Three-uniform Berge-C_4 saturation: linear bounds with coefficients $5/8$ and $3/4$, the cone spectrum and exact small values

Dev D. Goyal

31 October 2026; revised 1 November 2026

Abstract

Let $s(n) = \mathrm{sat}_3(n, \text{Berge-}C_4)$ be the least number of hyperedges of a three-uniform hypergraph on n vertices that contains no Berge four-cycle but contains one after any absent triple is added. Neither the hypergraph nor the added triple is required to be linear. English, Graber, Kirkpatrick, Methuku and Sullivan [1] showed $s(n) \le n-2$ for $n \ge 7$, and English, Gerbner, Methuku and Tait [3] showed that $s(n)$ grows linearly. We prove that for every $n \ge 9$

$$\max\left(\left\lceil\tfrac{n-1}{2}\right\rceil, \left\lceil\tfrac{5n-3623}{8}\right\rceil, \left\lceil\tfrac{5n+89}{8} - \tfrac{\sqrt{270n+1891}}{4}\right\rceil\right) \le s(n) \le \left\lceil\tfrac{3n-9}{4}\right\rceil,$$

so that $5/8 \le \liminf s(n)/n \le \limsup s(n)/n \le 3/4$. We also classify the saturated hypergraphs in which one vertex lies in every hyperedge (cones) for $n \ge 5$ and determine their sizes: exactly the integers from $\lceil(3n-5)/4\rceil$ to $n-2$, together with $n-1$ when $n \equiv 1 \pmod 3$. Hence every integer in $[\lceil(3n-9)/4\rceil, n-2]$ is the size of a saturated hypergraph. A vertex r such that every other vertex has both a one-edge and a two-edge Berge path to r forces at least $3(n-1)/4$ hyperedges, with equality exactly for bouquets of three-edge blocks. Exhaustive certificates give $s(n)$ for $n \le 13$. Lean 4 proves the upper bound for every $n \ge 9$, for actual hypergraphs, Berge cycles and saturation under every absent triple, and the values $s(n)$ for $n \le 5$; the lower bound, the cone results and the values for $6 \le n \le 13$ have conventional proofs or checked certificates.

1 Definitions and results

All hypergraphs are simple and three-uniform, and the vertex set may contain isolated vertices. A *Berge-C_4* in H consists of four distinct vertices v_0, v_1, v_2, v_3 and four distinct hyperedges e_0, e_1, e_2, e_3 of H with $\{v_i, v_{i+1}\} \subseteq e_i$, indices modulo 4; a hyperedge may contain further cycle vertices. A *Berge path of length ℓ* has distinct vertices $v_0, \dots, v_\ell$ and distinct hyperedges $e_1, \dots, e_\ell$ with $\{v_{i-1}, v_i\} \subseteq e_i$. A hypergraph H is *saturated* if it contains no Berge-C_4 but $H + f$ contains one for every three-element set $f \notin E(H)$. This is the unrestricted notion of [1]: f may share two vertices with a hyperedge of H. We write $s(n)$ for the least number of hyperedges of a saturated hypergraph on n vertices.

English, Graber, Kirkpatrick, Methuku and Sullivan [1, Theorem 5.7] gave the upper bound $s(n) \le n-2$ for $n \ge 7$, and English, Gerbner, Methuku and Tait [3] proved $\mathrm{sat}_k(n, \text{Berge-}F) = O(n)$ for every graph F and $3 \le k \le 5$. The elementary lower bound $s(n) \ge \lceil(n-1)/2\rceil$ follows from connectivity (Lemma 7). Wang and Zhang [5] studied the different *linear* saturation problem, in which H is linear and only additions preserving linearity are tested.

Theorem 1. *For every $n \ge 4$,*

$$5n \le 8s(n) + 3623 \qquad \textit{and} \qquad 5n \le 8s(n) + 2\sqrt{432\,s(n) + 1} + 19, \tag{1}$$

and $s(n) \geq \lceil (n-1)/2 \rceil$. *For every* $n \geq 9$, $s(n) \leq \lceil (3n-9)/4 \rceil$. *Consequently, for* $n \geq 9$, $s(n)$ *lies between the maximum displayed in the abstract and* $\lceil (3n-9)/4 \rceil$, *and* $5/8 \leq \liminf s(n)/n \leq \limsup s(n)/n \leq 3/4$.

