

Ramsey numbers of trees

Richard Montgomery*

Matías Pavez-Signé†

Jun Yan‡

September 9, 2025

Abstract

We show that there exists a constant $c > 0$ such that every n -vertex tree T with $\Delta(T) \leq cn$ has Ramsey number $R(T) = \max\{2t_1, t_1 + 2t_2\} - 1$, where $t_1 \geq t_2$ are the sizes of the bipartition classes of T . This improves an asymptotic result of Haxell, Łuczak, and Tingley from 2002, and shows that, though Burr’s 1974 conjecture on the Ramsey numbers of trees has long been known to be false for certain ‘double stars’, it is true for trees with up to small linear maximum degrees.

1 Introduction

The Ramsey number of a graph G , denoted by $R(G)$, is the smallest positive integer N such that every red/blue edge colouring of the complete N -vertex graph K_N contains a monochromatic copy of G . The existence of $R(G)$ follows from Ramsey’s foundational result in 1930 [32], but determining good bounds on Ramsey numbers has since proved extremely challenging. The most notorious and natural case is where G is the complete n -vertex graph K_n . The famous upper bound by Erdős and Szekeres [14] in 1935 and lower bound by Erdős [11] in 1947 showed that the rate of growth of $R(K_n)$ is exponential in n . Since then, these bounds saw only modest improvements until the recent remarkable breakthrough of Campos, Griffiths, Morris, and Sahasrabudhe [8] finally gave an exponential improvement to the upper bound of Erdős and Szekeres (see also [1, 18]).

Away from complete graphs, the sparser G is, the more ambitious we can reasonably be in bounding $R(G)$. For example, a classical result of Chvátal, Rödl, Szemerédi, and Trotter [9] from 1983 states that the Ramsey number of every n -vertex graph with bounded maximum degree is linear in n . That is, for every Δ , there is some c_Δ such that any n -vertex graph G with maximum degree at most Δ satisfies $R(G) \leq c_\Delta n$. Burr and Erdős [6] had conjectured in 1975 that, moreover, this should hold with maximum degree replaced by degeneracy, and this was proved by Lee [26] in 2017.

There are not many graphs G for which we can muster any hope of determining $R(G)$ exactly. Aside from the smallest of graphs, the main candidates are trees and cycles. In 1967, Gerencsér and Gyárfás [16] determined the Ramsey number of the n -vertex path P_{n-1} , showing that $R(P_{n-1}) = \lfloor 3n/2 \rfloor - 1$. For the n -vertex star $K_{1,n-1}$, note that $R(K_{1,n-1}) - 1$ is the size of the largest graph such that both it and its complement have maximum degree at most $n - 2$. Thus, as shown by Harary [20] in 1972, $R(K_{1,n-1}) = 2n - 2$ if n is even, and $R(K_{1,n-1}) = 2n - 3$ if n is odd. The Ramsey number of the n -vertex cycle C_n is known due to independent work in the early 1970’s by Bondy and Erdős [4], Faudree and Schelp [15], and Rosta [33], where we have $R(C_3) = R(C_4) = 6$, $R(C_n) = 2n - 1$ for odd $n \geq 5$, and $R(C_n) = 3n/2 - 1$ for even $n \geq 6$.

*Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK. richard.montgomery@warwick.ac.uk

†Departamento de Ingeniería Matemática, Universidad de Chile, and Centro de Modelamiento Matemático, CNRS IRL2807, mpavez@dim.uchile.cl

‡Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK. jun.yan@warwick.ac.uk.

RM and MPS supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No. 947978). MPS additionally supported by ANID-FONDECYT Regular grant No. 1241398 and by ANID Basal Grant CMM FB210005. JY supported by the Warwick Mathematics Institute Centre for Doctoral Training, and by funding from the UK EPSRC (Grant number: EP/W523793/1).

For a general tree T , two constructions of Burr [5] in 1974 (see Figure 1) show that if T has bipartition classes of sizes t_1 and t_2 , where $t_1 \geq t_2$, then

$$R(T) \geq \max\{t_1 + 2t_2, 2t_1\} - 1. \quad (1.1)$$

From the results quoted above, this bound is tight when T is a path or a star of odd size, and Burr conjectured [5] that this bound is tight for every tree T with $t_1 \geq t_2 \geq 2$. However, this was disproved in 1979 by Grossman, Harary, and Klawe [17] for certain trees called *double stars*. For each $t_1 \geq t_2 \geq 2$, let S_{t_1, t_2} be the tree formed by joining the central vertices of the stars K_{1, t_1-1} and K_{1, t_2-1} with an edge, noting that S_{t_1, t_2} has bipartition classes with sizes t_1 and t_2 . Grossman, Harary, and Klawe [17] showed that if $t_1 \geq 3t_2 - 2$, then $R(S_{t_1, t_2}) = 2t_1$, and thus the bound at (1.1) is off by 1 in this case. In 1982, Erdős, Faudree, Rousseau, and Schelp [12] attempted to rescue Burr's conjecture by conjecturing that the bound at (1.1) is tight when $t_1 = 2t_2$. However, this was strongly disproved by Norin, Sun, and Zhao [29] in 2016, who showed in particular that $R(S_{2t, t}) \geq (4.2 - o(1))t$ (see also [10]), and thus the bound at (1.1) can be off by a multiplicative factor.



Figure 1: Burr's extremal constructions for $R(T)$ when T is a tree with bipartition classes of sizes $t_1 \geq t_2$. **I:** Disjoint blue cliques on U_1 and U_2 , with $|U_1| = t_1 + t_2 - 1$, $|U_2| = t_2 - 1$, and every edge between U_1 and U_2 coloured red. Any connected blue subgraph has at most $t_1 + t_2 - 1 < |T|$ vertices, and any connected red subgraph is bipartite with fewer than t_2 vertices in one class.

II: Disjoint blue cliques on U_1 and U_2 , with $|U_1| = |U_2| = t_1 - 1$, and every edge between U_1 and U_2 coloured red. Any connected blue subgraph has at most $t_1 - 1 < |T|$ vertices, and any connected red subgraph is bipartite with fewer than t_1 vertices in each class.

Thus, in both **I** and **II** there is no monochromatic copy of T .

All the known counterexamples to Burr's conjecture, however, have large maximum degree, and thus the bound at (1.1) may still be tight for trees with small maximum degree. Towards this, Haxell, Łuczak, and Tingley [21] showed in 2002 that the bound at (1.1) is approximately tight for trees with up to small linear maximum degree. That is, they showed that, for every $\varepsilon > 0$, there exists some $c > 0$ such that any n -vertex tree T with maximum degree $\Delta(T) \leq cn$ and bipartition classes of sizes $t_1 \geq t_2$ satisfies $R(T) \leq (1 + \varepsilon) \max\{t_1 + 2t_2, 2t_1\}$.

In this paper, we will show that Burr's bound at (1.1) is tight for all trees with up to small linear maximum degree, as follows.

Theorem 1.1. *There exists a constant $c > 0$ such that the following holds. Any n -vertex tree T with $\Delta(T) \leq cn$ and bipartition classes of sizes $t_1 \geq t_2$ satisfies $R(T) = \max\{2t_1, t_1 + 2t_2\} - 1$.*

The existence of such a constant c in Theorem 1.1 answers in the positive a question asked explicitly by Stein [34] in 2020. Our value of c is very small due to the use of regularity methods, and is likely very far from optimal. It follows from the double star examples given in [29] that c cannot be improved beyond $7/11 + o(1)$. For a tree T with large maximum degree we do not have a good conjecture for the exact value of $R(T)$, though Burr and Erdős [7] conjectured in 1976 that for any n -vertex tree T , $R(T) \leq 2n - 2$ when n is even and $R(T) \leq 2n - 3$ when n is odd, or in other words $R(T) \leq R(K_{1, n-1})$. In 2011, Zhao [36] showed that this is true for all large even n , as a consequence of his resolution for large n of Loeb's $n/2 - n/2 - n/2$ conjecture [13]. Burr and Erdős's conjecture follows directly from the Erdős-Sós conjecture, and thus for large n follows from the proof of the Erdős-Sós conjecture for large trees announced by Ajtai, Komlós, Simonovits, and Szemerédi in the early 1990s (see [30] for a discussion of this result). A wide-ranging discussion of further results on Ramsey numbers can be found in the dynamic survey by Radziszowski [31].

To prove Theorem 1.1 we will conduct what is known as a stability analysis. When the red/blue coloured host graph is not close to one of the extremal constructions in Figure 1, we will develop the work of Haxell, Luczak, and Tingley [21], and find a monochromatic copy of our tree T using methods involving Szemerédi’s regularity lemma. If instead the colouring is close to an extremal construction, then we will analyse the structure more closely to still find a monochromatic copy of T , often using randomised embeddings. We call these two parts of the proof the ‘stability part’ and ‘extremal part’, and both of them will be rather involved. For the stability part of the argument, we will be able to start from a certain monochromatic structure found by Haxell, Luczak, and Tingley [21] in the reduced graph, but will still need to do much more work to cover all the non-extremal cases, with the main problem to overcome being a deficit of vertices in this initial structure. For the extremal part, proving Theorem 1.1 when the colouring approximates either extremal construction turns out to be surprisingly delicate. For instance, looking at the first extremal construction in Figure 1, one might expect that a blue copy of an n -vertex tree T with $\Delta(T) \leq cn$ would appear once U_1 contains one more vertex, even if a small linear proportion of the edges within U_1 are red. However, a famous example of Komlós, Sárközy and Szemerédi [23] shows that this is not true. As such, proving Theorem 1.1 in the extremal part will require a careful consideration of both the structure of the tree T and the presence of edges of the ‘wrong’ colour in the extremal colouring, to decide to where, and in which colour, the tree should be embedded in different cases.

This paper is organised as follows. In Section 2 we first give a brief overview of our proof of Theorem 1.1, focusing on how it can be divided into the stability part (Sections 4 and 5) and the extremal part (Sections 6 and 7), then collect all the basic notations and preliminary results. Then, in Section 3, we give a detailed outline of the stability part of our proof of Theorem 1.1. In Section 4, we prove a series of technical regularity embedding lemmas, each of which allows us to embed a monochromatic copy of the tree T into a red/blue coloured reduced graph that contains a certain suitable structure. In Section 5, we use these embedding lemmas to move through 4 stages of embedding attempts, and eventually conclude either that we can find a monochromatic copy of T using regularity, or that the reduced graph and thus the original graph are both extremal, in the sense that they approximate one of the extremal constructions. Depending on which of the two extremal constructions our original graph approximates, we show in Section 6 and Section 7 respectively that a monochromatic copy of T can still be found.

2 Proof overview and preliminaries

In this section, we begin by giving a short overview of our proof of Theorem 1.1 in Section 2.1, specifically on how it divides into the stability part and the extremal part, and formalising what it means for a colouring to approximate an extremal construction. Then, we record the basic notations we use in Section 2.2, and collect a series of preliminary results in Sections 2.3 to 2.6.

2.1 Division of the proof of Theorem 1.1 into stability and extremal parts

We start by recapping the situation in Theorem 1.1. Let t_1, t_2 be positive integers so that $n = t_1 + t_2$ and $t_1 \geq t_2$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 , such that $|V_1| = t_1$ and $|V_2| = t_2$. Note that $\Delta(T) \leq cn$ implies $t_2 \geq c^{-1}$. Let $N = \max\{t_1 + 2t_2, 2t_1\} - 1$ and let G be a red/blue coloured complete graph on N vertices. Our aim, then, is to find a monochromatic copy of T in G .

Our proof of Theorem 1.1 consists of two main parts, the stability part (Sections 4 and 5) and the extremal part (Sections 6 and 7). In the stability part, we show that either G contains a monochromatic copy of T , or G is close to one of the two extremal constructions in Figure 1. Then in the extremal part we show that a monochromatic copy of T still exists even if G approximates an extremal construction. These parts are quite separate, and both are quite involved, so we will sketch their proofs later (in Section 3 for the stability part, and in Sections 6.1 and 7.1 for the two different cases of the extremal part). Here, we state the main results for both parts, and put them together to prove Theorem 1.1. We start with the following definition of what it means for G to be close to one of the extremal constructions.

Definition 2.1. Let $0 < \mu < 1$ and let G be a red/blue coloured complete graph. We say G is *Type I* (μ, t_1, t_2) -extremal if with $n = t_1 + t_2$, there are disjoint subsets $U_1, U_2 \subset V(G)$ such that

- $|U_1| \geq (1 - \mu)n$ and $|U_2| \geq (1 - \mu)t_2$,
- for every $u \in U_1$, $d_{\text{red}}(u, U_1) \leq \mu n$, and
- for every $i \in [2]$ and every $u \in U_i$, $d_{\text{blue}}(u, U_{3-i}) \leq \mu n$,

or with red and blue swapped. On the other hand, we say G is *Type II* (μ, t_1, t_2) -extremal if with $n = t_1 + t_2$, there are disjoint subsets $U_1, U_2 \subset V(G)$ such that

- $|U_1|, |U_2| \geq (1 - \mu)t_1$, and
- for each $i \in [2]$ and $u \in U_i$, $d_{\text{red}}(u, U_i) \leq \mu n$ and $d_{\text{blue}}(u, U_{3-i}) \leq \mu n$,

or with red and blue swapped. If G is either or Type I or Type II (μ, t_1, t_2) -extremal, we say it is (μ, t_1, t_2) -extremal.

Using this, we can now state the main result of the stability part of our proof, as follows.

Theorem 2.2. *Let $1/n \ll c \ll \mu \ll 1$ and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$ and $t_1 \geq t_2$. Let G be a red/blue coloured complete graph with $\max\{t_1 + 2t_2, 2t_1\} - 1$ vertices. Then, at least one of the following is true.*

- G contains a monochromatic copy of every n -vertex tree T with $\Delta(T) \leq cn$ and bipartition class sizes t_1 and t_2 .
- G is Type I (μ, t_1, t_2) -extremal.
- $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Theorem 2.2 will be proved in Section 5, and reduces the proof of Theorem 1.1 to the following two results, which find a monochromatic copy of T even when G is close to one of the extremal constructions. We will prove them in Sections 6 and 7, respectively.

Theorem 2.3. *Let $1/n \ll c \ll \mu \ll 1$ and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$ and $t_1 \geq t_2$. If G is a Type I (μ, t_1, t_2) -extremal graph with $\max\{t_1 + 2t_2, 2t_1\} - 1$ vertices, then G contains a monochromatic copy of every n -vertex tree T with $\Delta(T) \leq cn$ and bipartition class sizes t_1 and t_2 .*

Theorem 2.4. *Let $1/n \ll c \ll \mu \ll 1$ and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$ and $t_1 \geq (2 - \mu)t_2$. If G is a Type II (μ, t_1, t_2) -extremal graph with $\max\{t_1 + 2t_2, 2t_1\} - 1$ vertices, then G contains a monochromatic copy of every n -vertex tree T with $\Delta(T) \leq cn$ and bipartition class sizes t_1 and t_2 .*

Given Theorems 2.2–2.4, Theorem 1.1 follows essentially immediately, but we will conclude this overview by formally making this deduction, as follows.

Proof of Theorem 1.1. Let μ satisfy $c \ll \mu \ll 1$. If G is not (μ, t_1, t_2) -extremal, then G contains a monochromatic copy of T by Theorem 2.2. If G is Type I (μ, t_1, t_2) -extremal, then G contains a monochromatic copy of T by Theorem 2.3, while if $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal, then G contains a monochromatic copy of T by Theorem 2.4. $\square$

Finally, we remark that in these proofs, we can often assume that $t_1 \leq 2t_2 + 1$. Indeed, if $t_1 \geq 2t_2 + 2$, then we can take $2/c$ vertices in V_1 with degree 1, which are guaranteed to exist by Lemma 2.10, and attach $\lfloor (t_1 - 2t_2)/2 \rfloor$ new leaves to them, with none of them receiving more than cn new leaves. Let T' be the new tree obtained in this way, and note that the bipartition classes of T' have sizes $t'_1 = t_1$ and $t'_2 = \lfloor t_1/2 \rfloor$, satisfying $t'_1 \leq 2t'_2 + 1$ and $\max\{t'_1 + 2t'_2, 2t'_1\} - 1 = 2t'_1 - 1 = 2t_1 - 1 = \max\{t_1 + 2t_2, 2t_1\} - 1$. Thus, the number of vertices in G remains unchanged, and it is clear that if G contains a monochromatic copy of T' , then G also contains a monochromatic copy of T .

2.2 Notation

For a positive integer $n \in \mathbb{N}$, we write $[n] = \{1, \dots, n\}$ and $[n]_0 = [n] \cup \{0\}$. We will use the standard hierarchy notation, that is, for $a, b \in (0, 1]$, we will use $a \ll b$ to mean that there exists a non-decreasing function $f : (0, 1] \rightarrow (0, 1]$ such that if $a \leq f(b)$ then the following statement holds. For $a, b \geq 1$, we write $a \ll b$ if $1/b \ll 1/a$. Hierarchies with more constants are defined in a similar way. For simplicity we will sometimes ignore floor and ceiling signs when doing so does not affect the argument.

Given a graph G , we use $V(G)$ and $E(G)$ to denote the set of vertices and edges of G , respectively, and write $|G| = |V(G)|$ and $e(G) = |E(G)|$. For not necessarily disjoint subsets $A, B \subset V(G)$, we denote the number of edges in G with one endpoint in A and one in B by $e(A, B)$. For a subset $S \subset V(G)$, we use $G[S]$ to denote the graph with vertex set S and all the edges from G with both endpoints in S , and we write $G - S$ for the graph $G[V(G) \setminus S]$. Given two disjoint subsets $S, S' \subset V(G)$, we use $G[S, S']$ to denote the bipartite graph with parts S and S' , and all edges of the form $ss' \in E(G)$ with $s \in S$ and $s' \in S'$.

For a vertex $v \in V(G)$, the set of neighbours of v is denoted by $N(v)$, and $d(v) = |N(v)|$ denotes the degree of v . The maximum degree and the minimum degree of G are denoted by $\Delta(G)$ and $\delta(G)$, respectively. Given a subset $S \subset V(G)$, its external neighbourhood is $N(S) = (\cup_{s \in S} N(s)) \setminus S$. For a vertex $v \in V(G)$ and subsets $S, U \subset V(G)$, we write $N(v, S) = N(v) \cap S$, $d(v, S) = |N(v, S)|$, and $N(U, S) = N(U) \cap S$. When working with more than one graph, we add subscripts to denote which graph we are working with. For example, $d_G(x)$ refers to the degree of x in the graph G .

Say G is a red/blue coloured graph if every edge in $E(G)$ is coloured with either red or blue. We let G_{red} and G_{blue} denote the graphs spanned by the red edges and the blue edges, respectively. For brevity, we write $d_{\text{red}}(x)$ instead of $d_{G_{\text{red}}}(x)$ and $d_{\text{blue}}(x)$ instead of $d_{G_{\text{blue}}}(x)$, and use similar notations for the red and blue neighbourhoods of a vertex or a set of vertices.

For $\mu \in [0, 1]$, a graph G is μ -almost complete if $\delta(G) \geq (1 - \mu)|G|$, and is μ -almost empty if $\Delta(G) \leq \mu|G|$. A bipartite graph H with bipartition classes A, B is μ -almost complete if $d(a) \geq (1 - \mu)|B|$ for every $a \in A$ and $d(b) \geq (1 - \mu)|A|$ for every $b \in B$, and is μ -almost empty if $d(a) \leq \mu|B|$ for every $a \in A$ and $d(b) \leq \mu|A|$ for every $b \in B$.

2.3 Concentration results

We will need the following well-known concentration results.

Lemma 2.5 (Chernoff's Bound [22, Corollary 2.3, Theorem 2.10]). *Let X be either a binomial random variable or a hypergeometric random variable. Then, for all $0 < \varepsilon \leq 3/2$,*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon \mathbb{E}[X]) \leq 2 \exp(-\varepsilon^2 \mathbb{E}[X]/3).$$

Lemma 2.6 (Azuma's Inequality [35, Lemma 4.2]). *Let $X_1, \dots, X_m$ be a sequence of random variables such that for each $i \in [m]$, there exist constants $a_i \in \mathbb{R}$ and $c_i > 0$ with $|X_i - a_i| \leq c_i$.*

- *If $\mathbb{E}[X_i | X_1, \dots, X_{i-1}] \geq a_i$ for every $i \in [m]$, then for every $t > 0$,*

$$\mathbb{P}(\sum_{i=1}^m (X_i - a_i) \leq -t) \leq \exp\left(-\frac{t^2}{2 \sum_{i=1}^m c_i^2}\right).$$

- *If $\mathbb{E}[X_i | X_1, \dots, X_{i-1}] \leq a_i$ for every $i \in [m]$, then for every $t > 0$,*

$$\mathbb{P}(\sum_{i=1}^m (X_i - a_i) \geq t) \leq \exp\left(-\frac{t^2}{2 \sum_{i=1}^m c_i^2}\right).$$

Lemma 2.7 (McDiarmid's Inequality [27, Lemma 1.2]). *Let $X_1, \dots, X_m$ be independent random variables taking values in a set Ω . Let $c_1, \dots, c_m \geq 0$ and suppose $f : \Omega^m \rightarrow \mathbb{R}$ is a function such that for every $i \in [m]$ and every $x_1, \dots, x_m, x'_i \in \Omega$, we have $|f(x_1, \dots, x_i, \dots, x_m) - f(x_1, \dots, x'_i, \dots, x_m)| \leq c_i$. Then, for all $t > 0$,*

$$\mathbb{P}\left(|f(X_1, \dots, X_m) - \mathbb{E}[f(X_1, \dots, X_m)]| \geq t\right) \leq 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^m c_i^2}\right).$$

2.4 Matchings in bipartite graphs

In many of our later tree embedding arguments, we will first embed all but a small set of vertices in T with degrees 1 or 2. To finish the embedding, the following well-known Hall's matching theorem and its generalisation are useful.

Lemma 2.8 (Hall's matching theorem [19, Theorem 1]). *Let G be a bipartite graph with bipartition classes A and B . If $|N(S)| \geq |S|$ for every $S \subset A$, then G contains a matching covering all vertices in A .*

Lemma 2.9 ([3, Corollary 11]). *Let G be a bipartite graph with bipartition classes A and B , and let $(f_a)_{a \in A}$ be a tuple of non-negative integers indexed by elements of A . Suppose that $|N(S)| \geq \sum_{a \in S} f_a$ for every $S \subset A$. Then, there exists a collection of vertex-disjoint stars $(S_a)_{a \in A}$ in G , such that for each $a \in A$, S_a is centred at a and has exactly f_a leaves.*

The conditions in Lemma 2.8 and Lemma 2.9 will both be referred to as *Hall's matching condition*.

2.5 Trees

We now record several useful results on tree embeddings and tree decompositions.

Lemma 2.10. *If T is an n -vertex tree with bipartition classes V_1 and V_2 such that $|V_1| = t_1$, $|V_2| = t_2$, and $t_1 \geq t_2$, then T contains at least $t_1 - t_2 + 1$ leaves in V_1 .*

Proof. Let L be the set of leaves of T in V_1 . Then

$$n - 1 = e(V_1, V_2) = \sum_{v \in V_1} d(v) \geq |L| + 2|V_1 \setminus L| = 2t_1 - |L|,$$

from which it follows that $|L| \geq 2t_1 - n + 1 = t_1 - t_2 + 1$. $\square$

A path P in a tree T is a *bare path* if all of its internal vertices have degree 2 in T . By the following well-known result, every tree has either many leaves or many bare paths.

