and w is the only good neighbour of u.

Proof. For $v \in H$ with $\varepsilon_v = 0$, equation (2.4) gives $u_v = a_v - 1 \ge 0$. If $a_v = 1$, then $c_{vw} = 1$ for every $w \neq v$, contrary to

$$2e(G[N(v)]) = \sum_{w \in N(v)} c_{vw} = q + 1,$$

since $q + 1$ is odd. Hence $a_v \ge 2$.

If v is good, then $u_v = 1$, and (2.2) gives $Q_v = 0$. If its unique uncovered partner were outside $N(v)$, the same parity identity would again have right-hand side $q + 1$. Thus $vf(v)$ is an edge. The row and column of R at a good vertex are zero, and all good vertices have the same degree. Taking the (u, v) entry of (2.7) therefore gives (2.8).

If $f(u) = w$ and w is good, symmetry of U gives $f(w) = u$. Suppose that $x \neq w$ is another good neighbour of u. Equation (2.8), applied to w, x, yields $A_{w,f(x)} = A_{u,x} = 1$. The vertices $u, x, f(x), w$ are distinct: $f(x) = u$ would contradict $f(u) = w$, and $f(x) = w$ would contradict $f(w) = u$. They form a four-cycle, contrary to $Q_u = 0$. □

2.2 C_4-free cores and full-degree surplus

A *core at deficit b* is a C_4-free graph F on n vertices with $e(F) = m_0 - b$, where $b \ge 0$ is an integer. An *extremal core* means such a graph with $b = 0$. The degree sum and Lemma 2.1 give

$$\sum_v \delta_v = 2q + 2b, \qquad u_v = q\delta_v - 1 + a_v.$$

For even q, Lemma 2.2 gives $a_v \geq 2$ at every full-degree vertex. The surplus in (1.4) is therefore

$$\eta(F) = \sum_{v \in H} (a_v - 2) \geq 0.$$

When $\Delta(F) \leq q+1$, the sets H and L partition $V(F)$, and all deficits are nonnegative. The extremal hypotheses used in Section 6 are

$$|V(F)| = q^2 + q, \qquad e(F) = m_0, \qquad F \text{ is } C_4\text{-free}, \qquad \Delta(F) \leq q+1. \tag{2.9}$$

2.3 Polarity graphs and the upper construction

A projective plane of order q has q^2+q+1 points and the same number of lines. Every point lies on $q+1$ lines, every line contains $q+1$ points, and two distinct points or lines determine a unique common incident object. A polarity is an incidence-reversing involution that interchanges points and lines. A point x is absolute if it lies on its polar line. A polarity is orthogonal when it has exactly $q+1$ absolute points.

Given a polarity π, let H be its simple polarity graph: distinct points x, y are adjacent precisely when $x \in \pi(y)$. Index the incidence matrix by points, using $\pi(y)$ as the line indexed by y. The resulting matrix B is symmetric and satisfies

$$B\mathbf{1} = (q+1)\mathbf{1}, \qquad B^2 = qI + J. \tag{2.10}$$

If $D = \operatorname{diag}(\alpha_x)$, where α_x indicates that x is absolute, then the adjacency matrix of H is $A = B - D$. Thus

$$d_H(x) = q + 1 - \alpha_x, \qquad (A^2)_{xy} = 1 - A_{xy}(\alpha_x + \alpha_y) \quad (x \neq y). \tag{2.11}$$

In particular, H is C_4-free, its absolute points are independent, and every nonadjacent pair has a unique common neighbour.

Proposition 2.3 (Four-cycles created by one edge)**.** *Let P be the simple polarity graph of a projective plane of order $q \geq 2$. For a nonedge uv, let z be its unique common neighbour in P, and let α_x indicate that x is absolute. Then*

$$N_{C_4}(P + uv) = q + 1 - \alpha_u - \alpha_v - \alpha_z. \tag{2.12}$$

For any point $p \notin \{u, v\}$, put $K = P - p$. Then

$$N_{C_4}(K + uv) \geq q - 2; \qquad N_{C_4}(K + uv) \geq q - 1 \quad \text{if } p \text{ is absolute.} \tag{2.13}$$

For each fixed added edge, the old edge sets of the resulting four-cycles are pairwise disjoint, both in P and in K.

Proof. Let A be the adjacency matrix of P, put $D = \operatorname{diag}(\alpha_x)$, and set $B = A + D$. The incidence identities (2.10) apply to B. Since uv is a nonedge, every length-three walk from u to v has distinct vertices. Hence its number is exactly $N_{C_4}(P + uv)$. Expand $A^3 = (B - D)^3$. The terms containing B_{uv} vanish. By (2.10),

$$(B^3)_{uv} = q + 1, \qquad (B^2 D)_{uv} = \alpha_v, \qquad (DB^2)_{uv} = \alpha_u.$$

Moreover $(BDB)_{uv} = \alpha_z$, because z is the unique common neighbour and the terms indexed by u, v vanish. This proves (2.12).

If z is absolute, both endpoints are nonabsolute. Thus the sum of the three indicators is at most two, and (2.12) is at least $q - 1$.

Suppose first that $p = z$ and p is nonabsolute. The paths using p as the first internal vertex number $|N_P(p) \cap N_P(v)| = 1 - \alpha_v$, and those using it as the second internal vertex number $1 - \alpha_u$, by (2.11). These two path families are disjoint. Deleting p therefore removes exactly $2 - \alpha_u - \alpha_v$ paths and leaves exactly $q - 1$. If $p = z$ is absolute, the two corresponding codegrees are zero, so the deletion removes no path and leaves q.

If $p \neq z$, it is adjacent to at most one of u, v. If it is adjacent to neither, it occurs in no such path. If it is adjacent to exactly one, the codegree bound shows that it occurs in at most one path. Thus at most one path is lost, proving (2.13). When p is absolute, the only case in which the original count is $q - 1$ has both endpoints absolute; independence of the absolute points then means no path is lost. In all other cases the original count is at least q. This proves the stronger bound for an absolute deletion.

Finally, two distinct length-three u–v paths in a C_4-free graph cannot share their first or last edge: such a shared edge would give two common neighbours to another pair. If their middle edge is shared, they either coincide or traverse it in opposite directions; the latter possibility also gives a four-cycle. An end edge of one path cannot be an end edge of the other at the opposite endpoint, or its middle edge, because u, v do not occur internally. Thus the three-edge sets of these paths are pairwise disjoint. Deleting p preserves this property. □

The identity (2.12) is the standard single-edge count for polarity graphs; compare [9, Lemma 2.5]. The point-deletion bounds are used in the construction below and in Section 9.

Proposition 2.4 (An attaining construction)**.** *For every power of two $q \geq 4$, there is a graph G on $q^2 + q$ vertices with $m_0 + 1$ edges and exactly $q - 1$ four-cycles. It has an edge e such that $F = G - e$ satisfies (2.9) and $\eta(F) = 0$.*

Proof. Take the points of $\mathrm{PG}(2, q)$ as vertices, and join distinct projective points x, y when

$$x_0y_0 + x_1y_1 + x_2y_2 = 0.$$

The nondegenerate symmetric bilinear form defines a polarity. In characteristic two its absolute points form the line

$$\mathcal{A} = \{[x_0 : x_1 : x_2] : x_0 + x_1 + x_2 = 0\}.$$

There are $q + 1$ absolute points, each of degree q; the other q^2 points have degree $q + 1$. The resulting graph H_q is C_4-free and has

$$e(H_q) = \frac{(q+1)q + q^2(q+1)}{2} = M.$$

The point $c = [1 : 1 : 1]$ is not absolute and has neighbourhood $\mathcal{A}$.

Choose distinct $p, u, v \in \mathcal{A}$, delete p, and add the edge uv. This gives $n = N - 1$ vertices and $M - q + 1 = m_0 + 1$ edges. The pair u, v has common neighbour c. Equation (2.12) gives $q - 1$ four-cycles before deleting p. Since p is adjacent to neither u nor v, the deletion does not change this count.

It remains to verify the surplus assertion. Put $X = \mathcal{A} \setminus \{p\}$ and $Y = N_{H_q}(p)$. These sets are disjoint and both have size q. In $F = H_q - p$, precisely the vertices of $X \cup Y$ have degree q; all other vertices have degree $q + 1$. Every full-degree vertex w differs from c, because $c \in Y$. It therefore has exactly one absolute neighbour, since its polar line meets $\mathcal{A}$ in one point. This neighbour belongs to X, as $w \notin N_{H_q}(p)$. Also p, w have a unique common neighbour, which belongs to Y. Thus w has exactly two neighbours of deficit one, and all its other neighbours have deficit zero. Therefore $\eta(F) = 0$. □

Proof of Corollary 1.2. For $q = 2^k$, the theorem of Firke, Kosek, Nash, and Williford [6] gives $\mathrm{ex}(q^2 + q, C_4) = m_0$. The lower bound follows from Theorem 1.1 and the integrality of the number of four-cycles. Proposition 2.4 gives the upper bound. Since $q \sim \sqrt{q^2 + q}$, the stated order of magnitude follows. □

2.4 Elementary geometric bounds

Lemma 2.5 (Concurrency from a bounded exceptional set)**.** *Let $\mathcal{F}$ be a family of r distinct lines of a projective plane. Suppose that the intersection point of every two distinct lines of $\mathcal{F}$ belongs to a set B, and every line of $\mathcal{F}$ has at most a points in B. If $r-1 > a(a-1)$, all lines of $\mathcal{F}$ are concurrent.*

Proof. If the family is not concurrent, fix $\ell \in \mathcal{F}$. For every point $p \in \ell \cap B$ used as an intersection, there is a line $\ell' \in \mathcal{F}$ not through p. All members of $\mathcal{F}$ through p meet ℓ' in distinct points of B, so at most a of them exist, including ℓ. Partitioning $\mathcal{F} \setminus \{\ell\}$ by its intersection with ℓ gives $r-1 \le a(a-1)$. □

Lemma 2.6 (An elementary edge bound for even-order polarities)**.** *Let π be a polarity of a projective plane of even order q. Its loopless graph P_π has minimum degree at least q and at most M edges. If it has M edges, its absolute points form a line of $q+1$ points.*

Proof. An absolute point has degree q and any other point has degree $q+1$. Distinct absolute points cannot be adjacent: otherwise their two distinct polar lines both contain both points. For a nonabsolute vertex x, a neighbour y is isolated in the matching induced by $N(x)$ precisely when y is absolute. Indeed, the distinct lines $\pi(x), \pi(y)$ meet at a unique point; this supplies a common neighbour unless that point is y itself, which is exactly the absolute case. Since $q+1$ is odd, x has an odd, positive number of absolute neighbours.

If a is the number of absolute points, counting their edges to the $N-a$ other points gives $qa \ge N-a$ and thus $a \ge q+1$. Consequently

$$e(P_\pi) = \frac{N(q+1)-a}{2} \le M.$$

If equality holds, $a = q+1$. Every line meets the absolute set: an absolute point's polar line contains that point, and every other polar line contains the positive odd number of absolute neighbours just proved. A blocking set of $q+1$ points is a line. To see this, take a line through two of its points. If that line contained a point outside the set, the other q lines through that outside point would require at least q additional points of the set off the chosen line, whereas at most $q-1$ are available. Hence the chosen line consists entirely of the $q+1$ absolute points. □

3 Supersaturation one edge above the bound

We prove Theorem 1.1 by treating all possible degree configurations at $m_+ = m_0 + 1$ edges. Throughout this section, $Q = N_{C_4}(G)$. The estimates are summarized below; the final subsection combines them and passes to larger edge counts.

Degree configuration	Lower bound
$\Delta(G) \le q+1$	$12Q \ge q+5$
$\Delta(G) \ge q+3$	$Q > q/2$
Exactly one vertex of degree $q+2$, all others at most $q+1$	$12Q \ge q-1$
At least two vertices of degree $q+2$, maximum degree $q+2$	$Q > q/12$

All four estimates hold for even $q \ge 1600$; the individual statements below give their respective ranges. In the last case, we also prove the asymptotic improvement $Q > q/4 - 3q^{3/4}$.

3.1 Maximum degree at most $q+1$

Theorem 3.1. *Let $q \geq 2$ be even. If G has q^2+q vertices, m_+ edges, and maximum degree at most $q+1$, then*

$$12N_{C_4}(G) \geq q+5. \tag{3.1}$$

Proof. In this proof $V(G) = L \cup H$. Write $h = |H|$, and put

$$\sigma = \sum_{v\in L}(\delta_v - 1), \qquad \chi = \sum_{v\in L}\delta_v(\delta_v-1),$$
$$\omega = \sum_{uv\in E(G[L])}(\delta_u+\delta_v), \qquad \beta = \sum_{v\in H}\varepsilon_v,$$
$$\mathcal{D} = 4 + 2\sigma + \chi + \omega.$$

All these parameters are nonnegative. The degree sum gives

$$|L| = 2q-2-\sigma, \qquad h = q^2-q+2+\sigma. \tag{3.2}$$

Let $B = H \setminus Z_0$ and $b = |B|$. Weighted counting of the edges between L and H gives

$$\sum_{v\in H} a_v = \sum_{w\in L}\delta_w d_H(w) = q(2q-2) - \chi - \omega.$$

Indeed, $d_H(w) = q+1-\delta_w - d_L(w)$, $\sum_{w\in L}\delta_w^2 = \chi + 2q - 2$, and $\sum_{w\in L}\delta_w d_L(w) = \omega$. Consequently

$$\mathcal{D} = \sum_{v\in H}(2-a_v), \qquad \sum_{v\in B} u_v = b - \mathcal{D} + \beta. \tag{3.3}$$

For the second identity, sum $u_v = a_v - 1 + \varepsilon_v$ over H and subtract the contribution of the $h-b$ good vertices.

We shall use two bounds for b. Let

$$C = \{v\in H : \varepsilon_v > 0\}, \qquad T = \{v\in H : \varepsilon_v = 0,\ a_v \geq 3\}, \qquad p = \sum_{v\in T}(a_v-2).$$

By Lemma 2.2, $B = C \cup T$ is a disjoint union. Since $|C| \leq \beta$, $|T| \leq p$, and

$$\mathcal{D} = 2|C| - \sum_{v\in C} a_v - p,$$

we have

$$b \leq 3|C| - \sum_{v\in C} a_v - \mathcal{D} \leq 3\beta - \mathcal{D}. \tag{3.4}$$

All a_v are nonnegative under the degree hypothesis. Thus, for each $w \in C$,

$$b + a_w \leq 3\beta - \mathcal{D}. \tag{3.5}$$

A vertex of B with at most one uncovered partner. Suppose that $v \in B$ and $u_v \leq 1$. Then $v \in C$, since $\varepsilon_v = 0$ would imply $a_v = u_v + 1 \leq 2$ and hence $v \in Z_0$. Choose w such that $R_{vw} > 0$. For every good vertex x, equation (2.7) has the form

$$A_{v,f(x)} + \sum_y R_{vy}A_{yx} = \sum_{z:U_{vz}=1} A_{zx}. \tag{3.6}$$

If $u_v = 0$, this forces w to have no good neighbour. If $u_v = 1$, let z be the unique uncovered partner of v. Every good neighbour of w is also adjacent to z. The vertices w, z are distinct

because $R_{vw} > 0$ and $U_{vz} = 1$. Two such neighbours would form a four-cycle with w, z, contradicting Lemma 2.2. In either case,

$$d_{Z_0}(w) \le 1. \tag{3.7}$$

If $w \in H$, symmetry of R gives $w \in C$. The vertex w has at most $b-1$ neighbours in B and at most a_w neighbours in L. Equations (3.5) and (2.3) give

$$q+1 = d(w) \le 1 + (b-1) + a_w \le 3\beta - \mathcal{D} \le 12Q - 4.$$

This proves (3.1).

If $w \in L$, put $\delta = \delta_w \ge 1$ and $s = d_L(w)$. The contributions of w to σ, χ, and of the edges of $G[L]$ incident with w to ω, give

$$\mathcal{D} \ge 4 + 2(\delta - 1) + \delta(\delta - 1) + (\delta + 1)s.$$

In particular,

$$\mathcal{D} - \delta - s \ge \delta^2 + 2 + \delta s \ge 3. \tag{3.8}$$

Using (3.7) and (3.4), we obtain

$$q+1 = d(w) + \delta \le 1 + b + s + \delta \le 1 + 3\beta - \mathcal{D} + s + \delta \le 3\beta - 2.$$

Here $\varepsilon_w \ge 1$ and $w \notin H$, so $\beta \le 4Q - 1$. Thus $q + 1 \le 12Q - 5$, again implying (3.1).

Every vertex of B has at least two uncovered partners. We may now assume that $u_v \ge 2$ for all $v \in B$. Equation (3.3) yields

$$b \le \beta - \mathcal{D}. \tag{3.9}$$

Suppose first that u, w are good and $f(u) = w$. Lemma 2.2 shows that w is the only good neighbour of u. Since $a_u = 2$, at most two neighbours of u belong to L. Hence $b \ge q - 2$, and

$$q - 2 \le b \le \beta - \mathcal{D} \le 4Q - 4.$$

The resulting bound $4Q \ge q + 2$ is stronger than (3.1).

It remains to consider the case where no two good vertices are uncovered. Put

$$Z = \{v \in Z_0 : f(v) \in L\}.$$

Every good vertex outside Z has its uncovered partner in B. Therefore, by (3.3) and (3.9),

$$\begin{aligned} |Z| &\ge h - b - \sum_{v \in B} u_v = h - 2b + \mathcal{D} - \beta \\ &\ge h - 3\beta + 3\mathcal{D}. \end{aligned} \tag{3.10}$$

Let $X = f(Z) \subseteq L$ and $Y = L \setminus X$. For distinct $a, b \in X$, choose $u, v \in Z$ with $f(u) = a$ and $f(v) = b$. If u were adjacent to b, then (2.8) would make v adjacent to a, producing the four-cycle a, u, b, v, a. Thus each $u \in Z$ has exactly one neighbour in X, namely $f(u)$.

Because $a_u = 2$, every $a \in X$ has deficit one or two. If $\delta_{f(u)} = 1$, then u has precisely one other neighbour $b \in L$, which lies in Y and has deficit one. For each pair $(a, b) \in X \times Y$, at most one such vertex u is possible, since two would form a four-cycle with a, b. If $\delta_{f(u)} = 2$, there is no other neighbour of u in L, and the fibre $f^{-1}(f(u)) \cap Z$ has size at most $d(f(u)) = q - 1$. Write

$$x_i = |\{a \in X : \delta_a = i\}| \quad (i = 1, 2), \qquad y_1 = |\{b \in Y : \delta_b = 1\}|.$$

Then

$$|Z| \le x_1 y_1 + (q-1)x_2, \qquad x_1 + y_1 + 2x_2 \le 2q-2.$$

Since $0 \le x_2 \le q-1$, it follows that

$$\begin{aligned} |Z| &\le (q-1-x_2)^2 + (q-1)x_2 \\ &= (q-1)^2 - x_2(q-1-x_2) \le (q-1)^2. \end{aligned} \tag{3.11}$$

Combining (3.10), (3.11), and (3.2) gives

$$3\beta \ge h + 3\mathcal{D} - (q-1)^2 = q + 13 + 7\sigma + 3\chi + 3\omega.$$

As $\beta \le 4Q$, this proves $12Q \ge q+13$ and completes the proof. □

3.2 Vertices of degree at least $q+3$

We use the following form of the neighbourhood-complement argument of Firke et al. [6, Lemma 1].

Lemma 3.2. *Let G be a graph on $\nu \ge 3$ vertices with m edges and Q four-cycles. If a vertex u has degree D, put*

$$s = 2m - D - (\nu-1) - Q.$$

If $s/(\nu-1) \ge 1/2$, then

$$(\nu-1)\binom{s/(\nu-1)}{2} \le \binom{\nu-D}{2} + 2Q. \tag{3.12}$$

Here $\binom{x}{2} = x(x-1)/2$ for real x.

Proof. Set $S = N(u)$ and $B = V(G) \setminus S$. For $v \ne u$, put

$$p_v = |N(v) \cap S| = c_{uv}, \qquad b_v = |N(v) \cap B| = d(v) - p_v.$$

Counting four-cycles containing u gives $\sum_{v \ne u} \binom{p_v}{2} = Q_u \le Q$. Consequently

$$\sum_{v \ne u} p_v \le (\nu-1) + \sum_{v \ne u} (p_v - 1)_+ \le (\nu-1) + Q,$$

and $\sum_{v \ne u} b_v \ge s$.

Since $N(u) \cap B = \varnothing$, the number of two-edge paths with endpoints in B is

$$\sum_{v \ne u} \binom{b_v}{2} = \sum_{\{x,y\} \subseteq B} c_{xy}.$$

It is at most

$$\begin{aligned} \binom{|B|}{2} + \sum_{\{x,y\} \subseteq B} (c_{xy}-1)_+ &\le \binom{|B|}{2} + \sum_{\{x,y\} \subseteq V(G)} \binom{c_{xy}}{2} \\ &= \binom{\nu-D}{2} + 2Q. \end{aligned}$$

Jensen's inequality, together with the monotonicity of $x(x-1)/2$ on $[1/2, \infty)$, bounds the same number below by $(\nu-1)\binom{s/(\nu-1)}{2}$. □

Theorem 3.3. *Let $q \geq 8$ be an integer. If G has q^2+q vertices, at least m_+ edges, and maximum degree at least $q+3$, then*

$$N_{C_4}(G) > q/2.$$

Proof. Suppose that $Q \leq q/2$, and choose a vertex of degree $D \geq q+3$. Put $a = n-1 = q^2+q-1$. For $D \leq n-1$ and $e(G) \geq m_+$, the quantity s in Lemma 3.2 satisfies

$$s \geq 2m_+ - 2(n-1) - Q = q^3 - 3q + 4 - Q.$$

Moreover,

$$q^3 - 3q + 4 - Q - (q-2)a = q^2 + 2 - Q > 0.$$

Thus $s/a > q-2 \geq 6$, and Lemma 3.2 implies

$$F(D,Q) := a\binom{(2m_+ - D - a - Q)/a}{2} - \binom{n-D}{2} - 2Q \leq 0. \tag{3.13}$$

The replacement of $e(G)$ by m_+ is valid by the same monotonicity.

For fixed Q, the coefficient of D^2 in $F(D,Q)$ is $1/(2a) - 1/2 < 0$. Its minimum on $[q+3, n-1]$ is therefore attained at an endpoint. Expanding at the first endpoint gives

$$2F(q+3,Q) = q^2 + 3q - 12 - (2q+3)Q + \frac{(2q+Q)^2}{a}. \tag{3.14}$$

The polynomial part is at least $3q/2-12 \geq 0$, and the last term is positive. Hence $F(q+3,Q) > 0$. At the other endpoint,

$$F(n-1,Q) \geq a\binom{q-2}{2} - 2Q \geq a\binom{q-2}{2} - q > 0.$$

This contradicts (3.13). □

3.3 A unique vertex of degree $q+2$

The remaining degree configurations have maximum degree $q+2$. We first consider graphs with a unique vertex of this degree.

Theorem 3.4. *Let $q \geq 6$ be even. Suppose that G has n vertices and m_+ edges, exactly one vertex of degree $q+2$, and all other degrees at most $q+1$. Then*

$$12N_{C_4}(G) \geq q - 1.$$

Throughout the proof, z denotes the unique vertex of degree $q+2$, so $\delta_z = -1$. We use the notation, codegree identities, and good-vertex properties from Section 2.1. In addition to (2.3), every pair of distinct vertices satisfies

$$\binom{c_{uv}}{2} \leq Q, \tag{3.15}$$

because the cycles counted on the left have u, v as opposite vertices.

