Combinatorics, Probability and Computing (2025), **34**, pp. 927–931
doi:10.1017/S0963548325100084

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Embedding nearly spanning trees*

Bruce Reed[1] and Maya Stein[2]

[1]Institute of Mathematics, Academica Sinica, Taipei, Taiwan and [2]Department of Mathematical Engineering and Center for Mathematical Modeling (CNRS IRL2807), University of Chile, Santiago, Chile
Corresponding author: Maya Stein; Email: mstein@dim.uchile.cl

(Received 25 June 2024; revised 23 April 2025; accepted 23 April 2025; first published online 11 August 2025)

Abstract
The Erdős-Sós Conjecture states that every graph with average degree exceeding $k-1$ contains every tree with k edges as a subgraph. We prove that there are $\delta > 0$ and $k_0 \in \mathbb{N}$ such that the conjecture holds for every tree T with $k \geq k_0$ edges and every graph G with $|V(G)| \leq (1+\delta)|V(T)|$.

Keywords Trees; Graphs; Degree; spanning; average degree
2020 MSC Codes: Primary: 05C35, 05C05, 05C07

1. Introduction

One of the best known conjectures in extremal graph theory is the Erdős-Sós Conjecture (see [3]).

Conjecture 1.1 (Erdős-Sós Conjecture). *Every graph G with average degree $d(G) > k-1$ contains every tree T with k edges as a subgraph.*

The conjecture clearly holds for stars and is known to hold for paths [4]. It also holds for large trees whose maximum degree is bounded [9]: this result relies on earlier results for dense host graphs [1, 12]. Further, Conjecture 1.1 holds for host graphs that are bipartite [15] or have no 4-cycles [13], and it holds whenever $k \geq |V(G)| - c$, where c is any given constant and k is sufficiently large depending on c [5]. For more background see [14].

We prove the high density case of Conjecture 1.1 for large k, with no restrictions at all on the host graph or on the tree. Our result reads as follows:

Theorem 1.2. *There are $k_0 \in \mathbb{N}$ and $\delta > 0$ such that for all $k \geq k_0$ every graph G with $|V(G)| \leq (1+\delta)k$ and with average degree $d(G) > k-1$ contains every tree T with k edges as a subgraph.*

In our proof of this theorem, we take δ to be 10^{-10}, and we make no effort to optimise this number. Our proof of Theorem 1.2 relies crucially on an earlier result by the authors. This result, shown in [10, 11], is the (surprisingly hard to prove) spanning tree case of a more general conjecture of Havet, Wood and the authors from [6], which states that every graph of minimum degree at least $2k/3$ and maximum degree at least k contains all trees with k edges as subgraphs.

Theorem 1.3. ([10, 11]) *There is an $m_0 \in \mathbb{N}$ such that for every $m \geq m_0$ every graph on $m+1$ vertices that has minimum degree at least $\lfloor 2m/3 \rfloor$ and a vertex of degree m contains every tree T with m edges as a subgraph.*

*Dedicated to the memory of Vera T. Sós.

© The Author(s), 2025. Published by Cambridge University Press.

https://doi.org/10.1017/S0963548325100084 Published online by Cambridge University Press

Let us quickly give some insight into the main ideas of our proof of Theorem 1.2. Since G has high average degree, but no subgraph fulfilling the conditions of Theorem 1.3, we are able to find a relatively large set $H \subseteq V(G)$ having degree at least k. If T has many leaves, we use H to embed parents of leaves while embedding the rest of the tree almost greedily, leaving it to the end to embed leaves from H.

The more difficult case is when T has few leaves. Then T has many vertices of degree 2, which we embed at the end. The problem is that we have to choose the images of the neighbours of these vertices very carefully, as we will need that at the very end of the embedding process, they are adjacent to most of the yet unused vertices. In order to achieve this, we embed most of the tree into a randomly selected set. More precisely, we embed most of the tree in an ordered way into a randomly selected permutation of most of $V(G)$. For this approach to work, it is crucial that we have used up all vertices of G having rather low degree in a different way in the beginning of the embedding process.

