Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Beyond the Erdős-Sós conjecture

(Extended abstract)

Akbar Davoodi* Diana Piguet* Hanka Řada*†
Nicolás Sanhueza-Matamala‡

Abstract

We prove an asymptotic version of a tree-containment conjecture of Klimošová, Piguet and Rozhoň [European J. Combin. 88 (2020), 103106] for graphs with quadratically many edges. The result implies that the asymptotic version of the Erdős-Sós conjecture in the setting of dense graphs is correct.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-045

1 Introduction

One of the most classical questions in graph theory is to determine the number of edges in a host graph G that forces the existence of a copy of another guest graph H. For H with chromatic number at least 3, this is well understood thanks to the Erdős–Stone–Simonovits theorem [9, 8]. For bipartite H much less is known, and even the case of trees is widely open. A seminal conjecture by Erdős and Sós [7] says that graphs with average degree larger than $k-1$ should contain all k-edge trees.

Conjecture 1 (Erdős-Sós conjecture)**.** Every graph G with average degree $d(G) > k-1$ contains every tree with k edges.

*The Czech Academy of Sciences, Institute of Computer Science, Pod Vodárenskou věží 2, 182 07 Prague, Czech Republic. The research leading to these results has received funding from The Czech Science Foundation, grant number 19-08740S.

†Department of Applied Mathematics, Faculty of Information Technology, Czech Technical University in Prague.

‡Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Chile. Partly supported by ANID-Chile through the FONDECYT Iniciación N°11220269 grant.

The conjecture has been verified for certain families of trees (see [14]). A solution for all large enough trees was announced by Ajtai, Komlós, Simonovits and Szemerédi [1, 3, 2], but remains unpublished.

It is well-known that any graph that satisfies $d(G) > k - 1$ contains a subgraph G' with $\Delta(G') \geq k$ and $\delta(G') \geq k/2$. However, this weaker condition on the host graph is not enough to ensure the containment of all trees. It fails for trees of diameter four, as shown by examples of Havet, Reed, Stein and Wood [10, Section 1]. It is natural to ask how many vertices of degree at least k in G together with $\delta(G) \geq k/2$ would guarantee the containment of all k-edge trees. A conjecture along these lines was proposed by Klimošová, Piguet, and Rozhoň.

Conjecture 2 (Klimošová, Piguet, Rozhoň, [13, Conjecture 1.4])**.** Every n-vertex graph G with $\delta(G) \geq k/2$ and at least $n/(2\sqrt{k})$ vertices of degree at least k contains all k-edge trees.

Our main result is an approximate version of Conjecture 2 for dense graphs.

Theorem 3. *For any $\eta, q > 0$, there exists an $n_0 \in \mathbb{N}$ so that for every $n \geq n_0$ and all $k \geq qn$, any n-vertex graph G with minimal degree $\delta(G) \geq (1+\eta)k/2$ and with at least ηn vertices of degree at least $(1+\eta)k$ contains all k-vertex trees.*

Theorem 3 implies an approximate dense version of the Erdős-Sós conjecture.

Corollary 4. *For any $\eta, q > 0$ there exists an $n_0 \in \mathbb{N}$ so that for every $n \geq n_0$ and all $k \geq qn$ any n-vertex graph with average degree at least $(1+\eta)k$ contains any tree on at most k vertices as its subgraph.*

Corollary 4 strengthens similar results by Rozhoň [13] and Besomi, Pavez-Signé and Stein [4, Theorem 1.3]. They give the same result as our Corollary 4 but only for trees T on k vertices which in addition satisfy $\Delta(T) = o(k)$; in contrast, our result works for all trees. It also gives a proof independent of the one proposed by Ajtai, Komlós, Simonovits, and Szemerédi [1, 3, 2] in the case the host graph is dense with the very mild strengthening that its average degree is required to be slightly larger than k. We also remark that Besomi, Pavez-Signé and Stein [4, Theorem 1.1] proved a version of the Erdős–Sós conjecture for k-edge trees which is sharp in the average degree condition (it only needs $d(G) > k - 1$) but works only for large bounded-degree trees (it needs that G is an n-graph, $\Delta(T) \leq \Delta$, and $k \geq qn$, with n large with respect to q and Δ).