3.3.1 Signed mass and short uncovered rows

Put $h = |H|$. Define

$$\sigma = \sum_{v\in L}(\delta_v - 1), \qquad \chi = \sum_{v\in L}\delta_v(\delta_v - 1),$$
$$\omega = \sum_{uv\in E(G[L])}(\delta_u + \delta_v), \qquad \tau = \sum_{v\in N(z)\cap L}(\delta_v - 1),$$
$$\beta = \sum_{v\in H}\varepsilon_v, \qquad k = d_H(z), \qquad \mathcal{D} = 2 + 2\sigma + \chi + \omega + \tau.$$

Write $B = H \setminus Z_0$ and $b = |B|$. The degree sum and the local deficit identity give

$$\sum_{v\in L}\delta_v = 2q - 1, \qquad |L| = 2q - 1 - \sigma, \qquad h = q(q-1) + \sigma, \tag{3.16}$$
$$k \le 1 + \tau + \varepsilon_z, \qquad \beta + \varepsilon_z \le 4Q, \tag{3.17}$$
$$\mathcal{D} = \sum_{v\in H}(2 - a_v), \qquad \sum_{v\in B}u_v = b - \mathcal{D} + \beta. \tag{3.18}$$

For the first relation in (3.17), use $a_z = q + 2 - k + \tau$ in (2.4). For (3.18), weighted edge counting gives

$$\sum_{v\in H}a_v = \sum_{w\in L}\delta_w d_H(w) - k = q(2q-1) - \chi - \omega - a_z - k = 2q^2 - 2q - 2 - \chi - \omega - \tau.$$

Summation of (2.4) over B gives the other identity.

Let $C = \{v \in H : \varepsilon_v > 0\}$ and

$$D_- = \sum_{v\in H}(2 - a_v)_+, \qquad D_+ = \sum_{v\in H}(a_v - 2)_+.$$

If $a_v < 2$, then $\varepsilon_v \ge 1$ and $2 - a_v \le 1 + \varepsilon_v \le 2\varepsilon_v$. Thus $D_- \le 2\beta$ and $D_+ \le 2\beta - \mathcal{D}$. Vertices of $B \setminus C$ have $a_v \ge 3$, and consequently

$$b \le 3\beta - \mathcal{D}, \qquad b + a_w \le 3\beta - \mathcal{D} + 2 \quad (w \in C). \tag{3.19}$$

To obtain the second inequality without assuming $a_v \ge 0$, observe that

$$b \le |C| + D_+ - \sum_{v\in C}(a_v - 2)_+.$$

For $w \in C$, the quantity $a_w - (a_w - 2)_+$ is at most two.

Lemma 3.5. *If there exist $v \in B$ and $w \ne z$ such that $u_v \le 1$ and $R_{vw} > 0$, then $12Q \ge q - 1$.*

Proof. The vertex v belongs to C. For every good x, equation (2.7) becomes

$$A_{v,f(x)} + \sum_y R_{vy}A_{yx} = \sum_{a:U_{va}=1}A_{ax}. \tag{3.20}$$

If $u_v = 0$, then w has no good neighbour. If $u_v = 1$, all good neighbours of w are also adjacent to the unique uncovered partner of v. That partner differs from w, and two such neighbours would give a four-cycle containing good vertices. In all cases, $d_{Z_0}(w) \le 1$.

If $w \in H$, then $w \in C$. Writing $t = A_{wz}$, we have $d_L(w) \le a_w + t$ and hence

$$q + 1 = d(w) \le 1 + (b-1) + a_w + 2t \le b + a_w + 2 \le 3\beta - \mathcal{D} + 4 \le 12Q + 2.$$

This gives the assertion. If $w \in L$, put $\delta = \delta_w$, $s = d_L(w)$, and $t = A_{wz}$. The contributions of w and its incident low edges imply

$$\mathcal{D} \geq 2 + 2(\delta - 1) + \delta(\delta - 1) + (\delta + 1)s + (\delta - 1)t.$$

Thus $\mathcal{D} - \delta - s - t \geq \delta^2 + \delta s + (\delta - 2)t \geq 0$. It follows that

$$q + 1 = d(w) + \delta \leq 1 + b + s + t + \delta \leq 1 + 3\beta.$$

Since $\varepsilon_w \geq 1$ and $w \notin H$, $\beta \leq 4Q - 1$. Therefore $12Q \geq q + 3$ in this case. □

Zero-row pairing. The remaining short rows have their excess entirely at z. The extra capacity furnished by these rows is controlled by a pairing inside $N(z) \cap H$.

Lemma 3.6. *Suppose that every $v \in B$ with $u_v \leq 1$ has $R_{vw} = 0$ for all $w \neq z$. Put*

$$S = \{v \in B : u_v \leq 1\}, \qquad S_0 = \{v \in B : u_v = 0\}, \qquad s_0 = |S_0|, \qquad \lambda = \sum_{v \in S}(2 - u_v).$$

Then

$$\lambda \leq \varepsilon_z + s_0, \qquad 2s_0 \leq k, \qquad 2s_0 + g \leq k, \quad g = |N(z) \cap Z_0|. \tag{3.21}$$

Moreover, every $v \in S_0$ is adjacent to z, has $\varepsilon_v = 1$ and $a_v = 0$, and has exactly one low neighbour, of deficit one.

Proof. Every member of S has positive excess, so $\varepsilon_v = R_{vz} \geq 1$ for $v \in S$. By symmetry of R, $\sum_{v \in S} \varepsilon_v \leq \varepsilon_z$. Writing $s_1 = |S \setminus S_0|$ gives $\lambda = 2s_0 + s_1 \leq \varepsilon_z + s_0$.

Let $v \in S_0$. If v were not adjacent to z, every vertex of $N(v)$ would have codegree one with v, contrary to the odd order $q + 1$ of this neighbourhood. Hence vz is an edge. Equation (2.4) gives $a_v = 1 - \varepsilon_v$. Since $a_v \geq -1$, we have $1 \leq \varepsilon_v \leq 2$. Also

$$2e(G[N(v)]) = \sum_{w \in N(v)} c_{vw} = q + 1 + \varepsilon_v,$$

so ε_v is odd. Thus $\varepsilon_v = 1$ and $a_v = 0$. The positive neighbour deficits of v sum to one, since the neighbour z contributes -1. This proves the assertion about its low neighbour. In particular $c_{vz} = 2$.

For every $w \in N(v) \setminus \{z\}$, one has $c_{zw} = 1$. The vertex v is a common neighbour of z, w; another common neighbour x would make z, w two distinct common neighbours of v, x, contradicting the support condition on the vth row of R. The two common neighbours of v, z cannot both be low, because v has only one low neighbour. At least one therefore belongs to $N(z) \cap H$, and it does not belong to S_0, since its codegree with z is one. These choices are distinct for distinct members of S_0: two members of S_0 already have common neighbour z, and their mutual codegree is exactly one. We have obtained an injection from S_0 into $(N(z) \cap H) \setminus S_0$, proving $2s_0 \leq k$.

If S_0 is nonempty, apply (3.20) to one of its members. The right-hand side is zero, and $R_{vz} = 1$, so $A_{zx} = 0$ for every good x; that is, $g = 0$. If S_0 is empty, $g \leq k$ is immediate. These two cases prove $2s_0 + g \leq k$. □

3.3.2 The fibre deficit and the neighbourhood union

Proof of Theorem 3.4. Suppose, for a contradiction, that $12Q < q - 1$. Lemma 3.5 shows that the hypotheses of Lemma 3.6 hold. From (3.18),

$$2b - \lambda \leq \sum_{v \in B} u_v = b - \mathcal{D} + \beta, \qquad b \leq \beta - \mathcal{D} + \lambda.$$

If good vertices u, v are uncovered, they are adjacent and have no other good neighbours. A good vertex x has at most $2 + A_{xz}$ low neighbours and at most A_{xz} high neighbours. Since $c_{uv} = 0$, at most one of u, v is adjacent to z; also their bad-neighbour sets are disjoint. Consequently $b \geq 2q - 6$. On the other hand, $s_0 \leq \varepsilon_z$, so

$$b \leq \beta + 2\varepsilon_z - \mathcal{D} \leq 5Q - 2.$$

Thus $5Q \geq 2q - 4$, contradicting $12Q < q - 1$ for $q \geq 6$. There is therefore no uncovered pair of good vertices.

Exceptional full-degree vertices. Set

$$Z = \{v \in Z_0 : f(v) \in L,\ vz \notin E(G)\}, \qquad \mu = h - |Z|.$$

Every good vertex omitted from Z either has its uncovered partner in B or is adjacent to z. Using (3.18) and (3.21), we obtain

$$\begin{aligned} \mu &\leq b + \sum_{v \in B} u_v + g = 2b - \mathcal{D} + \beta + g \\ &\leq 3\beta - 3\mathcal{D} + 2\lambda + g \\ &\leq 3\beta + 2\varepsilon_z + k - 3\mathcal{D} \\ &\leq 12Q - 5 - 6\sigma - 3\chi - 3\omega - 2\tau. \end{aligned} \tag{3.22}$$

Put $X = f(Z)$ and $Y = L \setminus X$. Every $u \in Z$ has exactly one neighbour in X, namely $f(u)$: another such neighbour, together with (2.8), would give a four-cycle containing good vertices. Since u is not adjacent to z, its positive neighbour deficits sum to two. If $\delta_{f(u)} = 1$, it has one further low neighbour, of deficit one, in Y, and no two such vertices use the same pair in $X \times Y$. If $\delta_{f(u)} = 2$, its fibre has size at most $q - 1$. With

$$x_i = |\{a \in X : \delta_a = i\}|\ (i = 1, 2), \qquad y_1 = |\{b \in Y : \delta_b = 1\}|,$$

it follows that

$$|Z| \leq x_1 y_1 + (q-1)x_2, \qquad x_1 + y_1 + 2x_2 \leq 2q - 1. \tag{3.23}$$

If $\sigma \geq 1$ and $x_2 \geq 1$, the right-hand side is at most

$$(q - x_2)(q - 1 - x_2) + (q-1)x_2 = q(q-1) - x_2(q - x_2) \leq q(q-1) - (q-1).$$

If $\sigma \geq 1$ and $x_2 = 0$, an inactive vertex of deficit at least two gives $x_1 + y_1 \leq 2q - 3$, and the same upper bound holds. Since $h = q(q-1) + \sigma$, either case yields $\mu \geq q - 1 + \sigma$, contrary to (3.22) and $12Q < q - 1$. Thus $\sigma = 0$.

All low deficits are now one, $|L| = 2q - 1$, $h = q(q-1)$, $\chi = \tau = 0$, and $\mathcal{D} = 2 + \omega \geq 2$. The vertices of Z give distinct edges of a simple bipartite support graph on X, Y, whose maximum degree is at most q. If one part had size at most $q - 2$, then $|Z| \leq q(q-2) = h - q$, contradicting (3.22). Consequently

$$\{|X|, |Y|\} = \{q, q-1\}, \qquad \mu \leq 3\beta + 2\varepsilon_z + k - 3\mathcal{D} \leq 12Q - 5 < q. \tag{3.24}$$

Here μ is also the number of missing support pairs. Since $\tau = 0$, (3.17) gives $k \leq 1 + \varepsilon_z \leq 1 + Q$.

Let $r = |N(z) \cap X|$, $s = |N(z) \cap Y|$, and $\ell = r + s = q + 2 - k$. Every pair in $(N(z) \cap X) \times (N(z) \cap Y)$ is missing from the support graph, so $rs \leq \mu$. If $k \leq 1$, then $\ell \geq q + 1$, both r, s are positive, and $rs \geq \ell - 1 \geq q$, a contradiction. Thus $k \geq 2$. If $r, s \geq 2$, then

$$rs \geq 2\ell - 4 = 2q - 2k \geq 2q - 2 - 2Q > \frac{11}{6}(q-1) \geq q,$$

again a contradiction. Thus $\min(r,s) \le 1$.

For every $w \in N(z) \cap H$, the support pairs belonging to $N(w) \cap Z$ form a matching, and neither coordinate of such a pair lies in $N(z) \cap L$. The first assertion follows from the absence of four-cycles through good vertices; the second follows by considering the cycle z, w, u, a, z for $u \in N(w) \cap Z$ and a low neighbour a of u. Since $\min(r,s) \le 1$, one of r, s is at least $q+1-k$, while its support part has at most q vertices. Hence

$$d_Z(w) \le \min\{|X \setminus N(z)|, |Y \setminus N(z)|\} \le k-1. \tag{3.25}$$

Suppose that $k \ge 3$, and choose distinct w_1, w_2, w_3 in $N(z) \cap H$. Put $A_i = N(w_i) \setminus \{z\}$, so $|A_i| = q$. Each common point of A_i, A_j gives a distinct four-cycle z, w_i, x, w_j, z. Thus

$$\sum_{i<j} |A_i \cap A_j| \le Q, \qquad \left|\bigcup_{i=1}^{3} A_i\right| \ge 3q - Q.$$

The union contains at most $3(k-1)$ points of Z, at most $\sum_{i=1}^{3} c_{zw_i} \le 3+\varepsilon_z$ points of $N(z) \cap L$, at most $|L \setminus N(z)| = q-3+k$ other low points, and at most μ full-degree points outside Z. Therefore

$$\mu \ge 2q + 3 - 4k - \varepsilon_z - Q.$$

Combining this with (3.24) gives

$$2q + 3 + 3\mathcal{D} \le 3\beta + 3\varepsilon_z + 5k + Q \le 18Q + 5.$$

Hence $18Q \ge 2q+4$, contrary to $12Q < q-1$. We conclude that $k = 2$.

Joint-codegree cancellation. Write $N(z) \cap H = \{u, v\}$. The union of $N(u) \setminus \{z\}$ and $N(v) \setminus \{z\}$ has size $2q + 1 - c_{uv}$. By (3.25), it contains at most two vertices of Z, at most $c_{zu} + c_{zv}$ vertices of $N(z) \cap L$, at most $q-1$ vertices of $L \setminus N(z)$, and at most μ full-degree vertices outside Z. Comparing these counts gives

$$\mu \ge q - c_{zu} - c_{zv} - c_{uv}. \tag{3.26}$$

To bound the three codegrees together, note that z is a common neighbour of u, v. Every other common neighbour x differs from u, v, z and has $c_{zx} \ge 2$. These contributions occupy different positions of the zth row of R from those indexed by u, v, and consequently

$$(c_{zu} - 1)_+ + (c_{zv} - 1)_+ + c_{uv} - 1 \le \varepsilon_z.$$

Thus $c_{zu} + c_{zv} + c_{uv} \le 3 + \varepsilon_z$. Equation (3.26) gives $\mu \ge q - 3 - \varepsilon_z$. On the other hand, (3.24), $k = 2$, and $\mathcal{D} \ge 2$ imply

$$\mu \le 3\beta + 2\varepsilon_z - 4 \le 12Q - \varepsilon_z - 4.$$

The excess at z cancels, giving $12Q \ge q+1$. This is the final contradiction. □

3.4 Several vertices of degree $q+2$

In this section G has n vertices and m_+ edges, and $\Delta(G) \le q+2$. Let $T = \{v : d(v) = q+2\}$, $t = |T|$, and

$$\alpha = \sum_{v \in L} \delta_v^2.$$

The improvement uses the squared deficit α, rather than the sum of the positive deficits, to count possible common neighbours of pairs in T.

Lemma 3.7. *Let $q \geq 8$ be even. Then*

$$\sum_{v\in L} \delta_v = 2q - 2 + t, \qquad \alpha \leq 2q + 4t - 6 + 8Q. \tag{3.27}$$

If $Q \leq q/4$, every pair of vertices of T has a common neighbour. For such a pair u, v, write

$$C = N(u)\cap N(v), \quad c = |C|, \quad W = \sum_{x\in C} \delta_x, \quad \xi = \varepsilon_u+\varepsilon_v-2(c-1), \quad \Theta = (2-c)q+c+W+\xi.$$

Then $0 \leq \xi \leq 2Q$, and

$$\begin{aligned}(2q-3)W \geq q^2 + (5-3c)q - c^2 + 12c - 24 \\ - (2q-3)\xi - 4Q + \frac{\Theta^2}{n-2}.\end{aligned} \tag{3.28}$$

In particular, if $c = 1$, then

$$W \geq q/2 - 2Q. \tag{3.29}$$

Proof. Write $h = |H|$, $\ell = |L|$, and $k_v = d_T(v)$. The degree sum gives $\sum_v \delta_v = 2q - 2$, while each member of T has deficit -1; this proves the first assertion. For $v \in L$, the local identity gives $u_v \geq q\delta_v - 1 - k_v$. For $v \in H$, parity gives $u_v \geq 1 - \varepsilon_v$: if $\varepsilon_v = 0$, the graph on its odd neighbourhood is a matching with an isolated vertex. We weaken this to $u_v \geq 1 - k_v - \varepsilon_v$. For $v \in T$, use $u_v \geq 0 \geq -k_v$. Since $\sum_v k_v = (q + 2)t$, summation yields

$$\sum_v u_v \geq q(2q-2) - n + 2h - t - \sum_{v\in H} \varepsilon_v.$$

If $P_2 = \sum_v \binom{d(v)}{2}$, then

$$P_2 = \binom{n}{2} - \frac{1}{2}\sum_v u_v + \sum_{\{x,y\}} (c_{xy} - 1)_+ \leq qm_+ - h + \frac{t}{2} + 4Q.$$

Here $\sum_{\{x,y\}}(c_{xy} - 1)_+ \leq 2Q$ and $\sum_{v\in H} \varepsilon_v \leq 4Q$. Expanding $2P_2 - 2qm_+$ gives

$$\alpha + t - (q + 1)(2q - 2) \leq -2h + t + 8Q.$$

Substitute $h = n - \ell - t$ and $\ell \leq 2q - 2 + t$ to obtain (3.27).

For the common-neighbour assertions, put $S = N(u) \cup N(v)$ and $B = V(G) \setminus S$. If $c = 0$, then $|B| = q^2 - q - 4$. For $w \neq u, v$, let $b_w = |N(w) \cap B|$. The excess-codegree bound gives

$$\sum_{w\neq u,v} b_w \geq 2m_+ - 2(q + 2) - 2(n - 2) - 2Q = q^3 - 5q + 2 - 2Q =: s.$$

There are no two-edge paths with endpoints in B and middle vertex u or v. The number of such paths is at most $\binom{|B|}{2} + 2Q$. Jensen's inequality therefore gives

$$(n-2)\binom{s/(n-2)}{2} \leq \binom{q^2-q-4}{2} + 2Q.$$

The required monotonicity holds because $s = (q - 1)(n - 2) - 2q - 2Q \geq (q - 2)(n - 2)$ for $q \geq 8$ and $Q \leq q/4$. Twice the left side minus twice the right side equals

$$q^2 + 5q - 24 - (4q - 2)Q + \frac{4(q + Q)^2}{n-2}.$$

Its polynomial part is at least $11q/2 - 24 > 0$, a contradiction. Hence $c \geq 1$.

For this case, $|B| = q^2 - q - 4 + c$ and

$$|N(w) \cap S| = c_{uw} + c_{vw} - |N(w) \cap C|.$$

The contributions of the pair $\{u, v\}$ show that $0 \leq \xi \leq 2Q$ and

$$\sum_{w \neq u,v} (c_{uw} + c_{vw}) \leq 2(n-2) + \xi, \qquad \sum_{w \neq u,v} |N(w) \cap C| = \sum_{x \in C} d(x) - 2c.$$

Consequently

$$\sum_{w \neq u,v} b_w \geq q^3 - 5q + 2 + \sum_{x \in C} d(x) - 2c - \xi = (q-1)(n-2) - \Theta =: s.$$

The bound $c \leq q + 2$ and $\xi \leq q/2$ gives $s \geq q^3 - 15q/2 - 2 > (q-2)(n-2)$ for $q \geq 8$. The same two-path count gives

$$(n-2)\binom{s/(n-2)}{2} \leq \binom{q^2 - q - 4 + c}{2} + 2Q.$$

Expanding proves (3.28). When $c = 1$, this inequality and $\xi \leq 2Q$ imply

$$W \geq \frac{q}{2} - 2Q + \frac{(7/2)q - 13 - 4Q}{2q - 3} \geq \frac{q}{2} - 2Q,$$

where the last numerator is positive for $q \geq 8$ and $Q \leq q/4$. □

Theorem 3.8. *Let $q \geq 1600$ be even. If G has n vertices, m_+ edges, maximum degree at most $q + 2$, and at least two vertices of degree $q + 2$, then*

$$N_{C_4}(G) > q/12.$$

Proof. Suppose that $Q \leq q/12$. Every pair of vertices in T has a common neighbour by Lemma 3.7. When the common neighbour is unique, its deficit is at least $q/3$, by (3.29). There are at most $9\alpha/q^2$ vertices of deficit at least $q/3$, each of degree at most q and therefore accounting for at most $q^2/2$ pairs in T. At most $2Q$ pairs have two or more common neighbours. It follows that

$$\binom{t}{2} \leq \frac{9}{2}\alpha + 2Q, \qquad t^2 - 37t \leq \frac{73}{3}q.$$

The second inequality follows from (3.27). It implies $t < q/2$: at $t = q/2$, the difference of the two sides is

$$\frac{q^2}{4} - \frac{257}{6}q > 0,$$

and the difference is increasing for $t \geq q/2$. Hence

$$\alpha \leq 2q + 4t - 6 + 8Q < \frac{14}{3}q < 5q. \tag{3.30}$$

Choose any two vertices in T and retain the notation C, c, W of Lemma 3.7. Equation (3.15) gives

$$c \leq 1 + \sqrt{2Q} \leq 1 + \sqrt{q/6} \leq \frac{1}{2}\sqrt{q}.$$

Discarding the nonnegative square in (3.28), and using $\xi \leq 2Q$, gives

$$W \geq \frac{q^2 - 3cq - c^2 - 24 - q/3}{2q - 3} - \frac{q}{6} \geq \frac{q}{4}.$$

For completeness, the last inequality follows from $c \le \sqrt{q}/2$ because the difference of its numerators is at least

$$\frac{q^2}{6} - \frac{3}{2}q^{3/2} + \frac{2}{3}q - 24 \ge 0 \qquad (q \ge 1600).$$

Replacing negative deficits in W by zero and applying Cauchy–Schwarz now yields

$$\alpha \ge \frac{W^2}{c} \ge \frac{1}{8}q^{3/2} \ge 5q,$$

contrary to (3.30). □

Theorem 3.9. *Let $q \ge 8$ be even, and suppose that G has n vertices, m_+ edges, maximum degree at most $q+2$, and at least two vertices of degree $q+2$. Then*

$$N_{C_4}(G) > \frac{q}{4} - 3q^{3/4}.$$

In particular, this degree configuration forces $(1/4 - o(1))q$ four-cycles as $q \to \infty$.

Proof. The inequality is immediate if its right-hand side is negative. Suppose otherwise that $Q \le q/4 - 3q^{3/4}$. Then $q \ge 12^4 = 20736$. Put $d = q/2 - 2Q$, so $d \ge 6q^{3/4}$ and $d^2 \ge q$. For a pair in T with a unique common neighbour, (3.29) gives deficit at least d there. There are at most α/d^2 such vertices, each accounting for at most $q^2/2$ pairs. Therefore

$$t(t-1) \le \frac{\alpha q^2}{d^2} + 4Q \le 4q^2 + 4tq + q.$$

Here we used $\alpha \le 4q + 4t - 6$ from (3.27). At $t = 5q$, the left side minus the final right side is $q^2 - 6q > 0$, and this difference increases for larger t. Thus $t < 5q$ and

$$\alpha < 24q. \tag{3.31}$$

For any pair in T, let C, c, W be as above. Then $c \le 1 + \sqrt{2Q} \le \sqrt{q}$. Equation (3.28) gives

$$\begin{aligned} W &\ge d - \frac{3cq + c^2 + 24 + 4Q}{2q - 3} \\ &\ge d - \frac{3q^{3/2} + 2q + 24}{2q - 3} \ge d - 3\sqrt{q} \ge 5q^{3/4}. \end{aligned}$$

All displayed numerical inequalities hold for $q \ge 20736$; for example, the fraction is at most $2\sqrt{q} + 4/3 + 16/q \le 3\sqrt{q}$. Cauchy–Schwarz gives

$$\alpha \ge \frac{W^2}{c} \ge 25q,$$

contradicting (3.31). □

3.5 The unconditional supersaturation bound

Proof of Theorem 1.1. Let G_0 be a spanning subgraph of G with exactly m_+ edges. Theorems 3.1 and 3.3 handle the cases $\Delta(G_0) \le q+1$ and $\Delta(G_0) \ge q+3$. If $\Delta(G_0) = q+2$, apply Theorem 3.4 when it occurs once and Theorem 3.8 when it occurs at least twice. In every case $12N_{C_4}(G_0) \ge q - 1$. Since $G_0 \subseteq G$, the same bound holds for G. □

4 Degree bounds for near-extremal cores

We derive the degree restrictions for a core at edge deficit b. The concentration estimate below first bounds individual deficits; neighbourhood counts then leave at most one vertex of degree $q+2$. The incidence-completion theorem in Appendix A excludes that last vertex when $q \geq 10^9(b+1)^8$.