2. The proof of Theorem 1.2

Set $\delta := 10^{-10}$. Let m_0 be given by Theorem 1.3 and set $k_0 := \max\{m_0, \delta^{-2}\}$. Let T be a tree with $k \geq k_0$ edges, and let G be a graph with $n := |V(G)| \leq (1+\delta)k$ and $d(G) > k-1$. We can assume that G is minimal with these properties. In particular, G has no vertex v of degree less than $\frac{k}{2}$, as deleting such a vertex would lead to a smaller graph of at least the same average degree. So,

$$\delta(G) \geq \frac{k}{2}, \tag{1}$$

where $\delta(G)$, as usual, denotes the minimum degree of G. Set $a := n-k$, $S \subseteq V(G)$ to be the set of all vertices of degree at most $\frac{2k}{3}+a$ (the letter S stands for Small), and $b = |S|$.

If $d(v) \geq k+b$ for some $v \in V(G)$, then $G_v := G\big[N[v]-S\big]$ (the subgraph of G induced by v and all its neighbours outside S) has at least $k+1$ vertices, each of which has at most $n - |V(G_v)| \leq a-1$ neighbours outside G_v. So the minimum degree of G_v is at least $\frac{2k}{3}+a-(a-1) > \frac{2k}{3}$. Note that v has degree $|V(G_v)|-1$ in G_v, and hence we may apply Theorem 1.3 to find that $T \subseteq G_v \subseteq G$, and we are done. Therefore, we assume from now on that

$$\Delta(G) < k+b. \tag{2}$$

Let H be the set of all vertices of G having degree at least k (the letter H stands for High). Then

$$|H| > \frac{k}{6}, \tag{3}$$

as otherwise, (2) together with our choice of δ, ensures that

$$\sum_{v \in V(G)} d(v) \leq \frac{k}{6}(k+b-1) + (n - \frac{k}{6} - b)(k-1) + b(\frac{2k}{3}+a) \leq n(k-1),$$

a contradiction since $d(G) > k-1$.

Observe that the number of vertices of G having degree less than $k-\sqrt{ak}$ is at most $2\sqrt{ak} \leq 2\sqrt{\delta}k$ (as otherwise there are more than $ak \geq \frac{an}{2}$ non-edges and thus $d(\overline{G}) > a = n-k$, a contradiction). We let S' be the set of the $\lceil 2\sqrt{\delta}k \rceil$ vertices of lowest degrees. Then for each $v \in V(G) \setminus S'$, we have $d(v) \geq k - \sqrt{ak}$, and in particular, setting $G' := G[V(G) \setminus S']$, we have

$$\delta(G') \geq k - \sqrt{ak} - |S'| \geq k - 4\sqrt{\delta}k. \tag{4}$$

Note that for each two vertices $u, v \in V(G')$, we have that

$$|N(u) \cap N(v) \cap V(G')| \geq k - 9\sqrt{\delta}k. \tag{5}$$

https://doi.org/10.1017/S0963548325100084 Published online by Cambridge University Press

According to whether the tree has many or few leaves, we will either use S' for embedding leaves at the end of our embedding procedure, or fill S' as early as possible. For this, we distinguish two cases.

Case 1. *T contains at least $10\sqrt{\delta}k$ leaves.*

Choose a set L of exactly $\lceil 10\sqrt{\delta}k \rceil$ leaves, and let P_1 be the set of their parents (note that not all children of P_1 need to belong to L). Root T in an arbitrary vertex r and construct a set P_2 as follows. We start by setting $P_2 := P_1 \cup \{r\}$. Then, while there is a vertex $p \in V(T) \setminus P_2$ having at least two children in P_2, we add p to P_2. If there is no such p, we stop the process. We note that $|P_2| \leq 2|P_1| \leq 2|L| \leq 21\sqrt{\delta}k$.

Let P_3 be the set of all parents of vertices from P_2. We embed $T[P_2]$ greedily into H, which is possible by (3) and the fact that the vertices of H have degree at least k in G (and thus degree greater than $|P_2|$ into H). We then embed $T - P_2 - L$ into G', going through $T - L$ in a top down fashion, starting with the root r, which, as it belongs to P_2, is already embedded. For each subsequent vertex v that is not yet embedded, we choose its image arbitrarily among the available neighbours of the image p_v of the parent of v, unless $v \in P_3$, in which case we embed v in a vertex that is adjacent to both p_v and the image of the unique neighbour of v in P_2. This is possible by (5), because $|L| \geq 10\sqrt{\delta}k$, and since so far, we only used G' for the embedding.

It only remains to embed the vertices of L. Note that the parents of these vertices were embedded in H, and the vertices of H have degree at least k. Thus we can embed L greedily into G.