Lemma 2.11 ([25, Lemma 2.1]). *Let $k, \ell, n \in \mathbb{N}$ and let T be an n -vertex tree with at most ℓ leaves. Then T contains a collection of at least $\frac{n}{k+1} - (2\ell - 2)$ vertex-disjoint bare paths, each of length k .*

In many of our tree embeddings, we will first divide the tree into two parts that are then embedded with different methods and different aims. For this, we use the following definition.

Definition 2.12. For a tree T , we say that subgraphs T_1, T_2 of T form a *decomposition* of T if they are edge-disjoint subforests of T such that $E(T) = E(T_1) \cup E(T_2)$.

We will use the following result to decompose a tree into two subtrees so that each subtree in the decomposition contains a large proportion of a set chosen in advance.

Lemma 2.13 ([28, Proposition 3.19]). *Let T be a tree and let $Q \subset V(T)$. Then, T has a decomposition into subtrees T_1 and T_2 with a unique common vertex such that $|Q \cap V(T_1)| \geq |Q|/3$ and $|Q \cap V(T_2)| \geq |Q|/3$.*

The following almost immediate corollary is obtained by taking $Q = V(T)$ in Lemma 2.13.

Corollary 2.14. *Every n -vertex tree T decomposes into subtrees T_1 and T_2 with a unique common vertex such that $\lceil n/3 \rceil \leq |T_1| \leq |T_2| \leq \lceil 2n/3 \rceil$.*

Proof. Apply Lemma 2.13 with $Q = V(T)$, we get a decomposition of T into subtrees T_1 and T_2 with a unique common vertex v such that $\lceil n/3 \rceil \leq |T_1| \leq |T_2| \leq n - \lceil n/3 \rceil + 1$. If n is congruent to 1 or 2 modulo 3, then $n - \lceil n/3 \rceil + 1 = \lceil 2n/3 \rceil$, so we are done. If $n = 3k$ for some integer $k \geq 1$, then the only situation where the result does not follow immediately is when $|T_1| = k$ and $|T_2| = 2k + 1$. Assume that this holds.

If $d_T(v, T_2) = 1$, let v' be the unique neighbour of v in T_2 , then $T_1 + vv'$ and $T_2 - v$ are subtrees decomposing T with a unique common vertex v' , and contains $k + 1$ and $2k$ vertices, respectively, as required. If $d_T(v, T_2) \geq 2$, let S be the smallest component in $T_2 - v$, so $1 \leq |S| \leq k$. Then $V(T_1) \cup S$ and $V(T_2) \setminus S$ induce two subtrees decomposing T with a unique common vertex v , and contain $k + 1 \leq k + |S| \leq 2k$ and $k + 1 \leq 2k + 1 - |S| \leq 2k$ vertices, respectively, finishing the proof. $\square$

The following results show that we can cut a tree into smaller subtrees using few vertices.

Lemma 2.15. *For every n -vertex tree T , there exists a vertex $v \in T$ so that each component of $T - v$ has size at most $n/2$.*

Proof. Choose an arbitrary vertex as the root of T . Let v be a vertex at a maximal distance from the root subject to the condition that the tree T' induced by v and all of its descendents has size at least $n/2$. By the choice of v , each component of $T' - v$ has size less than $n/2$. The only component of $T - v$ that is not a component of $T' - v$ is $T - T'$, which has size at most $n/2$ as $|T'| \geq n/2$. $\square$

Lemma 2.16 ([2, Proposition 4.1]). *Let $1/n \ll \xi \ll 1$ and let T be an n -vertex tree. Then, there exists a subset $X \subset V(T)$ with $|X| \leq 2\xi^{-1}$, such that every component of $T - X$ has size at most ξn .*

The following two results state that trees can be greedily embedded into graphs with large minimum degrees, and will be used throughout the paper without any further reference.

Lemma 2.17. *Let T be an n -vertex tree containing a vertex t . If G is a graph with $\delta(G) \geq n - 1$, then, for any vertex $v \in G$, there is a copy of T in G with t copied to v .*

Lemma 2.18. *Let T be a tree with bipartition classes V_1 and V_2 of sizes t_1 and t_2 , respectively. Suppose that G is a bipartite graph with bipartition classes U_1 and U_2 , such that*

- *every vertex in U_1 has at least t_2 neighbours in U_2 , and*
- *every vertex in U_2 has at least t_1 neighbours in U_1 .*

Then, for any $i \in [2]$ and any vertices $t \in V_i$ and $u \in U_i$, there exists a copy of T in G such that V_1 is copied to U_1 , V_2 is copied to U_2 , and t is copied to u .

2.6 Szemerédi's regularity lemma

Let G be a bipartite graph with bipartition classes A and B . For sets $X \subset A$ and $Y \subset B$, the *density* between X and Y is defined as

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}.$$

We say G is ε -regular if for every $X \subset A$ and every $Y \subset B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$, we have $|d(X, Y) - d(A, B)| \leq \varepsilon$. Furthermore, we say G is (ε, d) -regular if G is ε -regular and $d(A, B) \geq d$. The following results are standard.

Lemma 2.19. *Let $\varepsilon \leq 1/4$, and let G be a bipartite graph with bipartition classes A and B that is (ε, d) -regular. Suppose $X \subset A$ and $Y \subset B$ satisfy $|X| \geq \sqrt{\varepsilon}|A|$ and $|Y| \geq \sqrt{\varepsilon}|B|$, then $G[X, Y]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular.*

Lemma 2.20. *Let G be a bipartite graph with bipartition classes A and B that is (ε, d) -regular. Suppose $Y \subset B$ satisfies $|Y| \geq \varepsilon|B|$, then there are less than $\varepsilon|A|$ vertices $v \in A$ for which $d(v, Y) < (d - \varepsilon)|Y|$.*

Lemma 2.21. *Let G be a graph containing disjoint subsets $V_0, V_1, \dots, V_r \subset V(G)$, such that $G[V_0, V_i]$ is (ε, d) -regular for each $i \in [r]$. Let $U_i \subset V_i$ have size $|U_i| \geq \varepsilon|V_i|$ for each $i \in [r]$. Then, there are less than $\sqrt{\varepsilon}|V_0|$ vertices $v \in V_0$ such that $d(v, U_i) < (d - \varepsilon)|U_i|$ for at least $\sqrt{\varepsilon}r$ indices $i \in [r]$.*

The following colourful variant of Szemerédi's Regularity Lemma is well-known, and is the starting point of the stability part of our proof.

Theorem 2.22 (Coloured Regularity Lemma [24, Theorem 1.18]). *Let $1/k_2 \ll 1/k_1 \ll \varepsilon$. Every red/blue coloured graph G on $n \geq k_1$ vertices contains disjoint subsets $V_1, \dots, V_k \subset V(G)$ with $k_1 \leq k \leq k_2$ that satisfy the following.*

- (i) $|V(G) \setminus (V_1 \cup \dots \cup V_k)| \leq \varepsilon n$.
- (ii) $|V_1| = \dots = |V_k|$.

(iii) For all but at most εk^2 indices $1 \leq i < j \leq k$, both $G_{\text{red}}[V_i, V_j]$ and $G_{\text{blue}}[V_i, V_j]$ are ε -regular.

For technical reasons, we sometimes require the sets V_i to have different sizes, but do not necessarily need them to cover all but εn vertices in G . As this is a minor point, we do not introduce more notation and instead use the standard term ε -regular partition under the following more relaxed definition.

Definition 2.23. Let $1/n \ll \varepsilon \ll d \leq 1$, and let G be a red/blue coloured graph on n vertices. An ε -regular partition in G is a collection of disjoint subsets $V_1, \dots, V_k \subset V(G)$, such that for all but at most εk^2 pairs of indices $1 \leq i < j \leq k$, both $G_{\text{red}}[V_i, V_j]$ and $G_{\text{blue}}[V_i, V_j]$ are ε -regular. Each set V_i is called a *cluster*.

Given an ε -regular partition $V_1 \cup \dots \cup V_k$ in G , its corresponding (ε, d) -reduced graph R is a red/blue coloured graph with vertex set $[k]$, such that for each $* \in \{\text{red}, \text{blue}\}$ and any distinct $i, j \in [k]$, there is an ij edge of colour $*$ in R if and only if $G_*[V_i, V_j]$ is (ε, d) -regular.

Note that if $V_1, \dots, V_k$ form an ε -regular partition in a red/blue coloured complete graph and $\varepsilon \ll d \leq 1/2$, then for all but at most εk^2 pairs of indices $1 \leq i < j \leq k$, there is either a red edge ij or a blue edge ij (or both) in the corresponding (ε, d) -reduced graph R .

Finally, we prove the following refinement result that will be used later.

Lemma 2.24. Let $1/k, 1/m \ll \varepsilon \ll \eta \ll \alpha \ll d \leq 1$. Suppose G is a graph containing disjoint subsets $V_1, \dots, V_k \subset V(G)$, each of size m . Let R be a graph on $[k]$ such that for every $ij \in E(R)$, $G[V_i, V_j]$ is (ε, d) -regular. Suppose there exists a partition $[k] = I_1 \cup I_2$, with $|I_1| = k_1, |I_2| = k_2$, and $k_1, k_2 \geq \alpha k$, such that $R[I_1, I_2]$ is η -almost complete. Then, there exist two collections of disjoint sets $\{U_i : i \in J\}$ and $\{W_i : i \in J\}$ such that

- $|U_i| = |U_j|$ and $|W_i| = |W_j|$ for any $i, j \in J$,
- $\sum_{i \in J} |U_i| \geq (1 - \alpha) \sum_{i \in I_1} |V_i|$, $\sum_{i \in J} |W_i| \geq (1 - \alpha) \sum_{i \in I_2} |V_i|$, and
- $G[U_i, W_i]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular for every $i \in J$.

Proof. Let $\eta \ll \gamma \ll \alpha$. For each $i \in I_1$, pick a largest collection of disjoint subsets of V_i of size $\gamma k_1 m / (k_1 + k_2)$. For each $i \in I_2$, pick a largest collection of disjoint subsets of V_i of size $\gamma k_2 m / (k_1 + k_2)$. Let $\{U_i : i \in J_1\}$ and $\{W_i : i \in J_2\}$ be the collections of refined subsets coming from $\{V_i : i \in I_1\}$ and $\{V_i : i \in I_2\}$, respectively. Note that at most a γ -proportion of vertices are lost from each V_i in this refinement process. Let R' be a graph with vertex set $J_1 \cup J_2$, such that $ij \in E(R')$ if $G[U_i, W_j]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular. Then, as $R[I_1, I_2]$ is η -almost complete, $R'[J_1, J_2]$ is η -almost complete as well by Lemma 2.19. Therefore, we can greedily find a matching M of size $(1 - \eta) \min\{|J_1|, |J_2|\}$ in $R'[J_1, J_2]$. To finish, observe that $\sum_{i \in J_1 \cap V(M)} |U_i| \geq (1 - \eta)(1 - \gamma) \sum_{i \in I_1} |V_i| \geq (1 - \alpha) \sum_{i \in I_1} |V_i|$, and similarly $\sum_{i \in J_2 \cap V(M)} |W_i| \geq (1 - \alpha) \sum_{i \in I_2} |V_i|$. $\square$

3 Outline of the proof of Theorem 2.2: Stability

Simplifications for the discussion. The main technical tool for the stability part of the proof of Theorem 1.1, i.e. the proof of Theorem 2.2, is Szemerédi's regularity lemma. The following outline of the proof of Theorem 2.2 assumes a working knowledge of the regularity lemma and simple embeddings using it, and further suppresses two technical details that we will explain momentarily. Readers less familiar with regularity techniques may find it useful to start with Section 2.6, and readers finding this outline too scant in detail may find it valuable instead as a blueprint when reading the formal proofs in Section 4 and Section 5.

There are two main technicalities that we will suppress in the following outline. Most notably, due to the imbalance in the sizes t_1 and t_2 of the bipartition classes of the tree T , we will sometimes work with regularity partitions whose clusters have different sizes. In several cases, clusters will have two different sizes that are in the ratio $t_1 : t_2$. Moreover, in many of our more intricate arguments, the same cluster may change its role throughout the proof, having vertices from either the larger or the smaller side of the bipartition embedded into it. To facilitate this, we sometimes need to refine the regularity partition that

we started with, partitioning all clusters into smaller clusters of suitable sizes, before pairing them up again to form regular pairs with the right size ratio (see e.g. Lemma 2.24). In this outline, we will skim over this aspect of the proof. That is, we will only work with a fixed regularity partition here and not worry about the technicalities regarding cluster sizes and refinements. The crucial point we focus on in the outline here is the number of vertices in the original graph that are covered by a certain set of clusters, where for any set $I \subset [k]$, we say that I *covers* the vertices $\cup_{i \in I} V_i$.

The other technicality is one common to many uses of the regularity lemma: we will need to have many constants of decreasing sizes in some hierarchy. To avoid this burden here, we will informally use m^+ or m^- to denote a number equal to $m + \alpha n$ or $m - \alpha n$, respectively, for some small and suitable constant $\alpha > 0$. In particular, 0^+ will represent αn for some $\alpha > 0$. The constants α involved in different instances of these notations are all different and will be chosen later carefully in the formal proofs. To give a rough idea of the relation of parameters, we can expect these α to satisfy $\varepsilon \ll \alpha \ll 1$, where ε is the regularity parameter.

Set-up in the proof of Theorem 2.2. In Theorem 2.2, we have an n -vertex tree T satisfying $\Delta(T) \leq cn$ that has bipartition classes U_1 and U_2 with sizes t_1 and t_2 respectively, where $t_1 \geq t_2$. We also have a red/blue coloured complete graph G with $\max\{t_1 + 2t_2, 2t_1\} - 1$ vertices, and wish to either find a monochromatic copy of T in G or show that the colouring of G is close to one of the two extremal constructions, where this proximity is controlled with the parameter μ . As mentioned in Section 2.1, by adding leaves to the t_2 -side of the tree if necessary, we may assume that $t_1 \leq 2t_2 + 1$. In fact, as will be justified in Section 5.7, we can even assume that $t_1 \leq 2t_2$ in the proof of Theorem 2.2, which we will do from now on. In particular then, the graph G has $t_1 + 2t_2 - 1$ vertices.

Stages, situations and embedding methods. Let ε be a suitable regularity parameter satisfying $1/n \ll c \ll \varepsilon \ll \mu \ll 1$. We begin by applying a result (Theorem 5.1) of Haxell, Łuczak, and Tingley [21] to find an ε -regular partition in $V(G)$ that contains a certain monochromatic structure in the reduced graph (see the top left of Figure 2). Having found this, we say we have an **A-situation**. We then work through a sequence of 4 stages. At each stage, we either find a monochromatic copy of the tree, or deduce that G must be close to an extremal construction, or find more useful structure in the reduced graph. If the last of these is true at the end of a stage, we reach another named situation (see Figure 2), and if we reach the end of these 4 stages, then we have an **E-situation** (see the right of Figure 2), which will imply that G is close to an extremal construction.

At each stage, our deductions will often say that if there is a certain structure in the reduced graph R , then we can find a monochromatic copy of the tree T in G using a named embedding method. These embedding methods include one due to Haxell, Łuczak, and Tingley [21] that we will refer to as **HLT**, as well as a series of new ones denoted by **EM1a-c** and **EM2a-d** that we will prove in Section 4. The structure required for each of these methods is depicted in either Figure 3 or Figure 4, marked with relevant references to the corresponding lemmas and the sections they are proved in. Formal definitions of the required structures can be found in the relevant sections, though informal descriptions of these structures are provided in the captions. Which embedding methods are used for which stages are noted in Figure 2. To give a rough idea of how these embedding lemmas are proved, we will discuss and sketch a proof of simplest embedding method **HLT** below, and briefly relate it to the other methods. All of these proofs use the same general framework relying on a technical embedding lemma using regularity proved in Section 4.2.

Embedding method HLT. As shown by Haxell, Łuczak, and Tingley [21], if the left-most structure in Figure 3 can be found in the reduced graph R in red, say, then we can find a red copy of T in G . More precisely, the structure is defined as follows. There is an ε -regular partition $V_1 \cup \dots \cup V_{2k+1}$ in G with a corresponding (ε, d) -reduced graph R , an index $i \in [2k+1]$ and a partition $[2k+1] \setminus \{i\} = I_A \cup I_B$, such that ia is an edge in R_{red} for each $a \in I_A$, and there is a perfect matching M in R_{red} between I_A and I_B . Furthermore, I_A covers t_2^+ vertices of G while I_B covers t_1^+ vertices of G , and the ratio between the sizes of the clusters indexed by I_B and by I_A is around $t_1 : t_2$.

Observe that the structure described here is a bipartite subgraph of the reduced graph, so if we are to embed the tree T into the ε -regular pairs corresponding to edges in this structure, it is necessary that I_A covers t_2^+ vertices of G so that there is enough room for vertices in U_2 to be embedded among them, and similarly that I_B covers t_1^+ vertices of G for the embedding of vertices in U_1 . To prove their approximate

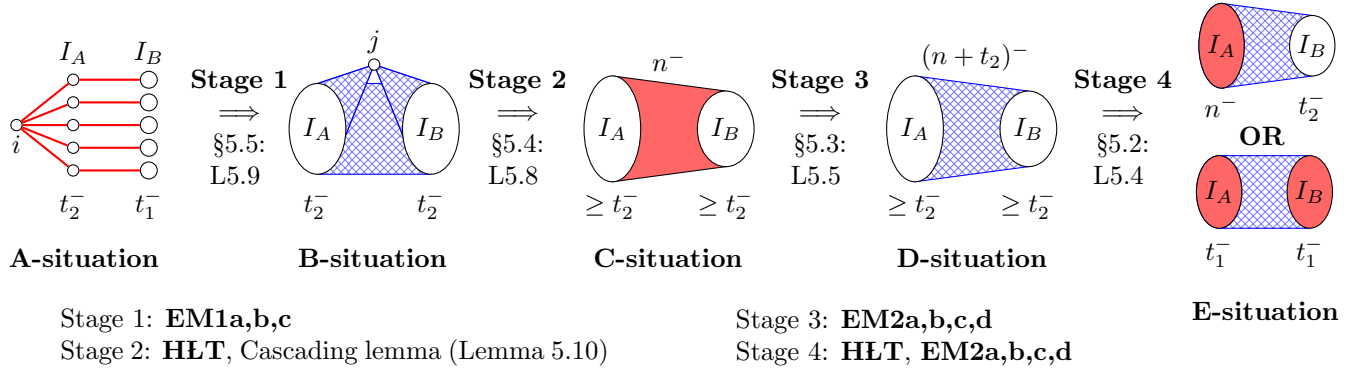


Figure 2: The different situations we find in the reduced graph, and the stages we use to work through them, along with the sections they are carried out in and the corresponding lemmas. Which embedding methods are used at which stages is recorded underneath. The numbers n^- and $(n + t_2)^-$ on top refer to the total number of vertices in G covered by $I_A \cup I_B$, while the numbers beneath refer to the number of vertices of G covered by I_A or I_B as appropriate. The shaded areas indicate that the corresponding edges in the reduced graph are mostly that colour.

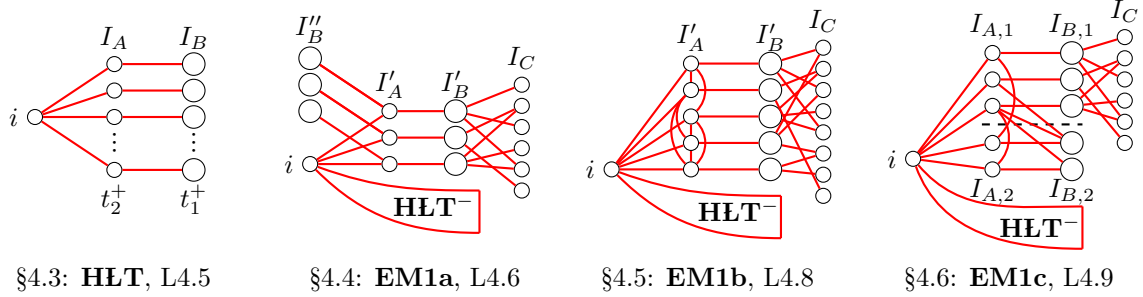


Figure 3: The structure in the reduced graph required for embedding methods **HLT** and **EM1a-c**, along with their corresponding sections, names, and lemmas. On the left, t_2^+ and t_1^+ refer to the number of vertices in G covered by I_A and I_B . In each other structure, **HLT**⁻ refers to the same structure as **HLT** attached to i but covering t_2^- and t_1^- vertices, and comprises the majority of the required structure, while the remaining structure pictured covers 0^+ vertices in G . In **EM1a-c**, each vertex in I_B' or $I_{B,1}$ has some red neighbours in I_C . In **EM1b** there are some red edges within I_A' , while in **EM1c** there are sets $I_{A,2}$ and $I_{B,2}$ matched together in red and every vertex in these sets has some red neighbours in $I_{A,1}$.

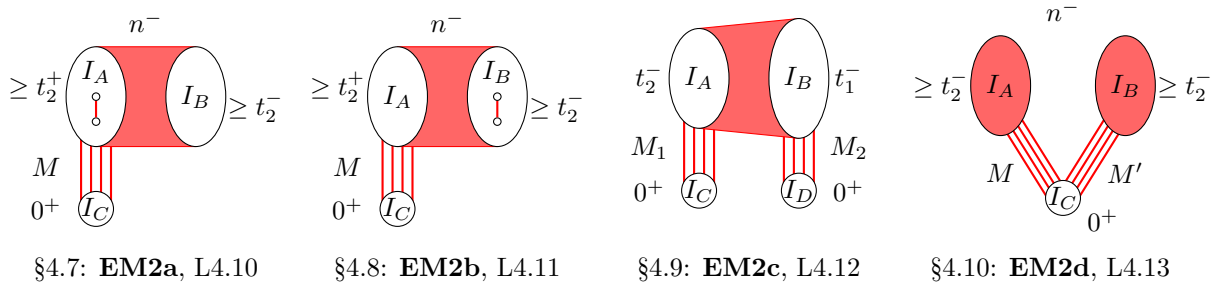


Figure 4: The structure in the reduced graph required for **EM2a-d**, along the corresponding sections, names, and lemmas. In all cases I_A and I_B together cover n^- vertices, and lower bounds or sizes for the number of vertices covered by I_A and by I_B are given in each case. In **EM2a-c**, almost all of the edges in R between I_A and I_B are red. In **EM2a** and **EM2b**, there is a red edge in $R[I_A]$ and $R[I_B]$, respectively. In **EM2d** almost all edges in R within I_A and within I_B are red.

version of Theorem 1.1 in [21], Haxell, Łuczak, and Tingley started with a red/blue coloured complete graph on $(t_1 + 2t_2)^+$ vertices, and showed that its reduced graph will always contain, in red or blue, the structure required to apply **HLT**. However, as our graph G has only $t_1 + 2t_2 - 1$ vertices, we cannot find this structure in full. Instead, we use their result to show that we can find a slightly scaled down version of **HLT** with the corresponding sets I_A and I_B covering t_2^- and t_1^- vertices respectively (see Theorem 5.1). This is the **A-situation** depicted in Figure 2, and we denote this structure by **HLT**⁻.

Roughly speaking, in each of our other embedding methods we will be able to embed most of the tree relatively easily, but need to somehow make up for a lack of vertices in both I_A and I_B in the **HLT**⁻ structure. This can be seen in the structures required for **EM1a-c** depicted in Figure 3, where part of the structure in each case is **HLT**⁻, and we need to find some additional structure to complete the embedding. A major driver of the complexity in our embedding is that the tree T could have a linear maximum degree, so the diameter of T could be as small as 5. This means that structures in the reduced graph with larger diameters are often not very useful for us.