We use the notation of Section 2.1 and the parameters in (1.1). For a C_4-free graph on n vertices,

$$u_v = q\delta_v - 1 + a_v. \tag{4.1}$$

When q is even and $d(v) = q+1$, neighbourhood parity gives $u_v \geq 1$ and hence $a_v \geq 2$.

4.1 Deficit concentration below the extremal bound

The following estimate permits vertices of degree $q+2$ and does not require an embedding into a polarity graph.

Lemma 4.1 (Deficit concentration)**.** *Let $q \geq 2$ be even, and let F be a C_4-free graph on q^2+q vertices with $\Delta(F) \leq q+2$. Put $T = \{v : d_F(v) = q+2\}$ and $h = |T|$. For a nonempty set $Z \subseteq L$, write*

$$k = |Z|, \qquad W = \sum_{z\in Z} \delta_z, \qquad R = \sum_{v\in L\setminus Z} \delta_v, \qquad \chi_R = \sum_{v\in L\setminus Z} \delta_v(\delta_v - 1).$$

Then

$$\begin{aligned}(q+k+2)(W-R) + \chi_R &\leq -2q^2 + (k^2+2k-2)q + 2k(k+2) + (k+2)h + e_F(Z,T)\\ &\leq -2q^2 + (k^2+2k-2)q + 2k(k+2) + 2(k+1)h. \end{aligned}\tag{4.2}$$

Proof. Set

$$B = L \cup T \cup \bigcup_{z\in Z} N_F(z), \qquad X = V(F) \setminus B.$$

Every vertex of X has degree $q+1$ and has no neighbour in Z. The graph induced by a neighbourhood in a C_4-free graph is a matching together with isolated vertices. Since $q+1$ is odd, each $v \in X$ has an uncovered partner in $N_F(v)$, and this partner does not belong to Z. Thus

$$|\{z \in Z : N_F(v) \cap N_F(z) = \varnothing\}| \leq u_v - 1 = a_v - 2.$$

On summing over X, the positive-deficit contributions come only from $L \setminus Z$. Discarding the negative contributions from T gives

$$\sum_{v\in X} (a_v - 2) \leq \sum_{w\in L\setminus Z} \delta_w d_F(w) - 2|X| = qR - \chi_R - 2|X|.$$

For each $z \in Z$, at most $|B|-1$ uncovered partners lie in B. Also $a_z \geq -d_T(z)$. Hence

$$-k + qW - e_F(Z,T) \leq \sum_{z\in Z} u_z \leq qR - \chi_R - 2|X| + k(|B|-1).$$

Using $|X| = n - |B|$, we obtain

$$q(W-R) + \chi_R + 2n \leq (k+2)|B| + e_F(Z,T). \tag{4.3}$$

Every positive deficit is an integer, so

$$|B| \leq |L| + h + \sum_{z\in Z} d_F(z) \leq k + R + h + k(q+1) - W = kq + 2k + h - (W-R).$$

Substitution in (4.3), together with $n = q^2+q$, proves the first inequality in (4.2). The second follows from $e_F(Z,T) \leq kh$. □

Corollary 4.2 (An individual deficit bound)**.** *Under the hypotheses of Lemma 4.1, suppose* $e(F) = m_0 - b$*, where* $b \geq 0$*. Every* $z \in L$ *satisfies*

$$2\delta_z \leq 2b + h + 7 + \frac{4h - 15}{q + 3}. \tag{4.4}$$

In particular, if $e(F) = m_0$ *and* $\Delta(F) \leq q + 1$*, then* $d_F(v) \geq q - 2$ *for every vertex. For even* $q \leq 10$*, the latter bound improves to* $d_F(v) \geq q - 1$*.*

Proof. The degree sum gives

$$\sum_v \delta_v = 2q + 2b, \qquad \sum_{v \in L} \delta_v = 2q + 2b + h.$$

Take $Z = \{z\}$ in Lemma 4.1. Then $W - R = 2\delta_z - 2q - 2b - h$. After discarding $\chi_R \geq 0$, we obtain

$$(q + 3)(2\delta_z - 2q - 2b - h) \leq -2q^2 + q + 6 + 4h.$$

The identity $-2q^2 + q + 6 = (-2q + 7)(q + 3) - 15$ yields (4.4). When $b = h = 0$, its right side is strictly less than 7, so the integer δ_z is at most three. For $q \leq 10$ the right side is strictly less than 6, giving $\delta_z \leq 2$. □

4.2 At most one high vertex in a linear edge window

Lemma 4.3 (A near-extremal maximum-degree bound)**.** *Let* $q \geq 8$ *be an integer and* $0 \leq b \leq q/16$*. If* F *is* C_4*-free with* n *vertices and* $m_0 - b$ *edges, then* $\Delta(F) \leq q + 2$*.*

Proof. Suppose a vertex has degree $D \geq q + 3$, and put $a = n - 1$ and $m = m_0 - b$. Count two-edge paths whose endpoints lie outside its neighbourhood. For each other middle vertex, the number of neighbours outside that neighbourhood is at least its degree minus one. Consequently, Jensen's inequality gives

$$a\binom{(2m - D - a)/a}{2} \leq \binom{n - D}{2}. \tag{4.5}$$

Here and below $\binom{x}{2} = x(x - 1)/2$ for real x. The monotonicity step in this use of Jensen is legitimate: for $D \leq n - 1$, the argument's numerator is at least

$$2m - 2(n - 1) = q^3 - 3q + 2 - 2b > (q - 2)(n - 1),$$

because the difference is $q^2 - 2b > 0$.

Let $J(D)$ be the left side of (4.5) minus its right side. It is concave as a function of D, since its quadratic coefficient is $1/(2a) - 1/2 < 0$. At the first endpoint, direct expansion gives

$$2J(q + 3) = q^2 - (4b + 1)q + 2b - 10 + \frac{4(q + b + 1)^2}{q^2 + q - 1} > 0.$$

Indeed, its polynomial part is at least $3q^2/4 - q - 10 > 0$. At $D = n - 1$, the right side of (4.5) is zero, while its left side is positive. Concavity makes J positive on the entire interval, a contradiction. □

Lemma 4.4 (Two high vertices force a large common deficit)**.** *Let* $q \geq 8$ *and* $0 \leq b \leq q/16$*, and let* F *be as in Lemma 4.3. Any two vertices of degree* $q + 2$ *have exactly one common neighbour* z*, and*

$$\delta_z \geq \frac{3q}{8} - 1. \tag{4.6}$$

Proof. Write x, y for the high vertices and $B = V(F) \setminus (N(x) \cup N(y))$. There is no two-edge path with endpoints in B and middle vertex x or y. For $v \neq x, y$, let $r_v = |N(v) \cap B|$.

First suppose $N(x) \cap N(y) = \varnothing$. Then $|B| = q^2 - q - 4$, and the codegree bound gives

$$\sum_{v \neq x,y} r_v \geq 2m - 2(q+2) - 2(n-2) = q^3 - 5q - 2b =: S_0.$$

Also $S_0 > (q-2)(n-2)$, since the difference is $q^2 - q - 2b - 4 > 0$. Thus the path count and Jensen imply

$$(n-2)\binom{S_0/(n-2)}{2} \leq \binom{q^2-q-4}{2}.$$

Twice the left side minus twice the right side is

$$q^2 + (1-4b)q + 6b - 18 + \frac{4(q+b+1)^2}{n-2} > 0.$$

Its polynomial part is at least $3q^2/4 + q - 18 > 0$. This contradiction rules out disjoint neighbourhoods.

Since F is C_4-free, there is now a unique common neighbour z. Put $w = \delta_z$. Inclusion–exclusion gives

$$\begin{aligned}\sum_{v \neq x,y} r_v &\geq 2m - 2(q+2) - 2(n-2) + d(z) - 2 \\ &= q^3 - 4q - 1 - 2b - w \\ &= (q-1)(n-2) - (q+3+2b+w) =: S_1.\end{aligned}$$

Because $d(z) \geq 2$, one has $w \leq q-1$ and $S_1 \geq S_0$. The same Jensen argument, with $|B| = q^2 - q - 3$, yields

$$(2q-3)w \geq q^2 - (4b+2)q + 6b - 7 + \frac{(q+3+2b+w)^2}{n-2}. \tag{4.7}$$

For $q \geq 8$ and $b \leq q/16$, the polynomial on the right is at least $3q^2/4 - 13q/8 - 7$. Subtracting $(2q-3)(3q/8-1)$ from this lower bound gives $3q/2 - 10 > 0$. Discarding the square term proves (4.6). □

Lemma 4.5 (Signed deficit-square budget)**.** *Let q be even and suppose F is C_4-free with n vertices, $m_0 - b$ edges, and maximum degree at most $q+2$. Let T, H, L be the vertices of degree $q+2$, degree $q+1$, and degree at most q, respectively. Put $h = |T|$ and $\alpha = \sum_{v \in L} \delta_v^2$. Then*

$$\sum_{v \in L} \delta_v = 2q + 2b + h, \qquad \alpha \leq 2bq + 6b + 4q + 4h. \tag{4.8}$$

Proof. Let $D = \sum_v \delta_v = 2q + 2b$ and $k_v = |N(v) \cap T|$. Equation (4.1) gives $u_v \geq q\delta_v - 1 - k_v$ on L. On H, neighbourhood parity gives $u_v \geq 1 \geq 1 - k_v$; on T, use $u_v \geq 0 \geq -k_v$. Since $\sum_v k_v = (q+2)h$, summation gives

$$\sum_v u_v \geq qD - n + 2|H| - h.$$

On the other hand, summing the exact local identity gives

$$\sum_v u_v = (2q+1)D - n - (\alpha + h),$$

because $\sum_v a_v = \sum_v \delta_v d(v) = (q+1)D - \sum_v \delta_v^2$. Comparison yields $\alpha \le (q+1)D - 2|H|$. Finally $|H| = n - h - |L|$ and $|L| \le \sum_L \delta_v = D + h$, so

$$\alpha \le (q+3)D - 2n + 4h = 2bq + 6b + 4q + 4h.$$

The first identity in (4.8) follows directly from the negative deficit -1 at each member of T. □

Theorem 4.6 (One-edge degree reduction in a linear window)**.** *Let $q \ge 2048$ be even, and let $0 \le b \le q/16$ be an integer. If F is C_4-free with n vertices and $m_0 - b$ edges, then*

$$\Delta(F) \le q+2, \qquad |\{v : d(v) = q+2\}| \le 1.$$

In particular, deleting at most one edge gives a spanning C_4-free graph of maximum degree at most $q+1$ and with at least $m_0 - b - 1$ edges. Every vertex of F has degree at least $q - b - 2$.

Proof. The maximum-degree assertion is Lemma 4.3. Suppose $h \ge 2$. Since $q \ge 24$, every pair of high vertices has its unique common neighbour in $L_* := \{v : \delta_v \ge q/3\}$. The positive deficit sum bounds $|L_*| \le 3(2q + 2b + h)/q$. Every vertex of L_* has degree at most $2q/3 + 1 \le q$, so it serves as the common neighbour of at most $q^2/2$ high pairs. Therefore

$$\binom{h}{2} \le \frac{3}{2} q(2q + 2b + h) \le \frac{51}{16} q^2 + \frac{3}{2} hq.$$

This implies $h < 5q$: at $h = 5q$ the left side minus the final right side is $29q^2/16 - 5q/2 > 0$, and that difference increases for $h \ge 5q$. Lemma 4.5 consequently gives

$$\alpha < \frac{1}{8} q^2 + \frac{195}{8} q.$$

Choose any high pair and its common neighbour z. Equation (4.6) gives

$$\alpha \ge \delta_z^2 \ge \left(\frac{3q}{8} - 1\right)^2.$$

The last lower bound minus the preceding strict upper bound is

$$\frac{q^2}{64} - \frac{201q}{8} + 1 > 0 \qquad (q \ge 2048),$$

a contradiction.

The edge-deletion assertion is immediate. For the minimum degree, Corollary 4.2 gives

$$2\delta_v \le 2b + h + 7 + \frac{4h - 15}{q+3}.$$

Since $h \le 1$, this is strictly less than $2b + 8$. The integer δ_v is therefore at most $b + 3$. □

4.3 Excluding the last high vertex for a fixed deficit

Theorem 4.7 (Automatic degree cap for a fixed edge deficit)**.** *For every fixed nonnegative integer b, there is $q_0(b)$ such that the following holds for every even $q \ge q_0(b)$. If F is C_4-free with $q^2 + q$ vertices and $m_0 - b$ edges, then*

$$\Delta(F) \le q + 1. \tag{4.9}$$

One sufficient explicit choice, using the numerical sufficient conditions in the fixed-parameter incidence-completion theorem, is

$$q_0(b) = 10^9 (b+1)^8. \tag{4.10}$$

Proof. Put $B = b + 1$ and suppose $q \geq 10^9 B^8$. Then $q \geq 2048$ and $b \leq q/16$, so Theorem 4.6 applies. The incidence-completion inequalities are checked at the end of the proof with this same explicit bound. Suppose its one remaining exceptional vertex y exists, with $d(y) = q + 2$. Let H be the degree-$(q + 1)$ vertices, L the positive-deficit vertices, and put

$$\sigma = \sum_{v \in L} (\delta_v - 1), \qquad \tau = \sum_{v \in N(y) \cap L} (\delta_v - 1), \qquad k = |N(y) \cap H|.$$

The degree sum gives

$$\sum_{v \in L} \delta_v = 2q + 2b + 1, \qquad |H| = q^2 - q - 2 - 2b + \sigma.$$

Because y is the only high vertex, its neighbourhood consists of k full vertices and $q + 2 - k$ low vertices. Its local identity is

$$u_y = -q - 1 + a_y = -q - 1 + (q + 2 - k + \tau) = 1 - k + \tau. \tag{4.11}$$

Thus $k \leq 1 + \tau \leq 1 + \sigma$.

Delete y and call the resulting graph F_0. It has $N - 2$ vertices. Every full vertex outside $N(y)$ remains of degree $q + 1$, so F_0 has at least

$$|H| - k \geq q^2 - q - 3 - 2b = N - 2(q + 1) + \binom{2}{2} - (2b + 3) \tag{4.12}$$

full-degree vertices. Choose exactly this many and call their set C.

Consider the incidence system with point set C and indexed lines

$$S_v = N_{F_0}(v) \cap C \qquad (v \in V(F_0)).$$

Each point is on exactly $q + 1$ indexed lines, and two different indexed lines meet in at most one point by C_4-freeness. Its parameters are those of the fixed-parameter completion theorem with

$$s = 2, \qquad t = 2b + 3, \qquad K = s + t = 2b + 5. \tag{4.13}$$

For completeness, the required uncovered-degree bound follows automatically. The shortest-line packing lemma gives $|S_v| \geq q + 1 - K$ for every v. For this paragraph, recompute the deficits in the smaller graph: $\delta_x^{(0)} = q + 1 - d_{F_0}(x)$ and $a_x^{(0)} = \sum_{v \in N_{F_0}(x)} \delta_v^{(0)}$. The line-size bound gives $\delta_x^{(0)} \leq K$. Every $x \in C$ has at most K neighbours outside C, while all neighbours in C have deficit zero. Its weighted neighbour deficit is at most K^2. The exact local identity at the order $N - 2$ is therefore

$$u_x^{(0)} = a_x^{(0)} - 2 \leq K^2 - 2,$$

where $u_x^{(0)}$ counts uncovered partners in F_0. Thus the uncovered-pair graph on C has maximum degree at most $D = K^2$. The fixed-parameter incidence-completion theorem embeds all the indexed lines, with their exact incidences on C, in a projective plane Π of order q. Denote their distinct image lines by ℓ_v.

There are exactly two plane lines not among the indexed lines. Every point of C already lies on its full pencil of $q + 1$ indexed lines, so no point of C lies on either missing line. Their union has $2q + 1$ points, whereas

$$|V(\Pi) \setminus C| = 2q + 4 + 2b.$$

There are therefore exactly $t = 2b + 3$ further points outside C and outside the union of the two missing lines. An indexed line meets that union in at most two points and hence contains at most $t + 2 = K$ points outside C.

Now take the $q+2$ distinct indexed lines

$$\mathcal{F} = \{\ell_v : v \in N_F(y)\}.$$

They intersect pairwise outside C. Indeed, a point $x \in C$ on both ℓ_v and ℓ_w would give the four-cycle y, v, x, w in F. Lemma 2.5 makes all $q+2$ lines concurrent once $q+1 > K(K-1)$. But a point of a plane of order q lies on exactly $q+1$ lines. This contradiction excludes y and proves the fixed-b assertion.

We finally verify the explicit sufficient choice (4.10). For the constants in the proof of the fixed-parameter incidence-completion theorem, the parameters in (4.13) satisfy

$$K \le 5B, \quad A_0 \le 12B, \quad B_0 \le 48B^2, \quad Y_0 \le 96B^2,$$
$$\mu = D = K^2 \le 25B^2, \quad C_1 \le 198B^3, \quad a_0 \le 106B^2.$$

The packing constant is $C_{2,t} = t(t-1) \le 9B^2$. The bad-pair deletion threshold is at most $352800B^7$. The two Metsch thresholds are at most $952B^3$ and $1938B^4$, and the extra exact-clique threshold is at most $2388B^4$. The restoration-of-indices condition follows from $q > 9221B^4$. Finally, for the Beutelspacher–Metsch polynomial $f(a) = 6a^4 + 9a^3 + 19a^2 + 8a$,

$$f(a_0) \le f(106)B^8 = 768419652B^8 < 4 \cdot 10^9 B^8.$$

The remaining inequalities in (A.4)–(A.11), including the shortest-line and clique-size bounds, hold for $q \ge 10^9 B^8$. This range also gives $q+1 > K(K-1)$, as required for the final pencil argument. □

5 Polarity stability at a prescribed edge deficit

We prove Theorem 1.3. Put $B = b+1$. The degree bounds of Section 4 reduce the argument to nonnegative deficits. A dichotomy for their total excess and an exact second-moment identity produce a bounded set of vertices whose deletion permits incidence completion. We then restore those vertices and their incident edges inside the resulting polarity graph.

5.1 A dichotomy for the total excess deficit

Throughout the preliminary argument let $b \ge 0$ be an integer, suppose q is even with $q \ge 100B^2$, and let F be a C_4-free graph satisfying

$$|V(F)| = N-1, \qquad e(F) = m_0 - b, \qquad q-b-2 \le d_F(v) \le q+1. \tag{5.1}$$

Set

$$\delta_v = q+1-d_F(v), \quad H = \{v : \delta_v = 0\}, \quad L = \{v : \delta_v > 0\},$$
$$r_v = |N_F(v) \cap L|, \qquad a_v = \sum_{w \in N_F(v)} \delta_w.$$

Write u_v for the number of other vertices having no common neighbour with v, and u_{HH} for the number of uncovered unordered pairs inside H. Define

$$\sigma = \sum_{v \in L} (\delta_v - 1), \quad \chi = \sum_{v \in L} \delta_v(\delta_v - 1), \quad \Omega = \sum_{v \in L} \delta_v r_v, \quad \eta = \sum_{v \in H} (a_v - 2).$$

The degree sum, the local two-path identity, and weighted counting between L and H give

$$\sum_L \delta_v = 2q + 2b, \qquad |L| = 2q + 2b - \sigma, \qquad |H| = q^2 - q - 2b + \sigma, \tag{5.2}$$

$$u_v = q\delta_v - 1 + a_v, \tag{5.3}$$

$$\eta + \chi + 2\sigma + \Omega = 2Bq + 4b. \tag{5.4}$$

Indeed, $\sum_H a_v = q(2q+2b) - \chi - \Omega$, and subtracting $2|H|$ proves the budget. Every full vertex has $a_v \geq 2$: its odd neighbourhood induces a matching and isolated vertices, so at least one neighbour is uncovered from it. In particular all terms on the left side of (5.4) are nonnegative.

A full vertex is good if $a_v = 2$ and bad otherwise. A good vertex has exactly one uncovered partner $f(v)$, and $vf(v)$ is an edge. For good vertices v, w, the common-neighbour matrix identity yields

$$A_{v,f(w)} = A_{f(v),w}. \tag{5.5}$$

To see this directly, write U for the uncovered-pair adjacency matrix. Since F is C_4-free, $A^2 = qI + J - U - \mathrm{diag}(\delta)$; commuting with A gives $[A, U]_{vw} = (\delta_w - \delta_v)(1 - A_{vw})$. At two good vertices, the corresponding rows of U are unit vectors, giving (5.5).

Lemma 5.1 (Bounded exceptional capacity)**.** *Let g be the number of good vertices whose uncovered partner is good, and let*

$$G_0 = \{v \in H : v \text{ is good and } f(v) \in L\}, \qquad p = |H \setminus G_0|.$$

Then

$$g \leq 2B, \qquad p \leq 2\eta + 2B, \qquad u_{HH} \leq 2\eta + B. \tag{5.6}$$

If $t = |\{v : \delta_v = 2\}|$ and $A_* = \sum_{\delta_v \geq 3} \delta_v$, then

$$|G_0| \leq (q + b - t - A_*/2)^2 + t(q-1). \tag{5.7}$$

Proof. If good vertices u, v are uncovered from one another, each is the other's partner. If another good vertex w were adjacent to u, equation (5.5), applied to v, w, would give $vf(w) \in E(F)$. The vertices $u, w, f(w), v$ would form a four-cycle. They are distinct because $f(w) \neq w$ and uniqueness of the partners at u, v excludes $f(w) \in \{u, v\}$. Since $a_u = 2$, at most two neighbours of u are low, and it has at least $q - 2$ bad full neighbours. If $2B + 1$ good vertices had good partners, their bad-neighbour sets would have union of size at least

$$(2B+1)(q-2) - \binom{2B+1}{2} > 2Bq + 4b \geq \eta.$$

Here intersections of any two neighbourhoods have size at most one, and the strict inequality follows from $q > 2B^2 + 9B - 2$. The number of bad full vertices is at most η, so this is impossible and $g \leq 2B$.

For a bad full vertex write $s_v = a_v - 2 \geq 1$; then $u_v = s_v + 1$. The number of its adjacent uncovered partners is odd. It is at most s_v if s_v is odd and at most $s_v + 1$ if s_v is even. Let B_0 be the number of bad vertices and B_{even} the number with positive even s_v. Good vertices outside G_0 either have a bad partner or are among the g vertices just counted. Therefore

$$p \leq B_0 + \eta + B_{\mathrm{even}} + g \leq 2\eta + 2B,$$

because $B_0 + B_{\mathrm{even}} \leq \sum s_v = \eta$. Every uncovered full–full pair either touches a bad vertex or is one of the $g/2$ good–good pairs. Thus $u_{HH} \leq \eta + B_0 + g/2 \leq 2\eta + B$.

