Exact Berge saturation numbers of the subdivided star in every uniformity and its minimality among trees

Dev D. Goyal

1 November 2026

Abstract

Let B_a be the star $K_{1,a}$ with one edge subdivided. For all $a, r \geq 3$ and $n \geq (r-1)a+1$ we prove

$$\mathrm{sat}_r(n, \text{Berge-}B_a) = \left\lfloor \frac{a(n+1)}{(r-1)a+1} \right\rfloor,$$

where $\mathrm{sat}_r(n, \text{Berge-}F)$ is the least number of hyperedges of an r-uniform hypergraph on n vertices that contains no Berge copy of F but contains one after the addition of any absent r-set. Disjoint loose stars attain the value; for smaller n the saturation number is a or $\binom{n}{r}$. Saturated hypergraphs within k hyperedges of the minimum consist, for $n \geq ((r-1)a+1)((r-1)(k+1)+1)$, of loose stars with at least a spokes and at most $(r-1)((r-1)a+1)(k+1)-1$ further vertices; the constant is sharp for $k=0$. In the three-uniform case, for every $a \geq 5$ every tree $T \ncong B_a$ on $a+2$ vertices satisfies $\mathrm{sat}_3(n, \text{Berge-}T) \geq \lceil (a+1)(n-2)/(2a+3) \rceil$, so B_a is the unique tree of its order with the least asymptotic lower saturation coefficient, $a/(2a+1)$, and the unique pointwise minimizer for $n \geq 3(a+1)(2a+1)$. This is the Berge analogue of the theorem of Faudree, Faudree, Gould and Jacobson [9] that the subdivided star minimizes graph saturation among trees; the Berge saturation problem is due to English, Graber, Kirkpatrick, Methuku and Sullivan [5]. Lean 4 verifies the exact formula, with the saturation number defined as a minimum over all hypergraphs on n labelled vertices, and the near-minimum structure theorem; the smaller orders, the sharpness constructions and the comparison among trees have conventional proofs.

1 Introduction and results

All hypergraphs are finite, simple and r-uniform with $r \geq 3$; the vertex set is part of the data, so isolated vertices count. A *Berge copy* of a graph F in a hypergraph H consists of an injection ϕ of $V(F)$ into $V(H)$ and an injection ψ of $E(F)$ into $E(H)$ with $\{\phi(u), \phi(v)\} \subseteq \psi(uv)$ for every edge uv. A hypergraph is *Berge-F-saturated* if it contains no Berge copy of F but $H+e$ contains one for every r-set $e \notin E(H)$; no condition such as linearity is imposed on e. The saturation number $\mathrm{sat}_r(n, \text{Berge-}F)$ is the least number of hyperedges of a Berge-F-saturated hypergraph on n vertices. English, Graber, Kirkpatrick, Methuku and Sullivan introduced the problem and studied paths, cycles, stars, triangles and matchings [5, 6]; English, Gerbner, Methuku and Tait proved linearity for $3 \leq r \leq 5$ and for every r when F has a tree component [7], and a recent preprint of Xie proves linearity for every graph and every uniformity [10]. Exact values are known for few families: Berge stars [1, 5], matchings and triangles [5], and recently three-uniform K_4 [4, 3]; for paths [5] gives bounds that differ by at most three. Clique asymptotics are due to English, Kritschgau, Nahvi and Sprangel [8], and the saturation spectrum of Berge stars was studied by Bushaw, English, Heath, Johnston and Rombach [2].

For graphs, Faudree, Faudree, Gould and Jacobson [9] showed that among all trees of order $k \geq 5$ the star $K_{1,k-2}$ with one edge subdivided has the least saturation number, and that it is

the only tree attaining the minimum for all n. Write B_a for that tree with $a+2$ vertices: centre c, stem b, terminal x and leaves $y_1, \dots, y_{a-1}$, with edges cb, bx and cy_i. It has $a+1$ edges, and for $a \geq 3$ it is not a star. Throughout put

$$s = r - 1, \qquad A = sa + 1.$$

A *loose star with t spokes* has hyperedges $\{c\} \cup P_i$ $(1 \leq i \leq t)$ for pairwise disjoint s-sets P_i not containing c; it has $st + 1$ vertices.

