ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.)
ARS MATHEMATICA CONTEMPORANEA 24 (2024) #P4.01
https://doi.org/10.26493/1855-3974.3109.e4b
(Also available at http://amc-journal.eu)

Connected Turán number of trees

Yair Caro
Department of Mathematics, University of Haifa-Oranim, Israel

Balázs Patkós *
HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary

Zsolt Tuza †
HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary and
University of Pannonia, Veszprém, Hungary

Received 24 April 2023, accepted 18 September 2023, published online 23 September 2024

Abstract

The connected Turán number is a variant of the much studied Turán number, $\mathrm{ex}(n, F)$, the largest number of edges that an n-vertex F-free graph may contain. We start a systematic study of the connected Turán number $\mathrm{ex}_c(n, F)$, the largest number of edges that an n-vertex *connected* F-free graph may contain. We focus on the case where the forbidden graph is a tree. Prior to our work, $\mathrm{ex}_c(n, T)$ was determined only for the case T is a star or a path. Our main contribution is the determination of the exact value of $\mathrm{ex}_c(n, T)$ for small trees, in particular for all trees with at most six vertices, as well as some trees on seven vertices and several infinite families of trees. We also collect several lower-bound constructions of connected T-free graphs based on different graph parameters.

The celebrated conjecture of Erdős and Sós states that for any tree T, we have $\mathrm{ex}(n, T) \leq (|T| - 2)\frac{n}{2}$. We address the problem how much smaller $\mathrm{ex}_c(n, T)$ can be, what is the smallest possible ratio of $\mathrm{ex}_c(n, T)$ and $(|T| - 2)\frac{n}{2}$ as $|T|$ grows.

Keywords: Extremal graph theory, connected host graphs, trees.

Math. Subj. Class. (2020): 05C35

*Corresponding author. Partially supported by NKFIH grants SNN 129364 and FK 132060.
†Partially supported by NKFIH grant SNN 129364.
E-mail addresses: yacaro@kvgeva.org.il (Yair Caro), patkos@renyi.hu (Balázs Patkós), tuza.zsolt@mik.uni-pannon.hu (Zsolt Tuza)

This work is licensed under https://creativecommons.org/licenses/by/4.0/

1 Introduction

For a graph G, we write $e(G)$ and $|G|$ to denote the number of edges and vertices in G. For a pair U, V of disjoint sets of vertices in G, we use $e_G(U, V)$ to denote the number of edges in G with one endpoint in U and the other in V. G will be omitted from the subscript if it is clear from context.

One of the most studied problems in extremal graph theory is to determine the Turán number $\mathrm{ex}(n, F)$, the largest number of edges that an n-vertex graph can have without containing a subgraph isomorphic to F. In this paper, we start a systematic study of a variant of this parameter: the connected Turán number $\mathrm{ex}_c(n, F)$ is the largest number of edges that a *connected* n-vertex graph can have without containing F as a subgraph. Observe that if F is 2-edge-connected, then any maximal F-free graph G is connected, as if G has at least two connected components, then adding an edge between them would not create any copy of F. Also, if the chromatic number of F is at least 3, then by the famous theorem by Erdős, Stone, and Simonovits [6, 7], we know that $\mathrm{ex}(n, F)$ is attained asymptotically (and for some graphs precisely) at the Turán graph that is connected. These two observations imply the following proposition.

Proposition 1.1.

(1) *If all connected components of F are 2-edge-connected, then* $\mathrm{ex}(n, F) = \mathrm{ex}_c(n, F)$.

(2) *If* $\chi(F) \geq 3$, *then* $\mathrm{ex}_c(n, F) = (1 + o(1))\,\mathrm{ex}(n, F)$.

The asymptotics of $\mathrm{ex}(n, F)$ is unknown for most bipartite F (for a general overview of the so-called degenerate Turán problems, see the survey by Füredi and Simonovits [8]). However, it is known that for any graph F that contains a cycle, $\mathrm{ex}(n, F)$ grows superlinearly. If $\mathrm{ex}(n, F)$ is attained at a non-connected graph with a connected component of size m, then we have $\mathrm{ex}(n, F) \leq \mathrm{ex}(m, F) + \mathrm{ex}(n - m, F)$, which does not hold for 'nice' superlinear functions. There is a relatively large literature on the Turán number of forests (see e.g. [3, 11, 12, 14, 15]), and in many cases the extremal graphs turned out to be connected, so for those forests F, we have $\mathrm{ex}(n, F) = \mathrm{ex}_c(n, F)$. In this paper, we concentrate on the family of trees. A famous conjecture of Erdős and Sós (that appeared in print first in [4]) states that any n-vertex graph with more than $\frac{(k-2)n}{2}$ edges contains any tree T on k vertices. A proof was announced in the early 1990's by Ajtai, Komlós, Simonovits, and Szemerédi, but only arguments of special cases have appeared. A recent survey of these and other degree conditions that imply embeddings of trees is given in [13]. The universal construction that shows the tightness of the Erdős–Sós conjecture is the union of vertex-disjoint cliques of size $k - 1$. This is not a connected graph and we are only aware of two explicit results concerning $\mathrm{ex}_c(n, T)$ (but there exist results on Turán problems in connected host graphs, see e.g. [2]). The connected Turán number of stars follows from the existence of (nearly) regular connected graphs. Apart from stars, paths on k vertices, denoted by P_k, have been considered. The value of $\mathrm{ex}_c(n, P_k)$ was determined by Kopylov, and independently by Balister, Győri, Lehel, and Schelp with the latter group also showing the uniqueness of extremal constructions.

Theorem 1.2 ([1, Balister, Győri, Lehel, Schelp], [10, Kopylov])**.** *If G is an n-vertex connected graph that does not contain any paths on $k + 1$ vertices, then*

$$e(G) \leq \max\left\{\binom{k-1}{2} + n - k + 1, \binom{\lceil \frac{k+1}{2} \rceil}{2} + \left\lfloor \frac{k-1}{2} \right\rfloor \left(n - \left\lceil \frac{k+1}{2} \right\rceil\right)\right\}$$

holds.

In the remainder of the introduction, we shall present the various results obtained concerning $\mathrm{ex}_c(n,T)$. Lower bound constructions are given in Section 2 and exact determination of $\mathrm{ex}_c(n,T)$ including all trees on up to six vertices and some trees having seven vertices is included in Section 3.

Our first result gathers several constructions, all based on some graph parameters, that provide lower bounds on $\mathrm{ex}_c(n,T)$. For those parameters we use the following notation.

Definition 1.3.

- $p(G)$ denotes the maximum number of vertices in a path P of G such that for all $x \in V(P)$ we have $d_G(x) \leq 2$.
- $\Delta(G)$ and $\delta(G)$ denote the maximum and the minimum degree in G.
- $\nu(G)$ denotes the number of edges in a largest matching of G.
- $\delta_2(T)$ denotes the smallest degree in T that is larger than 1.
- For a vertex $v \in V(T)$ let $m_T(v)$ be the size of largest component of $T - v$ and let $m(T) = \min\{m_T(v) : v \in V(T)\}$.
- For a vertex $v \in V(T)$ let $m_{T,2}(v)$ be the sum of the sizes of two largest components of $T - v$ and let $m_2(T) = \min\{m_{T,2}(v) : v \in V(T)\}$.
- For an edge $e = xy \in E(G)$ we write $w(e) = \min\{d_G(x), d_G(y)\}$ and define $w(G) = \max\{w(e) : e \in E(G)\}$.

Notation. For graphs H and G, their disjoint union is denoted by $H \cup G$. The join of H and G, denoted by $H + G$, is $H \cup G$ with all edges hg $h \in H, g \in G$ added. For a graph H and a positive integer k, kH denotes the pairwise vertex-disjoint union of k copies of H. S_k denotes the star with k leaves, P_k, C_k, K_k, E_k denote the path, the cycle, the complete graph and the empty graph on k vertices, respectively. The complete bipartite graph with parts of size a and b is denoted by $K_{a,b}$.

In the following remark, we gather the consequences of known constructions (mostly used for distinct purposes).

Remark 1.4.

(1) The existence of connected (nearly)-regular graphs show $\mathrm{ex}_c(n,T) \geq \lfloor \frac{n(\Delta(T)-1)}{2} \rfloor$.

(2) The construction of Theorem 1.2 shows that if T has diameter d, then $\mathrm{ex}_c(n,T) \geq \binom{\lceil \frac{d+1}{2} \rceil}{2} + \lfloor \frac{d-1}{2} \rfloor (n - \lceil \frac{d+1}{2} \rceil)$.

