Ramsey goodness of bounded degree trees

Igor Balla [*]   Alexey Pokrovskiy [†]   Benny Sudakov [‡]

Abstract

Given a pair of graphs G and H, the Ramsey number $R(G, H)$ is the smallest N such that every red-blue coloring of the edges of the complete graph K_N contains a red copy of G or a blue copy of H. If a graph G is connected, it is well known and easy to show that $R(G, H) \geq (|G| - 1)(\chi(H) - 1) + \sigma(H)$, where $\chi(H)$ is the chromatic number of H and $\sigma(H)$ is the size of the smallest color class in a $\chi(H)$-coloring of H. A graph G is called H-*good* if $R(G, H) = (|G| - 1)(\chi(H) - 1) + \sigma(H)$. The notion of Ramsey goodness was introduced by Burr and Erdős in 1983 and has been extensively studied since then.

In this paper we show that if $n \geq \Omega(|H| \log^4 |H|)$ then every n-vertex bounded degree tree T is H-good. The dependency between n and $|H|$ is tight up to log factors. This substantially improves a result of Erdős, Faudree, Rousseau, and Schelp from 1985, who proved that n-vertex bounded degree trees are H-good when $n \geq \Omega(|H|^4)$.

MSC: 05C05, 05C55

1   Introduction

For a pair of graphs G and H, the *Ramsey number* $R(G, H)$ is defined to be the minimum N such that every red-blue coloring of the edges of the complete graph K_N contains a red copy of G or a blue copy of H. An old theorem of Ramsey states that $R(K_n, K_n)$ is finite and therefore $R(G, H)$ is well-defined for any G, H. It is sometimes quite difficult to compute the Ramsey number. Indeed, the inequalities

$$2^{n/2} \leq R(K_n, K_n) \leq 4^n$$

were proven by Erdős and Szekeres [11] in 1935, and Erdős [7] in 1947, and there have not been any improvements to the constant in the exponent for either bound since then.

However, there are graphs for which we can compute the Ramsey number exactly. Erdős [7] showed that for a path P_n on n vertices, we have $R(P_n, K_m) = (n-1)(m-1) + 1$. The lower bound comes from considering the graph composed of $m - 1$ disjoint red cliques of size $n - 1$, with all edges between them blue. This lower bound construction was generalized by Burr [2], who observed that for any connected graph G and any graph H,

$$R(G, H) \geq (|G| - 1)(\chi(H) - 1) + \sigma(H). \tag{1}$$

where $\chi(H)$ is the chromatic number of H and $\sigma(H)$ is the size of the smallest color class in a $\chi(H)$-coloring of H. To see that eq. (1) holds, consider the graph composed of $\chi(H) - 1$ disjoint

[*]Department of Mathematics, ETH, 8092 Zurich. igor.balla@math.ethz.ch.

[†]Department of Mathematics, ETH, 8092 Zurich. dr.alexey.pokrovskiy@gmail.com.

[‡]Department of Mathematics, ETH, 8092 Zurich. benjamin.sudakov@math.ethz.ch. Research supported in part by SNSF grant 200021-149111.

red cliques of size $|G| - 1$ and one additional red clique of size $\sigma(H) - 1$, with all edges between the cliques blue. This graph has no red copy of G because every red connected component has size at most $|G| - 1$, and it has no blue copy of H because otherwise this copy would be partitioned, via the red cliques, into $\chi(H)$ parts with one part having size $\sigma(H) - 1$, contradicting the minimality of $\sigma(H)$.

We say that G is H-*good* when equality holds in eq. (1). The notion of Ramsey goodness was introduced by Burr and Erdős [3] in 1983, and has been studied extensively since then, see e.g., [1, 6, 12, 18, 19] and their references. Note that Erdős' argument which gives a lower bound on $R(K_n, K_n)$ can be used to show that if we have relatively dense graphs G, H, then the Ramsey number is super-polynomial in $|G|$ and hence G is not H-good. Thus we restrict our attention to sparse and connected G. In 1977, Chvátal [4] showed that any tree is K_m-good. Recently, Pokrovskiy and Sudakov [20] showed that any path P with $|P| \geq 4|H|$ is H-good, verifying a conjecture of Allen, Brightwell, and Skokan [1] in a strong sense.

Since paths are a special case of trees, it is natural to consider whether trees are Ramsey good for all graphs H. In [8] Erdős, Faudree, Rousseau and Schelp ask "What is the behavior of $R(T, K(n,n))$ when T has bounded degree?" Erdős, Faudree, Rousseau and Schelp [8, 9, 10] wrote several papers on this topic. The result in their 1985 paper [9] implies that for any H, all sufficiently large bounded degree trees T are H-good. Though they do not give an explicit dependency between $|T|$ and $|H|$, their proof method can be used to show that any bounded degree tree T with $|T| \geq \Omega(|H|^4)$, is H-good. In this paper, we improve their result as follows.

Theorem 1.1. *For all Δ and k there exists a constant $C_{\Delta,k}$ such that for any tree T with maximum degree at most Δ and any H with $\chi(H) = k$ satisfying $|T| \geq C_{\Delta,k}|H|\log^4|H|$, T is H-good.*

The dependency between $|T|$ and $|H|$ in the above theorem is tight up to the $\log|H|$ factors. Indeed for $|T| \leq m = |K_m^k|/k$, no tree T is K_m^k-good for the balanced complete multipartite graph K_m^k. To see this, consider, an edge colouring of a complete graph on $(2k-1)(|T|-1)+1$ vertices consisting of $2k-1$ red cliques of size $|T|-1$, with all other edges blue. It is easy to check that this graph has no red T and no blue K_m^k showing that $R(T, K_m^k) \geq (2k-1)(|T|-1)+1 > (k-1)(|T|-1)+m$.

In the proof of Theorem 1.1, we first consider the case where our tree T has many leaves. In this case, we are able to obtain the following stronger result.

Theorem 1.2. *Let T be a tree with l leaves and maximum degree at most Δ, and let H be a graph satisfying $l \geq 13\Delta|H|+1$. Then T is H-good.*

The results in this paper cannot be extended to trees without a bounded degree assumption. In particular a star S_n is not $K_{m,m}$-good for $n \leq O\left(2^{\frac{m+1}{2}}\right)$. To see this, recall that there are N-vertex graphs with minimum degree $\Omega\left(N^{1-\frac{2}{m+1}}\right)$ containing no $K_{m,m}$ (see for example [14], Theorem 2.26). The complement of such a graph is S_n free for $n = N - N^{1-\frac{2}{m+1}}$. From this, it can be shown that for $n \geq \Omega\left(2^{\frac{m+1}{2}}\right)$ there is such a graph on $\geq 2n$ vertices which is $K_{m,m}$ free and has minimum degree $N - n$. Equivalently we get that $R(S_n, K_{m,m}) \geq 2n$ for $n \geq \Omega\left(2^{\frac{m+1}{2}}\right)$. In particular this shows that S_n is not $K_{m,m}$-good for $n \geq \Omega\left(2^{\frac{m+1}{2}}\right)$ which shows that Theorem 1.1 cannot hold for an arbitrary tree T.

Remark 1.3. The condition $l \geq 13\Delta|H| + 1$ can be replaced with $l \geq 13\Delta m + 1$ where m is the size of the largest color class in a $\chi(H)$ coloring of H. Indeed, this is what we actually prove in Lemma 3.4.

2 Overview

Notation

For a graph G, we let $E(G)$ denote the set of edges of G. We define K_m^k to be the complete k-partite graph with parts of size m, i.e. a graph on km vertices partitioned into k classes of m vertices each, with edges between any pair of vertices from different classes. Note that in particular, K_m^1 denotes the empty graph on m vertices. Also let $K_{m_1,\ldots,m_k}$ be the complete multipartite graph with parts of size $m_1, \ldots, m_k$. For a graph G and vertex x, we let $N(x) = N_G(x) = \{y \in G : xy \in E(G)\}$ denote the neighborhood of x. We analogously let $d_G(x) = |N_G(x)|$ denote the degree of x and $\Delta(G)$ denote the maximum degree of a vertex in G. For any subset $S \subseteq G$, we define the neighborhood $N(S) = N_G(S) = \bigcup_{x \in S} N_G(x) \backslash S$.

Proof outline

We are given a tree T with n vertices and a graph H with $\chi(H) = k$ and $\sigma(H) = m_1$, and we would like to show that any graph G on $(n-1)(k-1) + m_1$ vertices either has a copy of T, or G^c has a copy of H. Note that as long as k and m_1 are fixed, adding more edges to H only makes the problem more difficult. Indeed, if we let $m_1 \leq \ldots \leq m_k$ be the sizes of the parts in a k-coloring of H, then a graph not containing H also doesn't contain $K_{m_1,\ldots,m_k}$. Because of this we will actually prove the following slightly stronger version of Theorem 1.1.

Theorem 2.1. *For all Δ and k, there exists a constant $C_{\Delta,k}$ such that for any tree T with max degree at most Δ and numbers $m_1 \leq m_2 \leq \cdots \leq m_k$ with $|T| \geq C_{\Delta,k} m_k \log^4 m_k$, the tree T is $K_{m_1,m_2,\ldots,m_k}$-good.*

Assume that we are given a graph G on $(n-1)(k-1) + m_1$ vertices such that G^c has no copy of $K_{m_1,\ldots,m_k}$. To prove Theorem 2.1 we need to show that G has a copy of T. Notice that since G^c has no copy of $K_{m_1,\ldots,m_k}$, we have that G^c has no copy of $K_{m_k}^k$ and most of the time we will only use this weaker assumption. For simplicity, in this proof outline we will deal only with the case when $m_1 = \cdots = m_k = m$.

Finding trees in expanders

The basic technique in this paper is to use results about finding large trees in graphs which are expanders. Here "expander" means a graph G in which $N(S)$ is suitably large for every set of vertices S. Expanders are closely related to graphs G with G^c containing no copy of $K_{m,m}$. Indeed it is easy to see that G^c being $K_{m,m}$-free is equivalent to every set S with $|S| = m$ satisfying $|N(S)| > |G| - 2m$.