Put $X = f(G_0)$. A vertex $v \in G_0$ cannot have a second neighbour in X: applying (5.5) to v and a vertex partnered to that neighbour would produce a four-cycle. If $\delta_{f(v)} = 1$, the two low neighbours of v are unit-deficit vertices, one in X and one outside X. Different such full vertices give different pairs. There are $2q + 2b - 2t - A_*$ unit-deficit vertices, so their number is at most $(q + b - t - A_*/2)^2$. If $\delta_{f(v)} = 2$, its partner is the sole low neighbour of v and supports at most $q - 1$ vertices. No deficit at least three can neighbour a good full vertex. This proves (5.7). □

Lemma 5.2 (Small or nearly full excess deficit)**.** *Put $C = 8b + 10$. Under (5.1), either*

$$\sigma < C \qquad \text{or} \qquad \sigma > q - C. \tag{5.8}$$

In the second case, F has at least $q^2 - 10B$ full-degree vertices.

Proof. The definitions imply $t + A_* \geq \sigma$ and $\sigma/2 \leq t + A_*/2 \leq \sigma$. For $\sigma \geq 2b$, expansion of (5.7) gives

$$\begin{aligned} p &\geq (t + A_* - 2b - 1)q - (t + A_*/2 - b)^2 - 2b + \sigma + t \\ &\geq (\sigma - 2b - 1)q - (\sigma - b)^2 - 2b. \end{aligned}$$

Use $p \leq 2\eta + 2B \leq 4Bq + 8b + 2B$ to obtain

$$(\sigma - 6b - 5)q - (\sigma - b)^2 \leq 12b + 2. \tag{5.9}$$

The left side is concave as a function of σ. At the endpoints of $[C, q - C]$ it equals, respectively,

$$(2b + 5)q - (7b + 10)^2, \qquad (4b + 5)q - (9b + 10)^2.$$

Both exceed $12b + 2$ for $q \geq 100B^2$: each is at least $5q - 100B^2 \geq 400B^2$. Also $C \geq 2b$ and $q \geq 2C$. Thus this entire interval is excluded. In the second alternative, (5.2) gives $|H| > q^2 - 2b - C = q^2 - 10B$. □

5.2 Signed moments and a bounded deletion set

Define

$$w_v = \begin{cases} r_v + \delta_v - 1, & v \in L, \\ r_v - 2, & v \in H, \end{cases} \qquad T = \sum_{v \in L} (\delta_v - 1) r_v.$$

For every vertex,

$$d_H(v) = q - w_v - \mathbf{1}_H(v). \tag{5.10}$$

Lemma 5.3 (General signed moments)**.** *Under (5.1),*

$$\sum_v w_v = 2Bq + 4b - (q + 2)\sigma, \tag{5.11}$$

$$\begin{aligned} \sum_v w_v^2 &= q^2 + (6b + 1 - \sigma)q + (2b - \sigma)^2 + 2b + \sigma \\ &\quad - 2\chi - 2\eta - 2T - 2u_{HH}. \end{aligned} \tag{5.12}$$

If $\sigma < C = 8b + 10$ and $w_v^+ = \max\{w_v, 0\}$, then

$$\sum_v w_v^+ \leq (2B + 1)q, \qquad \sum_v (w_v^+)^2 \geq q^2 - 32Bq. \tag{5.13}$$

Proof. Write $\ell = |L|$, $h = |H|$, and $e_L = e(F[L])$. Since $\sum_v r_v = \sum_L d(v) = q\ell - \sigma$, summing the definitions of w_v gives $\sum w_v = q\ell - 2h$, which is (5.11).

Let u_{LL} and u_{LH} count uncovered pairs of the indicated types. Summing (5.3) over the two layers gives

$$2u_{LL} + u_{LH} = q(2q + 2b) - \ell + \Omega, \qquad 2u_{HH} + u_{LH} = h + \eta.$$

Hence

$$2u_{LL} = q^2 + (2b - 1)q + \Omega - \eta + 2u_{HH}. \tag{5.14}$$

Counting covered pairs in L by their unique common neighbour yields

$$\sum_v r_v^2 = q\ell - \sigma + \ell(\ell-1) - 2u_{LL}.$$

Now $\sum_L(\delta_v-1)^2 = \chi - \sigma$, $\sum_H r_v = q\ell - \sigma - 2e_L$, and $\Omega = 2e_L + T$. Replacing r_v by its defining weight therefore gives

$$\begin{aligned}\sum_v w_v^2 &= q\ell - \sigma + \ell(\ell-1) - 2u_{LL} + 2T + \chi - \sigma - 4(q\ell - \sigma - 2e_L) + 4h \\ &= \ell(\ell-1) - 3q\ell + 2\sigma + \chi + 4\Omega - 2T + 4h - 2u_{LL}.\end{aligned}$$

Substitute (5.14), (5.4), and (5.2). Direct expansion gives (5.12).

Every low weight is nonnegative. A full vertex has $r_v \geq 1$, so the only negative value is -1. Such a vertex has just one low neighbour, of deficit at least two. If a_- counts the negative weights, then

$$a_- \leq \sum_{\delta_v \geq 2} d_H(v) \leq \sigma q.$$

Thus $\sum w_v^+ \leq 2Bq + 4b \leq (2B+1)q$. Also $\sum (w_v^+)^2 = \sum w_v^2 - a_-$. Use $\delta_v \leq b+3 = B+2$, $\chi \leq (B+2)\sigma$, $T \leq \Omega$, and $u_{HH} \leq 2\eta + B$ in the second moment. Discarding its nonnegative constant terms gives

$$\begin{aligned}\sum_v (w_v^+)^2 &\geq q^2 - (6b + 11 + 2\sigma)q - 2(B+2)\sigma - 24b - 2B \\ &\geq q^2 - 31Bq - 78B^2 \geq q^2 - 32Bq.\end{aligned}$$

The first inequality uses $6\eta + 2T \leq 6(\eta + \Omega) \leq 12Bq + 24b$. The last two use $\sigma < C = 8B+2$ and $q \geq 100B^2$. This proves (5.13). □

Lemma 5.4 (The largest k weights)**.** *Let $x_1 \geq x_2 \geq \cdots \geq 0$, let $k \geq 1$, and suppose $\sum_i x_i \leq kq$. Put $W = \sum_{i=1}^k x_i$, adjoining zero terms if necessary. If $W \leq q$, then*

$$\sum_i x_i^2 \leq qW.$$

Proof. Put $a = x_k$. For fixed a, W, convexity bounds the sum of the first k squares by $(W-(k-1)a)^2 + (k-1)a^2$. The remaining terms are at most a, so their squared sum is at most $a(kq - W)$. The resulting bound is the convex quadratic

$$W^2 + \{kq - (2k-1)W\}a + k(k-1)a^2, \qquad 0 \leq a \leq W/k.$$

At its endpoints the values are W^2 and qW. Since $W \leq q$, both are at most qW. The same calculation, or the bound $\sum x_i^2 \leq x_1 \sum x_i$, covers $k = 1$. □

Proposition 5.5 (Concentration in a bounded vertex set)**.** *If (5.1) holds and $\sigma < C$, put $k = 2B+1$. There is a set S of k vertices such that*

$$\sum_{v \in S} w_v \geq q - 32B.$$

In $F_0 = F - S$, at least

$$q^2 - kq - (34B-2) = N - (k+1)(q+1) + \binom{k+1}{2} - t_0, \qquad t_0 = 2B^2 + 35B - 2, \quad (5.15)$$

vertices have degree $q+1$ and were adjacent to no vertex of S.

Proof. Apply Lemma 5.4 and (5.13). If the largest-k sum is at most q, it is at least $q - 32B$; otherwise this conclusion is immediate. The chosen positive-part weights can be realized by distinct vertices of nonnegative actual weight. Indeed, $\sigma < 8b + 10$ and $q \geq 100(b+1)^2$ imply

$$|L| = 2q + 2b - \sigma > 2q - 6b - 10 > 2b + 3 = k.$$

The low vertices all have nonnegative weight and can replace any zero-weight choices at vertices of negative actual weight.

Every vertex in $H \setminus (S \cup N_F(S))$ remains full in F_0. Using (5.10), their number is at least

$$|H| - |S \cap H| - \sum_{v \in S} d_H(v) = |H| - kq + \sum_{v \in S} w_v \geq q^2 - kq - 2b + \sigma - 32B.$$

Discarding $\sigma \geq 0$ gives (5.15). □

5.3 Completion and reinsertion

Lemma 5.6 (A polarity with few extra absolute points)**.** *Let P be the graph of a polarity of a plane of even order q. If $e(P) \geq M - b$ and $q > 4b^2$, then its absolute points form a line of exactly $q + 1$ points, and $e(P) = M$.*

Proof. The elementary even-order polarity lemma shows that its absolute set $\mathcal{A}$ is a blocking set, that distinct absolute points are nonadjacent, and that $|\mathcal{A}| \geq q + 1$. The exact edge formula gives $|\mathcal{A}| \leq q + 1 + 2b$. Write $|\mathcal{A}| = q + d$, so $1 \leq d \leq 2b + 1$. We recall the standard counting argument for the blocking-set bound of Bruen [3].

Suppose no whole line lies in $\mathcal{A}$, and put $r_\ell = |\ell \cap \mathcal{A}|$. Choose a point of ℓ outside $\mathcal{A}$. The other q lines through that point each require a different point of $\mathcal{A}$ outside ℓ, so $1 \leq r_\ell \leq d$. Counting incidences and ordered collinear pairs gives

$$\sum_\ell r_\ell = (q + d)(q + 1), \qquad \sum_\ell r_\ell(r_\ell - 1) = (q + d)(q + d - 1).$$

Consequently

$$0 \geq \sum_\ell (r_\ell - 1)(r_\ell - d) = q\{q - (d-1)^2\} > 0,$$

a contradiction. Thus $\mathcal{A}$ contains a line. There is no additional absolute point x outside it: the polar line of x would meet this absolute line at another absolute point, making two distinct absolute points adjacent. It follows that $|\mathcal{A}| = q+1$, and the edge formula gives $e(P) = M$. □

Lemma 5.7 (Uniform completion bounds)**.** *Let $B \geq 1$ be an integer and suppose $q \geq 10^{18}B^{16}$. For either parameter choice*

$$(s, t) = (1, 10B) \quad \text{or} \quad (s, t) = (2B + 2,\ 2B^2 + 35B - 2), \tag{5.16}$$

put $K = s + t$ and $D = K^2$. All numerical conditions (A.4)–(A.11) for Theorem A.2 hold. The polarity-extension condition $q > (K + 2)^2$ in Lemma B.2 also holds. Moreover,

$$q \geq 10^9 B^8, \qquad q \geq 100B^2, \qquad q \geq 2048, \qquad B - 1 \leq q/16, \tag{5.17}$$

$$q > 4(B-1)^2, \qquad q - 3B - 2 > K(K-1), \qquad q - 6B - 4 > 2B^2 + 6B + 1. \tag{5.18}$$

Proof. Both choices in (5.16) satisfy

$$s \leq 4B, \quad t \leq 35B^2, \quad K \leq 39B^2, \quad |\beta| \leq 34B,$$

where $\beta = (s-1)(s-2)/2 - t$. The constants in the proof of Theorem A.2 therefore satisfy

$$A_0 \le 80B^2, \quad B_0 \le 2880B^4, \quad Y_0 \le 5760B^4,$$
$$\mu = D \le 1521B^4, \quad C_1 \le 65100B^6, \quad a_0 \le 7129B^4.$$

In particular, $s^4 + 2s^3 + 1 \le 385B^4$ and $|C_{s,t}| \le 1492B^3$. Substitution gives the following bounds for the sufficient thresholds in the incidence-completion proof.

Condition	Upper bound on its threshold for q
(A.4)	$385B^4$ and $1492B^3$
(A.5)	$6 \cdot 10^{11} B^{14}$
(A.6)	$3 \cdot 10^5 B^6$
(A.7)	$3 \cdot 10^6 B^8$
(A.8)	$8000B^4$
(A.9)	$6 \cdot 10^6 B^8$
(A.10)	$4 \cdot 10^7 B^8$
(A.11)	$4 \cdot 10^{15} B^{16}$

For the last row, with $f(a) = 6a^4 + 9a^3 + 19a^2 + 8a$,

$$f(a_0) \le f(7129)B^{16} < 1.6 \cdot 10^{16} B^{16}.$$

Every threshold in the table is strictly below $10^{18}B^{16}$. Finally, $K \le 39B^2$ gives $(K+2)^2 \le 1681B^4$ and $K(K-1) + 3B + 2 \le 1526B^4$. The degree, blocking-set, and remaining reinsertion conditions in (5.17) and (5.18) follow directly from $B \ge 1$ and $q \ge 10^{18}B^{16}$. □

Proof of Theorem 1.3. Write $B = b + 1$ and $C = 8b + 10$.

Degree reduction. The numerical conditions used below are supplied by Lemma 5.7. The automatic degree cap theorem applies because $q \ge 10^9 B^8$. The minimum-degree consequence of the linear-window degree theorem gives $d_F(v) \ge q - b - 2$. Thus (5.1) holds.

Large excess. Suppose $\sigma > q - C$. Lemma 5.2 gives at least $q^2 - 10B$ full vertices. Apply intrinsic recognition (Theorem B.3) with $s = 1$, $t = 10B$. Its conclusion is exactly $F \subseteq P - p$. The first parameter choice in (5.16) applies.

A polarity after deletion. Suppose $\sigma < C$. Choose S, F_0 as in Proposition 5.5. Intrinsic recognition applies using a prescribed clean full set C_0 of the size in (5.15), with parameters

$$s = k + 1 = 2B + 2, \qquad t = t_0 = 2B^2 + 35B - 2, \qquad K = s + t = 2B^2 + 37B. \tag{5.19}$$

Its construction supplies a polarity graph P and a set Q of s omitted points with $F_0 \subseteq P - Q$. It also supplies exact incidences at C_0 and

$$|N_{F_0}(x) \cap C_0| \ge q + 1 - K \qquad (x \in V(F_0)). \tag{5.20}$$

Every old polar line has at most K points outside C_0. Indeed, the s missing lines contain no point of C_0, and their union has at least $s(q+1) - \binom{s}{2}$ points. At most t further points lie outside C_0. An old polar line meets the missing-line union in at most s points, giving the claimed bound.

Restoring vertices and their old neighbours. For each $v \in S$, put $B_v = N_F(v) \setminus S$. These old vertices avoid C_0, and

$$|B_v| \geq q - b - 2 - (k-1) = q - 3B - 1. \tag{5.21}$$

The lines $\{\pi(x) : x \in B_v\}$ meet pairwise outside C_0; an intersection in C_0 would give a four-cycle in F. The bounded-pencil lemma makes them concurrent at a point p_v outside C_0. This centre cannot be an old point x: (5.20) would supply at least $q + 1 - K$ further lines through x, indexed by points of C_0, disjoint from the line indices in B_v. The combined number is at least $(q - 3B - 1) + (q + 1 - K) = 2q - 3B - K > q + 1$, by (5.18). Thus $p_v \in Q$. If two distinct deleted vertices had the same centre, their line families would have at most one common member by C_4-freeness, but would together have at least $2(q - 3B - 1) - 1 > q + 1$ members. The points p_v are therefore distinct. These assignments respect every old–new edge.

Restoring edges between the deleted vertices. By the elementary even-order polarity bound (Lemma 2.6), $e(P) \leq M$ and all degrees of P are at least q. Hence

$$e(P - Q) \leq M - sq + \binom{s}{2}.$$

On the other hand,

$$e(F_0) = M - (k+1)q - k - b + \sum_{v \in S} \delta_v + e(F[S]).$$

Consequently the number D_0 of old host edges absent from F_0 satisfies

$$D_0 \leq k + b + \binom{k+1}{2} = 2B^2 + 6B + 1. \tag{5.22}$$

Suppose $uv \in E(F[S])$ but p_u, p_v are nonadjacent in P. Write $\ell = \pi(p_u)$ and $j = \pi(p_v)$. The rule $x \mapsto \pi(x) \cap j$ is a bijection from the points of ℓ to those of j. It is well defined because $p_v \notin \ell$; if distinct points had the same image y, the line $\pi(y)$ would contain both points and hence equal ℓ, forcing $y = p_u \notin j$, a contradiction. Since $B_u \subseteq \ell$ and $B_v \subseteq j$, at least $|B_u| + |B_v| - (q+1)$ of the matched pairs lie in $B_u \times B_v$. At most one is a loop. No nonloop edge occurs in both orientations, because $|B_u \cap B_v| \leq 1$. Thus there are at least

$$|B_u| + |B_v| - q - 2 \geq q - 6B - 4$$

distinct old host edges between these two sets. All are absent from F_0, since an edge present there would make a four-cycle with uv in F. This contradicts (5.22). All edges inside S are respected as well. Since $|Q| = |S| + 1$, we have obtained a spanning embedding $F \subseteq P - p$.

Orthogonality and the exact deficit. For either branch, $e(P) \geq e(F) + d_P(p) \geq M - b$. Lemma 5.6 therefore makes the host orthogonal. If p is absolute, $e(P - p) = M - q = m_0$; otherwise $e(P - p) = M - q - 1 = m_0 - 1$. The asserted numbers of deleted edges follow by subtracting $e(F) = m_0 - b$.

□

Remark 5.8 (The quantitative range)**.** The bound is uniform in b and q subject to $q \geq 10^{18}(b+1)^{16}$. In particular, it includes every fixed edge deficit as q tends to infinity. The linear deficit range b proportional to q remains beyond this estimate.

6 Degree deficits at the extremal bound

We now refine the structure of cores at the edge count m_0. The local estimates in this section give a smaller threshold for extremal recognition, unique completion at zero surplus, and surplus gaps at smaller orders. We use the deficit notation from Section 2.1. For a C_4-free graph on n vertices, Lemma 2.1 gives

$$u_v = q\delta_v - 1 + a_v. \tag{6.1}$$

Throughout, $H = \{v : d_F(v) = q+1\}$ and $L = \{v : d_F(v) \le q\}$.

6.1 The weighted surplus and uncovered partners

From now until the end of this section, unless otherwise stated, $q \ge 4$ is even and F satisfies

$$|V(F)| = q^2 + q, \qquad e(F) = m_0, \qquad F \text{ is } C_4\text{-free}, \qquad \Delta(F) \le q+1. \tag{6.2}$$

Thus $V(F) = H \sqcup L$, all deficits are nonnegative, and

$$\sum_{v\in L} \delta_v = 2q.$$

For $v \in H$, put $s_v = a_v - 2$, and define

$$\eta = \sum_{v\in H} s_v, \qquad \omega = \sum_{v\in L} \delta_v d_L(v) = \sum_{uv\in E(F[L])} (\delta_u + \delta_v). \tag{6.3}$$

Lemma 2.2 gives $s_v \ge 0$. As in that lemma, a vertex of H is good when $s_v = 0$. Every good vertex v has a unique uncovered partner $f(v)$, and $vf(v) \in E(F)$. A vertex of H that is not good is called bad. Write $\mathcal{B}$ for the bad vertex set and $B_0 = |\mathcal{B}|$.

Lemma 6.1 (The weighted deficit identity)**.** *Under* (6.2)*,*

$$\eta + \sum_{v\in L} (\delta_v - 1)(\delta_v + 2) + \omega = 2q. \tag{6.4}$$

Proof. Put

$$\sigma = \sum_{v\in L} (\delta_v - 1), \qquad \chi = \sum_{v\in L} \delta_v(\delta_v - 1).$$

Then $|L| = 2q - \sigma$ and $|H| = q^2 - q + \sigma$. Weighted counting across L, H gives

$$\sum_{v\in H} a_v = \sum_{w\in L} \delta_w(q + 1 - \delta_w - d_L(w)) = 2q^2 - \chi - \omega.$$

Subtracting $2|H|$ gives $\eta = 2q - 2\sigma - \chi - \omega$, which is (6.4). □

The good-vertex commutator (2.8) has a useful immediate consequence. If u, v are good and $f(u) = v$, then v is the only good neighbour of u. Since $a_u = 2$, at most two neighbours of u lie in L. Consequently

$$\text{a good–good uncovered pair implies } |N_F(u) \cap \mathcal{B}| \ge q - 2 \text{ and } \eta \ge q - 2. \tag{6.5}$$

Suppose now that there is no uncovered pair consisting of two good vertices. Define

$$G_0 = \{v \in H : v \text{ is good and } f(v) \in L\}, \qquad D = H \setminus G_0, \qquad p = |D|,$$

and put $X = f(G_0) \subseteq L$. We call the vertices of X active. For $a \in X$, let $Z_a = \{v \in G_0 : f(v) = a\}$.

Lemma 6.2 (Exceptional vertices and support pairs)**.** *With the preceding notation,*

$$p \le 2\eta. \tag{6.6}$$

Every vertex $v \in G_0$ has $f(v)$ as its only neighbour in X. If $\delta_{f(v)} = 1$, then v has exactly one further low neighbour, of deficit one, outside X. No two such vertices have the same pair of low neighbours. If $\delta_{f(v)} = 2$, then $f(v)$ is the only low neighbour of v.

In particular, writing $\ell_i = |\{v : \delta_v = i\}|$,

$$|G_0| \le \left\lfloor \frac{\ell_1^2}{4} \right\rfloor + \ell_2(q-1). \tag{6.7}$$

For every vertex w of deficit two, with $Z_w = \varnothing$ if $w \notin X$, one also has

$$N_F(w) \cap G_0 = Z_w, \qquad |Z_w| = q - 1 - d_L(w) - d_D(w). \tag{6.8}$$

Proof. For a bad vertex b, its uncovered degree is $u_b = s_b + 1$. The number of its uncovered partners that are also neighbours is odd, since its neighbourhood of order $q+1$ induces a matching with an odd number of isolated vertices. This number is at most s_b if s_b is odd, and at most $s_b + 1$ if s_b is even. Every good vertex in D has a bad uncovered partner. If B_{even} counts the bad vertices with even s_b, then

$$p \le B_0 + \eta + B_{\text{even}} \le 2\eta,$$

where the last inequality follows from $B_0 + B_{\text{even}} \le \sum_{b\in\mathcal{B}} s_b = \eta$.

Let $v \in G_0$. If $a \in X \setminus \{f(v)\}$ were another neighbour of v, choose $x \in G_0$ with $f(x) = a$. Equation (2.8) would give $xf(v) \in E(F)$, producing the four-cycle $f(v), v, a, x$. Thus $f(v)$ is the only neighbour of v in X. The statements about the other low neighbours now follow from $a_v = 2$. Two vertices with the same pair of low neighbours would form a four-cycle with that pair.

The vertices whose partner has deficit one therefore correspond injectively to pairs consisting of a deficit-one point of X and a deficit-one point of $L \setminus X$. Their number is at most $\lfloor \ell_1^2/4 \rfloor$. A deficit-two point has degree $q-1$, so it can be the partner of at most $q-1$ vertices of G_0. No point of larger deficit is adjacent to a good vertex. This proves (6.7). Finally, if a deficit-two point w is adjacent to $v \in G_0$, then $a_v = 2$ makes w the only low neighbour of v, hence $f(v) = w$. Partitioning $N_F(w)$ into L, G_0, D proves (6.8). □

Lemma 6.3 (Localizing the bad partners)**.** *Assume that there is no good–good uncovered pair. Let $K \subseteq \mathcal{B}$ be the set of bad vertices that have a good vertex of H as an uncovered partner, and put $k = |K|$. Then*

$$d_{G_0}(b) \le 2 \qquad (b \in K), \tag{6.9}$$

$$p \le \eta + 2k, \qquad p \le 2\eta + k - B_0, \tag{6.10}$$

$$\eta \ge \frac{kq}{k+1} - 2k. \tag{6.11}$$

Every member of K has at least $q-1$ neighbours in $D \cup L$. If a member of K has at least two good uncovered partners in H, then it has no neighbour in G_0.

