A LOCAL APPROACH TO THE ERDŐS–SÓS CONJECTURE[*]

VÁCLAV ROZHOŇ[†]

Abstract. A famous conjecture of Erdős and Sós states that every graph with average degree more than $k-1$ contains all trees with k edges as subgraphs. We prove that the Erdős–Sós conjecture holds approximately, if the size of the embedded tree is linear in the size of the graph, and the maximum degree of the tree is sublinear.

1. Introduction. Typical problems in extremal graph theory ask how many edges in a graph force it to contain a given subgraph. A classical example of a result in this area is Turán's theorem, which determines the average degree that guarantees the containment of the complete graph K_r. A more complex example is the Erdős–Stone theorem [12], which essentially determines the average degree condition guaranteeing that the host graph contains a fixed nonbipartite graph. On the other hand, for a general bipartite graph the problem is wide open. If the embedded graph is a tree, the celebrated conjecture of Erdős and Sós asserts that an average degree greater than $k-1$ forces a copy of any tree of order $k+1$.

CONJECTURE 1.1 (the Erdős–Sós conjecture). *Every graph G with $\mathrm{d}\overline{\mathrm{eg}}(G) > k-1$ contains any tree on $k+1$ vertices.*

Here $\mathrm{d}\overline{\mathrm{eg}}(G)$ means the average degree of G; similarly, we denote the minimum and the maximum degree of G by $\delta(G)$ and $\Delta(G)$, respectively.

Observe that the conjecture is optimal, since a graph with average degree at most $k-1$ may have only k vertices. Also observe that if we replace the condition on the average degree by a stronger condition $\delta(G) > k-1$, the conjecture becomes trivial, since we can embed any tree on $k+1$ vertices in G in a greedy manner—every time we embed a vertex of the tree such that its neighbor is already embedded; since the neighborhood of the already embedded vertex is sufficiently large, we may always do that. Note that each graph with average degree $\mathrm{d}\overline{\mathrm{eg}}(G) \geq 2k$ contains a subgraph with $\delta(G) \geq k$. Such a subgraph can be found by repeatedly deleting vertices of degree smaller than k from G. Hence, the conjecture also holds trivially if we allow ourselves to lose a factor of 2.

After one verifies that the Erdős–Sós conjecture is true for both trees of diameter at most three and for paths (this was done already by Erdős and Gallai in 1959 [13]) one can observe that such trees can be embedded even in the case when the host graph contains a vertex of degree at least k and its minimum degree is at least $k/2$. This

---

[†]Faculty of Mathematics and Physics, Charles University, and Czech Academy of Sciences, Institute of Computer Science, Pod Vodárenskou věží 2, 182 07 Prague, Czech Republic (vaclavrozhon@gmail.com).
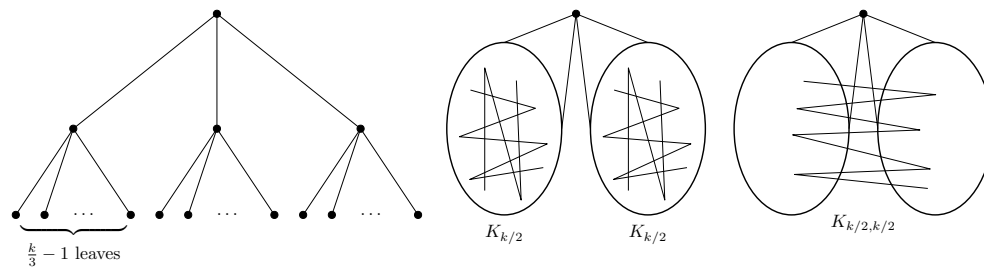
FIG. 1. *A tree on $k + 1$ vertices and two host graphs of the same size showing that there are graphs with $\Delta(G) = k$ and $\delta(G) \geq k/2$ that do not contain a tree on $k + 1$ edges. The tree consists of a vertex connected to centers of three stars with $\frac{k}{3} - 1$ leaves. The first host graph consists of a vertex complete to two complete graphs on $\frac{k}{2}$ vertices, while the second graph consists of a vertex complete to a complete bipartite graph with two color classes with $\frac{k}{2}$ vertices. The example is taken from* [18].

is trivial for trees of diameter at most three, while for the case of paths this follows from the mentioned proof of Erdős and Gallai.

While this local condition on the minimum and maximum degree of G suffices for both of these special cases, it already fails for trees of diameter four, as is demonstrated by the following example from [18]. Let T be a tree consisting of a vertex connected to centers of three stars on $k/3$ vertices and let G be a graph consisting of a vertex complete to either two cliques of size $k/2$, or $K_{k/2,k/2}$. Then $\Delta(G) \geq k$ and $\delta(G) \geq k/2$, but T is not contained in G (see Figure 1). This example shows that it would be naïve to try to prove the Erdős–Sós conjecture in the most general setting using only the local consequence of the bound on the average degree on the maximum and minimum degree of G. We will actually show in section 2.1 that trees of diameter at most three and paths are special cases. Proposition 2.1 states that with high probability, a random tree on $k + 1$ vertices cannot be embedded in the host graph with two cliques from Figure 1.

Despite this fact, we devote this paper to this *local* approach to the Erdős–Sós conjecture. We are interested in the following question: if the host graph contains a certain amount of vertices of degree at least roughly k and its minimum degree is at least roughly $k/2$, can we embed all trees of order k in the graph? We partially answer this question for dense graphs, which in turn implies an approximate version of the Erdős–Sós conjecture.

1.1. Main result of this paper. The main positive result of this paper is the following theorem.

THEOREM 1.2. *For any $\eta > 0$ there exists n_0 and $\gamma > 0$ such that for every $n > n_0$ and $k > 0$, any graph of order n with average degree $\overline{\deg}(G) \geq k + \eta n$ contains every tree on k vertices with maximum degree $\Delta(T) \leq \gamma k$.*

Another way to state this result is the following. Suppose that we have a class of trees $\mathcal{T}$ such that there exists a function f, $f(n) \in o(n)$ with $\forall T \in \mathcal{T} : \Delta(T) \leq f(|T|)$. Then there is another function g, $g(n) \in o(n)$ such that every graph with average degree $k + g(n)$ contains any tree from $\mathcal{T}$ on at most k vertices.

Theorem 1.2 is trivial if $k \leq \eta n/2$, since, as we already mentioned, a graph with average degree at least ηn contains a subgraph with minimum degree at least $\eta n/2$ and we can then embed T greedily in the subgraph. Hence, we interpret this result as one for trees of size linear in the size of the host graph: only for trees of size linear in

the size of the host graph this result is nontrivial. Another viewpoint is to consider this as a result for dense graphs: the additive error term $+\eta n$ ensures that G contains at least $\eta n^2/2$ edges, hence it is dense.

We in fact prove a stronger result—Theorem 1.3—that we first describe less formally. If we fix $r = 1/2$ in the statement of Theorem 1.3, then it states that one can embed any tree with k vertices from a class of trees with sublinear maximum degree in every large enough host graph that fulfills two conditions:

1. there is positive proportion of vertices of degree at least roughly k,
2. the minimum degree is at least roughly $k/2$.

Additionally we get the following trade-off. If our class of trees satisfies that the size of the smaller color class of any tree T from the class is bounded by $r|T|$, it suffices to assume that the minimum degree of the host graph is at least roughly rk. Also, only the vertices of the bigger color class are required to have bounded maximum degree.

THEOREM 1.3. *For any $r, \eta > 0$ there exist n_0 and $\gamma > 0$ such that the following holds. Let G be a graph of order $n > n_0$ and T a tree of order k with two color classes V_1, V_2 such that $|V_1| \leq rk$ and $\forall v \in V_2 : \deg(v) \leq \gamma k$. If $\delta(G) \geq rk + \eta n$, and at least ηn vertices of G have degree at least $k + \eta n$, then G contains T.*

As in the case of Theorem 1.2, this theorem is nontrivial only for trees of size linear in the size of the host graph.

We postpone a simple reduction of Theorem 1.2 to Theorem 1.3, as well as some further remarks regarding the theorem, to section 2.

We believe that Theorem 1.2 for $r = 1/2$ can be substantially generalized and put this generalization as a conjecture. Motivation of this conjecture is the question of how many vertices of degree at least k are needed to embed any tree on $k+1$ vertices if we moreover assume that the minimum degree is at least $k/2$.

CONJECTURE 1.4 (Klimošová, Piguet, Rozhoň). *Every graph G on n vertices with $\delta(G) \geq k/2$ and at least $\frac{1}{2} \cdot \frac{n}{\sqrt{k}}$ vertices of degree at least k contains every tree of order $k+1$.*

We postpone to section 2 the construction that shows that the constant $\frac{1}{2}$ in Conjecture 1.4 cannot be improved.

1.2. Relation of this paper to other work. There are many partial results concerning the Erdős–Sós conjecture. It has been verified for some special families of host graphs [4, 9, 11, 28, 32], special families of trees embedded [14, 15, 25], or when the size of the host graph is only slightly larger than the size of the tree [17, 30, 33].

A solution of this conjecture for large k, based on an extension of the regularity lemma, was announced in the early 1990s by Ajtai, Komlós, Simonovits, and Szemerédi. This result will be published as a sequence of three papers [1, 2, 3]. Although Theorem 1.2 is only a special case of this announced result, we still believe that it is of interest, since its proof is relatively straightforward.

A similar approach to ours was recently independently used in [7] to obtain a result very similar to Theorem 1.2. The only difference is that the authors of [7] require the maximum degree of the respective tree to be less than $k^{\frac{1}{67}}$ instead of $o(k)$ as in Theorem 1.2. The follow-up work [5, 6] then independently proves Theorem 1.2.

The idea to study embedding of trees under conditions on the minimum and maximum degree comes from the paper [18] and was later developed in [5, 6, 7, 8]. In a preliminary version of this paper we conjectured in this direction, together with Klimošová and Piguet, that any graph G that satisfies $\Delta(G) \geq 4k/3$ and $\delta(G) \geq k/2$ embeds any tree on $k+1$ vertices, but this was shown to be false in [8].

To prove Theorem 1.3 we employ the so-called regularity method, which is a relatively standard method used for embedding trees. This method was successfully used for proofs of similar results, notably for the sequence of results [10, 19, 20, 21, 22, 23, 26, 34] on the so-called Loebl–Komlós-Sós conjecture that asserts that each graph with at least half of vertices of degree at least k contains any tree on $k + 1$ vertices.

Simonovits conjectured that one can generalize the statement of the Loebl–Komlós-Sós conjecture. The Simonovits conjecture states that if one assumes that only $r|G|$ vertices ($0 \leq r \leq \frac{1}{2}$) of the host graph G have degree at least k, one can still embed in G all trees of order $k + 1$ such that their smaller color class has size at most $r(k+1)$. The dense approximate version (i.e., the necessary degree is replaced by $k+\eta n$ instead of just k) of this conjecture was proven in [24]. The authors used the word *skew* to denote the ratio r of the size of the smaller color class of the tree and the size of the tree itself. We adopt this notation.

Note that the spirit of the Simonovits conjecture is the same as the spirit of Theorem 1.3, i.e., one can generalize an embedding theorem by considering the skew of the embedded tree as an additional parameter. Highly skewed trees (trees with very small r) can then (at least in the dense approximate setting) be embedded under much milder conditions than general trees. This suggests that the skew of the embedded tree could be considered as a natural parameter showing how hard it is to embed the given tree. For example, the star is an extremely skewed tree that is very easy to embed. On the other hand, the example tree in Figure 1 is also highly skewed, although it is a hard example for certain embedding setting. The usefulness of this parameter is thus yet to be determined.

1.3. Organization of the paper. The paper is organized as follows. In the next section we deduce Theorem 1.2 from Theorem 1.3 and provide several remarks and constructions relevant to the results mentioned in the introduction. In section 3 we explain standard tools that we later use for the proof of Theorem 1.3. Finally, in section 4 we prove Theorem 1.3.

2. Proof of Theorem 1.2 and further remarks. In this section we prove Theorem 1.2 and then further elaborate on several topics already mentioned in the introduction.

Proof of Theorem 1.2. Let $\eta' = \eta/2$ and let G be a graph on $n \geq n_0 = \frac{n_{0,T1.3}(\eta')}{\eta}$ vertices. Here $n_{0,T1.3}(\eta')$ means the output of Theorem 1.3 with input η' and $r = 1/2$. Suppose that $k \geq \eta n/2$.

We choose a subgraph $G' \subseteq G$ such that $\mathrm{d\overline{eg}}(G') \geq k+\eta n$ and $\delta(G') \geq k/2+\eta n/2$. Hence, we know that the size of G' is at least $k + \eta n \geq \eta n \geq n_{0,T1.3}$.