(3) $K_{a-1,n-a+1}$ shows that if the bipartition of T consists of classes of sizes a and b with $a \leq b$, then $\mathrm{ex}_c(n,T) \geq (a-1)(n-a+1)$. In particular, we have $\mathrm{ex}_c(n,T) \geq (w(T)-1)(n-w(T)+1)$ and $\mathrm{ex}_c(n,T) \geq (\nu(T)-1)(n-\nu(T)+1)$. The latter can be improved to $\mathrm{ex}_c(n,T) \geq (\nu(T)-1)(n-\nu(T)+1) + \binom{\nu(T)-1}{2}$ shown by $K_{\nu(T)-1} + E_{n-\nu(T)+1}$, the largest graph with matching number less than $\nu(T)$ if n is large as proved by Erdős and Gallai [5].

Observe that if T is balanced, i.e. $a = b$ in its bipartition, then the number of edges in $K_{a-1,n-a+1}$ is just a constant smaller than the number of edges in $\frac{n}{k-1}K_{k-1}$, the extremal graph of the Erdős-Sós conjecture. The next proposition states lower bounds due to new constructions. The proof of Proposition 1.5 will be given in Scetion 2.

Proposition 1.5. *Suppose T is a tree on $k \geq 4$ vertices.*

(1) *If T is not a path and thus $p(T) \leq k-3$, then $\mathrm{ex}_c(n,T) \geq (\binom{k-2p(T)-3}{2} + p(T)+2)\lfloor \frac{n}{k-p(T)-2} \rfloor$. Furthermore, if T contains at least two vertices of degree at least three, then $\mathrm{ex}_c(n,T) \geq \frac{\binom{k-p(T)-1}{2}+p(T)+2}{k} n - O(k)$.*

(2) *If T is not a star and $\delta_2(T) > 2$, then $\mathrm{ex}_c(n,T) \geq \lfloor \frac{n-1}{k-1} \rfloor (\binom{k-2}{2} + \delta_2(T) - 1)$.*

(3) *If T is not a path, then $\mathrm{ex}_c(n,T) \geq n - 1 + \lfloor \frac{n-1}{m(T)-1} \rfloor \binom{m(T)-1}{2}$.*

(4) *$\mathrm{ex}_c(n,T) \geq \lfloor \frac{n}{k-m_2(T)} \rfloor (1 + \binom{k-m_2(T)}{2})$.*

Next, we determine $\mathrm{ex}_c(n,T)$ for all trees on k vertices with $4 \leq k \leq 6$ (note that there do not exist P_3-free connected graphs), some trees on 7 vertices and for some infinite families of trees. We need some notation first.

Let $D_{a,b}$ denote the *double star* on $a+b+2$ vertices such that the two non-leaf vertices have degree $a+1$ and $b+1$. $S_{a_1,a_2,\dots,a_j}$ with $j \geq 3$ denotes the *spider* obtained from j paths with $a_1, a_2, \dots, a_j$ edges by identifying one endpoint of all paths. So $S_{a_1,a_2,\dots,a_j}$ has $1 + \sum_{i=1}^{j} a_i$ vertices and maximum degree j. The only vertex of degree at least 3 is the *center* of the spider, the maximal paths starting at the center are the *legs* of the spider. M_n denotes the matching on n vertices (so if n is odd, then an isolated vertex and $\lfloor \frac{n}{2} \rfloor$ isolated edges).

The values of $\mathrm{ex}_c(n, P_{k+1})$ were determined by Theorem 1.2, and for $k \geq 3$, the statement $\mathrm{ex}_c(n,S_k) = \lfloor \frac{n(k-1)}{2} \rfloor$ follows from Remark 1.4(1) and that the degree-sum of an S_k-free graph is at most $n(k-1)$. So in the next theorem, we only list those trees that are neither paths nor stars. In particular, all trees have 5 or 6 vertices. Proofs of the following theorems will be given in Section 3.

Theorem 1.6. *For non-star, non-path trees with 5 or 6 vertices, the following exact results are valid.*

(1) *For any $T = S_{2,1,\dots,1}$ we have $\mathrm{ex}_c(n,T) = \lfloor \frac{n(\Delta(T)-1)}{2} \rfloor$ if $n \geq |T|$. In particular, $\mathrm{ex}_c(n, S_{2,1,1}) = n$ if $n \geq 5$ and $\mathrm{ex}_c(n, S_{2,1,1,1}) = \lfloor \frac{3n}{2} \rfloor$ if $n \geq 6$.*

(2) *We have $\mathrm{ex}_c(n, D_{2,2}) = 2n-4$ if $n \geq 6$.*

(3) *We have $\mathrm{ex}_c(n, S_{3,1,1}) = \lfloor \frac{3(n-1)}{2} \rfloor$ if $n \geq 7$ and $\mathrm{ex}(6, S_{3,1,1}) = 9$.*

(4) *We have $\mathrm{ex}_c(n, S_{2,2,1}) = 2n-3$ if $n \geq 6$.*

Let $D^*_{2,2}$ be the tree obtained from $D_{2,2}$ by attaching a leaf to one leaf of $D_{2,2}$.

Theorem 1.7. *$ex_c(n, D^*_{2,2}) = 2n-3$ for all $n \geq 7$, and $ex_c(n, D^*_{2,2}) = \binom{n}{2}$ for $1 \leq n \leq 6$.*

Theorem 1.8. *$ex_c(n, S_{2,2,2}) = 2n-2$ for all $n \geq 7$, and $ex_c(n, S_{2,2,2}) = \binom{n}{2}$ for $1 \leq n \leq 6$.*

Theorem 1.9. *$ex_c(n, S_{3,2,1}) = 2n-3$ for all $n \geq 7$, and $ex_c(n, S_{3,2,1}) = \binom{n}{2}$ for $1 \leq n \leq 6$.*

Theorem 1.10. *For any $T = S_{3,1,\dots,1}$ with $\Delta(T) \geq 4$, if n is large enough, then* $\mathrm{ex}_c(n,T) = \lfloor \frac{(\Delta(T)-1)n}{2} \rfloor$.

The broom, which we denote by $B(k,a)$, is the special spider $S_{a-1,1,1,\dots,1}$ on k vertices. So its maximum degree is $k-a+1$ and its diameter is a.

Theorem 1.11.

(1) *For any $a \leq k-2$,* $\mathrm{ex}_c(n,B(k,a)) \geq \max\{\lfloor \frac{(k-a)n}{2} \rfloor, \binom{\lceil \frac{a+1}{2} \rceil}{2} + \lfloor \frac{a-1}{2} \rfloor (n - \lfloor \frac{a+1}{2} \rfloor)\}$ *holds.*

(2) *For any $a \leq k/3$,* $\mathrm{ex}_c(n,B(k,a)) = \lfloor \frac{(k-a)n}{2} \rfloor$ *holds if n is large enough.*

For a better overview, we include tables with previous results, our results and open cases for trees on up to 7 vertices. $SD_{2,2}$ denotes the tree on 7 vertices obtained from the double star $D_{2,2}$ by subdividing the edge connecting its two centers.

Number of vertices	Tree	$\mathrm{ex}_c(n,T)$	Construction
4	P_4	$n-1$	S_{n-1}
	S_3	n	C_n
5	P_5	n	$K_1 + (K_2 \cup E_{n-3})$
	S_4	$\lfloor \frac{3n}{2} \rfloor$	(nearly) 3-regular
	$S_{2,1,1}$	n	C_n
6	P_6	$2n-3$	$K_2 + E_{n-2}$
	S_5	$2n$	4-regular
	$S_{2,1,1,1}$	$\lfloor \frac{3n}{2} \rfloor$	(nearly) 3-regular
	$S_{2,2,1}$	$2n-3$	$K_2 + E_{n-2}$
	$S_{3,1,1}$	$\lfloor \frac{3(n-1)}{2} \rfloor$	$K_1 + M_{n-1}$
	$D_{2,2}$	$2n-4$	$K_{2,n-2}$

Table 1: The value of $\mathrm{ex}_c(n,T)$ for all trees up to 6 vertices.

Tree	$\mathrm{ex}_c(n,T)$	Construction	Tree	$\mathrm{ex}_c(n,T)$	Construction
S_6	$\lfloor \frac{5n}{2} \rfloor$	(nearly) 5-regular	P_7	$2n-2$	$K_2 + (E_{n-4} \cup K_2)$
$S_{4,1,1}$	$\geq 2n-3$	$K_2 + E_{n-2}$	$S_{3,2,1}$	$2n-3$	$K_2 + E_{n-2}$
$S_{3,1,1,1}$	$\lfloor \frac{3n}{2} \rfloor$	(nearly) 3-regular	$S_{2,1,1,1,1}$	$2n$	4-regular
$S_{2,2,2}$	$2n-2$	$K_2 + (E_{n-4} \cup K_2)$	$S_{2,2,1,1}$	$\geq 2n-3$	$K_2 + E_{n-2}$
$D^*_{2,2}$	$2n-3$	$K_2 + E_{n-2}$	$D_{2,3}$	$\geq 2n-4$	$K_{2,n-2}$
$SD_{2,2}$	$\geq \frac{13n}{7} - O(1)$	Proposition 1.5(1)	$D_{2,3}$	$\geq 2n-2$ if $6 \mid n-1$	Proposition 1.5(2)

Table 2: Exact values and lower bounds on $\mathrm{ex}_c(n,T)$ for trees with 7 vertices.