Theorem 1. *Let $a, r \geq 3$ and $n \geq A$, and write $n = Aq + t$ with $0 \leq t < A$. Then*

$$\mathrm{sat}_r(n, \textit{Berge-}B_a) = \left\lfloor \frac{a(n+1)}{A} \right\rfloor = aq + \left\lfloor \frac{t}{s} \right\rfloor = \left\lceil \frac{a(n-s+1)}{A} \right\rceil.$$

The value is attained by q disjoint loose stars, one with $a + \lfloor t/s \rfloor$ spokes and the others with a, together with $t \bmod s$ isolated vertices.

Theorem 2. *Let $a, r \geq 3$ and $N_0 = \max(a+2, r+1)$. Then $\mathrm{sat}_r(n, \textit{Berge-}B_a) = \binom{n}{r}$ for $n < N_0$ and $\mathrm{sat}_r(n, \textit{Berge-}B_a) = a$ for $N_0 \leq n < A$.*

For a saturated hypergraph H let $R(H)$ be the number of vertices outside the components of H that are loose stars with at least a spokes.

Theorem 3. *Let $a, r \geq 3$, $k \geq 0$ and $n \geq A(s(k+1)+1)$. If H is a Berge-B_a-saturated hypergraph on n vertices with $|E(H)| \leq \mathrm{sat}_r(n, \textit{Berge-}B_a) + k$, then H has a component that is a loose star with at least a spokes, and $R(H) \leq sA(k+1) - 1$.*

For $k = 0$ the threshold is $n \geq rA$ and the bound is $R(H) \leq sA - 1$; both are attained (Proposition 11).

Theorem 4. *Let $a \geq 5$ and let T be a tree on $a+2$ vertices. For every $n \geq 1$,*

$$\mathrm{sat}_3(n, \textit{Berge-}T) \geq \left\lceil \frac{a(n-2)}{2a+1} \right\rceil, \qquad \textit{and if } T \ncong B_a, \qquad \mathrm{sat}_3(n, \textit{Berge-}T) \geq \left\lceil \frac{(a+1)(n-2)}{2a+3} \right\rceil.$$

Consequently $\mathrm{sat}_3(n, \textit{Berge-}B_a)/n \to a/(2a+1)$, while $\liminf_n \mathrm{sat}_3(n, \textit{Berge-}T)/n \geq (a+1)/(2a+3)$ for every other tree T of the same order; the gap is $1/((2a+1)(2a+3))$. Moreover $\mathrm{sat}_3(n, \textit{Berge-}T) > \mathrm{sat}_3(n, \textit{Berge-}B_a)$ for every $T \ncong B_a$ and every $n \geq 3(a+1)(2a+1)$.

For $a = 3$ and $a = 4$ a direct enumeration of the incidence trees with a triples finds further repeatable blocks, for the paths on five and six vertices, so the restriction $a \geq 5$ is used in Section 6.

Theorems 1 and 3 are verified in Lean 4 (Section 7); Theorems 2 and 4 and Proposition 11 have the conventional proofs given here.

2 Components and incidence counting

The *incidence graph* of H is the bipartite graph on $V(H) \sqcup E(H)$ joining v to e when $v \in e$. Components of H are the vertex classes of its components; an isolated vertex is a component without hyperedges. A connected hypergraph with v vertices and $m \geq 1$ hyperedges has an incidence graph with $v + m$ nodes and rm edges, so

$$v \leq sm + 1, \qquad \text{with equality exactly when the incidence graph is a tree.}$$

Call such a component an *incidence tree*. Single hyperedges and loose stars are incidence trees. In an incidence tree with m hyperedges the sum of $d(v) - 1$ over its vertices is $rm - v = m - 1$,

so at most $m-1$ vertices have degree at least two, and at least $(r-2)m+2$ vertices have degree one. Two hyperedges of an incidence tree meet in at most one vertex.

Every Berge copy of a connected graph lies in a single component, and the centre of a Berge copy of B_a has degree at least a in H, because the a hyperedges representing the edges at the centre are distinct and contain it; its stem has degree at least two.