Proof. Fix $b \in K$, and choose a good vertex $x \in H$ with $f(x) = b$. For every $v \in G_0$, equation (2.8) gives

$$A_{b,v} = A_{x,f(v)}. \tag{6.12}$$

If $a \in X \cap N_F(x)$, then every member of Z_a is adjacent to both a and b. Thus $|Z_a| \le 1$. Since $a_x = 2$, there are at most two low neighbours of x, proving (6.9).

For $b \in K$, let g_b count the good vertices of H whose uncovered partner is b. The absence of a good–good uncovered pair gives

$$p = B_0 + \sum_{b \in K} g_b.$$

Since $g_b \le s_b + 1$ and every bad vertex outside K contributes at least one to η,

$$p \le B_0 + \sum_{b \in K} (s_b + 1) \le \eta + 2k.$$

Let k_{even} count the members of K with even s_b. The parity argument in Lemma 6.2 gives $g_b \le s_b$ for odd s_b, and $g_b \le s_b + 1$ for even s_b. Therefore

$$p \le B_0 + \sum_{b \in K} s_b + k_{\mathrm{even}} \le \eta + k + k_{\mathrm{even}}.$$

Also $k_{\mathrm{even}} \le \sum_{b \in \mathcal{B}} (s_b - 1) = \eta - B_0$, which proves the second bound in (6.10).

For $b \in K$, at most $p-1$ neighbours lie in D and at most two lie in G_0. Hence $d_L(b) \ge q - p$ and $s_b = a_b - 2 \ge q - p - 2$. Summing over K and using $p \le \eta + 2k$ gives

$$\eta \ge k(q - p - 2) \ge k(q - \eta - 2k - 2),$$

which rearranges to (6.11).

For the last assertion, suppose that x_1, x_2 are two good uncovered partners of b and that $v \in G_0 \cap N_F(b)$. Putting $a = f(v)$ in (6.12) shows that both $x_1 a$ and $x_2 a$ are edges. Both x_1, x_2 are also adjacent to b, producing the four-cycle b, x_1, a, x_2. □

6.2 The minimum degree of an extremal core

For $q \ge 6$, the degree-sequence restriction of Firke, Kosek, Nash and Williford [6, Theorem 3] gives degrees between $q - 2$ and $q + 1$ under (6.2), with at most one vertex of degree $q - 2$. Corollary 4.2 gives this initial lower bound independently and strengthens it to $q - 1$ when $q = 4$. The next theorem excludes degree $q - 2$ throughout.

Theorem 6.4 (Minimum degree)**.** *Let $q \ge 4$ be even. Every graph satisfying (6.2) has*

$$d_F(v) \in \{q - 1, q, q + 1\} \qquad (v \in V(F)).$$

Proof. Corollary 4.2 gives $\delta_v \le 3$ for every vertex, and already gives $\delta_v \le 2$ for $q \le 10$. It remains to exclude deficit three when $q \ge 12$.

First, there is at most one vertex of deficit three. Indeed, let $W_3 = \{v : \delta_v = 3\}$, let $r = |W_3|$, and put $J_3 = \sum_{w \in W_3} d_L(w)$. If a full-degree vertex has $j \ge 1$ neighbours in W_3, its surplus is at least $3j - 2 \ge j$. Consequently

$$\eta \ge e_F(W_3, H) = r(q - 2) - J_3, \qquad \omega \ge 3J_3.$$

The weighted budget therefore gives

$$2q \ge \eta + 10r + \omega \ge r(q + 8) + 2J_3,$$

which is impossible for $r \ge 2$. Suppose henceforth that w is the unique deficit-three vertex.

We claim that there is no good–good uncovered pair. Suppose that u, v are such a pair. By (6.5), u has at least $q-2$ bad full-degree neighbours. Put $j = d_L(w)$ and $r_w = d_H(w) = q-2-j$.

All full-degree neighbours of w are bad, and their set intersects $N_F(u)\cap\mathcal{B}$ in at most one point. Thus $\eta \geq q + r_w - 3$. Every edge of $F[L]$ incident with w contributes at least four to ω, so

$$2q \geq \eta + 10 + \omega \geq q + r_w - 3 + 10 + 4(q - 2 - r_w) = 5q - 1 - 3r_w \geq 2q + 5,$$

a contradiction. We may consequently use G_0, D, p, K, k from Lemmas 6.2 and 6.3.

Let t be the number of deficit-two vertices. The deficit sum gives

$$|L| = 2q - t - 2, \qquad |H| = q^2 - q + t + 2, \qquad \ell_1 = 2q - 2t - 3.$$

In particular, $0 \leq t \leq q - 2$. By the support-pair bound,

$$q^2 - q + t + 2 - p \leq (q - t - 1)(q - t - 2) + t(q - 1),$$

and hence

$$p \geq 2q + t(q - t - 1). \tag{6.13}$$

The weighted budget is

$$\eta + 4t + 10 + \omega = 2q. \tag{6.14}$$

Since $p \leq 2\eta$, equation (6.13) implies $\eta \geq q$.

Put $T_0 = D \cup L$. Equations (6.6) and (6.14) give

$$|T_0| \leq 2\eta + |L| \leq 6q - 22 - 9t.$$

Every member of K has at least $q-1$ neighbours in T_0. Any two such neighbourhoods intersect in at most one point. Thus, for any $K' \subseteq K$,

$$|T_0| \geq |K'|(q - 1) - \binom{|K'|}{2}. \tag{6.15}$$

Seven members would give $|T_0| \geq 7q - 28 > 6q - 22 - 9t$, so $k \leq 6$. It follows that $p \leq \eta + 12 \leq 2q + 2 - 4t$. Combining this with (6.13) yields

$$t(q - t + 3) \leq 2.$$

If $t \geq 1$, then $t \leq q - 2$ makes the left side at least five. Therefore $t = 0$.

We now have

$$p \geq 2q, \qquad q \leq \eta \leq 2q - 10, \qquad |T_0| \leq \eta + 2k + 2q - 2 \leq 4q - 12 + 2k. \tag{6.16}$$

If $k = 6$, equation (6.15) gives $6q - 21 \leq |T_0| \leq 4q$, impossible for $q \geq 12$. If $k = 5$ and $q \geq 14$, it gives $5q - 15 \leq |T_0| \leq 4q - 2$, again impossible. Thus, for $q \geq 14$, one has $k \leq 4$, whence

$$2q \leq p \leq \eta + 8 \leq 2q - 2.$$

Only $q = 12$ remains. By (6.16), $12 \leq \eta \leq 14$ and $\omega = 14 - \eta \leq 2$. Since every low–low edge incident with w contributes at least four to ω, one has $d_L(w) = 0$. All ten neighbours of w are therefore bad full-degree vertices, so $B_0 \geq 10$. The preceding six-vertex argument gives $k \leq 5$, and the sharper bound in (6.10) yields

$$p \leq 2\eta + k - B_0 \leq 28 + 5 - 10 = 23 < 24,$$

contrary to $p \geq 2q$. □

6.3 Deficit two and the zero-surplus case

The next capacity alternative will be used both to recognize the zero-surplus case and to prove the quantitative gap.

Lemma 6.5 (A capacity alternative for deficit two)**.** *Suppose that F satisfies (6.2) and has no good–good uncovered pair. If a vertex of deficit two exists, then at least one of*

$$2p \geq q, \qquad 3p - \eta \geq q + 1 \tag{6.17}$$

holds.

Proof. By Theorem 6.4, all positive deficits are one or two. Let $t \geq 1$ count the deficit-two vertices. Then

$$|H| = q^2 - q + t, \qquad \ell_1 = 2q - 2t, \qquad \eta + 4t + \omega = 2q.$$

Equation (6.7) gives

$$p \geq (t-1)q - t(t-2).$$

For $2 \leq t \leq q$, this concave quadratic is at least q: its values at $t = 2$ and $t = q$ are both q. Thus the first alternative in (6.17) holds whenever $t \geq 2$.

It remains to consider a unique deficit-two vertex w. Put $J_w = d_L(w)$ and $n_D = d_D(w)$. If w is inactive, equation (6.8) gives $J_w = q - 1 - n_D \geq q - 1 - p$. Its low–low edges contribute at least $3J_w$ to ω. The budget therefore yields

$$2q \geq \eta + 4 + 3(q - 1 - p), \qquad 3p - \eta \geq q + 1.$$

Suppose instead that w is active, and write $Z = Z_w$. Every vertex of $G_0 \setminus Z$ is represented by a pair of deficit-one low neighbours, one active and one inactive. There are $2q-2$ deficit-one vertices, so

$$q^2 - q + 1 - p = |G_0| \leq (q-1)^2 + |Z| = (q-1)^2 + q - 1 - J_w - n_D.$$

Hence

$$J_w + n_D \leq p - 1. \tag{6.18}$$

The set Z is independent: an edge between two of its vertices would give one of them a common neighbour with its uncovered partner w. Each member of Z has w as its only low neighbour and has q full-degree neighbours. Among its neighbours in $G_0 \setminus Z$, at most one can use any given active low point and at most one can use any given inactive low point, by C_4-freeness. The two sets of deficit-one points have total size $2q - 2$. Thus at most $q - 1$ of its neighbours lie in G_0, and it has at least one neighbour in D.

Every member of D has at most one neighbour in Z, because all vertices of Z are adjacent to w. Moreover a member of D adjacent to w has no neighbour in Z, because w is the uncovered partner of every vertex of Z. Consequently $|Z| + n_D \leq p$. Together with (6.18), this gives

$$q - 1 = J_w + |Z| + n_D \leq J_w + p \leq 2p - 1,$$

and hence $2p \geq q$. □

Theorem 6.6 (Zero surplus and unique polarity completion)**.** *Let F satisfy (6.2). If $\eta = 0$, then all its degrees belong to $\{q, q+1\}$, and there is a partition*

$$L = X \sqcup Y, \qquad |X| = |Y| = q,$$

with a vertex $c \in Y$ such that the following hold.

1. *The graph $F[L]$ is the star with centre c and leaf set X, together with $q-1$ isolated vertices.*
2. *Every vertex of H has one neighbour in X and one in $Y\setminus\{c\}$. Each pair in $X\times(Y\setminus\{c\})$ is the neighbourhood in L of exactly one vertex of H.*
3. *Adjoining a new vertex p with neighbourhood Y gives a C_4-free graph $\widehat{F}$ on q^2+q+1 vertices with M edges. If $\widehat{A}$ is its adjacency matrix and $D_{\mathcal{A}}$ is the diagonal indicator matrix of $\mathcal{A}=X\cup\{p\}$, then*

$$B=\widehat{A}+D_{\mathcal{A}}\quad \textit{satisfies}\quad B^2=qI+J. \tag{6.19}$$

The matrix B defines a projective plane of order q and a polarity with exactly $q+1$ absolute points, namely $\mathcal{A}$. In particular, F is obtained by deleting an absolute point from its polarity graph. The one-point polarity completion is unique on the labelled vertex set $V(F)\cup\{p\}$.

Proof. Since $\eta=0$, every vertex of H is good. Equation (6.5) excludes a good–good uncovered pair, and (6.6) gives $D=\varnothing$. Lemma 6.5 excludes deficit two. Theorem 6.4 therefore shows that every positive deficit is one. It follows that

$$|L|=2q,\qquad |H|=q(q-1),\qquad \eta=2\bigl(q-e(F[L])\bigr)=0.$$

Thus $F[L]$ has q edges and every vertex of H has exactly two low neighbours.

All full-degree vertices belong to G_0. Put $X=f(H)$ and $Y=L\setminus X$. By Lemma 6.2, every vertex of H has one neighbour in each of X,Y, and no pair is repeated. Let T_* be the resulting simple bipartite support graph on X,Y: its $q(q-1)$ edges correspond to the vertices of H. Write $J_*=F[L]$ and $j_v=d_{J_*}(v)$. For every $v\in L$,

$$d_{T_*}(v)=d_H(v)=q-j_v. \tag{6.20}$$

For $a\in X$, the full-degree vertices uncovered from a are exactly the $d_{T_*}(a)$ members of its fibre. Equation (6.1) shows that a has $q-1+j_a$ uncovered partners in total, hence exactly $2j_a-1$ uncovered partners in L. This number is nonnegative, so $j_a\geq 1$ and $d_{T_*}(a)\leq q-1$. Since vertices of Y have support degree at most q,

$$|X|\geq q,\qquad |Y|\geq q-1.$$

As $|X|+|Y|=2q$, only

$$(|X|,|Y|)=(q,q)\quad\text{or}\quad(q+1,q-1) \tag{6.21}$$

is possible.

We need the following identity in the low-degree graph:

$$\sum_{b\in N_{J_*}(a)} j_b=q\qquad (a\in X). \tag{6.22}$$

Indeed, precisely $2q-2j_a$ points of $L\setminus\{a\}$ are covered from a. Of these, $q-j_a$ have their common neighbour with a in H, one for each support edge at a. The others are counted by $\sum_{b\in N_{J_*}(a)}(j_b-1)$. There are no repetitions or overlaps, by C_4-freeness. Rearrangement proves (6.22).

Suppose first that $(|X|,|Y|)=(q+1,q-1)$. Every vertex of Y has support degree q, so $j_y=0$. All q edges of J_* then lie in X, and every vertex of X has positive J_*-degree. The graph $J_*[X]$ has $q+1$ vertices and q edges, so it has a tree component; otherwise each component would have at least as many edges as vertices. This component is nontrivial and

has a leaf a. By (6.22), its unique neighbour has J_*-degree q, hence support degree zero. This contradicts its membership in $X = f(H)$.

Thus $|X| = |Y| = q$. Every point of X attains the support-degree bound $q-1$, and so has J_*-degree one. If c is the unique low neighbour of a point of X, (6.22) gives $j_c = q$. Since J_* has exactly q edges, all its edges are incident with c. The point c lies in Y, and the q points of X are exactly the leaves of this star. All other points of Y are isolated in J_*. Equation (6.20) now gives

$$T_* = K_{X,Y\setminus\{c\}}.$$

This proves the first two assertions.

Every pair of distinct vertices of Y is uncovered in F. A full-degree vertex has only one neighbour in Y, and the star description of J_* excludes a common neighbour in L. Adjoining p with neighbourhood Y therefore preserves C_4-freeness. The resulting graph has M edges; its degree-q vertices are exactly $\mathcal{A} = X \cup \{p\}$, while all others have degree $q+1$.

The set $\mathcal{A}$ is independent, and every edge incident with $\mathcal{A}$ has codegree zero. For an edge av with $a \in X$ and $v \in H$, this follows from $f(v) = a$; adding p does not change that codegree. The star edges ac have codegree zero, as do the edges py because Y is independent. The total number of uncovered pairs in $\widehat{F}$ is

$$\binom{q^2+q+1}{2} - q^2\binom{q+1}{2} - (q+1)\binom{q}{2} = q(q+1).$$

This equals the number of edges incident with $\mathcal{A}$. Every other pair of distinct vertices consequently has codegree one. Adding a diagonal one at each point of $\mathcal{A}$ changes the off-diagonal entries of $\widehat{A}^2$ precisely at those incident edges, and gives diagonal entries $q+1$. This proves (6.19).

The rows and columns of B have size $q+1$, and any two distinct rows or columns meet in exactly one position. They define a projective plane of order q. Symmetry gives a polarity by exchanging the point and line with the same index. Its absolute points are exactly the diagonal ones of B, namely $\mathcal{A}$. The row indexed by c is exactly $\mathcal{A}$, so the absolute points lie on the line indexed by c.

For uniqueness, the uncovered-pair graph of F consists exactly of

$$K_Y, \qquad \{ca : a \in X\}, \qquad \{vf(v) : v \in H\}. \tag{6.23}$$

Pairs within X are covered by c, pairs in $X \times (Y \setminus \{c\})$ are covered by their support vertex, and every vertex of H has its unique uncovered partner in X; this proves (6.23). A neighbourhood for a new vertex preserves C_4-freeness exactly when it is a clique in this uncovered-pair graph. Every clique of size at least three is contained in Y: vertices of H have uncovered degree one, and a vertex of X has only c as an uncovered partner in L. Thus there is a unique C_4-free extension in which the new vertex has degree q, namely $N(p) = Y$, and no such extension has a larger degree. Every vertex of a polarity graph of a plane of order q has degree q or $q+1$. Any one-point polarity completion is therefore this extension, with the new point absolute. Its incidence matrix is then determined by its adjacency matrix and its degree-q vertices, proving the asserted uniqueness. □

6.4 The gap for two-valued degree sequences

We next give a quantitative version of the low-star criterion. Its square identity also records the interaction between surplus and uncovered pairs.

Theorem 6.7 (The two-degree gap)**.** *Let F satisfy (6.2), and suppose that all its degrees belong to $\{q, q+1\}$. Write*

$$t = e(F[L]), \qquad \kappa = q - t.$$

Then $t \le q$, $\eta = 2\kappa$, and

$$\kappa = 0 \quad \text{or} \quad \kappa \ge \left\lceil \frac{2q-11}{5} \right\rceil. \tag{6.24}$$

Consequently, either condition

$$e(F[L]) = q \qquad \text{or} \qquad e(F[L]) > \frac{3q+11}{5}$$

forces the unique polarity completion in Theorem 6.6.

Proof. The degree sum gives $|L| = 2q$ and $|H| = q(q-1)$. The weighted budget reduces to

$$\eta + 2e(F[L]) = 2q,$$

so $t \le q$ and $\eta = 2\kappa$. Put $h_0 = q(q-1)$ and $r_v = d_L(v)$ for $v \in L$. Let u_{HH}, u_{LL}, u_{HL} count uncovered unordered pairs of the indicated types. Summing (6.1) over the two parts gives

$$2u_{HH} + u_{HL} = h_0 + 2\kappa, \qquad 2u_{LL} + u_{HL} = 2q(q-1) + 2t = 2q^2 - 2\kappa.$$

Hence

$$2u_{LL} = q^2 + q - 4\kappa + 2u_{HH}.$$

Counting covered pairs in L by their unique common neighbour gives

$$\begin{aligned} \binom{2q}{2} - u_{LL} &= \sum_{v \in L} \binom{r_v}{2} + \sum_{v \in H} \binom{2+s_v}{2} \\ &= \frac{1}{2} \sum_{v \in L} r_v^2 + h_0 - q + 5\kappa + \sum_{v \in H} \binom{s_v}{2}. \end{aligned}$$

Substitution yields the identity

$$\sum_{v \in L} r_v^2 + 2 \sum_{v \in H} \binom{s_v}{2} + 2u_{HH} = q^2 + q - 6\kappa. \tag{6.25}$$

If a good–good uncovered pair exists, (6.5) gives

$$\kappa \ge \frac{q-2}{2} > \frac{2q-11}{5},$$

as required. Otherwise every uncovered pair within H has a bad endpoint. Since $u_v = s_v + 1$ for $v \in H$,

$$u_{HH} \le \sum_{v \in \mathcal{B}} u_v = B_0 + 2\kappa \le 4\kappa.$$

A graph with t edges satisfies $\sum_v d(v)^2 \le t(t+1)$. Indeed, for every edge uv the incident edge sets have union of size $d(u) + d(v) - 1 \le t$, and summing $d(u) + d(v) \le t + 1$ over the edges proves the assertion. Also $\sum_{v \in H} s_v = 2\kappa$ gives

$$\sum_{v \in H} \binom{s_v}{2} \le \binom{2\kappa}{2}.$$

Using these bounds in (6.25), we obtain

$$\begin{aligned} \kappa(2q - 5 - \kappa) &= q^2 + q - 6\kappa - (q-\kappa)(q-\kappa+1) \\ &\le 2u_{HH} + 2 \sum_{v \in H} \binom{s_v}{2} \\ &\le 8\kappa + 2\binom{2\kappa}{2} = 6\kappa + 4\kappa^2. \end{aligned}$$

If $\kappa > 0$, division gives $5\kappa \ge 2q - 11$. This proves (6.24). The stated recognition conditions force $\kappa = 0$ and therefore $\eta = 0$. □

6.5 A gap for the weighted surplus

The support and partner estimates give a quantitative separation between zero and positive surplus.

Theorem 6.8 (The surplus gap)**.** *Let $q \geq 16$ be even, and let F satisfy (6.2). Then*

$$\eta = 0 \quad \text{or} \quad \eta \geq \frac{q}{2} - 2. \tag{6.26}$$

If F has a vertex of degree $q-1$, then the second alternative holds. In particular, $\eta < q/2 - 2$ forces the unique absolute-point polarity completion of Theorem 6.6.

Proof. Suppose $\eta < q/2 - 2$. Equation (6.5) excludes a good–good uncovered pair, so the notation and bounds of Lemmas 6.2 and 6.3 apply. Since $|L| \leq 2q$ and $p \leq 2\eta$,

$$|D \cup L| \leq 2\eta + 2q < 3q - 4.$$

Four members of K would have at least $4(q-1) - \binom{4}{2} = 4q - 10$ distinct neighbours in this set. This is impossible for $q \geq 16$, so $k \leq 3$. For $k = 1, 2, 3$, the lower bounds in (6.11) are

$$\frac{q}{2} - 2, \qquad \frac{2q}{3} - 4, \qquad \frac{3q}{4} - 6,$$

respectively. All are at least $q/2 - 2$ when $q \geq 16$. Hence $k = 0$, and every good full-degree vertex belongs to G_0. Thus $D = \mathcal{B}$ and $p \leq \eta$.

Theorem 6.4 leaves only deficits one and two. If a deficit-two vertex existed, Lemma 6.5 would give either $2p \geq q$ or $3p - \eta \geq q + 1$. But

$$2p \leq 2\eta < q - 4, \qquad 3p - \eta \leq 2\eta < q - 4,$$

contradicting both alternatives. All positive deficits therefore equal one. By Theorem 6.7, either $\eta = 0$ or

$$\eta \geq 2 \left\lceil \frac{2q - 11}{5} \right\rceil \geq \frac{q}{2} - 2,$$

where the last inequality holds for $q \geq 16$. Thus $\eta = 0$. The remaining assertions follow from Theorem 6.6. □

6.6 An unrestricted degree conclusion

Corollary 6.9 (Unrestricted degree control for extremal cores)**.** *For every even $q \geq 10^9$, every C_4-free graph F on $q^2 + q$ vertices with m_0 edges satisfies*

$$d_F(v) \in \{q - 1, q, q + 1\} \qquad (v \in V(F)).$$

Moreover, with

$$\delta_v = q + 1 - d_F(v), \quad H = \{v : d_F(v) = q + 1\}, \quad \eta(F) = \sum_{v \in H} \left(\sum_{w \in N_F(v)} \delta_w - 2 \right),$$

one has $\eta(F) = 0$ or $\eta(F) \geq q/2 - 2$. In the zero case, F has the unique absolute-point polarity completion already proved for capped cores.

Proof. Apply Theorem 4.7 with $b = 0$, then Theorems 6.4, 6.8, and 6.6. □

7 Graphs attaining the extremal bound

Write

$$N = q^2 + q + 1, \qquad n = N - 1, \qquad M = \frac{1}{2}q(q+1)^2, \qquad m_0 = M - q.$$

At zero edge deficit, concentration in two vertices gives the sufficient threshold $q \geq 10^{15}$, and the resulting recognition theorem determines the missing absolute point uniquely. The classification statement was announced by Firke, Kosek, Nash, and Williford [6, Theorem 2], who stated that its proof would appear in subsequent work. The argument below derives this classification and the unique completion from the degree and incidence-completion results.

We first assume that $q \geq 10$ is even and that F is a C_4-free graph satisfying

$$|V(F)| = q^2 + q, \qquad e(F) = m_0, \qquad q - 1 \leq d_F(v) \leq q + 1 \quad (v \in V(F)). \tag{7.1}$$

Theorems 4.7 and 6.4 will remove both degree hypotheses in the final result. Put

$$\delta_v = q + 1 - d_F(v), \quad H = \{v : \delta_v = 0\}, \quad L = \{v : \delta_v > 0\}, \quad Z = \{v : \delta_v = 2\}, \quad t = |Z|.$$

For every vertex write

$$a_v = \sum_{w \in N_F(v)} \delta_w, \qquad r_v = |N_F(v) \cap L|,$$

and let u_v be the number of other vertices having no common neighbour with v. For subsets I, J of the partition $V(F) = L \sqcup H$, write u_{IJ} for the number of uncovered unordered pairs of the indicated type. In particular u_{HH} counts unordered pairs, not the sum of their degrees.