The square-root form of (1) inverts to the third term of the abstract (Section 4); it is the better of the two for $n \leq 12751$. We do not know whether $s(n)/n$ converges. The constants in (1) are not optimized.

A *cone* is a hypergraph with a vertex r in every hyperedge. Its *link* is the graph G on $V \setminus \{r\}$ with $xy \in E(G)$ when $rxy \in E(H)$. Write $D(p,q)$ for the double star with adjacent centres carrying $p \geq 1$ and $q \geq 1$ leaves, $K_{1,t}$ for the star with t edges and K_3 for the triangle.

Theorem 2. *Let* $n \geq 5$ *and let* H *be a cone on* n *vertices with link* G. *Then* H *is saturated if and only if either* G *has no isolated vertex and its components are double stars, triangles and at most one star* $K_{1,t}$ *with* $t = 1$ *or* $t \geq 3$, *or* G *has exactly one isolated vertex and all other components are triangles. The sizes of saturated cones on* n *vertices are exactly the integers in* $[\lceil (3n-5)/4 \rceil, n-2]$, *together with* $n-1$ *when* $n \equiv 1 \pmod 3$.

Corollary 3. *For* $n \geq 9$ *every integer in* $[\lceil (3n-9)/4 \rceil, n-2]$, *a set of* $\lfloor (n+5)/4 \rfloor$ *integers, is the size of a saturated hypergraph on* n *vertices. The construction of Theorem 1 has one hyperedge fewer than the smallest saturated cone.*

Theorem 4. *Let* H *be a three-uniform hypergraph on* n *vertices with a vertex* r *such that every other vertex has a one-edge and a two-edge Berge path to* r. *Then* $|E(H)| \geq 3(n-1)/4$, *with equality exactly when* H *is a union of blocks* $\{r,b,c\}, \{r,b,x\}, \{r,c,y\}$ *that are pairwise disjoint outside* r.

Theorem 4 needs neither freeness nor saturation. It shows that the gluing interface of Section 3 cannot beat the coefficient $3/4$; it is not a lower bound for $s(n)$.

Theorem 5. *For* $n = 0, 1, \ldots, 13$,

$$s(n) = 0, 0, 0, 1, 3, 3, 4, 4, 5, 5, 6, 6, 7, 8.$$

At $n = 11$ *there are exactly two saturated hypergraphs with six hyperedges up to isomorphism,* $\{023, 034, 123, 156, 278, \{3, 9, 10\}\}$ *and* $\{012, 023, 045, 126, 178, \{2, 9, 10\}\}$.

The values agree with $\lceil (3n-9)/4 \rceil$ for $9 \leq n \leq 13$; whether $s(n) = \lceil (3n-9)/4 \rceil$ for all $n \geq 9$ is open.

2 Basic facts

Lemma 6 (path criterion)**.** *Let* H *contain no Berge-*C_4 *and let* $f \notin E(H)$ *be a triple. Then* $H + f$ *contains a Berge-*C_4 *if and only if two vertices of* f *are the ends of a Berge path of length three in* H.

Proof. A Berge-C_4 in $H + f$ uses f; deleting f from it leaves a path of length three whose ends lie in f. Conversely, closing such a path by f gives a Berge-C_4, since f differs from the hyperedges of H; the third vertex of f may lie on the path. □

A *leaf* is a vertex of degree one. For a saturated H on $n \geq 4$ vertices no hyperedge has three leaves: two of them with an outside vertex form an absent triple, and no Berge path of length three can end at a leaf of such a hyperedge and at a vertex outside it, nor join two of its leaves. A hyperedge with exactly two leaves is *pendant*; t denotes the number of pendant hyperedges.

Lemma 7. *Let* H *be saturated on* $n \geq 4$ *vertices with* m *hyperedges. Then* H *has at most one isolated vertex and at most two components. If* H *is disconnected, it has no leaf and* $3m \geq 2(n-1)$. *In all cases* $m \geq \lceil (n-1)/2 \rceil$.