We claim that at least $\eta'|G'|$ vertices of G' have degree at least $k + \eta' n$ and hence we may apply Theorem 1.3. If this was not true, most of the vertices of G' would have degree less than $k + \eta' n$ and hence

$$\mathrm{d\overline{eg}}(G') \leq \eta' \cdot n + (1 - \eta') \cdot (k + \eta' n) < \eta' n + (k + \eta' n) = k + 2\eta' n = k + \eta n,$$

a contradiction.   □

2.1. A graph in Figure 1 fails to embed a random tree. We observe that the example graph with two cliques in Figure 1 not only fails to embed the tree from the same figure but it actually fails to embed most trees.
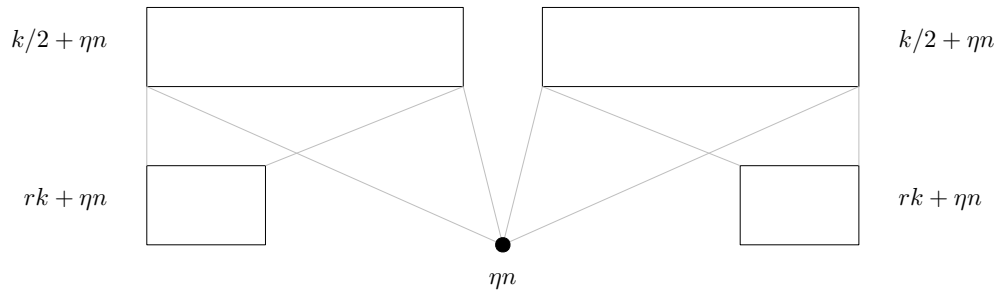
FIG. 2. *Example showing that the condition on bounded degree is needed in the statement of Theorem* 1.3.

PROPOSITION 2.1 (see [31]). *For even* k *the probability that a random unlabeled tree of size* $k + 1$ *can be embedded in the graph* G *consisting of a vertex complete to two cliques of size* $k/2$ *is in* $O(k^{-1/2})$.

Proof. We at first classify trees on $k + 1$ vertices that can be embedded in G. A vertex $u \in T$ is a centroid if after removing it from T we obtain a family of trees such that each tree is of size at most $k/2$. Since the size of the graph is the same as the size of the tree that we embed, only a centroid of T can be embedded in the vertex of G complete to all other vertices. Since $k + 1$ is odd, the centroid of the tree is unique. Hence, T can be embedded if and only if the subtrees created after removing its centroid can be partitioned into two classes such that the number of vertices in each class is $k/2$. We call such trees *balanced*.

Let r_k be the number of unlabeled rooted trees with k vertices. A formula of Otter (see, e.g., [16, p. 481]) states that $r_k = \Theta(k^{-3/2} \cdot B^k)$ for some positive constant B. Similarly, the number of unlabeled unrooted trees s_k is in $\Theta(k^{-5/2} \cdot B^k)$ for the same constant B (again [16, p. 481]).

Note that the number of balanced trees of order $k + 1$ is at most $r_{k/2+1}^2$, since each such tree can be decomposed into two rooted trees with $k/2 + 1$ vertices each. Hence the number of balanced trees is in $O(k^{-3} B^k)$. Comparing this with the sequence s_k, we conclude that the probability that a random unlabeled tree is balanced goes to 0 at a rate of at least $k^{-1/2}$. □

2.2. Remark about tightness of Theorem 1.3. Although the condition on the maximum degree $\Delta(T)$ in Theorem 1.2 is probably not necessary, it is crucial for Theorem 1.3. We show in the following claim that Theorem 1.3 fails when we drop the assumption on the sublinear degree of T.

CLAIM 2.2. *Suppose that* $0 < r < \frac{1}{3}$. *Then there is* $\eta > 0$ *such that there is a graph* G *on* n *vertices and a tree* T *on* k *vertices with the following properties. The minimum degree of* G *is at least* $rk + \eta n$ *and it contains at least* ηn *vertices of degree* $k + \eta n$. *One color class of* T *contains at most* rk *vertices. Finally,* G *does not embed* T.

Proof. Suppose $0 < r < 1/3$ and pick $\eta > 0$ to be sufficiently small depending on the value of r.

Let G be a graph on n vertices consisting of two disjoint copies of complete bipartite graphs with color classes of sizes $rk + \eta n$ and $k/2 + \eta n$. Moreover, ηn additional vertices are complete to both larger color classes of the two bipartite graphs (see Figure 2). This implies that $k = \frac{1 - 5\eta}{1 + 2r} n$; for simplicity we do not address the rounding issues regarding k.

Let T be a tree on k vertices consisting of a vertex x complete to centers of rk stars of sizes $\lfloor \frac{1}{r} \rfloor$ and $\lceil \frac{1}{r} \rceil$.

The smaller color class of T has size rk. Note that for fixed r the maximum degree of this smaller color class of T is constant. However, it is not true for the larger color class, hence Theorem 1.3 does not apply. We claim that the tree T is not contained in G if we have chosen η to be sufficiently small.

Suppose that there is an embedding of T in G. The graph G is bipartite with one color class of size at most $2rk + 3\eta n$. Since k is linear in n and $r < \frac{1}{3}$, we can choose η small enough depending on r so that this expression is less than $(1 - r)k$. Hence, the vertex x must be embedded in the larger color class. Out of $(1 - r)k - 1$ leaves at least $(1 - r)k - 1 - \eta n \cdot \lceil \frac{1}{r} \rceil > k/2 + \eta n$ have to be embedded in the same set of size $k/2 + \eta n$ as x, a contradiction with the assumption that G embeds T. □

Theorem 1.3 is thus an example of an asymptotic result that does not seem to have a natural exact strengthening. On the other hand, we believe that the assumption on the sublinear maximum degree in Theorem 1.3 can be dropped in the case $r = 1/2$. This would mean that the dense asymptotic version of the Erdős–Sós conjecture could be proven by this local approach.

2.3. The constant $\frac{1}{2}$ in Conjecture 1.4 cannot be improved. The following example shows that the constant $\frac{1}{2}$ in Conjecture 1.4 is best possible.

Let $k > 1$ be an odd square and T be a tree of order $k + 1$ consisting of a vertex connected to centers of $\sqrt{k}$ stars on $\sqrt{k}$ vertices. Let G be a graph consisting of two disjoint cliques of order $\frac{k-1}{2}$ and $\frac{k+1}{2}$ and an independent set of $\frac{\sqrt{k}-1}{2}$ vertices complete to both cliques. A simple calculation shows that the proportion of high degree vertices of G is

$$\frac{\frac{\sqrt{k}-1}{2}}{k + \frac{\sqrt{k}-1}{2}} < \frac{1}{2\sqrt{k}} \ .$$

On the other hand, note that for any $c < 1$ the left-hand side is larger than $\frac{c}{2\sqrt{k}}$ for sufficiently large k. We will check that G does not contain T, which in turn shows that the expression $\frac{n}{2\sqrt{k}}$ in the conjecture cannot be strengthened to $\frac{cn}{2\sqrt{k}}$ for any $c < 1$.

Suppose that G embeds T. If the central vertex of T is embedded in the independent set, at least $\frac{\sqrt{k}+1}{2}$ stars neighboring with that vertex have to be embedded in one of the cliques together with the independent set. This means that we have to embed at least $1 + \frac{\sqrt{k}+1}{2} \cdot \sqrt{k} = \frac{k}{2} + \frac{\sqrt{k}}{2} + 1$ vertices in part of G consisting of at most $\frac{k+1}{2} + \frac{\sqrt{k}-1}{2} = \frac{k}{2} + \frac{\sqrt{k}}{2}$ vertices, which is not possible.

Suppose that the central vertex of T is embedded in one of the cliques. We can embed at most $\frac{\sqrt{k}-1}{2} \cdot (\sqrt{k} - 1) = \frac{k-2\sqrt{k}+1}{2}$ vertices of T in the other clique, since each vertex in the independent set enables us to embed $\sqrt{k} - 1$ leaves in the other clique. We cannot use at least $\frac{k-1}{2} - \frac{k-2\sqrt{k}+1}{2} = \sqrt{k} - 1$ vertices of the host graph for the embedding of T. Only $|V(G)| - (\sqrt{k} - 1) = k + \frac{\sqrt{k}-1}{2} - (\sqrt{k} - 1) < k$ vertices in G can be used to embed T, hence it is again not possible to embed T.

3. The regularity method. In this section we state several preparatory results that will be used later for the proof of Theorem 1.3.

The basic idea of using the regularity lemma for embedding trees is that it is easy to embed trees when we know that the host graph is (pseudo)random, because then we can use its expansion properties. The regularity lemma (subsection 3.2) enables us to partition the host graph into a bounded number of clusters such that

the edges between them are behaving in a pseudorandom fashion. We cannot exploit this property to easily embed the whole tree, but we may partition the tree into small subtrees (subsection 3.3) and then it is reasonably easy, though technical, to embed any such small subtree in basically any pair of clusters with a nontrivial amount of edges between them (the second technical lemma in subsection 3.4).

The problem that we are left with (proof in section 4) is to embed the macroscopic structure of the tree that we get after its partitioning in the clusters of the host graph. This problem is quite similar to the problem that we started with; it is indeed tempting to think about this problem as of a fractional embedding, since we may embed several small subtrees in overlapping clusters subject to cardinality constraints. The reality is more complicated and technical, though, so we do not pursue this intuition later in the paper.

3.1. Notation. Throughout the paper we will use the following notation. The edge density of a bipartite graph with color classes X, Y is the fraction $\frac{e(X,Y)}{|X||Y|}$, where $e(X,Y) = |E(X,Y)|$ and $E(X,Y)$ is the set of edges with one endpoint in X and the other in Y. The average degree is defined as $\overline{\deg}(X,Y) = \frac{e(X,Y)}{|X|}$. When we work with a fixed graph G, we use $V(G)$ and $E(G)$ to denote the set of its vertices or edges, respectively. For $X \subseteq V(G)$ we then also write $\overline{\deg}(X)$ instead of $\overline{\deg}(X, G \setminus X)$—this is the average degree of a vertex of X in the subgraph of G induced by edges between X and $V(G) \setminus X$. The neighborhood of a vertex v is the set of vertices $N_G(v) = \{u \in G | \{u,v\} \in E(G)\}$. The neighborhood of a set S is $N_G(S) = \{u \in G | \exists v \in S : \{u,v\} \in E(G)\}$. If T is a tree and $x, y \in T$, $\mathrm{dist}_T(x,y)$ is the length of the unique path between x and y in T, i.e., the number of edges on that path.

3.2. Regularity lemma. We say that (X,Y) is an *ε-regular pair* if for every $X' \subseteq X$ and $Y' \subseteq Y$, $|X'| \geq \varepsilon|X|$, and $|Y'| \geq \varepsilon|Y|$ it holds that $|d(X',Y') - d(X,Y)| \leq \varepsilon$.

We say that a partition $\{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_m\}$ of $V(G)$ is an *ε-regular partition* if $|\mathbf{v}_0| \leq \varepsilon|V(G)|$, and all but at most εm^2 pairs $(\mathbf{v}_i, \mathbf{v}_j)$, $1 \leq i < j \leq m$, are ε-regular. Each set of the partition is called a cluster. We call the cluster $\mathbf{v}_0$ the *garbage set*. We call a regular partition *equitable* if $|\mathbf{v}_i| = |\mathbf{v}_j|$ for every $1 \leq i < j \leq m$.

THEOREM 3.1 (Szemerédi's regularity lemma [29]). *For every $\varepsilon > 0$ there is n_0 and M such that every graph of size at least n_0 admits an ε-regular equitable partition $\{\mathbf{v}_0, \ldots, \mathbf{v}_m\}$ with $1/\varepsilon \leq m \leq M$.*

Given an ε-regular pair (X,Y), we call a vertex $x \in X$ *typical* with respect to a set $Y' \subseteq Y$ if $\deg(x, Y') \geq (d(X,Y) - \varepsilon)|Y'|$. Note that from the definition of regularity it follows that all but at most $\varepsilon|X|$ vertices of X are typical with respect to any subset of Y of size at least $\varepsilon|Y|$.

3.3. Partitioning trees. Here we state a crucial lemma from [22] that allows us to partition the tree into a controllable number of small subtrees that we also informally call *microtrees*. These trees are neighboring with a set of vertices of bounded size consisting of vertices that we informally call *seeds*. Moreover, we need to work separately with seeds from different color classes of T.

In the following definition, as well as in the example partition in Figure 3, the set $W_A \cup W_B$ is the set of seeds of T and the set $\mathcal{D}_A \cup \mathcal{D}_B$ is the set of its microtrees.