Lemma 5. *Let H be Berge-B_a-saturated with $a, r \geq 3$. Every incidence-tree component of H with at least two hyperedges is a loose star with at least a spokes.*

Proof. Let C be such a component with $m \geq 2$ hyperedges. Suppose some vertex c of C has degree at least a and shares a hyperedge e with a vertex b of degree at least two. Let $f \neq e$ contain b and choose $x \in f \setminus \{b\}$; then $c \notin f$, and x lies in no hyperedge at c, for otherwise the incidence graph would contain a cycle. For the same reason the other hyperedges at c, of which there are at least $a-1$, meet one another and $e \cup f$ only in c, so choosing one vertex other than c from $a-1$ of them gives a Berge copy of B_a with centre c, stem b and terminal x, which is impossible. Hence if C has a vertex of degree at least a, all its neighbours have degree one and C is a loose star with at least a spokes. Otherwise every degree in C is at most $a-1$, and C has at least $(r-2)m+2 \geq r$ vertices of degree one. An r-set of them is absent, since every hyperedge of C contains a vertex of degree at least two; adding it raises those degrees to at most two and leaves every degree below a, so no Berge copy of B_a appears, a contradiction. □

For a saturated hypergraph H let h be the number of loose-star components with at least a spokes, e_0 the number of components consisting of one hyperedge, and z the number of isolated vertices. For each other component C, which contains a cycle in its incidence graph, put $j(C) = s|E(C)| - |V(C)| \geq 0$, and let S be the sum of the $j(C)$. Summing $s|E(C)|$ over components gives the exact identity

$$s\,|E(H)| = n - h - e_0 - z + S. \tag{1}$$

Lemma 6. *If H is Berge-B_a-saturated with $a, r \geq 3$, then $e_0 \leq 1$, $z \leq s$, $e_0 z = 0$, and $z \leq s-1$ when $h \geq 1$.*

Proof. In each case we add an absent r-set and find no Berge copy of B_a. If $e_0 \geq 2$, add an r-set meeting two single-hyperedge components: the new component has maximum degree two, below a. The same works for an isolated vertex and a single-hyperedge component. If $z \geq r = s+1$, add an r-set of isolated vertices, which forms a new single-hyperedge component. If $h \geq 1$ and $z \geq s$, add the centre of a loose star together with s isolated vertices; the result is a loose star with one more spoke, in which only the centre has degree above one, and a Berge copy of B_a needs two vertices of degree at least two. □

3 Constructions

Lemma 7. *Let $a, r \geq 3$. A disjoint union of loose stars with at least a spokes each and at most $s-1$ isolated vertices is Berge-B_a-saturated.*

Proof. Each component has at most one vertex of degree above one, so there is no Berge copy of B_a. Let e be an absent r-set. If e contains a non-centre vertex b of a star with centre c, choose $x \in e \setminus \{b, c\}$. Represent cb by the spoke containing b, bx by e, and $a-1$ further edges at c by other spokes, choosing in each a vertex other than c and x; this is possible because each spoke has $s \geq 2$ vertices besides c. Otherwise e consists of centres and at most $s-1$ isolated vertices, so it contains two centres c, d. Represent cd by e, the pendant edge at d by a spoke at d, and the remaining $a-1$ edges at c by spokes at c. □

Upper bound in Theorem 1. The hypergraph described in the theorem has $Aq + s\lfloor t/s \rfloor + (t \bmod s) = n$ vertices and $aq + \lfloor t/s \rfloor$ hyperedges, and it is saturated by Lemma 7. For the identities write $t = sd + \varepsilon$ with $0 \le \varepsilon < s$, so $0 \le d \le a$. Then $a(n+1) - A(aq+d) = a(\varepsilon+1) - d$ lies in $[0, A-1]$, which gives the floor form, and the ceiling form follows from $a(n-s+1) + A - 1 = a(n+1)$. □

For Theorem 2 we use cones. A family L of distinct s-sets is a *hyperforest* if every nonempty subfamily J satisfies $|\bigcup J| \ge |J| + 1$. The *cone* over L with a new root c has the hyperedges $\{c\} \cup P$ for $P \in L$.