Proof. If x, y, z lie in three different components (isolated vertices included), the triple xyz is absent and no path joins two of its vertices, contradicting Lemma 6. Suppose H has two components, and let u, v lie in one of them and w in the other. The triple uvw is absent and only u, v can be joined, so every pair of vertices in a common component is joined by a path of length three. Then no component contains a leaf x with hyperedge xab: a path of length three from a to b cannot use xab (at an end it would continue from the leaf x, in the middle it would need two internal vertices in $\{x\}$), and a path avoiding xab closes with xab to a Berge-C_4 in H. So at most one vertex (an isolated one) has degree below two, and $3m \geq 2(n-1)$. If H is connected, its incidence graph is connected, so $n + m - 1 \leq 3m$, that is, $m \geq (n-1)/2$. □

Lemma 8. *If H contains no Berge-C_4, its shadow (the graph of pairs covered by hyperedges) contains no K_5.*

Proof. First, four distinct triples J inside a five-set S always contain a Berge-C_4. The degrees sum to 12. If a vertex r has degree four, its link has four edges on four vertices, hence contains a path $abcd$; the fourth link edge is ac, ad or bd, and the triples for ac (or ad), ab, bc, cd give the cycle r, a, b, c, while for bd the reversed path gives r, d, c, b. If some vertex has degree zero, all four triples on the other four vertices are present and contain a Berge-C_4. If z has degree one, the other three triples are, up to renaming, abc, abd, acd, and z's triple contains ab or bc; the cores a, b, c, d and b, c, a, d use z's triple, abc, acd, abd in that order. The remaining degree sequence is $(3, 3, 2, 2, 2)$; the complements of the four triples in S form a graph with degrees $(1, 1, 2, 2, 2)$, which is P_5 or a triangle plus a disjoint edge. For $P_5 = 12345$ the triples are $345, 145, 125, 123$ and the core $1, 2, 3, 4$ uses $125, 123, 345, 145$; for the triangle 123 and edge 45 the triples are $345, 145, 245, 123$ and the core $1, 2, 4, 5$ uses $123, 245, 345, 145$.

Now let the shadow contain K_5 on $S = \{1, \dots, 5\}$, let T be the set of hyperedges inside S, and assume $|T| \leq 3$. Every pair of S not covered by T is covered by a hyperedge meeting S in exactly that pair, and distinct pairs give distinct hyperedges. Up to relabelling, the following cases exhaust T; each row gives a core and four distinct covering hyperedges in cyclic order, where a two-digit entry is the hyperedge chosen for that pair:

T	core	hyperedges
$\emptyset$	$1, 2, 3, 4$	$12; 23; 34; 14$
123	$1, 2, 4, 5$	$123; 24; 45; 15$
$123, 124$	$1, 2, 3, 5$	$124; 123; 35; 15$
$123, 145$	$2, 3, 4, 5$	$123; 34; 145; 25$
$123, 124, 125$	$1, 2, 3, 4$	$125; 123; 34; 124$
$123, 124, 134$	$1, 2, 3, 5$	$124; 123; 35; 15$
$123, 124, 135$	$1, 2, 4, 3$	$123; 124; 34; 135$
$123, 124, 345$	$1, 2, 3, 5$	$124; 123; 345; 15$

(For $|T| = 3$ the complements of the triples form a triangle, a star, a path P_4 or a path P_3 plus an edge.) Each case contains a Berge-C_4. □

Lemma 9. *A saturated hypergraph on $n \geq 4$ vertices has at most four pendant hyperedges.*

Proof. Let axy be pendant with leaves x, y. No path of length three joins x and y. For $z \notin \{a, x, y\}$ the absent triple xyz therefore needs a path of length three from x or y to z, which starts with axy and passes to a. If z is a leaf of a second pendant hyperedge bzw, this path ends with bzw, so its middle hyperedge contains a and b, and $a \neq b$. The attachment vertices of pendant hyperedges are therefore distinct and pairwise covered in the shadow, so Lemma 8 allows at most four. □

Four is attained: the triples $012, 013, 023$ with one pendant triple at each of $0, 1, 2, 3$ form a saturated hypergraph on 12 vertices.