The starting point of our final subtopic is the Erdős–Sós conjecture, $\mathrm{ex}(n,T) = \frac{k-2}{2}n + O_k(1)$. We would like to know how much smaller $\mathrm{ex}_c(n,T)$ can be than $\mathrm{ex}(n,T)$. For any tree T we introduce

$$\gamma_T := \limsup_n \frac{2}{|T|-2} \cdot \frac{\mathrm{ex}_c(n,T)}{n}$$

where $|T|$ denotes the number of vertices in T. It is well-known that any graph with average degree at least $2d$ contains a subgraph with minimum degree at least d. Also, any tree on k vertices can be embedded to any graph with minimum degree at least k. This shows that $\gamma_T \leq 2$ for any tree T on k vertices. The Erdős–Sós conjecture would imply $\gamma_T \leq 1$.

Let $\mathcal{T}_k$ denote the set of trees on at least k vertices. We write $\gamma_k := \inf\{\gamma_T : T \in \mathcal{T}_k\}$ and $\gamma := \lim_{k\to\infty} \gamma_k$. Observe that γ_k is monotone increasing as $\mathcal{T}_2 \supset \mathcal{T}_3 \supset \mathcal{T}_4 \supset \dots$, and thus the limit γ exists.

Theorem 1.12. *The following upper and lower bounds hold:* $\frac{1}{3} \leq \gamma \leq \frac{2}{3}$.

2 Constructions

Proof of Proposition 1.5. For all lower bounds we need constructions.

For the general lower bound of (1), we construct a graph $G(V, E)$ as follows: let $s := \lfloor \frac{n}{k-p(T)-2} \rfloor$ and let V be partitioned into $\bigcup_{i=1}^{s}(A_i \cup Q_i)$ with $|A_i| = k - 2p(T) - 3$ for all $1 \leq i \leq s$, $|Q_i| = p(T) + 1$ for all $1 \leq i < s$. $G[A_i]$ is a clique for all i. Every clique A_i contains a special vertex x_i, and $G[\{x_i, x_{i+1}\} \cup Q_i]$ is a path with end vertices x_i and x_{i+1} (with $x_{s+1} = x_1$). Then G cannot contain T, as a copy of T could contain the vertices of an A_i and then at most $p(T)$ vertices from both of Q_{i-1} and Q_i, so at least one vertex of T cannot be embedded.

To see the furthermore part of (1), we have the following construction G: we partition the vertex set of G into $\{v\} \cup \bigcup_{i=1}^{s}(A_i \cup Q_i)$, where $s = \lceil \frac{n-1}{k} \rceil$ with $|A_i| = k - p(T) - 1$, $|Q_i| = p(T) + 1$ for all $1 \leq i < s$, and $|Q_i| \leq p(T) + 1$ and if $|A_i| > 0$, then $|Q_i| = p(T) + 1$. The edges of G are defined such that $G[\{v\} \cup \bigcup_{i=1}^{s} Q_i]$ is a spider with center v and legs Q_i, $G[A_i]$ is a clique and exactly one vertex of A_i is connected to the leaf of the leg in Q_i. The number of edges adjacent to $A_i \cup Q_i$ is $\binom{k-p(T)-1}{2} + p(T) + 2$, therefore $e(G)$ is as claimed. Finally, to see that G is T-free, observe that as T contains at least two vertices of degree at least 3, if G contained a copy of T, then this copy should contain a vertex u from one of the A_is. Also, such a copy cannot contain all vertices of Q_i as $p(T) < |Q_i|$. Therefore, the vertices of the copy of T should be contained in $|A_i| + |Q_i| - 1 < k$ vertices — a contradiction.

The lower bound of (2) is shown by the following construction of a connected n-vertex T-free graph G: we partition the vertex set of G into $\{v\} \cup \bigcup_{i=1}^{\lceil \frac{n-1}{k-1} \rceil} A_i$ with $|A_i| = k - 1$ for all $i = 1, 2, \dots, \lfloor \frac{n-1}{k-1} \rfloor$ and every A_i containing a special vertex x_i. The edges of G are defined as follows: $G[A_i \setminus \{x_i\}]$ is a clique, v is adjacent to all x_i, and x_i is adjacent to $\delta_2(T) - 2$ other vertices of A_i, so $d_G(x_i) = \delta_2(T) - 1$. We claim that G is T-free. Indeed, as $G - v$ has components of size at most $k - 1$, a copy of T must contain v. As T is not a star, at least one of v's neighbors is not a leaf and so its degree should be at least $\delta_2(T)$. But all v's neighbors are x_i vertices that have degree $\delta_2(T) - 1$ in G. The number of edges in $G[\{v\} \cup \bigcup_{i=1}^{\lfloor \frac{n-1}{k-1} \rfloor} A_i]$ is $\lfloor \frac{n-1}{k-1} \rfloor (\binom{k-2}{2} + \delta_2(T) - 1)$.

The construction yielding the lower bound of (3) is $G = K_1 + (rK_{m(T)-1} \cup K_s)$, where $r = \lfloor \frac{n-1}{m(T)-1} \rfloor$ and $s \geq 0$. Indeed, if G contained a copy of T, then this copy should contain the vertex v of K_1 as otherwise T would be contained in $m(T) - 1$ vertices. But then we cannot embed the largest branch pending on v as it has size at least $m(T)$.

To obtain the construction yielding the lower bound of (4), we partition the vertex set to $A_1, A_2, \dots, A_s, A_{s+1}$ with $s = \lfloor \frac{n}{k-m_2(T)} \rfloor$ and $|A_i| = k - m_2(T)$ for all $i = 1, 2, \dots, s$. As T is not a path, we have $k - m_2(T) \geq 2$, so in each A_i we can pick two distinct vertices

x_i, y_i, maybe with the exception of A_{s+1}. Then we define G as a "cycle of cliques", so $G[A_i]$ is a clique for all i, and $x_i y_{i+1}$ is an edge (formally there should be three cases depending whether A_{s+1} has size 0, 1, or at least 2). To see that G is T-free, consider the vertex v with $m_2(T) = m_{T,2}(v)$, i.e. the largest two connected components B_1, B_2 in $T - v$ have a total size of $m_2(T)$. Suppose G contains a copy of T and the vertex playing the role of v belongs to A_i. Then, as there are only two edges leaving A_i, T apart from two components of $T - v$ must be embedded into A_i. Moreover, since the two edges leave from distinct vertices, at least one vertex of the two exceptional components must also be embedded to A_i. So A_i should contain at least $k - m_2(T) + 1$ vertices — a contradiction. (If $i = s + 1$ and $x_i = y_i$, then we have the same contradiction, as then A_{s+1} should contain at least $k - m_2(T)$ vertices, but A_{s+1} is strictly smaller than that.) □

3 Proofs

We start by proving Theorem 1.6. We restate and prove its parts separately.

Theorem 3.1. *For $T = S_{2,1,\dots,1}$, the equality $\mathrm{ex}_c(n,T) = \lfloor \frac{n(\Delta(T)-1)}{2} \rfloor$ holds if $n \geq |T|$.*

Proof. The constructions giving the lower bounds are connected (nearly) regular graphs of degree $\Delta(T) - 1$.

If $T = S_{2,1,1,\dots,1}$, then the upper bound proof is a special case of Theorem 1.11, but for completeness, we give a simpler proof of this case. If G is a connected, n-vertex, T-free graph and for some x we have $d_G(x) \geq \Delta(T)$, then G is the star. Indeed, the neighbors of x can be adjacent only to other neighbors of x, otherwise T would be a subgraph of G. So by connectivity $N_G[x] = V(G)$. But then if there is at least one edge between two neighbors of x, then, as $|V(G)| \geq |V(T)|$, again T would be a subgraph of G. The star has fewer edges than the claimed maximum, so to have $\mathrm{ex}_c(n,T)$ edges, G must be (nearly) $(\Delta(T) - 1)$-regular. □

Theorem 3.2. *For any $n \geq 6$, $\mathrm{ex}_c(n, D_{2,2}) = 2n - 4$ holds.*

Proof. To see the lower bound, observe that $K_{2,n-2}$ is $D_{2,2}$-free as $w(K_{2,n-2}) = 2$, while $w(D_{2,2}) = 3$.

To see the upper bound, observe first that all connected graphs with 6 vertices and at least 9 edges contain a copy of $D_{2,2}$ as can be checked in the table of graphs of [9] on pages 222–224.

Suppose there exists a minimum counterexample: a connected graph G on $n \geq 7$ vertices and $e(G) \geq 2n - 3$ edges with no copy of $D_{2,2}$. We consider several cases.