1.1 Notation

As is somewhat standard, we write $a \ll b$ in statements to mean "for all $b > 0$, there exists $a > 0$ such that the following is valid". Longer chains of constants are interpreted similarly, choosing the constants from right to left. We always assume those constants are positive, and if $1/n$ appears in such a chain of constants we assume that n is a positive integer.

For two disjoint subsets X and Y of $V(G)$, the *bipartite density* of the pair (X, Y) is given $d(X, Y) = |E(X, Y)|/(|X||Y|)$, where $|E(X, Y)|$ denotes the number of edges between X and Y. For a graph G, we denote by $d(G)$ the *average degree* of G, i.e. $d(G) := 2|E(G)|/|V|$. For a vertex $v \in V(G)$, let $N_G(v)$ denote the set of neighbours of v in G. We will omit G from the notation if the graph is clear from context.

A *digraph* is a graph in which every edge is oriented, meaning that it consists of an ordered pair of vertices. We admit cycles of length 2 (where the pairs of edges $\overrightarrow{uv}$ and $\overrightarrow{vu}$ are both present), but we do not allow for parallel edges in the same direction, and we also forbid loops.

2 Sketch of the proof and Main lemmas

Our proof has three main steps. First, we describe a way to cut the tree to be embedded into suitable chunks. Secondly, we prepare the host graph to embed the tree. For this, we use the Szemerédi's Regularity Lemma, which is somewhat standard in this type of proofs. A crucial definition, and our main innovation, in this step is what we call *skew matching pairs*, which are required to describe the structure which we wish to find in the host graph. The outcome of this step is summarised in what we call the Structural Lemma (Lemma 12), which is the main technical lemma of our work. In the third and final step, we construct an embedding given the structures in both the tree and the host graph, and this process is summarised in the Tree Embedding Lemma (Lemma 13).

The rest of this extended abstract is structured as follows. First, we state the aforementioned lemmas in more detail. Next, assuming the validity of those lemmas, we give short proofs of our main results: Section 3.1 contains the proof of Theorem 3 and Section 3.2 contains the proof of Corollary 4.

2.1 Preparing the tree

To prepare the embedding, we use the following handy concept used by Hladký, Komlós, Piguet, Simonovits, Stein, and Szemerédi [11, Definition 3.3]. It gives a partition of a tree into vertex-disjoint smaller trees which also satisfy several additional useful properties.

If T is a tree rooted at r, and $\widetilde{T} \subseteq T$ is a subtree with $r \notin V(\widetilde{T})$, the *seed of* $\widetilde{T}$ is the unique vertex $x \in V(T) \setminus V(\widetilde{T})$ which is farthest from r and also belongs to every (r, v)-path in T, for every $v \in V(\widetilde{T})$.

Definition 5 (ℓ-fine partition)**.** Let T be a tree on k vertices rooted at a vertex r. An *ℓ-fine partition of T* is a quadruple $(W_A, W_B, \mathcal{F}_A, \mathcal{F}_B)$, where $W_A, W_B \subseteq V(T)$ and $\mathcal{F}_A, \mathcal{F}_B$ are families of subtrees of T such that

(FP1) the three sets W_A, W_B, and $\{V(T^*)\}_{T^* \in \mathcal{F}_A \cup \mathcal{F}_B}$ partition $V(T)$ (in particular, the trees in $\mathcal{F}_A \cup \mathcal{F}_B$ are pairwise vertex-disjoint),

(FP2) $r \in W_A \cup W_B$,

(FP3) $\max\{|W_A|, |W_B|\} \leq 336k/\ell$, (FP4) $|V(T^*)| \le \ell$ for every $T^* \in \mathcal{F}_A \cup \mathcal{F}_B$,