3 The upper construction

Let r, b, c be vertices and X, Y disjoint nonempty sets of further vertices. For $Q \subseteq \{r, b, c\}$ let $H(X, Y, Q)$ have the hyperedges $E_0 = rbc$, $E_x = rbx$ $(x \in X)$, $F_y = rcy$ $(y \in Y)$ and, for each $u \in Q$, one pendant triple $u\, p_u\, q_u$ with two new vertices. The *block* is $B = H(\{x\}, \{y\}, \emptyset) = \{rbc, rbx, rcy\}$.

Lemma 10. *$H(X, Y, Q)$ is saturated. Every non-root vertex of $H(X, Y, Q)$ has a one-edge or a two-edge path to r, and every non-root vertex of B has both.*

Proof. Only r, b, c have degree at least two, and the four vertices of a Berge-C_4 have degree at least two, so there is none. Call the other vertices leaves. Two leaves in different hyperedges are joined through E_0 by choosing distinct core vertices in their own hyperedges (two pendant hyperedges have different attachments). A vertex $x \in X$ reaches the three core vertices by $x, E_x, b, E_0, c, F_y, r$; $x, E_x, r, F_y, c, E_0, b$; $x, E_x, b, E_0, r, F_y, c$, and symmetrically for Y. A pendant leaf at u reaches each other core vertex z through u, w, z with w the third core vertex, using E_0 for one step and E_x or F_y for the other. Every absent triple contains one of these pairs, or consists of a pendant leaf, its sibling and its attachment (a present hyperedge), or is rbc (present). Lemma 6 applies. The path statements are immediate: b, c, X, Y and leaves at r are adjacent to r; a leaf at b or c reaches r through its attachment and E_0; in B the two-edge paths are b, c, r; c, b, r; x, b, r; y, c, r. □

Lemma 11 (rooted gluing)**.** *Glue one $H(X, Y, Q)$ and any number of copies of B at r, disjointly otherwise. The result is saturated.*

Proof. Every hyperedge has its non-root vertices in one piece. Around a Berge-C_4 at most one vertex is r, and each other cycle vertex lies in two consecutive cycle hyperedges, so all four hyperedges lie in one piece; pieces are free (a block has only three hyperedges). An absent triple inside one piece and r is handled by that piece. Otherwise it contains non-root vertices u, w of different pieces; one of them lies in a block, so one has a one-edge and the other a two-edge path to r, and the concatenation is a path of length three because the pieces share only r. □

Proof of the upper bound in Theorem 1. For $n \geq 9$ let $n_0 \in \{9, 10, 11, 12\}$ with $n_0 \equiv n \pmod 4$ and take $|X|, |Y|, |Q| = (1, 1, 2), (2, 1, 2), (1, 1, 3), (2, 1, 3)$, a base with n_0 vertices and $5, 6, 6, 7$ hyperedges. Gluing $(n - n_0)/4$ blocks adds four vertices and three hyperedges each, which gives $\lceil (3n - 9)/4 \rceil$ hyperedges. □

The construction is formalized for every $n \geq 9$ (Section 8). The four bases used there are $\{012, 013, 024, 056, 178\}$, $\{012, 013, 014, 025, 067, 189\}$, $\{012, 013, 024, 056, 178, \{2, 9, 10\}\}$ and $\{012, 013, 014, 025, 067, \{1, 8, 9\}, \{2, 10, 11\}\}$ with root 0.

4 The lower bound

Let H be saturated on $n \geq 4$ vertices with m hyperedges. A *one-leaf hyperedge abx* has a leaf x and $\deg a, \deg b \geq 2$. Call x *bad* if $\operatorname{codeg}(a, b) = 1$ and *good* otherwise; q counts bad leaves. A good leaf is *supported* if some hyperedge containing a, b has no leaf, and *unsupported* otherwise; A and B count supported and unsupported good leaves. There are $A + B + q + 2t$ leaves.