As an example, we now give a sketch of how to embed T into the structure **HLT** described above. We start by finding a constant-sized set of vertices $X \subset V(T)$ such that $T - X$ has only small components (see Lemma 2.16 and Lemma 4.1). These components could be linear-sized, but must be much smaller compared to the regularity clusters. Let $i_1 = i$. Remove an arbitrary edge from M , let i_2 be the endpoint of this edge in I_A , and note that $i_1 i_2$ is an edge in R_{red} . This small modification is depicted in Figure 9.

Our aim is to embed T into G_{red} so that vertices in $X \cap U_2$ are embedded into V_{i_2} , vertices in $(X \cap U_1) \cup N_T(X \cap U_2)$ are embedded into V_{i_1} , and for each component K of $T - X$, we can assign to it some edge ab in M with $a \in I_A$ and $b \in I_B$, such that all vertices in K not mentioned so far are embedded into either V_a or V_b depending on if they are in U_2 or U_1 , respectively. As X is constant-sized, by choosing the maximum degree parameter c to be small enough, $N_T(X \cap U_2)$ will be a linear-sized set small enough to be embedded along with $X \cap U_1$ into the regularity cluster V_{i_1} . We are only embedding $X \cap U_2$ into the cluster V_{i_2} , so there is plenty of room there. To make sure that we have enough room for the rest of the embedding, for each component of $T - X$ we will decide which clusters to embed it into by picking an edge ab in M independently and uniformly at random. As each component of $T - X$ is small, with high probability this will distribute the components of $T - X$ across the edges of M without too many vertices assigned to any one edge. This ensures we have enough room in each cluster, so standard regularity techniques now apply to find a red copy of T .

Overview of the 4 stages in the proof of Theorem 2.2. We start the 4-stage proof of Theorem 2.2 by applying the aforementioned Haxell, Łuczak, and Tingley [21] result (see Theorem 5.1) to our red/blue coloured complete graph G to find a monochromatic **HLT**⁻ structure in the reduced graph R . This is the **A-situation** depicted in the left of Figure 2, and we can now proceed to **Stage 1**.

Stage 1

Given an **A-situation** in R , say in red, I_A and I_B each does not cover enough vertices to let us use the embedding method **HLT**. Let $I_C = V(R) \setminus (I_A \cup I_B \cup \{i\})$, so that I_C covers $(n + t_2)^- - t_2^- - t_1^- = t_2^-$ vertices. If almost all of the edges in R between I_B and I_C are blue, then we would have a **C-situation** in blue with I_C in place of I_A , and can skip ahead to **Stage 3**. Suppose, then, that there are at least some red edges between I_B and I_C . In particular, if $I_{B,1}$ is the set of vertices in I_B with at least some red neighbours in I_C , then $I_{B,1}$ is non-empty. Let $I_{A,1}$ be the set of vertices matched with $I_{B,1}$ by the matching M . Then, take $I_{A,3}$ to be the set of vertices in $I_A \setminus I_{A,1}$ with at least some red neighbours in I_C , and let $I_{B,3}$ be the set of vertices matched with $I_{A,3}$ by M . Finally, let $I_{A,2} = I_A \setminus (I_{A,1} \cup I_{A,3})$ and $I_{B,2} = I_B \setminus (I_{B,1} \cup I_{B,3})$. See the left of Figure 5 for a depiction of these vertex sets.

Suppose there is no copy of T in red, and thus the structure required to use any of **EM1a-c** does not exist in red. We will be able to show, then, that **i)** the edges between $I_{A,1}$ and I_C are almost all blue, **ii)** the edges between $I_{A,1}$ and $I_{B,3}$ are almost all blue, **iii)** most of the edges in $I_{A,1}$ are blue, and **iv)** for most of the edges $i_A i_B \in M[I_{A,2}, I_{B,2}]$, one of i_A or i_B will have mostly blue neighbours in $I_{A,1}$. These deductions are commented on below, but assuming **i)–iv)** hold, we find a **B-situation** as follows.

From **iv)**, we can find a subset $I_{AB,2} \subset I_{A,2} \cup I_{B,2}$ containing a vertex in almost every edge in $M[I_{A,2}, I_{B,2}]$, so that the edges between $I_{AB,2}$ and $I_{A,1}$ are mostly blue. Combined with **ii)** and **iii)**, the edges between $I_D := I_{AB,2} \cup I_{B,3} \cup I_{A,1}$ and $I_{A,1}$ are mostly blue. Furthermore, since I_D contains a vertex in almost every edge in M , and every cluster indexed by I_B contains more vertices than one indexed by I_A , we see that I_D covers at least close to the same number of vertices as I_A does, and thus I_D covers

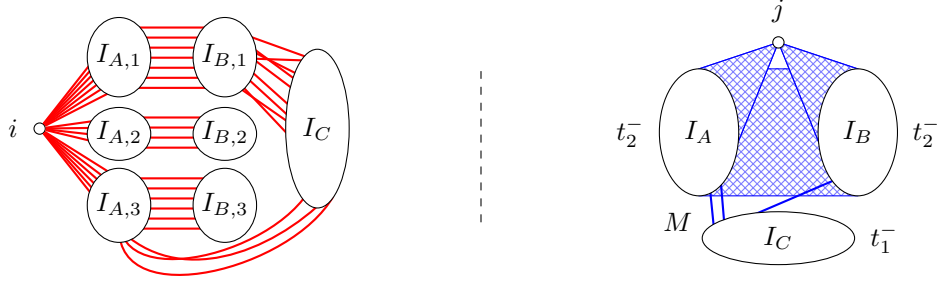


Figure 5: On the left, the main structure in Stage 1. On the right, the main structure in Stage 2.

at least t_2^- vertices. Moreover, as $I_{A,1}$ is non-empty and **i)** holds, the edges from I_C to I_D are almost all blue, and we can select some $j \in I_{A,1}$ with almost all blue edges to $I_C \cup I_D$. Thus, j , I_C and $I_D \setminus \{j\}$ give the structure required for a **B-situation** in blue.

We finish this discussion of **Stage 1** by commenting briefly on the four deductions mentioned above and the embedding methods required for them, using the same labels as in Figure 3 where possible.

i) $R[I_{A,1}, I_C]$ is almost all blue. If this does not hold, then let $I'_{A,1}$ be a small set of vertices in $I_{A,1}$ with some red neighbours in I_C , and let $I'_{B,1}$ be the vertices matched with $I'_{A,1}$ by M . We can then find a perfect red matching between $I'_{A,1}$ and some $I''_{B,1} \subset I_C$, and set $I'_C = I_C \setminus I''_{B,1}$ to obtain the structure required for **EM1a**.

ii) $R[I_{A,1}, I_{B,3}]$ is almost all blue. If this does not hold, then, similar to **i)**, let $I'_{A,1}$ be a small set of vertices in $I_{A,1}$ with some red neighbours in $I_{B,3}$, let $I'_{B,1}$ be the vertices matched with $I'_{A,1}$ by M , and find a perfect red matching between $I'_{A,1}$ and some $I''_{B,1} \subset I_{B,3}$. Unlike **i)**, using $I''_{B,1}$ for the structure in **EM1a** will ‘orphan’ the vertices in $I_{A,3}$ matched to $I''_{B,1}$ by M , say those in $I'_{A,3}$. However, from the definition of $I_{A,3}$, we can find a perfect red matching between $I'_{A,3}$ and some $I'_{B,3} \subset I_C$, which can be used to replace $I''_{B,1}$ to complete the structure required for **EM1a**.

iii) $R[I_{A,1}]$ is almost all blue. If this does not hold, then we have the structure required for **EM1b** – some red edges within a set $I_{A,1}$ whose neighbours under M (i.e., the vertices in $I_{B,1}$) all have some red neighbours in I_C .

iv) For most edges in $M[I_{A,2}, I_{B,2}]$, one endpoint has mostly blue edges to $I_{A,1}$. If this does not hold, then we have the structure required for **EM1c** – a small red submatching of $M[I_{A,2}, I_{B,2}]$ whose vertices all have some neighbours within $I_{A,1}$.

Stage 2

Suppose, now, we have a blue **B-situation** in R : a vertex j with blue edges to almost every vertex in two disjoint sets I_A and I_B , with both I_A and I_B covering t_2^- vertices of G , and almost every edge between them being blue. Let M be a maximum blue matching between $I_A \cup I_B$ and I_C (see the right of Figure 5). If $I_C \cap V(M)$ covers more than $(t_1 - t_2)^+$ vertices in G , then using this matching and the **B-situation** structure, we can find the structure required to embed T in blue using **HLT**, where we use j as the vertex i in the **HLT** structure and use that $I_A \cup I_B \cup (I_C \cap V(M))$ covers $2t_2^- + (t_1 - t_2)^+ = n^+$ vertices in G .

Therefore, we can assume that $I_C \cap V(M)$ covers at most $(t_1 - t_2)^+$ vertices, so $I_C \setminus V(M)$ covers at least $(n + t_2)^- - 2t_2^- - (t_1 - t_2)^+ = t_2^-$ vertices. If almost all edges between a subset $I \subset I_A \cup I_B$ and $I_C \setminus V(M)$ are red, with $I \cup (I_C \setminus V(M))$ covering n^- vertices in G , then we have a red **C-situation**. Unfortunately, the maximality of M only immediately gives that almost all edges between $(I_A \cup I_B) \setminus V(M)$ and $I_C \setminus V(M)$ are red, and they might cover only $(n + t_2)^- - 2(t_1 - t_2)^+ = (4t_2 - t_1)^-$ vertices together, which could be as small as $2t_2^-$ if $t_1 \approx 2t_2$.

To combat this, we exploit the maximality of the matching M in a more sophisticated way using what we call a ‘cascading argument’ (see Lemma 5.10). Note that any blue edge between some $c \in I_C \setminus V(M)$ and $(I_A \cup I_B) \cap V(M)$ would allow us to exchange that edge into the matching M to create a matching M' that has the same intersection with $I_A \cup I_B$ as M , but whose intersection with I_C includes c and omits some vertex $c' \in I_C \cap V(M)$. The maximality of M then implies the edges between c' and $(I_A \cup I_B) \setminus V(M)$

are mostly red. Iterating such an argument will eventually allow us to find two large subsets X and Y with $(I_A \cup I_B) \setminus V(M) \subset X \subset I_A \cup I_B$ and $I_C \setminus V(M) \subset Y \subset I_C$, such that the edges between X and Y are mostly red, X and Y both cover at least t_2^- vertices in G , and $X \cup Y$ cover at least n^- vertices in G . This gives a red **C-situation**.

Stage 3

Suppose then we have a red **C-situation** in R , which consists of two disjoint vertex sets I_A and I_B , each covering at least t_2^- vertices and together covering n^- vertices, such that $R[I_A, I_B]$ is mostly red. Let $I_C = V(R) \setminus (I_A \cup I_B)$, which covers $(n + t_2)^- - n^- = t_2^-$ vertices. In the rest of **Stage 3** we will use two different sequences of deductions (**Claim A** and **Claim B**) several times. Before continuing then, we state roughly what they are and summarise the arguments for them.

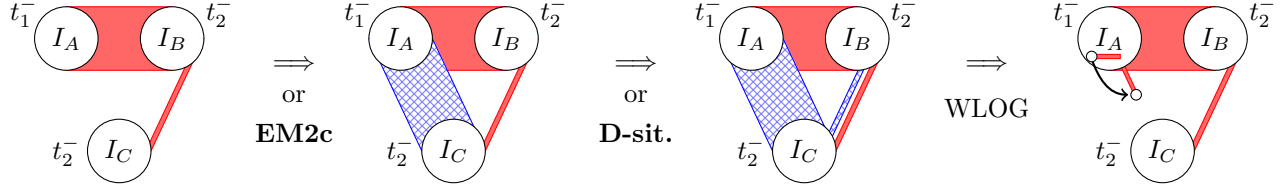


Figure 6: The deductions for **Claim A**, before finally **EM2c** is applied to get a red copy of T .

Claim A: If I_A and I_B cover t_1^- and t_2^- vertices respectively, and there are some red edges between I_B and I_C , then we can either find a monochromatic copy of T or reach a **D-situation**.

Argument for Claim A: If there are some red edges between I_A and I_C then **EM2c** applies. Thus, we can assume $R[I_A, I_C]$ is almost all blue. If $R[I_B, I_C]$ is mostly red, then we have a **D-situation** in red using $I_A \cup I_C$ and I_B , so there must be some blue edges between I_B and I_C , as well as some red edges as part of the assumption. If there are some red edges in I_A , then we can find a small red matching in I_A and move one side of this matching out of I_A to get the structure required for **EM2c** in red. Finally, if there are some blue edges in I_A then we can similarly apply **EM2c** in blue.

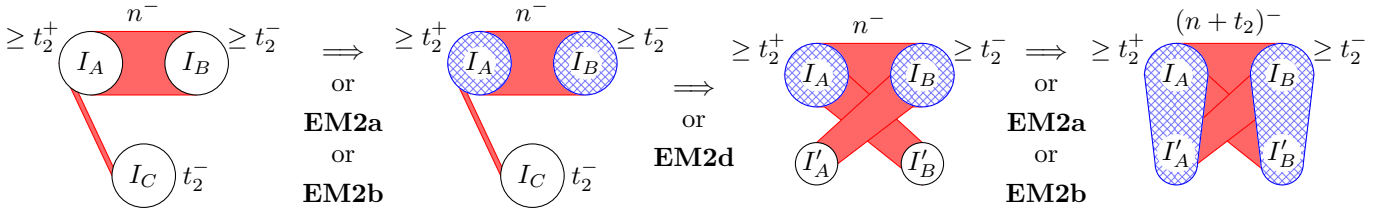


Figure 7: The deductions for **Claim B**, before finally either **EM2d** applies or we have a **D-situation**.

Claim B: Suppose I_A and I_B cover at least t_2^+ and t_2^- vertices respectively, and n^- vertices in total. If there are some red edges between I_A and I_C , then we can either find a monochromatic copy of T or reach a **D-situation**.

Argument for Claim B: If there are some red edges in $R[I_A]$ or $R[I_B]$, then **EM2a** or **EM2b** applies respectively, so assume that $R[I_A]$ and $R[I_B]$ are both mostly blue. If some vertices in I_C have some blue neighbours in both I_A and I_B , then we can use **EM2d**. Thus, we can partition most of I_C into $I'_A \cup I'_B$, such that $R[I_A, I'_B]$ and $R[I'_A, I_B]$ are both mostly red.

Like above, if there are some red edges in $R[I_A \cup I'_A]$ or $R[I_B \cup I'_B]$, then **EM2a** or **EM2b** applies respectively, so we can assume that both $R[I_A \cup I'_A]$ and $R[I_B \cup I'_B]$ are mostly blue. If there is some vertex in I'_A with some blue neighbours in $I_B \cup I'_B$, or if there is some vertex in I'_B with some blue neighbours in $I_A \cup I'_A$, then we can apply **EM2d**. Thus, we can assume that $R[I_A \cup I'_A, I_B \cup I'_B]$ is mostly red, which gives a **D-situation**.

Stage 3 using Claim A and B. Using these two claims, we can now carry out **Stage 3**. Note first that if $R[I_A \cup I_B, I_C]$ is mostly blue, then they form a **D-situation**, so assume that there are some red edges either between I_A and I_C or between I_B and I_C . If both I_A and I_B cover at least t_2^+ vertices, then we can use **Claim B** to find a monochromatic copy of T or reach a **D-situation**.

Thus, we can assume without loss of generality that I_B covers at most t_2^+ vertices, so I_A covers at least t_1^- vertices. If there are some red edges between I_B and I_C , then we can use **Claim A** to find a monochromatic copy of T or reach a **D-situation**. Otherwise, there must be some red edges between I_A and I_C . If $t_1 \leq t_2^+$, then we can apply **Claim A** with I_A and I_B swapped as $t_1 \approx t_2$, while if $t_1 \geq t_2^+$, then we can apply **Claim B**.

Stage 4

Suppose finally that we have a blue **D-situation** in R , which consists of two disjoint vertex sets I_A and I_B , each covering at least t_2^- vertices and together covering $(n + t_2)^-$ vertices, such that $R[I_A, I_B]$ is mostly blue. First, suppose in addition that both I_A and I_B cover at least t_2^+ vertices. Then, if there is a blue edge in either $R[I_A]$ or $R[I_B]$, we can take some vertices out of I_B to form I_C , which allows us to apply **EM2a** or **EM2b**, respectively, to find a blue copy of T . Thus, we can assume that $R[I_A]$ and $R[I_B]$ are both mostly red. If the larger of I_A and I_B , which we can assume is I_A , covers at least t_1^+ vertices, then we can take some vertices out of I_A to form I_D and some vertices out of I_B to form I_C , so that we can apply **EM2c** to get a blue copy of T . If I_A covers at most t_1^+ vertices, then we have $(n + t_2)^- \leq 2t_1^+$, and so $t_1 \approx 2t_2$ and both I_A and I_B must cover at least t_1^- vertices in G . This gives an **E-situation**, and will imply that G is close to a Type II extremal construction.

Now suppose that the smaller of I_A and I_B , which we can assume is I_B , covers at most t_2^+ vertices. Then, I_A covers $(n + t_2)^- - t_2^+ = n^-$ vertices in G . If there are some blue edges in $R[I_A]$, then we can use them to take some vertices out of I_A to form I_C , and then apply **EM2a**. Thus, we can assume that $R[I_A]$ is mostly red. If I_A covers at least n^+ vertices, then we can easily find the structure required to apply **EM2c** in $R[I_A]$. Therefore, we can assume that I_A covers at most n^+ vertices, and so we have an **E-situation** that will imply that G is close to a Type I extremal construction.

4 Embedding methods for the proof of Theorem 2.2: Stability

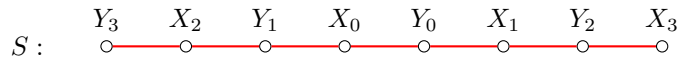
In this section, we prove a series of embedding lemmas using regularity, each of which says that if a certain structure exists in the reduced graph R , then we can embed T into G . In Section 4.1, we prove a tree decomposition lemma phrased in terms of graph homomorphisms, which cuts the tree T into small pieces by removing very few vertices. Then, in Section 4.2, we prove our main technical lemma, Lemma 4.2. Roughly speaking, it says that given a suitable structure in the reduced graph and an appropriate assignment of each small piece of the tree T to a part of the structure, we can find a copy of T by embedding each piece between a randomly chosen regular pair within the part it is assigned to. This is then applied to prove embedding methods **HLT** in Section 4.3, **EM1a-c** in Sections 4.4–4.6, and **EM2a-d** in Sections 4.7–4.10. In each application, the structure in the reduced graph provided by the assumption is transformed into a substructure of the one used in Lemma 4.2, then we find a proper assignment of each piece of the tree T to a part of this structure, so that on average no cluster has too many vertices assigned to it.

As mentioned at the end of Section 2.1, by adding leaves to the t_2 -side of the tree if necessary, we can assume that $t_1 \leq 2t_2 + 1$. In fact, as we will show later in Section 5 when we prove Theorem 2.2, it can even be assumed that $t_1 \leq 2t_2$. As such, all of the embedding methods we prove below in this section will have the assumption that $t_2 \leq t_1 \leq 2t_2$.

4.1 Tree decomposition

In this subsection, we prove the following lemma phrased in terms of graph homomorphisms that cuts the tree T into small pieces by removing very few vertices. Recall that for graphs H_1 and H_2 , a function $\phi : H_1 \rightarrow H_2$ is a *graph homomorphism* if for any edge uv in H_1 , $\phi(u)\phi(v)$ is also an edge in H_2 .

Lemma 4.1. *Let $1/n \ll c \ll \xi$. Let S be the following graph:*



Let T be an n -vertex tree. Then, there is a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$.

Proof. Let the bipartition classes of T be V_1 and V_2 . By Lemma 2.16, there exists $Z \subset V(T)$ such that each component of $T - Z$ has size at most ξn and $|Z| \leq 2\xi^{-1}$. It follows that $T - Z$ contains at most $|Z| \cdot \Delta(T) \leq 2cn\xi^{-1} \leq \xi n/10$ components. Arbitrarily pick $t_1 \in Z$, view T as being rooted at t_1 , and extend t_1 to an ordering $t_1, t_2, \dots, t_n$ of the vertices of T , such that for each $2 \leq i \leq n$, t_i has a unique neighbour, namely its parent in T , to its left in this ordering, and vertices in the same component of $T - Z$ appear consecutively.

Let $A \subset V(T - Z)$ be the set consisting of each vertex in a component of $T - Z$ that appears first in the ordering. Let B be the set of parents of vertices in Z , and let C be the set of parents and children of vertices in B . Note that the sets A, B, C could overlap. For each $* \in \{A, B, C, Z\}$ and $j \in [2]$, let $*_j = * \cap V_j$.

We now define a homomorphism $\phi : T \rightarrow S$, so that the following conditions are maintained.

A1 $\phi(t) = X_0$ for each $t \in Z_1$ and $\phi(t) = Y_0$ for each $t \in Z_2$.

A2 $\phi(t) \in \{X_0, X_1\}$ for each $t \in B_1$ and $\phi(t) \in \{Y_0, Y_1\}$ for each $t \in B_2$.

A3 $\phi^{-1}(\{X_0, X_1, Y_0, Y_1\}) \subset A \cup B \cup C \cup Z$.

To initialise, let $\phi(t_1) = X_0$ if $t_1 \in Z_1$ and $\phi(t_1) = Y_0$ if $t_1 \in Z_2$. Suppose we have just finished defining ϕ for a vertex in Z or a component in $T - Z$.

Suppose first that the next vertex in the ordering is a vertex $t_i \in Z$. If $t_i \in Z_1$, then we can define $\phi(t_i) = X_0$ as the image of its parent in T under ϕ is adjacent to X_0 in S by **A2**. Similarly, if $t_i \in Z_2$, then we can safely define $\phi(t_i) = Y_0$.

Now suppose the next vertex in the ordering is a vertex $t_i \in A$ within a component K of $T - Z$. Assume that $t_i \in A_1$, then ϕ sends its parent to Y_0 by **A1**, so we can define $\phi(t_i) = X_1$. Note that **A2** is also maintained if t_i happens to be in B_1 . Define ϕ on the remaining vertices $t \in K$ as follows.

$t \in$	A_1	B_1	B_2	C_1	$V_1 \setminus (A_1 \cup B_1 \cup C_1)$	$V_2 \setminus B_2$
$\phi(t) =$	X_1	X_1	Y_0	X_1	X_3	Y_2

One can check that this defines a valid homomorphism using the definitions of A, B, C . For example, if $t \in B_2 \cap K$, then $\phi(t)$ is defined to be Y_0 . The parent and children of t are in C_1 from definition, and are sent by ϕ to X_1 , which is valid. Moreover, both **A2** and **A3** are maintained. The case when $t_i \in A_2$ is symmetric so is omitted.

