Spanning Trees in Graphs of High Minimum Degree which have a Universal Vertex II: A Tight Result

Bruce Reed[*] Maya Stein[†]

May 23, 2019

Abstract

We prove that, if m is sufficiently large, every graph on $m+1$ vertices that has a universal vertex and minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ contains each tree T with m edges as a subgraph. Our result confirms, for large m, an important special case of a conjecture by Havet, Reed, Stein, and Wood.

The present paper builds on the results of a companion paper in which we proved the statement for all trees having a vertex that is adjacent to many leaves.

1 Introduction

This is the second in a series of two papers dedicated to a conjecture relating the minimum and the maximum degree of a graph to the occurence of certain trees as subgraphs. If we only condition on the minimum degree, it is easy to see that any graph of minimum degree at least m contains a copy of each tree with m edges, and that this bound is sharp. Possible strengthenings have

[*]School of Computer Science McGill University. Research supported by NSERC.

[†]Department of Mathematical Engineering and Centro de Modelamiento Matemático, Universidad de Chile, UMI 2807 CNRS (mstein@dim.uchile.cl). Research supported by CONICYT + PIA/Apoyo a centros científicos y tecnológicos de excelencia con financiamiento Basal, Código AFB170001, and by Fondecyt Regular Grant 1183080.

been conjectured in the form of the Erdős–Sós conjecture from 1963, which replaces the minimum degree with the average degree and whose proof for large graphs has been announced by Ajtai, Komlós, Simonovits and Szemerédi in the early 1990's, and in the form of the Loebl-Komlós-Sós conjecture from 1995, which replaces the minimum degree with the median degree, and which has been approximately solved in [HKP$^+$17a, HKP$^+$17b, HKP$^+$17c, HKP$^+$17d].

If instead, one sticks to conditioning on the minimum degree of the host graph but tries to weaken the imposed bound, it is still possible to embed bounded degree trees. Komlós, Sarközy and Szemerédi show in [KSS01] that for every $\delta > 0$, every large enough $(m+1)$-vertex graph of minimum degree at least $(1+\delta)\frac{m}{2}$ contains each tree with m edges whose maximum degree is bounded by $O(\frac{n}{\log n})$. An extension of this result to non-spanning trees is given in [BPS18].

It is clear, though, that only working with a condition on the minimum degree, but allowing it to be smaller than the size of the trees we are looking for, will never be enough to guarantee one can embed *all* trees, regardless of their maximum degree. So, it seems natural to seek an additional condition to impose on the host graph. The following conjecture in this respect has been put forward recently.

Conjecture 1.1 (Havet, Reed, Stein, and Wood [HRSW16])**.** *Let $m \in \mathbb{N}$. If a graph has maximum degree at least m and minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ then it contains every tree with m edges as a subgraph.*

This conjecture holds if the minimum degree condition is replaced by the much stronger bound $(1-\gamma)m$, for a tiny constant γ [HRSW16]. It also holds if the maximum degree condition is replaced by a large function in m [HRSW16]. Furthermore, an approximate version of the conjecture holds for bounded degree trees and dense host graphs [BPS18].

As further evidence for Conjecture 1.1, we prove that it holds when the graph has $m+1$ vertices, if m is large enough. That is, we show the conjeture for the case when we are looking for a spanning tree in a large graph.

Theorem 1.2. *There is an $m_0 \in \mathbb{N}$ such that for every $m \geq m_0$ every graph on $m+1$ vertices that has minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ and a universal vertex contains every tree T with m edges as a subgraph.*

The proof of Theorem 1.2 builds on results obtained in the companion paper [RS19a]. There, we showed the following lemma.

Lemma 1.3. **[RS19a, Lemma 1.3]** *For every $\delta > 0$, there is an m_δ such that for any $m \geq m_\delta$ the following holds for every graph G on $m+1$ vertices that has minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ and a universal vertex.*
If T is a tree with m edges, and some vertex of T is adjacent to at least δm leaves, then T embeds in G.

Lemma 1.3 covers the proof of our main result for all trees which have a vertex with many leaves, namely at least δm leaves, for some fixed δ, but is of no help for trees which have no such vertex. This latter case is covered by the next lemma which will be proved in the present paper.

Lemma 1.4. *There are $m_1 \in \mathbb{N}$, and $\delta > 0$ such that the following holds for every $m \geq m_1$, and every graph G on $m+1$ vertices that has minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ and a universal vertex.*
If T is a tree with m edges such that no vertex of T is adjacent to more than δm leaves, then T embeds in G.

With these two lemmas at hand, the proof of our main result, Theorem 1.2, is straightforward.

Proof of Theorem 1.2. We choose our output m_0 for Theorem 1.2 by taking the maximum value of m_1 and m_δ, where m_1 and δ are given by Lemma 1.4, and m_δ is given for input δ by Lemma 1.3. Given now T and G as in the theorem, Lemma 1.4 covers the case that T has no vertex adjacent to more than δm leaves, and Lemma 1.3 covers the remaining case. $\square$

So all that remains is to prove Lemma 1.4. The idea of the proof is to first reserve a random set $S \subseteq V(G)$ for later use. Then, we embed into $G - S$ a very small subtree T^* of the tree T we wish to embed. Actually, we will only embed $T^* - L$, having chosen a subset $L \subseteq V(T^*)$ of some low degree vertices (either leaves or vertices of degree 2). The vertices from L will be left out of the embedding for now, as they will only be embedded at the very end.

The set L is slightly larger than the set S. This gives us some free space when we embed $T - T^*$, which will be useful. In fact, this freedom makes it possible for us to use a lemma from [RS19a] (stated as Lemma 2.5 in the present paper) for embedding $T - T^*$, unless the graph G has a very special structure, in which case an ad-hoc embedding is provided by Lemma 2.6. After this, there is a small leftover set of vertices of G, which, together with the set S, serves for embedding the vertices from L, by using an absorption argument.

We formally organise the proof of Lemma 1.4 by splitting it up into four auxiliary lemmas, namely Lemma 2.2, Lemma 2.3, Lemma 2.5, and Lemma 2.6 (where Lemma 2.2 provides the subtree T^* from above, and Lemma 2.3 is responsible for absorbing the leftover vertices). The four lemmas will be stated in Section 2. That section also contains the proof of Lemma 1.4, under the assumption that Lemmas 2.2–2.6 hold, and the easy proof of Lemma 2.2. Lemma 2.5 was proved in [RS19a], so there are only two lemmas left we need to prove in the present paper. In the following two sections we state and prove two new lemmas, Lemma 3.1 and Lemma 4.1, which together imply Lemma 2.3. The last section of the paper is devoted to the proof of Lemma 2.6.

2 The Proof of Lemma 1.4

In the present section, we present our four auxiliary lemmas, Lemma 2.2, Lemma 2.3, Lemma 2.5, and Lemma 2.6, and then show how together, they imply Lemma 1.4.

We start with the simplest of our lemmas, Lemma 2.2, which will be proved at the end of the present section. This lemma enables us to find a convenient subtree T^* of a tree T. We need a quick definition before we give the lemma.

Definition 2.1 (γ-nice subtree)**.** *Let T be a tree with m edges. Call a subtree T^* of T with root t^* a γ-nice subtree if*

(i) $|V(T^*)| \leq \gamma m$;

(ii) every component of $T - T^*$ is adjacent to t^*,

and furthermore, one of the following conditions holds:

(1) T^* contains at least $\lceil \frac{\gamma m}{20} \rceil$ disjoint paths of length 5 and all vertices on these paths have degree at most 2 in T.

(2) T^* contains at least $\lceil \frac{\gamma m}{40} \rceil$ leaves from T.

If the former condition holds, we say T^* is of type 1, and if the latter condition holds, we say T^* is of type 2.

We are now ready to state the lemma that finds the γ-nice subtree.

Lemma 2.2. *For all $0 < \gamma \leq 1$, any tree with at least $\frac{200}{\gamma}$ edges has a γ-nice subtree.*

The proof of Lemma 2.2 is straightforward, but we prefer to leave it to the end of the present section, because our focus here is the proof of the main result.

Next, we exhibit a lemma that will enable us to transfer the embedding problem of the tree to an embedding problem of almost all of the tree, under the condition that we already embed a small part of it, namely a γ-nice subtree, beforehand.

For convenience, from now on let us call a graph *m-good* if it has $m+1$ vertices, minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ and a universal vertex.