Lemma 12. *Each vertex a lies in at most six hyperedges ab_ix_i with x_i a bad leaf.*

Proof. The b_i are distinct. Join b_i and b_j in an auxiliary graph T when some hyperedge contains both. If $b_ib_j, b_jb_k \in T$ through different hyperedges f, g, then a, b_i, b_j, b_k with ab_ix_i, f, g, ab_kx_k is a Berge-C_4 (the hyperedge ab_ix_i contains no other b_j). So $f = g = b_ib_jb_k$, and no further vertex is adjacent to this triangle: the components of T have at most three vertices. A path of length

three from x_i to x_j starts with ab_ix_i, ends with ab_jx_j, and its middle hyperedge must contain b_ib_j (a middle hyperedge through a and b_i or b_j would repeat an end hyperedge because the codegrees are one); conversely $b_ib_j \in T$ gives such a path. Bad leaves from three components of T would form an absent triple with no joined pair. So T has at most two components and at most six vertices. □

Lemma 13. *$q(q-2) \le 432m$ and $q \le 1803$.*

Proof. Let P be the graph on bad leaves joined by paths of length three. Any three bad leaves form an absent triple, so the complement of P is triangle-free and Mantel's theorem gives $|E(P)| \ge \binom{q}{2} - \lfloor q^2/4 \rfloor \ge q(q-2)/4$. A joined pair has a middle hyperedge g and internal vertices $a, b \in g$, and each end leaf lies in a bad-leaf hyperedge through a or b. By Lemma 12 a pair a, b accounts for at most 36 pairs of leaves, a hyperedge for at most 108, so $|E(P)| \le 108m$ and $q(q-2) \le 432m$.

For the absolute bound, let $w_a \in [1, 6]$ be the number of bad-leaf hyperedges through a core endpoint a; then $D = \sum w_a = 2q$. Process the shadow pairs among core endpoints in order of decreasing weight w_aw_b, assigning each to a still unused hyperedge containing it when one exists. The graph F of assigned pairs has no four-cycle (its assigned hyperedges would form a Berge-C_4). An unassigned pair p lies only in hyperedges already assigned to pairs of weight at least that of p, and a hyperedge contains three pairs; hence the total shadow weight is at most $3W$ with $W = \sum_{ab \in F} w_aw_b$. With $z_v = \sum_{u \sim v} w_u$, the weighted Cauchy–Schwarz inequality gives $\sum_v w_v z_v^2 \ge 4W^2/D$. Two vertices of F have at most one common neighbour, of weight at most six, so $\sum_v w_v(z_v^2 - \sum_{u \sim v} w_u^2) \le 12 \sum_{a<b} w_aw_b \le 6D^2$, while $\sum_v w_v \sum_{u \sim v} w_u^2 = \sum_{ab \in F} w_aw_b(w_a + w_b) \le 12W$. Thus $W^2 - 6qW - 12q^3 \le 0$ and $W \le 3q + q\sqrt{9 + 12q}$. Every edge of P has its middle pair in the shadow of the core endpoints, and a pair ab accounts for at most w_aw_b edges of P, so $q(q-2)/4 \le 3W$. Dividing by q gives $q \le 38 + 12\sqrt{12q + 9}$, and for $q > 38$ squaring gives $q^2 - 1804q + 148 \le 0$, which fails for every integer $q \ge 1804$. □

For connected H let $\beta = 2m - n + 1$ be the cycle rank of the incidence graph over $\mathbb{F}_2$, and $\Delta = \sum_{\deg v \ge 2} (\deg v - 2)$.

Lemma 14. *If H is connected, $\beta \ge A + 2B/3$.*

Proof. Group good-leaf hyperedges by their pair ab. In a supported group fix a leafless hyperedge on ab as reference; in an unsupported group (all hyperedges on ab are good-leaf hyperedges, at least two) fix one of them. Delete all non-reference good-leaf hyperedges with their leaves. Each deletion removes two nodes and three edges of the incidence graph without disconnecting it, so it lowers the cycle rank by one: with g unsupported groups and reduced hypergraph R, $\beta(H) = \beta(R) + A + B - g$.