Therefore, we can define a homomorphism $\phi : T \rightarrow S$ satisfying **A1–A3**. Every component of $T - \phi^{-1}(\{X_0, Y_0\})$ has size at most ξn as it is contained in a component of $T - Z$. Finally, by **A3**, $|\phi^{-1}(\{X_0, X_1, Y_0, Y_1\})| \leq |A| + |B| + |C| + |Z| \leq \xi n/10 + 2\xi^{-1} + 2\xi^{-1} \cdot \Delta(T) + 2\xi^{-1} \leq \xi n$, as required. $\square$

4.2 Main technical embedding lemma

In this subsection, we prove our main technical embedding lemma. Roughly speaking, it says that given a reduced graph structure R and a tree T cut into many small pieces, if each piece can be assigned appropriately into a part of R (represented below as a homomorphism $\varphi : T \rightarrow R'$), so that on average no cluster has too many vertices assigned to it (see **B2**), then we can find a copy of T in G .

Lemma 4.2. *Let $0 < 1/n \ll c \ll \xi \ll 1/k \ll \varepsilon \ll \alpha \ll d \leq 1$. Let $i_1, i_2, i_3 \in [k]$ be distinct and let $I_{0,1}, I_{0,2}, I_{0,3}, I_{1,1}, I_{1,2}, I_{1,3}, I_{2,1}, I_{2,2}, I_{2,3}, I_{3,1}, I_{3,2}, I_{3,3}$ partition $[k] \setminus \{i_1, i_2, i_3\}$ such that $|I_{0,1}| = |I_{0,2}| = |I_{0,3}|$, $|I_{1,1}| = |I_{1,2}| = |I_{1,3}|$, $|I_{2,1}| = |I_{2,2}| = |I_{2,3}|$, and $|I_{3,1}| = |I_{3,2}| = |I_{3,3}|$.*

As depicted in Figure 8, let R be a graph with vertex set $[k]$ and edge set consisting of the edges in

$$\{i_1 i_2, i_2 i_3\} \cup \{i_1 i : i \in I_{0,1}\} \cup \{i_a i : a \in [3], i \in I_{a,1}\},$$

along with a perfect matching between each of the eight pairs of vertex sets $(I_{a,1}, I_{a,2})$ and $(I_{a,2}, I_{a,3})$ for every $a \in [3]_0$. Let R' be the graph with vertex set $\{i_1, i_2, i_3\} \cup \{I_{a,b} : a \in [3]_0, b \in [3]\}$ and edge set

$$\{i_1 i_2, i_2 i_3, i_1 I_{0,1}, i_1 I_{1,1}, i_2 I_{2,1}, i_3 I_{3,1}\} \cup \{I_{a,1} I_{a,2} : a \in [3]_0\} \cup \{I_{a,2} I_{a,3} : a \in [3]_0\}.$$

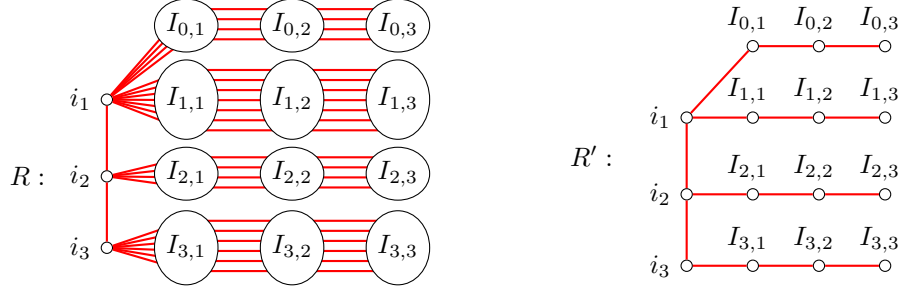


Figure 8: Auxiliary graphs R and R' used in the statement of Lemma 4.2

Let G be a graph on at most $2n$ vertices with a vertex partition $V_1 \cup V_2 \cup \dots \cup V_k$, such that for each $ij \in E(R)$, $G[V_i, V_j]$ is (ε, d) -regular, and for each $I \in V(R') \setminus \{i_1, i_2, i_3\}$, the sets V_i with $i \in I$ all have the same size. Assume also that $n/10k \leq |V_i| \leq \varepsilon n$ for each $i \in \{i_1, i_2, i_3\}$.

Let T be an n -vertex tree with $\Delta(T) \leq cn$. Suppose $\varphi : T \rightarrow R'$ is a homomorphism such that the following hold.

- B1** $\varphi^{-1}(\{i_1, i_2, i_3\}) \neq \emptyset$, and every component of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$ has size at most ξn .
- B2** For each $I \in V(R') \setminus \{i_1, i_2, i_3\}$ with $\varphi^{-1}(I) \neq \emptyset$, $|\varphi^{-1}(I)| + \alpha n \leq \sum_{i \in I} |V_i|$.
- B3** For each $i \in \{i_1, i_2, i_3\}$, $|\varphi^{-1}(i)| \leq \xi n$.

Then, G contains a copy of T .

Proof. Since φ is a homomorphism, the image of every component of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$ under φ is entirely contained in $I_a := I_{a,1} \cup I_{a,2} \cup I_{a,3}$ for some $a \in [3]_0$. For each $a \in [3]_0$ then, let r_a be the number of components of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$ whose images are contained in $I_{a,1} \cup I_{a,2} \cup I_{a,3}$, and label these components as $T_{a,1}, \dots, T_{a,r_a}$. Let $x_1 \in \varphi^{-1}(\{i_1, i_2, i_3\})$, view T as being rooted at x_1 , and extend it to an ordering $x_1, \dots, x_n$ of $V(T)$ so that $T[x_1, \dots, x_j]$ is a tree for each $j \in [n]$, and vertices in the same component of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$ appear consecutively. Moreover, we can ensure that for every $x \in \varphi^{-1}(\{i_1, i_2, i_3\})$, the components of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$ that directly descends from x appear right after x in this ordering. For each $a \in [3]_0$ and $\ell \in [r_a]$, let $p_{a,\ell}$ be the smallest index such that $x_{p_{a,\ell}}$ is a vertex in $T_{a,\ell}$. By relabelling if necessary, assume that $p_{a,1} < p_{a,2} < \dots < p_{a,r_a}$ for each $a \in [3]_0$.

We now provide a random algorithm that, with positive probability, produces an assignment function $\sigma : V(T) \rightarrow V(R)$ consistent with φ that guides an embedding $\psi : T \rightarrow G$. To initialise, let $\sigma(x) = i$ for every $x \in \varphi^{-1}(i)$ and $i \in \{i_1, i_2, i_3\}$, and let ψ be the empty function. Now, for each $s \in [n]$ in turn, if $s = p_{a,\ell}$ for some $a \in [3]_0$ and $\ell \in [r_a]$, then we extend the definition of σ to include all vertices in $T_{a,\ell}$ in a random manner defined below, while we do nothing to σ otherwise. Then, if possible, we extend ψ by embedding x_s into G so that the following properties hold, otherwise we stop this process. For notational convenience, let $i_0 = i_1$, and let $k_a = |I_{a,1}|$ for every $a \in [3]_0$.

- C1** $\psi(x_j) \in V_{\sigma(x_j)}$ for each $j \in [s]$.
- C2** For every $j \in [s]$ and $j' > j$ satisfying $j' \notin \{p_{a,\ell} : a \in [3]_0, \ell \in [r_a]\}$ and $x_j x_{j'} \in E(T)$, we have $d_G(\psi(x_j), V_{\sigma(x_{j'})} \setminus \psi(\{x_1, \dots, x_{j-1}\})) \geq d|V_{\sigma(x_{j'})}|/4$ if $\sigma(x_{j'}) \in \{i_1, i_2, i_3\}$, and $d_G(\psi(x_j), V_{\sigma(x_{j'})} \setminus \psi(\{x_1, \dots, x_{j-1}\})) \geq d\alpha n/4k_a$ if $\sigma(x_{j'}) \in I_a$ for some $a \in [3]_0$.
- C3** For every $a \in [3]_0$ and $j \in [s]$, if $\varphi(x_j) = i_a$, then for all but at most $\alpha k_a/100$ values of $i \in I_{a,1}$, there exists a set $W_{i,j} \subset N(\psi(x_j), V_i \setminus \psi(\{x_1, \dots, x_{j-1}\}))$ with size $d\alpha n/8k_a$, such that if $j < j' \leq s$ and $x_{j'}$ is a vertex in a component that directly descends from x_j , then $\psi(x_{j'})$ avoids these sets $W_{i,j}$ unless $x_{j'} \in N_T(x_j)$.

First, we show how σ is randomly extended when $s = p_{a,\ell}$ for some $a \in [3]_0$ and $\ell \in [r_a]$. Note that $s > 1$ as $x_1 \in \varphi^{-1}(\{i_1, i_2, i_3\})$. Let $1 \leq s' < s$ be the unique index satisfying $x_{s'} x_s \in E(T)$, and observe

that as $T_{a,\ell}$ is a component of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$, we have $\varphi(x_s) = I_{a,1}$, and thus $\sigma(x_{s'}) = i_a$. By **C3**, there exists $I_{a,1,\ell} \subset I_{a,1}$ with $|I_{a,1,\ell}| \geq (1 - \alpha/100)k_a$, such that for every $i \in I_{a,1,\ell}$, there exists $W_{i,s'} \subset N(\psi(x_{s'}), V_i \setminus \psi(\{x_1, \dots, x_{s'-1}\}))$ with size $d\alpha n/8k_a$. Pick $i_{a,1,\ell} \in I_{a,1,\ell}$ uniformly at random. Let $i_{a,2,\ell} \in I_{a,2}$ and $i_{a,3,\ell} \in I_{a,3}$ be such that $i_{a,1,\ell}i_{a,2,\ell}, i_{a,2,\ell}i_{a,3,\ell} \in E(R)$. For each $b \in [3]$ and each $x \in V(T_{a,\ell}) \cap \varphi^{-1}(I_{a,b})$, set $\sigma(x) = i_{a,b,\ell}$. This extends the definition of σ to include vertices in $V(T_{a,b})$.

Now, for each $a \in [3]_0$, $b \in [3]$, $i \in I_{a,b}$ and $\ell \in [r_a]$, if $i_{a,b,\ell} = i$, then let

$$Z_{a,b,i,\ell} = |V(T_{a,\ell}) \cap \varphi^{-1}(I_{a,b})|,$$

otherwise, including if the process above stops early, let $Z_{a,b,i,\ell} = 0$.

Claim 4.3. *For each $a \in [3]_0$, $b \in [3]$, and $i \in I_{a,b}$, with probability at least $1 - 1/20k$, we have*

$$\sum_{\ell \in [r_a]} Z_{a,b,i,\ell} \leq \frac{|\varphi^{-1}(I_{a,b})|}{k_a} + \frac{\alpha n}{2k_a}.$$

Proof of Claim 4.3. Fix $a \in [3]_0$, $b \in [3]$, and $i \in I_{a,b}$. For each $\ell \in [r_a]$, we have

$$\mathbb{E}(Z_{a,b,i,\ell} \mid Z_{a,b,i,1}, \dots, Z_{a,b,i,\ell-1}) \leq \frac{|V(T_{a,\ell}) \cap \varphi^{-1}(I_{a,b})|}{|I_{a,1,\ell}|} \leq \frac{|V(T_{a,\ell}) \cap \varphi^{-1}(I_{a,b})|}{(1 - \alpha/100)k_a},$$

and

$$\frac{1}{(1 - \alpha/100)k_a} \sum_{\ell \in [r_a]} |V(T_{a,\ell}) \cap \varphi^{-1}(I_{a,b})| \leq \frac{|\varphi^{-1}(I_{a,b})|}{k_a} + \frac{\alpha n}{4k_a}.$$

Furthermore, as $|T_{a,\ell}| \leq \xi n$ for each $\ell \in [r_a]$, we have $Z_{a,b,i,\ell} \leq |T_{a,\ell}| \leq \xi n$. Thus, using $\sum_{\ell \in [r_a]} |T_{a,\ell}| \leq n$, we have

$$\sum_{\ell \in [r_a]} |T_{a,\ell}|^2 \leq \frac{n}{\xi n} \cdot (\xi n)^2 = \xi n^2.$$

Therefore, applying Lemma 2.6 with $t = \alpha n/4k_a$, we have that

$$\mathbb{P} \left(\sum_{\ell \in [r_a]} Z_{a,b,i,\ell} > \frac{|\varphi^{-1}(I_{a,b})|}{k_a} + \frac{\alpha n}{2k_a} \right) \leq \exp \left(-\frac{(\alpha n/4k_a)^2}{2\xi n^2} \right) \leq \exp \left(-\frac{\alpha^2}{32\xi k^2} \right) \leq \frac{1}{20k},$$

as required, where we have used that $k_a \leq k$ and $\xi \ll 1/k, \alpha \ll 1$. $\square$

Claim 4.4. *Suppose for some $0 \leq s < n$, $\psi(x_1), \dots, \psi(x_s)$ satisfy **C1–C3**, and for every $a \in [3]_0$, $b \in [3]$ and $i \in I_{a,b}$, we have*

$$\sum_{\ell \in [r_a]: p_{a,\ell} \leq s} Z_{a,b,i,\ell} \leq \frac{|\varphi^{-1}(I_{a,b})|}{k_a} + \frac{\alpha n}{2k_a}. \quad (4.1)$$

*Then, we can extend the embedding ψ to include x_{s+1} so that **C1–C3** still hold with $s+1$ in place of s .*

Proof of Claim 4.4. Let $s' \in [s]$ be such that $x_{s'}x_{s+1} \in E(T)$.

If $s+1 = p_{a,\ell}$ for some $a \in [3]_0$ and $\ell \in [r_a]$, then from above we have $\sigma(x_{s+1}) = i$ for some $i \in I_{a,1}$, and by **C3** there exists $W_{i,s'} \subset N(\psi(x_{s'}), V_i \setminus \psi(\{x_1, \dots, x_{s'-1}\}))$ with size $d\alpha n/8k_a$. Using **C3**, $\Delta(T) \leq cn$, and that components directly descending from $x_{s'}$ appear right after $x_{s'}$ in the ordering, we see that at most cn vertices in $W_{i,s'}$ have been used, thus the set $Y_{s+1} := W_{i,s'} \setminus \psi(\{x_1, \dots, x_s\})$ has size at least $d\alpha n/10k_a$.

If $s+1 \notin \{p_{a,\ell} : a \in [3]_0, \ell \in [r_a]\}$, then $\sigma(x_{s'}) \notin \{i_1, i_2, i_3\}$. If $\sigma(x_{s+1}) = i_a$ for some $a \in [3]$, then by **C2**, we have $d_G(\psi(x_{s'}), V_{i_a} \setminus \psi(\{x_1, \dots, x_{s'-1}\})) \geq d|V_{i_a}|/4$. By **B3**, $Y_{s+1} := N_G(\psi(x_{s'}), V_{i_a} \setminus \psi(\{x_1, \dots, x_s\}))$ has size at least $d|V_{i_a}|/4 - \xi n \geq d|V_{i_a}|/5$. If instead $\sigma(x_{s+1}) = i \in I_a$ for some $a \in [3]_0$, then $x_{s'}$ and x_s are in the same component of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$. Say this component is directly descended from $x_{s''} \in \varphi^{-1}(\{i_1, i_2, i_3\})$. By **C2**, we have $d_G(\psi(x_{s'}), V_i \setminus \psi(\{x_1, \dots, x_{s'-1}\})) \geq d\alpha n/4k_a$. Since vertices in the same component appear consecutively in the ordering, by **B1**, at most ξn vertices

are embedded between $x_{s'}$ and x_s . Thus, the set $Y_{s+1} := N_G(\psi(x_{s'}), V_i \setminus (\psi(\{x_1, \dots, x_s\}) \cup W_{i,s''}))$ has size at least $d\alpha n/4k_a - \xi n - d\alpha n/8k_a \geq d\alpha n/10k_a$.

We now embed x_{s+1} to a suitable vertex in Y_{s+1} by splitting into the following two cases, depending on whether $\sigma(x_{s+1}) \in \{i_1, i_2, i_3\}$.

Case I. $\sigma(x_{s+1}) = i_a$ for some $a \in [3]$. For every $i \in N_R(i_a) \setminus \{i_1, i_2, i_3\} \subset I_{a,1}$, by **C1**, (4.1) and **B2**, we have

$$|V_i \cap \psi(\{x_1, \dots, x_s\})| \leq \frac{|\varphi^{-1}(I_{a,1})|}{k_a} + \frac{\alpha n}{2k_a} \leq \frac{\sum_{i' \in I_{a,1}} |V_{i'}|}{k_a} - \frac{\alpha n}{2k_a}.$$

Since $|V_i| = \frac{1}{k_a} \sum_{i' \in I_{a,1}} |V_{i'}|$, we have $|V_i \setminus \psi(\{x_1, \dots, x_s\})| \geq \alpha n/2k_a \gg \varepsilon |V_i|$. Since $i_a i \in E(R)$, $G[V_{i_a}, V_i]$ is (ε, d) -regular, so by Lemma 2.21, for all but at most $\sqrt{\varepsilon} |V_{i_a}|$ vertices $y \in Y_{s+1}$,

$$d(y, V_i \setminus \psi(\{x_1, \dots, x_s\})) \geq d|V_i \setminus \psi(\{x_1, \dots, x_s\})|/2 \geq d\alpha n/4k_a > d\alpha n/8k_a$$

for all but at most $\sqrt{\varepsilon} k_a \leq \alpha |I_{a,1}|/100$ indices $i \in I_{a,1}$. Similarly, when $a = 1$, for all but at most $\sqrt{\varepsilon} |V_{i_1}|$ vertices $y \in Y_{s+1}$, $d(y, V_i \setminus \psi(\{x_1, \dots, x_s\})) \geq d\alpha n/8k_0$ for all but at most $\alpha |I_{0,1}|/100$ indices $i \in I_{0,1}$.

Furthermore, for each $i \in N_R(i_a) \cap \{i_1, i_2, i_3\}$, by using **B3** and Lemma 2.20 instead of **B2** and Lemma 2.21, we have that for all but at most $\varepsilon |V_{i_a}|$ vertices $y \in Y_{s+1}$, $d(y, V_i \setminus \psi(\{x_1, \dots, x_s\})) \geq d|V_i|/4$. As $|Y_{s+1}| \geq d|V_{i_a}|/5 \geq 10\sqrt{\varepsilon} |V_{i_a}|$, we can pick $\psi(x_{s+1}) \in Y_{s+1}$ such that all of the above hold, so **C1–C3** hold with $s+1$ in place of s , as required.

Case II. $\sigma(x_{s+1}) \in I_a$ for some $a \in [3]_0$. Similar to **Case I**, we can deduce from Lemma 2.20, **C1**, (4.1), and either **B2** or **B3** that for every $i \in N_R(\sigma(x_{s+1}), I_a)$, all but at most $\varepsilon |V_{\sigma(x_{s+1})}|$ vertices $y \in Y_{s+1}$ satisfy $d(y, V_i \setminus \psi(\{x_1, \dots, x_s\})) \geq d\alpha n/4k_a$, and for every $i \in N_R(\sigma(x_{s+1}), \{i_1, i_2, i_3\})$, all but at most $\varepsilon |V_{\sigma(x_{s+1})}|$ vertices $y \in Y_{s+1}$ satisfy $d(y, V_i \setminus \psi(\{x_1, \dots, x_s\})) \geq d|V_i|/4$. Then, using that $d_R(\sigma(x_{s+1})) \leq 2$, and $|Y_{s+1}| \geq d\alpha n/10k_a \geq 20\varepsilon n/k_a \geq 10\varepsilon |V_{\sigma(x_{s+1})}|$, we can pick $\psi(x_{s+1}) \in Y_{s+1}$ such that all of the above hold, so **C1–C3** hold with $s+1$ in place of s , as required. $\square$

Finally, note that by a union bound over all $a \in [3]_0, b \in [3]$, and $i \in I_{a,b}$, Claim 4.3 and Claim 4.4 combine to show that the process above embeds T into G with strictly positive probability, and thus G contains a copy of T . $\square$

4.3 Embedding method HLT

The following result appeared in the work of Haxell, Luczak, and Tingley [21]. For completion and to illustrate our method, we include a proof using our framework.

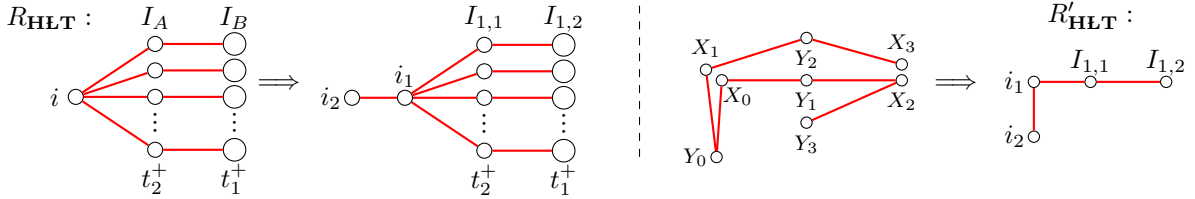


Figure 9: On the left, the slight refinement of the initial reduced graph R_{HLT} used in the proof of Lemma 4.5. On the right, a depiction of the rule of the embedding used at (4.3), condensing S from Lemma 4.1 into a subgraph R'_{HLT} of R' from Figure 8.

Lemma 4.5 (HLT). *Let $1/n \ll c \ll 1/k \ll \varepsilon \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 satisfying $t_2 \leq t_1 \leq 2t_2$. Let G be a graph on at most $2n$ vertices with a partition $V(G) = V_1 \cup \dots \cup V_{2k+1}$. Let R_{HLT} be a graph with vertex set $[2k+1]$, such that if $ij \in E(R_{\text{HLT}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Let $i \in [2k+1]$ and suppose there is a partition $[2k+1] \setminus \{i\} = I_A \cup I_B$, with $|I_A| = |I_B| = k$, such that the following hold for some m_A, m_B (see Figure 9).*

D1 $|V_a| = m_A$ for each $a \in I_A$, $|V_b| = m_B$ for each $b \in I_B$, and $|V_i| \geq n/10k$.

D2 $km_A \geq t_2 + \alpha n$ and $km_B \geq t_1 + \alpha n$.

D3 In $R_{\mathbf{HLT}}$, i is adjacent to each vertex in I_A , and there is a perfect matching M between I_A and I_B .

Then, there is a copy of T in G .

Proof. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1. Using this lemma, we can take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Assume, by relabelling if necessary, that $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$.

Let $i_1 = i$. Pick some $i_2 \in I_A$ and suppose it is matched with $i_3 \in I_B$ by M . Let $I_{1,1} = I_A \setminus \{i_2\}$ and $I_{1,2} = I_B \setminus \{i_3\}$ (see Figure 9). Let $R'_{\mathbf{HLT}}$ be the graph on the right in Figure 9, and note that it is a subgraph of R' in Figure 8. Then, for each $v \in V(T)$, as depicted in Figure 9, let

$$\varphi(v) = \begin{cases} i_1 & \text{if } \phi(v) \in \{X_0, X_1\} \\ i_2 & \text{if } \phi(v) = Y_0 \\ I_{1,1} & \text{if } \phi(v) \in \{Y_1, Y_2, Y_3\} \\ I_{1,2} & \text{if } \phi(v) \in \{X_2, X_3\} \end{cases},$$

so that φ is a homomorphism from T to $R'_{\mathbf{HLT}}$, and thus to R' , with $|\varphi^{-1}(\{i_1, i_2\})| \leq \xi n$, $|\varphi^{-1}(I_{1,1})| \leq t_2$, and $|\varphi^{-1}(I_{1,2})| \leq t_1$.