CASE I: $\delta(G) \leq 2$ and there is a vertex v of degree at most 2 which is not a cut vertex.

Delete this vertex v of degree 1 or 2 to obtain a connected $H = G \setminus v$ with $|H| \geq 6$. By minimality $e(H) \leq 2(n-1)–4$ and $2n-3 \leq e(G) \leq e(H)+2 \leq 2(n-1)–4+2 = 2n-4$ — a contradiction.

CASE II: $\delta(G) = 2$ and every vertex of degree 2 is a cut vertex.

Consider v of degree 2 such that in $H = G - v$ out of the two components A and B, $|A|$ is as small as possible. Let w be the vertex in A adjacent to v and let z be the vertex in B adjacent to v.

If $|A| \geq 6$ then by minimality of G, $2n - 3 \leq e(G) \leq 2|A| - 4 + 2|B| - 4 + 2 = 2(|A| + |B| + 1) - 8 = 2n - 8$ — a contradiction. Otherwise $3 \leq |A| \leq 5$ as $|A| \leq 2$

would imply $\delta(G) = 1$ and we were in Case I. Also, $|A| \geq 4$ as $|A| = 3$ would imply that A must contain a vertex of degree 2 which is not a cut vertex and we were in Case I again.

Suppose $|A| = 5$. If $d_G(w) = 2$ then $|A|$ is not minimum, so in the induced subgraph on A all vertices have degree at least 2 and $d_G(w) \geq 3$. But then the induced graph on A either contains a vertex of degree 2 which is not a cut vertex and we are in Case I or all degrees in $G[A \cup \{v\}]$ (except for v) are at least 3. Then one can find a copy of $D_{2,2}$ with w being one of the centers and v being a leaf pending from w. Indeed, by the degree condition, $G[A \setminus \{w\}]$ contains a C_4, so if $N(w)$ contains two non-neighbor vertices x, y of this C_4, then x can be the other center of the copy of $D_{2,2}$ and y the other leaf pending from w. Otherwise w has exactly two neighbors in A, and then by the degree condition $G[A \setminus \{w\}]$ is K_4 and it is trivial to embed $D_{2,2}$.

Finally suppose $|A| = 4$. As $|B| \geq |A| = 4$, it follows that $B^* = B \cup \{v, w\}$ has at least 6 vertices and $|B^*| = n - 3$, and hence by minimality of G, $e(B^*)$ contains at most $2(n-3)-4$ edges and together with at most 6 edges in A gives $e(G) \leq 2n-10+6 = 2n-4$ — a contradiction.

CASE III: $\delta(G) \geq 3$.

If all vertices are of degree 3, we have $3n/2$ edges, which is at most $2n - 4$ for $n \geq 8$. For $n = 7$ this is impossible by parity, hence $\delta(G) \geq 3$ and $\Delta(G) \geq 4$. Consider an edge $e = xy$ with $d_G(y) = \Delta(G) \geq 4$ and $d_G(x) \geq 3$.

If $d_G(y) \geq 5$, then for $u, u' \in N(x)$ we have $|N(y) \setminus \{x, u, u'\}| \geq 2$, so x and y are centers of a copy of $D_{2,2}$. If $d_G(y) = 4$ and $d_G(x) = 4$ then either x and y have distinct neighbors s not in $N[y]$ and t not in $N[x]$ and we find a copy of $D_{2,2}$ with centers x, y, or x and y are twins having the same neighbors a, b, c excluding themselves. But as $|G| \geq 7$, at least one vertex, say a, has a neighbor d not adjacent to the other 4 vertices and then a and x can be centers of $D_{2,2}$ with y and d pending from a.

So we can assume that all vertices have degree 3 or 4 and vertices of degree 4 form an independent set Q. Let $P = V \setminus Q$, and consider the bipartite $G[P, Q]$ where $p + q = n$, $|P| = p$ and $|Q| = q$. Clearly, $4q = e(P, Q) \leq 3p$. Hence $3n = 3q + 3p \geq 7q$ and $q \leq 3n/7, p \geq 4n/7$. But then

$$e(G) = \frac{4q + 3p}{2} \leq \frac{12n/7 + 12n/7}{2} = \frac{12n}{7} < 2n - 3$$

for $n \geq 11$. So we are left with $n = 7, 8, 9, 10$.

For $n = 7$: $q \leq 3n/7 = 3$ and q must be an integer. If $q = 3$, then $G = K_{4,3}$ containing $D_{2,2}$. The case $q = 2$ is impossible as the degree sum would be odd (by the number p of odd-degree vertices). Hence $q = 1$ and $p = 6$. Consider a vertex v of degree 4 and its neighbors a, b, c, d all of degree 3. If say a is adjacent to a vertex outside $\{v, b, c, d\}$, then there is $D_{2,2}$. But as this holds for all of a, b, c, d it means $A = \{v, a, b, c, d\}$ has no neighbor in $V \setminus A$ and G is not connected.

For $n = 8$, we still have $q \leq \lfloor \frac{3n}{7} \rfloor = 3$ and $p \geq 5$. But $p = 5, 7$ are impossible, again due to parity, hence $q = 2$ and $p = 6$. Let $Q = \{a, b\}$ be the set of vertices of degree 4. If some vertex x in P is adjacent to both a and b, then consider the only neighbor z of x in P. Here a is adjacent to x and three more vertices in P, so at least two vertices except x and z are neighbors of a and x can use z and b to obtain a copy of $D_{2,2}$ with centers x and a. Hence every vertex in P is adjacent to at most one vertex in Q, yielding $|P| \geq e(P, Q) = 4|Q|$ — a contradiction.

For $n = 9$, we have $q \le \lfloor \frac{3n}{7} \rfloor = 3$. The case $q = 2$ is impossible by parity and $q = 1$, $p = 8$ implies $e(G) = (4 + 24)/2 = 14 = 2n - 4$ as stated by the theorem. So only $q = 3$, $p = 6$ is to be checked. Let $Q = \{a, b, c\}$ be the set of vertices of degree 4. If some vertex v in P has at least two neighbors in Q, say a, b, then we have a copy of $D_{2,2}$ with centers v and a, as all the four neighbors of a are in P and at most two of them belong to $N[x]$. So every vertex in P can have at most one neighbor in Q and as in the previous case we have $|P| \ge e(P, Q) = 4|Q|$ — a contradiction.

For $n = 10$, $q \le \lfloor \frac{3n}{7} \rfloor = 4$, and so parity of the degree sum implies $q = 4$ or $q = 2$. If $q = 2$ then $e(G) = (8 + 24)/2 = 16 = 2n - 4$ as stated in the theorem, so only $q = 4$, $p = 6$ remains to be checked.

Let $Q = \{a, b, c, d\}$ be the set of vertices of degree 4. If some vertex v in P has all its neighbors in Q, say a, b, c, then we obtain a copy of $D_{2,2}$ with centers v and a. Otherwise, we have $4|Q| = e(P, Q) \le 2|P|$ — a contradiction. □

Theorem 3.3. $\mathrm{ex}_c(n, S_{3,1,1}) = \lfloor \frac{3(n-1)}{2} \rfloor$ *if* $n \ge 7$ *and* $ex_c(6, S_{3,1,1}) = 9$.

Proof. The lower bounds are shown by $K_1 + M_{n-1}$ for $n \ge 7$ and by $K_{3,3}$ for $n = 6$. The former is $S_{3,1,1}$-free as shown in Proposition 1.5(3) with $m(S_{3,1,1}) = 3$. The graph $K_{3,3}$ is $S_{3,1,1}$-free as the bipartition of $S_{3,1,1}$ has a part of size 4.

To obtain the upper bound, we consider an $S_{3,1,1}$-free connected graph G. The general idea is to choose a longest cycle $C = v_1 v_2, \dots, v_k$ in G, and argue depending on its length k.

If $k = n$, then C is a Hamiltonian cycle. It cannot have short chords; e.g. if $v_2 v_4$ is an edge, then $S_{3,1,1}$ can have center v_2 and legs $v_2 v_1$, $v_2 v_3$, $v_2 v_4 v_5 v_6$. Moreover if $n > 6$, then longer chords cannot occur either. Indeed, if $v_2 v_j$ with $j = 5, \dots, n-2$ is an edge, then v_2 with v_j and its two successors can form the leg of length 3. Likewise for $j = 6, \dots, n-1$ such a leg can be formed using the two predecessors of v_j, still keeping the legs $v_2 v_1$ and $v_2 v_3$. This excludes all chords if $n > 6$, hence $|E(G)| = n$. If $n = 6$, then antipodal vertices can be adjacent without creating any copy of $S_{3,1,1}$, but no other chords may occur. In this way we obtain the extremal graph $K_{3,3}$.

Assume next that $4 < k < n$. We show that this is impossible whenever $n \ge 6$. Since G is connected, there is a vertex x not in C but having at least one neighbor in C. If e.g. xv_2 is an edge, we find $S_{3,1,1}$ with center v_2 and legs xv_2, $v_2 v_1$, $v_2 v_3 v_4 v_5$.