(FP5) $V(T^*) \cap N(W_B) = \emptyset$ for every $T^* \in \mathcal{F}_A$, and $V(T^*) \cap N(W_A) = \emptyset$ for every $T^* \in \mathcal{F}_B$;

(FP6) each tree of $\mathcal{F}_A \cup \mathcal{F}_B$ has its seeds in $W_A \cup W_B$,

The crucial fact, proven in [11, Lemma 3.5], is that any tree T admits an ℓ-fine partition, for any $1 \le \ell \le |V(T)|$. We denote by $\mathcal{T}^{\rho}_{a_1,a_2,b_1,b_2}$ the set of trees T, so that there is a $(\rho|V(T)|)$-fine partition $(W_A, W_B, \mathcal{F}_A, \mathcal{F}_B)$ of T so that $|V_i(\mathcal{F}_A)| = a_i$, $|V_i(\mathcal{F}_B)| = b_i$, for $i \in \{1, 2\}$, where $V_1(\mathcal{F}_A)$ (resp. $V_2(\mathcal{F}_A)$) is the set of vertices of $\mathcal{F}_A$ that are at odd (resp. even) distance from W_A, and $V_i(\mathcal{F}_B)$ are defined analogously with respect to W_B.

2.2 Preparing the host graph

In this step, we find a suitable structure in the host graph to embed the tree, using the information about the fine partition found in the previous step. The description of this step requires Szemerédi's Regularity Lemma. Before stating it, we recall the standard notions involved in its statement.

Definition 6 (Regular pair and regular partitions)**.** A pair (X, Y) with $X, Y \subseteq V(G)$ is said to be *ε-regular*, if for any sets $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \ge \varepsilon|X|$ and $|Y'| \ge \varepsilon|Y|$ we have that $|d(G[X', Y']) - d(G[X, Y])| < \varepsilon$.

We say that a partition $\{V_0, \dots, V_t\}$ of $V(G)$ is an *ε-regular partition* if $|V_0| \le \varepsilon|V(G)|$, and for every $1 \le i \le t$, all but at most εt values of $1 \le j \le t$ are such that the pair (V_i, V_j) is not ε-regular.[1] We call the cluster V_0 the *garbage set*. We call a regular partition *equitable* if $|V_i| = |V_j|$ for every $1 \le i < j \le t$.

Szemerédi's Regularity Lemma ensures that regular partitions exist for every graph.

Theorem 7 (Szemerédi's Regularity Lemma, [15])**.** *Let $1/n \ll 1/M_0 \ll \varepsilon$. Any n-vertex graph has an ε-regular equitable partition $\{V_0, V_1, \dots, V_t\}$ with $1/\varepsilon \le t \le M_0$.*

We capture the structure of a regular partition in a so-called *reduced graph*.

Definition 8 (Reduced graph)**.** Given a graph G, $d > 0$, and a ε-regular equitable partition $\mathcal{P} = \{V_0, \dots, V_t\}$ of $V(G)$, we define the *d-reduced graph* Γ as follows. The vertex set of Γ is $\{1, \dots, t\}$, and there is an edge $ij \in E(\Gamma)$ if and only if the pair (V_i, V_j) is ε-regular and $d(V_i, V_j) \ge d$.

As mentioned before, the Structural Lemma (Lemma 12) will yield a useful structure in the host graph G; more precisely, this structure will be defined in a reduced graph of G. In essence, the structure we want is an "allocation" of the tree in the clusters of the reduced

[1] We point out that ε-regular partitions are most commonly defined in a slightly different way, with the property that at most εt^2 pairs of the partition are not ε-regular. But the version we use is also common, and in fact the existence of such partitions can be deduced from the well-known 'degree form' of Szemerédi's Regularity Lemma, see e.g. [12, Theorem 1.10].

graph which respects the sizes of the various parts of T given by the fine partition. If T is in $\mathcal{T}^{\rho}_{a_1,a_2,b_1,b_2}$, the structure will incorporate the values a_1, a_2, b_1, b_2 to describe where the parts of the tree will be allocated.