The degree sum and the exact local path count give

$$\sum_{v \in L} \delta_v = 2q, \qquad |L| = 2q - t, \qquad |H| = q^2 - q + t, \tag{7.2}$$

$$u_v = q\delta_v - 1 + a_v. \tag{7.3}$$

At every vertex of H, the neighbourhood has odd order $q + 1$ and induces a matching with isolated vertices. Thus $u_v \geq 1$ and $a_v \geq 2$. Set

$$s_v = a_v - 2 \quad (v \in H), \qquad \eta = \sum_{v \in H} s_v, \qquad \omega = \sum_{v \in L} \delta_v r_v.$$

Weighted counting of the edges between L and H gives the exact budget

$$\eta + 4t + \omega = 2q. \tag{7.4}$$

Indeed,

$$2|H| + \eta = \sum_{v \in L} \delta_v(q + 1 - \delta_v - r_v) = 2q^2 - 2t - \omega,$$

which is equivalent to (7.4).

7.1 A bounded number of vertices of deficit two

A full-degree vertex is called good when $s_v = 0$, and bad otherwise. Write $\mathcal{B}$ for the bad set and $B_0 = |\mathcal{B}|$. A good vertex has precisely one uncovered partner, denoted by $f(v)$; neighbourhood parity also gives $vf(v) \in E(F)$.

For clarity, the commutation relation used below follows directly from the C_4-free hypothesis. Let A be the adjacency matrix and U the zero-diagonal uncovered-pair matrix. Then

$$A^2 = qI + J - U - \operatorname{diag}(\delta_v),$$

and therefore

$$(AU - UA)_{uv} = (\delta_v - \delta_u)(1 - A_{uv}). \tag{7.5}$$

For good vertices u, v, this becomes

$$A_{u,f(v)} = A_{f(u),v}. \tag{7.6}$$

Lemma 7.1 (The good–good pairs and the exceptional capacity)**.** *Under* (7.1)*, at most one uncovered pair consists of two good vertices. Consequently*

$$u_{HH} \le 2\eta + 1. \tag{7.7}$$

Define

$$G_0 = \{v \in H : v \text{ is good and } f(v) \in L\}, \qquad D = H \setminus G_0, \qquad p = |D|.$$

Then

$$p \le 2\eta + 2, \qquad |G_0| \le (q-t)^2 + t(q-1). \tag{7.8}$$

In particular,

$$t \le 4. \tag{7.9}$$

Proof. Suppose u, v are good and $f(u) = v$. If $w \neq v$ were another good neighbour of u, then (7.6), applied to v, w, would give $vf(w) \in E(F)$. The four distinct vertices $u, v, f(w), w$ would form a four-cycle. Distinctness follows from the uniqueness of the uncovered partners at u and v. Thus v is the only good neighbour of u. Since $a_u = 2$, at most two neighbours of u lie in L, even when one of them has deficit two. Hence u has at least $q - 2$ bad full neighbours.

If three good vertices participated in good–good uncovered pairs, their bad-neighbour sets would have union of size at least $3(q-2) - 3 = 3q - 9$: each pair of such neighbourhoods intersects in at most one vertex. This is impossible because $B_0 \le \eta \le 2q$ and $q \ge 10$. At most two good vertices therefore participate, giving at most one such pair. Every other uncovered H–H pair touches $\mathcal{B}$, so

$$u_{HH} \le 1 + \sum_{b \in \mathcal{B}} u_b = 1 + \eta + B_0 \le 2\eta + 1.$$

For each bad vertex b, the number of uncovered partners that are also adjacent to b is odd, because it equals the number of isolated vertices in the matching induced by $N_F(b)$. As $u_b = s_b + 1$, this number is at most s_b when s_b is odd and at most $s_b + 1$ when s_b is even. Every good vertex in D either has a bad partner or belongs to the one possible good–good pair. If B_{even} denotes the number of bad vertices with even s_b, it follows that

$$p \le B_0 + \eta + B_{\text{even}} + 2 \le 2\eta + 2.$$

The final inequality uses $B_0 + B_{\text{even}} \le \sum_{b \in \mathcal{B}} s_b = \eta$.

Put $X = f(G_0) \subseteq L$. For $v \in G_0$, its only neighbour in X is $f(v)$. Indeed, if $a \in X \setminus \{f(v)\}$ were another neighbour, choose $x \in G_0$ with $f(x) = a$. Equation (7.6) gives $xf(v) \in E(F)$, and $f(v), v, a, x$ form a four-cycle. If $\delta_{f(v)} = 1$, the equality $a_v = 2$ says that v has exactly one further low neighbour, also of deficit one, in $L \setminus X$. No two vertices of G_0 have the same pair of low neighbours. There are $2q - 2t$ vertices of deficit one, so these pairs number at most $(q-t)^2$. If $\delta_{f(v)} = 2$, then $f(v)$ is the only low neighbour of v; each such partner has degree $q - 1$ and supports at most $q - 1$ vertices. This proves the second assertion in (7.8).

Since $|H| = q^2 - q + t$, we obtain

$$p \ge (t-1)q - t(t-2).$$

Together with $p \le 2\eta + 2$ and (7.4), this gives

$$(t-5)q - t^2 + 10t - 2 \le 0. \tag{7.10}$$

The budget also gives $0 \le t \le q/2$. As a function of t, the left side of (7.10) is concave. On $[5, q/2]$ its values at the two endpoints are 23 and $q^2/4 - 2$, both positive. This interval is therefore excluded, and $t \le 4$. □

7.2 Signed moments and two useful deletions

The weights below express the number of full-degree neighbours of each vertex through (7.12). For two deleted vertices, the sum of their weights controls the number of full-degree vertices that remain. The weights may be negative, so we distinguish their signed moments from those of their positive parts.

Define

$$w_v = \begin{cases} r_v + \delta_v - 1, & v \in L, \\ r_v - 2, & v \in H, \end{cases} \qquad R_Z = \sum_{z \in Z} r_z. \tag{7.11}$$

For every vertex,

$$d_H(v) = q - w_v - \mathbf{1}_H(v). \tag{7.12}$$

Lemma 7.2 (The signed moment identities)**.** *Under* (7.1)*,*

$$\sum_v w_v = 2q - t(q+2), \tag{7.13}$$

$$\sum_v w_v^2 = q^2 + (1-t)q + t^2 - 3t - 2\eta - 2R_Z - 2u_{HH}. \tag{7.14}$$

Writing $w_v^+ = \max\{w_v, 0\}$*, one has*

$$\sum_v w_v^+ \le 2q, \qquad \sum_v (w_v^+)^2 \ge q^2 - 27q - 4. \tag{7.15}$$

Proof. Put $\ell = |L| = 2q - t$, $h = |H| = q^2 - q + t$, and $e_L = e(F[L])$. Since the vertices of $L \setminus Z$ have degree q and those of Z have degree $q-1$,

$$\sum_v r_v = \sum_{v \in L} d_F(v) = q\ell - t.$$

Adding $\sum_L (\delta_v - 1) = t$ and subtracting $2h$ proves (7.13).

We derive the second identity without an assumption about the signs of the weights. Summing (7.3) over L and then over H gives

$$2u_{LL} + u_{LH} = 2q^2 - \ell + \omega, \qquad 2u_{HH} + u_{LH} = h + \eta.$$

Subtracting yields

$$2u_{LL} = q^2 - q + \omega - \eta + 2u_{HH}. \tag{7.16}$$

The number of covered pairs in L is $\sum_v \binom{r_v}{2}$, because each such pair has at most one common neighbour. Therefore

$$\sum_v r_v^2 = q\ell - t + \ell(\ell-1) - 2u_{LL}. \tag{7.17}$$

Also

$$\omega = 2e_L + R_Z, \qquad \sum_{v \in H} r_v = q\ell - t - 2e_L.$$

Relative to $\sum_v r_v^2$, replacing r_v by $r_v - 2$ on H adds $-4\sum_H r_v + 4h$, and replacing r_z by $r_z + 1$ on Z adds $2R_Z + t$. Thus

$$\begin{aligned}\sum_v w_v^2 &= q\ell - t + \ell(\ell-1) - 2u_{LL} - 4(q\ell - t - 2e_L) + 4h + 2R_Z + t \\ &= \ell(\ell-1) - 3q\ell + 4t + 4h + 4\omega - 2R_Z - 2u_{LL} \\ &= \ell(\ell-1) - 3q\ell + 4t + 4h - q^2 + q + 3\omega + \eta - 2R_Z - 2u_{HH}.\end{aligned}$$

Now substitute $\ell = 2q - t$, $h = q^2 - q + t$, and $\omega = 2q - 4t - \eta$. This gives (7.14) exactly.

Every low weight is nonnegative. At a full vertex $r_v \geq 1$, because $a_v \geq 2$ and the only positive deficits are one and two. Thus a negative weight is exactly -1, and it occurs precisely at a full vertex whose only low neighbour belongs to Z. Let a_- be the number of such vertices. Then

$$a_- \leq e(Z,H) = t(q-1) - R_Z \leq t(q-1). \tag{7.18}$$

Consequently

$$\sum_v w_v^+ = 2q - t(q+2) + a_- \leq 2q - 3t \leq 2q, \qquad \sum_v (w_v^+)^2 = \sum_v w_v^2 - a_-.$$

By Lemma 7.1, $t \leq 4$ and $u_{HH} \leq 2\eta + 1$. We may use the coarse bounds

$$(1-t)q \geq -3q, \quad t^2 - 3t \geq -2, \quad \eta \leq 2q, \quad R_Z \leq 4q, \quad u_{HH} \leq 4q+1, \quad a_- \leq 4q.$$

Their substitution into (7.14) proves the second inequality in (7.15). □

Lemma 7.3 (A two-largest-weights estimate)**.** *Let $x_1 \geq x_2 \geq \cdots \geq 0$ be a finite sequence with $\sum_i x_i \leq 2q$, and put $W = x_1 + x_2$ (adjoining zero terms if necessary). If $W \leq q$, then*

$$\sum_i x_i^2 \leq qW. \tag{7.19}$$

Proof. Put $a = x_1$ and $b = x_2 = W - a$. Every remaining term is at most b, so

$$\sum_i x_i^2 \leq a^2 + b(2q - a).$$

Subtracting qW from the last expression gives $(2a - W)(a - q) \leq 0$, since $W/2 \leq a \leq W \leq q$. □

Proposition 7.4 (Two deletions leave many full vertices)**.** *Under (7.1), there are two distinct vertices u, v with $w_u, w_v \geq 0$ and*

$$w_u + w_v \geq q - 27. \tag{7.20}$$

For $F_0 = F - \{u,v\}$, at least

$$q^2 - 2q - 27 = N - 3(q+1) + \binom{3}{2} - 28 \tag{7.21}$$

vertices have degree $q+1$ and are adjacent to neither deleted vertex.

Proof. Apply Lemma 7.3 to the positive weights. If their largest-two sum is at most q, then (7.15) gives $W \geq q - 27 - 4/q$. The weights are integers and $q > 4$, so $W \geq q - 27$. If $W > q$, this conclusion is immediate. We can realize the two chosen positive-part weights by distinct vertices of nonnegative actual weight: if at most one weight is positive, use any other low vertex as the second vertex. There are at least $2q - 4$ low vertices, and their weights are nonnegative.

Every full vertex outside $\{u,v\} \cup N_F(u) \cup N_F(v)$ retains its degree in F_0. By (7.12), their number is at least

$$\begin{aligned}|H| - |\{u,v\} \cap H| - d_H(u) - d_H(v) &= |H| - 2q + w_u + w_v \\ &\geq q^2 - 2q + t - 27 \geq q^2 - 2q - 27.\end{aligned}$$

This proves (7.21). □

7.3 From the incidence completion to a polarity host

The incidence-completion and polarity-extension theorems apply with three omitted points and the clean full-degree set constructed above.

Lemma 7.5 (A host after the two deletions)**.** *Let $q \geq 10^{15}$ be even, let F satisfy (7.1), and choose u, v, F_0 as in Proposition 7.4. There are a projective plane Π of order q, a polarity π of Π, and a three-point set Q such that*

$$F_0 \subseteq P_\pi - Q$$

as a spanning subgraph, where P_π denotes the loopless polarity graph. A set C of $q^2 - 2q - 27$ full vertices can be chosen among the vertices adjacent to neither u nor v so that every old polar line has at most 31 points outside C.

Proof. Choose C as in (7.21), of exactly the displayed size. On the point set C take the indexed lines

$$S_x = N_{F_0}(x) \cap C \qquad (x \in V(F_0)).$$

Every point lies on exactly $q+1$ indexed lines, and any two lines meet in at most one point. The parameters of the fixed-parameter incidence-completion theorem are

$$s = 3, \qquad t_{\mathrm{inc}} = 28, \qquad K = s + t_{\mathrm{inc}} = 31. \tag{7.22}$$

The shortest-line packing lemma gives $|S_x| \geq q - 30$ for every x. In particular the degree deficits $q + 1 - d_{F_0}(x)$ are at most 31. A full vertex $x \in C$ has at most 31 neighbours outside C, and all its neighbours in C have deficit zero. Its weighted neighbour deficit is therefore at most 31^2. At the order $N - 3$, the exact local identity for this full vertex is $u_x = a_x - 3$. Hence the uncovered-pair graph on C has maximum degree at most 31^2. Theorem A.2, with $D = 961$, embeds the entire indexed incidence system into a projective plane Π of order q. Write p_x for the point images of $x \in C$ and ℓ_y for the distinct line images of all $y \in V(F_0)$. The incidences are exact:

$$p_x \in \ell_y \quad \Longleftrightarrow \quad xy \in E(F_0) \qquad (x \in C,\ y \in V(F_0)). \tag{7.23}$$

There are three missing lines. Every point of C is already incident with its full pencil of $q+1$ indexed lines and therefore lies on none of them. Their union has at least $3(q+1)-3 = 3q$ points. The complement of C has $3q+28$ points, so at most 28 of these avoid all three missing lines. Every indexed line meets the missing-line union in at most three points; it therefore has at most 31 points outside C.

Apply Lemma B.2 with $s = 3$ and $t = 28$. The chosen set C has the required size, all its vertices have degree $q+1$ in F_0, and $q \geq 10^{15} > (3+28+2)^2$. The lemma extends (7.23) to distinct point images p_x for all $x \in V(F_0)$ and a polarity π with $\pi(p_x) = \ell_x$. In particular, the old incidences are symmetric:

$$p_x \in \ell_y \quad \Longleftrightarrow \quad p_y \in \ell_x. \tag{7.24}$$

For this old assignment, the pairing of the three omitted points and lines is uniquely determined by Lemma B.1. Let Q be the three unassigned points. The respected-edge conclusion of Lemma B.2 gives $F_0 \subseteq P_\pi - Q$.

We verify the stated numerical range of the completion theorem. For (7.22) its constants are

$$K = 31, \quad A_0 = 64, \quad B_0 = 1856, \quad Y_0 = 3712, \quad \mu = 961, \quad C_1 = 32515, \quad a_0 = 4525.$$

Its packing constant is 644 and its bad-pair threshold is 105989597184. The other polynomial thresholds are at most 13775263, except for the Beutelspacher–Metsch condition $4q > f(a_0)$, where

$$f(a) = 6a^4 + 9a^3 + 19a^2 + 8a, \qquad f(4525) = 2516341573744950 < 4 \cdot 10^{15}.$$

Thus all completion conditions hold for $q \geq 10^{15}$. This range also covers all pencil, injectivity, and Gram-extension inequalities used above. □

7.4 Reinserting the two deleted vertices

Theorem 7.6 (Recognition of all extremal cores)**.** *Let $q \geq 10^{15}$ be even. If F is a C_4-free graph with*

$$|V(F)| = q^2 + q, \qquad e(F) = \frac{1}{2}q(q+1)^2 - q,$$

then F is obtained by deleting an absolute point from an orthogonal polarity graph of a projective plane of order q. In particular all its degrees belong to $\{q, q+1\}$, its surplus $\eta(F)$ is zero, and its completion by the missing absolute point is unique.

Proof. Theorem 4.7, with $b = 0$, gives $\Delta(F) \leq q+1$ for $q \geq 10^9$. The minimum-degree theorem for capped extremal cores gives $\delta(F) \geq q-1$. Thus (7.1) holds. Choose u, v by Proposition 7.4, and apply Lemma 7.5 to $F_0 = F - \{u, v\}$. Identify its vertices with their old points in the plane, and put

$$B_u = N_F(u) \setminus \{u, v\}, \qquad B_v = N_F(v) \setminus \{u, v\}.$$

Each set has size at least $q-2$, and both avoid C by the choice of the clean full vertices.

For $s \in \{u, v\}$, the distinct polar lines $\{\pi(x) : x \in B_s\}$ intersect pairwise outside C; otherwise two neighbours of s would have another common neighbour in C. They are concurrent at a point p_s outside C, by Lemma 2.5 and the bound of 31 outside points on each old polar line. The centre p_s cannot be an old point. If it were the old point $x \notin C$, its at least $q-30$ neighbours in C would index that many lines through p_s. These line indices lie in C, whereas B_s avoids C. Together the two families would have at least $(q-30)+(q-2) > q+1$ distinct lines through p_s. Thus $p_u, p_v \in Q$. Moreover $|B_u \cap B_v| \leq 1$ by C_4-freeness. Equality of the two centres would put at least $2(q-2)-1 > q+1$ distinct lines through one point. Therefore $p_u \neq p_v$. The assignments $u \mapsto p_u$, $v \mapsto p_v$ are injective and respect all old–new edges.

It remains to check the possible edge between the two new points. Write $\varepsilon = A_{uv}$ and retain the original deficits $\delta_u, \delta_v \in \{0, 1, 2\}$. The exact edge count after deleting the two vertices is

$$e(F_0) = M - 3q - 2 + \delta_u + \delta_v + \varepsilon. \tag{7.25}$$

Lemma 2.6 gives

$$e(P_\pi - Q) \leq M - 3q + 3,$$

because the three deleted points have degrees at least q and span at most three edges. Consequently the number D_0 of old host edges absent from F_0 satisfies

$$D_0 \leq 5 - \delta_u - \delta_v - \varepsilon \leq 5. \tag{7.26}$$

Suppose $\varepsilon = 1$ but p_u, p_v are nonadjacent in the polarity graph. Put $\ell = \pi(p_u)$ and $k = \pi(p_v)$. Then $p_v \notin \ell$ and $p_u \notin k$. The rule

$$x \in \ell \longmapsto \pi(x) \cap k$$

is a bijection between the $q+1$ points of ℓ and k. Indeed, it is well defined because $x \neq p_v$; two points with the same image y would force $\pi(y) = \ell$, hence $y = p_u \notin k$, a contradiction. Since $B_u \subseteq \ell$ and $B_v \subseteq k$, at least $|B_u| + |B_v| - (q+1)$ matching pairs lie in $B_u \times B_v$. At most one pair is a loop, because a loop lies in $\ell \cap k$. No nonloop edge is counted in both orientations, since that would require two different vertices in $B_u \cap B_v$. There are therefore at least

$$|B_u| + |B_v| - q - 2 \geq q - 6$$

distinct nonloop old host edges between these two neighbour sets. All must be absent from F_0, since any present one makes a four-cycle with uv. This contradicts $D_0 \leq 5 < q - 6$. Thus the edge uv, if present, is also respected.

Let p be the remaining point of Q. We have embedded all of F as a spanning subgraph of $P_\pi - p$. The edge bound and the minimum degree in Lemma 2.6 give

$$M - q = e(F) \leq e(P_\pi - p) \leq M - q.$$

All inequalities are equalities. Hence $F = P_\pi - p$, $e(P_\pi) = M$, and $d_{P_\pi}(p) = q$. The polarity has exactly $q+1$ absolute points, and p is one of them. Such a polarity is orthogonal; its absolute points form a line by Lemma 2.6.

For the surplus assertion, let $\mathcal{A}$ be this absolute line and let $c = \pi(\mathcal{A})$ be its nucleus. Then $c \notin \mathcal{A}$ and $N_{P_\pi}(c) = \mathcal{A}$. In $F = P_\pi - p$, the low-degree set is the disjoint union

$$X = \mathcal{A} \setminus \{p\}, \qquad Y = N_{P_\pi}(p), \qquad |X| = |Y| = q.$$

Every vertex of this set has degree q, and all others have degree $q+1$. A full vertex differs from c, because $c \in Y$. Its polar line meets $\mathcal{A}$ in exactly one point, which belongs to X; it also has exactly one common neighbour with the nonadjacent absolute point p, and this lies in Y. It therefore has exactly two low neighbours. This proves $\eta(F) = 0$. The zero-surplus completion theorem (Theorem 6.6) now gives uniqueness of the missing absolute-point completion. □

8 Zero full-degree surplus

Put $n = q^2 + q$, $M = q(q+1)^2/2$, and $m_0 = M - q$. For a C_4-free graph F on n vertices with $e(F) = m_0 - b$ and $\Delta(F) \leq q+1$, write

$$\delta_v = q + 1 - d(v), \qquad H = \{v : \delta_v = 0\}, \qquad L = V(F) \setminus H, \qquad a_v = \sum_{w \in N(v)} \delta_w.$$

We consider the nonnegative full-degree surplus

$$\eta = \sum_{v \in H} (a_v - 2).$$

Here q is even and b is a nonnegative integer. Nonnegativity follows from the odd order of every full-degree neighbourhood.

Theorem 8.1 (Classification at zero surplus)**.** *Let $b \geq 0$ and let $q \geq 8192(b+1)^4$ be even. Suppose that F is C_4-free and*

$$|V(F)| = q^2 + q, \qquad e(F) = m_0 - b, \qquad \Delta(F) \leq q + 1, \qquad \eta(F) = 0.$$

Then precisely one of the following holds.

1. *$b = 0$, and F is obtained by deleting an absolute point from a polarity graph of a projective plane of order q.*

2. *$b = 1$, and F is obtained by deleting the point whose polar line consists of the $q + 1$ absolute points of a polarity of a projective plane of order q.*

In either case the one-point polarity completion is unique on the labelled vertex set. In particular, zero surplus is impossible for $b \geq 2$ in the stated range.

The $b = 0$ conclusion holds already for even $q \geq 4$ by the zero-surplus completion theorem, Theorem 6.6. The proof below gives a uniform exclusion of every mixed degree pattern when $b \geq 1$.

8.1 Partners and the low-degree identities

Let U denote the graph of uncovered pairs, and let $u_v = d_U(v)$. Two-path counting and the degree sum give

$$u_v = q\delta_v - 1 + a_v, \qquad \sum_{v \in L} \delta_v = 2q + 2b. \tag{8.1}$$

If $\eta = 0$, every vertex $v \in H$ has $a_v = 2$, so $u_v = 1$. Its unique uncovered partner $f(v)$ is adjacent to v. The codegree commutator gives, for $u, v \in H$,

$$A_{u,f(v)} = A_{f(u),v}. \tag{8.2}$$

These identities follow directly by commuting $A^2 = qI + J - \mathrm{diag}(\delta) - U$ with A.

There is no uncovered pair within H when $q \geq 4$. Indeed, if $f(u) = v \in H$, symmetry gives $f(v) = u$. Equation (8.2) and C_4-freeness show that v is the only full-degree neighbour of u. Since $a_u = 2$, at most two neighbours of u belong to L, contradicting $d(u) = q + 1 \geq 5$. Thus $f(H) \subseteq L$. Call these low vertices active. Each full-degree vertex has its partner as its only active neighbour: an additional active neighbour, together with a full-degree vertex having that partner, would give a four-cycle by (8.2).