Lemma 2.3. *There is an $m_0 \in \mathbb{N}$ such that the following holds for all $m \geq m_0$, and all γ with $\frac{2}{10^7} \leq \gamma < \frac{1}{30}$.*

Let G be an m-good graph, with universal vertex w. Let T be a tree with m edges, such that no vertex of T is adjacent to more than $\frac{m}{10^{23}}$ leaves. Let T^* be a γ-nice subtree of T, rooted at vertex t^*.
Then there are sets $L \subseteq V(T^*) \setminus \{t^*\}$ and $S \subseteq V(G)$ satisfying

$$|S| \leq |L| - \lceil (\frac{\gamma}{2})^4 m \rceil.$$

Furthermore, for any $w' \in V(G) - S$, with $w' \neq w$, there is an embedding of $T^* - L$ into $G - S$, with t^* embedded in w', such that the following holds. Any embedding of $T - L$ into $G - S$ extending our embedding of $T^* - L$ can be extended to an embedding of all of T into G.

Below, we shall split Lemma 2.3 into two lemmas, Lemma 3.1 and Lemma 4.1, depending on the type of the γ-nice subtree. We will state and prove Lemma 3.1 in Section 3, and state and prove Lemma 4.1 in Section 4.

In order to state the remaining two of our four auxiliary lemmas, we need a simple definition. This definition describes the extremal case, where the graph G has a very specific structure (and therefore, the approach from the companion paper [RS19a] does not work).

Definition 2.4. *Let $\gamma > 0$. We say a graph G on $m+1$ vertices is γ-special if $V(G)$ consists of three mutually disjoint sets X_1, X_2, X_3 such that*

- $\frac{m}{3} - 3\gamma m \leq |X_i| \leq \frac{m}{3} + 3\gamma m$ *for each* $i = 1, 2, 3$*; and*
- *there are at most* $\gamma^{10}|X_1| \cdot |X_2|$ *edges between* X_1 *and* X_2*.*

The following lemma, which excludes the extremal situation, was proved[1] in the companion paper [RS19a].

Lemma 2.5. [RS19a, Lemma 7.3] *For all* $\gamma < \frac{1}{10^6}$ *there are* $m_0 \in \mathbb{N}$ *and* $\lambda > 0$ *such that the following holds for all* $m \geq m_0$*.*
Let G *be an* m*-good graph, which is not* γ*-special. Let* T *be a tree with* m *edges such that no vertex in* T *is adjacent to more than* λm *leaves. Let* T^* *be a* γ*-nice subtree of* T*, with root* t^**, let* $L \subseteq V(T^*) \setminus \{t^*\}$*, and let* $S \subseteq V(G)$ *such that* $|S| \leq |L| - \lceil (\frac{\gamma}{2})^4 m \rceil$*.*
If there is an embedding of $T^* - L$ *into* $G - S$*, then there is an embedding of* $T - L$ *into* $G - S$ *extending the embedding of* $T^* - L$*.*

Our last auxiliary lemma deals with the extremal case as described in Definition 2.4.

Lemma 2.6. *There are* $m_0 \in \mathbb{N}$*,* $\beta \leq \frac{1}{10^{10}}$*, and* $\gamma_0, \gamma_1 \leq \frac{1}{50}$ *such that the following holds for all* $m \geq m_0$*.*
Suppose G *is a* γ_0*-special* $(m+1)$*-vertex graph of minimum degree at least* $\lfloor \frac{2}{3} m \rfloor$*, and suppose* T *is a tree with* m *edges such that none of its vertices is adjacent to more than* βm *leaves. Let* T^* *be a* γ_1*-nice subtree of* T*, with root* t^**, and let* $L \subseteq V(T^*) \setminus \{t^*\}$*. Assume there is a set* $S \subseteq V(G)$ *such that* $|S| \leq |L| - \lceil (\frac{\gamma_1}{2})^4 m \rceil$*.*
If, for any $W \subseteq V(G) - S$ *with* $|W| \geq \gamma_0 m$*, there is an embedding of* $T^* - L$ *into* $G - S$*, with* t^* *embedded in* W*, then there is an embedding of* $T - L$ *into* $G - S$ *extending the given embedding of* $T^* - L$*.*

We prove Lemma 2.6 in Section 5.

Let us close the section by showing how our four auxiliary lemmas imply Lemma 1.4.

Proof of Lemma 1.4. First, we apply Lemma 2.6 to obtain four numbers $\beta, \gamma_0, \gamma_1 > 0$ and $m_0^{Lem\ 2.6} \in \mathbb{N}$. Next, we apply Lemma 2.3 to obtain a

[1] We remark that for simplicity, we used a slightly weaker definition of a γ-nice tree in [RS19a], namely we did not require one of the conditions (1) and (2) to hold. Clearly, the lemma still holds with the stronger definition given here.

number $m_0^{Lem\ 2.3}$. Finally, we apply Lemma 2.5 with input γ_0 to obtain another integer $m_0^{Lem\ 2.5}$ as well as a number $\lambda > 0$.

For the output of Lemma 1.4, we will take

$$m_1 := \max\{m_0^{Lem\ 2.6}, m_0^{Lem\ 2.3}, m_0^{Lem\ 2.5}, \frac{200}{\gamma_0}\},$$

and

$$\delta := \min\{\beta, \lambda, 10^{-23}\}.$$

Now, consider an m-good graph G, and a tree T with m edges as in the statement of Lemma 1.4. Use Lemma 2.2 together with Lemma 2.3, once for each input γ_0, γ_1, to obtain, for $i = 0, 1$, a γ_i-nice tree T_i^* with root t_i^*, and sets S_i, L_i satifying

$$|S_i| \ \leq \ |L_i| - (\frac{\gamma_i}{2})^4 m.$$

Moreover, for $i = 0, 1$, there are embeddings of $T_i^* - L_i$ into $G - S_i$ that map the vertex t_i^* to any given vertex, except possibly the universal vertex of G. Furthermore, Lemma 2.3 guarantees that, in order to embed T into G, we only need to extend, for either $i = 0$ or $i = 1$, the embedding of $T_i^* - L_i$ given by the lemma to an embedding of all of $T - L_i$ into $G - S$.

For this, we will use Lemmas 2.5 and 2.6. More precisely, if G is not γ_0-special, then we can apply Lemma 2.5 to G with sets S_0 and L_0, together with the tree T_0^*. If G is γ_0-special, we can apply Lemma 2.6 to G with sets S_1 and L_1, together with the tree T_1^*. This finishes the proof of the lemma. □

We now give the short proof of Lemma 2.2.

Proof of Lemma 2.2. As an auxiliary measure, we momentarily fix any leaf v_L of the given tree T as the root of T. Next, we choose a vertex t^* in T having at least $\lfloor \frac{\gamma m}{2} \rfloor$ descendants, such that it is furthest from v_L having this property.

Then, each component of $T - t^*$ that does not contain v_L has size at most $\lfloor \frac{\gamma m}{2} \rfloor$. So, there is a subset S^* of these components such that

$$\left\lceil \frac{\gamma m}{2} \right\rceil \ \leq \ \sum_{S \in S^*} |S| \ \leq \ \gamma m.$$

Now, consider the tree T^* formed by the union of the trees in S^* and the vertex t^*. The tree T^* clearly fulfills items (a) and (b) of Definition 2.1. If T^*

contains at least $\lceil \frac{\gamma m}{40} \rceil$ leaves of T, then T^* is γ-nice of type 2, and we are done.

Otherwise, delete from T^* all of the at most $\lfloor \frac{\gamma m}{40} \rfloor$ vertices whose degree is greater than 2. It is easy to see that this leaves us with a set of at most $\frac{\gamma m}{20}$ paths, together containing at least $\frac{19}{40}\gamma m$ vertices. All vertices of these paths have degree at most 2 in T. Deleting at most four vertices on each path we can ensure all paths have lengths divisible by five, and together contain at least $\frac{19}{40}\gamma m - 4 \cdot \frac{\gamma m}{20} \geq \frac{\gamma m}{4} + 5$ vertices. Dividing each of the paths into paths of length five we obtain a set $\mathcal{P}$ of at least $\lceil \frac{\gamma m}{20} \rceil$ disjoint paths in T^*. So, T^* is γ-nice of type 1.