Trees in expanders are well studied. By results of Friedmann-Pippenger [13] and Haxell [15], expanders contain all suitably large trees. See Lemma 3.1 for the specific instance of this which we apply in our paper. By applying this to graphs G with G^c containing no copy of K_m^k, one can

immediately prove something quite similar to Theorem 2.1—In Lemma 4.9 we prove that

$$R(K_m^k, T) \leq (k-1)(|T| + 13\Delta m) + m \text{ for any tree } T \text{ with } \Delta(T) \leq \Delta. \tag{2}$$

When $|T| \gg m$, this is quite close to the bound "$R(K_m^k, T) \leq (k-1)(|T|-1) + m$" which we want to prove in Theorem 2.1. Using (2) we obtain that for any subtree $T' \subseteq T$ with $|T'| \leq |T| - 13\Delta m$, we have $R(K_m^k, T') \leq (k-1)(|T|-1) + m$. This shows that we can find any large subtree T' of T in a graph G with $|G| = (k-1)(|T|-1) + m$ and G^c K_m^k-free. The bulk of the proof of Theorem 2.1 consists of extending T' into a copy of T. This extension is performed by different methods depending on whether T has many leaves or many bare paths (A *bare* path in a graph is a path such that all interior vertices have degree 2). It is a well known result (see eg. Lemma 2.1 in [16]) that a tree either has many leaves or many long bare paths.

Lemma 2.2. *For any integers $n, r > 2$, a tree on n vertices either has at least $n/4r$ leaves or a collection of at least $n/4r$ vertex disjoint bare paths, each having length r.*

Theorem 2.1 is proved by different methods depending on which case of Lemma 2.2 holds for the tree T.

Case 1: Many leaves

In section 3, we suppose our tree T has many leaves. Here "many" means that T has $> 13\Delta m$ leaves. In this case if we let T' be T with $13\Delta m$ leaves deleted, then using (2) any G with G^c K_m^k-free contains a copy of T'. With some extra work, it is possible to find such a copy of T' with all subsets of $V(T')$ expanding outside $V(T')$. Once we have this, it is easy to find all the required leaves using a variant of Hall's Theorem (Lemma 3.3).

Case 2: Many bare paths

In section 4, we consider the case where our tree has few leaves, and therefore many long bare paths by Lemma 2.2. In this case we will often need to find disjoint paths of prescribed length between pairs of vertices, so we make the following definition.

Definition 2.3. *For two sets X and W in a graph, we say that (X, W) is (s, d^-, d^+)-linked system if the following holds. Suppose that we have distinct vertices $x_1, y_1, \ldots, x_s, y_s \in X$, and integers $d_1, \ldots, d_s$ with $d^- \leq d_i \leq d^+$ for all i. Then there are disjoint paths $P_1, \ldots, P_s$ with P_i going from x_i to y_i, P_i internally contained in W, and P_i having length d_i.*

We then follow the approach of Montgomery [17], who shows that an expander is a $(s, d^-, d+)$-linked system for some appropriate choices of s, d^-, d^+ (Lemma 4.5 and Theorem 4.7.) Thus we first apply results like (2) in order to embed the tree with the bare paths removed and then apply Montgomery's result to find the required bare paths, completing the embedding. This strategy works to prove Theorem 2.1 for the case $k = 2$ (see section 4.1.)

When $k \geq 3$, the proof is substantially more complicated and is dealt with in section 4.2. The first step is to note that by (2), if G has no copy of T, then G^c contains a copy of K_M^{k-1} for $M \gg m$ (see Claim 4.18.) Then there are two subcases depending on whether there are a lot of short (length at most 3) paths in G between some pairs of parts of the copy of K_M^k.

If there are a lot of paths between pairs of parts of K_M^k, then these parts together with the paths between them form a large linked system. Using (2) and techniques from Montgomery [17], we find a copy of the tree T similar to how we did in the $k = 2$ case.

If there are few paths between all pairs of parts of K_M^k, then we show that the entire graph G must be close to Burr's extremal construction for showing (1). Specifically, we show that in this case G has $k - 1$ disjoint sets $H_1, \ldots, H_{k-1}$ of size $\geq 0.9n$, which have no edges between them. In Lemma 4.16 we analyse graphs with this structure and prove Theorem 2.1 for them.

3 Embedding a tree with many leaves

To deal with the case where our tree has many leaves, we will need a result of Haxell [15], which lets us embed a bounded degree tree with prescribed root into a graph with sufficient expansion. In section 4.2, we will actually need a generalization of this result to forests, so we state the more general version in the following lemma. For a proof of Lemma 3.1, we refer the reader to the appendix.

Lemma 3.1. *Let Δ, M, t and m be given. Let $X = \{x_1, \ldots, x_t\}$ be a set of vertices in a graph G. Suppose that we have rooted trees $T_{x_1}, \ldots, T_{x_t}$ satisfying $\sum_{i=1}^t |T_{x_i}| \leq M$ and $\Delta(T_{x_i}) \leq \Delta$ for all i. Suppose that for all S with $m \leq |S| \leq 2m$ we have $|N(S)| \geq M + 10\Delta m$, and for S with $|S| \leq m$ we have $|N(S) \setminus X| \geq 4\Delta|S|$.*

Then we can find disjoint copies of the trees $T_{x_1}, \ldots, T_{x_t}$ in G such that for each i, T_{x_i} is rooted at x_i. In addition for all $S \subseteq T_{x_1} \cup \ldots \cup T_{x_t}$ with $|S| \leq m$, we have

$$\left| N(S) \setminus \left(T_{x_1} \cup \cdots \cup T_{x_t} \right) \right| \geq \Delta|S|.$$

As a corollary, we can embed a large bounded degree tree into a graph whose complement does not contain K_{m_1, m_2}.

Corollary 3.2. *Let Δ, m_1, m_2 be integers, T be a forest with $\Delta(T) \leq \Delta$, and G a graph with $|G| \geq |T| + 13\Delta m_1 + m_2$ such that G^c does not contain K_{m_1, m_2}. Then G contains a copy of T. Additionally, for all $S \subseteq T$ with $|S| \leq m_1$, we have*

$$|N(S) \backslash T| \geq \Delta|S|.$$

Proof. Since every forest F is a subgraph of some tree on $|F|$ vertices, without loss of generality we may suppose that T is a tree.

Since G^c does not contain K_{m_1, m_2}, we have that for any $S \subseteq G$ with $m_1 \leq |S| \leq 2m_1$, $|N_G(S)| \geq |G| - 2m_1 - m_2$. Indeed, otherwise $|G \backslash (N_G(S) \cup S)| > m_2$ and so if we choose $A \subseteq S$, $B \subseteq G \backslash (N_G(S) \cup S)$ with $|A| = m_1, |B| = m_2$, then $A \cup B$ is a copy of K_{m_1, m_2} in G^c.

Now if we choose $|X| \leq m_1 - 1$ maximal so that $|N_G(X)| \leq 4\Delta|X|$, then we claim that $G' = G \backslash X$ satisfies that for all $S \subseteq G'$ with $1 \leq |S| \leq m_1$, $|N_{G'}(S)| \geq 4\Delta|S| + 1$. Indeed, for any $S \subseteq G'$ with $1 \leq |S| \leq m_1$, if $|N_{G'}(S)| \leq 4\Delta|S|$ then $|N_G(X \cup S)| \leq |N_G(X) \cup N_G(S)| \leq 4\Delta|X \cup S|$, so we must have $m_1 \leq |X \cup S| \leq 2m_1$ by maximality of X. But then

$$8\Delta m_1 \geq 4\Delta|X \cup S| \geq |N_G(X \cup S)| \geq |G| - 2m_1 - m_2,$$

contradicting the assumption of the lemma. Also, for any $S \subset G'$ with $m_1 \leq |S| \leq 2m_1$, we have

$$|N_{G'}(S)| \geq |N_G(S)| - |X| \geq |N_G(S)| - m_1 \geq |G| - 3m_1 - m_2 \geq |T| + 10\Delta m_1.$$

Thus we may apply Lemma 3.1 with the graph G', $m = m_1$, $X = \{x\}$ for any vertex x, and the tree $T_x = T$, to obtain that G' contains a copy of T. Moreover, for all $S \subseteq T$ with $|S| \leq m_1$, we have

$$|N_G(S) \backslash T| \geq |N_{G'}(S) \backslash T| \geq \Delta |S|. \qquad \square$$

We will also need the following extension of Hall's theorem (see eg. exercise 25.4 in [22].)

Lemma 3.3. *Given a bipartite graph (A, B) and a function $l : A \to \mathbb{N}$, if $|N(S)| \geq \sum_{v \in S} l(v)$ for all $S \subseteq A$ then the graph contains a forest F such that $d_F(v) = l(v)$ for all $v \in A$ and $d_F(v) \leq 1$ for all $v \in B$.* $\qquad \square$

We are now ready to prove that a bounded degree tree with sufficiently many leaves is $K_{m_1,\ldots,m_k}$-good.

Lemma 3.4. *Let $l, \Delta, k \in \mathbb{N}$ and $m_1 \leq \ldots \leq m_k$ be given with $l \geq 13\Delta m_k + 1$. Then any tree T with l leaves and $\Delta(T) \leq \Delta$ is $K_{m_1,\ldots,m_k}$ good.*

Proof. Let $n = |T|$. We proceed by induction on k. For $k = 1$, any graph on m_1 vertices trivially contains $K_{m_1}^1$ as a subgraph (since K_m^1 is the graph with m vertices and no edges.) Now suppose $k \geq 2$ and let G be a graph with $(k-1)(n-1)+m_1$ vertices such that G^c does not contain $K_{m_1,\ldots,m_k}$.

First suppose there exists $S \subseteq G$ with $|S| \geq m_k$, such that $|N_G(S)| + |S| \leq n - 1$. Then letting $H = G \backslash (N_G(S) \cup S)$, we have $|H| \geq (k-2)(n-1) + m_1$ and H does not contain a $K_{m_1,\ldots,m_{k-1}}$, or else we could take it together with an m_k vertex subset of S to get a copy of $K_{m_1,\ldots,m_k}$ in G^c. Thus we may apply induction to H to conclude that it contains a copy of T.