Lemma 8.2. *If $\eta = 0$, $q \geq 4$ is even, and $q > 2b$, then every positive deficit is one or two, and every deficit-two vertex is active.*

Proof. A vertex w with $\delta_w \geq 3$ has no full-degree neighbour, because such a neighbour would have $a_v > 2$. It also has no full-degree uncovered partner, since every such partner would have to be adjacent to it. Thus all its uncovered partners lie in L, and

$$3q - 1 \leq u_w \leq |L| - 1 \leq 2q + 2b - 1,$$

contrary to $q > 2b$.

If a deficit-two point w is inactive, it has no full-degree neighbour: any full-degree neighbour has w as its only low neighbour, so its uncovered partner must be w. Consequently $a_w \geq q - 1$ and $u_w \geq 3q - 2$. As before $u_w \leq 2q + 2b - 1$, giving $q \leq 2b + 1$. Evenness and $q > 2b$ exclude this possibility. □

Write X_1 for the active deficit-one vertices, X_2 for the deficit-two vertices, and Y for the inactive vertices. Put

$$x = |X_1|, \qquad y = |Y|, \qquad t = |X_2|, \qquad J = F[L], \qquad j_v = d_J(v), \qquad \lambda_v = j_v + \delta_v - 1.$$

All vertices of Y have deficit one. Every full-degree vertex has either a pair of low neighbours in $X_1 \times Y$, or a single low neighbour in X_2. No pair in $X_1 \times Y$ occurs twice. Write $Z_w = f^{-1}(w)$ for an active fibre. Then

$$x + y = 2q + 2b - 2t, \qquad |H| = q^2 - q - 2b + t, \tag{8.3}$$

$$|Z_w| = q - 1 - j_w \qquad (w \in X_2). \tag{8.4}$$

Lemma 8.3 (Low-degree identities)**.** *Put $s = 2b - t$. Then*

$$\sum_{v \in N_J(a)} \lambda_v = q + s \qquad (a \in X_1), \tag{8.5}$$

$$\sum_{v \in N_J(y)} (\lambda_v - 1) = s \qquad (y \in Y), \tag{8.6}$$

$$\sum_{v \in N_J(w)} (\lambda_v + 1) = q + s - 1 \qquad (w \in X_2). \tag{8.7}$$

Moreover, writing $X = X_1 \cup X_2$,

$$d_X(w) = x + t - q - 1 \qquad (w \in X_2). \tag{8.8}$$

Proof. For an active deficit-one point a, its full-degree uncovered partners are precisely its $q - j_a$ fibre members. Equation (8.1) therefore gives $a_a + j_a - 1$ uncovered partners in L. The covered partners in L are counted by

$$(q - j_a) + \sum_{v \in N_J(a)} (j_v - 1).$$

There is no repetition, by C_4-freeness. Since $|L| = 2q + 2b - t$, subtraction and rearrangement give (8.5).

An inactive point y has no full-degree uncovered partner. Its full-degree neighbours each contribute one distinct covered point of X_1, so the same covered-pair count gives (8.6). For $w \in X_2$, the full-degree uncovered partners number $q - 1 - j_w$, and its full-degree neighbours contribute no covered low point. Applying the count once more gives (8.7).

To prove (8.8), choose $z \in Z_w$. It has w as its sole low neighbour, and exactly q full-degree neighbours. For every $a \in X \setminus \{w\}$, the pair z, a is covered exactly once. If $wa \in E(J)$, its common neighbour is w, so z has no neighbour in Z_a. If $wa \notin E(J)$, its common neighbour is a full-degree vertex in Z_a, giving exactly one neighbour there. There is no neighbour in Z_w, since z, w are uncovered. Counting the full-degree neighbours of z gives $q = x + t - 1 - d_X(w)$, as required. □

8.2 The pure deficit-two case

If Y is empty, then X_1 is empty as well: a full-degree vertex with a deficit-one partner requires a second low neighbour in Y. Thus $L = X_2$, $t = q + b$, and (8.8) gives $j_w = b - 1$ for all w. Every fibre has size $q - b$. Consequently

$$q^2 - b = |H| = (q + b)(q - b) = q^2 - b^2.$$

It follows that $b = 0$ or $b = 1$; the former would give $j_w = -1$, so $b = 1$. The low set has $q + 1$ vertices and is independent; all its vertices have degree $q - 1$, and all other vertices have degree $q + 1$.

Adjoin a point p adjacent to the whole low set. Distinct low vertices have no common neighbour in F, so the extension is C_4-free and has M edges. Its degree-q set is exactly L. Every edge incident with L has codegree zero, by the partner property and the independence of L. There are exactly $q(q + 1)$ uncovered pairs, by the degree count, and exactly this many edges incident with L. Thus, if $\widehat{A}$ is the adjacency matrix of the extension,

$$B = \widehat{A} + \operatorname{diag}(\mathbf{1}_L) \quad \text{satisfies} \quad B^2 = qI + J.$$

The rows of B define a projective plane, and symmetry defines a polarity with absolute set L. The polar line of p is L. Any one-point polarity completion must raise every degree-$q - 1$ vertex to degree at least q; it must therefore join the new point to all of L. This proves uniqueness.

8.3 Excluding mixed deficits at positive edge deficit

Assume henceforth that $b \geq 1$ and $Y \neq \varnothing$. Every summand in (8.6) is nonnegative, because it is evaluated at a neighbour. Thus

$$t \leq 2b, \qquad s = 2b - t \geq 0. \tag{8.9}$$

No vertex of X_1 has J-degree one. Indeed, (8.5) would force its unique neighbour to have $\lambda = q + s$. An active unit-deficit point has $j \leq q - 1$, and an active deficit-two point has $j \leq q - 2$, because their fibres are nonempty. Thus every low point has $\lambda \leq q$, with equality possible only at an inactive unit-deficit point of degree q in J. It follows that $s = 0$, $t = 2b$, and the neighbour is such an inactive point. Equation (8.6) then says that all its q neighbours have deficit one and J-degree one. None can be inactive, since the inactive identity at such a leaf would read $q - 1 = 0$. The q active leaves have fibres of size $q - 1$, accounting for all $|H| = q^2 - q$ full-degree vertices. This leaves no fibre for the $t = 2b > 0$ deficit-two points, a contradiction. Equation (8.5) also excludes degree zero. Consequently

$$j_a \geq 2 \qquad (a \in X_1). \tag{8.10}$$

Let E be the number of full-degree vertices with two low neighbours, and put $J_2 = \sum_{w \in X_2} j_w$. The fibre count gives

$$E = q^2 - (t + 1)q + 2t - 2b + J_2. \tag{8.11}$$

Since $E \leq x(q - 2)$ and $q > 2b$, we get $x \geq q - t + 1$. Indeed, replacing x by $q - t$ would make the upper bound smaller than the right side of (8.11) by at least $q - 2b > 0$. Also $E \leq yq$ gives $y \geq q - t - 1$, and hence $x \leq q + 2b - t + 1$. Writing $k = x + t - q$, we have

$$1 \leq k \leq 2b + 1, \qquad x \geq q - 2b + 1. \tag{8.12}$$

For a deficit-two point w, equation (8.8) gives $d_X(w) = k - 1$. If it has a neighbour in Y, the inactive identity at that neighbour gives $j_w \leq s$; otherwise $j_w = k - 1$. Thus

$$j_w \leq 2b, \qquad \lambda_w \leq 2b + 1 \qquad (w \in X_2). \tag{8.13}$$

Every inactive point y satisfies $j_y \leq s+1$. For if $j_y > s+1$, it cannot have a neighbour with $\lambda = 1$. Such a neighbour would have to be a unit-deficit leaf. Active leaves have been excluded; an inactive leaf would force $j_y - 1 = s$ by (8.6). All j_y summands in that identity would consequently be at least one, giving $j_y \leq s$, a contradiction. In particular $\lambda_y = j_y \leq 2b + 1$.

If an active unit-deficit point a has $j_a > s+1$, it has no neighbour in Y, since its contribution $j_a - 1$ would exceed the inactive budget s. Every neighbour of a then has $\lambda \geq 2$, by (8.10) and the fact that any adjacent deficit-two point has positive J-degree. Equation (8.5) yields $2j_a \leq q + s$. Combining all cases, and using $q \geq 2b + 2$, gives

$$\max_{v \in L} \lambda_v \leq q/2 + b. \tag{8.14}$$

The weighted edge budget at edge deficit b is

$$\eta + \sum_{v \in L} (\delta_v - 1)(\delta_v + 2) + \sum_{v \in L} \delta_v j_v = 2q + 2(q + 2)b.$$

Here $\eta = 0$ and the middle sum is $4t$. Therefore

$$\sum_{v \in L} \lambda_v = 2(b + 1)q + 4b - 3t - J_2 \leq 2(b + 1)q + 4b \leq 3(b + 1)q. \tag{8.15}$$

Put $D = 4(b+1)$. By the sharper bound in (8.15), fewer than $q/2+1$ active unit points have $j_a > D$. Equations (8.12) and $q \geq 8b$ show that at least $q/4$ points of X_1 have $j_a \leq D$. Let

$$S = \{v \in L : \lambda_v \geq q/(4D)\}.$$

Equation (8.15) gives $|S| \leq 12D(b+1) = 48(b+1)^2$. Every one of the preceding $q/4$ active points has at least two neighbours in S. Otherwise (8.14) would give

$$\sum_{v \in N_J(a)} \lambda_v < q/2 + b + (D-1)\frac{q}{4D} < 3q/4 + b < q,$$

contradicting (8.5). Two vertices cannot have the same pair of neighbours in S, since J is C_4-free. Hence

$$\frac{q}{4} \leq \binom{|S|}{2} < 1152(b+1)^4,$$

contrary to $q \geq 8192(b+1)^4$. This excludes the mixed case for every $b \geq 1$ and completes the proof of Theorem 8.1.

8.4 One edge below the bound

Theorem 8.4 (Nucleus completion one edge below the bound)**.** *Let $q \geq 12$ be even. Suppose that F is a C_4-free graph on q^2+q vertices with $m_0 - 1$ edges, maximum degree at most $q+1$, and $\eta = 0$. Then F has exactly $q+1$ vertices of degree $q-1$, all other vertices have degree $q+1$, and the vertices of degree $q-1$ are independent.*

Adding one new vertex adjacent to these $q+1$ vertices gives the polarity graph of a projective plane of order q. Its absolute points are exactly the $q+1$ old vertices of degree $q-1$, and their common neighbour is the new vertex. This is the unique completion of F to a polarity graph of order q by adjoining one vertex and preserving the induced graph on $V(F)$.

Proof. Apply Lemmas 8.2 and 8.3 with $b = 1$, using the partner and fibre descriptions in Section 8. If $Y = \varnothing$, the asserted degree distribution and independence already follow. We exclude $Y \neq \varnothing$. In this case $t \leq 2$, and

$$x + y = 2q + 2 - 2t, \qquad |H| = q^2 - q - 2 + t.$$

Let E_T be the number of edges of the bipartite support graph on X_1, Y. Put $J_2 = \sum_{w \in X_2} j_w$. Counting full vertices by their fibres gives

$$E_T = q^2 - (t+1)q + 2t - 2 + J_2. \tag{8.16}$$

An active unit vertex has $j_a \leq q-1$, while an inactive unit vertex has $j_y \leq q$. An active deficit-two vertex has $j_w \leq q-2$. Consequently $\lambda_v \leq q$ for every low vertex.

The case $t = 1$. Write $X_2 = \{w\}$ and $j = j_w$. Equation (8.7) gives $j \geq 1$. An active unit vertex cannot have degree one in J, because (8.5) would require its unique neighbour to have $\lambda = q+1$. Thus every vertex of X_1 has $j_a \geq 2$ and support degree at most $q-2$. From

$$E_T = q(q-2) + j$$

we obtain $x \geq q+1$ and $y \geq q-1$. Since $x + y = 2q$, equality holds in both. Furthermore,

$$\sum_{y \in Y} j_y = q|Y| - E_T = q - j.$$

Equation (8.6) has right side one, so every $y \in Y$ has $j_y \geq 1$. Hence $j \leq 1$, and therefore $j = 1$ and all $y \in Y$ have $j_y = 1$. The unique neighbour a of w in J satisfies $\lambda_a = q - 1$ by (8.7). Thus $a \in X_1$ and $j_a = q - 1$. It has no neighbour in Y, because such a neighbour's equation (8.6) would contain the term $\lambda_a - 1 = q - 2 > 1$. All neighbours of a therefore have $\lambda \geq 2$. Equation (8.5) implies $q + 1 \geq 2(q - 1)$, contrary to $q \geq 12$.

The case $t = 2$. Here (8.6) has right side zero. Every neighbour of a vertex of Y is consequently a unit-deficit vertex of degree one in J. We first show that no active unit vertex a has $j_a = 1$. Such a vertex would have a unique neighbour y with $\lambda_y = q$, necessarily an inactive unit vertex with $j_y = q$. All q neighbours of y have deficit one and degree one in J. None is inactive, since its inactive equation would contain $\lambda_y - 1 = q - 1 > 0$. Thus $x \geq q$ and $y \leq q - 2$. But a has support degree $q - 1$, requiring at least $q - 1$ inactive unit vertices, a contradiction.

It follows that all active unit vertices have $j_a \geq 2$. For each $w \in X_2$, equation (8.7) has right side $q - 1 > 0$, so $j_w \geq 1$ and $J_2 \geq 2$. From

$$E_T = (q - 1)(q - 2) + J_2$$

and the support-degree bounds, we obtain $x \geq q$ and $y \geq q - 2$. Since $x + y = 2q - 2$, we have $x = q$, $y = q - 2$. There are no edges in J between Y and X, because all active unit vertices have $j_a \geq 2$ and every active deficit-two vertex has $\lambda_w \geq 2$. Therefore (8.8) gives $j_w = d_X(w) = 1$ for both deficit-two vertices. For either one, its unique low neighbour a has $\lambda_a = q - 2$ by (8.7). As $q - 2 > 2$, this neighbour lies in X_1 and has $j_a = q - 2$. Every neighbour of a is active and has $\lambda \geq 2$. Equation (8.5) now gives $q \geq 2(q - 2)$, again a contradiction.

The case $t = 0$. All low deficits are one, and

$$E_T = (q - 2)(q + 1), \qquad x + y = 2q + 2.$$

Since the active equation has right side $q+2$, every active vertex has $j_a \geq 2$. The support-degree bounds give $x \geq q + 1$ and $y \geq q - 1$. Thus x is one of $q + 1, q + 2, q + 3$. Every inactive vertex has $j_y \geq 1$, because its equation has right side two.

If $x = q + 3$ and $y = q - 1$, then

$$\sum_{y \in Y} j_y = q(q - 1) - E_T = 2 < |Y|,$$

a contradiction.

If $x = y = q + 1$, then

$$\sum_{a \in X_1} j_a = q(q + 1) - E_T = 2(q + 1),$$

so all active vertices have degree two in J. An inactive vertex of degree at least four cannot have an inactive neighbour: the latter's equation would contain a term at least three. All its neighbours would be active vertices of degree two, in which case its own inactive equation would make its degree equal to two. Hence every inactive vertex has degree at most three. For an active vertex, the sum of the degrees of its two low neighbours is then at most six, whereas its active equation requires that sum to be $q + 2 > 6$.

It remains that $x = q + 2$, $y = q$. In this case

$$\sum_{y \in Y} j_y = q + 2, \qquad \sum_{a \in X_1} j_a = 3q + 2. \tag{8.17}$$

Since every inactive vertex has positive degree, its degree is at most three. If an active vertex has degree at least four, it has no inactive neighbour, and all its neighbours have degree at least two. Its active equation therefore bounds its degree by

$$d_* = (q+2)/2.$$

The same bound also holds for every other low vertex. Let n_2 be the number of active vertices of degree two in J. Equation (8.17) yields

$$n_2 = 4 + \sum_{\substack{a \in X_1 \\ j_a \geq 4}} (j_a - 3). \tag{8.18}$$

For each of these n_2 vertices, the two neighbouring degrees sum to $q+2 = 2d_*$. Both neighbours must therefore have degree exactly d_*, and both lie in X_1. There are at most five such heavy vertices, since their degree sum is at most $3q+2 < 6d_*$. No two degree-two vertices have the same heavy neighbour pair, by C_4-freeness. Consequently $n_2 \leq \binom{5}{2} = 10$. On the other hand, $n_2 \geq 4$ ensures the existence of at least two heavy vertices. Their contributions to (8.18) give

$$n_2 \geq 4 + 2(d_* - 3) = q \geq 12,$$

the final contradiction.

We have proved $Y = \varnothing$. The pure deficit-two argument in Section 8.2 gives the asserted degree distribution, completion, and uniqueness. □

9 Sharp supersaturation from polarity structure

Proposition 2.3 gives a lower bound for the cycles created by each new edge after a point deletion. Its edge-disjointness assertion allows missing edges of the containing polarity graph to be accounted for separately. We first apply this to modifications of an extremal core, and then to an arbitrary core supplied by Theorem 1.3.

9.1 Modifications of an extremal core

The old edges supporting the four-cycles created by one added edge in a polarity graph are pairwise disjoint. This permits several edge deletions before the sharp lower bound is lost.

Theorem 9.1. *Let $q \geq 10^{15}$ be even. Suppose G has q^2+q vertices and m_0+1 edges, and F is a C_4-free graph on the same vertex set with m_0 edges. If*

$$b = |E(F) \setminus E(G)| \leq q-2,$$

then

$$N_{C_4}(G) \geq (b+1)(q-1-b) \geq q-1.$$

The same conclusion holds for every $q \geq 2$, retaining $0 \leq b \leq q-2$, when F is already known to be an absolute-point deletion of a polarity graph.

Proof. By Theorem 7.6, write $F = \mathcal{P} - p$, with p absolute. Put $D = E(F) \setminus E(G)$ and $A = E(G) \setminus E(F)$. The edge counts give $|D| = b$ and $|A| = b+1$. For each $e \in A$, Proposition 2.3 supplies at least $q-1$ four-cycles in $F+e$, whose old edge sets are pairwise disjoint. Deletion of D destroys at most b of them. Every surviving cycle contains exactly one member of A. The cycles counted for different members of A are therefore distinct, and summation gives the first inequality. The second follows from

$$(b+1)(q-1-b) - (q-1) = b(q-2-b) \geq 0.$$

□

Corollary 9.2. *For even $q \geq 10^{15}$, a graph on q^2+q vertices with m_0+1 edges and fewer than $q-1$ four-cycles must satisfy*

$$|E(F) \setminus E(G)| \geq q-1$$

for every C_4-free graph F on the same vertex set with m_0 edges. Its four-cycles have no common edge.

Proof. The distance statement is the contrapositive of Theorem 9.1. If one edge met every four-cycle, deleting it would give a graph F as in that theorem with $b=0$. □

9.2 Counting the edges outside a polarity host

Proposition 9.3. *Let $q \geq 4$ be even, and let G have $N-1$ vertices and m_0+1 edges. Suppose that G has a spanning C_4-free subgraph F with m_0-b edges, where $0 \leq b \leq q-2$, and that F is a spanning subgraph of $K = P-p$ for a polarity graph P of a projective plane of order q. Then*

$$N_{C_4}(G) \geq q-1. \tag{9.1}$$

Proof. The even-order polarity bound (Lemma 2.6) gives $e(P) \leq M$ and $d_P(p) \geq q$. Thus $e(K) \leq m_0$. Put

$$d = m_0 - e(K), \qquad c = |E(K) \setminus E(G)|, \qquad a = |E(G) \setminus E(K)|.$$

Since $F \subseteq G \cap K$,

$$0 \leq d \leq b, \qquad 0 \leq c \leq b-d, \qquad a = c+d+1. \tag{9.2}$$

Fix $e \in E(G) \setminus E(K)$. If $K+e$ has at least L_0 four-cycles, Proposition 2.3 shows that deleting the c absent host edges destroys at most c of them. At least $L_0 - c$ such cycles remain in G. Each has exactly one edge outside K, namely e. The collections obtained for different e are therefore disjoint. We may sum these contributions over all a edges outside K.

If $d=0$, equality in $e(K) = e(P) - d_P(p) \leq M - q = m_0$ forces $e(P) = M$ and $d_P(p) = q$. In particular p is absolute. Proposition 2.3 and (9.2) give

$$N_{C_4}(G) \geq (c+1)(q-1-c) = q-1+c(q-2-c) \geq q-1,$$

because $0 \leq c \leq b \leq q-2$.

If $d \geq 1$, then $c \leq q-3$ and $a \geq c+2$. The arbitrary-puncture bound gives

$$N_{C_4}(G) \geq (c+d+1)(q-2-c) \geq (c+2)(q-2-c) = q-1+(c+1)(q-3-c) \geq q-1.$$

All lower bounds being summed are positive in the indicated ranges. This proves (9.1). □

9.3 The edge-transversal consequence

Proof of Corollary 1.4. Choose an edge transversal and enlarge it, if necessary, to an edge set T of size exactly t. There are enough edges for this enlargement under (1.2). Its deletion leaves a C_4-free spanning graph

$$F = G - T, \qquad e(F) = m_0 - (t-1).$$

Apply the fixed-deficit stability theorem (Theorem 1.3) with $b = t-1$. The required threshold is exactly $q \geq 10^{18}(b+1)^{16} = 10^{18}t^{16}$. It supplies a polarity graph P and a point p such that $F \subseteq P-p$ as a spanning subgraph. Also $b = t-1 \leq q-2$ under (1.2). Proposition 9.3 now proves the result. □

Corollary 9.4 (A necessary transversal size for a sharp counterexample)**.** *Let $q \geq 10^{18}$ be even, and let G have q^2+q vertices and m_0+1 edges. If $N_{C_4}(G) \leq q-2$, then*

$$\tau_4(G) > \left(\frac{q}{10^{18}}\right)^{1/16}, \qquad \textit{equivalently} \qquad \tau_4(G) \geq \left\lfloor \left(\frac{q}{10^{18}}\right)^{1/16} \right\rfloor + 1. \tag{9.3}$$

Proof. Set $t = \lfloor (q/10^{18})^{1/16} \rfloor \geq 1$. Then $q \geq 10^{18} t^{16}$. If $\tau_4(G) \leq t$, Corollary 1.4 contradicts $N_{C_4}(G) \leq q-2$. □

10 Further questions

Question 10.1. Does $h(q^2+q) = q-1$ hold for all sufficiently large powers of two q?

Theorem 1.1 gives the order of magnitude on this sequence. For even $q \geq 10^{18}$, Corollaries 9.2 and 9.4 impose two simultaneous conditions on a graph G with n vertices, m_0+1 edges and at most $q-2$ four-cycles:

$$|E(F) \setminus E(G)| \geq q-1 \quad \text{for every } C_4\text{-free } F \text{ on the same vertices with } m_0 \text{ edges},$$
$$\tau_4(G) > \left(\frac{q}{10^{18}}\right)^{1/16}.$$

Thus the sharp question concerns graphs far from every extremal core whose four-cycles require a growing edge transversal.

If $Q = N_{C_4}(G) \leq q-2$, choosing one edge from each four-cycle gives a transversal of size at most $q-2$. A minimum transversal T therefore leaves a C_4-free spanning graph

$$F = G - T, \qquad e(F) = m_0 - b, \qquad b = |T| - 1 \leq q-3.$$

This deficit may be proportional to q, whereas Theorem 1.3 applies under $q \geq 10^{18}(b+1)^{16}$. The cores arising in this way carry additional information: for every $e \in T$, minimality supplies a four-cycle whose only edge in T is e, hence a three-edge path in F joining the endpoints of e. Extending the polarity embedding to these cores, or exploiting the incidence of their associated paths to obtain a smaller transversal, would address the remaining range.