□

3 The Proof of Lemma 3.1

This section is devoted to the proof of the folloing lemma, which proves Lemma 2.3 for all γ-nice trees of type 1.

Lemma 3.1. *There is an $m_0 \in \mathbb{N}$ such that the following holds for all $m \geq m_0$, and for all $\gamma > 0$ with $\frac{2}{10^7} \leq \gamma < \frac{1}{30}$.*
Let G be an m-good graph. Let T be a tree with m edges, such that no vertex of T is adjacent to more than $\frac{m}{10^{23}}$ leaves. Let T have a γ-nice subtree T^* of type 1, with root t^*.
Then there are sets $L \subseteq V(T^*) \setminus \{t^*\}$ and $S \subseteq V(G)$ satisfying $|S| \leq |L| - (\frac{\gamma}{2})^4 m$. Furthermore, for any $w \in V(G) - S$, there is an embedding of $T^* - L$ into $G - S$, with t^* embedded in w, such that any embedding of $T - L$ into $G - S$ extending our embedding of $T^* - L$ can be extended to an embedding of all of T into G.

In the proof of Lemma 3.1, some random choices are going to be made, and in order to see we are not far from the expected outcome, it will be useful to have the well-known Chernoff bounds at hand (see for instance [McD89]). For the reader's convenience let us state these bounds here.

Let $X_1, \dots, X_n$ be independent random variables satisfying $0 \leq X_i \leq 1$. Let $X = X_1 + \dots + X_n$ and set $\mu := \mathbb{E}[X]$. Then for any $\varepsilon \in (0, 1)$, it holds that

$$\mathbb{P}[X \geq (1+\varepsilon)\mu] \leq e^{-\frac{\varepsilon^2}{2+\varepsilon}\mu} \quad \text{and} \quad \mathbb{P}[X \leq (1-\varepsilon)\mu] \leq e^{-\frac{\varepsilon^2}{2}\mu}. \tag{1}$$

We are now ready for the proof of Lemma 3.1.

Proof of Lemma 3.1. We choose $m_0 = 10^{25}$. Now assume that for some $m \geq m_0$, we are given an m-good graph G, and a tree T with m edges such that none of its vertices is adjacent to more than $10^{-23}m$ leaves. We are also given a γ-nice subtree T^* of T, with root t^*, and a set $\mathcal{P}$ of disjoint paths of length five such that

$$|\mathcal{P}| = \lceil \frac{\gamma m}{20} \rceil,$$

for some γ as in the lemma.

We now define L as the set that consists of the fourth vertex (counting from the vertex closest to t^*) of each of the paths from $\mathcal{P}$. Clearly,

$$|L| = \lceil \frac{\gamma m}{20} \rceil \geq \lceil \frac{m}{10^8} \rceil, \tag{2}$$

by our assumptions on γ.

In order to prove Lemma 3.1, we need to do three things. First of all, we need to find a set $S \subseteq V(G)$ of size at most $|L| - (\frac{\gamma}{2})^4 m$. Then, given any vertex $w \in V(G) - S$, we have to embed $T - L$ into $G - S$, with t^* going to w. Finally, we need to make sure that any extension of this embedding to an embedding of all of $T - L$ into $G - S$ can be completed to an embedding of all of T.

It is clear that for the last point to go through, it will be crucial to have chosen both S and the set N of the images of the neighbours of the vertices in L carefully, in order to have the necessary connections between N and S. Our solution is to choose both S and N randomly. More precisely, choose a set S of size

$$|S| = |L| - \lceil (\frac{\gamma}{2})^4 m \rceil \tag{3}$$

uniformly and independently at random in $V(G - w)$. Also, choose a set N of size

$$|N| = 2|L| \tag{4}$$

uniformly and independently at random in $V(G - w - S)$.

Now, we can proceed to embed $T' := T - L$ into $G - S$. We will start by embedding the neighbours of vertices in L arbitrarily into N. Let us keep track of these by calling $n_1(x)$ and $n_2(x)$ the images of the neighbours of x, for each $x \in L$.

Next, we embed t^* into w, and then proceed greedily, using a breadth-first order on T^* (skipping the vertices of L and those already embedded into N). Each vertex we embed has at most two neighbours that have been embedded earlier (usually this is just the parent, but parents of vertices embedded into N have two such neighbours, and the root of T' has none). So, since G has minimum degree at least $\lfloor \frac{2m}{3} \rfloor$ and given the small size of T', we can easily embed all of T' as planned.

It remains to prove that any extensions of this embedding can be completed to an embedding of all of T. This will be achieved by the following claim, which finishes the proof of Lemma 3.1.

Claim 3.2. *For any set R of $|L|-|S|$ vertices, there is a bijection between L and $S \cup R$ mapping each vertex $x \in L$ to a common neighbour of $n_1(x)$ and $n_2(x)$.*

In order to prove Claim 3.2, we define an auxiliary bipartite graph H having $V(G-w)$ on one side, and L on the other. We put an edge between $v \in V(G-w)$ and $x \in L$ if v is adjacent to both $n_1(x)$ and $n_2(x)$. We are interested in the subgraph H' of H that is obtained by restricting the $V(G-w)$-side to the set $S \cup R$ (but sometimes it is enough to consider degrees in H).

By the minimum degree condition on G, the expectation of the degree in H of any vertex $v \in V(G-w)$ is

$$\mathbb{E}(deg_H(v)) \geq (\frac{199}{300})^2|L|,$$

since v has at least $\lfloor \frac{2}{3}m - 1 \rfloor \geq \frac{199}{300}m$ neighbours in $G-w$, and thus, for any given $x \in L$, each $n_i(x)$ is adjacent to v with probability at least $\frac{199}{300}m$. Therefore, the probability that all vertices of G have degree at least

$$d := (\frac{198}{300})^2|L|$$

is bounded from below by

$$\begin{aligned} \mathbb{P}[\delta(G) \geq d] &\geq 1 - \sum_{v \in V(G-w)} \mathbb{P}[deg_H(v) < d] \\ &\geq 1-(m+1) \cdot e^{-(\frac{397}{199 \cdot 300})^2 \frac{|L|}{2}} \\ &\geq 0.9999, \end{aligned}$$

where we used (1) (Chernoff's bound) with $\varepsilon = \frac{199^2-198^2}{199^2} = \frac{397}{199^2}$, our bound on the size of L as given in (2) and the fact that $m \geq 10^{25}$.

Furthermore, since G has minimum degree at least $\lfloor \frac{2}{3}m \rfloor$, we know that for each $x \in L$, vertices $n_1(x)$ and $n_2(x)$ have at least $\frac{1}{3}m - 3$ common neighbours in $G - w$. Therefore, every vertex of L has degree at least $\frac{1}{3}m - 3$ in H. However, we are interested in the degree of these vertices into the set S. For a bound on this degree, first note that the expected degree of any vertex of L into the set S is bounded from below by $\frac{999}{3000}|S|$. Now again apply (1) (Chernoff's bound), together with the fact that $|S| \geq 10^{17}$, to obtain that with probability greater than 0.9999, every element of L is incident to at least $\frac{998}{3000}|S|$ vertices of S.

Resumingly, we can say that with probability greater than 0.999 we chose the sets S and N such that the resulting graph H obeys the following degree conditions:

(A) the minimum degree of $V(G - w)$ into L is at least $(\frac{198}{300})^2|L|$; and

(B) the minimum degree of L into S is at least $\frac{998}{3000}|S|$.

Let us from now on assume that we are in the likely situation that both (A) and (B) hold.

Further, assume there is no matching from $S \cup R$ to L in H'. Then by Hall's theorem, there is a partition of L into sets L' and L'' and a partition of $S \cup R$ into sets J' and J'' such that there are no edges from L' to J'', and such that

$$|J'| < |L'| \quad \text{and} \quad |L''| < |J''|.$$

Since $J'' \neq \emptyset$, and since by (A), each vertex in J'' has degree at least $(\frac{198}{300})^2|L|$ into L, and thus into L'', we deduce that

$$|J''| > |L''| \geq (\frac{198}{300})^2|L|. \tag{5}$$

Since also $L' \neq \emptyset$, and by (B), each of its elements has at least $\frac{998}{3000}|S|$ neighbours in $S \cap J'$, we see that

$$|L'| > |J'| \geq \frac{998}{3000}|S|.$$

Thus, using (2) and (3), as well as our upper bound on γ, we can calculate that

$$|L''| \;=\; |L| - |L'| \;\leq\; |S| + \lceil (\frac{\gamma}{2})^4 m \rceil - \frac{998}{3000}|S| \;\leq\; \frac{2003}{3000}|S|. \tag{6}$$

Let us iteratively define a subset S^* of $S \cap J''$ as follows. We start by putting an arbitrary vertex $v_0 \in S \cap J''$ into S^*, and while there is a vertex of $S \cap J''$ whose neighbourhood contains $\frac{m}{1000 \log m}$ vertices which are not in the neighbourhood of S^*, we augment S^* by adding any such vertex v that maximises $N(v) - N(S^*)$. We stop when there is no suitable vertex that can be added to S^*. Note that $|S^*| \leq 1000 \log m$.