Lemma 8. *Let L be a hyperforest, $b \in \bigcup L$ and $x \ne b$. There is a member $P \ni b$ of L such that the members of $L \setminus \{P\}$ have distinct representatives avoiding b and x.*

Proof. After x is deleted from every member, each nonempty subfamily still covers at least as many vertices as it has members, so by Hall's theorem the members of L have distinct representatives different from x. If b is not a representative, give it to one member containing b instead. The member P represented by b satisfies the lemma. □

Lemma 9. *Let L be a hyperforest of a members, all s-sets, without isolated vertices. The cone over L together with at most $s-1$ isolated vertices is Berge-B_a-saturated.*

Proof. It has only a hyperedges, so it has no Berge copy of B_a. An absent r-set e contains at most $s-1$ isolated vertices and possibly the root, so it contains a vertex b of $\bigcup L$; choose $x \in e \setminus \{b, c\}$. Lemma 8 gives $P \ni b$ and representatives of the other $a-1$ members avoiding b, x. Represent cb by $\{c\} \cup P$, bx by e and the other edges at c by the corresponding cone hyperedges. □

Lemma 10. *For every v with $\max(a+1, s+1) \le v \le sa$ there is a hyperforest of a distinct s-sets without isolated vertices on exactly v vertices.*

Proof. For $v_0 = \max(a+1, s+1)$ take a of the v_0 cyclic intervals of length s in a cyclically ordered v_0-set. They are distinct since $s < v_0$; any $k \ge 1$ of them cover at least $\min(v_0, k+s-1) \ge k+1$ vertices; and they cover every vertex, since if $v_0 = a+1$ only one interval is omitted and $s \ge 2$, and if $v_0 = s+1$ any two intervals cover the cycle. While the family has fewer than sa vertices, some vertex lies in two members; replace one of its occurrences by a new vertex. The members stay distinct, no vertex becomes isolated, every union grows by at most one or stays fixed, so the hyperforest condition persists, and the order grows by one. □

Proof of Theorem 2. If $n < a+2$ there is no injection of $V(B_a)$, and if $n \le r$ there is at most one hyperedge while B_a needs $a+1$; so for $n < N_0$ every hypergraph is Berge-B_a-free and only the complete one is saturated. Let $N_0 \le n < A$. A saturated hypergraph either has an absent r-set, whose addition creates a copy with $a+1$ hyperedges, so it has at least a hyperedges, or it is complete, with $\binom{n}{r} \ge n > a$ hyperedges. For the upper bound, $\max(a+1, s+1) \le n-1 \le sa$, so the cone over the hyperforest of Lemma 10 on $n-1$ vertices is saturated with a hyperedges by Lemma 9. □

4 The lower bound

Lower bound in Theorem 1. Let H be saturated on $n = Aq + t \ge A$ vertices with m hyperedges. By Lemma 5 the loose-star components with at least a spokes are exactly the incidence-tree components with at least two hyperedges, and each has at least A vertices, so $h \le q$. If $h \ge 1$, then $e_0 + z \le s-1$ by Lemma 6, and (1) gives $sm \ge n - q - s + 1 = saq + t - s + 1$, hence

$$m \ge aq + \left\lceil \frac{t-s+1}{s} \right\rceil = aq + \left\lfloor \frac{t}{s} \right\rfloor.$$

If $h = 0$, then $e_0 + z \le s$ and $sm \ge n - s$, so $m \ge aq + \lceil (q+t)/s \rceil - 1 \ge aq + \lfloor t/s \rfloor$, because $q + t > s\lfloor t/s \rfloor$ when $q \ge 1$. □

5 Near-minimum structure

Proof of Theorem 3. Let $m = |E(H)| \le aq + \lfloor t/s \rfloor + k$ and $\varepsilon = t \bmod s$, so $s\lfloor t/s \rfloor = t - \varepsilon$. By (1) with $S \ge 0$ and $e_0 + z \le s$,

$$h \ge n - sm - s \ge q + \varepsilon - s(k+1) \ge 1,$$

since $q \ge s(k+1) + 1$. Hence $e_0 + z \le s - 1$ and $h \ge n - sm - s + 1$, while $Ah \le n - R(H)$. Therefore

$$R(H) \le n - A(n - sm - s + 1) = sAm - san + A(s-1).$$

Inserting $m \le aq + \lfloor t/s \rfloor + k$ and $n = Aq + t$ gives $R(H) \le t - A\varepsilon + sAk + A(s-1) \le sA(k+1) - 1$, because $t \le A - 1$. □