Assume now $k = 4$, $C = v_1 v_2 v_3 v_4$, $n \ge 6$. If P is any path with one end in C and all its other vertices in $V(G) \setminus V(C)$, then P can have no more than two edges, otherwise $S_{3,1,1}$ would be found, with the long leg in P and the two short legs in C. We are going to prove that if P is shorter than 3, the number of edges in G is smaller than what is given in the theorem.

If P has length 2, let xyv_1 be a path attached to C. Then the edges xv_2, xv_3, xv_4, yv_2, yv_4 cannot be present because C is a longest cycle. Also the edges $v_1 v_3$ and $v_2 v_4$ are excluded because G is $S_{3,1,1}$-free. This implies $|E(G)| \le 8$ if $n = 6$. If $n > 6$, there should be a further vertex z adjacent to $C \cup P$, but any edge from z to $C \cup P$ would create an $S_{3,1,1}$. (For zx the center is v_1, and for any other edge the center is the neighbor of z.) Hence $n > 6$ is impossible in this case.

Suppose that $P = yv_1$ is a single edge not extendable to a longer path outside C. Then a sixth vertex x can only be adjacent to v_2 or v_4 (or both), otherwise an $S_{3,1,1}$ would occur. And also here, it is not possible to extend this graph to a connected graph of order 7 without creating an $S_{3,1,1}$ subgraph. Hence $n = 6$. Moreover, the diagonals of C must be missing;

e.g. the edges xv_2 and v_2v_4 would yield $S_{3,1,1}$ with center v_2 and legs xv_2, v_2v_3, $v_2v_4v_1y$. Thus the number of edges is only 4 plus the degree sum of x and y, which is at most 7 because the presence of all four edges xv_2, xv_4, yv_1, yv_3 would make G Hamiltonian, hence C would not be a longest cycle.

Finally we have to consider graphs without any cycles longer than 3. It means that each block of G is K_2 or K_3. Let $f(n)$ denote the maximum number of edges in such a graph. We clearly have $f(1) = 0$, $f(2) = 1$, $f(3) = 3$. Let B be an endblock of G, with cut vertex w. Deleting $B - w$ from G we obtain a $S_{3,1,1}$-free connected graph of order $n - |V(B)| + 1$, where $|V(B)|$ is 2 or 3. Hence

$$f(n) \leq \max\{f(n-1) + 1, f(n-2) + 3\}.$$

This recursion implies $f(n) \leq \lfloor 3(n-1)/2 \rfloor$ for every n, completing the proof of the upper bound for $n \geq 7$. □

Theorem 3.4. $\mathrm{ex}_c(n, S_{2,2,1}) = 2n - 3$ *if* $n \geq 6$.

Proof. The lower bound is shown by $K_2 + E_{n-2}$ as it has matching number 2, while $\nu(S_{2,2,1}) = 3$.

To obtain the upper bound on $\mathrm{ex}_c(n, S_{2,2,1})$, we proceed by induction: for $n = 6$ every connected graph on 6 vertices and 10 edges contains $S_{2,2,1}$ (by inspecting the table of graphs of [9] on pages 222–224).

For the induction step assume that the statement of the theorem holds for graphs of at most $n - 1$ vertices and assume on the contrary that G is a connected graph on n vertices and $2n - 2$ edges without $S_{2,2,1}$. Here $2n - 2$ suffices as otherwise if $e(G) \geq 2n - 1$, we can delete an edge on a cycle.

If $\delta(G) \leq 2$ and there is a vertex v of degree at most 2 which is not a cut vertex, then we can apply induction to $H = F - v$ to obtain $e(G) \leq e(H) + 2 \leq 2(n-1) - 3 + 2 = 2n - 3$ — a contradiction.

Suppose $\delta(G) = 2$ and every vertex of degree 2 is a cut vertex. Then let v be such a cut vertex with neighbors x and y. Consider $H = G - v + (xy)$. Here $|H| = n - 1$ and $e(H) = 2n - 2 - 2 + 1 = 2(n-1) - 2 + 1$, hence by induction H contains a copy S of $S_{2,2,1}$. If S does not use the edge xy, then S is also in G — a contradiction. If S uses xy such that one of x and y, say x, is a leaf in S, then replace x by v and the edge xy by vy to obtain a copy S' of $S_{2,2,1}$ in G — a contradiction. Finally, if xy is the edge of a 2-leg of S containing the center, say x and the leg is xyz, then replace this leg by xvy to obtain S' in G — a contradiction.

So we can assume $\delta(G) \geq 3$. If all vertices are of degree 3, then $e(G) = 3n/2 < 2n-2$. If all vertices are of degree at least 4, then $e(G) \geq 2n > 2n - 2$, hence there exists a vertex y of degree 3 adjacent to a vertex x of degree at least 4. Let u, v be the other two neighbors of y, and let $z \neq u, v, y$ be a neighbor of x. If u or v has a neighbor outside these 5 vertices, then we obtain a copy of $S_{2,2,1}$ with center y. If not and $N(x) = \{u, v, y, z\}$, then z must have a neighbor outside these 5 vertices and we obtain a copy of $S_{2,2,1}$ with center x. Finally, if $N(u) \cup N(v) \subseteq \{u, v, x, y, z\}$ and z' is another neighbor of x, then $d_G(z') \geq 3$ implies that z' must have a neighbor outside these 6 vertices, and we obtain a copy of $S_{2,2,1}$ with center x. This contradiction finishes the proof. □

Proof of Theorem 1.7. The assertion is trivial for $n < 7$. For larger n the split graph construction $K_2 + E_{n-2}$ shows that $2n - 3$ is a lower bound.

To derive the same as an upper bound, assume $n > 6$ and consider any $D^*_{2,2}$-free graph G of order n with more than $2n-4$ edges. Then, by Theorem 1.6(2), there is a $D = D_{2,2}$ subgraph in G; let the central edge of D be xy.

If some vertex not in D is adjacent to a leaf of D, then a copy of $D^*_{2,2}$ arises — a contradiction. More generally, there cannot exist any vertex at distance exactly 2 from $\{x, y\}$. By the connectivity of G, it follows that every vertex of G is adjacent to at least one of x and y. On this basis we partition $V(G) - \{x, y\}$, defining

$$X = N(x) - N[y], \qquad Y = N(y) - N[x], \qquad Z = N(x) \cap N(y).$$

Let us assume $|Y| \geq |X|$. Due to the presence of $D_{2,2}$ we know that $|X| + |Z| \geq 2$ holds. Moreover, $|Y| \geq |X|$ with $n \geq 7$ implies $|Y| + |Z| \geq 3$. Hence there cannot be any $X - Y$ edges, moreover $Y \cup Z$ is an independent set, both because G is $D^*_{2,2}$-free. For the same reason, if $|X| + |Z| > 2$, then also $X \cup Z$ is independent. In this case the entire $X \cup Y \cup Z$ is independent and G cannot have more than $2n-3$ edges, yielding just the extremal split graph $K_2 + E_{n-2}$. Otherwise, if $|X| + |Z| = 2$, there can be just one edge inside $X \cup Z$, hence we have 6 edges in the K_4 subgraph induced by $X \cup Z \cup \{x, y\}$, and there are further $n-4$ edges from Y to y. These are altogether $n+2$ edges only, i.e. fewer than the assumed $2n-3$. This contradiction completes the proof. □

Proof of Theorem 1.8. To simplify notation, let $f(n) = \mathrm{ex}_c(n, S_{2,2,2})$. The lower bound for $n \geq 7$ is obtained by the following construction that works for all n. Take a complete graph K_4 on the vertex set $\{v_1, v_2, v_3, v_4\}$ and join all v_i for $i = 5, 6, \ldots, n$ to v_1 and v_2. Equivalently, v_1 and v_2 are universal vertices, supplemented with the single edge v_3v_4. This connected graph with $2n-2$ edges does not contain $S_{2,2,2}$ because it is not possible to delete two vertices from $S_{2,2,2}$ to destroy all but one edges.

The argument for the upper bound applies induction on n, with base cases $n \leq 7$, from which only $n = 7$ is nontrivial. We note here that $n = 5$ and $n = 6$ are the only cases where $2n-2$ is not an upper bound on the formula given for $f(n)$.

For $n = 7$ the assertion is that every connected graph G with 7 vertices and at least 13 edges contains $S_{2,2,2}$ as a subgraph. To prove it, suppose first that G has a cut vertex x, and consider the vertex distribution between the components of $G - x$. If it is $(3, 3)$ — where we unite components if there are more than two, e.g. the distribution $(3, 2, 1)$ is also viewed as $(3, 3)$ — then already 9 nonadjacencies are found, hence G would have at most $21 - 9 = 12$ edges — a contradiction. If the distribution is $(2, 4)$, then it forces 8 nonadjacencies, hence G must be the graph in which the two blocks incident with x are K_3 and K_5. Obviously this graph contains $S_{2,2,2}$. If the distribution is $(1, 5)$, then x has a pendant neighbor, say y, and $G - y$ is a connected graph of order 6, having at least 12 edges. Routine inspection shows that all such graphs G contain $S_{2,2,2}$.