Our plan is to show next that the set S^* has certain properties which are unlikely to be had by *any* set having certain other properties that S^* has (for instance, having size at most $1000 \log m$). More precisely, the probability that a set like S^* exists will be bounded from above by 0.005. This will finish the proof of Claim 3.2, as we then know that with probability at least 0.99 we chose sets S and N such that in the resulting graph H', the desired matching exists, and thus Claim 3.2 holds.

So, let us define $\mathcal{Q}$ as the set of all subsets of $V(G-w)$ having size at most $1000 \log m$. For each $Q \in \mathcal{Q}$, let $V_1(Q)$ be the set consisting of all vertices of $G - w$ which have less than $\frac{m}{1000 \log m}$ neighbours outside $N(Q)$ (in the graph $G - w$).

Finally, let $\mathcal{Q}' \subseteq \mathcal{Q}$ contain all $Q \in \mathcal{Q}$ for which

$$\frac{m}{10^9} \;\leq\; |V_1(Q)| \;\leq\; \frac{m}{3} + \frac{m}{\log m} + 2. \tag{7}$$

Observe that, for $Q \in \mathcal{Q}'$ fixed, the expected size of $V_1(Q) \cap S$ is

$$\mathbb{E}[V_1(Q) \cap S] = |V_1(Q)| \cdot \frac{|S|}{m}$$

because S was chosen at random in $G - v$. So by (3) and (2), and by (7), we see that

$$\frac{1}{2} \cdot \frac{m}{10^{17}} \;\leq\; \mathbb{E}[V_1(Q) \cap S] \;\leq\; \frac{|S|}{3} + \frac{|S|}{\log m} + 2 \;\leq\; \frac{38}{100}|S|, \tag{8}$$

where the last inequality follows from the fact that $m \geq 10^{25}$. Now, we can use (1) (Chernoff's bound) and the first inequality of (8) to bound the probability that $|V_1(Q) \cap S|$ exceeds its expectation by a factor of at least $\frac{20}{19}$ as follows:

$$\mathbb{P}\Big[|V_1(Q) \cap S| \geq \frac{20}{19} \cdot \mathbb{E}[V_1(Q) \cap S]\Big] \;\leq\; e^{-\frac{\mathbb{E}[V_1(Q)\cap S]}{820}} \;\leq\; e^{-\frac{m}{164 \cdot 10^{18}}} \;\leq\; \frac{0.001}{m^{\log m}}.$$

Since by (8), we know that

$$\frac{20}{19} \cdot \mathbb{E}[V_1(Q) \cap S] \; < \; \frac{41}{100}|S|,$$

and since $|Q| \leq m^{\log m}$ for each $Q \in \mathcal{Q}$, we can deduce that

$$\mathbb{P}\Big[\; \exists Q \in \mathcal{Q}' \text{ with } |V_1(Q) \cap S| \geq \frac{41}{100}|S| \;\Big] \; \leq \; 0.001. \tag{9}$$

Now, let us turn back to the set S^*. First of all, we note that by the definition of S^*, we have $S \cap J'' \subseteq V_1(S^*)$. Thus, we can use (5) and (3) to deduce that

$$\begin{aligned} |V_1(S^*) \cap S| \; &\geq \; |J''| - |R| \\ &\geq \; (\frac{198}{300})^2|L| - \lceil (\frac{\gamma}{2})^4 m \rceil \\ &\geq \; (\frac{197}{300})^2|S| \\ &> \; \frac{43}{100}|S|. \end{aligned} \tag{10}$$

So, by (2) and (3), the first inequality of (7) holds for $Q = S^*$.

For a moment, assume that $N(S^*) \leq \frac{999}{1000}m$. Then, also the second inequality of (7) holds for $Q = S^*$, as otherwise, each of the at least $\frac{m}{1000}$ vertices of $V(G-w) \setminus N(S^*)$ sees at least $\frac{m}{\log m}$ vertices of $V_1(S^*)$, and so, by the definition of S^*, we have that

$$\begin{aligned} \frac{m}{1000} \cdot \frac{m}{\log m} \quad &\leq e(V_1(S^*), V(G-w) \setminus N(S^*)) \\ &< \frac{m}{1000 \log m} \cdot |V_1(S^*)| \\ &\leq \frac{m^2}{1000 \log m}, \end{aligned}$$

a contradiction. Hence $S^* \in \mathcal{Q}'$. But then, according to (9), we know that (10) is not likely to happen. So, with probability at least 0.998, we chose S in a way that all three of (A), (B), and

(C) $|N(S^*)| \geq \frac{999}{1000}m$

hold. We will from now on assume that we are in this likely case.

Consider the set $\mathcal{Q}''$ which consists of all sets $Q \in \mathcal{Q}$ for which the first inequality in (7) holds, and for which $|N(Q)| \geq \frac{999}{1000}m$. By (10) and by (C), $S^* \in \mathcal{Q}''$.

Call $\mathcal{Q}''_+$ the set of all $Q \in \mathcal{Q}''$ for which at least one of the following holds:

- Q has a vertex of degree at least $\frac{2m}{3} + \frac{m}{100}$; or
- Q has two vertices v, v' such that each sees at least $\frac{m}{100}$ vertices outside the neighbourhood of the other one.

We are going to show that the sets $Q \in \mathcal{Q}''_+$ typically have larger neighbourhoods in L than S^* has, and will thus be able to conclude that $S^* \notin \mathcal{Q}''_+$, which will be crucial for the very last part of the proof.

For this, let $X(Q)$ be the set of unordered pairs $\{v, v'\}$ of distinct vertices which have a common neighbour in Q, for each $Q \in \mathcal{Q}''$. Then, because of the minimum degree condition we imposed on the graph G, we know that each vertex $v \in N(Q)$ is in at least $\lfloor \frac{2m}{3} \rfloor - 2$ pairs of $X(Q)$. So, since N was chosen at random in $V(G-w)$, and because of the definition of $\mathcal{Q}''$, we know that for any fixed set $Q \in \mathcal{Q}''$, and any fixed vertex $x \in L$, the probability that $n_1(x)$ and $n_2(x)$ have a common neighbour in Q can be bounded as follows:

$$\mathbb{P}[\{n_1(x), n_2(x)\} \in X(Q)] \geq \frac{\frac{999m}{1000} \cdot (\lfloor \frac{2m}{3} \rfloor - 2)}{m^2}.$$

However, if we take any fixed $Q \in \mathcal{Q}''_+$, and any fixed $x \in L$, the bound becomes

$$\begin{aligned}\mathbb{P}[\{n_1(x), n_2(x)\} \in X(Q)] &\geq \frac{\frac{999m}{1000}(\lfloor \frac{2m}{3} \rfloor - 2) + \min\{(\frac{2m}{3} + \frac{m}{100})\frac{m}{100}, (\frac{m}{3} - 2)\frac{m}{100}\}}{m^2} \\ &\geq \frac{669}{1000},\end{aligned}$$

where the two entries in the minimum stand for the two scenarios that may cause the set Q to belong to $\mathcal{Q}''_+$. In order to to see the term for the second scenario, observe that vertices v and v' have at least $\frac{m}{3} - 2$ common neighbours, and each of these neighbours belongs to at least $\lfloor \frac{2m}{3} \rfloor - 2 + \frac{m}{100}$ pairs of $X(Q)$.