Otherwise, we have that for all $S \subseteq G$ with $|S| \geq m_k$, $|N_G(S)| + |S| \geq n$. For sets S with $|S| = m_k$, this is equivalent to G^c not containing $K_{m_k, m'}$ for $m' = (k-2)(n-1) + m_1$. Now let T' be the subtree of T with all leaves removed. Using $l \geq 13\Delta m_k + 1$ we have

$$(k-1)(n-1) + m_1 \geq n - l + 13\Delta m_k + (k-2)(n-1) + m_1 = n - l + 13\Delta m_k + m'.$$

Combining this with $|T'| = n - l$, we can apply Corollary 3.2 to conclude that G contains a copy of T'. Now let P be the vertices of T' to which we need to connect leaves in order to get T, and let $l(v)$ be the number of leaves to attach for each $v \in P$. From the last part of Corollary 3.2, we have that for any $S \subseteq P$ with $|S| \leq m_k$,

$$|N_G(S) \backslash T'| \geq \Delta |S| \geq \sum_{v \in S} l(v).$$

Moreover, for any $S \subseteq P$ with $|S| \geq m_k$, we have $|N_G(S)| + |S| - n \geq 0$ which implies

$$|N_G(S) \backslash T'| \geq |N_G(S)| - |T' \backslash S| = |N_G(S)| + |S| - n + l \geq l = \sum_{v \in P} l(v) \geq \sum_{v \in S} l(v).$$

Thus we may apply Lemma 3.3 to complete the embedding of T. $\qquad \square$

Theorem 1.2 now follows immediately.

Proof of Theorem 1.2. Let $n = |T|$, $k = \chi(H)$ and $m_1 \leq \ldots \leq m_k$ be the sizes of the color classes in a k-coloring of H, so that $m_1 = \sigma(H)$. Let G be a graph on $(n-1)(k-1)+m_1$ vertices such that G^c has no copy of H. Then G^c has no copy of $K_{m_1,\ldots,m_k}$, and we have that $\ell \geq 13\Delta|H|+1 \geq 13\Delta m_k+1$, so by Lemma 3.4 G must contain a copy of T. $\qquad \square$

4 Embedding a tree with few leaves

If a bounded degree tree doesn't have many leaves, then it has many long bare paths by Lemma 2.2, so it remains to embed such trees. We will need the following definitions and lemmas of Montgomery [17]. First we define a notion of expansion into a subset of a graph.

Definition 4.1. *For a graph G and a set $W \subseteq G$, we say G d-expands into W if*

1. $|N(X) \cap W| \geq d|X|$ for all $X \subseteq G$ with $1 \leq |X| < \left\lceil \frac{|W|}{2d} \right\rceil$.

2. $e(X, Y) > 0$ for all disjoint $X, Y \subseteq G$ with $|X| = |Y| = \left\lceil \frac{|W|}{2d} \right\rceil$.

Definition 4.2. *We call G an (n, d)-expander if $|G| = n$ and it d-expands into G.*

We state some basic properties of expansion.

Lemma 4.3. *Let $W \subseteq Z \subseteq G$ and suppose that G d-expands into W.*

(i) Z d-expands into W.

(ii) If $d \geq 2$ then G d-expands into Z.

(iii) If $d > 1$ and $d/(d-1) \leq c \leq d$ then G c-expands into W.

Proof. (i) follows directly from the definition.

For (ii), condition 2 follows immediately. For condition 1, let $X \subseteq G$ with $1 \leq |X| < \left\lceil \frac{|Z|}{2d} \right\rceil$ be given. If $|X| < \left\lceil \frac{|W|}{2d} \right\rceil$ then we have $|N(X) \cap Z| \geq |N(X) \cap W| \geq d|X|$. Otherwise if $\left\lceil \frac{|W|}{2d} \right\rceil \leq |X| < \left\lceil \frac{|Z|}{2d} \right\rceil$ then we know that $|Z \setminus (N(X) \cup X)| < \left\lceil \frac{|W|}{2d} \right\rceil$ by condition 2 of d-expansion. It follows that

$$|N(X) \cap Z| \geq |Z| - |X| - \frac{|W|}{2d} \geq |Z| - \frac{|Z|}{2d} - \frac{|W|}{2d} \geq \frac{|Z|}{2} \geq d|X|.$$

The proof of (iii) is similar to that of (ii). The interesting case to check is when $\left\lceil \frac{|W|}{2d} \right\rceil \leq |X| < \left\lceil \frac{|W|}{2c} \right\rceil$, which implies $|W \setminus (N(X) \cup X)| < \left\lceil \frac{|W|}{2d} \right\rceil$ by condition 2 of d-expansion. Notice that $d/(d-1) \leq c$ is equivalent to $c^{-1} + d^{-1} \leq 1$. Combining these gives

$$|N(X) \cap W| \geq |W| - |X| - \frac{|W|}{2d} \geq |W| - \frac{|W|}{2c} - \frac{|W|}{2d} \geq \frac{|W|}{2} \geq c|X|. \qquad \square$$

We will also need a useful decomposing property of this expansion.

Lemma 4.4 (Lemma 2.3 of Montgomery [17]). *There exists n_0 such that for $k, n \in \mathbb{N}$ with $n \geq n_0$ and $k \leq \log n$, if we have $m_1, \ldots, m_k \in \mathbb{N}$ with $m = m_1 + \ldots + m_k$ and $d_i = \frac{m_i}{5m} d \geq 2 \log n$, then for any graph G with n vertices which d-expands into W with $|W| = m$, we can partition W into k disjoint sets $W_1, \ldots, W_k$ of sizes $m_1, \ldots, m_k$ respectively, so that G d_i-expands into W_i.* $\qquad \square$

The following lemma will be crucial for section 4.2. It allows us to simultaneously find many paths of prescribed lengths between endpoints in an expander graph.

Lemma 4.5 (Lemma 3.2 of Montgomery [17])**.** *Let G be a graph with n vertices, where n is sufficiently large, and let $d = 160 \log n / \log \log n$. Suppose $r, k_1, \ldots, k_r$ are integers with $4 \left\lceil \frac{\log n}{\log \log n} \right\rceil \leq k_i \leq n/40$ for each i, and $\sum_i k_i \leq 3|W|/4$. Suppose G contains disjoint vertex pairs $(x_i, y_i), 1 \leq i \leq r$, and let $W \subset G$ be disjoint from these vertex pairs.*

If G d-expands into W, then we can find disjoint paths P_i, $1 \leq i \leq r$, with interior vertices in W, so that each path P_i is an x_i, y_i-path with length k_i. ☐

It will be convenient for us to restate the previous lemma using the definition of a linked system.

Corollary 4.6. *Let $n, s \in \mathbb{N}$ and $c = 160 \log n / \log \log n$. Suppose that G is a graph on n vertices and $W \subseteq G$ such that $n \geq |W| + 2s$ and G c-expands into W. Then $(G \backslash W, W)$ is (s, d^-, d^+)-linked system, for $d^- = 4 \left\lceil \frac{\log n}{\log \log n} \right\rceil$ and $d^+ = \frac{|W|}{40s}$.*

Proof. This follows immediately from Lemma 4.5 and the definition of (s, d^-, d^+)-linked system. ☐

Lemma 4.5 shows that if a graph G expands into a set W, then it is possible to cover $3/4$ of W by disjoint paths of prescribed length. The following theorem shows that, under similar assumptions to Lemma 4.5, it is possible to cover all of W by such paths.

Theorem 4.7 (Theorem 4.3 of Montgomery [17])**.** *Let n be sufficiently large and let $l \in \mathbb{N}$ satisfy $l \geq 10^3 \log^2 n$ and $l | n$. Let a graph G contain n/l disjoint vertex pairs (x_i, y_i) and let $W = G \backslash (\cup_i \{x_i, y_i\})$. Suppose G d-expands into W, where $d = 10^{10} \log^4 n / \log \log n$. Then we can cover G with n/l disjoint paths P_i of length $l - 1$, so that P_i is an x_i, y_i-path.* ☐

Montgomery uses the above theorem to embed a spanning tree with many long bare paths in an expander. The idea is to first find a copy of the tree with the bare paths removed, and then apply Theorem 4.7 to find the paths. We will use this theorem for the same purpose in section 4.1.

4.1 The case $k = 2$

If we have a graph with at least n vertices for which small sets expand and whose complement does not contain K_m^2, then we can find an embedding of the tree via Theorem 4.7, as in Montgomery [17].

Lemma 4.8. *Let $n, m, \Delta \in \mathbb{N}$ with n sufficiently large relative to Δ and let $d = 4 \cdot 10^{12} \frac{\log^4 n}{\log \log n}$, $r = \lceil 10^3 \log^2 n \rceil$, such that $n \geq 2(d+1)m$. Let T be a tree with n vertices, $\Delta(T) \leq \Delta$, and at least $n/(4r)$ disjoint bare paths of length r. If G is a graph with n' vertices such that $n' \geq n$, G^c does not contain K_m^2, and for all $S \subset G$ with $|S| \leq m$, $|N(S)| \geq d|S|$, then G contains a copy of T.*

Proof. If $n' \geq n + 13\Delta m + m$ then G contains a copy of T by Corollary 3.2. Otherwise we have $n \leq n' < n + 13\Delta m + m = n(1 + o(1))$.

We first note that G is an (n', d)-expander. Indeed, for any $S \subseteq G$ with $1 \leq |S| \leq m$ we have $|N(S)| \geq d|S|$ by assumption. For $S \subseteq G$ with $m \leq |S| < \lceil n'/(2d) \rceil$, using $n' \geq 2(d+1)m$ and the K_m^2-freeness of G^c we have

$$|N(S)| \geq n' - |S| - m \geq d|S|,$$

so the first condition holds. Moreover, since G^c does not have K_m^2 and $\lceil n'/(2d) \rceil \geq m$, the second condition holds as well.

Now let T' be T with the interior vertices of the $n/4r$ bare paths of length r deleted. Then $|T'| = 3n/4 + n/(4r)$. Let $n_1 = n' - n/8$ and $n_2 = n/8$. Then if we let $d_i = \frac{n_i}{5n'}d$, we can apply Lemma 4.4 to partition G into G_1 and G_2 such that $|G_i| = n_i$ and G d_i-expands into G_i. Note that $m = o(n)$ and hence

$$n_1 = n' - n/8 \geq \frac{7n}{8} \geq \frac{3n}{4} + \frac{n}{4r} + 13\Delta m + m = |T'| + 13\Delta m + m.$$

Morever, G_1^c has no K_m^2 so we conclude by Corollary 3.2 that G_1 contains a copy of T'. Let (x_i, y_i) be the disjoint vertex pairs in the copy of T' that need to be connected by paths to get T.