Proposition 11. *Let $a, r \ge 3$. Let U be the cone over a distinct s-sets of which two share exactly one vertex and which are otherwise pairwise disjoint; it has a hyperedges and sa vertices.*

1. *For $q \ge s$, the union of $q - s + 1$ loose stars with a spokes, s copies of U and $s - 1$ isolated vertices is a minimum saturated hypergraph on $Aq + A - 1$ vertices with $R = sA - 1$.*

2. *The union of r copies of U and s isolated vertices is a minimum saturated hypergraph on $rA - 1$ vertices with no loose-star component.*

Proof. The link of U has one repeated incidence, so it is a hyperforest. In both hypergraphs every component has at most a hyperedges, so there is no Berge copy of B_a. Let e be absent. If e contains a non-centre vertex of a loose star, argue as in Lemma 7; if it contains a non-root vertex b of a copy of U with root c, choose $x \in e \setminus \{b, c\}$ and argue as in Lemma 9, the vertex x possibly lying in another component. Otherwise e consists of centres, roots and isolated vertices. If it contains two centres or roots c, d, represent cd by e, the pendant edge by a hyperedge at d, and the other $a - 1$ edges at c by hyperedges at c with distinct representatives avoiding d and the terminal; these exist since every link is a hyperforest. In (1) there are at most $s - 1$ isolated vertices, so this covers every case. In (2) the remaining case is one root c and all s isolated vertices. Let w be the shared vertex of the two overlapping hyperedges f_1, f_2 of that copy of U, and $x \in f_2 \setminus \{c, w\}$. Represent cw by f_1, wx by f_2, one further edge at c by e, and the remaining $a - 2$ edges at c by the other $a - 2$ hyperedges of the copy, which meet $f_1 \cup f_2$ only in c.

The orders are $(q - s + 1)A + s^2a + s - 1 = Aq + A - 1$ and $rsa + s = rA - 1$, and the sizes are $(q + 1)a$ and ra, which equal $\mathrm{sat}_r(n, \text{Berge-}B_a)$ by Theorem 1 because $t = A - 1 = sa$ in both cases. The vertices outside loose stars number $s \cdot sa + s - 1 = sA - 1$ in (1). □

6 Minimality among trees

In this section $r = 3$, $a \ge 5$ and T is a tree with $a + 1$ edges. A Berge copy of T selects from each hyperedge it uses one pair, the image of an edge of T; conversely, distinct hyperedges with selected pairs forming a copy of T give a Berge copy. Core degrees are at most host degrees. If a host has exactly $a + 1$ hyperedges, a Berge copy of T uses all of them. Three-uniform incidence trees with m hyperedges have $2m + 1$ vertices, and connected components that are not incidence trees satisfy $v \le 2m$. For an ordinary tree R with at least two vertices, the handshake identity gives

$$L(R) = 2 + \sum_{x \in V(R)} (d_R(x) - 2)_+, \tag{2}$$

where $L(R)$ is the number of leaves.

6.1 Internally saturated incidence trees with a triples

Let H be an incidence tree with exactly a triples such that $H + e$ contains a Berge copy of T for every absent triple $e \subseteq V(H)$. (H itself has too few hyperedges to contain one.)

Lemma 12. *No triple of H has three vertices of degree at least two.*

Proof. Suppose h does. Removing h from the incidence tree leaves three branches meeting h in distinct vertices, each with at least one triple; the $a - 1 \geq 4$ other triples lie in these branches, so one branch has $m \geq 2$ triples and $2m + 1$ vertices, hence an absent triple e inside it. A Berge copy of T in $H + e$ uses every triple. Every triple other than h, including e, lies inside one branch, and each branch contains a used triple whose selected pair lies in that branch; the selected pair of h meets at most two branches, so the selected pairs do not form a connected graph. □

Discarding from every triple one vertex of degree one, keeping both vertices of degree at least two when there are two, turns H into an ordinary graph G with a edges and $a + 1$ vertices. It is connected and acyclic, so G is a tree and H is its *expansion*, obtained by adding a private vertex to each edge. The vertices of degree at least two in H are the nonleaf vertices of G, with the same degrees.