DEFINITION 3.2 (see [22, Definition 3.3]). *Let T be a tree on $k+1$ vertices. An ℓ-fine partition of T is a quadruple $(W_A, W_B, \mathcal{D}_A, \mathcal{D}_B)$, where $W_A, W_B \subseteq V(T)$ and $\mathcal{D}_A$ and $\mathcal{D}_B$ are families of subtrees of T such that*
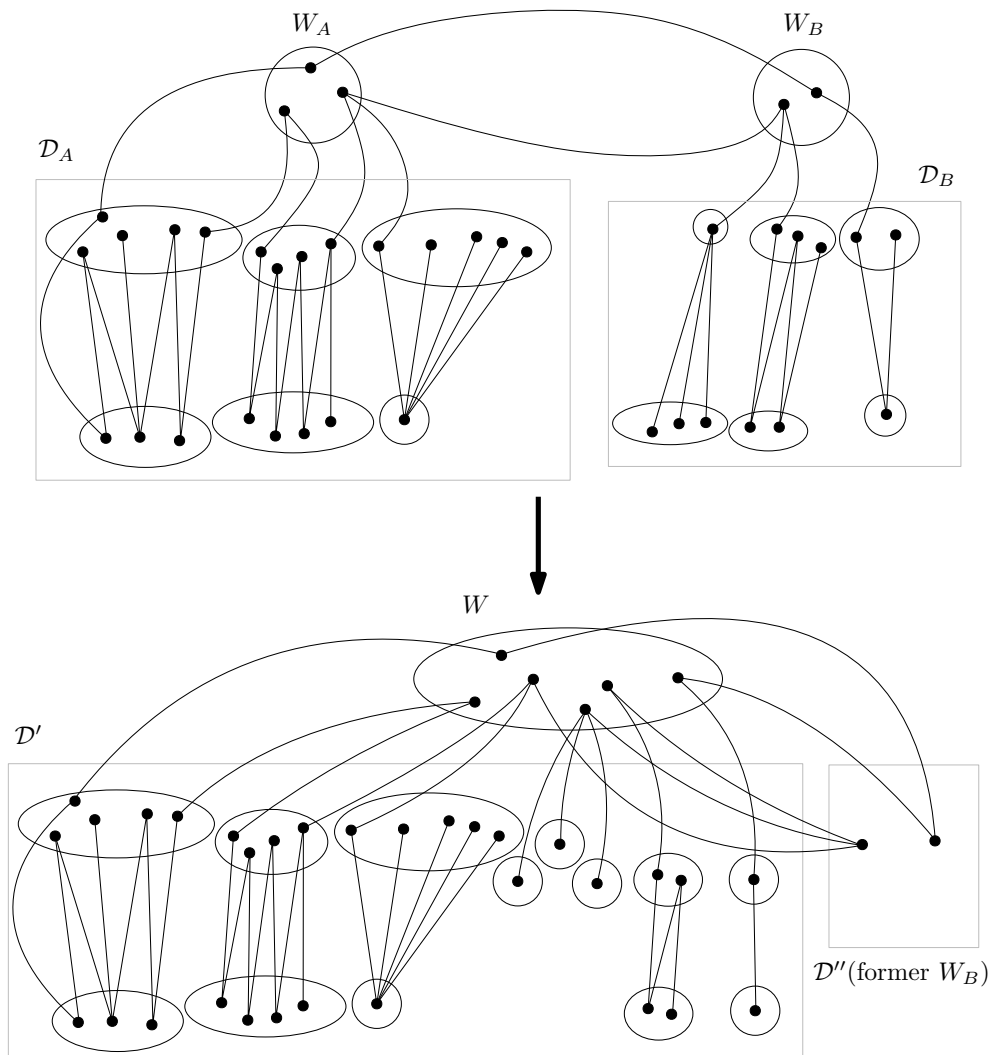
FIG. 3. *An 8-fine partition of a tree and the respective one-sided 8-fine partition of the same tree from the proof of Lemma 3.5.*

1. *the sets W_A, W_B and $\{V(K)\}_{K \in \mathcal{D}_A \cup \mathcal{D}_B}$ partition $V(T)$ (in particular, the trees in $K \in \mathcal{D}_A \cup \mathcal{D}_B$ are pairwise vertex disjoint),*
2. $\max\{|W_A|, |W_B|\} \le 336k/\ell$,
3. *for $w_1, w_2 \in W_A \cup W_B$ their distance in T is odd if and only if one of them lies in W_A and the other one in W_B,*
4. $|K| \le \ell$ *for every tree $K \in \mathcal{D}_A \cup \mathcal{D}_B$,*
5. *for each $K \in \mathcal{D}_A$ we have $N_T(V(K)) \setminus V(K) \subseteq W_A$, and similarly for each $K \in \mathcal{D}_B$ we have $N_T(V(K)) \setminus V(K) \subseteq W_B$,*
6. $|N_T(V(K)) \cap (W_A \cup W_B)| \le 2$ *for each $K \in \mathcal{D}_A \cup \mathcal{D}_B$,*
7. *if $N_T(V(K)) \cap (W_A \cup W_B)$ contains two vertices z_1, z_2 for some $K \in \mathcal{D}_A \cup \mathcal{D}_B$, then $\mathrm{dist}_T(z_1, z_2) \ge 6$.*

We did not list all properties of ℓ-fine partition from [22], only those we need.

LEMMA 3.3 (see [22, Lemma 3.5]).   *Let T be a tree on $k + 1$ vertices and let $\ell \in \mathbb{N}, \ell < k$. Then T has an ℓ-fine partition.*

In the subsequent applications we are always working with $\ell = \beta k$ for some small $\beta > 0$.

Since we work with trees with sublinear degree, we may further constrain the ℓ-fine partition in such a way that all of its seeds are only from one color class of T. We call this simpler structure a *one-sided ℓ-fine* partition.

We at first define precisely the notion of a one-sided ℓ-fine partition and in the subsequent lemma we observe that we may get the one-sided ℓ-fine partition from ℓ-fine partition by adding neighbors of seeds of W_B to W_A and then treating the former set W_B similarly to the set $\mathcal{D}_A \cup \mathcal{D}_B$ (the lower part of Figure 3).

DEFINITION 3.4. *Let T be a tree on $k + 1$ vertices and V_1, V_2 its color classes. Let $\Delta = \max_{v \in V_2} \deg(v)$. A one-sided ℓ-fine partition of T is a pair $(W, \mathcal{D})$, where $W \subseteq V(V_1)$ and $\mathcal{D}$ is a family of subtrees of T such that*
   1. *the sets W and $\{V(K)\}_{K \in \mathcal{D}}$ partition $V(T)$,*
   2. $|W| \leq 336k(1 + \Delta)/\ell$,
   3. $|K| \leq \ell$ *for every tree $K \in \mathcal{D}$,*
   4. *for each $K \in \mathcal{D}$ we have $N_T(V(K)) \setminus V(K) \subseteq W$,*
   5. *we can split $\mathcal{D}$ into two subfamilies, $\mathcal{D} = \mathcal{D}' \sqcup \mathcal{D}''$, in such a way that all trees from $\mathcal{D}'$ have at most two neighbors $z_1, z_2 \in W$ such that $\operatorname{dist}_T(z_1, z_2) \geq 4$, while $|\mathcal{D}''| \leq 336k/\ell$ and every tree from $\mathcal{D}''$ is a single vertex with at most Δ neighbors in W.*

LEMMA 3.5. *Let T be a tree on $k + 1$ vertices and let $\ell \in \mathbb{N}, \ell < k$. Then T has a one-sided ℓ-fine partition.*

Proof. Let $(W_A, W_B, \mathcal{D}_A, \mathcal{D}_B)$ be an ℓ-fine partition of T. Let V_1, V_2 be the partition of the vertices of T into color classes. Suppose that $W_B \subseteq V_2$. Let $W = W_A \cup N_T(W_B)$ and define $\mathcal{D}$ as the set of trees of the forest $T \setminus W$. Conditions 1, 2, and 4 are clearly satisfied. Each vertex from W_B is now a singleton tree in $\mathcal{D}$. Define $\mathcal{D}''$ as the family of these singleton trees and set $\mathcal{D}' = \mathcal{D} \setminus \mathcal{D}''$. Each tree in $\mathcal{D}''$ clearly satisfies conditions 3 and 5. Each tree from $\mathcal{D}'$ is either a tree from $\mathcal{D}_A$ or a subtree of a tree from $\mathcal{D}_B$, and all such trees satisfy condition 3. Finally recall that for each tree from $\mathcal{D}_A \cup \mathcal{D}_B$ with two neighbors z_1 and z_2 in $W_A \cup W_B$ we have $\operatorname{dist}_T(z_1, z_2) \geq 6$. Thus, all trees from $\mathcal{D}_A$ satisfy condition 5. Each tree from $\mathcal{D}_B$ with two neighbors $z_1, z_2 \in W_B$ was split into one tree with two neighbors in W, such that their distance in T is at least 4, and maybe several other trees with only one neighbor in W. All such trees also satisfy condition 5. ☐

3.4. Embedding in regular pairs. In this section we present two embedding lemmas. The first will be used to embed the seeds of a one-sided partition, together with the set $\mathcal{D}''$, in vertices of two neighboring clusters.

PROPOSITION 3.6. *For any $d, \beta, \varepsilon > 0, \varepsilon \leq d^2/100$ there exist k_0 and $\gamma > 0$ such that the following holds.*

Let T be a tree of order $k \geq k_0$ and V_2 one of its color classes such that $\forall v \in V_2 : \deg(v) \leq \gamma k$. Moreover, let $(W, \mathcal{D}), \mathcal{D} = \mathcal{D}' \sqcup \mathcal{D}''$ be its one-sided βk-fine partition. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two clusters of vertices of a graph G forming an ε-regular pair of density at least d. Suppose that $|\mathbf{v}_1| = |\mathbf{v}_2| \geq k/M_{T3.1}(\varepsilon)$, where $M_{T3.1}(\varepsilon)$ is the output of the regularity lemma (Theorem 3.1) with an input ε. Let $U \subseteq \mathbf{v}_1, |U| \leq 2\sqrt{\varepsilon}|\mathbf{v}_1|$. Then
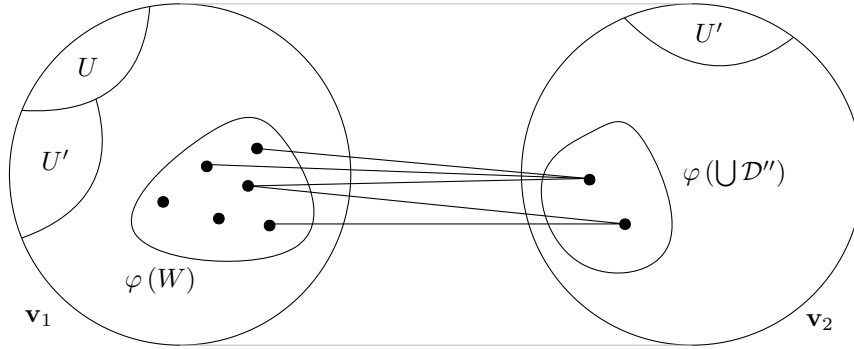
FIG. 4. *Embedding of W and $\bigcup \mathcal{D}''$ in the two clusters $\mathbf{v}_1$ and $\mathbf{v}_2$ in Proposition 3.6.*

there is an injective mapping φ of $W \cup \left(\bigcup \mathcal{D}'' \right)$ that embeds vertices of W in $\mathbf{v}_1 \setminus U$ and vertices of $\bigcup \mathcal{D}''$ in $\mathbf{v}_2$.

Proof. Choose $\gamma, k_0 > 0$ such that

$$\gamma = \frac{\beta d}{2000 M_{T3.1}(\varepsilon)},$$

$$k_0 = \frac{10}{\gamma}.$$

Note that in this case we have

$$\left| \bigcup D'' \right| + |W| \le \frac{336k}{\beta k} + \frac{336k(1 + \gamma k)}{\beta k}$$

$$= \frac{336(\gamma k + 2))}{\beta}$$

$$\boxed{k \ge 10/\gamma} \quad \le \frac{500\gamma k}{\beta}$$

$$\boxed{\text{definition of } \gamma} \quad = \frac{500\beta dk}{\beta \cdot 2000 M_{T3.1}(\varepsilon)} = \frac{dk}{4 M_{T3.1}(\varepsilon)}$$

$$\boxed{|\mathbf{v}_1| \ge k/M_{T3.1}(\varepsilon)} \quad \le \frac{d}{4}|\mathbf{v}_1|.$$

Take an arbitrary vertex $r \notin \bigcup \mathcal{D}''$ of T and root the tree at r. Order all vertices of $W \cup \left(\bigcup \mathcal{D}'' \right)$ according to an order, in which they are visited by a depth-first search starting at r. Let $U' \subseteq \mathbf{v}_1 \cup \mathbf{v}_2$ be the set of vertices of $\mathbf{v}_1$ not typical to $\mathbf{v}_2$ together with vertices of $\mathbf{v}_2$ not typical to $\mathbf{v}_1$. We will provide an algorithm that gradually defines a partial embedding φ of the vertices of $W \cup \left(\bigcup \mathcal{D}'' \right)$ such that $\varphi(W) \subseteq \mathbf{v}_1 \setminus (U \cup U')$ and $\varphi(\bigcup \mathcal{D}'') \subseteq \mathbf{v}_2 \setminus U'$. The situation is displayed in Figure 4.