In order to describe this structure precisely, we need some definitions. The basic building block is what we call a *skew fractional matching*, whose definition is inspired by the standard fractional matching.

Given a graph G, its associated digraph $G^{\leftrightarrow}$ is the digraph with the same vertex set as G and $\overrightarrow{uv}$ and $\overrightarrow{vu}$ are present for each undirected $uv \in E(G)$.

Definition 9. Let G be a graph and $\gamma \geq 0$.

(i) An *γ-skew fractional matching* is a function $\sigma : E(G^{\leftrightarrow}) \to [0, 1]$ such that for any vertex $u \in V(G)$,

$$\frac{1}{1+\gamma} \sum_{v \in N_G(u)} \sigma(\overrightarrow{uv}) + \frac{\gamma}{1+\gamma} \sum_{v \in N_G(u)} \sigma(\overrightarrow{vu}) \leq 1.$$

The *weight* of σ is $W(\sigma) := \sum_{uv \in E(G)} \sigma(\overrightarrow{uv}) + \sigma(\overrightarrow{vu})$.

(ii) Set $\sigma^1(u) := \frac{1}{1+\gamma} \sum_{v \in N(u)} \sigma(\overrightarrow{uv})$ and $\sigma^2(u) := \frac{\gamma}{1+\gamma} \sum_{v \in N(u)} \sigma(\overrightarrow{uv})$. Abusing notation, we shall use the symbol σ to describe the *total charge* of σ on u, defined as $\sigma(u) := \sigma^1(u) + \sigma^2(u)$.

(iii) If σ, σ' are γ-skew and γ'-skew matchings respectively, we say σ, σ' are *disjoint* if, for every $u \in V(G)$, $\sigma(u) + \sigma'(u) \leq 1$.

Intuitively, γ-skew fractional matchings can be understood as fractional matchings in graphs where the weight of the edge is distributed in an unbalanced way, meaning that one end of the edge gets γ times the weight of the other end. Here, the direction of this imbalance is given by the direction of the edge in the digraph.

Definition 10. Let G be a graph, $\gamma \geq 0$, and $u \in V(G)$. We will say a γ-skew fractional matching σ is *anchored in* $N(u)$ if $\sigma^1(v) > 0$ implies that $v \in N(u)$.

Definition 11. Let G be a graph, $\gamma_A, \gamma_B > 0$. Given an edge $cd \in E(G)$, a (γ_A, γ_B)*-skew matching pair anchored in* $\overrightarrow{cd}$ is a pair (σ_A, σ_B) such that

(i) σ_A and σ_B are disjoint,

(ii) σ_B is a γ_B-skew fractional matching anchored in $N(d)$, and

(iii) σ_A is a γ_A-skew fractional matching anchored in $N(c)$.

Lemma 12 (Structural Lemma: Simplified version)**.** *Let $k \in \mathbb{N}$ and let H be a graph such that $\delta(H) \geq k/2$ and $\Delta(H) \geq k$. Let $a_1, a_2, b_1, b_2 > 0$ such that $a_1 + a_2 + b_1 + b_2 = k$. Let $\gamma_A := \frac{a_2}{a_1}$ and $\gamma_B := \frac{b_2}{b_1}$. Then $H^{\leftrightarrow}$ admits a (γ_A, γ_B)-skew matching pair (σ_A, σ_B) anchored in some edge $\overrightarrow{cd} \in E(H^{\leftrightarrow})$ such that $W(\sigma_A) = a_1 + a_2$ and $W(\sigma_B) = b_1 + b_2$.*

2.3 Embedding the tree

Based on the structure given by the Structural Lemma, the next lemma ensures that the embedding of the tree T is possible.