Therefore, fixing $Q \in \mathcal{Q}''_+$, and letting $L(Q)$ denote the sets of all $x \in L$ with $\{n_1(x), n_2(x)\} \in X(Q)$, we know that the expected size of $L(Q)$ is bounded by

$$\mathbb{E}\Big[|L(Q)|\Big] \geq \frac{669}{1000}|L|.$$

As above, we can apply the Chernoff bound (1) to see that with very high probability, $|L(Q)|$ is not much smaller than its expectation:

$$\mathbb{P}\Big[|L(Q)| \leq \frac{668}{669} \cdot \mathbb{E}[|L(Q)|]\Big] \;\leq\; e^{-\frac{\mathbb{E}[|L(Q)|]}{2 \cdot 669^2}} \;\leq\; e^{-\frac{|L|}{2 \cdot 10^6}} \;\leq\; e^{-\frac{m}{2 \cdot 10^{14}}} \;\leq\; \frac{0.001}{m \log m},$$

where we use (2) and the fact that $m \geq 10^{25}$. So with probability at least 0.997, we have chosen N in a way that (A), (B), (C), and also

(D) $|L(Q)| \;>\; \frac{668}{1000}|L| \;=\; \frac{2004}{3000}|L|$ for every $Q \in \mathcal{Q}''_+$

hold.

Because of (6) (and (3)), and since $L'' \supseteq L(S^*)$, this means that

$$S^* \notin \mathcal{Q}''_+.$$

In particular, the degree of v_0 (in $G - w$) is less than $\frac{2m}{3} + \frac{m}{100}$, and each vertex of S^* has less than $\frac{m}{100}$ neighbours outside $N(v_0)$. Moreover, by the choice of S^*, we can deduce that

$$\text{every vertex in } S \cap J'' \text{ has less than } \frac{m}{100} \text{ neighbours outside } N(v_0). \tag{11}$$

By (3) and by (6), and since $|R| = |L| - |S|$, we know that

$$|S \cap J''| \;\geq\; |J''| - |R| \;\geq\; (\frac{198}{300})^2|L| - \lceil (\frac{\gamma}{2})^4 m \rceil \;>\; \frac{2}{5}|S|. \tag{12}$$

Fix a subset Z of size $\frac{m}{4}$ of $G - w - N(v_0)$, and let us look at the average degree d of the vertices of Z into $S \cap J''$. We have

$$d \cdot \frac{m}{4} = \sum_{v \in Z} deg(v, S \cap J'') = \sum_{v \in S \cap J''} deg(v, Z) \leq \frac{m \cdot |S \cap J''|}{100},$$

where for the last inequality we used (11). Thus

$$d \le \frac{|S \cap J''|}{25}.$$

Now use (12) to see that the average degree of the vertices of Z into S is bounded from above by $|S| - \frac{48}{125}|S| < (\frac{2}{3} - \frac{3}{100})|S|$. This means that there must be at least one vertex in Z, say the vertex z, which has degree at most $(\frac{2}{3} - \frac{3}{100})|S|$ into S. However, by Chernoff's bound (1), and since the expected degree of any vertex of $G - W$ into S is at least $(\frac{2}{3} - \frac{1}{1000})|S|$, we know that this would only happen with probability at most 0.001. So we can assume we are in a situation where no such vertex z exists, and reach a contradiction, as desired.

Resumingly, we know that with probability at least 0.995, our choice of S and N guarantee that a set S^* as above does not exist in the resulting auxiliary graph H', and thus, Hall's condition holds in H'. This means we find the desired matching, which finishes the proof of Claim 3.2, and with it the proof of Lemma 3.1.

□

4 The Proof of Lemma 4.1

This section is devoted to the proof of the following lemma, which proves Lemma 2.3 for all γ-nice trees of type 2. (So, since the other type of γ-nice trees are covered by Lemma 3.1, this finishes the proof of Lemma 2.3.)

Lemma 4.1. *There is an $m_0 \in \mathbb{N}$ such that the following holds for all $m \ge m_0$, and all $\gamma > 0$ with $\frac{2}{10^7} \le \gamma < \frac{1}{30}$.*
Let G be an m-good graph, with universal vertex w. Let T be a tree with m edges, such that no vertex of T is adjacent to more than $\frac{m}{10^{23}}$ leaves. Let T have a γ-nice subtree T^* of type 2, with root t^*.
Then there are sets $L \subseteq V(T^*) \setminus \{t^*\}$ and $S \subseteq V(G)$ satisfying $|S| \le |L| - (\frac{\gamma}{2})^4 m$. Furthermore, for any $w' \in V(G) - (S \cup \{w\})$, there is an embedding of $T^* - L$ into $G - S$, with t^* embedded in w', such that any embedding of $T - L$ into $G - S$ extending our embedding of $T^* - L$ can be extended to an embedding of all of T into G.

In the proof of Lemma 4.1 we will use Azuma's inequality which can be found for instance in [McD89]). This well-known inequality states that for any sub-martingale $\{X_0, X_1, X_2, \ldots\}$ which for each k almost surely satisfies $|X_k - X_{k-1}| < c_k$ for some c_k, we have that

$$\mathbb{P}[X_n - X_0 \leq -t] \;\leq\; e^{-\frac{t^2}{2\cdot\sum_{k=1}^{n} c_k^2}} \tag{13}$$

for all $n \in \mathbb{N}_+$ and all positive t.

Let us now give the proof of Lemma 4.1.

Proof of Lemma 3.1. We choose $m_0 \in \mathbb{N}$ large enough so that certain inequalities below are satisfied.

Let G be an m-good graph, with universal vertex w. Let T be a tree with m edges, such that no vertex of T is adjacent to more than $\frac{m}{10^{23}}$ leaves. We are also given a γ-nice subtree T^* of T, with root t^*, and since T^* is of type 2, there is a set $L \subseteq V(T^*) \setminus \{t^*\}$ of

$$|L| = \lceil \frac{\gamma m}{40} \rceil \geq \lceil \frac{m}{10^9} \rceil$$

leaves of T.

In order to prove Lemma 4.1, it suffices to find a set $S \subseteq V(G)$ satisfying $|S| \leq |L| - (\frac{\gamma}{2})^4 m$, to embed $T^* - L$ into $G - S$, and show that any extension of this embedding to an embedding of $T - L$ into $G - S$ can be completed to an embedding of all of T into G.

For this, let us define t as the vertex of T^* that is adjacent to most leaves from L. Define α so that t is incident to $\lceil \alpha m \rceil$ leaves and call L_t the set of these leaves. By the assumptions of the lemma,

$$\alpha \leq 10^{-23}. \tag{14}$$

We now randomly embed $T^* - L$ in a top down fashion, where we start by putting t^* in to w'. At each moment, when we embed a vertex $v \neq t$, we choose a uniformly random neighbour of the image of the (already embedded) parent $p(v)$ of v. When we reach t, we embed t into w, the universal vertex of G. (This gives us some leeway when we later have to embed L.) We do not have to worry about the connection of w to the image of $p(t)$ because of the universality of w.

For every $x \in L$, let us call $n(x)$ the image of $p(x)$.

Next, we pick a set S of size

$$|S| = |L| - \lceil (\frac{\gamma}{2})^4 m \rceil$$

uniformly and independently at random in what remains of G. It only remains to prove the following analogue of Claim 3.2 to finish the proof of Lemma 4.1.

Claim 4.2. *For any set R of $|L|-|S|$ vertices, there is a bijection between L and $S \cup R$ mapping each vertex $x \in L$ to a neighbour of $n(x)$.*

In order to prove Claim 4.2, consider a set R of size $|L| - |S|$ such that there is no matching from L to $S \cup R$ in the auxiliary bipartite graph H which is defined as follows. The bipartition classes of this graph H are L and $S \cup R$, and every vertex $x \in L$ is joined to all unoccupied neighbours of the image $n(x)$ of the parent of x in $S \cup R$. Our aim is to derive a contradiction from the assumption that such a set R exists.

Our first observation is that by Chernoff's bound (1) and by our assumption on the minimum degree of G, we know that with probability at least 0.999, every vertex of L has degree at least $(\frac{2}{3} - \frac{2}{10^4})|L|$ in H.