We iterate over the sequence $x_1, x_2, x_3, \ldots$ of vertices from $W \cup \left(\bigcup \mathcal{D}'' \right)$, where the vertices are ordered by the depth-first search. In the ith step we deal with the vertex $x = x_i$. First we deal with the case $x \in W$.

Suppose that $y \in \bigcup \mathcal{D}''$ is the already embedded parent of x (if $y \notin \bigcup \mathcal{D}''$, our task is simpler, since we do not have to embed x in the neighborhood of $\varphi(y)$). We want to embed x in an arbitrary neighbor of y in $\mathbf{v}_1 \setminus (U \cup \varphi(W) \cup U')$. To do so, it suffices to verify that $N_G(y) \setminus (U \cup \varphi(W) \cup U')$ is nonempty. This can be done with the

help of the fact that $\varphi(y)$ is typical to $\mathbf{v}_1$ and together with our bound $|W| \leq \frac{d}{4}|\mathbf{v}_1|$,

$$|N_G(y) \setminus (U \cup \varphi(W) \cup U')| \geq |\mathbf{v}_1| \left((d - \varepsilon) - 2\sqrt{\varepsilon} - \frac{d}{4} - \varepsilon \right) > 0.$$

Similarly, suppose that $x \in \bigcup \mathcal{D}''$. From the definition of $\mathcal{D}''$ we know that its parent y is certainly in W and $\varphi(y)$ is typical to $\mathbf{v}_2$. Now we similarly verify that

$$\left| N_G(y) \setminus \left(\varphi \left(\bigcup \mathcal{D}'' \right) \cup U' \right) \right| \geq |\mathbf{v}_2| \left((d - \varepsilon) - \frac{d}{4} - \varepsilon \right) > 0. \qquad \square$$

Next, we state a similar proposition that enables us to embed small trees from a fine partition of T in the regular pairs of the host graph. The proposition is a variation on a folklore result and is similar to, e.g., [24, Lemma 5].

PROPOSITION 3.7. *For all $0 < d, \varepsilon \leq 1$ such that $\varepsilon < d^2/100$ there exists $\beta > 0$ such that the following holds.*

Let $\mathbf{v}_1, \mathbf{u}, \mathbf{v}$ be three clusters of vertices of a graph G such that $\mathbf{v}_1\mathbf{u}$ and $\mathbf{u}\mathbf{v}$ are ε-regular pairs of density at least d. Let v_1, v_2 be two (not necessarily distinct) vertices of $\mathbf{v}_1$. Suppose that $|\mathbf{v}_1| = |\mathbf{u}| = |\mathbf{v}| \geq k/M_{T3.1}(\varepsilon)$. Let K be a tree of order at most βk and let x_1, x_2 be any of its two (not necessarily distinct) vertices from the same color class of K. Suppose that $v_1 \neq v_2$ if and only if $x_1 \neq x_2$. Let U be any subset of vertices of $\mathbf{u} \cup \mathbf{v}$ such that $|\mathbf{u} \setminus U| \geq 4\sqrt{\varepsilon}|\mathbf{u}|$ and $|\mathbf{v} \setminus U| \geq 4\sqrt{\varepsilon}|\mathbf{v}|$. Suppose that $|N_G(v_i) \cap (\mathbf{u} \setminus U)| \geq 3\varepsilon|\mathbf{u}|$ for $i = 1, 2$.

Then there is an injective mapping φ of K in $\mathbf{u} \cup \mathbf{v}$ such that $\varphi(V(K)) \cap U = \emptyset$. Moreover, $\varphi(x_1) \in \mathbf{u}$ is a neighbor of v_1 and $\varphi(x_2) \in \mathbf{u}$ is a neighbor of v_2.

Proof. We show the proof for the harder case when $v_1 \neq v_2$. Choose

$$\beta = \varepsilon/M_{T3.1}(\varepsilon).$$

From this we get

$$|\mathbf{v}_1| \geq \frac{k}{M_{T3.1}(\varepsilon)} = \beta \cdot \frac{M_{T3.1}(\varepsilon)}{\varepsilon} \cdot \frac{k}{M_{T3.1}(\varepsilon)} = \frac{\beta k}{\varepsilon}.$$

Let U' be the set of at most $\varepsilon|\mathbf{u}|$ vertices of $\mathbf{u}$ that are not typical to $\mathbf{v} \setminus U$ (note that $|\mathbf{v} \setminus U| \geq \varepsilon|\mathbf{v}|$) together with at most $\varepsilon|\mathbf{v}|$ vertices of $\mathbf{v}$ that are not typical to $\mathbf{u} \setminus U$. The situation is displayed in Figure 5.

Observe that for each vertex $u \in \mathbf{u} \setminus (U \cup U')$ we have

$$|N_G(u) \cap (\mathbf{v} \setminus (U \cup U'))| \geq (d - \varepsilon)|\mathbf{v} \setminus U| - \varepsilon|\mathbf{v}|$$
$$\boxed{|\mathbf{v} \setminus U| \geq 4\sqrt{\varepsilon}|\mathbf{v}|} \quad \geq (d - \varepsilon)4\sqrt{\varepsilon}|\mathbf{v}| - \varepsilon|\mathbf{v}|$$
$$\boxed{d \gg \sqrt{\varepsilon}} \quad \geq \sqrt{\varepsilon} \cdot 4\sqrt{\varepsilon}|\mathbf{v}| - \varepsilon|\mathbf{v}| \geq 2\varepsilon|\mathbf{v}|$$
$$\boxed{|\mathbf{v}| \geq \beta k/\varepsilon} \quad \geq \varepsilon|\mathbf{v}| + \beta k \geq \varepsilon|\mathbf{v}| + |K|,$$

and similar holds for any $u \in \mathbf{v} \setminus (U \cup U')$. This means that during embedding we may always find a neighbor of u in $\mathbf{v} \setminus (U \cup U')$ that was not yet used for embedding.

The same applies for both vertices v_1, v_2. We have

$$|N_G(v_i) \cap (\mathbf{u} \setminus (U \cup U'))| \geq |N_G(v_i) \cap (\mathbf{u} \setminus U)| - \varepsilon|\mathbf{u}|$$
$$\boxed{|N_G(v_i) \cap (\mathbf{u} \setminus U)| \geq 3\varepsilon|\mathbf{u}|} \quad \geq 2\varepsilon|\mathbf{u}|$$
$$\boxed{|\mathbf{u}| \geq \beta k/\varepsilon} \quad \geq \varepsilon|\mathbf{u}| + \beta k \geq \varepsilon|\mathbf{u}| + |K|.$$
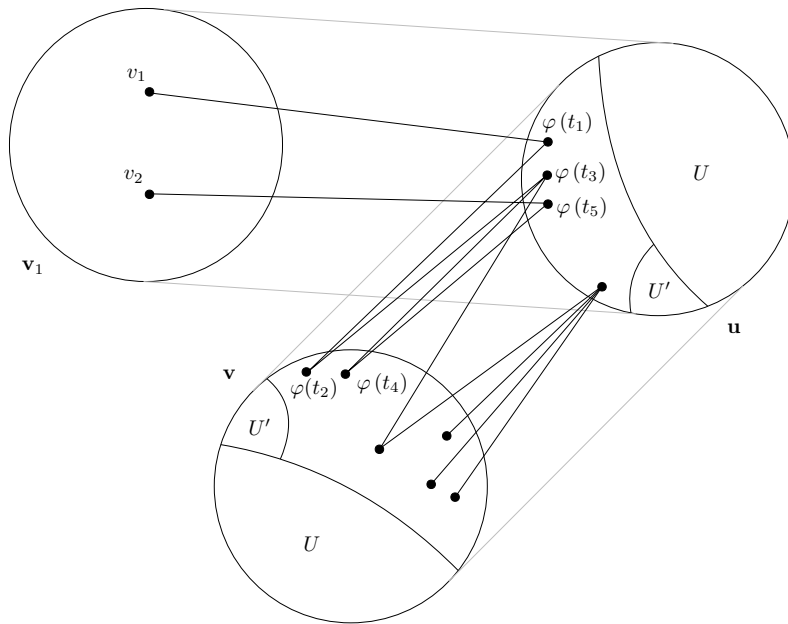
FIG. 5. *Embedding of a small tree K in two clusters $\mathbf{u}$ and $\mathbf{v}$ in Proposition* 3.7. *In this case* $r = 5$.

We start by embedding the path $t_1 = x_1, t_2, \ldots, t_r = x_2$ connecting x_1 with x_2 in K. We embed these vertices alternately in clusters $\mathbf{u}$ and $\mathbf{v}$. We embed x_1 in an arbitrary vertex of $N_G(v_1) \cap \big(\mathbf{u} \setminus (U \cup U')\big)$. Now for i going from 2 to $\ell - 2$ we always map t_i to a neighbor of $\varphi(t_{i-1})$ not lying in $U \cup U'$. Observe that both $N_G(v_2) \cap \big(\mathbf{u} \setminus (U \cup U')\big)$ and $N_G(t_{r-2}) \cap \big(\mathbf{v} \setminus (U \cup U')\big)$ have sizes at least $\varepsilon|\mathbf{v}_1|$; thus there is an edge connecting those two neighborhoods. Map t_{r-1} and t_r to the two endpoints of the edge. The rest of the tree can then be embedded in the greedy manner. $\quad\square$

4. Proof of Theorem 1.3. In this section we prove Theorem 1.3. We split the proof into three parts. At first we preprocess the host graph by applying the regularity lemma and we partition the tree by applying Lemma 3.5. In the second part we find a suitable matching structure in the host graph. In the last part we embed the tree in the host graph.

Preprocessing the host graph and the tree. Fix η, r. Suppose that $\eta < 1$. Choose $d, \varepsilon, \beta, n_0$ such that

$$d = \frac{(\eta r)^2}{1000},$$

$$\varepsilon = \frac{(\eta r d)^{20}}{10^{15}},$$

$$\beta = \min\left(\beta_{P3.7}(d, \varepsilon), \frac{\eta d}{10^5 \cdot M_{T3.1}(\varepsilon)}\right),$$

$$\gamma = \gamma_{P3.6}(d, \varepsilon, \beta),$$

$$n_0 = \max\left(n_{0,T3.1}(\varepsilon), \frac{2}{\eta} k_{0,P3.6}(d, \varepsilon, \beta)\right).$$

Let G be a fixed graph on $n \geq n_0$ vertices with at least ηn vertices of degree at least $k+\eta n$ and with $\delta(G) \geq rk+\eta n$. Suppose that $k \geq \eta n/2$; otherwise we embed the tree T greedily. We apply the regularity lemma (Theorem 3.1) on G with $\varepsilon_{T3.1} = \varepsilon$ and obtain an ε-regular equitable partition $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_m$ with $1/\varepsilon \leq m \leq M_{T3.1}(\varepsilon)$ clusters. Each cluster has average degree at least $rk + \eta n$.

Erase all edges within sets $\mathbf{v}_i$ of the partition, between irregular pairs, and between pairs of density less than d. We have erased at most $m \cdot \binom{n/m}{2} \leq \frac{n^2}{m} \leq \varepsilon n^2$ edges within the sets $\mathbf{v}_i$, at most $\varepsilon m^2 \cdot (n/m)^2 = \varepsilon n^2$ edges in irregular pairs, and at most $\binom{m}{2} \cdot d \cdot (n/m)^2 \leq d \cdot n^2$ edges in pairs of low density. Erase the garbage set $\mathbf{v}_0$ and all of the at most $\varepsilon n \cdot n$ incident edges. Note that we have erased at most $(3\varepsilon + d)n^2$ edges. We abuse notation and still call the resulting graph G.