Lemma 13. *G is a star or a double star, and $L(T) \leq L(G)$.*

Proof. Let G have L leaves and $I = a + 1 - L$ nonleaf vertices. H has $a + L \geq 3$ vertices of degree one, and no triple of H consists of such vertices, since every edge of the tree G with $a \geq 2$ edges has a nonleaf endpoint. Add a triple e of them. In $H + e$ they have degree at most two and the other degrees are unchanged, so by (2) applied to T and the monotonicity of $(d - 2)_+$, $L(T) \leq 2 + \sum_v (d_G(v) - 2)_+ = L$. If $I \geq 3$, add a triple f of nonleaf vertices of G, absent by Lemma 12. In $H + f$ only the I nonleaf vertices have degree at least two, so a Berge copy of T has at least $a + 2 - I = L + 1$ leaves, a contradiction. Thus $I \leq 2$. □

There is one star and there are $\lfloor (a-1)/2 \rfloor$ double stars with a edges, so at most $1 + \lfloor (a-1)/2 \rfloor = \lceil a/2 \rceil$ isomorphism types of H are possible.

6.2 Three disjoint copies

Lemma 14. *If the disjoint union $3H$ of three copies of H is Berge-T-saturated, then $T \cong B_a$ and H is a loose star.*

Proof. If a triple meeting three different copies is added, a new copy of T uses it, and its selected pair joins two of the copies; deleting that edge of T leaves two rooted subtrees, one in each of these copies, rooted at the endpoints of the pair, and nothing from the third copy.

Suppose first that H is a loose star with centre c and triples $\{c, x_i, y_i\}$. A connected Berge subtree containing c selects pairs through c, so it is a star centred at c. Adding the triple of the three centres shows that T is two stars with joined centres. T is not the star $K_{1,a+1}$: adding a triple of three vertices of degree one leaves the maximum degree a. So T is a double star. Now add $\{c, x_1, x_2\}$: only c has degree at least three, and x_1, x_2 have degree two. The two centres of T have degrees summing to $a + 2 \geq 7$, so one has degree at least three and is mapped to c, and the other has degree two. Hence $T \cong B_a$.

Suppose next that H is the expansion of a double star with hubs u, v, with $p \geq 1$ pendant triples at u and $q \geq 1$ at v, $p + q = a - 1$, and central triple $\{u, v, w\}$. A nontrivial Berge subtree rooted at w uses the central triple with pair wu or wv and then only pendant triples at that hub, each through the hub; it is a path w–hub followed by $\sigma \leq \max(p, q) \leq a - 2$ leaves at the hub. Add the triple of the three copies of w. If one rooted half were a single vertex, the other would contain all a old triples used, but it contains at most $a - 1$; so both halves are nontrivial, with $\sigma + \tau = a - 2$ leaves. If $\sigma, \tau > 0$, T contains a path with five edges. Adding the absent triple

$\{u, v, x\}$, with x in a pendant triple at u, leaves only three vertices of degree at least two, while such a path has four internal vertices; so this is impossible. Hence T has a vertex of degree $a-1$ with $a-2$ leaves and a pendant path of length three, and $a-2 \le \max(p,q)$ forces $\min(p,q) = 1$; say $p = a-2$, $q = 1$, and let f be the pendant triple at v. Add a triple e of three non-hub vertices of pendant triples at u (there are $2p \ge 6$ of them). In $H + e$ only u has degree at least three, namely $a-1$, so the vertex of degree $a-1$ of T is mapped to u and its edges use all $a-1$ triples containing u. The two remaining edges of T are consecutive on the pendant path but must be represented by the disjoint triples e and f, a contradiction. □

6.3 Proof of Theorem 4

Let K be Berge-T-saturated on n vertices with m hyperedges. A component with an absent triple inside it has at least a hyperedges, since adding the triple creates a connected copy with $a+1$ hyperedges within that component. An incidence-tree component with $b \ge 2$ triples has $\binom{2b+1}{3} > b$ triples available, so $b \ge a$. Adding a triple on three isolated vertices, on two vertices of one single-triple component and one of another, or on an isolated vertex and two vertices of a single-triple component creates components with at most three hyperedges and changes nothing else, so K has at most two isolated vertices, at most one single-triple component, and not both.