For an unsupported group abx, aby of size two, the triple axy is absent, x and y are not joined (the middle hyperedge would contain ab), so a is joined to x or y; removing the leaf's hyperedge gives a two-edge path from a to b avoiding the group. Replacing deleted hyperedges on this path by their references keeps a genuine two-edge path (a common reference would be a leafless hyperedge on ab). With the retained reference on ab it forms a cycle of R through three hyperedges. Project the cycle space of R onto one incidence coordinate of each of the g_2 references of size-two groups. Each chosen cycle contains its own coordinate and at most three coordinates, so the span has dimension at least $g_2/3$. Hence $\beta(H) \ge A + \sum(k-1) + g_2/3$ over unsupported groups of sizes k, and $k - 1 + [k = 2]/3 \ge 2k/3$. □

Lemma 15. *$\Delta \ge A/2 + 2B/3$.*

Proof. Let $E = abx$ be a good-leaf hyperedge and $F = aby$ another hyperedge on ab. If a and b both have degree two, no pair of the absent triple axy is joined: a path of length three from x starts with E and continues with F through a or b, after which only y or a vertex of E remains

and no third hyperedge is available; a path between a and y starts with E or F, passes through b and then through the other of E, F, and again stops. So one of a, b has degree at least three. In an unsupported group abx, aby of size two, if a had degree two, again axy would have no joined pair: x and y are not joined because a third hyperedge on ab is missing, and a path from x to a would run x, b and end with aby, needing a middle hyperedge from b to the leaf y; similarly for y. Larger unsupported groups give degree at least three directly. Charge each supported good leaf $1/2$ to its endpoints of degree at least three (split equally) and each unsupported good leaf $1/3$ to each endpoint. A vertex of degree $d \geq 4$ receives at most $d/2 \leq d - 2$. A vertex of degree three in at most two good-leaf hyperedges receives at most one; if all three of its hyperedges are good-leaf hyperedges, they form one unsupported group of size three and it receives exactly one. So no vertex receives more than $\deg v - 2$. □

Proof of the lower bound in Theorem 1. If H is disconnected, Lemma 7 gives $m \geq 2(n-1)/3$, which implies (1). If H is connected, it has no isolated vertex and counting incidences gives $2n = 3m + A + B + q + 2t - \Delta$. With Lemma 15 this gives $2n \leq 3m + A/2 + B/3 + q + 2t$, and Lemma 14 gives $n \leq 2m + 1 - A - 2B/3$. Adding twice the first to the second,

$$5n \leq 8m + 2q + 4t + 1.$$

Lemma 9 gives $t \leq 4$, and Lemma 13 gives $q \leq 1803$ and $q \leq 1 + \sqrt{432m + 1}$, which yield the two inequalities. For the third term of the abstract put $x = \sqrt{432m + 1}$; the second inequality becomes $(x + 54)^2 \geq 270n + 1891$, so $x \geq \sqrt{270n + 1891} - 54 > 0$, and squaring and substituting $m = (x^2 - 1)/432$ gives $m \geq (5n + 89)/8 - \sqrt{270n + 1891}/4$. □

5 Cones

Proof of Theorem 2. A Berge-C_4 in a cone with root r either avoids r (a C_4 in G) or has core r, a, b, c and link edges ax, ab, bc, cy with $x, y \neq b$; then $x = c$ or $y = a$ gives a paw, $x = y$ a C_4 and $x \neq y$ a P_5. Conversely P_5, C_4 and the paw give Berge-C_4s. So H is free if and only if G contains no P_5, C_4 or paw, that is, every component is a star, a double star, a triangle or an isolated vertex. A Berge path whose ends lie in $\{r\}$ and one component stays in that component.

A component $K_{1,2}$ with leaves x, y fails at the triple rxy. A star centre has no path of length two or three to r, so two star components fail at ruv for their centres. An isolated vertex z requires every pair of other vertices to be joined; this fails for a star centre and r, for a double-star centre and r, and for a second isolated vertex. These cases give necessity. For sufficiency, each allowed component with r is saturated: $K_{1,1}$ trivially, $K_{1,t}$ ($t \geq 3$) through a third leaf, K_3 because the cone is $K_4^{(3)}$ minus one triple, and $D(p, q)$ by Lemma 10 with $Q = \emptyset$. Every vertex of a double star or triangle has paths of lengths one and two to r, every vertex of a star one of length one, so triples meeting two components are closed through r; with an isolated vertex all pairs in the triangles are joined.