Note that the quantity $\sum_{1 \leq i \leq m} |\mathbf{v}_i| \cdot \overline{\deg}(\mathbf{v}_i)$ dropped down by at most $(6\varepsilon + 2d)n^2$. Thus there are at most $\sqrt{6\varepsilon + 2d} \cdot m$ clusters such that their average degree dropped down by more than $\sqrt{6\varepsilon + 2d} \cdot n$. Delete all such clusters and incident edges. We again call the resulting graph G. The average degree of each cluster of G that was not deleted at first dropped by at most $\sqrt{6\varepsilon + 2d} \cdot n$. Then we erased at most $\sqrt{6\varepsilon + 2d} \cdot m$ clusters, so now it is at least $rk + \eta n - 2 \cdot \sqrt{6\varepsilon + 2d} \cdot n > rk + \eta n/2$. Moreover, G contains at least $(\eta - \varepsilon - \sqrt{6\varepsilon + 2d})n \geq \eta n/2$ vertices of degree at least $k + \eta n - 2 \cdot \sqrt{6\varepsilon + 2d} \cdot n \geq k + \eta n/2$. Hence, there exists a cluster, which without loss of generality is $\mathbf{v}_1$, such that the proportion of vertices of degree at least $k + \eta n/2$ in that cluster is at least $\eta/2 \geq \varepsilon$. If we denote by L this set of high degree vertices of $\mathbf{v}_1$, we have $\overline{\deg}(\mathbf{v}_1, \mathbf{v}_i) \geq \overline{\deg}(L, \mathbf{v}_i) - \varepsilon|\mathbf{v}_i|$ from regularity of each pair $(\mathbf{v}_1, \mathbf{v}_i)$. This yields that $\overline{\deg}(\mathbf{v}_1) \geq \overline{\deg}(L) - \varepsilon n \geq k + \eta n/3$. Moreover, if it is the case that $\overline{\deg}(\mathbf{v}_1) > 2k$, we erase several regular pairs with one endpoint in $\mathbf{v}_1$ so as to achieve $\overline{\deg}(\mathbf{v}_1) \leq 2k$. After deletion the average degree of each cluster is still at least $k + \eta n/2 - n/m \geq k + \eta n/3$.

The cluster graph $\mathbf{G}$ of G is a graph such that its vertex set is the clusters of G and between any pair of vertices $\mathbf{u}, \mathbf{v} \in \mathbf{G}$ there is an edge with weight $d(\mathbf{u}, \mathbf{v})$ if and only if $\mathbf{uv}$ is a regular pair with density $d(\mathbf{u}, \mathbf{v}) > 0$. Since we already deleted all irregular pairs and pairs with low density, for any edge in the cluster graph we have $d(\mathbf{u}, \mathbf{v}) \geq d$.

We stick to using the boldface font whenever we may think about the corresponding object as a vertex or a set of vertices of $\mathbf{G}$. We use $N_{\mathbf{G}}(\mathbf{v})$ to denote the set of clusters of $\mathbf{G}$ that are neighbors of $\mathbf{v}$ in $\mathbf{G}$, while $N_G(v)$ denotes the neighborhood of a vertex v of G.

After preprocessing the host graph we turn our attention to the tree T. Let V_1, V_2 be its color classes such that $|V_1| \leq rk$ and $\forall v \in V_2 : \deg(v) \leq \gamma k$. We apply Lemma 3.5 with parameter $\ell_{L3.5} = \beta k$ and obtain its one-sided βk-fine partition $(W, \mathcal{D}), \mathcal{D} = \mathcal{D}' \sqcup \mathcal{D}''$ such that $|W| \leq 336(1 + \gamma k)/\beta$ and $|\bigcup \mathcal{D}''| \leq 336/\beta$. Moreover, for each $K \in \mathcal{D}'$ we have $|K| \leq \beta k$ and for each $K \in \mathcal{D}''$ we have $|K| = 1$. Also note that $W \subseteq V_1$.

Structure of the host graph. We now find a suitable structure in the cluster graph $\mathbf{G}$ that will be used for the embedding of T. It suffices to look at the cluster $\mathbf{v}_1$, which will serve for the embedding of the seeds of T, and its neighborhood.

Let $\mathbf{M}$ a maximal matching in $N_{\mathbf{G}}(\mathbf{v}_1)$. We will denote by $\mathbf{M}$ both the graph and its underlying vertex set. Suppose that $\mathbf{uv} \in \mathbf{M}$. Note that from the condition on maximality we get that there cannot be two vertices $\mathbf{x} \neq \mathbf{y} \in N_{\mathbf{G}}(\mathbf{v}_1) \setminus \mathbf{M}$ such that both $\mathbf{xu}$ and $\mathbf{yv}$ are edges of $\mathbf{G}$. Thus there are two possibilities for each edge $\mathbf{uv}$: either only one of its endpoints has neighbors in $N_{\mathbf{G}}(\mathbf{v}_1) \setminus \mathbf{M}$ or both of its endpoints
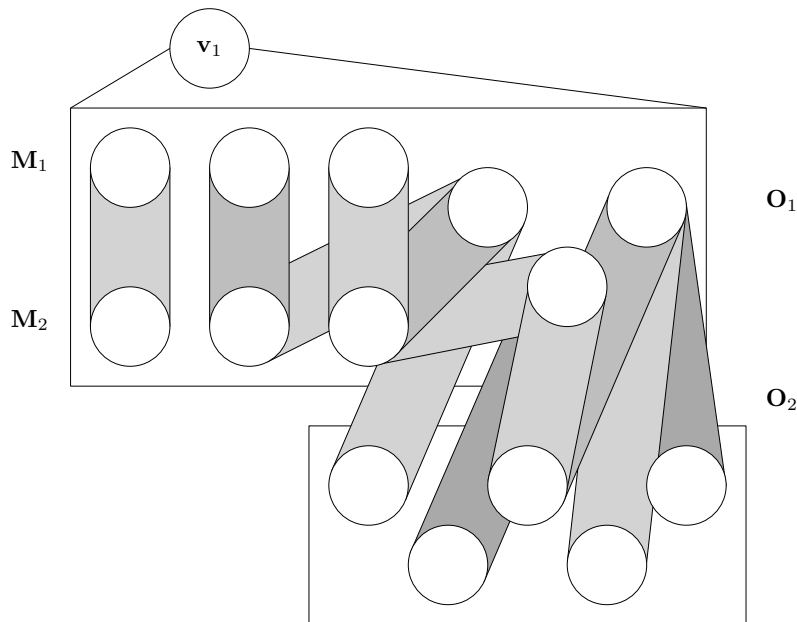
FIG. 6. *Cluster* $\mathbf{v}_1$ *and four sets of clusters* $\mathbf{M}_1, \mathbf{M}_2, \mathbf{O}_1, \mathbf{O}_2$ *that will be used for embedding. The regular pairs of different density are sketch in shades of gray (we omit pairs touching* $\mathbf{v}_1$).

have just one neighbor in $N_\mathbf{G}(\mathbf{v}_1) \setminus \mathbf{M}$. We can get rid of the second special case as follows. For each vertex in $N_\mathbf{G}(\mathbf{v}_1) \setminus \mathbf{M}$ either we delete it if it is a common neighbor of at least $\eta m/40$ matching pairs, or we delete all edges in at most $2 \cdot \eta m/40$ regular pairs connecting the vertex with these matching pairs. In this way we delete at most $40/\eta$ clusters and the degree of all remaining clusters of $\mathbf{G}$ drops down by at most $\eta m/20 \cdot |\mathbf{v}_1| + 40/\eta \cdot |\mathbf{v}_1| \leq (\eta/20 + 40\varepsilon/\eta) \cdot n \leq \eta n/10$. We abuse notation and still call the resulting graph $\mathbf{G}$. The degree of $\mathbf{v}_1$ is at least $k + \eta n/3 - \eta n/10 \geq k + \eta n/5$ and the average degree of every cluster is similarly at least $rk + \eta n/5$. The matching $\mathbf{M}$ is still maximal in $N_\mathbf{G}(\mathbf{v}_1)$. Moreover, we can split the vertices of $\mathbf{M}$ into two color classes, $\mathbf{M} = \mathbf{M}_1 \cup \mathbf{M}_2$, in such a way that only clusters from $\mathbf{M}_2$ have neighbors in $N_\mathbf{G}(\mathbf{v}_1) \setminus \mathbf{M}$. Let $\mathbf{O}_1 = N_\mathbf{G}(\mathbf{v}_1) \setminus \mathbf{M}$. Note that it is an independent set. Define $\mathbf{O}_2 = N_\mathbf{G}(\mathbf{O}_1) \setminus \{\{\mathbf{v}_1\} \cup \mathbf{M}\}$. Note that $N_\mathbf{G}(\mathbf{v}_1) \cap \mathbf{O}_2 = \emptyset$. Set $\mathbf{O} = \mathbf{O}_1 \cup \mathbf{O}_2$. All these sets are displayed in Figure 6.

Embedding. The final step of the proof is broken into three subparts. First we give an overview of the method that we use for the construction of the mapping φ. Then we formulate several preparatory technical claims. In the last part we propose the embedding algorithm.

Overview. We gradually construct an injective mapping φ from T to G. In each step φ denotes the partial embedding that we already constructed. The idea behind the embedding process is very straightforward—we will try to embed microtrees of $\mathcal{D}$ inside the regular pairs in $\mathbf{M}$ and through the vertices of $\mathbf{O}_1$. We will, however, have to overcome several technical difficulties.

One of the standard approaches of embedding trees (e.g., pursued in [24]) is to start by embedding the seeds of T in vertices of two clusters (one for each color class) such that the neighborhoods of these special clusters are sufficiently rich. Moreover, we

embed the seeds in such vertices that are typical to almost all neighboring clusters. We then split the microtrees in T into several subsets and embed each subset of microtrees in some part of the neighborhood of the special clusters. Here we take a different approach. We start in the same way by embedding the seeds W of T in $\mathbf{v}_1$—a high degree cluster of G. We then propose an algorithm that iterates over clusters in the neighborhood of $\mathbf{v}_1$, each time finding two clusters that can be used to embed a microtree.

There are two main technical difficulties that we have to overcome. Recall that each seed is embedded in a vertex that is typical to almost all clusters. This means that when we choose a pair of clusters that will be used for the embedding, we have to find a microtree that has not yet been embedded such that its adjacent seeds are embedded in vertices typical to the first cluster from the pair. We can ensure that there will be such a microtree, unless the number of vertices that remain to be embedded is very small, specifically $\sqrt[4]{\varepsilon}k$. To ensure that we can embed the whole tree T, we at first allocate a small fraction of vertices $F \subseteq \bigcup(\mathbf{M} \cup \mathbf{O})$ that we do not use for the embedding during the main embedding procedure. When only at most $\sqrt[4]{\varepsilon}k$ vertices remain to be embedded, we finally embed this small proportion of trees in the set F.

The second technical problem is that we cannot ensure that all the microtrees have the same skew. This complicates the main embedding procedure that would have been simpler in the case of microtrees with uniform skew. During the embedding procedure we behave against intuition and sometimes redefine the embedding of some microtrees.

Preparations. Note that there are at most $\sqrt{\varepsilon}|\mathbf{v}_1|$ vertices of $\mathbf{v}_1$ that are not typical to more than $\sqrt{\varepsilon}m$ clusters. Indeed, otherwise there would be more than $\sqrt{\varepsilon}|\mathbf{v}_1| \cdot \sqrt{\varepsilon}m = \varepsilon m|\mathbf{v}_1|$ pairs of a cluster and a vertex not typical to it, which in turn implies the existence of a cluster such that more than $\varepsilon|\mathbf{v}_1|$ vertices are not typical to it, a contradiction to the ε-regularity. For each cluster $\mathbf{v} \in \mathbf{M}_1 \cup \mathbf{O}_1$ fix an arbitrary subset $F_{\mathbf{v}}$ of size $\lfloor \eta rd|\mathbf{v}|/300 \rfloor$. Since $|F_{\mathbf{v}}| \geq \varepsilon|\mathbf{v}|$, we may apply the same reasoning to get that there are at most $\sqrt{\varepsilon}|\mathbf{v}_1|$ vertices of $\mathbf{v}_1$ that are not typical to more than $\sqrt{\varepsilon}m$ sets $\{F_{\mathbf{v}_i}\}_{[2,m]}$.

We invoke Proposition 3.6 with parameters $d_{P3.6} = d$, $\beta_{P3.6} = \beta$, $\varepsilon_{P3.6} = \varepsilon$. We also choose $\mathbf{v}_{2,P3.6} = \mathbf{v}_2$ to be any cluster from the neighborhood of $\mathbf{v}_{1,P3.6} = \mathbf{v}_1$. Finally, we define the set $U_{P3.6}$ to be the set of the at most $2\sqrt{\varepsilon}|\mathbf{v}_1|$ vertices not typical to more than $\sqrt{\varepsilon}m$ neighboring clusters $\mathbf{v}_i$, or their subsets $F_{\mathbf{v}_i}$. Note that due to our initial choice of constants all the conditions from the statement of the proposition are satisfied. Hence we embed the vertices of W in $\mathbf{v}_1$, while the vertices of $\bigcup \mathcal{D}''$ will be embedded in $\mathbf{v}_2$. Moreover, each vertex from W is embedded in a vertex typical to all but at most $\sqrt{\varepsilon}m$ clusters $\mathbf{v}_i$ and their fixed subsets $F_{\mathbf{v}_i}$ of size $\lfloor \eta rd|\mathbf{v}_1|/300 \rfloor$.