Put $\alpha = a/(2a+1)$ and compare each component with density α. Incidence-tree components with $b \ge a$ triples have $b/(2b+1) \ge \alpha$; other components have at least half as many hyperedges as vertices; two isolated vertices fall short by 2α and a single-triple component by $3\alpha - 1 < 2\alpha$. Hence $m \ge \alpha(n-2)$, the first inequality.

Now let $T \ncong B_a$, let h be the number of incidence-tree components with exactly a triples, and put $\beta = (a+1)/(2a+3)$. Each such component is internally saturated, so it is the expansion of a star or a double star, of at most $\lceil a/2 \rceil$ types. No type occurs three times: three isomorphic components form a Berge-T-free hypergraph in which every absent triple creates a copy (the copy lies in the union of the three components and the new triple), contradicting Lemma 14. So $h \le 2\lceil a/2 \rceil \le a+1$.

If $h \ge 1$, K has at most one isolated vertex. Indeed, let H be the expansion of G as above, c a nonleaf vertex of G (one exists since $a \ge 2$) and z_1, z_2 isolated vertices, and add $\{c, z_1, z_2\}$. A new copy of T uses the $a+1$ hyperedges of the augmented component, and its new pair is cz_1 or cz_2, as $z_1 z_2$ would be disconnected from the rest. Removing that leaf leaves a connected selection using every triple of H once. Such a selection keeps both nonleaf vertices of a central triple, since otherwise one side of the corresponding edge of G would be cut off, and the hub of every pendant triple; so it is isomorphic to G, and T is G with a leaf added at a nonleaf vertex, giving $L(T) = L(G) + 1$, against Lemma 13.

Compare with density β. Incidence-tree components with at least $a+1$ triples and the components that are not incidence trees have density at least β; each of the h components with a triples falls short by $\beta(2a+1) - a = 1/(2a+3)$. If $h = 0$ the small components fall short by at most 2β; if $h \ge 1$ by at most β, since a single-triple component falls short by $3\beta - 1 = a/(2a+3) < \beta$. So the total shortfall is at most 2β if $h = 0$, and at most $(h+a+1)/(2a+3) \le 2(a+1)/(2a+3) = 2\beta$ if $h \ge 1$; in both cases $m \ge \beta(n-2)$.

The limit for B_a follows from this lower bound and the value $\lfloor a(n+1)/(2a+1) \rfloor$ of Theorem 1 with $r = 3$, and $\beta - \alpha = 1/((2a+1)(2a+3))$. Finally

$$\frac{(a+1)(n-2)}{2a+3} - \frac{a(n+1)}{2a+1} = \frac{n - (6a^2+9a+2)}{(2a+1)(2a+3)} > 0$$

for $n \ge 6a^2 + 9a + 3 = 3(a+1)(2a+1)$, so the ceiling on the left exceeds $\lfloor a(n+1)/(2a+1) \rfloor = \mathrm{sat}_3(n, \text{Berge-}B_a)$. □

7 Formal verification

The Lean project in `lean/` pins Lean 4.35.0 and Mathlib revision `56a0058`. Hypergraphs are `UniformHypergraph V r`: a finite set of finite sets of vertices of cardinality r, on an ambient vertex type, so multiple hyperedges are excluded and isolated vertices are kept. `BroomCopy a E` consists of an injection $\texttt{Fin}(a+2) \hookrightarrow V$ and an injection of $\texttt{Fin}(a+1)$ into the hyperedges of E, with both endpoints of the ith edge of B_a in the ith hyperedge; the edges are 0–1, 1–2 and 0–$(i+1)$ for $i \geq 2$, which is B_a with centre 0, stem 1 and terminal 2. `Saturated a H` requires that there is no such copy and that one exists after inserting any r-set that is not a hyperedge. The saturation number is

$$\texttt{satNumber a r n} = \inf\{m : \exists H \text{ on } \texttt{Fin}\, n,\ \texttt{Saturated a H} \wedge |E(H)| = m\}.$$