Case 2. *T has fewer than $10\sqrt{\delta}k$ leaves.*

In this case, T has fewer than $10\sqrt{\delta}k$ vertices of degree at least 3. So the set D_2 of vertices of degree 2 has size greater than $k - 20\sqrt{\delta}k$. We embed the vertices of $D_1 := V(T) \setminus D_2$ arbitrarily into G' (respecting adjacencies) which is possible by (4). Let φ denote this embedding and all future extensions of it.

Let $\mathcal{R}$ be the set of all components of $T[D_2]$. Note that each such component is a path and that

$$|\mathcal{R}| \leq |D_1| \leq 20\sqrt{\delta}k. \tag{6}$$

Take a minimal subset $\mathcal{R}'_1$ of $\mathcal{R}$ such that $\bigcup \mathcal{R}'_1$ contains at least $100\sqrt{\delta}k$ vertices. Choose an arbitrary path Q from $\mathcal{R}'_1$ and delete one of its edges, giving us two subpaths Q_1, Q_2 of Q, in a way that $\mathcal{R}_1 := (\mathcal{R}'_1 \setminus \{Q\}) \cup \{Q_1\}$ covers exactly $\lfloor 100\sqrt{\delta}k \rfloor$ vertices. Set $\mathcal{R}_2 := (\mathcal{R} \setminus \mathcal{R}'_1) \cup \{Q_2\}$.

For each path $R \in \mathcal{R}_1$, we proceed as follows. Say $R = x_1 x_2 \dots x_m$. Set $X := \{x_{2+3i} : 0 \leq i \leq m/3 - 1\}$ and note that $|X| \geq \lfloor 100\sqrt{\delta}k \rfloor - 4|\mathcal{R}_1| \geq |S'|$. We embed an arbitrary subset $X' \subseteq X$ of size $|S'|$ arbitrarily into S'. Then we embed the vertices from $V(R) \setminus X'$ into G', in any order. Note that at the moment of being embedded, each such vertex v has at most two already embedded neighbours, at most one of which is embedded in S'. By (1) and (4), there are at least $k - 4\sqrt{\delta}k - (\frac{k}{2} + a) = \frac{k}{2} - a - 4\sqrt{\delta}k$ vertices that are common neighbours of the images of the neighbours of v. Moreover, we have embedded at most $|D_1| + \lfloor 100\sqrt{\delta}k \rfloor \leq 120\sqrt{\delta}k$ vertices so far. Thus, there are at least

$$\frac{k}{2} - a - 4\sqrt{\delta}k - 120\sqrt{\delta}k > 0$$

vertices of G' that can serve as an appropriate image for v. Hence we can embed $\bigcup \mathcal{R}_1$ as planned.

Let $U \subseteq V(G)$ be the set of all vertices used so far for the embedding. Note that

$$|U| \leq 120\sqrt{\delta}k \leq \frac{k}{300}. \tag{7}$$

https://doi.org/10.1017/S0963548325100084 Published online by Cambridge University Press

It remains to embed the paths from $\mathcal{R}_2$ into $G-U$. For this, let us introduce some notation. Given a permutation $\pi=(v_1,v_2,\dots,v_{n'})$ of $V(G)\setminus U$, we set $V_\pi:=\{v_1,v_2,\dots,v_{\lceil\frac{49}{50}k\rceil}\}$. Let J_π be the set of all indices $i<\lceil\frac{49}{50}k\rceil$ such that v_i is not adjacent to v_{i+1}. Let H_π be the set of all vertices in $H\setminus(U\cup V_\pi)$ having less than $\frac{a}{3}$ non-neighbours in $G'\setminus(U\cup V_\pi)$.

We claim that there is a permutation π of $V(G)\setminus U$ such that

(A) $|J_\pi|\le 30\sqrt{\delta}k$, and

(B) $|H_\pi|\ge 16\sqrt{\delta}k$.