Furthermore, as there is no matching from L to $S \cup R$ in H, we can apply Hall's theorem. This gives a partition of L into sets L' and L'' and a partition of $S \cup R$ into sets J' and J'' such that there are no edges from L' to J'', and such that furthermore,

$$|J'| < |L'| \text{ and } |L''| < |J''|.$$

As $L' \neq \emptyset$, we know that $|J'| \geq (\frac{2}{3} - \frac{2}{10^4})|L|$ and therefore,

$$|J''| \leq (\frac{1}{3} + \frac{2}{10^4})|L|. \tag{15}$$

Since L'' contains all the children of t (this follows from the definition of H and from the fact that $|J'| < m$), and because of the definition of α, we know that L'' has size at least $\lceil \alpha m \rceil$ and hence

$$|J''| > \lceil \alpha m \rceil. \tag{16}$$

We now consider the set V^* of vertices of G which are adjacent to at most $(\frac{1}{3} + \frac{2}{10^4})|L|$ vertices of L in H. (The vertices in V^* are those that serve only for relatively few leaves in L as a possible image.) Note that the size of V^* depends on how we embedded $T^* - L$ (which was done randomly). We plan to show that

$$\text{with probability} \geq 0.99, \text{ we embedded } T^* - L \text{ such that } |V^*| < \alpha m. \tag{17}$$

Then, by (16) there is a vertex $v \in J'' \setminus V^*$. As the neighbours of v in H are contained in L'', we get that

$$|J''| > |L''| \geq (\frac{1}{3} + \frac{2}{10^4})|L|,$$

which is a contradiction to (15). This would prove Claim 4.2.

So, it only remains to show (17). For this, we start by bounding the probability that a specific vertex v is in V^*. Consider any vertex p that is the parent of some subset L_p of L, and recall that p was embedded randomly in the neighbourhood N_p of the image of the parent of p. By our minimum degree condition on G, we know that v is incident to at least $\frac{499}{1000}|N_p|$ vertices of N_p.

Hence, the probability that v is adjacent to p in G, and thus to all of L_p in H, is bounded from below by $\frac{499}{1000}$. Since $T^* - L$ is very small, this bound actually holds independently of whether v is adjacent to $L_{p'}$ for some other parent p'. Therefore,

$$\text{the expected degree of } v \text{ into } L_p \text{ is at least } \frac{499}{1000}|L_p|, \tag{18}$$

for each p.

Our plan is to use Azuma's inequality (i.e., inequality (13) above). For this, order the set P of parents p of subsets L_p of L as above, writing

$$P = \{p_1, p_2, \dots, p_n\}.$$

For $1 \leq i \leq n$, write d_i for the degree of v into L_{p_i}. Now, define the random variable

$$X_k := \sum_{1 \leq i \leq k} d_i + \frac{499}{1000} \cdot \sum_{k < i \leq n} |L_{p_i}|.$$

By (18), this is a sub-martingale. Observe that

$$X_0 = \frac{499}{1000} \cdot |L|$$

and

$$X_n = \deg(v, L).$$

We set $c_k := |L_{p_k}|$ for all $k \leq n$. Then $\sum_{k=1}^n c_k = |L|$, and furthermore, by our choice of the vertex t in the beginning of the proof of Lemma 4.1, we know that

$$c_k \leq \alpha m, \text{ for all } k \leq n.$$

This, together with Azuma's inequality (13), tells us that the probability that v is in V^* can be bounded as follows:

$$\begin{aligned}\mathbb{P}[v \in V^*] &\leq \mathbb{P}\Big[\deg(v, L) \leq \frac{336}{1000}|L|\Big] \\ &= \mathbb{P}[X_n - X_0 \leq -\frac{163}{1000}|L|] \\ &\leq e^{-\frac{(\frac{163}{1000}|L|)^2}{2\alpha m \cdot \sum_{k=1}^{n} c_k}} \\ &\leq e^{-\frac{163^2}{2\alpha \cdot 10^{15}}} \\ &\leq e^{-\frac{1}{10^{11}\cdot\alpha}}.\end{aligned}$$

So, the expected size of V^* is at most $m \cdot e^{-\frac{1}{10^{11}\cdot\alpha}}$. Using Markov's inequality we see that the probability that V^* contains more than αm vertices is bounded from above by

$$\frac{e^{-\frac{1}{10^{11}\cdot\alpha}}}{\alpha} \leq 0.01,$$

where we used the fact that $\alpha \leq 10^{-23}$ by (14). This proves (17), and thus finishes the proof of Claim 4.2, and of Lemma 4.1.

□

5 The Proof of Lemma 2.6

The whole section is devoted to the proof of Lemma 2.6. We employ an ad-hoc strategy, which we briefly outline now.

First, we clean up the γ_0-special host graph G a bit, ensuring a convenient minimum degree between the three sets X_i (the witnesses to the fact that G is γ_0-special, see Definition 2.4).

Then, given the tree T with its γ_1-special subtree T^*, rooted at t^*, we preprocess the part $T - T^*$ we have to embed. We do this by strategically choosing some cutvertices in $T - T^*$, ensuring that most of the resulting components are not very large. This allows us to group the components into two sets A_1 and A_2, which each cover roughly half of the vertices (actually we might deviate a bit from covering $\frac{m}{2}$ vertices but then gain other important knowledge about our sets of components).

Finally, we embed $T - L$, extending the given embedding of $T^* - L$, using the two sets A_1 and A_2. Components from sets A_1 will be embedded into $X_1 \cup X_3$, and components from A_2 will be embedded into $X_2 \cup X_3$.

Let us now formally give the proof of Lemma 2.6.

Setting up the constants and resuming the situation. For the output of Lemma 2.6, we choose

$$\beta := \frac{1}{10^{40}} \text{ and } m_0 := \frac{1}{\beta^{100}},$$

and set

$$\gamma_0 := \frac{2}{10^7} \text{ and } \gamma_1 := \frac{1}{50}.$$

Now, assume we are given a γ_0-special $(m+1)$-vertex graph G of minimum degree at least $\lfloor\frac{2m}{3}\rfloor$, for some $m \geq m_0$, together with a tree T with m edges, such that none of the vertices of T is adjacent to more than βm leaves. Assume T has a γ_1-nice subtree T^* rooted at t^*, and there are sets $L \subseteq V(T^*) \setminus \{t^*\}$ and $S \subseteq V(G)$ such that $|S| \leq |L| - \lceil(\frac{\gamma_1}{2})^4 m\rceil$.

Furthermore, for any large enough set W, it is possible to embed $T^* - L$ into a subset $\varphi(T^* - L)$ of $V(G) - S$, with t^* going to W. We will specify below which set W we will use.

Once $T^* - L$ is embedded, our task is to embed the rest of $T - L$ into $G-(\varphi(T^*-L)\cup S)$. Observe that because of the discrepancy of the sizes of the sets L and S, we can count on an approximation of at least $\lceil(\frac{\gamma_1}{2})^4\rceil$, that is, we know our embedding will leave at least $\lceil(\frac{\gamma_1}{2})^4 m\rceil$ vertices of $G-(\varphi(T^*-L)\cup S)$ unused.

Preparing G for the embedding. Because G is γ_0-special, there are sets X_1, X_2, X_3 partitioning $V(G)$ such that

$$\frac{m}{3} - 3\gamma_0 m \leq |X_i| \leq \frac{m}{3} + 3\gamma_0 m \tag{19}$$

for each $i = 1, 2, 3$, and such that

$$\text{there are at most } \gamma_0^{10}|X_1| \cdot |X_2| \text{ edges between } X_1 \text{ and } X_2. \tag{20}$$

Using the minimum degree condition on G, and using (20), an easy calculation shows that we can eliminate at most $\gamma_0^5 m$ vertices from each of the sets X_i, for $i = 1, 2$, in a way that the vertices of the thus obtained subsets X_i' each have degree at least $\lfloor\frac{2m}{3}\rfloor - \gamma_0^5|X_{3-i}|$ in $X_i' \cup X_3$, for $i = 1, 2$.

Because of (19), we can deduce that the number of edges between the sets X'_i and X_3 is at least $(1-6\gamma_0)|X'_i||X_3|$, for $i=1,2$. Therefore, we can eliminate at most $2\cdot\sqrt{6\gamma_0}m$ vertices from X_3, obtaining a set X'_3, so that each of the vertices in X'_3 has degree at least $(1-6\sqrt{\gamma_0})|X'_i|$ into X'_i, for $i=1,2$.