Note that each microtree $K \in \mathcal{D}'$ has at most two neighbors in W (point 5 in Definition 3.4). We call a cluster $\mathbf{u} \neq \mathbf{v}_1$ *nice* with respect to $K \in \mathcal{D}'$ if the at most two neighbors of K in W are embedded in vertices of $\mathbf{v}_1$ typical to $\mathbf{u}$. We will now, yet again, employ a doublecounting argument to observe that most of the clusters are nice to most of the trees from $\mathcal{D}'$. We claim that there are at most $2\sqrt[4]{\varepsilon}m$ clusters such that if we take all trees such that the cluster is not nice to them, then the union of all such trees contains more than $\sqrt[4]{\varepsilon}k$ vertices.

Suppose that it is not true. Note that each vertex from W was mapped to a vertex that is typical to all but at most $\sqrt{\varepsilon}m$ clusters; thus for each tree K there are at most $2\sqrt{\varepsilon}m$ clusters that are not nice to K. Consider all pairs consisting of

a microtree from $\mathcal{D}'$ and a cluster that is not nice to the tree. Moreover, each such connection shall be weighted by the size of the tree. Consider the sum of all weights of all such pairs. Each tree K contributes to the overall sum by at most $|K| \cdot 2\sqrt{\varepsilon}m$; thus the overall sum is at most $k \cdot 2\sqrt{\varepsilon}m$. On the other hand, if our claim were not true, the overall sum would be bigger than $2\sqrt[4]{\varepsilon}m \cdot \sqrt[4]{\varepsilon}k = k \cdot 2\sqrt{\varepsilon}m$.

Delete all clusters such that the union of all trees not nice to them has more than $\sqrt[4]{\varepsilon}k$ vertices. If they are from $\mathbf{M}$, delete also their neighbors in $\mathbf{M}$. For simplicity also delete the cluster $\mathbf{v_2}$, because then we will not need to consider $\mathcal{D}''$ in further calculations.

Observe that the average degree of each cluster is still at least

$$rk + \eta n/10 - 4\sqrt[4]{\varepsilon}m|\mathbf{v_1}| - |\mathbf{v_2}|$$
$$\boxed{m \geq 1/\varepsilon} \quad \geq rk + \eta n/10 - 4\sqrt[4]{\varepsilon}n - \varepsilon n$$
$$\boxed{\varepsilon \ll \eta} \quad \geq rk + \eta n/20.$$

Similarly, the degree of $\mathbf{v}_1$ is still at least $\overline{\deg}(\mathbf{v}_1) \geq k + \eta n/20$. We still call the new graph $\mathbf{G}$. For each $\mathbf{u} \in N_{\mathbf{G}}(\mathbf{v}_1)$ it now holds that the number of vertices in microtrees such that $\mathbf{u}$ is not nice to them is at most $\sqrt[4]{\varepsilon}k$.

Our main embedding algorithm will work until fewer than $\sqrt[4]{\varepsilon}k$ vertices of T remain to be embedded. To embed the rest of the vertices of T, we now define a set F that intersects each cluster in $\mathbf{M} \cup \mathbf{O}$ in a small fraction of vertices.

CLAIM 4.1. *There is a set $F \subseteq \bigcup(\mathbf{M} \cup \mathbf{O})$ satisfying $|F| \leq \eta r\overline{\deg}(\mathbf{v}_1)/100$, $F_{\mathbf{u}} \subseteq F \cap \mathbf{u}$ for any $\mathbf{u} \in \mathbf{M}_1 \cup \mathbf{O}_1$ and $|F \cap \mathbf{u}| = |F \cap \mathbf{v}|$ for any $\mathbf{uv} \in \mathbf{M}$. Moreover, if we extend our partial mapping φ of $W \cup \bigcup\mathcal{D}''$ in such a way that the extended mapping satisfies $\varphi(T) \cap F = \emptyset$ and φ is defined on the whole T except of some $\bar{\mathcal{D}} \subseteq \mathcal{D}$ with $|\bigcup\bar{\mathcal{D}}| \leq \sqrt[4]{\varepsilon}k$, then we can injectively extend φ to the whole tree T.*

Proof. In the first part of the proof we propose a suitable procedure defining F. Then we show that during the defining procedure the size of F is bounded by $\eta r\overline{\deg}(\mathbf{v}_1)/100$ as desired in the statement of the claim. Finally we use this fact to argue that the defining procedure finishes only after defining the whole F. In the second part of the proof we use Proposition 3.7 to embed a small set of trees $\bar{\mathcal{D}}$ in F.

We define F as follows. For each $\mathbf{u} \in \mathbf{M}_1 \cup \mathbf{O}_1$ we add $F_{\mathbf{u}}$ to F. Then for each set $F_{\mathbf{u}}$ we find a set of the same size in some neighbor $\mathbf{v} \neq \mathbf{v}_1$ of $\mathbf{u}$ and also add this set to F. We call this set $G_{\mathbf{u}}$ and find it as follows. For $\mathbf{uv} \in \mathbf{M}$ we take $G_{\mathbf{u}} = F_{\mathbf{v}}$. For $\mathbf{u} \in \mathbf{O}_1$ we find its neighboring cluster in $\mathbf{O}_2 \cup \mathbf{M}_2$ with at least $\lfloor \eta rd|\mathbf{u}|/300 \rfloor$ vertices that were not yet added to F and we set $G_{\mathbf{u}}$ to be this set (we argue later why we always find a suitable neighboring cluster, i.e., why this defining procedure cannot finish sooner than required). In the case when $F_{\mathbf{u}} \subseteq \mathbf{u} \in \mathbf{O}_1$, but $G_{\mathbf{u}} \subseteq \mathbf{v}' \in \mathbf{M}_2$, it is no longer true that $|F \cap \mathbf{u}'| = |F \cap \mathbf{v}'|$ for some matching edge $\mathbf{u}'\mathbf{v}' \in \mathbf{M}$. We reestablish the condition by adding $\lfloor \eta rd|\mathbf{u}'|/300 \rfloor$ vertices from $\mathbf{u}'$ to F.

During the defining procedure and after it finishes we have

$$|F| \leq 3 \cdot \sum_{\mathbf{u} \in \mathbf{M}_1 \cup \mathbf{O}_1} \lfloor \eta rd|\mathbf{u}|/300 \rfloor \leq \eta r\overline{\deg}(\mathbf{v}_1)/100.$$

Now we argue that each cluster $\mathbf{u} \in \mathbf{O}_1$ has a neighbor in $\mathbf{M}_2 \cup \mathbf{O}_2$ with at least $\lfloor \eta rd|\mathbf{u}|/300 \rfloor$ vertices that are not yet in F. Since we know that

$$\overline{\deg}\left(\mathbf{u}, \bigcup(\mathbf{M}_2 \cup \mathbf{O}_2)\right) \geq rk$$
$$\boxed{|F|/r \ll \overline{\deg}(\mathbf{v}_1) \leq 2k} \quad > 2|F| > 2\overline{\deg}(\mathbf{u}, F),$$

there is certainly a cluster $\mathbf{v} \in \mathbf{M}_2 \cup \mathbf{O}_2$ such that $\overline{\deg}(\mathbf{u}, \mathbf{v})/2 > \overline{\deg}(\mathbf{u}, F \cap \mathbf{v})$, hence $\overline{\deg}(\mathbf{u}, \mathbf{v} \setminus F) = \overline{\deg}(\mathbf{u}, \mathbf{v}) - \overline{\deg}(\mathbf{u}, \mathbf{v} \cap F) > \overline{\deg}(\mathbf{u}, \mathbf{v})/2 \geq d|\mathbf{v}|/2$, meaning that there is a subset of at least $d|\mathbf{v}|/2 > \lfloor \eta r d|\mathbf{v}|/300 \rfloor$ vertices in $\mathbf{v}$ that can be used to define $G_{\mathbf{u}}$.

It remains to show how to embed any family of trees $\bar{\mathcal{D}}$ of small size in F. We define the embedding φ of all trees $K \in \bar{\mathcal{D}}$ in a step-by-step manner. Suppose that $\mathbf{u} \in \mathbf{M}_1 \cup \mathbf{O}_1$ and $G_{\mathbf{u}} \subseteq \mathbf{v}$. If the at most two neighbors z_1, z_2 of K in W are embedded to two vertices of $\mathbf{v}_1$ that are typical to set $F_{\mathbf{u}}$ and, moreover, $|\varphi(T) \cap F_{\mathbf{u}}| \leq \frac{d}{2}|F_{\mathbf{u}}|$ and $|\varphi(T) \cap G_{\mathbf{u}}| \leq \frac{d}{2}|G_{\mathbf{u}}|$, we can compute that for $i = 1, 2$ we have

$$|F_{\mathbf{u}} \setminus \varphi(T)| \geq \left(1 - \frac{d}{2}\right)|F_{\mathbf{u}}| \geq \frac{1}{2}\lfloor \eta r d|\mathbf{u}|/300 \rfloor$$
$$\boxed{\varepsilon \ll (\eta r d)^2} \quad \geq 4\sqrt{\varepsilon}|\mathbf{u}|$$

and similarly $|G_{\mathbf{u}} \setminus \varphi(T)| \geq 4\sqrt{\varepsilon}|\mathbf{v}|$. We also have

$$|N_G(v_i) \cap (F_{\mathbf{u}} \setminus \varphi(T))| \geq |N_G(v_i) \cap F_{\mathbf{u}}| - |\varphi(T) \cap F_{\mathbf{u}}|$$
$$\boxed{v_i \text{ is typical to } F_{\mathbf{u}}} \quad \geq (d - \varepsilon)|F_{\mathbf{u}}| - |\varphi(T) \cap F_{\mathbf{u}}|$$
$$\boxed{\varepsilon \ll d, \; |\varphi(T) \cap F_{\mathbf{u}}| \leq d|F_{\mathbf{u}}|/2} \quad \geq \frac{d}{3}|F_{\mathbf{u}}| \geq \frac{d}{3}\lfloor \eta r d|\mathbf{u}|/300 \rfloor$$
$$\boxed{\varepsilon \ll (\eta r d)^2} \quad \geq 3\varepsilon|\mathbf{u}|.$$

Hence in this case we can use Proposition 3.7 with parameters $U_{P3.7} = (V(G) \setminus F) \cup \varphi(T)$, $d_{P3.7} = d$, $\varepsilon_{P3.7} = \varepsilon$, $\beta_{P3.7} = \beta$, $\mathbf{v}_{1,P3.7} = \mathbf{v}_1$, $\mathbf{u}_{P3.7} = \mathbf{u}$, $\mathbf{v}_{P3.7} = \mathbf{v}$, $K_{P3.7} = K$, $v_{1,P3.7} = \varphi(z_1)$, $v_{2,P3.7} = \varphi(z_2)$. The proposition then allows us to embed K.