We will now check the conditions required for an application of Lemma 4.2. First, **B1** holds as every component of $T - \varphi^{-1}(\{i_1, i_2, i_3\})$ is contained in a component of $T - \phi^{-1}(X_0 \cup Y_0)$, which has size at most ξn . Next, using **D1**, **D2**, and $1/n \ll 1/k \ll \alpha$, we have

$$|\varphi^{-1}(I_{1,1})| + \alpha n/2 \leq t_2 + \alpha n/2 \leq (k-1)m_A = \sum_{j \in I_{1,1}} |V_j|,$$

and similarly $|\varphi^{-1}(I_{1,2})| + \alpha n/2 \leq \sum_{j \in I_{1,2}} |V_j|$, so **B2** holds as $\varphi^{-1}(I) = \emptyset$ for each $I \in V(R') \setminus \{i_1, i_2, I_{1,1}, I_{1,2}\}$. Finally, for each $j \in \{i_1, i_2, i_3\}$, $|\varphi^{-1}(j)| \leq |\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$, so **B3** holds. Therefore, we can apply Lemma 4.2 to find a copy of T in G , as required. $\square$

4.4 EM1a Embedding Method

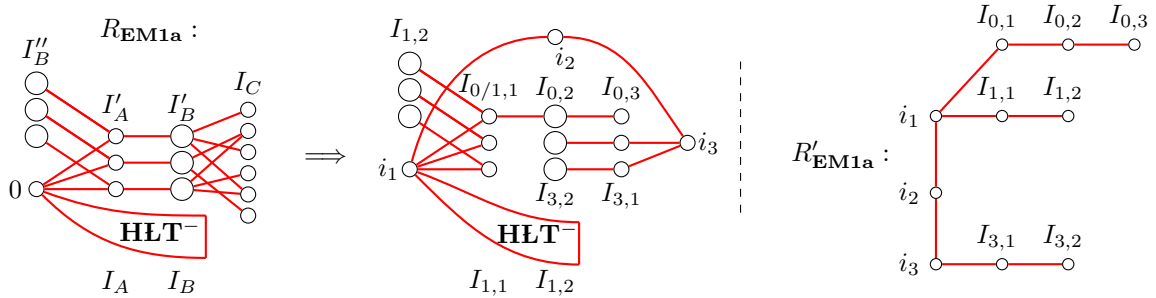


Figure 10: On the left, the initial reduced graph $R_{\mathbf{EM1a}}$ transformed into the subgraph used to embed the tree in Lemma 4.6. On the right, the auxiliary graph $R'_{\mathbf{EM1a}}$ used when applying Lemma 4.2.

Lemma 4.6 (EM1a). Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 satisfying $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with a partition $V(G) = V_0 \cup V_1 \cup \dots \cup V_k$. Let $R_{\mathbf{EM1a}}$ be a graph with vertex set $[k]_0$, such that if $ij \in E(R_{\mathbf{EM1a}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I'_A \cup I_B \cup I'_B \cup I''_B \cup I_C$, such that the following properties hold (see Figure 10).

E1 $|V_i| = m$ for all $i \in \{0\} \cup I_A \cup I'_A \cup I_C$ and $|V_i| = t_1 m / t_2$ for all $i \in I_B \cup I'_B \cup I''_B$.

E2 $|\cup_{i \in I_A \cup I'_A} V_i| \geq (1 - \alpha^2)t_2$, $|\cup_{i \in I_B \cup I'_B} V_i| \geq (1 - \alpha^2)t_1$, $|\cup_{i \in I'_A} V_i| = \alpha t_2$, $|\cup_{i \in I'_B} V_i| = |\cup_{i \in I''_B} V_i| = \alpha t_1$,
and $|\cup_{i \in I_C} V_i| \geq \frac{2}{3}t_2$.

E3 In $R_{\mathbf{EM1a}}$, 0 is adjacent to every vertex in $I_A \cup I'_A$.

E4 In $R_{\mathbf{EM1a}}$, there is a perfect matching between I_A and I_B , a perfect matching between I'_A and I'_B ,
and a perfect matching between I'_A and I''_B .

E5 In $R_{\mathbf{EM1a}}$, every vertex in I'_B is adjacent to at least $10\alpha|I_C|$ vertices in I_C .

Then, G contains a copy of T .

Proof. We begin with a claim that will also be used later in the proofs of Lemma 4.8 and Lemma 4.9.

Claim 4.7. For any $Z \subset I'_B$ with $\alpha^{-2} \ll |Z| \leq 5\alpha|I_C|$, there exists $z \in Z$ and $Z' \subset Z \setminus \{z\}$ with $|Z'| = 5\alpha|Z|$, and a matching M in $R_{\mathbf{EM1a}}[Z', I_C]$ covering Z' , with z adjacent to every $i \in I_C \cap V(M)$.

Proof of Claim 4.7. By **E5**, $e(R_{\mathbf{EM1a}}[Z, I_C]) \geq 10\alpha|Z||I_C|$, so there exists a set $I'_C \subset I_C$ with size $5\alpha|I_C|$ such that every $i \in I'_C$ has at least $5\alpha|Z| + 1$ neighbours in Z , as otherwise

$$e(R_{\mathbf{EM1a}}[Z, I_C]) < 5\alpha|I_C| \cdot |Z| + (1 - 5\alpha)|I_C| \cdot (5\alpha|Z| + 1) < 10\alpha|Z||I_C|,$$

a contradiction. Then, since $e(R_{\mathbf{EM1a}}[Z, I'_C]) \geq 5\alpha|Z||I'_C|$, by averaging, there exists $z \in Z$ with at least $5\alpha|I'_C|$ neighbours in I'_C . Let I''_C be a set of $5\alpha|I'_C|$ neighbours of z in I'_C . Greedily, and using $5\alpha|Z| \leq |I''_C|$, we can find a matching M between I''_C and $Z \setminus \{z\}$ with size $5\alpha|Z|$, which proves the claim. $\square$

Let $i_1 = 0$. Since $|I'_B| = \alpha t_2/m < 10\alpha t_2/3m \leq 5\alpha|I_C|$, we can apply Claim 4.7 to find $Z_B \subset I'_B$ with $|Z_B| = 5\alpha|I'_B| = 5\alpha^2 t_2/m$, $i_3 \in I'_B \setminus Z_B$, and a perfect matching M between Z_B and some $Z_C \subset I_C$, such that i_3 is adjacent to all $i \in Z_C$. Suppose that in the matchings given by **E4**, i_3 is matched with $i_2 \in I'_A$, and i_2 is matched with $i'_2 \in I''_B$.

Let $V'_i = V_i$ for all $i \in I_B \cup I''_B \cup Z_B \cup Z_C \cup \{i_1, i_2, i_3\}$. For every $i \in (I_A \cup I'_A) \setminus \{i_2\}$, let $V'_i \subset V_i$ have size $(1 - \alpha^2/2)|V_i|$. Let $I_{1,1} = (I_A \cup I'_A) \setminus \{i_2\}$ and $I_{1,2} = (I_B \cup I''_B) \setminus \{i'_2\}$. Partition Z_B as evenly as possible into two sets $I_{0,2}$ and $I_{3,2}$, and say they are matched by M with subsets $I_{0,3}$ and $I_{3,1}$ of Z_C , respectively. Note that i_3 is adjacent to every vertex in $I_{3,1}$. Finally, take a new index set $I_{0,1}$ with size $|I_{0,2}|$, say $I_{0,2}$ is matched with $I'_{0,1} \subset I'_A$ in the matching given by **E4**, and relabel the collection $\{V_i \setminus V'_i : i \in I'_{0,1}\}$ as $\{V'_j : j \in I_{0,1}\}$. Note that by Lemma 2.19, if ij is an edge in the graph depicted in the middle of Figure 10, then $G[V'_i, V'_j]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular.

Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1. Using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$, and note that each of these component has neighbours in exactly one of $\phi^{-1}(X_0)$ and $\phi^{-1}(Y_0)$. Thus, we can partition J as $J_X \cup J_Y$, such that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $J'_X \subset J_X$ and $J'_Y \subset J_Y$ both be random sets with each element being included independently with probability $2\alpha^2$. Then, by Lemma 2.7, with positive probability we have both of the following, so fix such a choice of J'_X, J'_Y .

$$\sum_{j \in J'_X \cup J'_Y} |K_j \cap \phi^{-1}(Y_1 \cup Y_2 \cup Y_3)| = 2\alpha^2 t_2 \pm \alpha^2 n/100. \quad (4.2)$$

$$\sum_{j \in J'_X \cup J'_Y} |K_j \cap \phi^{-1}(X_1 \cup X_2 \cup X_3)| = 2\alpha^2 t_1 \pm \alpha^2 n/100. \quad (4.3)$$

Let $R'_{\mathbf{EM1a}}$ be the graph on the right of Figure 10. Define a homomorphism $\varphi : T \rightarrow R'_{\mathbf{EM1a}}$ as follows. Let $\varphi(v) = i_1$ for every $v \in \phi^{-1}(X_0)$, and let $\varphi(v) = i_2$ for every $v \in \phi^{-1}(Y_0)$. For every K_j with $j \in J'_X$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{0,1}, I_{0,2}, I_{0,3}$, respectively; while if $j \in J_X \setminus J'_X$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively. For every K_j with $j \in J'_Y$, define φ on K_j by composing ϕ with the function

sending X_1, Y_2, X_3 to $i_3, I_{3,1}, I_{3,2}$, respectively; while for every K_j with $j \in J_Y \setminus J'_Y$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively.

Since $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$ and $\xi \ll 1/k \ll \alpha$, we have $|\varphi^{-1}(I)| \leq \xi n$ for each $I \in \{i_1, i_2, i_3\}$, and $|\varphi^{-1}(I_{0,1})| + \alpha^4 n/100 \leq \sum_{i \in I_{0,1}} |V'_i|$. From definition, $\sum_{i \in I_{0,2}} |V'_i|, \sum_{i \in I_{3,2}} |V'_i| \geq 2.4\alpha^2 t_1$ and $\sum_{i \in I_{0,3}} |V'_i|, \sum_{i \in I_{3,1}} |V'_i| \geq 2.4\alpha^2 t_2$. Thus, by (4.2), (4.3), and **E2**, we have $|\varphi^{-1}(I)| + \alpha^3 n/100 \leq \sum_{i \in I} |V'_i|$ for each $I \in \{I_{1,1}, I_{1,2}, I_{0,2}, I_{0,3}, I_{3,1}, I_{3,2}\}$. Therefore, we can apply Lemma 4.2 to find a copy of T in G . $\square$

4.5 EM1b Embedding Method

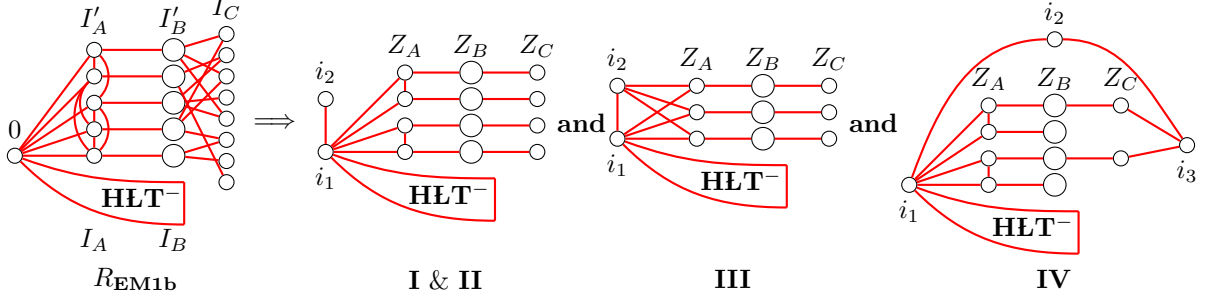


Figure 11: The initial reduced graph R_{EM1b} in Lemma 4.8 on the left, and the three substructures within that we use to embed the tree in **Cases I & II**, **Case III**, and **Case IV**, respectively.

Lemma 4.8 (EM1b). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \gamma \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 satisfying $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with a partition $V(G) = V_0 \cup V_1 \cup \dots \cup V_k$. Let R_{EM1b} be a graph with vertex set $[k]_0$, such that if $ij \in E(R_{\text{EM1b}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I'_A \cup I_B \cup I'_B \cup I_C$, such that the following properties hold (see Figure 11).*

F1 $|V_i| = m$ for all $i \in \{0\} \cup I_A \cup I'_A \cup I_C$ and $|V_i| = t_1 m/t_2$ for all $i \in I_B \cup I'_B$.

F2 $|\cup_{i \in I_A \cup I'_A} V_i| \geq (1 - \gamma)t_2, |\cup_{i \in I_B \cup I'_B} V_i| \geq (1 - \gamma)t_1, |\cup_{i \in I'_A} V_i| = 10\alpha t_2, |\cup_{i \in I'_B} V_i| = 10\alpha t_1$, and $|\cup_{i \in I_C} V_i| \geq \frac{2}{3}t_2$.

F3 In R_{EM1b} , 0 is adjacent to every vertex in $I_A \cup I'_A$.

F4 In R_{EM1b} , there exists a perfect matching between I_A and I_B , and a perfect matching between I'_A and I'_B .

F5 In R_{EM1b} , every vertex in I'_B is adjacent to at least $10\alpha|I_C|$ vertices in I_C .

F6 In R_{EM1b} , there exist at least $\alpha|I'_A|$ vertices with at least $\alpha|I'_A|$ neighbours in I'_A .

Then, G contains a copy of T .

Proof. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1. Using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$. Moreover, partition J as $J_X \cup J_Y$, so that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $\tau_{1,X} = |\phi^{-1}(X_2)|$, $\tau_{2,X} = |\phi^{-1}(Y_3)|$, $\tau_{1,Y} = |\phi^{-1}(X_3)|$, and $\tau_{2,Y} = |\phi^{-1}(Y_2)|$. Let $\gamma \ll \beta \ll \alpha$, and consider the following four cases. **I:** $\tau_{2,X} \geq 3\beta t_2$. **II:** $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 < (1 + 200\beta)t_2$. **III:** $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 \geq (1 + 200\beta)t_2$. **IV:** $\tau_{2,X} < 3\beta t_2$ and $\tau_{1,X} \geq 100\beta t_1$.

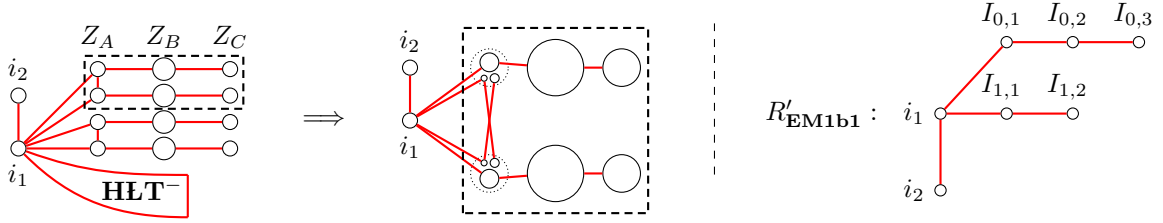


Figure 12: On the left, the transformation of the reduced graph structure used in **Cases I** and **II** to embed the tree. On the right, the auxiliary graph R'_{EM1b1} used when applying Lemma 4.2.

Cases I & II. By **F6**, we can greedily find a matching in $R_{\text{EM1b}}[I'_A]$ with size $\alpha|I'_A|/2 = 5\alpha^2 t_2/m$. Let $Z_A \subset I'_A$ be the set of vertices covered by this matching, and let $Z_B \subset I'_B$ be the vertices matched with Z_A . By **F5**, we can greedily find a perfect matching in R_{EM1b} between Z_B and some $Z_C \subset I'_C$.

Now, as depicted on the left of Figure 12, we further transform this structure into what we need to apply Lemma 4.2. Set $i_1 = 0$, pick i_2 arbitrarily from $I'_A \setminus Z_A$, and say it is matched with $i'_2 \in I'_B \setminus Z_B$. For each $a \in Z_A$, let $a' \in Z_A$ be its neighbour in the matching above. Let $\eta \geq \beta$ be a constant to be chosen later depending on whether we are in **Case I** or **Case II**. Pick disjoint subsets $V_{a,1}, V_{a,2} \subset V_a$ and $V_{a',1}, V_{a',2} \subset V_{a'}$ such that $|V_{a,1}| = |V_{a',1}| = \eta t_1 m/n$, $|V_{a,2}| = |V_{a',2}| = \eta t_2 m/n$. By Lemma 2.19, we can refine each of these new clusters, along with each cluster V_i with index i in $(I_A \cup I'_A) \setminus (Z_A \cup \{i_2\})$ and $(I_B \cup I'_B) \setminus (Z_B \cup \{i'_2\})$, into a maximum disjoint collection of smaller clusters with sizes γm or $\gamma t_1 m/t_2$ accordingly, then pair them up so that they form $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs. Note that at most $O(\gamma n)$ covered vertices are lost in this refinement process. Relabel these new refined clusters as $\{V'_i : i \in I_{1,1}\}$ and $\{V'_i : i \in I_{1,2}\}$ with $I_{1,1}$ indexing those with the smaller size. Then, we have

$$\begin{aligned} \sum_{i \in I_{1,1}} |V'_i| &\geq \sum_{i \in I_A \cup I'_A} |V_i| - m - |Z_A| \left(m - \frac{\eta t_2 m}{n} \right) - O(\gamma n) \\ &\geq (1 - O(\gamma))t_2 - 10\alpha^2 t_2 \left(1 - \frac{\eta t_2}{n} \right) \geq (1 - 10\alpha^2 + 3\alpha^2 \eta)t_2, \end{aligned}$$

and similarly $\sum_{i \in I_{1,2}} |V'_i| \geq (1 - 10\alpha^2 + 3\alpha^2 \eta)t_1$.

Then, relabel the subsets $\{V_i \setminus (V_{i,1} \cup V_{i,2}) : i \in Z_A\}$ as $\{V'_i : i \in I_{0,1}\}$, and relabel the subsets V_i with i in Z_B and Z_C as $\{V'_i : i \in I_{0,2}\}$ and $\{V'_i : i \in I_{0,3}\}$, respectively. Note that $\sum_{i \in I_{0,1}} |V'_i| = |Z_A|(1 - \eta)m = (10\alpha^2 - 10\alpha^2 \eta)t_2$, $\sum_{i \in I_{0,2}} |V'_i| = 10\alpha^2 t_1$, and $\sum_{i \in I_{0,3}} |V'_i| = 10\alpha^2 t_2$.

Let R'_{EM1b1} be the graph on the right of Figure 12. If we are in **Case I**, so $\tau_{2,X} \geq 3\beta t_2$, set $\eta = \beta$ and define a homomorphism $\varphi : T \rightarrow R'_{\text{EM1b1}}$ as follows. Let $\varphi(v) = i_1$ for every $v \in \phi^{-1}(X_0)$, and let $\varphi(v) = i_2$ for every $v \in \phi^{-1}(Y_0)$. For every K_j with $j \in J_X$, independently with probability $10\alpha^2 - \alpha^2 \beta$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{0,1}, I_{0,2}, I_{0,3}$, respectively; and with probability $1 - 10\alpha^2 + \alpha^2 \beta$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively. For every K_j with $j \in J_Y$, independently with probability $10\alpha^2 - \alpha^2 \beta$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $i_1, I_{0,1}, I_{0,2}$, respectively; and with probability $1 - 10\alpha^2 + \alpha^2 \beta$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively.

By Lemma 2.7, with positive probability we have $|\varphi^{-1}(I_{0,1} \cup I_{0,3})| = (10\alpha^2 - \alpha^2 \beta \pm \alpha^3 \beta)t_2$, $|\varphi^{-1}(I_{0,3})| \geq 20\alpha^2 \beta t_2$ using that $\tau_{2,X} \geq 3\beta t_2$, and $|\varphi^{-1}(I_{0,2})| = (10\alpha^2 - \alpha^2 \beta \pm \alpha^3 \beta)t_1$. Thus, $|\varphi^{-1}(I_{0,1})| \leq (10\alpha^2 - 20\alpha^2 \beta)t_2 \leq \sum_{i \in I_{0,1}} |V'_i| - 10\alpha^2 \beta t_2$, and similarly for $I_{0,2}$ and $I_{0,3}$. It also follows that

$$|\varphi^{-1}(I_{1,1})| \leq t_2 - |\varphi^{-1}(I_{0,1} \cup I_{0,3})| \leq (1 - 10\alpha^2 + 2\alpha^2 \beta)t_2 \leq \sum_{i \in I_{1,1}} |V'_i| - \alpha^2 \beta t_2,$$

and similarly for $I_{1,2}$. Therefore, we can apply Lemma 4.2 to find a copy of T in G .

If we are in **Case II** instead, so $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 < (1 + 200\beta)t_2$, then we proceed similarly to above with the role of X and Y swapped, and with $\eta = 1/2$. More specifically, we define a

homomorphism $\varphi : T \rightarrow R'_{\text{EM1b1}}$ as follows. Let $\varphi(v) = i_1$ for every $v \in \phi^{-1}(Y_0)$, and let $\varphi(v) = i_2$ for every $v \in \phi^{-1}(X_0)$. For every K_j with $j \in J_X$, independently with probability $9\alpha^2$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $i_1, I_{0,1}, I_{0,2}$, respectively; and with probability $1 - 9\alpha^2$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $i_1, I_{1,1}, I_{1,2}$, respectively. For every K_j with $j \in J_Y$, independently with probability $9\alpha^2$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $I_{0,1}, I_{0,2}, I_{0,3}$, respectively; and with probability $1 - 9\alpha^2$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively.

By Lemma 2.7, with positive probability we have $|\varphi^{-1}(I_{0,3})| \leq |\varphi^{-1}(I_{0,1} \cup I_{0,3})| = (9 \pm 0.1)\alpha^2 t_1$, $|\varphi^{-1}(I_{0,1})| \leq \xi n + 1000\alpha^2 \beta t_1$ using $\tau_{1,X} < 100\beta t_1$, and $|\varphi^{-1}(I_{0,2})| = (9 \pm 0.1)\alpha^2 t_2$. It follows using $t_1 < (1+200\beta)t_2$ that $|\varphi^{-1}(I)| \leq \sum_{i \in I} |V'_i| - \alpha^2 n/100$ for each $I \in \{I_{0,1}, I_{0,2}, I_{0,3}\}$. Using $t_1 < (1+200\beta)t_2$ again, we also get $|\varphi^{-1}(I_{1,1})| \leq (1 - 8.8\alpha^2)t_1 \leq (1 - 8.7\alpha^2)t_2 \leq \sum_{i \in I_{1,1}} |V'_i| - \alpha^2 n/100$, and similarly for $I_{1,2}$. Therefore, we can apply Lemma 4.2 to find a copy of T in G .

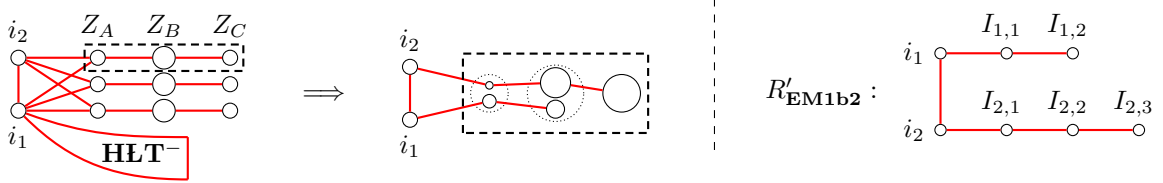


Figure 13: On the left, the transformation of the reduced graph structure used in **Case III** to embed the tree. On the right, the auxiliary graph R'_{EM1b2} used when applying Lemma 4.2.