A second question is the quantitative dependence in the stability theorem. Its proof supplies explicit polynomial sufficient bounds for the degree restriction and the geometric embedding. Determining the largest deficit interval in which every core is contained in a one-point deletion of a polarity graph would also clarify the relation between structural stability and the sharp supersaturation value.

A Incidence completion

We complete partial incidence structures with a fixed number of missing indexed lines. Deleting one indexed line and applying an integer packing identity supplies the minimum line size required by the completion argument.

A.1 The finite-geometric inputs

We use the following form of Metsch's clique theorem [12, Result 2.2], also stated explicitly in [10, Theorem 2.7]. Let $\lambda_1, \lambda_2, \mu$ be positive integers. Suppose a connected graph has between λ_1 and λ_2 common neighbours at every adjacent pair, at most μ common neighbours at every nonadjacent pair, and maximum degree at most k. If an integer $s \geq 1$ satisfies

$$2\lambda_1 - \lambda_2 > (2s-1)(\mu-1) - 1, \qquad k < (s+1)(\lambda_1+1) - \frac{s(s+1)}{2}(\mu-1), \tag{A.1}$$

then the maximal cliques of size at least $\lambda_1 + 2 - (s-1)(\mu-1)$ cover every edge exactly once, and each vertex lies in at most s of them. We apply this statement separately to connected components.

The embedding theorem of Beutelspacher and Metsch [2], in the form of [1, Theorem 2.2.1], says the following. The order of a finite linear space is its maximum point degree minus one. A finite linear space with maximum point degree $q+1$ and minimum line size at least $q+1-a$ embeds in a projective plane of order q whenever

$$4q > f(a) := 6a^4 + 9a^3 + 19a^2 + 8a.$$

Indeed, the semiaffine parameter, defined as the maximum of $r_P - k_\ell$ over nonincident point–line pairs, is at most a. In the application below every point has degree exactly $q+1$. The input is applied after every pair of points has been assigned its unique line.

A.2 Packing and completion

The packing argument uses indexed subsets as lines, allowing empty or singleton lines after restriction. Before applying the embedding theorem, we construct a finite linear space in which every pair of points lies on exactly one line and every line has at least two points.

Lemma A.1 (The general packing identity and the shortest line)**.** *Let $s \geq 1$ and $t \geq 0$ be integers, put $N = q^2+q+1$, and suppose a finite partial linear space has*

$$m = N - s, \qquad b = N - s(q+1) + \binom{s}{2} - t = q^2 - (s-1)q + \beta, \qquad \beta = \frac{(s-1)(s-2)}{2} - t,$$

indexed lines and points respectively. Every point is to lie on exactly $q+1$ indexed lines. Distinct indexed lines intersect in at most one point. Write r_v for their sizes, and u for the number of uncovered unordered point pairs. Then

$$2u + \sum_v (r_v - q + s - 1)(r_v - q + s - 2) = 2tq + C_{s,t}, \tag{A.2}$$

where

$$C_{s,t} = \beta^2 + (2s-3)\beta - (s-1)^2(s-2).$$

If $q \geq s^4 + 2s^3 + 1$, then every indexed line has size at least

$$r_v \geq q - s + 1 - t. \tag{A.3}$$

In particular, for fixed s, t and sufficiently large q, the number of uncovered pairs is at most $(t+1)q$.

Proof. We have $\sum_v r_v = b(q+1)$ and $\sum_v r_v(r_v-1) = b(b-1) - 2u$. Expanding the left side of (A.2) gives the stated identity. Every summand is a product of two consecutive integers, hence is nonnegative. This proves the last assertion once $q > |C_{s,t}|$.

Fix an indexed line of size r, delete this line and all its points, and restrict every other indexed line to the remaining points. The new counts are $m' = N - s - 1$ and $b' = b - r$. Each remaining point still has degree $q+1$. The pair count and Cauchy's inequality give

$$\frac{b'^2(q+1)^2}{m'} \leq \sum_v (r'_v)^2 \leq b'(b'+q).$$

If $b' = 0$, the claimed lower bound on r is immediate. Otherwise

$$b' \leq \frac{q(q^2+q-s)}{q+s+1} = q^2 - sq + s^2 - \frac{s^2(s+1)}{q+s+1} < q^2 - sq + s^2.$$

Suppose $b' > q^2 - sq + s(s-1)/2$, and write $b' = q^2 - sq + \beta'$. Then

$$\frac{s(s-1)}{2} + 1 \le \beta' \le s^2 - 1.$$

The same packing expansion, with $s+1$ in place of s, is

$$2u' + \sum_v (r'_v - q + s)(r'_v - q + s - 1) = [s(s-1) - 2\beta']q + \beta'^2 + (2s-1)\beta' - s^2(s-1).$$

Its right side is at most $-2q + s^4 + 2s^3 < 0$, while its left side is nonnegative. Thus $b' \le q^2 - sq + s(s-1)/2$, and subtracting this from b gives (A.3). □

Theorem A.2 (Fixed-parameter incidence completion)**.** *Fix integers $s \ge 1$, $t \ge 0$, and $D \ge 0$. For all sufficiently large integers q, the following holds. Let a finite partial linear space have*

$$|\mathcal{L}| = q^2 + q + 1 - s, \qquad |P| = q^2 + q + 1 - s(q+1) + \binom{s}{2} - t,$$

every point on exactly $q+1$ indexed lines, and distinct indexed lines intersecting in at most one point. If its uncovered-pair graph U has $\Delta(U) \le D$, then U is empty, and the whole incidence structure embeds in a projective plane of order q. The lower bound on q depends only on s, t, D. The inequalities (A.4)–(A.11) below give sufficient numerical conditions.

Proof. We retain the notation $b, \beta, C_{s,t}$ from Lemma A.1 and define

$$\begin{aligned} &K = s + t, && A_0 = 2(K+1), && D_* = \max\{D, 2\}, \\ &B_0 = A_0(t+1), && Y_0 = 2B_0, && \mu = K^2, \\ &C_1 = s + Y_0 + (K-1)(\mu - 1), && && a_0 = s + Y_0 + (K-1)\max\{t-1, 0\}. \end{aligned}$$

All these constants are independent of q. It suffices to assume

$$q \ge s^4 + 2s^3 + 1, \qquad q > |C_{s,t}|, \tag{A.4}$$

$$q > A_0 \binom{B_0 + 1}{2} D_*, \tag{A.5}$$

$$q > 2\{2(s + 2 + Y_0) + K^2 + (2K-1)(\mu - 1) + 1\}, \tag{A.6}$$

$$q > 2(K+1)(s + 1 + Y_0) + K(K+1)(\mu - 1), \tag{A.7}$$

$$q > \max\{s + Y_0,\ a_0,\ \beta + (t+1)(K-1)\}, \tag{A.8}$$

$$q > 2(K+1)(C_1 + 1), \tag{A.9}$$

$$q > \max\{2K,\ K + Y_0(Y_0 - 1)\}, \tag{A.10}$$

$$4q > f(a_0) = 6a_0^4 + 9a_0^3 + 19a_0^2 + 8a_0. \tag{A.11}$$

The first two conditions control packing and the deletion of badly separated line pairs. The next four give an exact clique decomposition at each remaining index. The last two permit the geometric embedding and restoration of the deleted indices.

By the packing lemma, $e(U) \le (t+1)q$ and every line has at least $q + 1 - K$ points. Write $S_v \subseteq P$ for an indexed line and $d_v = q + 1 - |S_v|$. A nonempty S_v meets exactly $1 + q|S_v|$ indexed lines, including itself. Since there are $N - s$ indices, $|S_v| \le q$ for $q > s$. Consequently

$$1 \le d_v \le K. \tag{A.12}$$

Removing the bounded set of badly separated line pairs. Let R be the graph on the line indices in which adjacency means disjointness. For disjoint point sets X, Y, let $h(X, Y)$ count uncovered pairs with one end in each. Direct inclusion–exclusion gives

$$d_R(v) = qd_v - s, \tag{A.13}$$

$$|N_R(v) \cap N_R(w)| = q - s - 1 + (d_v - 1)(d_w - 1) - h(S_v, S_w) \qquad (vw \in E(R)). \tag{A.14}$$

For distinct nonadjacent indices v, w, write $S_v \cap S_w = \{p\}$. Then

$$|N_R(v) \cap N_R(w)| = d_v d_w - s - h(S_v \setminus \{p\}, S_w \setminus \{p\}). \tag{A.15}$$

Thus nonadjacent codegrees are at most μ.

Call $vw \in E(R)$ bad if $h(S_v, S_w) \geq q/A_0$, and denote its set of cross-holes by F_{vw}. Two different such sets intersect in at most D_* pairs. If the edges share an index, every common hole is incident with the at most one intersection point of their other two lines, so the bound is D. With four distinct indices, the two possible ways to match the line pairs give at most two common holes. If there were $B_0 + 1$ bad edges, inclusion–exclusion would give

$$e(U) \geq (t+1)q + q/A_0 - \binom{B_0+1}{2} D_* > (t+1)q$$

by (A.5). Hence there are at most B_0 bad edges. Delete their endpoints, a set Y of size at most Y_0, and let R_0 be the induced graph on the remaining indices. For its vertices,

$$qd_v - s - Y_0 \leq d_{R_0}(v) \leq qd_v - s. \tag{A.16}$$

Its adjacent codegrees lie between

$$\lambda_1 = \lfloor (1 - 1/A_0)q \rfloor - (s + 1 + Y_0), \qquad \lambda_2 = q + K^2,$$

and its nonadjacent codegrees are at most μ.

The exact clique count at a line index. Apply Metsch's theorem, in the form stated in (A.1), with clique-count parameter K and maximum-degree bound Kq. The bound $\lambda_1 \geq (1 - 1/A_0)q - (s + 2 + Y_0)$ and $1 - 2/A_0 \geq 1/2$ show that (A.6) implies its first inequality. Since $(K+1)(1 - 1/A_0) - K = 1/2$, (A.7) implies its second inequality. The theorem applies componentwise; there are no isolated vertices by (A.8). The maximal cliques selected by the theorem form a family $\mathcal{K}$ covering each edge once, with at most K cliques at any vertex, and

$$|J| \geq (1 - 1/A_0)q - C_1 \qquad (J \in \mathcal{K}). \tag{A.17}$$

On the other hand, their original line sets are pairwise disjoint, so $|J|(q + 1 - K) \leq b$. Since

$$(q + t + 1)(q + 1 - K) - b = q - \beta - (t+1)(K-1),$$

we have

$$|J| \leq q + t \tag{A.18}$$

by (A.8).

Let k_v count cliques containing v. Edge uniqueness implies $d_{R_0}(v) = \sum_{J \ni v} (|J| - 1)$. If $k_v \leq d_v - 1$, then

$$d_{R_0}(v) \leq (d_v - 1)(q + t - 1) < qd_v - s - Y_0$$

by (A.8), a contradiction. If $k_v \geq d_v + 1$, then

$$d_{R_0}(v) \geq (d_v + 1)((1 - 1/A_0)q - C_1 - 1) > qd_v - s$$

by (A.9), also a contradiction. Thus

$$k_v = d_v. \tag{A.19}$$

Put $b_v = |N_R(v) \cap Y| \le Y_0$. The exact degree identity now is

$$\sum_{J \ni v} (q + 1 - |J|) = s + b_v.$$

Every summand is at least $1 - t$, by (A.18), and there are at most K summands. It follows that every clique has size at least $q + 1 - a_0$.

Embedding a genuine linear space and restoring the indices. On the point set $I \setminus Y$, include the original restricted point pencils

$$L_p = \{v \in I \setminus Y : p \in S_v\} \qquad (p \in P)$$

as lines, and also include the cliques of $\mathcal{K}$ as lines. Every pair of indices lies on exactly one line: intersecting original lines use their unique common point, and disjoint ones use their unique clique. All these lines are distinct, any two intersect in at most one point, and every point v has degree $|S_v| + d_v = q + 1$. Original pencil lines have size at least $q + 1 - Y_0$, and clique lines have size at least $q + 1 - a_0$. Since $a_0 \ge Y_0$, the resulting genuine linear space embeds in a projective plane of order q by the Beutelspacher–Metsch theorem [1, Theorem 2.2.1] and (A.11). The embedded incidences are exact: each assigned point already has its $q + 1$ prescribed lines, exhausting its ambient pencil.

Write p_v for the assigned point of $v \in I \setminus Y$ and ℓ_p for the plane line of L_p. Each ℓ_p contains at most Y_0 points outside the assigned set. For a deleted index $v \in Y$, the at least $q + 1 - K$ distinct lines ℓ_p, $p \in S_v$, intersect pairwise outside that set. Condition (A.10) permits Lemma 2.5 to assign their common point to v. Two deleted indices cannot receive the same point: their line families share at most one member, and together would have at least $2(q + 1 - K) - 1 > q + 1$ lines, again by (A.10). Thus all original indexed lines have distinct point images. For each $p \in P$, its $q + 1$ prescribed indexed-line points now fill ℓ_p. The intersection of two such plane lines is therefore an assigned indexed-line point, so the original two points are covered by a line. Hence U is empty. Dualizing the plane embeds the entire original incidence structure. The numerical conditions involve only s, t, D, proving the theorem. □

B Polarity extension

B.1 A matrix criterion for unique extension

Füredi's partial-symmetry criterion, recalled in [9, Lemma 2.1], assumes symmetry of every entry involving a sufficiently large prescribed set of indices. The following lemma starts with a symmetric principal block and determines the omitted point–line matching.

Lemma B.1 (Extending symmetric incidence across finitely many omissions)**.** *Let Π be a projective plane of order q, put $N = q^2 + q + 1$, and let $s \ge 1$ with $q \ge 2s$. Suppose $p_1, \dots, p_{N-s}$ are distinct points and $\ell_1, \dots, \ell_{N-s}$ are distinct lines such that*

$$p_i \in \ell_j \quad \Longleftrightarrow \quad p_j \in \ell_i \qquad (1 \le i, j \le N - s).$$

Then there is a unique polarity π of Π satisfying $\pi(p_i) = \ell_i$ for every i.

Proof. Let Q be the s omitted points and Λ the s omitted lines. With the prescribed old point and line orders, write the full incidence matrix of Π as

$$M = \begin{pmatrix} B & X \\ Y^{\mathsf{T}} & Z \end{pmatrix}.$$

Thus B is the old-point–old-line incidence matrix, X records incidences of old points with Λ, and Y records incidences of old lines with Q. The hypothesis says $B = B^{\mathsf{T}}$. Since every line and point has $q+1$ incidences and every two distinct points or lines have exactly one common incidence,

$$MM^{\mathsf{T}} = M^{\mathsf{T}}M = qI + J.$$

Comparing the upper-left blocks therefore gives

$$XX^{\mathsf{T}} = YY^{\mathsf{T}}. \tag{B.1}$$

For every $\lambda \in \Lambda$, there is an old point on λ and on no other omitted line. Indeed, among its $q+1$ points, at most s lie in Q and at most $s-1$ are intersections with other omitted lines, leaving at least $q - 2s + 2 > 0$ choices. Consequently X contains every standard unit row of $\{0,1\}^s$. For the corresponding s rows of Y, (B.1) gives squared norm one and pairwise inner product zero. They are therefore the s different standard unit rows, in some order. After a unique permutation of the columns of Y, these rows agree with the chosen rows of X. Inner products with these unit rows then force every row to agree, so $Y = X$ after that permutation. This fixes the required bijection from Q to Λ.

Order the omitted points accordingly. We now have

$$M = \begin{pmatrix} B & X \\ X^{\mathsf{T}} & Z \end{pmatrix}.$$

The upper-right blocks of $MM^{\mathsf{T}} = M^{\mathsf{T}}M$ give

$$BX + XZ^{\mathsf{T}} = BX + XZ, \qquad X(Z - Z^{\mathsf{T}}) = 0.$$

The unit rows show that X has column rank s, and hence $Z = Z^{\mathsf{T}}$. Thus the entire reordered incidence matrix is symmetric. The resulting point-to-line bijection is a polarity extending all prescribed assignments. The omitted bijection was forced by the unit rows, proving uniqueness. □

B.2 A bounded incidence defect

The symmetric incidence assignment extends even when the clean point set omits a bounded number of points in addition to those on the missing lines.

Lemma B.2 (Polarity extension with additional omitted points)**.** *Let Π be a projective plane of order q, and put $N = q^2+q+1$. Let $s \geq 1$ and $t \geq 0$ be integers, and suppose $q > (s+t+2)^2$. Let F be a C_4-free graph on a set V_0 of $N - s$ vertices, and let $C \subseteq V_0$ satisfy*

$$d_F(u) = q+1 \quad (u \in C), \qquad |C| \geq N - s(q+1) + \binom{s}{2} - t.$$

Suppose there are distinct points p_u for $u \in C$ and distinct lines ℓ_v for $v \in V_0$ in Π, with exact incidences

$$p_u \in \ell_v \quad \Longleftrightarrow \quad uv \in E(F) \qquad (u \in C,\ v \in V_0).$$

Then the assignment extends to distinct points p_v for every $v \in V_0$ and to a polarity π of Π with $\pi(p_v) = \ell_v$. In particular, F is a spanning subgraph of $H - Q$, where H is the loopless polarity graph of π and Q is the set of s unassigned plane points.

Proof. Write $a = s + t$, and identify C with its point images. There are exactly s missing lines, denoted by Λ. Every point of C is already incident with $q+1$ old lines, so it lies on no missing line. The union of s lines has at least $s(q+1) - \binom{s}{2}$ points. Hence at most t points outside C avoid all missing lines. Every old line therefore contains at most $s + t = a$ points outside C. Consequently

$$|N_F(v) \cap C| = |\ell_v \cap C| \geq q + 1 - a \qquad (v \in V_0). \tag{B.2}$$

For $v \in V_0 \setminus C$, consider the lines ℓ_u indexed by $u \in N_F(v) \cap C$. Any two intersect outside C; an intersection in C would give their two indices the common neighbours v and a point of C, contrary to C_4-freeness. There are at least $q + 1 - a$ such lines, and each has at most a points outside C. Since $q - a > a(a-1)$, the pencil lemma (Lemma 2.5) makes them concurrent at a point p_v outside C. If $p_v = p_w$ for distinct $v, w \in V_0 \setminus C$, their two families would be contained in a pencil of $q+1$ lines, while having at most one common line. This would imply $2(q+1-a) - 1 \leq q+1$, contrary to our bound on q. Thus all $N - s$ old vertices have distinct point images. Let Q consist of the s remaining plane points.

For $u \in C$, the $q+1$ distinct images of its neighbours exhaust ℓ_u. Hence ℓ_u avoids Q, and its incidences with all old points are exactly the adjacencies at u.

We first prove symmetry on old points:

$$p_v \in \ell_w \quad \Longleftrightarrow \quad p_w \in \ell_v. \tag{B.3}$$

The claim holds if either index is in C. Otherwise put $A_v = N_F(v) \cap C$ and $A_w = N_F(w) \cap C$. The incidence $p_v \in \ell_w$ implies that there is no edge between A_v and A_w: all lines ℓ_u, $u \in A_v$, meet ℓ_w at $p_v \notin C$. Conversely, if there is no such edge and $p_v \notin \ell_w$, those at least $q + 1 - a$ lines have distinct intersections with ℓ_w, all outside C. This contradicts the bound a on the number of outside points. The absence-of-edges condition is symmetric, proving (B.3). For an edge vw of F with $v, w \notin C$, an edge between A_v and A_w would produce a four-cycle. Thus every old edge is respected by the incidence assignment.

We now have $N - s$ distinct old points and $N - s$ distinct old lines with symmetric incidences. Since $q \geq 2s$, Lemma B.1 extends this assignment to a unique polarity of the plane. The old-edge conclusion proved above gives $F \subseteq H - Q$. □

B.3 Recognition from full-degree vertices

Theorem B.3 (Intrinsic recognition across any fixed deficient order)**.** *Fix integers $s \geq 1$ and $t \geq 0$. For all sufficiently large integers q, let F be a C_4-free graph on $N - s = q^2 + q + 1 - s$ vertices. If at least*

$$N - s(q+1) + \binom{s}{2} - t \tag{B.4}$$

vertices have degree exactly $q+1$, then there is a polarity graph H of a projective plane of order q and a set Q of s points such that F is a spanning subgraph of $H - Q$. No maximum-degree assumption and no uncovered-pair assumption is needed. In particular the conclusion forces $\Delta(F) \leq q + 1$.

Proof. Choose exactly the number in (B.4) of degree-$(q+1)$ vertices, and call their set C. Take C as the point set and $N_F(v) \cap C$, indexed by all $v \in V(F)$, as the lines. Point degrees are $q+1$, and C_4-freeness gives the two intersection conditions for a partial linear space. Put $K = s + t$. Lemma A.1 gives

$$r_v := |N_F(v) \cap C| \geq q + 1 - K \quad \text{and hence} \quad \delta_v := q + 1 - d_F(v) \leq K \qquad (v \in V(F)).$$

The deficits may be negative; no nonnegativity is used. For $x \in C$, there are at most K neighbours outside C, and all neighbours inside C have deficit zero. Consequently

$$a_x := \sum_{v \in N_F(x)} \delta_v \leq K^2.$$

Counting length-two paths from this full-degree vertex in the $N - s$ vertex graph gives the exact identity $u_x = a_x - s$. Thus the uncovered-pair graph restricted to C has maximum degree at most K^2. Theorem A.2, with $D = K^2$, embeds the complete indexed incidence system in a projective plane of order q, with exact incidences. For $q > (s + t + 2)^2$, the polarity-extension lemma (Lemma B.2) now applies to F, C with these same parameters. It supplies the claimed graph H and omitted set Q. □

References

[1] A. Beutelspacher. Linear spaces: history and theory. *Le Matematiche*, 45(1):25–38, 1990. Theorem 2.2.1. Published article.

[2] A. Beutelspacher and K. Metsch. Embedding finite linear spaces in projective planes, II. *Discrete Mathematics*, 66(3):219–230, 1987. doi:10.1016/0012-365X(87)90098-7.

[3] A. A. Bruen. Baer subplanes and blocking sets. *Bulletin of the American Mathematical Society*, 76:342–344, 1970. doi:10.1090/S0002-9904-1970-12470-3.

[4] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983. doi:10.1007/BF02579292.

[5] P. Erdős and M. Simonovits. Cube-supersaturated graphs and related problems. In *Progress in Graph Theory (Waterloo, Ont., 1982)*, pages 203–218. Academic Press, Toronto, ON, 1984. Published chapter.

[6] F. A. Firke, P. M. Kosek, E. D. Nash, and J. Williford. Extremal graphs without 4-cycles. *Journal of Combinatorial Theory, Series B*, 103(3):327–336, 2013. doi:10.1016/j.jctb.2013.01.001.

[7] Z. Füredi. Graphs without quadrilaterals. *Journal of Combinatorial Theory, Series B*, 34(2):187–190, 1983. doi:10.1016/0095-8956(83)90018-7.

[8] J. He, J. Ma, and T. Yang. Some extremal results on 4-cycles. *Journal of Combinatorial Theory, Series B*, 149:92–108, 2021. doi:10.1016/j.jctb.2021.01.007.

[9] J. He, J. Ma, and T. Yang. Some exact results on 4-cycles: stability and supersaturation. *CSIAM Transactions on Applied Mathematics*, 4(1):74–128, 2023. doi:10.4208/csiam-am.SO-2021-0006.

[10] B. Kivva. On the spectral gap and the automorphism group of distance-regular graphs. *Journal of Combinatorial Theory, Series B*, 149:161–197, 2021. Theorem 2.7. Published version.

[11] J. Ma and T. Yang. Upper bounds on the extremal number of the 4-cycle. *Bulletin of the London Mathematical Society*, 55(4):1655–1667, 2023. doi:10.1112/blms.12810.

[12] K. Metsch. A characterization of Grassmann graphs. *European Journal of Combinatorics*, 16(6):639–644, 1995. Result 2.2. doi:10.1016/0195-6698(95)90045-4.