For the sizes, put $N = n - 1$. Without an isolated vertex let τ be the number of tree components; then $m = N - \tau$, every tree has at least four vertices except at most one $K_{1,1}$, so $N \geq 4\tau - 2$, and $\tau = 0$ needs $3 \mid N$. With an isolated vertex $m = N - 1$. Conversely, for $1 \leq \tau \leq \lfloor (N+2)/4 \rfloor$ take $\tau - 1$ paths P_4 and one double star on the remaining $N - 4(\tau - 1) \geq 4$ vertices, or, when $N \in \{4\tau - 2, 4\tau - 1\}$, one $K_{1,1}$, $\tau - 2$ paths P_4 and a double star on four or five vertices; $N/3$ triangles give $m = N$. The minimum is $N - \lfloor (N+2)/4 \rfloor = \lceil (3n - 5)/4 \rceil$. □

Corollary 3 follows, since $\lceil (3n-5)/4 \rceil = \lceil (3n-9)/4 \rceil + 1$ and the smallest value is realized by Section 3.

6 The cost of a flexible root

Proof of Theorem 4. Let G be the link of r on $N = n-1$ vertices, $a = |E(G)|$, and b the number of hyperedges avoiding r. One-edge paths mean G has no isolated vertex. Let t_2 count components K_2 and t_3 components P_3; every other component C has at least $3|C|/4$ edges, so $a \geq 3N/4 - t_2/2 - t_3/4$. A vertex all of whose neighbours in G have degree one, and which lies only in hyperedges through r, has no two-edge path to r; so both vertices of each K_2 and the centre of each P_3 lie in hyperedges avoiding r, and $3b \geq 2t_2 + t_3$. Hence $m = a + b \geq 3N/4 + t_2/6 + t_3/12$. Equality forces $t_2 = t_3 = 0$, $b = 0$ and components with exactly $3|C|/4$ edges, that is trees on four vertices; a star $K_{1,3}$ has a centre without a two-edge path, so every component is P_4, which is the block. □

7 Exact small values

For $n \leq 2$ the empty hypergraph is vacuously saturated. For $n \geq 3$ a saturated hypergraph is nonempty, and for $n \geq 4$ it has at least three hyperedges, since adding a triple to at most two hyperedges leaves fewer than four. The witnesses $\{012\}$, $\{012, 013, 023\}$ and the block show $s(3) = 1$ and $s(4) = s(5) = 3$. Cones over the double star $D(1,2)$, over $K_{1,1} \cup P_4$ and over $K_{1,1} \cup D(1,2)$ give $s(6) \leq 4$, $s(7) \leq 4$ and $s(8) \leq 5$, and Section 3 gives the upper bounds for $9 \leq n \leq 13$.

The remaining lower bounds exclude saturated hypergraphs with m hyperedges on n vertices for $m = 3$, $n \geq 6$; $m = 4$, $n \geq 8$; $m = 5$, $n \geq 10$; $m = 6$, $n \geq 12$; and $m = 7$, $n \geq 13$. By Lemma 7, an isolated vertex forces $3m \geq 2(n-1)$, which fails in each range, and without isolated vertices $n \leq 3m$. It therefore suffices to treat hypergraphs without isolated vertices on $n \leq 3m$ vertices. Each is encoded, after labelling its hyperedges, by the multiset of incidence masks of its vertices; masks with at least two bits are listed in nondecreasing order and the remaining incidences are filled by degree-one vertices. Every hypergraph appears (possibly several times). The certificate gives, for every profile in the enumeration order, an exclusion: three vertices in different components, an explicit Berge-C_4 of H, or an absent triple whose addition leaves the hypergraph free. The verifier regenerates the enumeration independently, rejects missing or surplus records, and checks freeness of $H + f$ by testing all four-sets of hyperedges and all injective choices of cycle vertices. The record counts are 21, 261, 4664, 107740 and 10592111. A sixth certificate covers all 203895 profiles with $m = 6$, $n = 11$; it excludes all but 1080, each of which is matched, under an explicit permutation of the hyperedge labels, to one of the two hypergraphs of Theorem 5. The two are not isomorphic, having degree sequences $(4, 3, 2, 2, 1^7)$ and $(4, 3, 3, 1^8)$.