Assuming such a permutation π exists, we can finish the embedding as we will explain now. Choose any $H'_\pi\subseteq H_\pi$ of size exactly $\lceil 16\sqrt{\delta}k\rceil$ which is possible by (B). We start by successively embedding paths from $\mathcal{R}_2$ as follows until we have used all of V_π. We use the paths from $\mathcal{R}_2$ in non-decreasing order of their length. We embed each $R=x_1x_2\dots x_m\in\mathcal{R}_2$ vertex by vertex, avoiding H'_π. Say we are at vertex $x_j\in V(R)$ with $j\neq m$. If possible we embed x_j in the vertex v_i with lowest index i that has not been used yet. Otherwise we can and do embed x_j in a neighbour of $v_i\in V(G')\setminus H'_\pi$. Vertex x_m has two already embedded neighbours x,x' neither of which is embedded in S', and we embed x_m in a common neighbour of $\varphi(x)$ and $\varphi(x')$, avoiding H'_π. All of this is possible by (5), and since at any point, we have used at most $2|\mathcal{R}_2|+|J_\pi|\le 70\sqrt{\delta}k$ vertices outside V_π, where the inequality holds by (6) and (A). We stop once we have used all of V_π, and let R' be the remainder of the path we were currently embedding. Let $\mathcal{R}_3$ consist of R' and all remaining paths of $\mathcal{R}_2$. Observe that since $|V(G)\setminus(U\cup V_\pi)|\le\frac{k}{50}$ and because of the order in which we used the paths from $\mathcal{R}_2$,

$$|\mathcal{R}_3|\le\frac{|\mathcal{R}_2|}{50}+1\le\frac{\sqrt{\delta}k}{2}. \tag{8}$$

where the second inequality follows from (6).

By (7), the current total amount of vertices of G used for the embedding is at most $|U|+|V_\pi|+70\sqrt{\delta}k\le\frac{99}{100}k-1$. Therefore,

$$\Big|\bigcup\mathcal{R}_3\Big|\ge\frac{k}{100}\ge 2|H'_\pi|+3|\mathcal{R}_3|$$

and thus, there are sufficiently many vertices on the paths from $\mathcal{R}_3$ such that we can embed the paths from $\mathcal{R}_3$ as follows. For each path $x_1x_2\dots x_m\in\mathcal{R}_3$, we successively embed all vertices x_j with even index $j\neq m$ into H'_π, as long as there still are unused vertices in H'_π. For each odd index $j\notin\{1,m-1,m\}$ having the property that x_{j-1} and x_{j+1} are embedded in H'_π, we add vertex x_j to a set W, which is to be embedded at the very end. Observe that by construction, and by (8),

$$|W|\ge|H'_\pi|-|\mathcal{R}_3|\ge 15\sqrt{\delta}k.$$

We now use (5) to embed all remaining vertices from $V(\bigcup\mathcal{R}_3)\setminus W$ into G'.

Finally, we embed W. By construction, each vertex of W is an x_j from some path of $\mathcal{R}_3$, with x_{j-1},x_{j+1} embedded in vertices $u,v\in H'_\pi\subseteq H_\pi$. By definition of H_π, at most $\frac{2a}{3}$ vertices in $V(G)\setminus(U\cup V_\pi)$ are not common neighbours of u and v. So, as $|V(T)|=n-a+1$ and $U\cup V_\pi$ has been used, we are able to find a common neighbour of u and v in which to embed x_j. This finishes the embedding of T.

It only remains to prove our claim that there is a permutation of $V(G)\setminus U$ such that (A) and (B) hold. We take a random permutation $\pi=(v_1,v_2,\dots,v_{n'})$ of $V(G)\setminus U$, and show that with positive probability, it has both these properties. We note that by (4), and since $V(G)\setminus U\subseteq V(G')$, we have

$$\mathbb{E}\big[|J_\pi|\big]\le\sum_{v\in V_\pi}\frac{|V(G)\setminus N(v)|}{|V_\pi|}\le|V_\pi|\cdot\frac{4\sqrt{\delta}k+a}{|V_\pi|}\le 5\sqrt{\delta}k.$$

https://doi.org/10.1017/S0963548325100084 Published online by Cambridge University Press

Hence by Markov's inequality (see [8]), the probability that (A) fails is at most $\frac{1}{6}$.