Case III. In this case, we have $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 \geq (1 + 200\beta)t_2$. Thus, $\tau_{2,Y} \geq t_2 - \tau_{2,X} - \xi n \geq (1 - 4\beta)t_2$, and similarly $\tau_{1,Y} \geq (1 - 101\beta)t_1$. Set $i_1 = 0$. By **F6**, we can find $i_2 \in I'_A$ such that i_2 is adjacent to a set Z_A of $\alpha|I'_A| = 10\alpha^2 t_2/m$ vertices in I'_A . Let $i'_2 \in I'_B$ be the vertex matched with i_2 , and let $Z_B \subset I'_B$ be the vertices matched with Z_A . By **F5**, we can greedily find a perfect matching between Z_B and some subset Z_C of I_C .

For convenience, denote t_1/t_2 by ρ , so $1 + 200\beta \leq \rho \leq 2$ from assumptions. We now further transform this structure as depicted on the left of Figure 13. For each $a \in Z_A$, suppose it is matched with $b \in Z_B$, which is in turn matched with $c \in Z_C$. Take $U_a \subset V_a$ with size $2\sqrt{\varepsilon}m$, take $U_b \subset V_b$ with size m/ρ , and let $U_c = V_c$. Then, $G[U_a, U_b], G[U_b, U_c]$ are both $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular by Lemma 2.19. Relabel $\{U_a : a \in Z_A\}, \{U_b : b \in Z_B\}, \{U_c : c \in Z_C\}$ as $\{V'_i : i \in I_{2,1}\}, \{V'_i : i \in I_{2,2}\}, \{V'_i : i \in I_{2,3}\}$, respectively. Note that $\sum_{i \in I_{2,1}} |V'_i| = 20\alpha^2 \sqrt{\varepsilon} t_2 \gg \xi n$, $\sum_{i \in I_{2,2}} |V'_i| = 10\alpha^2 t_2/\rho$, and $\sum_{i \in I_{2,3}} |V'_i| = 10\alpha^2 t_2$.

Next, again for each $a \in Z_A$, suppose it is matched with $b \in Z_B$. Let $W_b = V_b \setminus U_b$, so $|W_b| = (\rho - 1/\rho)m$. Let $W_a \subset V_a \setminus U_a$ have size $(1 - 1/\rho^2)m$, possible as $\varepsilon \ll 1/\rho$. Observe that $|W_a| \geq (1 - (1 + 200\beta)^{-2})m \geq 200\beta m$. Refine the collections of clusters $\{W_a : a \in Z_A\}$ and $\{W_b : b \in Z_B\}$ above, and the clusters in $\{V'_i : i \in (I_A \cup I'_A) \setminus (Z_A \cup \{i_2\})\}$ and $\{V'_i : i \in (I_B \cup I'_B) \setminus (Z_B \cup \{i'_2\})\}$ down to clusters with sizes γm and $\gamma \rho m$ respectively. In the process, we lose $O(\gamma n)$ covered vertices, and the resulting refined clusters can be paired together again as $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs by Lemma 2.19. Relabel these refined clusters as $\{V'_i : i \in I_{1,1}\}$ and $\{V'_i : i \in I_{1,2}\}$, with $I_{1,1}$ indexing the smaller clusters. Note that

$$\begin{aligned} \sum_{i \in I_{1,1}} |V'_i| &\geq \sum_{i \in I_A \cup I'_A} |V_i| - \sum_{i \in Z_A \cup \{i_2\}} |V_i| + \sum_{a \in Z_A} |W_a| - O(\gamma n) \\ &\geq (1 - \gamma)t_2 - 10\alpha^2 t_2 - m + 10\alpha^2 t_2(1 - 1/\rho^2) - O(\gamma n) \geq (1 - (1 + \beta)10\alpha^2/\rho^2)t_2, \end{aligned}$$

and

$$\begin{aligned} \sum_{i \in I_{1,2}} |V'_i| &\geq \sum_{i \in (I_B \cup I'_B) \setminus \{i'_2\}} |V_i| - \sum_{b \in Z_B} |U_b| - O(\gamma n) \\ &\geq (1 - \gamma)t_1 - m - 10\alpha^2 t_2/\rho - O(\gamma n) \geq (1 - (1 + \beta)10\alpha^2/\rho^2)t_1. \end{aligned}$$

Since $\rho \geq 1 + 200\beta$, we have $\rho(1 - 10\beta)(1 - 101\beta) \geq 1 + 10\beta$, so we can find $p \in [0, 1]$ such that

$$\frac{10\alpha^2(1 + 10\beta)}{(1 - 101\beta)\rho^2} \leq p \leq \frac{10\alpha^2(1 - 10\beta)}{\rho}.$$

Let R'_{EM1b2} be the graph on the right of Figure 13. Define a homomorphism $\varphi : T \rightarrow R'_{\text{EM1b2}}$ as follows. Let $\varphi(v) = i_1$ for every $v \in \phi^{-1}(X_0)$, and let $\varphi(v) = i_2$ for every $v \in \phi^{-1}(Y_0)$. For every K_j with $j \in J_X$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively. For every K_j with $j \in J_Y$, independently with probability p , define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $I_{2,1}, I_{2,2}, I_{2,3}$, respectively; and with probability $1 - p$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively.

Using $\tau_{2,Y} \geq (1 - 4\beta)t_2$, $\tau_{1,Y} \geq (1 - 101\beta)t_1$, and Lemma 2.7, with positive probability we have $|\varphi^{-1}(I_{2,1})| \leq \xi n$,

$$(1 + 5\beta)10\alpha^2 t_2 / \rho^2 \leq p(1 - 4\beta)t_2 - 5\alpha^2 \beta t_2 \leq |\varphi^{-1}(I_{2,2})| = p\tau_{2,Y} \pm 5\alpha^2 \beta t_2 \leq (1 - 5\beta)10\alpha^2 t_2 / \rho,$$

and similarly $(1 + 5\beta)10\alpha^2 t_1 / \rho^2 \leq |\varphi^{-1}(I_{2,3})| \leq (1 - 5\beta)10\alpha^2 t_1 / \rho$. It follows that $|\varphi^{-1}(I)|$ is suitably smaller than $\sum_{i \in I} |V'_i|$ for each $I \in \{I_{2,1}, I_{2,2}, I_{2,3}\}$. Moreover, we have $|\varphi^{-1}(I_{1,1})| \leq (1 - (1 + 5\beta)10\alpha^2 / \rho^2)t_2 \leq \sum_{i \in I_{1,1}} |V'_i| - 40\alpha^2 \beta t_2 / \rho^2$, and similarly for $I_{1,2}$. Thus, we can apply Lemma 4.2 to find a copy of T in G .

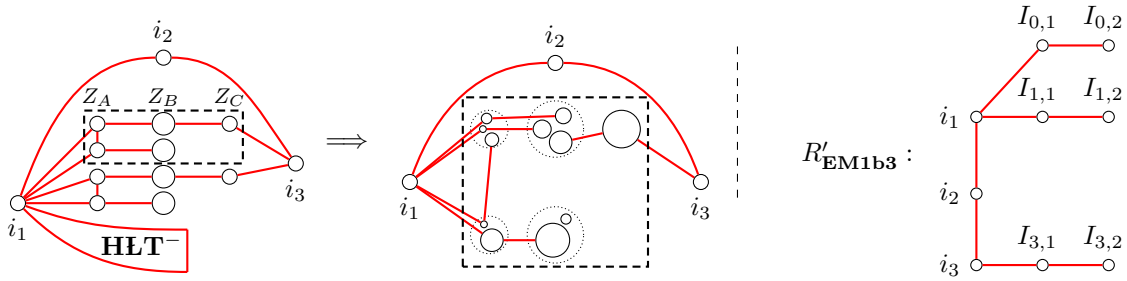


Figure 14: On the left, the transformation of the reduced graph structure used in **Case IV** to embed the tree. On the right, the auxiliary graph R'_{EM1b3} used when applying Lemma 4.2.

Case IV. In this case, we assume that $\tau_{2,X} < 3\beta t_2$ and $\tau_{1,X} \geq 100\beta t_1$. Then, $\tau_{2,Y} \geq (1 - 4\beta)t_2$ and $\tau_{1,Y} < (1 - 100\beta)t_1$. Let $\rho_Y = \tau_{1,Y} / \tau_{2,Y}$, and observe that $\rho_Y < (1 - 40\beta)t_1 / t_2$.

Let $i_1 = 0$. By **F6**, we can greedily find a matching M in $R_{\text{EM1b}}[I'_A]$ with size $\alpha|I'_A|/2 = 5\alpha^2 t_2 / m$. Let $Z'_A \subset I'_A$ be the set of vertices covered by this matching, so $|Z'_A| = 10\alpha^2 t_2 / m$, and let $Z'_B \subset I'_B$ be the vertices matched with Z'_A . By **F5**, and as in Claim 4.7, we can find a matching M' in $R_{\text{EM1b}}[Z'_B, I'_C]$ with size $5\alpha|Z'_B| = 50\alpha^3 t_2 / m$, and a vertex $i_3 \in Z'_B \setminus V(M')$ that is adjacent to every vertex in $I'_C \cap V(M')$. At the cost of halving the size, we can restrict M' to a perfect matching between some $Z_B \subset Z'_B$ and $Z_C \subset I'_C$ with $|Z_B| = |Z_C| = 25\alpha^3 t_2 / m$, such that if $Z_A \subset Z'_A$ is the set matched with Z_B , then vertices in Z_A all belong to the same side of the matching M . Let $i_2 \in Z'_A \setminus Z_A$ be the vertex matched with i_3 .

We now further transform this structure as depicted on the left of Figure 14. For every $a \in Z_A$, suppose it is matched with $a' \in Z'_A \setminus Z_A$ by M and with $b \in Z_B$. Say b is matched with $c \in Z_C$ by M' , and a' is matched with $b' \in Z'_B$. Let $U_c = V_c$, and pick $U_b \subset V_b$ with size $(\rho_Y + 2\sqrt{\varepsilon})m$. Note that $G[U_b, U_c]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular by Lemma 2.19. Relabel $\{U_c : c \in Z_C\}$ and $\{U_b : b \in Z_B\}$ as $\{V'_i : i \in I_{3,1}\}$ and $\{V'_i : i \in I_{3,2}\}$, respectively. Note that $\sum_{i \in I_{3,1}} |V'_i| = 25\alpha^3 t_2$ and $\sum_{i \in I_{3,2}} |V'_i| = 25\alpha^3 (\rho_Y + 2\sqrt{\varepsilon})t_2 \geq 25\alpha^3 \tau_{1,Y} + 50\alpha^3 \sqrt{\varepsilon} t_2$.

Since $\rho_Y < (1 - 40\beta)t_1 / t_2$, we have $|V_b \setminus U_b| \geq 39\beta t_1 m / t_2$, so $V_b \setminus U_b$ can be partitioned as $W_b \cup S_b$ with $|W_b| = 20\beta t_1 m / t_2$ and

$$|S_b| = (1 - 20\beta)t_1 m / t_2 - (\rho_Y + 2\sqrt{\varepsilon})m \geq 19\beta t_1 m / t_2.$$

Partition V_a into $W_a \cup S_a \cup L_a$ with $|W_a| = 20\beta m$, $|S_a| = 5\beta m$, and $|L_a| = (1 - 25\beta)m$. Recall that a' is the vertex in I'_A matched with a by M , and it is matched with $b' \in I'_B$. Take $L_{a'} \subset V_{a'}$ and $L_{b'} \subset V_{b'}$ with sizes $5\beta m$ and $5\beta t_1 m / t_2$, respectively. By shrinking exactly one of S_a and $L_{a'}$ if necessary, we can ensure $|L_a| / |L_{a'}| = |S_b| / |S_a|$, $\max\{|L_{a'}|, |S_a|\} = 5\beta m$, and $\min\{|L_{a'}|, |S_a|\} \geq \beta^2 m$.

Refine all pairs of clusters of the forms (W_a, W_b) and $(V_{a'} \setminus L_{a'}, V_{b'} \setminus L_{b'})$, and all other matched clusters in $\{V_i : i \in (I_A \cup I'_A) \setminus (Z_A \cup \{i_2\})\}$ and $\{V_i : i \in (I_B \cup I'_B) \setminus (Z_B \cup \{i_3\})\}$ down to clusters with sizes

γm and $\gamma t_1 m/t_2$ accordingly. In the refinement process, we lose $O(\gamma n)$ covered vertices, and the resulting refined clusters can still be paired together as $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs by Lemma 2.19. Relabel these refined clusters as $\{V'_i : i \in I_{1,1}\}$ and $\{V'_i : i \in I_{1,2}\}$, with $I_{1,1}$ indexing the smaller clusters. Note that

$$\begin{aligned} \sum_{i \in I_{1,1}} |V'_i| &\geq \sum_{i \in I_A \cup I'_A} |V_i| - \sum_{i \in Z_A \cup \{i_2\}} (1 + 5\beta) |V_i| + \sum_{a \in Z_A} |W_a| - O(\gamma n) \\ &\geq (1 - O(\gamma)) t_2 - 25\alpha^3 (1 + 5\beta) t_2 + 500\alpha^3 \beta t_2 \geq (1 - 25\alpha^3 + 370\alpha^3 \beta) t_2, \end{aligned}$$

and similarly $\sum_{i \in I_{1,2}} |V'_i| \geq (1 - 25\alpha^3 + 370\alpha^3 \beta) t_1$.

By Lemma 2.19, we can also refine all pairs of clusters of the forms (S_a, S_b) and $(L_{a'}, L_a)$ down to clusters with sizes γm and $\gamma |L_a| m / |L_{a'}|$, then pair the new clusters into $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs. Relabel the refined clusters as $\{V'_i : i \in I_{0,1}\}$ and $\{V'_i : i \in I_{0,2}\}$, with $I_{0,1}$ indexing the smaller clusters. Note that $\sum_{i \in I_{0,1}} |V'_i| \geq 5\beta m \cdot 25\alpha^3 t_2 / m - O(\gamma n) \geq 100\alpha^3 \beta t_2$. Moreover, using $\rho_Y = \tau_{1,Y} / \tau_{2,Y} \leq (t_1 - \tau_{1,X}) / \tau_{2,Y}$, we have

$$\begin{aligned} \sum_{i \in I_{0,2}} |V'_i| &\geq \sum_{a \in Z_A} |L_a| + \sum_{b \in Z_B} |S_b| - O(\gamma n) \\ &\geq 25\alpha^3 \left((1 - 25\beta) t_2 + (1 - 20\beta) t_1 - \frac{t_1 - \tau_{1,X}}{\tau_{2,Y}} t_2 - 2\sqrt{\varepsilon} t_2 \right) - O(\gamma n) \\ &\geq 25\alpha^3 \left((1 - 26\beta) t_2 + (1 - 20\beta) t_1 - \frac{t_1}{1 - 4\beta} + \tau_{1,X} \right) \geq 25\alpha^3 \tau_{1,X}. \end{aligned}$$

Let R'_{EM1b3} be the graph on the right of Figure 14. Define a homomorphism $\varphi : T \rightarrow R'_{\text{EM1b3}}$ as follows. Let $\varphi(v) = i_1$ for every $v \in \phi^{-1}(X_0)$, and let $\varphi(v) = i_2$ for every $v \in \phi^{-1}(Y_0)$. For every K_j with $j \in J_X$, independently with probability $p = 25\alpha^3 - \alpha^3 \beta$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{0,1}, I_{0,2}, I_{0,1}$, respectively; and with probability $1 - p$, define φ on K_j by composing ϕ with the function sending Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively. For every K_j with $j \in J_Y$, independently with probability $p = 25\alpha^3 - \alpha^3 \beta$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $i_3, I_{3,1}, I_{3,2}$, respectively; and with probability $1 - p$, define φ on K_j by composing ϕ with the function sending X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively.

By Lemma 2.7, with positive probability we have

$$(25\alpha^3 - 150\alpha^3 \beta) t_2 \leq 25\alpha^3 (1 - 2\beta) (1 - 4\beta) t_2 \leq |\varphi^{-1}(I_{3,1})| = p\tau_{2,Y} \pm \alpha^3 \beta t_2 / 2 \leq (25\alpha^3 - \alpha^3 \beta / 2) t_2,$$

$$|\varphi^{-1}(I_{0,2} \cup I_{3,2})| = p(\tau_{1,X} + \tau_{1,Y}) \pm \alpha^3 \beta t_1 \geq (25\alpha^3 - \alpha^3 \beta) (t_1 - \xi n) - \alpha^3 \beta t_1 \geq (25\alpha^3 - 3\alpha^3 \beta) t_1,$$

as well as $|\varphi^{-1}(I_{0,2})| \leq p\tau_{1,X} + \alpha^3 \beta \tau_{1,X} / 2 \leq (25\alpha^3 - \alpha^3 \beta / 2) \tau_{1,X}$, $|\varphi^{-1}(I_{3,2})| \leq p\tau_{1,Y} + \alpha^3 \sqrt{\varepsilon} t_2 \leq (25\alpha^3 - \alpha^3 \beta) \tau_{1,Y} + \alpha^3 \sqrt{\varepsilon} t_2$, and $|\varphi^{-1}(I_{0,1})| \leq p\tau_{2,X} \pm \alpha^3 \beta t_2 \leq 90\alpha^3 \beta t_2$.

It also follows that $|\varphi^{-1}(I_{1,1})| \leq t_2 - |\varphi^{-1}(I_{3,1})| \leq (1 - 25\alpha^3 + 150\alpha^3 \beta) t_2$ and $|\varphi^{-1}(I_{1,2})| \leq t_1 - |\varphi^{-1}(I_{0,2} \cup I_{3,2})| \leq (1 - 25\alpha^3 + 3\alpha^3 \beta) t_1$. Therefore, $|\varphi^{-1}(I)|$ is suitably smaller than $\sum_{i \in I} |V'_i|$ for each $I \in \{I_{1,1}, I_{1,2}, I_{0,1}, I_{0,2}, I_{3,1}, I_{3,2}\}$. Thus, we can apply Lemma 4.2 to find a copy of T in G . $\square$

4.6 EM1c Embedding Method

Lemma 4.9 (EM1c). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \gamma \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with a partition $V(G) = V_0 \cup V_1 \cup \dots \cup V_k$. Let R_{EM1c} be a graph with vertex set $[k]_0$, such that if $ij \in E(R_{\text{EM1c}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I_{A,1} \cup I_{A,2} \cup I_B \cup I_{B,1} \cup I_{B,2} \cup I_C$, such that the following properties hold (see Figure 15).*

- G1** $|V_i| = m$ for all $i \in \{0\} \cup I_A \cup I_{A,1} \cup I_{A,2} \cup I_C$ and $|V_i| = t_1 m / t_2$ for all $i \in I_B \cup I_{B,1} \cup I_{B,2}$.
- G2** $|\cup_{i \in I_A \cup I_{A,1} \cup I_{A,2}} V_i| \geq (1 - \gamma) t_2$, $|\cup_{i \in I_B \cup I_{B,1} \cup I_{B,2}} V_i| \geq (1 - \gamma) t_1$, $|\cup_{i \in I_{A,1}} V_i| \geq |\cup_{i \in I_{A,2}} V_i| = 10\alpha t_2$, $|\cup_{i \in I_{B,1}} V_i| \geq |\cup_{i \in I_{B,2}} V_i| = 10\alpha t_1$, and $|\cup_{i \in I_C} V_i| \geq \frac{2}{3} t_2$.
- G3** In R_{EM1c} , 0 is adjacent to every vertex in $I_A \cup I_{A,1} \cup I_{A,2}$.

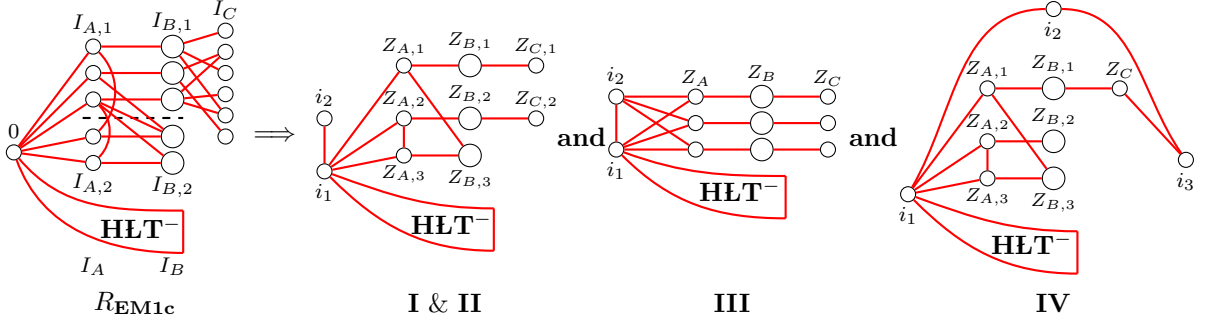


Figure 15: The initial reduced graph R_{EM1c} in Lemma 4.9 on the left, and the three substructures within used to embed the tree in **Cases I & II**, **Case III**, and **Case IV**, respectively.

G4 In R_{EM1c} , there exist perfect matchings between each of the three pairs of sets (I_A, I_B) , $(I_{A,1}, I_{B,1})$, and $(I_{A,2}, I_{B,2})$.

G5 In R_{EM1c} , every vertex in $I_{B,1}$ is adjacent to at least $10\alpha|I_C|$ vertices in I_C .

G6 In R_{EM1c} , for every $a \in I_{A,2}$ and $b \in I_{B,2}$ with $ab \in E(R_{\text{EM1c}})$, both a and b have at least $\alpha|I_{A,1}|$ neighbours in $I_{A,1}$.

Then, G contains a copy of T .

Proof. We proceed similarly to the proof of Lemma 4.8. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1 and, using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$. Moreover, partition J as $J_X \cup J_Y$, so that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $\tau_{1,X} = |\phi^{-1}(X_2)|$, $\tau_{2,X} = |\phi^{-1}(Y_3)|$, $\tau_{1,Y} = |\phi^{-1}(X_3)|$, and $\tau_{2,Y} = |\phi^{-1}(Y_2)|$. Let $\gamma \ll \beta \ll \alpha$, and consider the following four cases. **I:** $\tau_{2,X} \geq 3\beta t_2$. **II:** $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 < (1 + 200\beta)t_2$. **III:** $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 \geq (1 + 200\beta)t_2$. **IV:** $\tau_{2,X} < 3\beta t_2$ and $\tau_{1,X} \geq 100\beta t_1$.

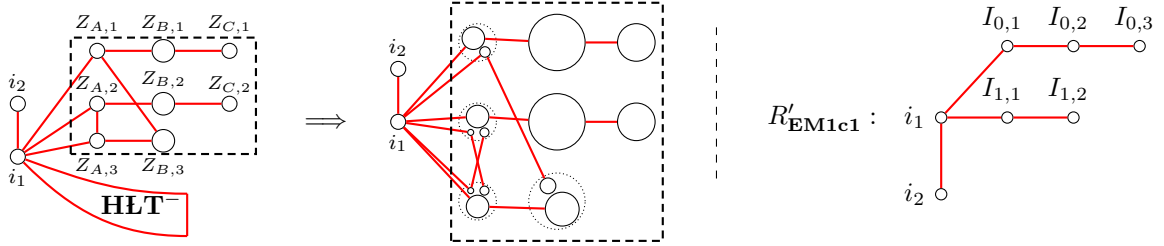


Figure 16: On the left, the transformation of the reduced graph structure used in **Cases I and II** to embed the tree. On the right, the auxiliary graph R'_{EM1c1} used when applying Lemma 4.2.