Let G' be any subgraph of G of size $(r+1)n/4r$ containing $G_2 \cup (\bigcup_i \{x_i, y_i\})$, and let $W = G' \setminus (\bigcup_i \{x_i, y_i\})$. Since $G_2 \subseteq W$, we may apply Lemma 4.3 $(i), (ii)$ to conclude that G' d_2-expands into W. We have

$$d_2 = dn/40n' \geq d/41 \geq 10^{10} \log^4 n / \log\log n \geq 10^{10} \log^4 |G'| / \log\log |G'|.$$

By Lemma 4.3 (iii), G' $10^{10} \log^4 |G'| / \log\log |G'|$-expands into W. Since $|G'| \leq n$, we have and $r + 1 \geq 10^3 \log^2 |G'|$. Combining these, we can apply Theorem 4.7 with $l = r + 1$, $G = G'$, and $d = 10^{10} \log^4 |G'| / \log\log |G'|$ to conclude that the pairs (x_i, y_i) can be connected by disjoint paths of length r in G', completing the embedding of T. $\qquad\square$

Putting Lemma 3.4 and Lemma 4.8 together, we may conclude the case $k = 2$ for all bounded degree trees as follows.

Proof of Theorem 2.1 for $k = 2$. Let $d = 4 \cdot 10^{12} \frac{\log^4 n}{\log\log n}$ and $r = \lceil 10^3 \log^2 n \rceil$. We can choose $C_{\Delta,k}$ such that n is sufficiently large relative to Δ, k and $n \geq (2d + 3)m_2$. Now let G be a graph with $n + m_1 - 1$ vertices such that G^c does not contain K_{m_1,m_2}. Notice that in particular, G^c doesn't contain $K_{m_2}^2$. If T has at least $n/4r \geq 13\Delta m_2 + 1$ leaves, then by Lemma 3.4 we are done. Otherwise, by Lemma 2.2 T has at least $n/4r$ disjoint bare paths of length r. Note that since G^c has no K_{m_1,m_2}, we have that for any $S \subseteq G$ with $|S| \geq m_1$, $|N(S)| \geq n - m_2 - |S|$. Now choose $X \subseteq G$ with $|X| \leq m_1 - 1$ maximal so that $|N(X)| < d|X|$ and let $G' = G \setminus X$. Then we claim that for all $S \subseteq G'$ with $|S| \leq m_1$, $|N_{G'}(S)| \geq d|S|$. Indeed, otherwise we would have $|N(X \cup S)| < d|X \cup S|$, so by maximality of X this would imply $m_1 \leq |X \cup S| \leq 2m_1$. But then

$$2dm_1 \geq d|X \cup S| > |N(X \cup S)| \geq n - m_2 - |X \cup S| \geq n - 2m_1 - m_2,$$

a contradiction. For S with $m_1 \leq |S| \leq m_2$ we have $|N(S)| \geq n - 2m_2 \geq dm_2 \geq d|S|$. Since $|X| \leq m_1 - 1$, we have $|G'| \geq n$. Thus we may apply Lemma 4.8 to G' with $m = m_2$ to conclude that G' contains a copy of T. $\qquad\square$

4.2 The case $k \geq 3$

We first extend Corollary 3.2 to show that we can embed a large bounded degree tree into a graph whose complement does not contain K_m^k.

Lemma 4.9. *Let $\Delta, k, m \in \mathbb{N}$ be given, T a tree with $\Delta(T) \leq \Delta$, and G a graph with $|G| \geq (k-1)(|T| + 13\Delta m) + m$ such that G^c does not contain K_m^k. Then G contains a copy of T.*

Proof. We proceed by induction on k. For $k = 1$, any graph on m vertices trivially contains K_m^1. Now suppose $k \geq 2$ and let $m' = (k-2)(|T| + 13\Delta m) + m$. If G^c does not contain $K_{m,m'}$ then by Corollary 3.2, G contains a copy of T.

Otherwise G contains disjoint A, B with $|A| = m, |B| = m'$ and no edges between A and B. Then B^c does not contain a copy of K_m^{k-1} or else taking this copy together with A would give a copy of K_m^k in G^c. But then by induction, B contains a copy of T. $\qquad\square$

Moreover, for $k \geq 3$, we observe that we can embed much larger bounded degree forests than trees. This makes sense in view of the Burr's construction showing (1) – it does not have a tree on n vertices, but it has a forest made of $k-1$ trees each of size $n-1$.

Corollary 4.10. *Let* $k, m, \Delta \in \mathbb{N}$ *be given with* $k \geq 3$, *and let* T_a, T_b *be trees with* $|T_a| \leq |T_b|$ *and* $\Delta(T_a), \Delta(T_b) \leq \Delta$. *Let* G *be a graph with* G^c *not containing* K_m^k. *If*

$$|G| \geq |T_a| + (k-1)(|T_b| + 13\Delta m) + m,$$

then G *contains a copy of the forest* $T_a \cup T_b$.

Proof. We first apply Lemma 4.9 to obtain a copy of T_a in G. Now we let $G' = G \backslash T_a$ and apply Lemma 4.9 to G' to obtain a copy of T_b in G'. $\qquad\square$

The following lemma lets us find a copy of T in a sufficiently large graph which contains a linked system and whose complement is K_m^k-free, but does contain K_u^{k-1}, for a sufficiently large u. The idea of the proof is to break up our tree into three parts—two forests T_a, T_b, and a collection of bare paths joining the forests. Then the forests T_a and T_b are found using Corollary 4.10, while the bare paths are found using the linked system.

Lemma 4.11. *Let* $n, m, k, \Delta \in \mathbb{N}$ *with* $k \geq 3$ *and* n *sufficiently large relative to* Δ, k *and let* $d = 4 \cdot 10^{12} \frac{\log^4 n}{\log \log n}$, $r = \lceil 10^3 \log^2 n \rceil$ *and* $y = \lceil \log n \rceil$, *such that* $n \geq 2(d+1)m$. *Let* X, W, Z *be disjoint subsets of a graph such that* $(Z \cup X)^c$ *is* K_m^k-*free with* $|Z| \geq 0.99(k-1)n$. *Let* T *be a tree on* n *vertices with* $\Delta(T) \leq \Delta$, *and at least* $n/4r$ *bare paths of length* r. *Suppose that* X^c *contains* K_u^{k-1} *for* $u \geq 2n/r$. *Suppose that* (X, W) *is a* $(n/2r, d^-, d^+)$-*linked system for* $d^- \leq y \leq d^+$. *Then* $Z \cup X \cup W$ *contains a copy of* T.

Proof. We first find a subset of Z with appropriate expansion properties.

Claim 4.12. *There exists* $Z' \subseteq Z$ *with* $|Z'| \geq 0.9(k-1)n$ *such that* $|N(S) \cap X| \geq |S|$ *for any* $S \subseteq Z'$ *with* $|S| \leq n/r$.

Proof. Let $U_1, \ldots, U_{k-1}$ be the parts of the K_u^{k-1} in X^c. If there exists $S \subseteq Z$ with $|S| \geq m$ and $|N_{G^c}(S) \cap U_i| \geq m$ for all i, then we can take subsets of size m from $S, N_{G^c}(S) \cap U_1, \ldots, N_{G^c}(S) \cap U_{k-1}$ to obtain a K_m^k in $(X \cup Z)^c$, a contradiction. Thus for all $S \subseteq Z$ with $n/r \geq |S| \geq m$, we have that $|N_G(S) \cap X| \geq u - m \geq n/r \geq |S|$ (using $m = o(n)$.)

Now let $A \subset Z$ with $|A| \leq m-1$ be maximal such that $|N_G(A) \cap X| < |A|$, and let $Z' = Z \backslash A$. We claim that for all $S \subseteq Z'$ with $|S| \leq m$, $|N_G(S) \cap X| \geq |S|$. Indeed, otherwise $|N_G(A \cup S) \cap X| < |A \cup S|$, so we must have $m \leq |A \cup S| \leq 2m$ by maximality of A. But then

$$2m \geq |A \cup S| > |N_G(A \cup S)| \geq n/r,$$

a contradiction to $n \geq 2(d+1)m$. $\qquad\square$

10

Now let T_a be a collection of $n/4r$ disjoint paths of length $r - 2y - 4$, so that $|T_a| = n(r - 2y - 3)/(4r) \leq n/4$ and let T_b be T without the interior vertices of the $n/4r$ bare paths of length r, so that $|T_b| = n - n(r-1)/4r = 3n/4 + n/(4r)$. Since we can always add edges to T_a and T_b to make them trees without increasing the maximum degree, and

$$|Z'| \geq 0.9(k-1)n \geq \frac{n}{4} + (k-1)\left(\frac{3n}{4} + \frac{n}{4r} + 13\Delta m\right) + m,$$

we may apply Corollary 4.10 to conclude that Z' has a copy of T_a and T_b. Let $x_a, y_a \in Z'$ for $1 \leq a \leq n/2r$ be the endpoints of those copies so that if we connect x_a with y_a by disjoint paths of length $y + 2$ for all i, we obtain an embedding of T. By Lemma 3.3 and the claim, there is a matching from $\{x_a : 1 \leq a \leq n/2r\} \cup \{y_a : 1 \leq a \leq n/2r\}$ to some set $\{x'_a : 1 \leq a \leq n/2r\} \cup \{y'_a : 1 \leq a \leq n/2r\}$ contained in X. Since (X, W) is a $(n/2r, d^-, d^+)$-linked system for $d^- \leq y \leq d^+$, there are disjoint x'_a to y'_a paths of length y in W as required. $\qquad\square$

Next we prove two lemmas which help us construct linked systems. Lemma 4.13 lets us combine 2 linked systems into a bigger linked system, provided that there are sufficiently many short paths between them. In Lemma 4.14, we combine several linked systems with many short paths between them into a big linked system, by making repeated use of Lemma 4.13.