Lemma 13 (Tree Embedding Lemma: Simplified version)**.** *Let $1/n \ll \rho \ll 1/M, \varepsilon \ll d \ll \eta, q$. Suppose G is an n-vertex graph, that $\mathcal{P} = \{V_0, V_1, \dots, V_t\}$ is a ε-regular equitable partition of G with $t \leq M$, and that Γ is the d-reduced graph obtained from G and $\mathcal{P}$. Suppose $k \geq qn$ and that we have numbers $a_1, b_1 \in \mathbb{N}$ and $\gamma_A, \gamma_B \geq 0$ such that $k = (1+\gamma_A)a_1 + (1+\gamma_B)b_1$. Suppose $\Gamma^{\leftrightarrow}$ admits a (γ_A, γ_B)-skew matching pair (σ_A, σ_B), anchored in $\overrightarrow{cd} \in E(\Gamma^{\leftrightarrow})$, with weights satisfying $W(\sigma_A)n \geq (1+\eta)(1+\gamma_A)a_1 t$ and $W(\sigma_B)n \geq (1+\eta)(1+\gamma_B)b_1 t$. Then G contains any k-vertex tree $T \in \mathcal{T}^{\rho}_{a_1, \gamma_A a_1, b_1, \gamma_B b_1}$.*

3 Proof of the main results

3.1 Proof of Theorem 3

Now we give the proof of Theorem 3, assuming the validity of the main lemmas (Lemma 12, Lemma 13).

Setting up the parameters. Suppose we are given input parameters $\eta > 0$, $q > 0$ and $k \geq qn$. We may assume that $q, \eta \ll 1$, or we just replace them with smaller values. We set the following parameters to satisfy

$$1/n \ll \rho \ll 1/M \ll \varepsilon \ll d \ll \eta, q \ll 1. \tag{1}$$

From now on we fix an arbitrary k-vertex tree T, and the goal is to show that $T \subseteq G$.

Processing the tree. By [11, Lemma 3.5], T has an $(\rho|V(T)|)$-fine partition. Let a_1, a_2, b_1, b_2 such that $T \in \mathcal{T}^{\rho}_{a_1, a_2, b_1, b_2}$. By assumption, G satisfies $\delta(G) \geq (1+\eta)k/2 = (1+\eta)|V(T)|/2$ and at least ηn vertices of G have degree at least $(1+\eta)k = (1+\eta)|V(T)|$.

Preparing the host graph. We apply Theorem 7 on G with parameters ε and $1/M$, and obtain an ε-regular equitable partition $\{V_0, V_1, \dots, V_t\}$, with $t \leq M$. Given G and $\mathcal{P}$, let Γ be the d-reduced graph Γ. Standard arguments show that the reduced graph inherits degree properties of the original graph, up to a small loss. In particular, it can be shown that $\delta(\Gamma) \geq (1 + \frac{\eta}{40})kt/(2n)$ and $\Delta(\Gamma) \geq (1 + \frac{\eta}{40})kt/n$.

We apply Lemma 12 with Γ, $(1 + \frac{\eta}{40})kt/n$, $(1 + \frac{\eta}{40})a_i t/n$ and $(1 + \frac{\eta}{40})b_i t/n$ playing the roles of H, k and a_i, b_i, for $i \in [2]$. This outputs an $(\frac{a_2}{a_1}, \frac{b_2}{b_1})$-skew matching pair (σ_A, σ_B) anchored in some edge $\overrightarrow{cd} \in E(\Gamma^{\leftrightarrow})$ with $W(\sigma_A) = (1 + \frac{\eta}{40})(a_1 + a_2)t/n$ and $W(\sigma_B) = (1 + \frac{\eta}{40})(b_1 + b_2)t/n$.

Embedding the tree. Finally, we can apply Lemma 13 with $a_2/a_1, b_2/b_1, \eta/40$ playing the roles of γ_A, γ_B, η respectively. This shows that $T \subseteq G$, as required. □

3.2 Proof of the approximate version of the Erdős–Sós conjecture

Now we derive Corollary 4 from Theorem 3. Let $k = rn$ with $r \geq q > 0$, and let G be a graph on n vertices with average degree at least $(1+\eta)k$. It is well-known [6, Proposition 1.2.2] that G contains an induced subgraph H such that $\delta(H) \geq d(H)/2 \geq d(G)/2 \geq (1+\eta)k/2$. Let m be the number of vertices of H, we clearly have $(1+\eta)k/2 \leq \delta(H) < m \leq n$.