Assume that G is 2-connected. If G has minimum degree 3, then G has a Hamiltonian cycle, say $C = v_1v_2v_3v_4v_5v_6v_7$. (More generally it is well known that a graph of order $2d + 1$ and minimum degree d is non-Hamiltonian if and only if either it is the complete bipartite graph $K_{d,d+1}$ or it has two blocks incident with a cut vertex, both blocks being K_{d+1}; in our case both of them would have only 12 edges.) The presence of any long chord in C, e.g. v_3v_6 immediately creates an $S_{2,2,2}$ with center v_3 and legs $v_3v_2v_1$, $v_3v_4v_5$, $v_3v_6v_7$. Moreover, any three consecutive short chords, e.g. v_2v_4, v_3v_5, v_4v_6 create an $S_{2,2,2}$ with center v_4 and legs $v_4v_2v_1$, $v_4v_3v_5$, $v_4v_6v_7$. And now at least one of these situations holds because in general a cycle of length n without three consecutive short chords and with no other chords at all can have no more than $n + 2n/3 < 2n - 2$ edges if $n \geq 7$.

Hence in the 2-connected case G has minimum degree exactly 2, and if we remove a vertex x of degree 2, we obtain a graph on 6 vertices with at least 11 edges. If it is K_5 with a pendant edge, then the pendant vertex must be adjacent to x and we immediately find $S_{2,2,2}$. Otherwise there can be at most one vertex of degree 2 in $G - x$, hence it contains a C_6, say $v_1v_2v_3v_4v_5v_6$ (as a rather particular corollary of Pósa's theorem). If the two neighbors of x are antipodal in C, e.g. v_3 and v_6, we find $S_{2,2,2}$ with center v_3 and legs v_3xv_6, $v_3v_2v_1$, $v_3v_4v_5$. If the two neighbors of x are consecutive in C, then C extends to C_7 which we already settled. Hence we can assume that the neighbors of x are v_2 and v_4. Since C has at least 5 chords, some of the five chords v_1v_3, v_1v_4, v_2v_5, v_3v_5, v_3v_6 must be present, and each of them creates $S_{2,2,2}$ with x and the edges of C. This completes the proof of $f(7) = 12$.

Turning now to the inductive step, assume that $n \geq 8$ and that the upper bound $2n - 2$ is valid for all smaller orders other than 5 and 6. Depending on the structure of the graph under consideration, we will apply one of the following upper bounds:

$$f(n-1)+2, \qquad f(n-3)+6, \qquad f(n-6)+12.$$

Suppose that G is an $S_{2,2,2}$-free connected graph of order $n \geq 8$, and G is $S_{2,2,2}$-saturated, i.e. the insertion of any new edge inside $V(G)$ would create an $S_{2,2,2}$ subgraph. Under the latter assumption we observe the following.

Claim 3.5. *If x is a vertex of degree 2, say with neighbors y and z, then yz is also an edge of G.*

Proof of Claim 3.5. Otherwise yxz would be an induced path in G. Let then G' be the graph obtained by the insertion of edge yz. By assumption there is an $S = S_{2,2,2}$ subgraph in G', which necessarily contains the edge yz. If yz is a leaf edge of S, then of course the degree-3 center of S cannot be x, it must be another vertex w adjacent to y or to z. But then z or y is a leaf vertex of S, and replacing yz with yx or zx we find another copy of $S_{2,2,2}$ which is a subgraph of G — a contradiction. The other possibility would be that y or z is the degee-3 vertex of S, and the edge yz is continued with a leaf edge zw or yw (allowing also $w = x$). But then x cannot be a mid-vertex of any leg of S since x does not have a neighbor other than y and z. Hence the leg yzw or zyw can be replaced with yxz or zxy, and we would again find a copy of $S_{2,2,2}$ as a subgraph of G. □

As a consequence of Claim 3.5, if G has a vertex of degree 1 or 2, then $|E(G)| \leq f(n-1)+2 \leq 2n-2$ follows by induction, because deleting a vertex of minimum degree the graph remains connected. Hence from now on we may assume that G has minimum degree at least 3.

Let $C = v_1v_2v_3v_4 \dots v_s$ be a longest cycle in G. We have already seen that if $s = n$, then $|E(G)| \leq 5n/3 < 2n - 2$. Next, we observe that if $n > s \geq 5$, then $V(G) \setminus V(C)$ is an independent set. Indeed, if xy is an edge outside C then there is a path P (possibly an edge) from $\{x, y\}$ to C and in this case a copy of $S_{2,2,2}$ is easily found using edges of C, with two edges from $P \cup \{xy\}$. E.g., if v_3x is an edge, then $S_{2,2,2}$ can have center v_3 and legs v_3xy, $v_3v_2v_1$, $v_3v_4v_5$. Thus, every vertex outside of C has at least three neighbors in C. Moreover, no two of those neighbors are consecutive in C, because C is longest. This immediately excludes $s = 5$. But also $s > 5$ is impossible because if e.g. v_2, v_4, v_6 are neighbors of x, then an $S_{2,2,2}$ can have center x and legs xv_2v_1, xv_4v_3, xv_6v_5.

As a consequence, investigations are reduced to $S_{2,2,2}$-free connected graphs with minimum degree at least 3 and without any cycles longer than 4. Such a graph G cannot be 2-connected (because due to Dirac's theorem, 2-connectivity would imply the presence of a cycle longer than 5). Hence G contains at least two endblocks.

Let B be an endblock of G, attached with cut vertex w to the other part of G. We argue that B induces K_4 in G. All vertices of B except w have degree at least 3 inside B, therefore B contains a 4-cycle, say $C' = wxyz$. If there is a vertex u in $V(B) \setminus V(C')$, then 2-connectivity of B and the exclusion of cycles longer than 4 imply that there are exactly two neighbors of u in C', either w and y, or x and z. But then there must exist a third neighbor v of u not in C', and v also has two neighbors in C'; and then a cycle longer than 4 would occur. Thus B is a K_4 indeed.

Now we are in a position to complete the proof of the theorem by induction on n. Consider any maximal $S_{2,2,2}$-free connected graph G of order $n > 7$ that has at least $2n-2$ edges. If G has a vertex of degree at most 2, then apply the upper bound $f(n-1)+2$.

If G has minimum degree at least 3, we know that G is not 2-connected. Then we distinguish cases according to n. If $n = 8$ or $n = 9$, remove all the 6 non-cutting vertices of two K_4 endblocks of G and apply the upper bound $f(n-6)+12$. This yields $|E(G)| \leq 13$ for $n = 8$ and $|E(G)| \leq 15$ for $n = 9$, both are smaller than $2n-2$.

If $n \geq 10$, remove the 3 non-cutting vertices of a K_4 endblock of G and apply the upper bound $f(n-3)+6$. This yields $|E(G)| \leq 2n-2$. □

Remark 3.6. The extremal graphs are not unique if $n \geq 7$. In the graph constructed at the beginning of the proof we can remove three vertices of degree 2 and attach a block K_4 to one of the two high-degree vertices. As another alternative for $n \geq 10$, we can remove six vertices of degree 2 and attach two blocks isomorphic to K_4, one block to each high-degree vertex. A further extremal graph of order 7 can be obtained from K_5 by attaching two pendant edges to a vertex of K_5.

Proof of Theorem 1.9. A lower bound for $n \geq 7$ is the split graph $K_2 + E_{n-2}$ with $2n-3$ edges which does not even contain $S_{2,2,1}$ and hence $S_{3,2,1}$ cannot be a subgraph either.

The proof of the upper bound proceeds by induction on n. The base case $n = 7$ is left to the Reader. Assume G is a minimum connected counterexample with $n \geq 8$ vertices and has at least $2n-2$ edges but no copy of $S_{3,2,1}$. If G contains a vertex v of degree at most 2 such that $H = G - v$ is connected, then, by minimality, $e(H) \leq 2(n-1)-3$ hence $2n-2 \leq e(G) \leq e(H)+2 \leq 2n-3$ — a contradiction.

Next, assume v is a cut vertex with neighbors x and y. Consider the graph H that we obtain from G by deleting v and adding the edge xy. We will show that if H contains $S_{3,2,1}$ then so does G. Let A be the component containing x and B the component containing y. By symmetry we may assume that if H contains a copy S of $S_{3,2,1}$, then its center is in A and so B can contain vertices of at most one leg of S. We consider cases according to the number of vertices in $S \cap B$. If A contains $S_{3,2,1}$ completely, then so does G.