Lemma 4.13. *Suppose that we have sets of vertices X_1, X_2, W_1, W_2 with $(X_1 \cup W_1) \cap (X_2 \cup W_2) = \emptyset$, such that (X_1, W_1) is a (s_1, d_1^-, d_1^+)-linked system and (X_2, W_2) is a (s_2, d_2^-, d_2^+)-linked system. Suppose that there are disjoint paths $P_1, \ldots, P_t$ of length ≤ 3 from X_1 to X_2 internally outside $X_1 \cup X_2 \cup W_1 \cup W_2$. Then $\left(X_1 \cup X_2, W_1 \cup W_2 \cup \bigcup_{i=1}^{t} P_t\right)$ is a (s, d^-, d^+)-linked system for $d^- = d_1^- + d_2^- + 3$, $d^+ = \min(d_1^+, d_2^+)$, and $s = \min(s_1, s_2, t/3)$.*

Proof. Let $x_1, y_1, \ldots, x_s, y_s$ be vertices in $X_1 \cup X_2$ and $d_1, \ldots d_s \in [d^-, d^+]$ as in the definition of (s, d^-, d^+)-linked system. To prove the lemma we need to find disjoint paths $Q_1, \ldots, Q_s$ with Q_i a length d_i path from x_i to y_i. Without loss of generality $x_1, y_1, \ldots, x_s, y_s$ are labeled so that $x_1, y_1, \ldots, x_a, y_a \in X_1$, $x_{a+1}, y_{a+1}, \ldots, x_b, y_b \in X_2$, $x_{b+1}, \ldots, x_s \in X_1$, and $y_{b+1}, \ldots, y_s \in X_2$ for some a and b.

Since the paths $P_1, \ldots, P_t$ are disjoint and have only 2 vertices each in $X_1 \cup X_2$, we have that $\leq 2s$ of the paths $P_1, \ldots, P_t$ intersect $\{x_1, \ldots, x_s, y_1, \ldots, y_s\}$. Since $t \geq 3s$, without loss of generality, we can suppose that the paths $P_{b+1}, \ldots, P_s$ are disjoint from $\{x_1, \ldots, x_s, y_1, \ldots, y_s\}$. For each $i = b+1, \ldots, s$, let y'_i be the endpoint of P_i in X_1, and x'_i the endpoint of P_i in X_2. For each $i = b+1, \ldots, s$, let $d_i^1 = d_1^-$ and $d_i^2 = d_i - d_1^- - |E(P_i)|$. Notice that by assumption we have $d_1^- + d_2^- + 3 = d^- \leq d_i \leq d^+ = \min(d_1^+, d_2^+)$ which combined with $|E(P_i)| \leq 3$ implies that $d_1^- \leq d_i^1 \leq d_1^+$ and $d_2^- \leq d_i^2 \leq d_2^+$.

Apply the definition of (X_1, W_1) being a (s_1, d_1^-, d_1^+)-linked system in order to find paths $Q_1, \ldots, Q_a, Q_{b+1}^1, \ldots, Q_s^1$ with Q_i a length d_i path from x_i to y_i internally contained in W_1, and Q_i^1 a length d_i^1 path from x_i to y'_i internally contained in W_1. Similarly, apply the definition of (X_2, W_2) being a (s_2, d_2^-, d_2^+)-linked system to find paths $Q_{a+1}, \ldots, Q_b, Q_{b+1}^2, \ldots, Q_s^2$ with Q_i a length d_i path from x_i to y_i internally contained in W_2, and Q_i^2 a length d_i^2 path from x'_i to y_i internally contained in W_2. For $i = b+1, \ldots, s$, let $Q_i = Q_i^1 + P_i + Q_i^2$ to get a length $d_i = d_i^1 + d_i^2 + |E(P_i)|$ path going from x_i to y_i. Now the paths $Q_1, \ldots, Q_s$ are paths from $x_1, \ldots, x_s$ to $y_1, \ldots, y_s$ internally contained in $W_1 \cup W_2 \cup \bigcup_{i=1}^{t} P_t$ as in the definition of (s, d^-, d^+)-linked system. $\qquad\square$

11

Lemma 4.14. *Let G be a graph and $k, s, d^-, d^+ \in \mathbb{N}$. For $i = 1, \ldots, k$ suppose that we have a (s, d^-, d^+)-linked system (X_i, W_i) with $(X_i \cup W_i) \cap (X_j \cup W_j) = \emptyset$ for $i \neq j$. Suppose that we have a connected graph F with vertex set $\{1, \ldots, k\}$ such that for all $uv \in E(F)$ there is a family $\mathcal{P}_{uv}$ of t disjoint paths of length ≤ 3 from X_u to X_v internally outside $\bigcup_{i=1}^k X_i \cup W_i$ with $t \geq 15ks$. Then (X, W) is a $(s, k(d^- + 3), d^+)$-linked system for $X = X_1 \cup \cdots \cup X_k$ and $W = W_1 \cup \cdots \cup W_k \cup \bigcup_{e \in H} \mathcal{P}_e$.*

Proof. Without loss of generality we can suppose that F is a tree with edges $e_2, \ldots, e_k$, and that the vertices of F are ordered so that for each i, the edge e_i goes from vertex i to some vertex in $\{1, \ldots, i-1\}$. Notice that this ensures that the induced subgraph $F[\{1, \ldots, i\}]$ is a tree for every i.

For all $e_i \in E(F)$, choose a subfamily $\mathcal{P}'_{e_i} \subseteq \mathcal{P}_{e_i}$ with $|\mathcal{P}'_{e_i}| = 3s$ such that the paths in $\mathcal{P}'_{e_i}$ are disjoint from those in $\mathcal{P}'_{e_j}$ for $i \neq j$. This is done by choosing the paths in $\mathcal{P}'_{e_i}$ one by one for each i always choosing them to be disjoint from $\bigcup_{j=2}^{i-1} \bigcup_{P \in \mathcal{P}'_{e_j}} P$. This is possible since $|\bigcup_{j=2}^{i-1} \bigcup_{P \in \mathcal{P}'_{e_j}} P| \leq 12is$ (using the fact that the paths in all $\mathcal{P}_{e_j}$ have length ≤ 3), and since there are $t \geq 15ks > 12is + 3s$ paths in P_{e_i} which are all disjoint.

We will use induction on i to prove that "$(X_1 \cup \cdots \cup X_i, W_1 \cup \cdots \cup W_i \cup \bigcup_{j=2}^i \mathcal{P}'_{e_j})$ is a $(s, i(d^- + 3), d^+)$-linked system." The initial case "$i = 1$" follows from (X_1, W_1) being a (s, d^-, d^+)-linked system. Suppose that $i \geq 2$, and (X', W') is a $(s, (i-1)(d^- + 3), d^+)$-linked system for $X' = X_1 \cup \cdots \cup X_{i-1}$ and $W' = W_1 \cup \cdots \cup W_{i-1} \cup \bigcup_{j=2}^{i-1} \mathcal{P}'_{e_j}$.

By construction of $\mathcal{P}'_{e_i}$ and the initial assumption that paths in $\mathcal{P}_{e_i}$ are internally disjoint from $\bigcup_{j=1}^k X_j \cup W_j$ we have that paths in $\mathcal{P}'_{e_i}$ are internally disjoint from $X' \cup W'$ and $X_i \cup W_i$. From the lemma's assumptions, for $a < b$ we have $(X_a \cup W_a) \cap (X_b \cup W_b) = \emptyset$ and we know that paths in $\mathcal{P}_{e_a}$ are disjoint from $X_b \cup W_b$. These imply $(X' \cup W') \cap (X_i \cup W_i) = \emptyset$. Also, since $e_i \in E(F)$, we have that every path in $\mathcal{P}'_{e_i}$ goes from X' to X_i and has length ≤ 3. By Lemma 4.13, we have that $(X' \cup X_i, W' \cup W_i \cup \bigcup_{j=1}^i \mathcal{P}'_{e_j})$ is a $(\min(s, |\mathcal{P}'_{e_i}|/3), (i-1)(d^- + 3) + d^- + 3, d^+)$-linked system. Since $|\mathcal{P}'_{e_i}|/3 = s$, this completes the induction step. $\qquad\square$

We will need the well known folklore result that every tree T can be separated into two parts of size $\leq 2|T|/3$ with one vertex (see e.g. [5], Corollary 2.1.)

Lemma 4.15. *The vertices of any tree T can be partitioned into a vertex u and two disjoint sets T_a and T_b such that $|T_a|, |T_b| \leq 2n/3$ and there are no edges between T_a and T_b.* $\qquad\square$

The following lemma shows that if we have a 2-edge-coloured complete graph on $(k-1)(n-1) + m_1$ vertices whose colouring is close to Burr's extremal construction, then it either contains a red copy of T or a blue copy of $K_{m_1, \ldots, m_k}$

Lemma 4.16. *Suppose that we have numbers $n, k, \Delta, m_1, \ldots, m_k \in \mathbb{N}$ with $k \geq 3$, $m_1 \leq m_2 \leq \ldots \leq m_k$, n large enough relative to Δ, k and $n \geq 2(d+1)m_k$ where $d = 4 \cdot 10^{12} \frac{\log^4 n}{\log \log n}$.*

Let T be a tree with $|T| = n$ and $\Delta(T) \leq \Delta$. Let G be a graph with $(k-1)(n-1) + m_1$ vertices that has disjoint vertex sets $H_1, \ldots, H_{k-1}$ with $|H_i| \geq 0.9n$, such that there are no edges between H_i and H_j for all $i \neq j$. If G^c has no $K_{m_1, \ldots, m_k}$, then G contains a copy of T.

Proof. Fix $m = m_k$ and $r = \lceil 10^3 \log^2 n \rceil$. Notice that we have $n \geq 2(d+1)m$ and G^c has no K_m^k. If T has $\geq n/4r$ leaves, then since $n/4r \geq 13\Delta|K_{m_1, \ldots, m_k}| + 1$ we are done by Theorem 1.2. Therefore, by Lemma 2.2, we may assume that T has at least $n/4r$ bare paths of length r.

We first need the following claim.