The declarations `Berge.satNumber_eq` and `Berge.exact_statement_expanded` state Theorem 1 for all $a, r \geq 3$ and $n \geq (r-1)a + 1$, the second with all definitions unfolded, and `Berge.near_minimum_structure` states Theorem 3 with the remainder defined as the total order of the components that are not loose stars with at least a spokes. The proofs follow Sections 2–5: the construction is built on an explicit vertex type and transported to $\texttt{Fin}\, n$, and the lower bound formalizes Lemmas 5 and 6 and identity (1) through the connected components of the incidence graph. A module-level audit enumerates every declaration of the fourteen project modules and checks that each depends only on `propext`, `Classical.choice` and `Quot.sound`; the sources contain no `sorry` and no `native_decide`. Theorem 2, Proposition 11 and Theorem 4 are not formalized.

8 Prior work and limitations

The definitions are those of [5]. Theorem 4 is the three-uniform Berge analogue of Theorem 5 of [9], with the same minimizer, and it uses the same strategy of comparing tree components; the ordinary-graph theorem does not imply it, since an added triple can meet three components and a minimum incidence tree need not be the expansion of an ordinary tree before Lemma 13 is proved. Linearity of $\mathrm{sat}_r(n, \text{Berge-}T)$ for trees is due to [7]. A search of the Berge-saturation literature through October 2026, including the papers cited here, found no exact value for B_a and no comparison of Berge saturation numbers across trees, but the search is not exhaustive and priority is not claimed. Theorem 4 is proved only for $r = 3$ and $a \geq 5$; we do not know whether B_a minimizes among trees in higher uniformities, nor the exact values for other trees.

9 Research process and AI disclosure

This is fully autonomous agentic research. The results, their proofs, the original Lean development (for Lean 4.19.0) and the literature search were produced by autonomous AI research agents in an earlier session. Its records describe a fresh internal review, which rederived the comparison among trees and reran the Lean build, as an Astra review; they do not otherwise name the system used.

Opus 5.5 by Anthropic used with Claude Code assessed the package for this release. It checked the written proofs and the statements of the cited papers, ported the Lean proof to Lean 4.35.0, added the definition of the saturation number as a minimum and the headline statements, ran the fresh build and axiom audit, reran the upload's finite checks, and wrote this note. The editorial revision of 1 November 2026, also made with Opus 5.5 by Anthropic used with Claude Code, changed the presentation only, not the mathematics or the formalization.

Dev D. Goyal is responsible for publication. The work has not been externally peer reviewed; no expert validation is claimed, AI checks are not human peer review, and the novelty assessment is provisional. Section 7 states the formal scope, and Section 8 the remaining limits.

References

[1] B. Austhof and S. English, *Nearly-regular hypergraphs and saturation of Berge stars*, Electronic Journal of Combinatorics **26**(4) (2019), P4.49. arXiv:1812.00472.

[2] N. Bushaw, S. English, E. Heath, D. P. Johnston and P. Rombach, *The saturation spectrum of Berge stars*, arXiv:2502.17686 (2025).

[3] X. Cao, Y. Lai and C. Wang, *Saturation in three-uniform Berge hypergraphs: the five-star spectrum and the K_4 saturation number*, arXiv:2610.32353 (2026).

[4] Y. Chen, J. He and T. Xie, *Saturation numbers for 3-uniform Berge-K_4*, arXiv:2601.18455 (2026).

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

[6] 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.

[7] 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.

[8] S. English, J. Kritschgau, M. Nahvi and E. Sprangel, *Saturation numbers for Berge cliques*, European Journal of Combinatorics **118** (2024), 103911. arXiv:2301.02973.

[9] J. Faudree, R. J. Faudree, R. J. Gould and M. S. Jacobson, *Saturation numbers for trees*, Electronic Journal of Combinatorics **16** (2009), R91. doi:10.37236/180.

[10] T. Xie, *A linear upper bound for Berge saturation numbers*, arXiv:2610.11409 (2026).