If A contains all of $S_{3,2,1}$ except for a leaf played by y, then the same copy with v replacing y is contained in G. If $S \cap B = \{y, w\}$, then the leg of S ending $x - y - w$ can be replaced in G with $x - v - y$ to obtain a copy S' of $S_{3,2,1}$. If $S \cap B = \{y, w, z\}$, then the leg of S ending $x - y - w - z$ can be replaced in G with $x - v - y - w$ to obtain a copy S' of $S_{3,2,1}$. So, as proved, H must be $S_{3,2,1}$-free, hence $2n-2 \leq e(G) \leq e(H)+1 \leq 2(n-1)-3+1 \leq 2n-4$ — a contradiction.

Therefore, from now on we may assume $\delta(G) \geq 3$. By Theorem 1.6(4), we know that G contains a copy S of $S_{2,2,1}$. Let v be the center of S with legs $v-u$, $v-x-y$, and $v-a-b$. If y or b has a neighbor not in S, then G contains a copy of $S_{3,2,1}$ — a contradiction.

Suppose x (or a) has a neighbor z not in S. Then z cannot be adjacent to any of v, y, a, b as a copy of $S_{3,2,1}$ would appear. Also, z cannot be adjacent to any vertex outside S as again a copy of $S_{3,2,1}$ would appear in G. By $\delta(G) \geq 3$, z must be adjacent to u, x, and a, but then a copy of $S_{3,2,1}$ (this time with center z) would appear in G.

We have shown so far that x, y, a, b cannot have neighbors outside S.

If u has at least two neighbors z and w outside S, then they cannot be adjacent (it would create the leg $v-u-z-w$ of a copy of $S_{3,2,1}$) and none of them can have a neighbor outside S as a copy of $S_{3,2,1}$ would appear in G. As shown above, they cannot be adjacent to any of x, y, a, b hence they have degree at most 2 (with neighbors u and possibly v) contradicting $\delta(G) \geq 3$.

If u has just one neighbor, say z outside S, then z cannot have a neighbor outside S as a copy of $S_{3,2,1}$ would appear, and as before, z cannot be adjacent to any of x, y, a, b hence z can be adjacent to at most u and v but then $d_G(z) \leq 2$ contradicts $\delta(G) \geq 3$.

So the only vertex of S that can have further neighbors outside S is v. We claim that there cannot exist a path $v-w-z$ with $w, z \notin S$. Indeed, if w, z existed, then any of the edges ax, ay would create a copy of $S_{3,2,1}$ with center a. Similarly, any of the edges xa, xb would create a copy of $S_{3,2,1}$ with center b. But then $\delta(G) \geq 3$ implies the presence of ua and ux in G creating a copy of $S_{3,2,1}$ with center u. Therefore all vertices outside S must have degree 1, which case has already been dealt with. This finishes the proof of the induction step. □

Proof of Theorem 1.10. It is enough to prove that if G is a connected n-vertex graph with $\Delta(G) \geq \Delta(T)$, then G contains T or $e(G) \leq \lfloor \frac{(\Delta(T)-1)n}{2} \rfloor$. So fix a vertex v with $d_G(v) = \Delta(G) \geq \Delta(T)$ and consider the partition $\{v\}, N(v), X := V(G) \setminus N[v]$.

If X contains an edge xy, then by connectivity of G, there must exist a path (maybe a single edge) from xy to $N(v)$ and we find a copy of T in G. So we may assume that X is independent, and thus by connectivity of G, every $x \in X$ is adjacent to at least one $u \in N(v)$.

Case I: $d_G(v) = \Delta(G) > \Delta(T)$.

Then any $x \in X$ is adjacent to exactly one vertex $u \in N(v)$ as if xu, xu' are edges in G, then uxu' can form the long leg of a copy of T with center v and other neighbors of v complete this copy of T. So $d_G(x) = 1$ for all $x \in X$. Let $u, u' \in N(v)$ be two vertices such that at least one of them has a neighbor in X. Then again if uu' is an edge, we find a copy of T. So if $U \subseteq N(v)$ is the set of neighbors of v that are adjacent to a vertex in X and $U' = N(v) \setminus U$, then $e(G) \leq |U \cup X| + e(U')$. If $|U'| \leq \Delta(T) + 1$, then $e(U') \leq \binom{\Delta(T)+1}{2}$ and so $e(G) \leq n - 1 + \binom{\Delta(T)+1}{2} \leq \lfloor \frac{(\Delta(T)-1)n}{2} \rfloor$ as $\Delta(T) - 1 \geq 3$. Finally, if $|U'| \geq \Delta(T) + 2$, then either $G[U']$ is a (partial) matching and thus $e(G) \leq (1 + |U| + |X| - 1) + \frac{3|U'|}{2} \leq \frac{3(n-1)}{2} \leq \lfloor \frac{(\Delta(T)-1)n}{2} \rfloor$ (here we use $\Delta(T) \geq 4$) or $G[U']$ contains a path on 3 vertices, and then by $|U'| \geq \Delta(T) + 2$ we find a copy of T in G.

Case II: $d_G(v) = \Delta(G) = \Delta(T)$.

As X is independent, we have $e(G) \leq (\Delta(G) + 1)\Delta(G) = (\Delta(T) + 1)\Delta(T) = O(1)$. □

Proof of Theorem 1.11. The lower bound $\lfloor\frac{(k-a)n}{2}\rfloor$ follows from Remark 1.4(1), while, as the diameter of $B(k,a)$ is a, Remark 1.4(2) yields the lower bound $\binom{\lceil\frac{a+1}{2}\rceil}{2}+\lfloor\frac{a-1}{2}\rfloor(n-\lfloor\frac{a+1}{2}\rfloor)$.

To see the upper bound of (2), let $G(V,E)$ be an n-vertex $B(k,a)$-free graph with $a\le k/3$. Assume first that there exists a vertex x with $d_G(x)\ge k-1$. We claim that $G[V\setminus\{x\}]$ does not contain a path on $2a-3$ vertices. Indeed, suppose to the contrary that $y_1,y_2,\dots,y_{2a-3}$ is a path in $G[V\setminus\{x\}]$. Then as G is connected, there exists a path P from x to some y_j that does not contain any other y_i. Then either $x,P,y_j,y_{j-1},\dots,y_1$ or $x,P,y_j,y_{j+1},\dots,y_{2a-3}$ contains at least a vertices. So x and the first $a-1$ of them together with the other neighbors of x form a copy of $B(k,a)$ — a contradiction. Theorem 1.2 implies that if n is large enough, then $e(G)\le n-1+e(G-x)\le n-1+\lfloor\frac{2a-5}{2}\rfloor n\le an\le\lfloor\frac{k-a}{2}n\rfloor$ as $a\le k/3$. This finishes the proof in this case.

Assume finally that $\Delta(G)\le k-2$. Then if n is large enough, every vertex x of G is the endpoint of a path on $a\cdot k$ vertices, since G is connected and have maximum degree at most $k-2$. Suppose towards a contradiction that G contains a vertex x with $d_G(x)=d\ge k-a+1$. Let $z_1,z_2,\dots z_d$ be the neighbors of x and let $x,y_2,y_3,\dots,y_{a\cdot k}$ be a path P. Then y_2 is one of the z_j's, and as $d\le k-2$, there must exist z_j such that $z_j\in P$, say $z_j=y_i$ and either $y_{i-1},y_{i-2},\dots,y_{i-a+2}$ or $y_{i+1},y_{i+2},\dots,y_{i+a-2}$ are not neighbors of x. Then x, these y_is and the neighbors of x form a $B(k,a)$.

We obtained that $\Delta(G)\le k-a$ must hold, which implies $e(G)\le\lfloor\frac{(k-a)n}{2}\rfloor$ as claimed. □

Theorem 1.11 will provide the upper bound of Theorem 1.12. The next statement gives a general lower bound on $\mathrm{ex}_c(n,T)$ and thus will help us obtain the lower bound of Theorem 1.12.

Theorem 3.7. *For any $\varepsilon>0$ there exists a positive integer $k_0=k_0(\varepsilon)$ such that for any tree T on $k\ge k_0$ vertices, we have $\mathrm{ex}_c(n,T)\ge(\frac{k}{6}-\varepsilon)n$ if $k\ge k_0$ and n is large enough.*

Proof. Case I: $m(T)>\lfloor k/3\rfloor$.

Then by Proposition 1.5(3) we have

$$\mathrm{ex}_c(n,T)\ge n-1+\left\lfloor\frac{n-1}{m(T)-1}\right\rfloor\binom{m(T)-1}{2}\ge(n-1)\left(1+\frac{\lfloor k/3\rfloor-1}{2}\right)$$
$$\ge nk\left(\frac{1}{6}-\varepsilon\right),$$

if k and n are large enough.

Case II: $m(T)\le\lfloor k/3\rfloor$.