Claim 4.17. *There exist $H_i' \subseteq H_i$ with $|H_i'| \geq 0.8n$ such that for all $S \subseteq H_i'$ with $|S| \leq m$, we have $|N_{H_i'}(S)| \geq 5\Delta|S|$ and for all $S \subseteq H_i'$ with $m \leq |S| \leq 2m$, we have $|N_{H_i'}(S)| \geq 2n/3 + 10\Delta m$.*

Proof. First observe that for each i, H_i^c has no copy of K_m^2, or else we could take such a copy together with m vertices from each $H_j : j \neq i$, to obtain a K_m^k in G^c, a contradiction. Thus for any $S \subseteq H_i$ with $m \leq |S| \leq 2m$ we have $|N_{H_i}(S)| \geq |H_i| - |S| - m \geq |H_i| - 3m \geq 0.8n$.

Now for each i, choose a maximal $X_i \subseteq H_i$ with $|X_i| \leq m - 1$ such that $|N_{H_i}(X_i)| < 5\Delta|X_i|$, and let $H_i' = H_i \backslash X_i$. Notice that we have $|H_i'| \geq |H_i| - m \geq 0.8n$ as required by the claim. Using $n \geq 2(d+1)m$ and the fact that n is sufficiently large relative to Δ, we have that for any $S \subseteq H_i'$ with $m \leq |S| \leq 2m$

$$|N_{H_i'}(S)| \geq |N_{H_i}(S)| - m \geq 0.8n - m \geq \frac{2}{3}n + 10\Delta m.$$

Finally, suppose for sake of contradiction that there exists $S \subseteq H_i'$ with $|S| \leq m$ such that $|N_{H_i'}(S)| < 5\Delta|S|$. Then we have $|N_{H_i}(X_i \cup S)| < 5\Delta|X_i \cup S|$ so that $m \leq |X_i \cup S| \leq 2m$ by maximality of X_i and hence

$$10\Delta m \geq 5\Delta|X_i \cup S| > |N_{H_i}(X_i \cup S)| \geq 0.8n,$$

a contradiction to $n \geq 2(d+1)m$ and n being sufficiently large relative to Δ. $\qquad\square$

Let $Z = G \backslash \bigcup_{i=1}^{k-1} H_i'$. Suppose there exists $v \in Z$ and $a \neq b$ such that $d_{H_a'}(v), d_{H_b'}(v) \geq \Delta$. Apply Lemma 4.15 to T in order to get a vertex u and two forests T_a and T_b with no edges between them and $|T_a|, |T_b| \leq 2n/3$. We think of the trees in the forests T_a and T_b as being rooted at the neighbours of u. Let $t_a, t_b \leq \Delta$ be the number of neighbors of u in T_a and T_b respectively. Now choose $X_a \subseteq H_a' \cap N(v)$ so that $|X_a| = t_a$ and $X_b \subseteq H_b' \cap N(v)$ so that $|X_b| = b$. We observe that for $i \in \{a, b\}$, for all $S \subseteq H_i'$ with $1 \leq |S| \leq m$, we have

$$|N_{H_i'}(S) \backslash X_i| \geq |N_{H_i'}(S)| - |X_i| \geq 5\Delta|S| - t_i \geq 4\Delta|S|. \tag{3}$$

Because of the claim and (3), H_a' satisfies the assumptions of Lemma 3.1 with $G = H_a$, $M = 2n/3$, $t = t_a$, $X = X_a$, and $\{T_{x_1}, \ldots, T_{x_t}\}$ the collection of trees in the forest T_a. Therefore we can apply Lemma 3.1 to H_a in order to find a copy of T_a with its trees rooted in X_a. By the same argument, H_b has a copy of T_b with its trees rooted in X_b. These copies of T_a and T_b together with the vertex v give a copy of T in G, so we are done.

Otherwise, for all $v \in Z$ there exists i_v such that for all $j \neq i_v$, $d_{H_j'}(v) < \Delta$. We partition G into $k-1$ parts via $G_i = H_i' \cup \{v \in Z : i_v = i\}$. Observe that for any $i \neq j$ and $S \subseteq G_i$, we have $|N(S) \cap H_j'| < \Delta|S|$. We claim that therefore G_i^c has no K_m^2. Indeed suppose without loss of generality that S_1 was a copy of K_m^2 in G_1^c. Then for $j = 2, \ldots, k-1$, observing that $|H_j' \backslash N(S_1)| \geq |H_j'| - |N(S) \cap H_j'| \geq 0.8n - 2\Delta m \geq m$, we can choose a set $S_j \subseteq H_j' \backslash N(S_1)$ of size m. Then $\bigcup_{i=1}^{k-1} S_i$ is a copy of K_m^k in G^c, a contradiction.

Now fix i and observe that since G_i^c has no K_m^2, we have that for any $S \subseteq G_i$ with $|S| \geq m$, $|N_{G_i}(S)| \geq |G_i| - |S| - m$. Now choose $Z_i \subseteq G_i$ with $|Z_i| \leq m-1$ maximal so that $|N_{G_i}(Z_i)| < d|Z_i|$ and let $G_i' = G_i \backslash Z_i$. Then we claim that for all $S \subseteq G_i'$ with $|S| \leq m$, $|N_{G_i'}(S)| \geq d|S|$. Indeed, otherwise we would have $|N_{G_i}(Z_i \cup S)| < d|Z_i \cup S|$, so by maximality of X this would imply $m \leq |Z_i \cup S| \leq 2m$. But then

$$2dm \geq d|Z_i \cup S| > |N(Z_i \cup S)| \geq n - |Z_i \cup S| - m \geq n - 3m,$$

13

a contradiction.

Now let $n'_i = |G'_i|$. If for some i, $n'_i \geq n$ then we can apply Lemma 4.8 to conclude that G'_i has a copy of T. Otherwise we have that $n'_i \leq n - 1$ for all $i \in [k-1]$, and therefore using $|G| = (n-1)(k-1) + m_1$ we conclude $\sum_{i=1}^{k-1} |Z_i| \geq m_1$. For each $j = 1, \ldots, k-1$, we observe that

$$\left| N\left(\bigcup_{i=1}^{k-1} Z_i \right) \cap H'_j \right| \leq |N(Z_j) \cap H'_j| + \sum_{i \neq j} |N(Z_i) \cap H'_j| \leq dm + k\Delta m,$$

and hence

$$\left| H'_j \backslash N\left(\bigcup_{i=1}^{k-1} Z_i \right) \right| \geq 0.8n - k\Delta m - dm \geq m.$$

Thus for each $j = 1, \ldots, k-1$ we can choose a set $S_j \subseteq H'_j \backslash N\left(\bigcup_{i=1}^{k-1} Z_i \right)$ of size $m_{j+1} \leq m$. But then by taking a subset $X \subseteq \bigcup_{i=1}^{k-1} Z_i$ of size m_1, we obtain that $X \cup \bigcup_{i=1}^{k-1} S_i$ is a copy of $K_{m_1, \ldots, m_k}$ in G^c, a contradiction. □

We can now complete the case $k \geq 3$ by using either Lemma 4.11 or Lemma 4.16.

Proof of Theorem 2.1 for $k \geq 3$. Fix $m = m_k$, $d = 4 \cdot 10^{12} \frac{\log^4 n}{\log \log n}$, $r = \lceil 10^3 \log^2 n \rceil$ and $y = \lceil \log n \rceil$. We can choose $C_{\Delta, k}$ such that n is sufficiently large relative to Δ, k and $n \geq 2(d+1)m$. Let G be a graph with $(k-1)(n-1) + m_1$ vertices such that G^c has no copy of $K_{m_1, \ldots, m_k}$. Notice that in particular G^c has no K_m^k. If T has at least $n/4r \geq 13\Delta m + 1$ leaves, then by Lemma 3.4 we are done. Otherwise, by Lemma 2.2 T has at least $n/4r$ disjoint bare paths of length r.

Claim 4.18. *There are disjoint sets $Q'_1, \ldots, Q'_{k-1}$ of size $\in [22yn/r, 23yn/r]$, and $W'_1, \ldots, W'_{k-1}$ of size $\in [20yn/r, 21yn/r]$ such that for all i, $W'_i \subseteq Q'_i$, Q'_i y-expands into W'_i, and there are no edges between Q'_i and Q'_j for $i \neq j$.*

Proof. Let $q = 23yn/r$ and $w = 21yn/r$. Since n is sufficiently large relative to k, Δ and $r = \lceil 10^3 \log^2 n \rceil$ we have $(n-1)(k-1) + m_1 \geq (k-2)(n + 13\Delta q) + q$. Therefore we can apply Lemma 4.9 to conclude that either G contains a copy of T so that we are done, or else there exists a copy of K_q^{k-1} in G^c. Label the parts of K_q^{k-1} by $Q_1, \ldots, Q_{k-1}$. Observe that clearly Q_i^c has no copy of K_m^2. For each i, let $W_i \subseteq Q_i$ be a set of size w. Now choose $X_i \subseteq Q_i$ with $|X_i| \leq m-1$ maximal so that $|N_{Q_i}(X_i) \cap W_i| < y|X_i|$ and let $Q'_i = Q_i \backslash X_i$, and $W'_i = W_i \backslash X_i$. We claim that for all $S \subseteq Q'_i$ with $|S| \leq m$, $|N_{Q'_i}(S) \cap W'_i| \geq y|S|$. Indeed, otherwise we would have $|N_{Q_i}(X_i \cup S) \cap W_i| < y|X_i \cup S|$ so that $m \leq |X_i \cup S| \leq 2m$ by maximality of X_i. But then since Q_i^c has no K_m^2,

$$2ym \geq y|X_i \cup S| > |N_{Q_i}(X_i \cup S) \cap W_i| \geq w - |X_i \cup S| - m \geq w - 3m,$$

a contradiction to $n \geq 2(d+1)m$. Note that since $m \leq yn/r$, we have $|Q'_i| \geq q - m \geq 22yn/r$ and $|W'_i| \geq w - m \geq 20yn/r$. We further conclude that Q'_i y-expands into W'_i. Indeed, since Q'^c_i does not have K_m^2 we have that for any $S \subseteq Q'_i$ with $m \leq |S| < \lceil \frac{w}{2y} \rceil$,

$$|N_{Q'_i}(S) \cap W'_i| \geq |W'_i| - |S| - m \geq w - 2m - \frac{w}{2y} \geq \frac{w}{2} \geq y|S|,$$

so the first condition holds. Moreover, since Q'^c_i does not have K_m^2 and $\lceil w/2y \rceil \geq m$, the second condition holds as well. □