Resumingly, we eliminated a few vertices from each of the sets X_1, X_2, X_3 to obtain three sets X'_1, X'_2, X'_3 satisfying

$$|X'_i| \geq |X_i| - 5\sqrt{\gamma_0}m \tag{21}$$

such that for $i=1,2$, for any vertex $v \in X'_i$, and for each $X \in \{X'_i, X'_3\}$, we have that

$$\text{the degree of } v \text{ into } X \text{ is at least } |X| - 10\sqrt{\gamma_0}m, \tag{22}$$

and for $i=1,2$, and any vertex v in X'_3,

$$\text{the degree of } v \text{ into } X'_i \text{ is at least } |X'_i| - 3\sqrt{\gamma_0}m. \tag{23}$$

Finding more cutvertices in $T-T^*$ (if necessary), and grouping the components. Let us next have a closer look at the to-be-embedded $T-T^*$. This forest might have relatively large components, which, for reasons that will become clearer below, might add unnecessary difficulties to our embedding strategy. For this reason, we will now find a set Z of up to four new cutvertices in $T-T^*$ so that all components in $T-T^*-Z$ have controlled sizes, and can be grouped into convenient sets.

More precisely, our aim is to prove the following statement.

Claim 5.1. *There are a set $Z \subseteq V(T)\setminus V(T^*)$ with $|Z|\leq 4$, a partition $A = A_1 \cup A_2$ of the set A of all components of $T-T^*-Z$, and an independent set $Z' \subseteq Z \cup \{t^*\}$ such that*

(a) no element of A_1 is adjacent to any vertex from $Z\cup\{t^*\}\setminus Z'$;

(b) no element of A is adjacent to more than three vertices from $Z\cup\{t^*\}$;

(c) for $i=1,2$, we have that $\frac{m}{3}+3\gamma_1 m \leq |\bigcup A_i| \leq \frac{2}{3}m - 3\gamma_1 m$; and

(d) for $i=1,2$, if $|\bigcup A_i| \geq (\frac{1}{2}+10\gamma_0)m$, then $B\cap A_i=\emptyset$.

where B is the set of all elements of A that have size at most $\frac{1}{\gamma_0}$.

Note that, in particular, (a) implies that each element of A_1 is adjacent to at least one vertex of Z'.

For proving Claim 5.1, we plan to use the following folklore argument, and for completeness, we include its proof.

Claim 5.2. *Every tree D has a vertex t_D such that each component of $D - t_D$ has size at most $\frac{|D|}{2}$.*

Proof. In order to see Claim 5.2, temporarily root D at any leaf vertex v_L. Let t_D be a vertex that is furthest from v_L having the property that t_D and its descendants constitute a set of at least $\frac{|D|}{2}$ vertices. Then each component of $D - t_D$ (including the one containing v_L, has at most $\frac{|D|}{2}$ vertices. □

We can now prove the claim that finds the cutvertices and groups the components, that is, Claim 5.1.

Proof of Claim 5.1. If there is a subset A_{t^*} of the set A of all components of $T - T^*$ such that $\frac{m}{3} + 3\gamma_1 m \leq |\bigcup A_{t^*}| \leq \frac{2}{3}m - 3\gamma_1 m$, then set $Z := \emptyset$ and $Z' := \{t^*\}$. Then (a), (b) and (c) hold with $A_1 := A_{t^*}$ and $A_2 := A \setminus A_{t^*}$. By shifting a few trees belonging to B from A_{t^*} to $A \setminus A_{t^*}$ or vice versa, if necessary, we can make sure that also (d) holds.

So let us assume from now on that there is no such subset A_{t^*}. It is easy to see that then, either $T - T^*$ has three components C_1, C_2, C_3 such that

$$\frac{m}{3} - 3\gamma_1 m \leq |C_i| \leq \frac{m}{3} + 3\gamma_1 m$$

for $i = 1, 2, 3$ (plus possibly a set $\mathcal{D}$ of very small components), or there is a component of $T - T^*$ that has size greater than $\frac{2}{3}m - 3\gamma_1 m$.

In the former case, we can apply Claim 5.2 to each of the three components C_1, C_2, C_3, to obtain three vertices, z_1, z_2, z_3, such that for $i = 1, 2, 3$, the components of $C_i - z_i$ have size at most $\frac{m}{6} + \frac{3}{2}\gamma_1 m$. We shall see that then, either setting $Z := Z' := \{z_1, z_2, z_3\}$ or setting $Z := \{z_1, z_2, z_3\}$ and $Z' := \{t^*, z_1, z_2, z_3\}$, we can easily find a partition $A = A_1 \cup A_2$ with

$$|\bigcup A_1| \geq \frac{m}{2}$$

that is as desired for properties (a), (b) and (c).

This is easiest if the set $\{t^*, z_1, z_2, z_3\}$ is independent, in which case we choose to include t^* in Z'. It is then easy to find sets A_1 and A_2, as all components haves their size bounded by roughly $\frac{m}{6}$.

Otherwise, we will choose $Z' := \{z_1, z_2, z_3\}$, which will oblige us to put into A_2 any component that is only adjacent to t^*, as well as any component C_i^* that is adjacent both to t^* and to one of the vertices z_i. But since the union of all components of the former type is very small, and since there are at most two components of the second type (because t^* is adjacent to one of the vertices z_i), this is not a problem.

Finally, property (d) can be ensured by shifting some components belonging to B from A_1 to A_2, or vice versa, if necessary.

So, from now on we will suppose that there is a component D of $T - T^*$ such that

$$|D| > \frac{2}{3}m - 3\gamma_1 m. \tag{24}$$

We choose a vertex t_D as in Claim 5.2. If this results in three components C_1, C_2, C_3 of $T - T^* - t_D$ that have size

$$\frac{m}{3} - 3\gamma_1 m \leq |C_i| \leq \frac{m}{3} + 3\gamma_1 m \tag{25}$$

for $i = 1, 2, 3$ (plus possibly a number of very small components), then we will be able to proceed similarly as above, finding three cutvertices z_1, z_2, z_3, and provisionally setting $Z := \{t_D, z_1, z_2, z_3\}$ and $Z' := \{z_1, z_2, z_3\}$. We will see that it is also possible to find a partition of A into A_1 and A_2 that is as desired.

For this, note that in the present situation, there might be a total of three components in $\bigcup_{i=1,2,3}(C_i - z_i)$ neighbouring t_D. However, if t_D is adjacent to one of the z_i, there are at most two such components.

Similarly, there might be a component in $C_i - z_i$ neighbouring t^*. However, if t^* is adjacent to one of the z_i, or adjacent to t_D, then there is no such component.

Now, let us add t_D to Z', if this is possible. Next, let us add t^* to Z', if this is possible. If we end up with both t_D and t^* in Z', then we can proceed as above to find A_1 and A_2, satisfying properties (a), (b), (c) and (d) of Claim 5.1. If we end up with $t_D \in Z'$ but $t^* \notin Z'$, then there is no component in $\bigcup_{i=1,2,3}(C_i - z_i)$ neighbouring t^* (for any i), and therefore the union of all components we are obliged to add to A_2 is very small. We have no problem finding A_1 and A_2 as above. If we end up with $t_D \notin Z'$ but $t^* \in Z'$, then there are at most two components in $\bigcup_{i=1,2,3}(C_i - z_i)$ neighbouring t_D, each having size roughly $\frac{m}{6}$ at worst. (Additionally, there might be components of $D - t_D$ apart from the C_i, but the union of these components is very small.) So again, we can find A_1 and A_2 as above, satisfying (a), (b), (c). Finally, if $t_D, t^* \notin Z'$, there are at most two problematic components and we can proceed as above to find A_1 and A_2.

As before, property (d) can be ensured by shifting some components belonging to B from A_1 to A_2, or vice versa, if necessary.

So assume there are no three components C_1, C_2, C_3 of $T - T^* - t_D$ of sizes as in (25). We next treat a very similar situation, namely, the case that there are three components C'_1, C'_2, C'_3 of $T - (T^* \setminus \{t^*\}) - t_D$ that have size

$$\frac{m}{3} - 3\gamma_1 m \leq |C'_i| \leq \frac{m}{3} + 3\gamma_1 m \tag{26}$$

each (there might also be a number of very small components). In this case, we can proceed analogously to the strategy given above. We use Claim 5.2 to find three cutvertices z'_1, z'_2, z'_3 in the components C'_1, C'_2, C'_3, and set $Z := \{t_D, z_1, z_2, z_3\} \setminus \{t_D\}$ and $Z' := \{z_1, z_2, z_3\}$ (so if t^* is one of the new cutvertices z'_j, then t^* will be added to Z'). We will again be able to find a partition of the set A of all components of $T - T^* - Z$ into sets A_1 and A_2 that is as desired: As before, we try to add vertices t_D and/or t^* to Z', and see that in any outcome, the number of components in $\bigcup_{i=1,2,3}(C_i - z_i)$ neighbouring vertices not in Z' is at most two.