For any $\lambda > 0$, let X_λ be the set of vertices of H whose degree in H is at least $(1+\lambda)k$. Then we have

$$(1+\eta)km \leq md(H) = \sum_{v \in V(H)} \deg_H(v) \leq |X_\lambda|m + (m - |X_\lambda|)(1+\lambda)k,$$

which, by rearranging, gives $|X_\lambda| \geq \frac{(\eta-\lambda)km}{m-(1+\lambda)k} \geq (\eta-\lambda)k$. From now on, fix $\lambda := \eta k/(m+k)$. This choice satisfies $\eta \geq \lambda \geq \eta r/(1+r)$, and from the previous calculations we deduce that H satisfies $\delta(H) > (1+\lambda)k/2$, and has at least $(\eta-\lambda)k = \lambda m$ vertices of degree at least $(1+\lambda)k$. Thus the statement follows by applying Theorem 3 to H, with λ, m playing the role of η and n, respectively. □

4 Final remarks

We stated our main technical lemmas (Lemma 12 and Lemma 13) in simplified versions which are enough to give a faithful version of the main ideas of our proof. In our actual proof, the statements are a bit more complicated since we need to consider *weighted* reduced graphs, where each edge $ij \in \Gamma$ receives a weight $d_{ij} \in [0,1]$ corresponding to the bipartite density $d(V_i, V_j)$ of the pair (V_i, V_j). Further details, and full proofs of the main lemmas, will be found in the full version of the paper [5].

Acknowledgment

We would like to thank Jan Hladký for drawing our attention to the fact that Theorem 3 implies the approximate Erdős-Sós Conjecture for dense graphs (Corollary 4).

References

[1] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. On the approximative solution of the Erdős-Sós conjecture on trees. In preparation.

[2] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. The solution of the Erdős-Sós conjecture for large trees. In preparation.

[3] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. Some elementary lemmas on the Erdős-Sós conjecture on trees. In preparation.

[4] Guido Besomi, Matías Pavez-Signé, and Maya Stein. On the Erdős-Sós conjecture for trees with bounded degree. *Combin. Probab. Comput.*, 30(5):741–761, 2021.

[5] Akbar Davoodi, Diana Piguet, Hanka Řada, and Nicolás Sanhueza-Matamala. The asymptotic version of the Erdős–Sós conjecture and beyond. In preparation.

[6] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, Berlin, fifth edition edition, 2017.

[7] P. Erdős. Extremal problems in graph theory. In *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963)*, pages 29–36. Publ. House Czech. Acad. Sci., Prague, 1964.

[8] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.

[9] P. Erdös and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[10] Frédéric Havet, Bruce Reed, Maya Stein, and David R. Wood. A variant of the Erdös-Sós conjecture. *J. Graph Theory*, 94(1):131–158, 2020.

[11] Jan Hladký, János Komlós, Diana Piguet, Miklós Simonovits, Maya Stein, and Endre Szemerédi. The approximate Loebl-Komlós-Sós conjecture IV: Embedding techniques and the proof of the main result. *SIAM J. Discrete Math.*, 31(2):1072–1148, 2017.

[12] J. Komlós and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 295–352. János Bolyai Math. Soc., Budapest, 1996.

[13] Václav Rozhoň. A local approach to the Erdős-Sós conjecture. *SIAM J. Discrete Math.*, 33(2):643–664, 2019.

[14] Maya Stein. Tree containment and degree conditions. In *Discrete Mathematics and Applications*, volume 165 of *Springer Optim. Appl.*, pages 459–486. Springer, Cham, 2020.

[15] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.