Now let $M_i = Q_i' \setminus W_i'$ and note that $yn/r \leq |M_i| \leq 3yn/r$. For $i \neq j$, fix a maximal family $\mathcal{P}_{i,j}$ of $\leq 8kn/r$ vertex-disjoint paths of length ≤ 3 from M_i to M_j internally outside $R_1 = \bigcup_{i=1}^{k-1} Q_i'$. Let F be an auxiliary graph on $[k-1]$ with ij an edge whenever $|\mathcal{P}_{i,j}| = 8kn/r$. Let $R_2 = \bigcup_{i \neq j} P_{i,j}$ and $R = R_1 \cup R_2$. Note that $|R_1| \leq 23kyn/r$ and $|R_2| \leq 8k^3n/r$ so that $|R| \leq 24kyn/r$ (since $y \geq 8k^2$ as a consequence of n being sufficeintly large relative to k.) Now let $M_i' = M_i \setminus R_2$ and note that $|M_i'| \leq |M_i| \leq 3yn/r$ and

$$|M_i'| \geq |M_i| - |R_2| \geq yn/r - 8k^3n/r \geq 2n/r \geq m.$$

Note that $|Q_i'| \leq 23yn/r$, so $160 \log |Q_i'| / \log\log |Q_i'| \leq \log n \leq y$ and hence by Lemma 4.3 (iii), we have that Q_i' $160 \log |Q_i'| / \log\log |Q_i'|$-expands into W_i'. Moreover

$$|Q_i'| \geq 22\frac{yn}{r} \geq 21\frac{yn}{r} + \frac{n}{r} \geq |W_i'| + 2\frac{n}{2r},$$

so we may apply Corollary 4.6 with $s = n/2r$. Since

$$y \leq \frac{|W_i'|}{40(n/2r)} \qquad \text{and} \qquad 4\left\lceil \frac{\log |Q_i'|}{\log\log |Q_i'|} \right\rceil \leq 4\left\lceil \frac{\log n}{\log\log n} \right\rceil \leq \frac{y}{k} - 3,$$

we conclude that (M_i, W_i') is a $(n/2r, y/k - 3, y)$-linked system and hence so is (M_i', W_i'). We now consider two cases depending on whether F is empty or not.

Case 1: Suppose that F is not empty. Let F' be the largest connected component of F and let $k' = |F'| + 1$. Since F is not empty we have $k' \geq 3$. Let $G' = G \setminus \bigcup_{i \in F'}(M_i' \cup N(M_i'))$.

Case 1.1: Suppose that $|G'| \geq (k - k')(n + 13\Delta m) + m$. Then G'^c has no $K_m^{k-k'+1}$ or else we could take it together with subsets of $M_i' : i \in F'$ of size m to obtain a K_m^k in G^c, a contradiction. But then G' contains a copy of T by Lemma 4.9.

Case 1.2: Suppose that $|G'| < (k - k')(n + 13\Delta m) + m$. Then since $m = o(n)$, we have

$$\left| \bigcup_{i \in F'} M_i' \cup N(M_i') \right| > (k-1)(n-1) + m_1 - (k-k')(n + 13\Delta m) - m = (k'-1)(n-1)(1 - o(1)).$$

So if we let $Z = \bigcup_{i \in F'} N(M_i') \setminus R$, we obtain

$$|Z| \geq \left| \bigcup_{i \in F'} N(M_i') \right| - |R| \geq \left| \bigcup_{i \in F'} M_i' \cup N(M_i') \right| - \left| \bigcup_{i \in F'} M_i' \right| - |R|$$

$$\geq (k'-1)(n-1)(1 - o(1)) - 3\frac{kyn}{r} - 24\frac{kyn}{r}$$

$$\geq 0.99(k'-1)n.$$

Moreover, if we let $X = \bigcup_{i \in F'} M_i'$ then we claim $(Z \cup X)^c$ has no $K_m^{k'}$. Indeed, since $ij \notin E(F)$ for any $i \in F', j \notin F'$, we could take subsets of $M_i' : i \notin F'$ of size m, together with a copy of $K_m^{k'}$ in $(Z \cup X)^c$ to obtain a copy of K_m^k in G. Since F' is connected, Lemma 4.14 applied with $d^- = y/k - 3$, $d^+ = y$, $s = n/2r$, and $k = k'$ implies that (X, W) is a $(n/2r, y, y)$-linked system for $W = R_2 \cup \bigcup_{i \in F'} W_i'$. Thus we may apply Lemma 4.11 to conclude that G contains a copy of T.

Case 2: Suppose that F is empty. Note that if $ij \notin E(F)$ then we must have no edges between $M_i' \cup (N(M_i') \setminus R)$ and $M_j' \cup (N(M_j') \setminus R)$ by the maximality of the family of paths $\mathcal{P}_{i,j}$. Thus if we

15

define $H_i = N(M_i')\backslash R$, then $H_1, \ldots, H_{k-1}$ are disjoint and there are no edges between H_i and H_j, for all $i \neq j$. Fix some $i \in \{1, \ldots, k-1\}$. Since $|M_i'| \geq m$, we have that $(G\backslash(N(M_i') \cup M_i'))^c$ does not contain K_m^{k-1} or else we could take it together with a subset of M_i' of size m to obtain a K_m^k in G^c, a contradiction. Thus if $|G\backslash(N(M_i') \cup M_i')| \geq (k-2)(n+13\Delta m) + m$, then $G\backslash(N(M_i') \cup M_i')$ has a copy of T by Lemma 4.9, so we are done. Otherwise we have

$$\begin{aligned} |N(M_i') \cup M_i'| &\geq (n-1)(k-1) + m_1 - ((k-2)(n+13\Delta m) + m) \\ &= n - (k-2)(13\Delta m + 1) + m_1 - m \\ &= n(1 - o(1)), \end{aligned}$$

so that $|N(M_i')| \geq n(1 - o(1)) - 3yn/r = n(1 - o(1))$ and hence for n sufficiently large,

$$|H_i| \geq |N(M_i')| - |R| \geq n(1 - o(1)) - 24kyn/r \geq 0.9n.$$

This holds for all i, so we can apply Lemma 4.16 to conclude that G contains a copy of T. $\qquad\square$

5 Concluding Remarks

In this paper we determined the range in which bounded degree trees are H-good, up to logarithmic factors. However, we conjecture that these factors can be removed to obtain the following.

Conjecture 5.1. *For all Δ and k there exists a constant $C_{\Delta,k}$ such that for any tree T with maximum degree at most Δ and any H with $\chi(H) = k$ satisfying $|T| \geq C_{\Delta,k}|H|$, T is H-good.*

As mentioned in the introduction, Conjecture 5.1 is best possible up to the constant factor $C_{\Delta,k}$. Pokrovskiy and Sudakov [20] showed that Conjecture 5.1 holds for paths, and our Theorem 1.2 shows that Conjecture 5.1 holds for trees with linearly (in $|H|$) many leaves.

References

[1] P. Allen, G. Brightwell and J. Skokan. Ramsey-goodness and otherwise. *Combinatorica* **33** (2013), 125–160.

[2] S. Burr. Ramsey numbers involving graphs with long suspended paths. *J. London Math. Soc.* **24** (1981), 405–413.

[3] S. Burr and P. Erdős. Generalizations of a Ramsey-theoretic result of Chvátal. *J. Graph Theory* **7** (1983), 39–51.

[4] V. Chvátal. Tree-complete graph Ramsey number. *J. Graph Theory* **1** (1977), 93.

[5] F. Chung. Separator theorems and their applications. *Forschungsinst. für Diskrete Mathematik* (1989).

[6] D. Conlon, J. Fox, C. Lee and B Sudakov. Ramsey numbers of cubes versus cliques. *Combinatorica*, **36** (2016), 37–70.

[7] P. Erdős. Some remarks on the theory of graphs. *Bull. Am. Math. Soc.* **53** (1947), 292–294.

[8] P. Erdős, R. Faudree, C. Rousseau and R. Schelp. Tree-multipartite graph Ramsey numbers. *Graph Theory and Combinatorics - A Volume in Honor of Paul Erdős.* Bela Bollobás, editor, Academic Press (1984), 155–160.

[9] P. Erdős, R. Faudree, C. Rousseau and R. Schelp. Multipartite graph-sparse graph Ramsey numbers. *Combinatorica* **5** (1985), 311–318.

[10] P. Erdős, R. Faudree, C. Rousseau and R. Schelp. Multipartite graph-tree graph Ramsey numbers. *Ann. NY Acad. Sci.* **576** (1989), 146–154.

[11] P. Erdős and G. Szekeres. Some remarks on the theory of graphs. *Bull. Am. Math. Soc.* **53** (1947), 292–294.

[12] G. Fiz Pontiveros, S. Griffiths, R. Morris, D. Saxton and J. Skokan. The Ramsey number of the clique and the hypercube. *J. Lond. Math. Soc.* **89** (2014), 680–702.

[13] J. Friedman and N. Pippenger. Expanding graphs contain all small trees. *Combinatorica* **7** (1987), 71–76.

[14] Z. Füredi and M. Simonovits. The history of degenerate (bipartite) extremal graph problems. *Erdős Centennial.* Springer Berlin Heidelberg, (2013) 169–264.

[15] P. Haxell. Tree embeddings. *J Graph Theory* **36** (2001), 121–130.

[16] M. Krivelevich. Embedding spanning trees in random graphs. *SIAM J. Discrete Math.* **24** (2010), 1495–1500.

[17] R. Montgomery. Embedding bounded degree spanning trees in random graphs. arXiv preprint arXiv:1405.6559 (2014).

[18] V. Nikiforov. The cycle-complete graph Ramsey numbers. *Combin. Probab. Comput.* **14** (2005), 349–370.

[19] V. Nikiforov and C. Rousseau. Ramsey goodness and beyond. *Combinatorica* **29** (2009), 227–262.