Therefore, from now on we assume that neither are there three components C_1, C_2, C_3 of $T - T^* - t_D$ of sizes as in (25), nor are there three components C'_1, C'_2, C'_3 of $T - (T^* \setminus \{t^*\}) - t_D$ of sizes as in (26). We will refer to this fact by saying that $T - T^* - t_D$ does not *3-split*.

Let A denote the set of all components of $T - T^* - t_D$. Set $Z := \{t_D\}$ If vertices t^* and t_D are not adjacent to each other, then we set $Z' := \{t^*, t_D\}$, and otherwise, we set $Z' := Z := \{t_D\}$.

Since A has no components of size larger than $\frac{m}{2}$, and since $T - T^* - t_D$ does not 3-split, it is clear that we can find a partition of A into sets A_1 and A_2 such that $\frac{m}{3} + 3\gamma_1 m \leq |\bigcup A_1| \leq \frac{2}{3}m - 3\gamma_1 m$. If necessary, we shift some components belonging to B either from A_1 to A_2 or from A_2 to A_1, and can thus make sure that properties (a), (b), (c) and (d) of Claim 5.1 are satisfied. This finishes the proof of Claim 5.1. □

Embedding $T - T^*$ into $S_1 \cup S_2 \cup S_3$. We now embed $T - T^*$ into $G - S$. We will make use of the sets A_1, A_2, Z and Z' and their properties as given in Claim 5.1.

If $t^* \in Z'$, then we choose W as the set X_3', that is, we let $T^* - L$ be embedded into

$$\varphi(T^* - L) \subseteq (X_1 \cup X_2 \cup X_3) \setminus S,$$

with vertex t^* embedded into a vertex $\varphi(t^*)$ from X_3'. If $t^* \notin Z'$, then we choose W as the set X_2', that is, we choose $\varphi(t^*)$ to belong to X_2'.

We also embed all vertices from $Z' \setminus \{t^*\}$ into X_3', and embed all vertices from $Z \setminus Z'$ in X_2' (taking into account the possible adjacencies between vertices from $Z \cup \{t^*\}$).

After doing this, we define, for $i = 1, 2, 3$,

$$S_i := X_i' \setminus \Big(\varphi(T^* - L) \cup \varphi(Z) \cup S\Big).$$

Then, because of (19) and (21), we have that for $i = 1, 2, 3$,

$$\begin{aligned} \frac{m}{3} + 3\gamma_0 m \ \geq \ |S_i| \geq \ & \frac{m}{3} - 5\sqrt{\gamma_0} m - \Big(\gamma_1 m - \lceil (\frac{\gamma_1}{2})^4 m \rceil\Big) - 4 \\ \geq \ & \frac{m}{3} - \frac{11}{10}\gamma_1 m. \end{aligned} \tag{27}$$

We proceed to embed the at most $5\beta m$ leaves adjacent to t^* or to vertices from Z anywhere in G, using properties (22) and (23). Since β is much smaller than γ_0, it will not matter for any future calculations where these leaves are embedded.

Before we start the actual embedding of any of the components from A, let us make some observations on how these components *could* be embedded.

For this, consider any tree $\bar{T}$ from A_1. Let $r_{\bar{T}}$ denote its root. Recall that the parent of $r_{\bar{T}}$ was embedded into S_3.

Therefore, in principle (that is, if there is enough space left), we could embed $r_{\bar{T}}$ into S_i, for either $i = 1, 2$, and then embed $\bar{T} - r_{\bar{T}}$ in a way that the even levels go into S_i, and the odd levels go into S_3, or we could embed $\bar{T} - r_{\bar{T}}$ the other way around. This means that for each $\bar{T} \in A_1$, we can embed the larger colour class of $\bar{T} - r_{\bar{T}}$ into S_3, and the rest into S_i. Even better, reembedding some of the vertices that went to S_3, and putting them instead into S_i, we can actually embed $\bar{T}$ such that for any given number t, which obeys

$$0 \leq t \leq \lceil \frac{|\bar{T}| - 1}{2} \rceil,$$

we embed t vertices into S_3, and the rest into S_i (always under the assumption that there is enough space).

This means that for the trees in $A_1 \setminus B$, we can basically work under the assumption that half of their vertices (or less, if desired) can be embedded into S_3. This is so because there are at most $\gamma_0 m$ trees in $A_1 \setminus B$, and hence at most $\gamma_0 m$ roots of such trees. So, these roots will take up little space, and it does not matter for our strategy where we embed them. For the trees in B, we can still assume that they can be embedded with a third of their vertices (or less, if desired) going to S_3.

All of this also holds for the trees in A_2, with the only difference that they have to be embedded into $S_2 \cup S_3$ (since we embedded $(Z \cup t^*) \setminus Z'$ into X_2', and therefore have no direct access to the set S_1).

So, by (27), and by properties (c) and (d) of Claim 5.1, we can embed $\bigcup A_2$ into $S_2 \cup S_3$, leaving at most $15\sqrt{\gamma_0} m$ vertices from S_2 unused. For the trees in A_2 containing only one neighbour of $Z \cup \{t^*\}$ this is straightforward. For those trees in A_2 that contain more than one neighbour of $Z \cup \{t^*\}$, we have to take some more care. We make sure that, when their time comes, each such neighbour v is embedded into a suitable vertex from S_2 (namely into a common neighbour of the images of the corresponding vertices in $Z \cup \{t^*\}$, which exists because of conditions (22) and (23), and because of Claim 5.1 (b)). This distorts our embedding plan a little, because v now goes to S_2 (while we might have accounted for v as a vertex going to S_3). However, in total there will be very few such vertices, since $|Z| \leq 4$, and hence there are only few components lying between vertices of Z, so this will not be a problem.

Next, we embed the trees from A_1. We can proceed in the same way as in the previous paragraph, the only difference being that we embed $\bigcup A_1$ into $S_1 \cup S_3$. We are aided, as before, by properties (b), (c) and (d) of Claim 5.1, and by inequalities (22), (23) and (27). Also, we use Claim 5.1 (a), which is crucial, since the roots of the trees from A_1 are embedded into S_1, and this set is not seen by the images of the vertices in $Z \cup \{t^*\} \setminus Z'$. This finishes the embedding, and thus the proof of Lemma 2.6.

References

[AKS95] M. Ajtai, J. Komlós, and E. Szemerédi. On a conjecture of Loebl. In *Graph theory, combinatorics, and algorithms, Vol. 1, 2 (Kalamazoo, MI, 1992)*, 1135–1146. Wiley, New York, 1995.

[AS08] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[BPS18] G. Besomi, M. Pavez-Signé, and M. Stein. Degree conditions for embedding trees. Accepted for publication in *SIAM Journal on Discrete Mathematics.*

[EFLS95] P. Erdős, Z. Füredi, M. Loebl, and V. T. Sós. Discrepancy of trees. *Studia Sci. Math. Hungar.*, 30(1-2):47–57, 1995.

[HRSW16] F. Havet, B. Reed, M. Stein, and D. Wood. A variant of the Erdős-Sós Conjecture. *Preprint 2016.* arXiv:1606.09343

[HKP$^+$17a] J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi. The approximate Loebl–Komlós–Sós Conjecture I: The sparse decomposition. *SIAM Journal on Discrete Mathematics* 31-2 (2017), pages 945-982.

[HKP$^+$17b] 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 Journal on Discrete Mathematics* 31-2 (2017), pages 983-1016.

[HKP$^+$17c] 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 Journal on Discrete Mathematics* 31-2 (2017), pages 1017-1071.

[HKP$^+$17d] 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 Journal on Discrete Mathematics* 31-2 (2017), pages 1072-1148.

[KSS01] J. Komlós, G. Sárközy and E. Szemerédi. Spanning Trees in Dense Graphs. *Combinatorics, Probability and Computing,* Vol. 5, 397–416 (2001).

[McD89] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.

[RS19a] B. Reed and M. Stein. Embedding Spanning Trees in Graphs of High Minimum Degree which have a Universal Vertex I: An approximate asymptotic version. *Preprint 2019.*