Now it suffices to show that for any K we always find a suitable $\mathbf{u}$ such that $\varphi(z_1), \varphi(z_2)$ are typical to $F_{\mathbf{u}}$ and both $F_{\mathbf{u}}$ and $G_{\mathbf{u}}$ do not contain many embedded vertices of T. Recall that vertices $\varphi(z_1), \varphi(z_2)$ are typical to all but at most $\sqrt{\varepsilon}m$ sets $F_{\mathbf{u}}$. If we cannot use for the embedding any other set $F_{\mathbf{u}}$ from remaining clusters of $\mathbf{M}_1 \cup \mathbf{O}_1$, it means that we have embedded more than $\frac{d}{2} \cdot \lfloor \eta r d|\mathbf{v}_1|/300 \rfloor$ vertices to any such set $F_{\mathbf{u}}$, or we have embedded at least the same number of vertices in the appropriate set $G_{\mathbf{u}}$. This means that the number of vertices we have embedded is at least

$$\left(|\mathbf{M}_1 \cup \mathbf{O}_1| - 2\sqrt{\varepsilon}m\right) \cdot \left(\frac{d}{2} \cdot \lfloor \eta r d|\mathbf{v}_1|/300 \rfloor\right) \geq \left(\frac{|\mathbf{M} \cup \mathbf{O}_1|}{2} - 2\sqrt{\varepsilon}m\right) \cdot \frac{d^2 r \eta}{700}|\mathbf{v}_1|$$
$$\boxed{\overline{\deg}(\mathbf{v}_1) \leq |\mathbf{M} \cup \mathbf{O}_1| \cdot |\mathbf{v}_1|} \quad \geq \left(\frac{\overline{\deg}(\mathbf{v}_1)}{2} - 2\sqrt{\varepsilon}m|\mathbf{v}_1|\right) \cdot \frac{d^2 r \eta}{700}$$
$$\boxed{\varepsilon \ll d^{10} r^5 \eta^5} \quad \geq \left(\frac{\overline{\deg}(\mathbf{v}_1)}{2} - 2\sqrt{\varepsilon}m|\mathbf{v}_1|\right) \cdot \sqrt[5]{\varepsilon}$$
$$\boxed{m|\mathbf{v}_1| \leq n} \quad \geq \left(\frac{k}{2} - 2\sqrt{\varepsilon}n\right) \cdot \sqrt[5]{\varepsilon}$$
$$\boxed{k \geq \eta n/2} \quad \geq \left(\frac{1}{2} - \frac{4\sqrt{\varepsilon}}{\eta}\right)\sqrt[5]{\varepsilon}k$$
$$\quad > \sqrt[4]{\varepsilon}k,$$

a contradiction.   □

Embedding algorithm. So far we have embedded the set W in vertices of $\mathbf{v}_1$ that are typical to almost all clusters in the neighborhood of $\mathbf{v}_1$. We also embedded the small set $\mathcal{D}''$ in $\mathbf{v}_2$. We denote this partial embedding by φ_0 and we use the symbol φ to denote the successive extensions of φ_0 that we will construct.

We invoke Claim 4.1 to get a small set F. Now we will gradually embed microtrees from $\mathcal{D}$ in $\bigcup(\mathbf{M} \cup \mathbf{O}) \setminus F$, until the number of vertices of microtrees that were not embedded yet is at most $\sqrt[4]{\varepsilon}k$. Then we embed the remaining parts of T in F using Claim 4.1.

We will use the following notation for the sake of brevity.

DEFINITION 4.2. *Let φ be a fixed partial embedding of T in G extending φ_0. We say that a cluster $\mathbf{u}$ is full if*

$$\left| \mathbf{u} \cap \big(\varphi(V(T)) \cup F \big) \right| \geq |\mathbf{u}| - 4\sqrt{\varepsilon}\,|\mathbf{u}|.$$

We say that a cluster $\mathbf{u} \in N_{\mathbf{G}}(\mathbf{v}_1)$ is saturated, if

$$\left| \mathbf{u} \cap \big(\varphi(V(T)) \cup F \big) \right| \geq \overline{\deg}(\mathbf{v}_1, \mathbf{u}) - 4\sqrt{\varepsilon}\,|\mathbf{u}|.$$

We say that a matching edge $\mathbf{u}\mathbf{v} \in \mathbf{M}$ is saturated if

$$\left| (\mathbf{u} \cup \mathbf{v}) \cap \big(\varphi(V(T)) \cup F \big) \right| \geq \overline{\deg}(\mathbf{v}_1, (\mathbf{u} \cup \mathbf{v})) - 8\sqrt{\varepsilon}\,|\mathbf{u}| - \beta k.$$

Note that every full cluster is also saturated. The intuition behind these definitions will be clear from the statements of the following claims. Recall that V_1, V_2 denote the two color classes of T and that $W \subseteq V_1$.

CLAIM 4.3. *Let φ be a partial embedding of T in G extending φ_0. If $\mathbf{u} \in N_{\mathbf{G}}(\mathbf{v}_1)$ is not saturated and $\mathbf{v} \in N_{\mathbf{G}}(\mathbf{u}) \setminus \{\mathbf{v}_1\}$ is not full, then, unless φ satisfies $|\mathrm{dom}(\varphi)| \geq k - \sqrt[4]{\varepsilon}k$, we may extend φ injectively to some $K \in \mathcal{D}'$ that was not yet embedded in such a way that $\varphi(V(K) \cap V_2) \subseteq \mathbf{u}$, $\varphi(V(K) \cap V_1) \subseteq \mathbf{v}$, and $\varphi(V(K)) \cap F = \emptyset$.*

Proof. We have ensured that the union of all trees of $\mathcal{D}$ such that $\mathbf{u}$ is not nice to them has at most $\sqrt[4]{\varepsilon}k$ vertices. Hence there is a yet nonembedded tree $K \in \mathcal{D}$ such that its at most two neighbors t_1, t_2 in W are embedded in vertices of $\mathbf{v}_1$ that are typical to $\mathbf{u}$. Note that this implies that

$$\left| N_G(t_i) \cap \Big(\mathbf{u} \setminus \big(\varphi(V(T)) \cup F \big) \Big) \right| \geq \overline{\deg}(\mathbf{v}_1, \mathbf{u}) - \varepsilon|\mathbf{u}| - |\mathbf{u} \cap \big(\varphi(V(T)) \cup F \big)|$$

$$\boxed{\mathbf{u} \text{ is not saturated}} \quad \geq \overline{\deg}(\mathbf{v}_1, \mathbf{u}) - \varepsilon|\mathbf{u}| - \big(\overline{\deg}(\mathbf{v}_1, \mathbf{u}) - 4\sqrt{\varepsilon}\,|\mathbf{u}| \big)$$

$$\geq 3\varepsilon|\mathbf{u}|.$$

Recall that neither $\mathbf{u}$ nor $\mathbf{v}$ is full. Together with the inequality above this enables us to apply Proposition 3.7 with $d_{P3.7} := d, \varepsilon_{P3.7} = \varepsilon, \beta_{P3.7} = \beta, \mathbf{v}_{1,P3.7} = \mathbf{v}_1, \mathbf{u}_{P3.7} = \mathbf{u}, \mathbf{v}_{P3.7} = \mathbf{v}, K_{P3.7} = K, v_{i,P3.7} = \varphi(t_i), x_{i,P3.7} = N_T(t_i) \cap K, U_{P3.7} = \varphi(V(T)) \cup F$. The proposition then allows us to extend injectively φ to K. □

CLAIM 4.4. *Let φ be a partial embedding of T in G extending φ_0.*
  1. *There exists either an unsaturated cluster of $\mathbf{O}_1$ or an unsaturated edge of $\mathbf{M}$.*
  2. *Suppose that $\varphi(V_2) \cap \bigcup \mathbf{M_2} = \emptyset$ and let $\mathbf{u} \in \mathbf{O}_1$. There exists a cluster in $N_{\mathbf{G}}(\mathbf{u}) \setminus \{\mathbf{v}_1\}$ that is not full.*

Proof.

1. Suppose that each edge of $\mathbf{M}$ is saturated. For every $\mathbf{uv} \in \mathbf{M}$ we then have

$$
\begin{aligned}
\big|(\mathbf{u} \cup \mathbf{v}) \cap \big(\varphi(V(T)) \cup F\big)\big| & \\
\boxed{\mathbf{uv} \text{ is saturated}} \quad &\geq \overline{\deg}(\mathbf{v}_1, \mathbf{u} \cup \mathbf{v}) - 8\sqrt{\varepsilon}\,|\mathbf{u}| - \beta k \\
\boxed{\overline{\deg}(\mathbf{v}_1, \mathbf{u} \cup \mathbf{v}) \geq 2d|\mathbf{v}_1|} \quad &\geq \overline{\deg}(\mathbf{v}_1, \mathbf{u} \cup \mathbf{v})\left(1 - \frac{8\sqrt{\varepsilon}\,|\mathbf{v}_1| + \beta k}{2d|\mathbf{v}_1|}\right) \\
\boxed{|\mathbf{v}_1| \geq n/M_{T3.1}(\varepsilon),\ k \leq n} \quad &\geq \overline{\deg}(\mathbf{v}_1, \mathbf{u} \cup \mathbf{v})\left(1 - \frac{4\sqrt{\varepsilon}}{d} - \frac{\beta n}{2dn/M_{T3.1}(\varepsilon)}\right) \\
\boxed{\varepsilon \ll (\eta d)^2,\ \beta \ll d\eta/M_{T3.1}(\varepsilon)} \quad &\geq \overline{\deg}(\mathbf{v}_1, \mathbf{u} \cup \mathbf{v})(1 - \eta/100).
\end{aligned}
$$

Suppose that each cluster of $\mathbf{O}_1$ is saturated. After a similar calculation, we get that for each $\mathbf{u} \in \mathbf{O}_1$ we have

$$
|\mathbf{u} \cap (\varphi(V(T)) \cup F)| \geq \overline{\deg}(\mathbf{v}_1, \mathbf{u})(1 - \eta/100).
$$

Hence we have

$$
\begin{aligned}
\Big|\bigcup(\mathbf{M} \cup \mathbf{O}_1) \cap (\varphi(V(T)) \cup F))\Big| &\geq \overline{\deg}(\mathbf{v}_1)(1 - \eta/100) \\
&= \eta\overline{\deg}(\mathbf{v}_1)/100 + \overline{\deg}(\mathbf{v}_1)(1 - \eta/50) \\
\boxed{\overline{\deg}(\mathbf{v}_1) \geq k + \eta k/20} \quad &\geq \eta\overline{\deg}(\mathbf{v}_1)/100 + (k + \eta k/20)(1 - \eta/50) \\
\boxed{|F| \leq \eta\overline{\deg}(\mathbf{v}_1)/100} \quad &> |F| + k \geq |\varphi(V(T)) \cup F|,
\end{aligned}
$$

a contradiction.

2. Suppose that all clusters in $N_{\mathbf{G}}(\mathbf{u}) \setminus \{\mathbf{v}_1\}$ are full. For each full cluster $\mathbf{v}$ we have

$$
\begin{aligned}
|\mathbf{v} \cap (\varphi(V(T)) \cup F)| &\geq |\mathbf{v}|(1 - 4\sqrt{\varepsilon}) \\
\boxed{\varepsilon \ll \eta^2} \quad &\geq |\mathbf{v}|(1 - \eta/200) \\
&\geq \overline{\deg}(\mathbf{u}, \mathbf{v})(1 - \eta/200).
\end{aligned}
$$

Since this holds for any $\mathbf{v} \in N_{\mathbf{G}}(\mathbf{u}) \setminus \{\mathbf{v}_1\}$, we get that

$$
\begin{aligned}
\Big|\bigcup(\mathbf{M}_2 \cup \mathbf{O}_2) \cap (\varphi(V(T)) \cup F))\Big| &\geq \overline{\deg}\Big(\mathbf{u}, \bigcup(\mathbf{M}_2 \cup \mathbf{O}_2) \setminus \mathbf{v}_1\Big)(1 - \eta/200) \\
&\geq \Big(\overline{\deg}\Big(\mathbf{u}, \bigcup(\mathbf{M}_2 \cup \mathbf{O}_2)\Big) - |\mathbf{v}_1|\Big)(1 - \eta/200) \\
\boxed{N_{\mathbf{G}}(\mathbf{u}) \subseteq \{\mathbf{v}_1\} \cup \mathbf{M}_2 \cup \mathbf{O}_2} \quad &\geq \overline{\deg}(\mathbf{u})(1 - \eta/200) - |\mathbf{v}_1| \\
\boxed{\frac{\eta}{200}\overline{\deg}(u) \geq \frac{\eta r k}{200} \geq \frac{\eta^2 r n}{200} \gg \varepsilon n \geq |\mathbf{v}_1|} \quad &\geq \overline{\deg}(\mathbf{u})(1 - \eta/100).
\end{aligned}
$$

Hence, we have

$$\left|\bigcup(\mathbf{M_2} \cup \mathbf{O_2}) \cap (\varphi(V_1) \cup F))\right|$$

$$\boxed{\varphi(V_2) \cap (\bigcup(\mathbf{M_2} \cup \mathbf{O_2})) = \emptyset} \quad = \left|\bigcup(\mathbf{M_2} \cup \mathbf{O_2}) \cap (\varphi(V(T)) \cup F))\right|$$

$$\boxed{\text{inequality above}} \quad \geq \mathrm{d}\overline{\mathrm{eg}}(\mathbf{u})(1 - \eta/100)$$

$$\geq \eta \mathrm{d}\overline{\mathrm{eg}}(\mathbf{u})/50 + \mathrm{d}\overline{\mathrm{eg}}(\mathbf{u})(1 - \eta/30)$$

$$\boxed{\mathrm{d}\overline{\mathrm{eg}}(\mathbf{v_1}) \leq 2k,\ \mathrm{d}\overline{\mathrm{eg}}(\mathbf{u}) \geq rk + \eta k/20} \quad \geq \eta r \mathrm{d}\overline{\mathrm{eg}}(\mathbf{v_1})/100 + (rk + \eta k/20)(1 - \eta/30)$$

$$> |F| + rk$$

$$\geq |F| + |V_1| \geq |(\varphi(V_1) \cup F))|,$$

a contradiction.   □

We can now finish the proof of Theorem 1.3.