8 Formal verification

The Lean 4 project (Lean 4.35.0, Mathlib `56a005855ea17e0b0958f73ec801ae140b0a1679`) defines a hypergraph on a vertex type V as a finite set of finite subsets of V. A Berge-C_4 is given by an injective map $v : \mathrm{Fin}\, 4 \to V$ and an injective map e to hyperedges of H with $v_i, v_{i+1} \in e_i$; `IsBergeC4Saturated` requires three-uniformity, freeness and a Berge-C_4 after inserting any absent three-element set; and `satNumber n` is the least size of a saturated hypergraph on $\mathrm{Fin}\, n$. The project proves:

- `exists_saturated`: for every $n \geq 9$ there is a saturated hypergraph on $\mathrm{Fin}\, n$ with $(3n-6)/4 = \lceil (3n-9)/4 \rceil$ hyperedges, and hence `satNumber_le_upper`;
- `glued_saturated`: the rooted gluing theorem for an arbitrary base and block satisfying the hypotheses of Lemmas 10 and 11, with the freeness argument around the root formalized in general;

- `satNumber_small`: $s(n) = 0, 0, 0, 1, 3, 3$ for $n \leq 5$, and `satNumber_le_small`: $s(6) \leq 4$, $s(7) \leq 4$, $s(8) \leq 5$.

The properties of the explicit bases and the block (uniformity, the three vertices of degree at least two, paths for every absent triple and paths to the root) are Boolean searches evaluated by the kernel with `decide +kernel`, each with a soundness theorem. No `native_decide` is used, and a module-level audit shows that every declaration depends only on `propext`, `Classical.choice` and `Quot.sound`. The lower bound, Theorems 2 and 4, Corollary 3 and the values $s(n)$ for $6 \leq n \leq 13$ are not formalized.

9 Related work and scope

The definitions and the bound $s(n) \leq n - 2$ are from [1], whose journal version is [2]; the inspected statement is that of the preprint, as the journal text could not be accessed. Linear growth of $\mathrm{sat}_k(n, \text{Berge-}F)$ for $3 \leq k \leq 5$ is due to [3], and gluing at a vertex is used in [1]. Wang and Zhang [5] bound the linear saturation number of Berge-C_4, a different problem. The ordinary path saturation numbers are due to Kászonyi and Tuza [4]. Searches in October 2026 of arXiv, journal records and general web indexes did not find the coefficients 5/8 or 3/4, the cone classification or the values of Theorem 5; some recent full texts were inaccessible, and this does not establish priority.

10 Research process and AI disclosure

This is fully autonomous agentic research. Dev D. Goyal is responsible for publication. The work has not been externally peer reviewed, no expert validation is claimed, and the novelty assessment is provisional.

The upper construction was found by autonomous AI research agents in an earlier session. In a later session the agents proved the lower bound, the cone classification and spectrum and the root-cost theorem, built and independently verified the finite certificates and searched the literature. Their records do not name the system used; the research contract specifies Astra referees for the internal reviews, which were AI-assisted reviews of the proofs and not peer review, and the records state that no Lean formalization was attempted.

Opus 5.5 by Anthropic used with Claude Code assessed the package for this release. It checked the written proofs and the arithmetic of the lower bound, compared the statements with [1, 3, 5], rechecked the construction with its own exhaustive checker, reran the six certificates, wrote and ran the Lean formalization described in Section 8, and wrote this note. It also made the editorial revision of 1 November 2026, which did not change the mathematics. AI checks are not human peer review. Section 8 states the formal scope and Section 9 the limits of the literature comparison.

References

[1] S. English, N. Graber, P. Kirkpatrick, A. Methuku and E. C. Sullivan, *Saturation of Berge hypergraphs*, arXiv:1710.03735v1 (2017).

[2] S. English, P. Gordon, N. Graber, A. Methuku and E. C. Sullivan, *Saturation of Berge hypergraphs*, Discrete Mathematics **342** (2019), 1738–1761. doi:10.1016/j.disc.2019.01.031.

[3] S. English, D. Gerbner, A. Methuku and M. Tait, *Linearity of saturation for Berge hypergraphs*, European Journal of Combinatorics **78** (2019), 205–213. arXiv:1807.06947.

[4] L. Kászonyi and Zs. Tuza, *Saturated graphs with minimal number of edges*, Journal of Graph Theory **10** (1986), 203–210.

[5] C. Wang and J. Zhang, *Linear saturation numbers of Berge-C_3 and Berge-C_4*, arXiv:2306.06435 (2023); Applied Mathematics and Computation **474** (2024).