We will show that (B) fails with probability less than $\frac{5}{6}$, which will finish the proof of our claim. By definition of H, each vertex from H has less than a non-neighbours in $V(G') \setminus U$. Moreover, as $|U| \geq |S'| > a$, for any $v \in V(G') \setminus U$, the probability that $v \notin V_\pi$ is at most $\frac{1}{50}$. So, setting

$$s_\pi := \sum_{u \in H \setminus (U \cup V_\pi)} |\{v \mid u \notin U, uv \notin E(G')\} \setminus V_\pi|,$$

we have that

$$\mathbb{E}[s_\pi] \leq \frac{|H \setminus U|a}{2500} \leq \frac{ak}{2400},$$

and by Markov's inequality (see [8]), the probability that $s_\pi \geq \frac{ak}{1600}$ is at most $\frac{2}{3}$. In particular, the probability that $H \setminus (U \cup V_\pi)$ has more than $\frac{k}{500}$ vertices which each have at least $\frac{a}{3}$ non-neighbours in $V(G') \setminus (U \cup V_\pi)$ is at most $\frac{2}{3}$. So, if we can show that the probability that $|H \setminus (U \cup V_\pi)| < \frac{k}{500} + 16\sqrt{\delta}k$ is at most $\frac{1}{6}$, we are done.

For this, note that $|H \setminus U| > \frac{4}{25}k$ by (3) and (7). Also by (7), for each $v \in V(G') \setminus U$, the probability that $v \notin V_\pi$ is at least $\frac{1}{60}$. It follows that the expectation of $|H \setminus (U \cup V_\pi)|$ is at least $\frac{k}{375}$. Applying the Chernoff bound (see [8]), we deduce that the probability that $|H \setminus (U \cup V_\pi)| < \frac{k}{400}$ is (much) less than $\frac{1}{6}$. As $\frac{k}{400} > \frac{k}{500} + 16\sqrt{\delta}k$, we are done.

Funding statement

This work is Supported by FONDECYT Regular Grant 1221905, ANID Basal Grant CMM FB210005, and by MSCA-RISE-2020-101007705 project *RandNET*.

References

[1] Besomi, G., Pavez-Signé, M. and Stein, M. (2021) On the Erdős-Sós conjecture for bounded degree trees. *Comb. Prob. Comp.* **30**(5) 741–761.
[2] Brandt, S. and Dobson, E. (1996) The Erdős–Sós conjecture for graphs of girth 5. *Discrete Math.* **150** 411–414.
[3] Erdős, P. (1964) Extremal problems in graph theory. In Theory of Graphs and Its Applications, Proceedings of the Symposium, pp. 29–36.
[4] Erdős, P. and Gallai, T. (1959) On maximal paths and circuits of graphs. *Acta Mathematica Academiae Scientiarum Hungarica* **10** 337–356.
[5] Goerlich, A. and Zak, A. (2016) On Erdős-Sós conjecture for trees of large size. *Electron. J. Comb.* **23**(1) P1–52.
[6] Havet, F., Reed, B., Stein, M. and Wood, D. R. (2020) A Variant of the Erdős-Sós conjecture. *J. Graph Theory* **94**(1) 131–158.
[7] Haxell, P. E. (2001) Tree embeddings. *J. Graph Theory* **36**(3) 121–130.
[8] Molloy, M. and Reed, B. (2002) *Graph Colouring and the Probabilistic Method*. Springer, Berlin.
[9] Pokrovskiy, A. (2024) Hyperstability in the Erdős-Sós conjecture Preprint 2024, arXiv: 2409.15191.
[10] Reed, B. and Stein, M. (2023) Spanning trees in graphs of high minimum degree with a universal vertex II: A tight result. *J. Graph Theory* **102**(4) 797–821.
[11] Reed, B. and Stein, M. (2023) Spanning trees in graphs of high minimum degree with a universal vertex I: An asymptotic result. *J. Graph Theory* **102**(4) 737–783.
[12] Rozhoň, V. (2019) A local approach to the Erdős–Sós conjecture. *SIAM J. Discrete Math.* **33**(2) 643–664.
[13] Saclé, J.-F. and Woźniak, M. (1997) A note on the Erdős–Sós conjecture for graphs without C_4. *J. Comb. Theory (Series B)* **70**(2) 229–234.
[14] Stein, M. (2021) Tree containment and degree conditions. In *Discrete Mathematics and Applications, Springer Optimization and Its Applications*, Vol. **165**. Springer, Cham, pp. 459–486.
[15] Stein, M. (2024) Kalai's conjecture for r-partite r-graphs. *Eur. J. Comb.* **117** 103827.

Cite this article: Reed B and Stein M (2025). Embedding nearly spanning trees. *Combinatorics, Probability and Computing* **34**, 927–931. https://doi.org/10.1017/S0963548325100084

https://doi.org/10.1017/S0963548325100084 Published online by Cambridge University Press