Proof. We will gradually embed microtrees from $\mathcal{D}'$ in $\bigcup(\mathbf{M} \cup \mathbf{O})$ in a specified manner using Claim 4.3 (hence avoiding the set F) until $|\mathrm{dom}(\varphi)| \geq k - \sqrt[4]{\varepsilon}k$ or all edges of $\mathbf{M}$ and all vertices of $\mathbf{O}$ are saturated—from Claim 4.4.1 we know that the latter actually cannot be true. When $|\mathrm{dom}(\varphi)| \geq k - \sqrt[4]{\varepsilon}k$, we finish by applying Claim 4.1 on our set F. We split the embedding procedure into three phases:

1. *Phase 1—saturating the matching edges of* $\mathbf{M}$. In the first phase we embed gradually the microtrees of $\mathcal{D}'$ in the edges of $\mathbf{M}$ in such a way that for each $K \in \mathcal{D}'$ we have $\varphi(K \cap V_2) \subseteq \mathbf{M_1}$. We run the process of applying Claim 4.3 for each edge $\mathbf{uv}$ until either $\mathbf{u} \in \mathbf{M_1}$ is saturated, $\mathbf{v} \in \mathbf{M_2}$ is full, or $|\mathrm{dom}(\varphi)| \geq k - \sqrt[4]{\varepsilon}k$.

2. *Phase 2—saturating the clusters in* $\mathbf{O}$. We repeatedly pick an unsaturated cluster $\mathbf{v} \in \mathbf{O_1}$ and then embed trees from $\mathcal{D}'$ in it by repeatedly applying Claim 4.3 in such a way that for each embedded K we have $\varphi(K \cap V_2) \subseteq \mathbf{O_1}$ and $\varphi(K \cap V_1) \subseteq \mathbf{M_2} \cup \mathbf{O_2}$. Note that due to Claim 4.4.2 the cluster $\mathbf{v}$ always has a neighbor that is not full and therefore can be used for the embedding. Hence we can apply this procedure until all clusters from $\mathbf{O_1}$ are saturated, or $|\mathrm{dom}(\varphi)| \geq k - \sqrt[4]{\varepsilon}k$.

3. *Phase 3—finalizing the matching* $\mathbf{M}$. All clusters in $\mathbf{O_1}$ are now saturated. Our task is now to show how to saturate the remaining edges of $\mathbf{M}$. This may not be possible with the current φ as it is defined right now, since it could have happened, for example, that after the first phase we completely filled one cluster from a matching pair, while the other cluster remained almost empty. We solve this problem by potentially redefining the embedding of several microtrees that were embedded in $\mathbf{M_1} \cup \mathbf{M_2}$ in Phase 1.

   Note that for each edge $\mathbf{uv} \in \mathbf{M}$, $\mathbf{u} \in \mathbf{M_1}$, it is true that either $\mathbf{u}$ is saturated or $\mathbf{v}$ is full at the end of Phase 1. We deal with the first case in part (a). In the latter case we did not embed anything in $\mathbf{v}$ in Phase 2. We undefine embedding of all trees that were embedded in $\mathbf{uv}$ and saturate this edge in part (c).

   (a) If $\mathbf{u}$ is saturated, we repeatedly embed trees in $\mathbf{uv}$ in such a way that for each $K \in \mathcal{D}'$ we have $\varphi(K \cap V_2) \subseteq \mathbf{v}$. We do this until either $\mathbf{u}$ is full or $\mathbf{v}$ is saturated. In the latter case the whole edge is saturated. We deal with the first case in (b).

   (b) Suppose that $\mathbf{u}$ is full, but $\mathbf{v}$ is not saturated. Note that Claim 4.1 ensures that $|F \cap \mathbf{u}| = |F \cap \mathbf{v}|$. Hence it must be the case that $|\varphi(V(T)) \cap$

$\mathbf{u}| \geq |\varphi(V(T)) \cap \mathbf{v}|$. Moreover, in Phase 2 we did not embed trees in $\mathbf{u}$. This means that there exists a tree $K \in \mathcal{D}'$ that was embedded in the matching edge $\mathbf{uv}$ in such a way that $|\varphi(V(K)) \cap \mathbf{u}| \geq |\varphi(V(K)) \cap \mathbf{v}|$. As long as it is true that $|\varphi(V(T)) \cap \mathbf{u}| \geq |\varphi(V(T)) \cap \mathbf{v}|$, we find any tree K with this property and we undefine its embedding. When this procedure ends, we have $||\varphi(V(T)) \cap \mathbf{u}| - |\varphi(V(T)) \cap \mathbf{v}|| \leq \beta k$. We later refer to this inequality as the *balancing* condition.

(c) Finally, it suffices to show how to saturate an edge $\mathbf{uv}$ fulfilling the balancing condition (note that if $\varphi(V(T)) \cap \mathbf{uv} = \emptyset$, then the matching edge certainly fulfills the condition). We again embed the microtrees in $\mathbf{uv}$ one after another. Unless one of the clusters is saturated, we choose to embed $K \in \mathcal{D}'$ in such a way that the color class of K with fewer vertices is embedded in the cluster such that more of its vertices were already used for the embedding of T. In this way we ensure that the balancing condition still holds.

After one cluster, say, $\mathbf{u}$, becomes saturated, we continue by embedding only in such a way that for each $K \in \mathcal{D}'$ we have $\varphi(K \cap V_2) \subseteq \mathbf{v}$. We do this until either $\mathbf{v}$ becomes saturated or $\mathbf{u}$ is full. In the first case the whole edge $\mathbf{uv}$ is clearly saturated. In the other case note that we have

$$|(\varphi(V(T)) \cup F) \cap \mathbf{u}|$$
$$\boxed{\mathbf{u} \text{ is full}} \quad \geq |\mathbf{u}| - 4\sqrt{\varepsilon}|\mathbf{u}|$$
$$\boxed{|\mathbf{u}| = |\mathbf{v}| \geq \mathrm{d\overline{eg}}(\mathbf{v}_1, \mathbf{v})} \quad \geq \mathrm{d\overline{eg}}(\mathbf{v}_1, \mathbf{v}) - 4\sqrt{\varepsilon}|\mathbf{v}|$$

and hence

$$|(\varphi(V(T)) \cup F) \cap \mathbf{v}|$$
$$\boxed{\text{the balancing condition}} \quad \geq |(\varphi(V(T)) \cup F) \cap \mathbf{u}| - \beta k$$
$$\boxed{\mathbf{u} \text{ is full, hence saturated}} \quad \geq \mathrm{d\overline{eg}}(\mathbf{v}_1, \mathbf{u}) - 4\sqrt{\varepsilon}|\mathbf{u}| - \beta k.$$

Hence, the whole matching edge is saturated.

We described an algorithm that terminates when $|\mathrm{dom}(\varphi)| \geq k - \sqrt[4]{\varepsilon}k$, or all edges of $\mathbf{M}$ and all vertices of $\mathbf{O}$ are saturated. But the latter cannot happen due to Claim 4.4.1. We finish by invoking Claim 4.1. □

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] S. BALASUBRAMANIAN AND E. DOBSON, *Constructing trees in graphs with no $K_{2,s}$*, J. Graph Theory, 56 (2007), pp. 301–310, https://doi.org/10.1002/jgt.20261.

[5] G. Besomi, *Tree Embeddings in Dense Graphs*, Master's thesis, Universidad de Chile, Santiago, Chile, 2018.

[6] G. Besomi, M. Pavez-Signé, and M. Stein, *On the Erdős-Sós conjecture for bounded degree trees*, in preparation.

[7] G. Besomi, M. Pavez-Signé, and M. Stein, *Degree Conditions for Embedding Trees*, preprint, arXiv:1805.07338, 2018.

[8] G. Besomi, M. Pavez-Signé, and M. Stein, *Maximum and Minimum Degree Conditions for Embedding Trees*, preprint, arXiv:1808.09934, 2018.

[9] S. Brandt and E. Dobson, *The Erdős-Sós conjecture for graphs of girth* 5, Discrete Math., 150 (1996), pp. 411–414, https://doi.org/10.1016/0012-365X(95)00207-D.

[10] O. Cooley, *Proof of the Loebl-Komlós-Sós conjecture for large, dense graphs*, Discrete Math., 309 (2009), pp. 6190–6228.

[11] E. Dobson, *Constructing trees in graphs whose complement has no* $K_{2,s}$, Combin. Probab. Comput., 11 (2002), p. 343–347, https://doi.org/10.1017/S0963548302005102.

[12] P. Erdös and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. (N.S.), 52 (1946), pp. 1087–1091.

[13] P. Erdös and T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Hungar., 10 (1959), pp. 337–356, https://doi.org/10.1007/BF02024498.

[14] G. Fan, *The Erdös–Sós conjecture for spiders of large size*, Discrete Math., 313 (2013), pp. 2513–2517, http://www.sciencedirect.com/science/article/pii/S0012365X13003324.

[15] G. Fan and L. Sun, *The Erdős–Sós conjecture for spiders*, Discrete Math., 307 (2007), pp. 3055–3062, http://www.sciencedirect.com/science/article/pii/S0012365X07001070.

[16] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, New York, 2009.

[17] A. Görlich and A. Zak, *On Erdős–Sós conjecture for trees of large size*, Electron. J. Combin., 23 (2016), P1.52.

[18] F. Havet, B. Reed, M. Stein, and D. R. Wood, *A Variant of the Erdős–Sós Conjecture*, preprint, arXiv:1606.09343, 2016.

[19] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi, *The approximate Loebl–Komlós–Sós Conjecture* I: *The sparse decomposition*, SIAM J. Discrete Math., 31 (2017), pp. 945–982.

[20] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi, *The approximate Loebl–Komlós–Sós Conjecture* II: *The rough structure of LKS graphs*, SIAM J. Discrete Math., 31 (2017), pp. 983–1016.

[21] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi, *The approximate Loebl–Komlós–Sós Conjecture* III: *The finer structure of LKS graphs*, SIAM J. Discrete Math., 31 (2017), pp. 1017–1071.

[22] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi, *The approximate Loebl–Komlós–Sós Conjecture* IV: *Embedding techniques and the proof of the main result*, SIAM J. Discrete Math., 31 (2017), pp. 1072–1148.

[23] J. Hladký and D. Piguet, *Loebl-Komlós-Sós Conjecture: Dense case*, J. Combin. Theory. Ser. B, 116 (2016), pp. 123–190.

[24] T. Klimošová, D. Piguet, and V. Rozhoň, *A Version of the Loebl-Komlós-Sós Conjecture for Skewed Trees*, preprint, arXiv:1802.00679, 2018.

[25] A. McLennan, *The Erdős-Sós conjecture for trees of diameter four*, J. Graph Theory, 49 (2005), pp. 291–301, https://doi.org/10.1002/jgt.20083.

[26] D. Piguet and M. J. Stein, *An approximate version of the Loebl-Komlós-Sós conjecture*, J. Combin. Theory. Ser. B, 102 (2012), pp. 102–125.

[27] V. Rozhoň, *Sufficient Conditions for Embedding Trees*, Bachelor's thesis, Czech Academy of Sciences, http://is.cuni.cz/webapps/zzp/search/, 2018.

[28] J.-F. Saclé and M. Woźniak, *The Erdős-Sós conjecture for graphs without* C_4, J. Combin. Theory. Ser. B, 70 (1997), pp. 367–372, https://doi.org/10.1006/jctb.1997.1758.

[29] E. Szemerédi, *Regular Partitions of Graphs*, in Problemes combinatoires et theorie des graphes, Colloq. Internat. CNRS 260, CNRS, Paris, 1978, pp. 399–401.

[30] G. Tiner, *On the Erdos-Sos conjecture for graphs on* $n = k + 3$ *vertices*, Ars Combin., 95 (2010).

[31] S. Wagner, *private communication*, 2018.

[32] M. Wang, G.-j. Li, and A.-d. Lui, *A result of Erdős-Sós conjecture*, Ars Combin., 55 (2000).

[33] M. Wozniak, *On the Erdős-Sós conjecture*, J. Graph Theory, 21 (1996), pp. 229–234, https://doi.org/10.1002/(SICI)1097-0118(199602)21:2⟨229::AID-JGT13⟩3.0.CO;2-E.

[34] Y. Zhao, *Proof of the* $(n/2 - n/2 - n/2)$ *conjecture for large n*, Electron. J. Combin., 18 (2011), P27.