Cases I & II. Fix a submatching in $R_{\text{EM1c}}[I_{A,2}, I_{B,2}]$ with size $\alpha|I_{A,2}|/2 = 5\alpha^2 t_2/m$, say between $Z_{A,3}$ and $Z_{B,3}$. By **G6**, we can greedily find two disjoint matchings M_1 and M_2 in $R_{\text{EM1c}}[I_{A,1}, Z_{B,3}]$ and $R_{\text{EM1c}}[I_{A,1}, Z_{A,3}]$ covering $Z_{B,3}$ and $Z_{A,3}$, respectively. Let $Z_{A,1} = V(M_1) \cap I_{A,1}$ and $Z_{A,2} = V(M_2) \cap I_{A,1}$. Let $Z_{B,1}, Z_{B,2} \subset I_{B,1}$ be the vertices matched with $Z_{A,1}$ and $Z_{A,2}$, respectively. By **G5**, we can greedily find disjoint perfect matchings in $R_{\text{EM1c}}[Z_{B,1} \cup Z_{B,2}, I_C]$ between $Z_{B,1}$ and some $Z_{C,1} \subset I_C$, and between $Z_{B,2}$ and some $Z_{C,2} \subset I_C$. Let $Z_A = Z_{A,1} \cup Z_{A,2} \cup Z_{A,3}$ and $Z_B = Z_{B,1} \cup Z_{B,2} \cup Z_{B,3}$.

We now further transform this structure as depicted on the left of Figure 16. Set $i_1 = 0$, pick i_2 arbitrarily from $I_{A,2} \setminus Z_{A,3}$, and let i'_2 be the vertex in $I_{B,2} \setminus Z_{B,3}$ that i_2 is matched with. For each $a_3 \in Z_{A,3}$ matched with $b_3 \in Z_{B,3}$, say a_3 is matched with $a_2 \in Z_{A,2}$ under M_2 , and b_3 is matched with $a_1 \in Z_{A,1}$

under M_1 . Let $\eta \geq \beta$ be a constant to be chosen later depending on whether we are in **Case I** or **Case II**. Pick disjoint subsets $V_{a_3,1}, V_{a_3,2} \subset V_{a_3}$ and $V_{a_2,1}, V_{a_2,2} \subset V_{a_2}$ such that $|V_{a_3,1}| = |V_{a_2,1}| = \eta t_1 m/n$ and $|V_{a_3,2}| = |V_{a_2,2}| = \eta t_2 m/n$. Partition V_{a_1} as $V_{a_1,1} \cup V_{a_1,2}$ with $|V_{a_1,1}| = \eta m$ and $|V_{a_1,2}| = (1-\eta)m$. Partition V_{b_3} as $V_{b_3,1} \cup V_{b_3,2}$ with $|V_{b_3,1}| = \eta t_1 m/t_2$ and $|V_{b_3,2}| = (1-\eta)t_1 m/t_2$. By Lemma 2.19, we can refine all pairs of clusters of the forms $(V_{a_3,2}, V_{a_2,1}), (V_{a_2,2}, V_{a_3,1}), (V_{a_1,1}, V_{b_3,1}), (V_{a_3} \setminus (V_{a_3,1} \cup V_{a_3,2}), V_{b_3,2})$, along with all pairs of matched clusters indexed by $(I_A \cup I_{A,1} \cup I_{A,2}) \setminus (Z_A \cup \{i_2\})$ and $(I_B \cup I_{B,1} \cup I_{B,2}) \setminus (Z_B \cup \{i'_2\})$ into smaller clusters with sizes γm or $\gamma t_1 m/t_2$ accordingly, then pair them up again into $(\sqrt{\varepsilon}, d-\varepsilon)$ -regular pairs. Note that $O(\gamma n)$ covered vertices are lost in this refinement process. Relabel these new clusters as $\{V'_i : i \in I_{1,1}\}$ and $\{V'_i : i \in I_{1,2}\}$ with $I_{1,1}$ indexing those with the smaller size. Then, we have

$$\begin{aligned} \sum_{i \in I_{1,1}} |V'_i| &\geq \sum_{i \in I_A \cup I_{A,1} \cup I_{A,2}} |V_i| - m - 2|Z_{A,2}| \left(m - \frac{\eta t_2 m}{n} \right) - O(\gamma n) \\ &\geq (1 - O(\gamma))t_2 - 10\alpha^2 t_2 \left(1 - \frac{\eta t_2}{n} \right) \geq (1 - 10\alpha^2 + 3\alpha^2 \eta)t_2, \end{aligned}$$

and similarly $\sum_{i \in I_{1,2}} |V'_i| \geq (1 - 10\alpha^2 + 3\alpha^2 \eta)t_1$.

Then, let $U_i = V_{i,2}$ for every $i \in Z_{A,1}$ and let $U_i = V_i \setminus (V_{i,1} \cup V_{i,2})$ for every $i \in Z_{A,2}$. Relabel the subsets $\{U_i : i \in Z_{A,1} \cup Z_{A,2}\}$ as $\{V'_i : i \in I_{0,1}\}$, and similarly use $I_{0,2}$ and $I_{0,3}$ to relabel the subsets V_i with $i \in Z_{B,1} \cup Z_{B,2}$ and $i \in Z_{C,1} \cup Z_{C,2}$, respectively. Note that $\sum_{i \in I_{0,1}} |V'_i| = 2|Z_{A,1}|(1-\eta)m = (10\alpha^2 - 10\alpha^2 \eta)t_2$, $\sum_{i \in I_{0,2}} |V'_i| = 10\alpha^2 t_1$, and $\sum_{i \in I_{0,3}} |V'_i| = 10\alpha^2 t_2$. We are now in the same situation as **Cases I & II** in the proof of Lemma 4.8, so we can proceed in the same way to find a copy of T in G .

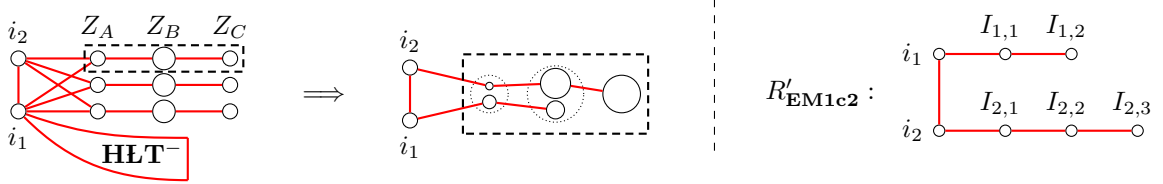


Figure 17: On the left, the transformation of the reduced graph structure used in **Case III** to embed the tree. On the right, the auxiliary graph R'_{EM1c2} used when applying Lemma 4.2.

Case III. $\tau_{2,X} < 3\beta t_2$, $\tau_{1,X} < 100\beta t_1$, and $t_1 \geq (1 + 200\beta)t_2$. Set $i_1 = 0$ and pick $i_2 \in I_{A,2}$ arbitrarily. By **G6**, there is a set $Z_A \subset N_{R_{EM1c}}(i_2, I_{A,1})$ with size $\alpha|I_{A,1}| = 10\alpha^2 t_2/m$. Let $Z_B \subset I_{B,1}$ be the indices matched with Z_A , and use **G5** to greedily find a perfect matching between Z_B and some $Z_C \subset I_C$. This is the same structure used in **Case III** of the proof of Lemma 4.8, so we can proceed in the same way to find an embedding of T in G .

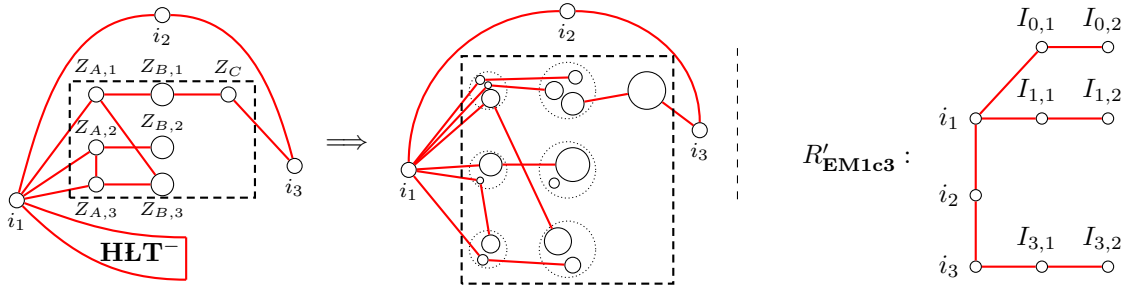


Figure 18: On the left, the transformation of the reduced graph structure used in **Case IV** to embed the tree. On the right, the auxiliary graph R'_{EM1c3} used when applying Lemma 4.2.

Case IV. $\tau_{2,X} < 3\beta t_2$ and $\tau_{1,X} \geq 100\beta t_1$. Let $\rho_Y = \tau_{1,Y}/\tau_{2,Y}$, and observe that $\rho_Y < (1 - 40\beta)t_1/t_2$. Fix a submatching in $R_{EM1c}[I_{A,2}, I_{B,2}]$ with size $\alpha|I_{A,1}|/2 = 5\alpha^2 t_2/m$, say between $Z'_{A,3}$ and $Z'_{B,3}$. By **G6**, we can greedily find two disjoint matchings M_1 and M_2 in $R_{EM1c}[I_{A,1}, Z'_{B,3}]$ and $R_{EM1c}[I_{A,1}, Z'_{A,3}]$ covering $Z'_{B,3}$ and $Z'_{A,3}$, respectively. Let $Z'_{A,1} = V(M_1) \cap I'_{A,1}$ and $Z'_{A,2} = V(M_2) \cap I'_{A,1}$, and let $Z'_{B,1}, Z'_{B,2} \subset I_{B,1}$

be the vertices matched with $Z'_{A,1}$ and $Z'_{A,2}$, respectively. Using **G5**, and as in Claim 4.7, we can find a matching M in $R_{\mathbf{EM1c}}[Z'_{B,1}, I_C]$ with size $5\alpha|Z'_{B,1}| = 25\alpha^3 t_2/m$, and a vertex $i_3 \in Z'_{B,1} \setminus V(M)$ that is adjacent to every vertex in $I_C \cap V(M)$. Say M matches $Z_{B,1} \subset Z'_{B,1}$ with $Z_C \subset I_C$, let $Z_{A,1} \subset Z'_{A,1}$ be the set matched with $Z_{B,1}$, and let $i_2 \in Z'_{A,1} \setminus Z_{A,1}$ be the vertex matched with i_3 . Let $Z_{B,3} \subset Z'_{B,3}$ be matched with $Z_{A,1}$ by M_1 , and suppose $Z_{B,3}$ is matched with $Z_{A,3} \subset Z'_{A,3}$, which is in turn matched with $Z_{A,2} \subset Z'_{A,2}$ by M_2 . Finally, let $Z_{B,2} \subset Z'_{B,2}$ be matched with $Z_{A,2}$, set $Z_A = Z_{A,1} \cup Z_{A,2} \cup Z_{A,3}$, and set $Z_B = Z_{B,1} \cup Z_{B,2} \cup Z_{B,3}$.

We now further transform this structure as depicted on the left of Figure 18. Set $i_1 = 0$. For each $a_3 \in Z_{A,3}$ matched with $b_3 \in Z_{B,3}$, say a_3 is matched with $a_2 \in Z_{A,2}$ by M_2 , which is matched with $b_2 \in Z_{B,2}$. Say b_3 is matched with $a_1 \in Z_{A,1}$ by M_1 , which is matched with $b_1 \in Z_{B,1}$, which is in turn matched with $c \in Z_C$ by M . Let $U_c = V_c$, and pick $U_{b_1} \subset V_{b_1}$ with size $(\rho_Y + 2\sqrt{\varepsilon})m$. Note that $G[U_{b_1}, U_c]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular by Lemma 2.19. Relabel $\{U_c : c \in Z_C\}$ and $\{U_{b_1} : b_1 \in Z_{B,1}\}$ as $\{V'_i : i \in I_{3,1}\}$ and $\{V'_i : i \in I_{3,2}\}$, respectively. Note that $\sum_{i \in I_{3,1}} |V'_i| = 25\alpha^3 t_2$ and $\sum_{i \in I_{3,2}} |V'_i| = 25\alpha^3 (\rho_Y + 2\sqrt{\varepsilon}) t_2 \geq 25\alpha^3 \tau_{1,Y} + 50\alpha^3 \sqrt{\varepsilon} t_2$.

Since $\rho_Y < (1 - 40\beta)t_1/t_2$, we have $|V_{b_1} \setminus U_{b_1}| \geq 39\beta t_1 m/t_2$, so $V_{b_1} \setminus U_{b_1}$ can be partitioned as $W_{b_1} \cup L_{b_1}$ with $|W_{b_1}| = 20\beta t_1 m/t_2$ and

$$|L_{b_1}| = (1 - 20\beta)t_1 m/t_2 - (\rho_Y + 2\sqrt{\varepsilon})m \geq 19\beta t_1 m/t_2.$$

Partition V_{a_1} into $W_{a_1} \cup W'_{a_1} \cup L_{a_1}$ with $|W_{a_1}| = 20\beta m$, $|W'_{a_1}| = (1 - 25\beta)m$, and $|L_{a_1}| = 5\beta m$. Partition V_{b_3} as $W_{b_3} \cup W'_{b_3}$ with $|W_{b_3}| = 25\beta t_1 m/t_2$ and $|W'_{b_3}| = (1 - 25\beta)t_1 m/t_2$. Partition V_{a_3} as $W_{a_3} \cup L_{a_3}$ with $|W_{a_3}| = 25\beta m$ and $|L_{a_3}| = (1 - 25\beta)m$. Partition V_{a_2} as $W_{a_2} \cup L_{a_2}$ with $|W_{a_2}| = (1 - 5\beta)m$ and $|L_{a_2}| = 5\beta m$. Partition V_{b_2} as $W_{b_2} \cup L_{b_2}$ with $|W_{b_2}| = (1 - 5\beta)t_1 m/t_2$ and $|L_{b_2}| = 5\beta t_1 m/t_2$. By shrinking exactly one of L_{a_1} or L_{a_2} if necessary, we can ensure $|L_{b_1}|/|L_{a_1}| = |L_{a_3}|/|L_{a_2}|$, $\max\{|L_{a_1}|, |L_{a_2}|\} = 5\beta m$, and $\min\{|L_{a_1}|, |L_{a_2}|\} \geq \beta^2 m$.

Refine all pairs of clusters of the forms (W_{a_1}, W_{b_1}) , (W'_{a_1}, W'_{b_1}) , (W_{a_3}, W_{b_3}) , (W_{a_2}, W_{b_2}) , and all matched clusters in $\{V_i : i \in (I_A \cup I_{A,1} \cup I_{A,2}) \setminus (Z_A \cup \{i_2\})\}$ and $\{V_i : i \in (I_B \cup I_{B,1} \cup I_{B,2}) \setminus (Z_B \cup \{i_3\})\}$ down to clusters with sizes γm or $\gamma t_1 m/t_2$ accordingly. In the refinement process, we lose $O(\gamma n)$ covered vertices, and the resulting refined clusters can be paired together again as $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs by Lemma 2.19. Relabel these refined clusters as $\{V'_i : i \in I_{1,1}\}$ and $\{V'_i : i \in I_{1,2}\}$, with $I_{1,1}$ indexing the smaller clusters. Note that

$$\begin{aligned} \sum_{i \in I_{1,1}} |V'_i| &\geq \sum_{i \in I_A \cup I_{A,1} \cup I_{A,2}} |V_i| - m - |Z_{A,1}|(5\beta + 5\beta + (1 - 25\beta))m - O(\gamma n) \\ &\geq (1 - O(\gamma))t_2 - 25\alpha^3(1 - 15\beta)t_2 \geq (1 - 25\alpha^3 + 370\alpha^3\beta)t_2, \end{aligned}$$

and similarly $\sum_{i \in I_{1,2}} |V'_i| \geq (1 - 25\alpha^3 + 370\alpha^3\beta)t_1$.

Next, by Lemma 2.19, we can refine all pairs of clusters of the forms (L_{a_2}, L_{a_3}) and (L_{a_1}, L_{b_1}) down to clusters with sizes γm and $\gamma|L_{b_1}|m/|L_{a_1}|$, and pair them up again into $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs. Relabel the refined clusters as $\{V'_i : i \in I_{0,1}\}$ and $\{V'_i : i \in I_{0,2}\}$, with $I_{0,1}$ indexing the smaller clusters. Note that $\sum_{i \in I_{0,1}} |V'_i| \geq 5\beta m \cdot 25\alpha^3 t_2/m - O(\gamma n) \geq 100\alpha^3 \beta t_2$. Moreover, using $\rho_Y = \tau_{1,Y}/\tau_{2,Y} \leq (t_1 - \tau_{1,X})/\tau_{2,Y}$, we have

$$\begin{aligned} \sum_{i \in I_{0,2}} |V'_i| &\geq \sum_{a_3 \in Z_{A,3}} |L_{a_3}| + \sum_{b_1 \in Z_{B,1}} |L_{b_1}| - O(\gamma n) \\ &\geq 25\alpha^3 \left((1 - 25\beta)t_2 + (1 - 20\beta)t_1 - \frac{t_1 - \tau_{1,X}}{\tau_{2,Y}} t_2 - 2\sqrt{\varepsilon} t_2 \right) - O(\gamma n) \\ &\geq 25\alpha^3 \left((1 - 26\beta)t_2 + (1 - 20\beta)t_1 - \frac{t_1}{1 - 4\beta} + \tau_{1,X} \right) \geq 25\alpha^3 \tau_{1,X}. \end{aligned}$$

We are now in the same situation as **Case IV** in the proof of Lemma 4.8, so we can proceed in the same way to find an embedding of T in G . $\square$

4.7 EM2a Embedding Method

Lemma 4.10 (EM2a). *Let $1/n \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with at most*

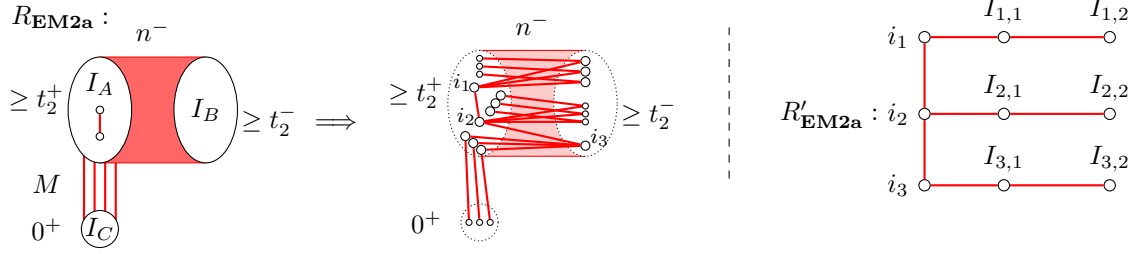


Figure 19: On the left, the initial reduced graph R_{EM2a} transformed into the substructure used to embed the tree in Lemma 4.10. On the right, the auxiliary graph R'_{EM2a} used when applying Lemma 4.2.

$2n$ vertices and a vertex partition $V_1 \cup \dots \cup V_k$ such that $|V_1| = |V_2| = \dots = |V_k| = m$. Let R_{EM2a} be a graph with vertex set $[k]$, such that if $ij \in E(R_{\text{EM2a}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I_B \cup I_C$ such that the following properties hold (see Figure 19).

H1 $|\cup_{i \in I_A} V_i| \geq t_2 + 200\alpha n, |\cup_{i \in I_B} V_i| \geq t_2 - \alpha n, |\cup_{i \in I_A \cup I_B} V_i| = n - \alpha n$, and $|\cup_{i \in I_C} V_i| = 100\alpha n$.

H2 $R_{\text{EM2a}}[I_A, I_B]$ is an η -almost complete bipartite graph.

H3 $R_{\text{EM2a}}[I_A, I_C]$ contains a matching M covering I_C .

H4 $E(R_{\text{EM2a}}[I_A] - V(M)) \neq \emptyset$.

Then, G contains a copy of T .

Proof. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1 and, using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$. Moreover, partition J as $J_X \cup J_Y$, so that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $\tau_{1,X} = |\phi^{-1}(X_2)|$, $\tau_{2,X} = |\phi^{-1}(Y_3)|$, $\tau_{1,Y} = |\phi^{-1}(X_3)|$ and $\tau_{2,Y} = |\phi^{-1}(Y_2)|$. Let $n_A = |\cup_{i \in I_A} V_i|$, $n_B = |\cup_{i \in I_B} V_i|$ and $n_C = |\cup_{i \in I_C} V_i|$. We now separate into the following two cases. **I:** $\tau_{1,X} + \tau_{2,Y} \geq n_B + 20\alpha n$. **II:** $\tau_{1,X} + \tau_{2,Y} < n_B + 20\alpha n$.

Case I. $\tau_{1,X} + \tau_{2,Y} \geq n_B + 20\alpha n$. Then, $(1 - 20\alpha)(\tau_{1,X} + \tau_{2,Y}) \geq n_B - \alpha n$. Note that $\tau_{2,X} + \tau_{2,Y} \leq t_2 \leq n_B + \alpha n$, so $(1 - 20\alpha)(\tau_{2,X} + \tau_{2,Y}) \leq n_B - \alpha n$. Thus, we can find $p \in [0, 1 - 20\alpha]$ such that $(1 - 20\alpha)\tau_{2,Y} + (1 - 20\alpha - p)\tau_{2,X} + p\tau_{1,X} = n_B - \alpha n$.

Let R'_{EM2a} be the graph depicted on the right in Figure 19, which is a subgraph of the graph R' in Lemma 4.2. Define a random homomorphism $\varphi : T \rightarrow R'_{\text{EM2a}}$ as follows. First, for vertices in $\phi^{-1}(X_0 \cup Y_0)$, set $\varphi(v) = i_3$ if $\phi(v) = Y_0$, and set $\varphi(v) = i_2$ if $\phi(v) = X_0$.

For each $j \in J_X$ independently at random, with probability 20α , define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $i_3, I_{3,1}, I_{3,2}$, respectively; with probability p define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $i_1, I_{1,1}, I_{1,2}$, respectively; and with probability $1 - 20\alpha - p$ define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $I_{2,1}, I_{2,2}, I_{2,1}$, respectively.

For each $j \in J_Y$ independently at random, with probability 20α , define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $I_{3,1}, I_{3,2}, I_{3,1}$, respectively; and with probability $1 - 20\alpha$ define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_2, I_{2,1}, I_{2,2}$, respectively.

Since $|K_j| \leq \xi n$ for all $j \in J$, and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$, we can use Lemma 2.7 to conclude that with strictly positive probability, $|\varphi^{-1}(I_{3,1})| = 20\alpha t_1 \pm \alpha t_1$, $|\varphi^{-1}(I_{3,2})| = 20\alpha t_2 \pm \alpha t_2 \leq n_C - \alpha n$, and $|\varphi^{-1}(I_{1,1} \cup I_{2,1})| = n_B - \alpha n \pm \alpha n/2$. It follows that $|\varphi^{-1}(\{I_{1,2}, I_{2,2}, I_{3,1}\})| \leq n - 19\alpha t_2 - (n_B - 3\alpha n/2) \leq n_A + 3\alpha n - 19\alpha t_2 \leq n_A - \alpha n$.