[20] A. Pokrovskiy and B. Sudakov. Ramsey goodness of paths. *J. Combinatorial Theory Ser. B* **122** (2017), 384–390.

[21] A. Pokrovskiy and B. Sudakov. Ramsey goodness of cycles. preprint.

[22] R. J. Wilson. **Introduction to graph theory**. Fourth edition. Longman, Harlow, 1996.

Appendix

Our goal will be to prove Lemma 3.1. This is a generalization of Haxell's theorem [15], and the proof follows the method of Friedman and Pippenger [13]. The idea is to prove a stronger statement from which Lemma 3.1 will follow as a corollary. For this, we will also need a slightly different definition of neighborhood. For a vertex x in a graph G, let $\Gamma(x) = N(x)$ be the neighborhood of x and for a set of vertices S in G, define $\Gamma(S) = \bigcup_{x \in S} \Gamma(x)$. Also, for a tree T rooted at v, we define $d_{root}(T) = d_T(v)$.

Lemma 5.2. *Let Δ, M, t and m be given. Let $X = \{x_1, \ldots, x_t\}$ be a set of vertices in a graph G. Suppose that we have rooted trees $T_{x_1}, \ldots, T_{x_t}$ satisfying $\sum_{i=1}^{t} |T_{x_i}| \leq M$ and $\Delta(T_{x_i}) \leq \Delta$ for all i. Suppose that for all S with $m \leq |S| \leq 2m$ we have $|\Gamma(S)| \geq M + 10\Delta m$, and for S with $|S| \leq m$ we have*

$$|\Gamma(S) \setminus X| \geq 4\Delta|S \setminus X| + \sum_{x \in S \cap X} \Big(d_{root}(T_x) + \Delta\Big). \tag{4}$$

Then we find disjoint copies of the trees $T_{x_1}, \ldots, T_{x_t}$ in G such that for each i, T_{x_i} is rooted at x_i. In addition for all $S \subseteq G$ with $|S| \leq m$, we have

$$\Big|\Gamma(S) \setminus \big(T_{x_1} \cup \cdots \cup T_{x_t}\big)\Big| \geq \Delta|S|. \tag{5}$$

Proof. The proof is by induction on $\sum_{i=1}^{t} e(T_{x_i})$. The initial case is when each tree is just a single vertex which holds by embedding T_{x_i} to x_i. Then (5) holds as a consequence of (4). Now suppose that the lemma holds for all families of trees with $\sum_{i=1}^{t} e(T_{x_i}) < e$ and we have a family with $\sum_{i=1}^{t} e(T_{x_i}) = e > 0$. Without loss of generality, we may assume that $e(T_{x_1}) \geq 1$. Let r be the root of T_{x_1} and c one of its children. For every $v \in \Gamma(x_1)$ we define a set $X^v = X \cup \{v\}$ and a corresponding family of rooted trees $\{T_x^v : x \in X^v\}$ as follows. Let $T_{x_1}^v$ be the subtree of T_{x_1} rooted at r formed by deleting c and its children. Let T_v^v be the subtree of T_{x_1} rooted at c formed by c and its children. For all $x \in X^v - x_1 - v$, let $T_x^v = T_x$.

Suppose that there is a vertex $v \in \Gamma(x_1) \setminus X$ such that the set X^v together with the family of trees $\{T_x^v : x \in X^v\}$ satisfy the following for every $C \subseteq G$ with $|C| \leq m$.

$$|\Gamma(C) \setminus X^v| \geq 4\Delta|C \setminus X^v| + \sum_{x \in C \cap X^v} \Big(d_{root}(T_x^v) + \Delta\Big).$$

Then, by induction we have an embedding of $T_{x_1}^v, \ldots, T_{x_t}^v, T_v^v$ into G which satisfies (5). By adding the edge $x_1 v$, we can join the trees $T_{x_1}^v$ and T_v^v in order to obtain a copy of T_{x_1} rooted at x_1. This gives an embedding of $T_{x_1}, \ldots, T_{x_t}$ into G which satisfies (5).

Otherwise, for every $v \in \Gamma(x_1) \setminus X$, there is a set C_v with $|C_v| \leq m$ and

$$|\Gamma(C_v) \setminus X^v| \leq 4\Delta|C_v \setminus X^v| + \sum_{x \in C_v \cap X^v} \Big(d_{root}(T_x^v) + \Delta\Big) - 1. \tag{6}$$

Notice that taking $S = \{x_1\}$, (4) implies that x_1 has at least one neighbour outside of X. Define a set of vertices S to be *critical* if it has order $\leq m$ and equality holds in (4).

Claim 5.3. *For every $v \in \Gamma(x_1) \setminus X$, the set C_v is critical, and also $v \in \Gamma(C_v)$ and $x_1 \notin C_v$.*

Proof. Notice that the following hold.

$$|\Gamma(C_v) \setminus X| - 1 \leq |\Gamma(C_v) \setminus X^v|, \tag{7}$$

$$4\Delta|C_v \setminus X^v| + \sum_{x \in C_v \cap X^v} \Big(d_{root}(T^v(x)) + \Delta\Big) \leq 4\Delta|C_v \setminus X| + \sum_{x \in C_v \cap X} \Big(d_{root}(T(x)) + \Delta\Big). \tag{8}$$

Adding (6), (7), (8), and (4) applied with $S = C_v$ gives "$0 \leq 0$" which implies that equality holds in each of these inequalities. In particular equality holds in (4), which implies that C_v is critical. For equality in (7) to hold, we must have $v \in \Gamma(C_v)$. For equality in (8) to hold, we must have $x_1 \notin C_v$ (since $d_{root}(T_{x_1}^v) = d_{root}(T_{x_1}) - 1$.) $\qquad\square$

We remark that the above proof also gives $v \notin C_v$, although this will not be needed in the proof. We'll also need the following claim.

Claim 5.4. *For two critical sets S and T, the union $S \cup T$ is critical.*

Proof. First we show that the reverse of the inequality (4) holds for $S \cup T$. We have the following

$$|\Gamma(S) \setminus X| = 4\Delta|S \setminus X| + \sum_{x \in S \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big). \tag{9}$$

$$|\Gamma(T) \setminus X| = 4\Delta|T \setminus X| + \sum_{x \in T \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big). \tag{10}$$

$$|\Gamma(S \cap T) \setminus X| \geq 4\Delta|S \cap T \setminus X| + \sum_{x \in S \cap T \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big). \tag{11}$$

Equations (9) and (10) come from S and T being critical, whereas (11) is just (4) applied to $S \cap T$ (which is smaller than m since S is critical.) Also, note that by inclusion-exclusion, we have

$$|S \cup T \setminus X| = |S \setminus X| + |T \setminus X| - |S \cap T \setminus X|, \tag{12}$$

$$\sum_{x \in (S \cup T) \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big) = \sum_{x \in S \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big) + \sum_{x \in T \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big)$$
$$- \sum_{x \in S \cap T \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big). \tag{13}$$

Moreover, we observe that

$$|\Gamma(S \cup T) \setminus X| = |(\Gamma(S) \cup \Gamma(T)) \setminus X|,$$
$$|\Gamma(S \cap T) \setminus X| \leq |(\Gamma(S) \cap \Gamma(T)) \setminus X|,$$

which together with inclusion-exclusion implies

$$|\Gamma(S \cup T) \setminus X| \leq |\Gamma(S) \setminus X| + |\Gamma(T) \setminus X| - |\Gamma(S \cap T) \setminus X|. \tag{14}$$

Plugging (9), (10), and (11) into (14), and then using (12) and (13) gives

$$|\Gamma(S \cup T) \setminus X| \leq 4\Delta|S \cup T \setminus X| + \sum_{x \in (S \cup T) \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big). \tag{15}$$

Since both S and T are critical we have $|S \cup T| \leq 2m$, which together with (15) implies that $|\Gamma(S \cup T)| \leq |X| + |\Gamma(S \cup T) \setminus X| \leq |X| + 8\Delta m < M + 10\Delta m$. By the assumption of the lemma we have $|S \cup T| \leq m$. Therefore (4) holds for the set $S \cup T$ which together with (15) implies that $S \cup T$ is critical. $\square$

Let $C = \bigcup_{v \in \Gamma(x_1) \backslash X} C_v$. By the two claims, we have that C is critical. Since from the first claim $\Gamma(x_1) \backslash X \subseteq \Gamma(C)$ and $x_1 \notin C$, we have that

$$
\begin{aligned}
|\Gamma(C \cup \{x_1\}) \setminus X| &= |\Gamma(C) \setminus X| \\
&= 4\Delta|C \setminus X| + \sum_{x \in C \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big) \\
&< 4\Delta|C \setminus X| + \sum_{x \in C \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big) + d_{root}(T(x_1)) + \Delta \\
&= 4\Delta|(C \cup \{x_1\}) \setminus X| + \sum_{x \in (C \cup \{x_1\}) \cap X} \Big(d_{root}\big(T(x)\big) + \Delta \Big).
\end{aligned}
$$

By (4) we have that $|C \cup \{x_1\}| > m$, which combined with C being critical means that $|C \cup \{x_1\}| = m + 1$. But then $|\Gamma(C \cup \{x_1\})| \le |X| + |\Gamma(C \cup \{x_1\}) \setminus X| \le |X| + 8\Delta m$ contradicts the assumption of the lemma that $|\Gamma(C \cup \{x_1\})| \ge M + 10\Delta m$. $\qquad\square$

Proof of Lemma 3.1. Note that since $|\Gamma(S)| \ge |N(S)|$ and $\sum_{x \in S \cap X} (d_{root}(T_x) + \Delta) \le 4\Delta|S \cap X|$ for all S, we may apply Lemma 5.2 to obtain copies of $T_{x_1}, \ldots, T_{x_t}$ rooted at $x_1, \ldots, x_t$ respectively so that (5) holds for all S with $|S| \le m$. In particular, if $S \subseteq T_{x_1} \cup \ldots \cup T_{x_t}$ and $|S| \le m$ then

$$
|N(S) \setminus (T_{x_1} \cup \ldots \cup T_{x_t})| = |\Gamma(S) \setminus (T_{x_1} \cup \ldots \cup T_{x_t})| \ge \Delta|S|. \qquad\square
$$