Then $m_2(T)\le 2m(T)\le 2\lfloor\frac{k}{3}\rfloor$, and thus $k-m_2(T)\ge\lceil\frac{k}{3}\rceil$. Proposition 1.5(4) yields

$$\mathrm{ex}_c(n,T)\ge\left\lfloor\frac{n}{\lceil\frac{k}{3}\rceil}\right\rfloor\left(1+\binom{\lceil\frac{k}{3}\rceil}{2}\right)\ge nk\left(\frac{1}{6}-\varepsilon\right),$$

if k and n are large enough. □

Proof of Theorem 1.12. The lower bound follows from Theorem 3.7, the upper bound from Theorem 1.11(2) with taking $a=\lfloor k/3\rfloor$. □

4 Concluding remarks

Theorem 1.12 gave upper and lower bounds on γ. If the lower bound from Theorem 1.11(1) turned out to be (asymptotically) sharp (which we believe to be the case) for $a = (1/2-\varepsilon)k$ or $a = (1/2+\varepsilon)k$, then the upper bound on γ would improve from $2/3$ to $1/2$. Note that a special case of Theorem 1.10 yields $\mathrm{ex}_c(n, S_{3,1,1,1}) = \lfloor\frac{(\Delta(S_{3,1,1,1})-1)n}{2}\rfloor$, so a small case when $a = \lfloor k/2 \rfloor$. We have no evidence to believe that the lower bound of $1/3$ on γ is best possible.

In Remark 1.4 and Proposition 1.5, we enumerated several graph parameters based on which one could define general constructions avoiding trees T for which these parameters have small value. It would be nice to add other parameters to this list, and would be wonderful to prove that it is enough to consider a finite set of parameters to determine the asymptotics of $\mathrm{ex}_c(n, T)$ for all trees T. Of particular interest is the characterization of those trees for which $\mathrm{ex}(n, T) - c(T) \leq \mathrm{ex}_c(n, T) \leq \mathrm{ex}(n, T)$ holds for some constant $c(T)$. As we have seen after Remark 1.4, balanced trees share this property assuming the Erdős-Sós conjecture.

Proposition 1.5 gave constructions that do not contain any tree T on k vertices with given $p(T)$, $\delta_2(T)$, $m(T)$, and $m_2(T)$. It would be interesting to figure out whether these constructions are best possible. For a family $\mathcal{G}$ of graphs, we write $\mathrm{ex}_c(n, \mathcal{G})$ to denote the maximum number of edges in an n-vertex connected graph that does not contain any $G \in \mathcal{G}$ as a subgraph.

Problem 4.1. (1) For any k and p let $\mathcal{T}^0_{k,p}$ denote the set of trees T on k vertices with $p(T) \leq p$. Determine $\mathrm{ex}_c(n, \mathcal{T}^0_{k,p})$.

(2) For any k and $d \geq 3$ let $\mathcal{T}^0_{k,d}$ denote the set of trees T on k vertices with $\delta_2(T) \geq d$. Determine $\mathrm{ex}_c(n, \mathcal{T}^0_{k,d})$.

(3) For any k and m let $\mathcal{T}^1_{k,m}$ denote the set of trees T on k vertices with $m(T) \geq m$. Determine $\mathrm{ex}_c(n, \mathcal{T}^1_{k,m})$.

(4) For any k and m let $\mathcal{T}^2_{k,m}$ denote the set of trees T on k vertices with $m_2(T) \leq m$. Determine $\mathrm{ex}_c(n, \mathcal{T}^2_{k,m})$.

As for special tree classes, one such class that could give some insight is the set of spiders with all legs of at most 2 vertices. For the spider $S = S_{2,2,\ldots,2,1,1,\ldots,1}$ with t legs of two vertices and s legs consisting of a single vertex, we have $|S| = 2t + s + 1$, and

- $\nu(T) = t + 1$ if $s > 0$,
- $\Delta(T) = t + s$,
- $m_2(T) = 4$ if $t \geq 2$.

The construction of Remark 1.4(1) based on maximum degree outperforms the one based on the matching number in Remark 1.4(3) if $s > t$. But the one based on m_2 in Proposition 1.5(4) is better than both previous ones once $s \geq 5$ and $t \geq 2$. It would be interesting to see whether these constructions achieve the asymptotics of $\mathrm{ex}_c(n, S)$.

Classical Turán numbers are monotone with two respects: Firstly, if H is a subgraph of F then $\mathrm{ex}(n, H) \leq \mathrm{ex}(n, F)$. This inequality is preserved for the connected Turán number

$\mathrm{ex}_c(n, F)$ (excluding the small "undefined" cases K_2 and P_3). Secondly, if $m < n$, then $\mathrm{ex}(m, F) \leq \mathrm{ex}(n, F)$. This property is not necessarily preserved by connected Turán numbers for small values of n with respect to $|T|$. There are several examples given by our results, of the following type: $\mathrm{ex}_c(|T| - 1, T) = \binom{|T|-1}{2} > \mathrm{ex}_c(|T|, T)$; see e.g. $T = S_{3,2,1}$.

Problem 4.2. Is it true that there exists a threshold $n_0(F)$ such that $\mathrm{ex}_c(m, F) \leq \mathrm{ex}_c(n, F)$ holds whenever $n_0(F) \leq m < n$?

ORCID iDs

Yair Caro https://orcid.org/0000-0002-9687-5770
Balázs Patkós https://orcid.org/0000-0002-1651-2487
Zsolt Tuza https://orcid.org/0000-0003-3235-9221

References

[1] P. N. Balister, E. Győri, J. Lehel and R. H. Schelp, Connected graphs without long paths, *Discrete Math.* **308** (2008), 4487–4494, doi:10.1016/j.disc.2007.08.047, https://doi.org/10.1016/j.disc.2007.08.047.

[2] N. Bougard and G. Joret, Turán's theorem and k-connected graphs, *J. Graph Theory* **58** (2008), 1–13, doi:10.1002/jgt.20289, https://doi.org/10.1002/jgt.20289.

[3] N. Bushaw and N. Kettle, Turán numbers of multiple paths and equibipartite forests, *Combin. Probab. Comput.* **20** (2011), 837–853, doi:10.1017/S0963548311000460, https://doi.org/10.1017/S0963548311000460.

[4] P. Erdős, Extremal problems in graph theory, in: *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963)*, Publ. House Czech. Acad. Sci., Prague, pp. 29–36, 1964, https://users.renyi.hu/~p_erdos/Erdos.html.

[5] P. Erdős and T. Gallai, On the minimal number of vertices representing the edges of a graph, *Publ. Math. Inst. Hung. Acad. Sci., Ser. A* **6** (1961), 181–203, https://users.renyi.hu/~p_erdos/Erdos.html.

[6] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57, https://users.renyi.hu/~p_erdos/Erdos.html.

[7] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091, doi:10.1090/S0002-9904-1946-08715-7, https://doi.org/10.1090/S0002-9904-1946-08715-7.

[8] Z. Füredi and M. Simonovits, The history of degenerate (bipartite) extremal graph problems, in: L. Lovász, I. Z. Ruzsa and V. T. Sós (eds.), *Erdős Centennial*, Springer, Berlin, Heidelberg, volume 25 of *Bolyai Soc. Math. Stud.*, pp. 169–264, 2013, doi:10.1007/978-3-642-39286-3_7, https://doi.org/10.1007/978-3-642-39286-3_7.

[9] F. Harary, *Graph Theory*, CRC Press, Boca Raton, FL, 2018.

[10] G. N. Kopylov, Maximal paths and cycles in a graph, *Dokl. Akad. Nauk SSSR* **234** (1977), 19–21.

[11] Y. Lan, T. Li, Y. Shi and J. Tu, The Turán number of star forests, *Appl. Math. Comput.* **348** (2019), 270–274, doi:10.1016/j.amc.2018.12.004, https://doi.org/10.1016/j.amc.2018.12.004.

[12] H. Liu, B. Lidický and C. Palmer, On the Turán number of forests, *Electron. J. Comb.* **20** (2013), Paper 62, 13 pp., doi:10.37236/3142, https://doi.org/10.37236/3142.

[13] M. Stein, Tree containment and degree conditions, in: *Discrete Mathematics and Applications*, Springer, Cham, volume 165 of *Springer Optim. Appl.*, pp. 459–486, 2020, doi:10.1007/978-3-030-55857-4_19, https://doi.org/10.1007/978-3-030-55857-4_19.

[14] L.-T. Yuan and X.-D. Zhang, The Turán number of disjoint copies of paths, *Discrete Math.* **340** (2017), 132–139, doi:10.1016/j.disc.2016.08.004, https://doi.org/10.1016/j.disc.2016.08.004.

[15] L.-P. Zhang and L. Wang, The Turán numbers of special forests, *Graphs Comb.* **38** (2022), Paper No. 84, 16 pp., doi:10.1007/s00373-022-02479-x, https://doi.org/10.1007/s00373-022-02479-x.