Using **H2–H4**, we can find $i_1, i_2 \in I_A \setminus V(M)$ and $i_3 \in I_B$, such that $i_1 i_2, i_2 i_3 \in E(R_{\text{EM2a}})$, and i_3 is adjacent to all but at most ηk vertices in $I_A \cap V(M)$. Thus, we can find $I_{3,1} \subset N_{R_{\text{EM2a}}}(i_3, I_A \cap V(M))$ covering $|\varphi^{-1}(I_{3,1})| + \alpha n/10$ vertices. Let $I_{3,2} \subset I_C$ denote the set of vertices matched with $I_{3,1}$ by M , then $|\varphi^{-1}(I_{3,b})| \leq \sum_{i \in I_{3,b}} |V_i| - \alpha n/10$ for each $b \in [2]$.

Let $I'_A = I_A \setminus (I_{3,1} \cup \{i_1, i_2\})$ and $I'_B = (N_{R_{\text{EM2a}}}(i_1, I_B) \cap N_{R_{\text{EM2a}}}(i_2, I_B)) \setminus \{i_3\}$. Note that from above, $|\varphi^{-1}(\{I_{1,2}, I_{2,2}\})| \leq n_A - \alpha n - |\varphi^{-1}(I_{3,1})| \leq \sum_{i \in I'_A} |V_i| - \alpha n/5$, and $|\varphi^{-1}(I_{1,1} \cup I_{2,1})| \leq n_B - \alpha n/2$, so we can find partitions $I'_A = I'_{1,2} \cup I'_{2,2}$ and $I'_B = I'_{1,1} \cup I'_{2,1}$, such that $\sum_{i \in I'_{a,b}} |V_i| \geq |\varphi^{-1}(I_{a,b})| + \alpha n/20$ for each $a, b \in [2]$. Moreover, as $|I'_{a,b}| \geq \alpha n/20m$ by construction for each $a, b \in [2]$, and $R_{\text{EM2a}}[I_A, I_B]$ is η -almost complete, by choosing these two partitions randomly and applying Lemma 2.5, we can find a realisation such that both $R_{\text{EM2a}}[I'_{1,1}, I'_{1,2}]$ and $R_{\text{EM2a}}[I'_{2,1}, I'_{2,2}]$ are 10η -almost complete. This allows us to apply Lemma 2.24 to refine the clusters V_i with $i \in I'_{1,1} \cup I'_{1,2} \cup I'_{2,1} \cup I'_{2,2}$ to obtain, for each $a, b \in [2]$, a set $\{V'_i : i \in I_{a,b}\}$ of clusters of the same size, such that $\sum_{i \in I_{a,b}} |V'_i| \geq |\varphi^{-1}(I_{a,b})| + \alpha n/100$. Moreover, for each $a \in [2]$, the refined clusters indexed by $I_{a,1}$ and $I_{a,2}$ can be matched up, so that each pair is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular (see Figure 19). This allows us to use Lemma 4.2 to find a copy of T in G .

Case II. $\tau_{1,X} + \tau_{2,Y} < n_B + 20\alpha n$. Then, $(1 - 70\alpha)(\tau_{1,X} + \tau_{2,Y}) \leq n_B - \alpha n$. Note that $\tau_{1,X} + \tau_{1,Y} \geq t_1 - \xi n \geq n_B + 100\alpha n$, so $(1 - 70\alpha)(\tau_{1,X} + \tau_{1,Y}) \geq n_B - \alpha n$. Therefore, there exists $p \in [0, 1 - 70\alpha]$ such that $(1 - 70\alpha)\tau_{1,X} + (1 - 70\alpha - p)\tau_{1,Y} + p\tau_{2,Y} = n_B - \alpha n$.

Similar to above, we define a random homomorphism $\varphi : T \rightarrow R'_{\text{EM2a}}$ as follows. First, for vertices in $\varphi^{-1}(X_0 \cup Y_0)$, set $\varphi(v) = i_2$ if $\phi(v) = Y_0$, and set $\varphi(v) = i_3$ if $\phi(v) = X_0$.

For each $j \in J_X$ independently at random, with probability 70α , define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $I_{3,1}, I_{3,2}, I_{3,1}$, respectively; and with probability $1 - 70\alpha$ define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $i_2, I_{2,1}, I_{2,2}$, respectively.

For each $j \in J_Y$ independently at random, with probability 70α , define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_3, I_{3,1}, I_{3,2}$, respectively; with probability p define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively; and with probability $1 - 70\alpha - p$ define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $I_{2,1}, I_{2,2}, I_{2,1}$, respectively.

Again, we can use Lemma 2.7 to conclude that with positive probability, $|\varphi^{-1}(I_{3,1})| = 70\alpha t_2 \pm \alpha t_2$, $|\varphi^{-1}(I_{3,2})| = 70\alpha t_1 \pm \alpha t_1 \leq n_C - \alpha n$, and $|\varphi^{-1}(I_{1,1} \cup I_{2,1})| = n_B - \alpha n \pm \alpha n/2$. It follows that $|\varphi^{-1}(\{I_{1,2}, I_{2,2}\})| + |\varphi^{-1}(I_{3,2})| \leq n - 69\alpha t_2 - (n_B - 3\alpha n/2) \leq n_A + 3\alpha n - 69\alpha t_2 \leq n_A - \alpha n$.

Similar to **Case I** above, by **H2-H4** and after refining, we can find $i_1, i_2 \in I_A$ and $i_3 \in I_B$, along with three matchings of refined clusters of suitable sizes in G attached to i_1, i_2, i_3 respectively, as depicted in Figure 19, which allow us to apply Lemma 4.2 to find a copy of T in G . $\square$

4.8 EM2b Embedding Method

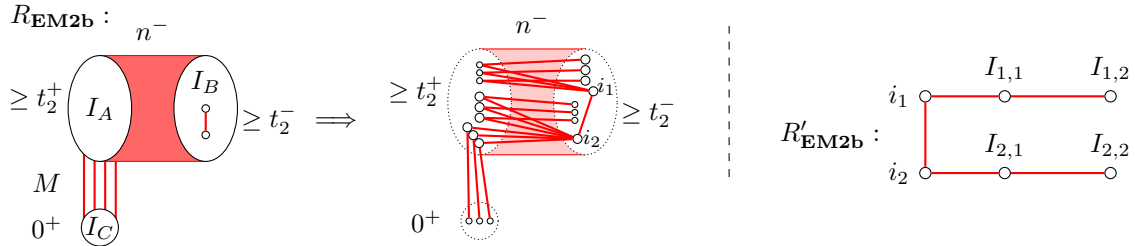


Figure 20: On the left, the initial reduced graph R_{EM2b} transformed into the substructure used to embed the tree in Lemma 4.11. On the right, the auxiliary graph R'_{EM2b} used when applying Lemma 4.2.

Lemma 4.11 (EM2b). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with at most $2n$ vertices and a vertex partition $V_1 \cup \dots \cup V_k$ with $|V_1| = \dots = |V_k| = m$. Let R_{EM2b} be a graph with vertex set $[k]$, such that if $ij \in E(R_{\text{EM2b}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I_B \cup I_C$ such that the following properties hold (see Figure 20).*

- I1** $|\cup_{i \in I_A} V_i| \geq t_2 + 200\alpha n$, $|\cup_{i \in I_B} V_i| \geq t_2 - \alpha n$, $|\cup_{i \in I_A \cup I_B} V_i| = (1 - \alpha)n$, and $|\cup_{i \in I_C} V_i| = 100\alpha n$.
- I2** $R_{\text{EM2b}}[I_A, I_B]$ is an η -almost complete bipartite graph.

I3 $R_{\text{EM2b}}[I_A, I_C]$ contains a matching M covering I_C .

I4 $E(R_{\text{EM2b}}[I_B]) \neq \emptyset$.

Then, G contains a copy of T .

Proof. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1. Using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$, and partition J as $J_X \cup J_Y$, so that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $\tau_1 = |\phi^{-1}(X_2 \cup X_3)|$ and $\tau_2 = |\phi^{-1}(Y_2 \cup Y_3)|$. Let $n_A = |\cup_{i \in I_A} V_i|$, $n_B = |\cup_{i \in I_B} V_i|$, and $n_C = |\cup_{i \in I_C} V_i|$. Note that $\tau_2 \leq t_2 \leq n_B + \alpha n$, so $(1 - 20\alpha)\tau_2 \leq n_B - \alpha n$. On the other hand, $(1 - 20\alpha)\tau_1 \geq (1 - 20\alpha)(t_1 - \xi n) \geq n_B - \alpha n$ as $n_B \leq t_1 - 201\alpha n$, so there exists $p \in [0, 1 - 20\alpha]$ such that $p\tau_1 + (1 - 20\alpha - p)\tau_2 = n_B - \alpha n$.

Let R'_{EM2b} be the graph depicted on the right in Figure 20, which is a subgraph of the graph R' in Lemma 4.2. Define a random homomorphism $\varphi : T \rightarrow R'_{\text{EM2a}}$ as follows. First, for vertices in $\phi^{-1}(X_0 \cup Y_0)$, set $\varphi(v) = i_1$ if $\phi(v) = X_0$, and set $\varphi(v) = i_2$ if $\phi(v) = Y_0$.

For each $j \in J_X$ independently at random, with probability $1 - p$, define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $i_2, I_{2,1}, I_{2,2}$, respectively; and with probability p define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively.

For each $j \in J_Y$ independently at random, with probability $1 - p$, define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $I_{2,1}, I_{2,2}, I_{2,1}$, respectively; and with probability p define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively.

Since $|K_j| \leq \xi n$ for all $j \in J$ and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$, we can use Lemma 2.7 to conclude that with strictly positive probability, $|\varphi^{-1}(I_{1,2} \cup I_{2,2})| = p\tau_1 + (1 - p)\tau_2 \pm \alpha n \leq n_B + 20\alpha\tau_2 \leq n_B + n_C - 10\alpha n$, and $|\varphi^{-1}(I_{1,1} \cup I_{2,1})| = (1 - p)\tau_1 + p\tau_2 \pm \alpha n \leq n_A - \alpha n$. Therefore, like in the proof of Lemma 4.10, by refining and using **I2–I4**, we can find $i_1, i_2 \in I_B$ such that $i_1 i_2 \in E(R_{\text{EM2b}})$, along with two matchings of refined clusters of suitable sizes forming $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs attached to i_1, i_2 as depicted in Figure 20, which allow us to apply Lemma 4.2 to find a copy of T in G . $\square$

4.9 EM2c Embedding Method

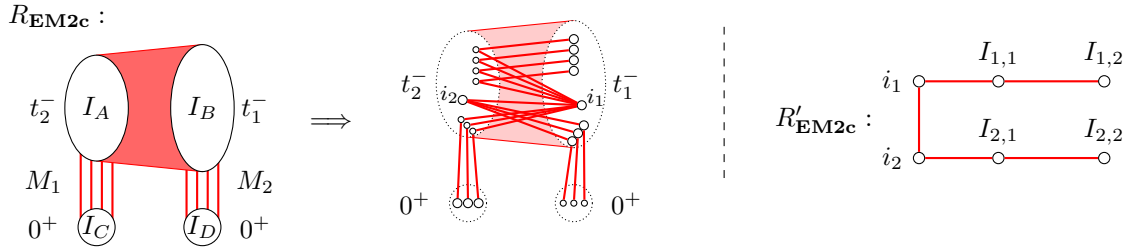


Figure 21: On the left, the initial reduced graph R_{EM2c} transformed into the substructure used to embed the tree in Lemma 4.12. On the right, the auxiliary graph R'_{EM2c} used when applying Lemma 4.2.

Lemma 4.12 (EM2c). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with a partition $V(G) = V_1 \cup \dots \cup V_k$ satisfying $|V_1| = \dots = |V_k| = m$. Let R_{EM2c} be a graph with vertex set $[k]$, such that if $ij \in E(R_{\text{EM2c}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I_B \cup I_C \cup I_D$ such that the following properties hold (see Figure 21).*

J1 $|\cup_{i \in I_A} V_i| \geq (1 - \alpha)t_2$, $|\cup_{i \in I_B} V_i| \geq (1 - 11\alpha)t_1$, $|\cup_{i \in I_C} V_i| = 100\alpha t_2$, and $|\cup_{i \in I_D} V_i| = 10\alpha t_2$.

J2 $R_{\text{EM2c}}[I_A, I_B]$ is an η -almost complete bipartite graph.

J3 $R_{\text{EM2c}}[I_A, I_C]$ contains a matching M_1 covering I_C .

J4 $R_{\text{EM2c}}[I_B, I_D]$ contains a matching M_2 covering I_D .

Then, G contains a copy of T .

Proof. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1. Using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$, and partition J as $J_X \cup J_Y$, so that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $\tau_1 = |\phi^{-1}(X_2 \cup X_3)|$ and $\tau_2 = |\phi^{-1}(Y_2 \cup Y_3)|$. For each $* \in \{A, B, C, D\}$, let $n_* = |\cup_{i \in I_*} V_i|$. Let R'_{EM2c} be the graph depicted on the right in Figure 21, which is a subgraph of the graph R' in Lemma 4.2. Define a random homomorphism $\varphi : T \rightarrow R'_{\text{EM2c}}$ as follows. First, for vertices in $\phi^{-1}(X_0 \cup Y_0)$, set $\varphi(v) = i_1$ if $\phi(v) = X_0$, and set $\varphi(v) = i_2$ if $\phi(v) = Y_0$.

For each $j \in J_X$ independently at random, with probability 4α , define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $i_2, I_{2,1}, I_{2,2}$, respectively; and with probability $1 - 4\alpha$ define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively.

For each $j \in J_Y$ independently at random, with probability 4α , define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $I_{2,1}, I_{2,2}, I_{2,1}$, respectively; and with probability $1 - 4\alpha$ define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively.

By Lemma 2.7, with positive probability, $|\varphi^{-1}(I_{2,2})| = 4\alpha\tau_2 \pm \alpha\tau_2/2 \leq n_D - \alpha n/10$, $|\varphi^{-1}(I_{2,1})| = 4\alpha\tau_1 \pm \alpha\tau_1/2 \leq 10\alpha t_2 - \alpha n/100$, $|\varphi^{-1}(I_{1,2} \cup I_{2,1})| \leq t_1 \leq n_B + n_C - 2\alpha n$, and $|\varphi^{-1}(I_{1,1})| = (1 - 4\alpha)\tau_2 \pm \alpha\tau_2 \leq n_A - \alpha n/10$. Therefore, like in Lemma 4.10, by **J2–J4** and after refining, we can find $i_1 \in I_B$ and $i_2 \in I_A$ such that $i_1 i_2 \in E(R_{\text{EM2c}})$, along with two matchings of refined clusters of suitable sizes forming $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs attached to i_1, i_2 as depicted in Figure 21, which allow us to apply Lemma 4.2 to find a copy of T in G . $\square$

4.10 EM2d Embedding method

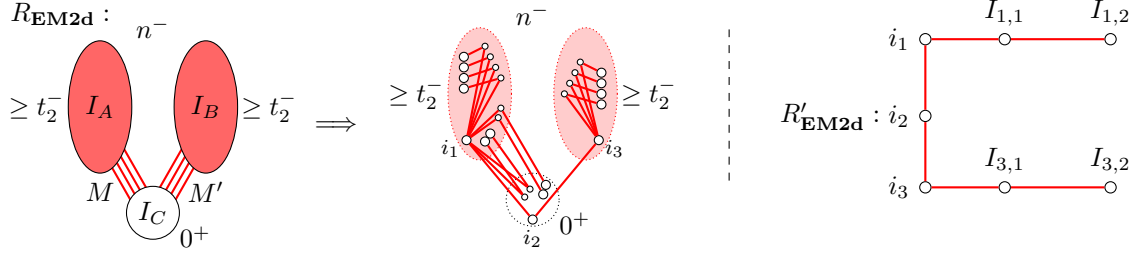


Figure 22: On the left, the initial reduced graph R_{EM2d} transformed into the substructure used to embed the tree in Lemma 4.13. On the right, the auxiliary graph R'_{EM2d} used when applying Lemma 4.2.

Lemma 4.13 (EM2d). Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a graph with a partition $V(G) = V_1 \cup \dots \cup V_k$ satisfying $|V_1| = \dots = |V_k| = m$. Let R_{EM2d} be a graph with vertex set $[k]$, such that if $ij \in E(R_{\text{EM2d}})$ then $G[V_i, V_j]$ is (ε, d) -regular. Suppose there is a partition $[k] = I_A \cup I_B \cup I_C$ such that the following properties hold (see Figure 22).

K1 $|\cup_{i \in I_A} V_i| \geq |\cup_{i \in I_B} V_i| \geq (1 - \alpha)t_2$, $|\cup_{i \in I_A \cup I_B} V_i| \geq (1 - \alpha)n$, and $|\cup_{i \in I_A \cup I_B \cup I_C} V_i| = (1 + 100\alpha)n$.

K2 $R_{\text{EM2d}}[I_A]$ and $R_{\text{EM2d}}[I_B]$ are both η -almost complete graphs.

K3 $R_{\text{EM2d}}[I_A, I_C]$ contains a matching M covering I_C , and a vertex in $I_A \cap V(M)$ that is adjacent to every vertex in $I_C \cap V(M)$.

K4 $R_{\text{EM2d}}[I_B, I_C]$ contains a matching M' covering I_C .

Then, G contains a copy of T .

Proof. Let ξ satisfy $c \ll \xi \ll 1/k$. Let S be the graph defined in Lemma 4.1. Using that lemma, take a homomorphism $\phi : T \rightarrow S$ such that each component of $T - \phi^{-1}(X_0 \cup Y_0)$ has size at most ξn and $|\phi^{-1}(X_0 \cup Y_0 \cup X_1 \cup Y_1)| \leq \xi n$. Without loss of generality, say $|\phi^{-1}(X_0 \cup X_1 \cup X_2 \cup X_3)| = t_1$ and $|\phi^{-1}(Y_0 \cup Y_1 \cup Y_2 \cup Y_3)| = t_2$. Let the components of $T - \phi^{-1}(X_0 \cup Y_0)$ be $\{K_j : j \in J\}$, and partition J as $J_X \cup J_Y$, so that $N_T(K_j) \subset \phi^{-1}(X_0)$ for each $j \in J_X$, and $N_T(K_j) \subset \phi^{-1}(Y_0)$ for each $j \in J_Y$.

Let $\tau_X = |\phi^{-1}(X_2 \cup Y_3)|$ and $\tau_Y = |\phi^{-1}(Y_2 \cup X_3)|$. For each $*$ in $\{A, B, C\}$, let $n_* = |\cup_{i \in I_*} V_i|$. Let R'_{EM2d} be the graph depicted on the right in Figure 22, which is a subgraph of the graph R' in Lemma 4.2. Since

$$\tau_X + \tau_Y \leq n \leq n_A + n_B + 2n_C - 20\alpha n \leq 2n_A + 2n_C - 20\alpha n,$$

we can separate into the following two cases. **I:** $\tau_X \leq n_A + n_C - 10\alpha n$. **II:** $\tau_Y \leq n_A + n_C - 10\alpha n$.

Case I. As $\tau_X + \tau_Y \geq n - \xi n \geq n_A + n_C - 10\alpha n$, there exists $p \in [0, 1]$ such that $\tau_X + p\tau_Y = n_A + n_C - 10\alpha n$. It follows that $(1-p)\tau_Y \leq n - \tau_X - p\tau_Y \leq n_B - 50\alpha n$. Define a random homomorphism $\varphi : T \rightarrow R'_{\text{EM2d}}$ as follows. First, for vertices in $\phi^{-1}(X_0 \cup Y_0)$, set $\varphi(v) = i_1$ if $\phi(v) = X_0$, and set $\varphi(v) = i_2$ if $\phi(v) = Y_0$. For each $j \in J_X$, define φ on K_j by composing ϕ with the map that sends Y_1, X_2, Y_3 to $I_{1,1}, I_{1,2}, I_{1,1}$, respectively. For each $j \in J_Y$ independently at random, with probability p , define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_1, I_{1,1}, I_{1,2}$, respectively; and with probability $1-p$ define φ on K_j by composing ϕ with the map that sends X_1, Y_2, X_3 to $i_3, I_{3,1}, I_{3,2}$, respectively.

By Lemma 2.7, with positive probability, $|\varphi^{-1}(I_{1,1} \cup I_{1,2})| = \tau_X + p\tau_Y \pm \alpha n \leq n_A + n_C - 5\alpha n$ and $|\varphi^{-1}(I_{3,1} \cup I_{3,2})| = (1-p)\tau_Y \pm \alpha n \leq n_B - 10\alpha n$.

We now further transform the structure as depicted on the left of Figure 22. Let $i_1 \in I_A \cap V(M)$ be the vertex given by **K3** that is adjacent to every vertex in $I_C \cap V(M)$, suppose it is matched with $i_2 \in I_C$ by M , and let $i_3 \in I_B$ be the neighbour of i_2 in M' . By **K2**, we can find a matching M_1 in $R_{\text{EM2d}}[I_A \cup I_C]$ containing the matching $M - \{i_1, i_2\}$, such that i_1 is adjacent to every $i \in V(M_1)$ and $\sum_{i \in V(M_1)} |V_i| = n_A + n_C - \alpha n$. Since $|\varphi^{-1}(I_{1,1} \cup I_{1,2})| + 4\alpha n \leq n_A + n_C - \alpha n$ from above, we can find $m_1, m_2 \geq \sqrt{\varepsilon}m$ such that $m_1 + m_2 = m$ and $|\varphi^{-1}(I_{1,b})| + \alpha n \leq (n_A + n_C - \alpha n)m_b/m$ for each $b \in [2]$. For each $i \in V(M_1)$, partition V_i as $V_{i,1} \cup V_{i,2}$, such that $|V_{i,1}| = m_1$ and $|V_{i,2}| = m_2$. Then, for every edge ii' in M_1 , both $G[V_{i,1}, V_{i',2}]$ and $G[V_{i,2}, V_{i',1}]$ are $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular by Lemma 2.19. Relabel $\{V_{i,1} : i \in V(M_1)\}$ and $\{V_{i,2} : i \in V(M_1)\}$ as $\{V'_i : i \in I_{1,1}\}$ and $\{V'_i : i \in I_{1,2}\}$, respectively. Note that $\sum_{i \in I_{1,1}} |V'_i| \geq |\varphi^{-1}(I_{1,1})| + \alpha n$ and $\sum_{i \in I_{1,2}} |V'_i| \geq |\varphi^{-1}(I_{1,2})| + \alpha n$.

Similarly, we can use **K2** to find a matching M_2 in $R_{\text{EM2d}}[I_B]$ with all vertices in $V(M_2)$ adjacent to i_3 , then refine them accordingly to obtain clusters $\{V'_i : i \in I_{3,1}\}$ and $\{V'_i : i \in I_{3,2}\}$ that can be matched together to form $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pairs, such that $\sum_{i \in I_{3,1}} |V'_i| \geq |\varphi^{-1}(I_{3,1})| + \alpha n$ and $\sum_{i \in I_{3,2}} |V'_i| \geq |\varphi^{-1}(I_{3,2})| + \alpha n$. Thus, we can apply Lemma 4.2 to find a copy of T in G .

Case II. In this case, we use essentially the same argument except that the role of X and Y are flipped, so we omit the full details. That is, we set $\varphi(v) = i_1$ if $\phi(v) = Y_0$, and $\varphi(v) = i_2$ if $\phi(v) = X_0$. Then, every component K_j with $j \in J_Y$ will be embedded into $I_A \cup I_C$, while every component K_j with $j \in J_X$ will be embedded into $I_A \cup I_C$ with some suitable probability p , and into I_B otherwise. $\square$

5 Proof of Theorem 2.2: Stability

In this section, we will prove Theorem 2.2 by following the 4-stage process outlined in Section 3 and depicted in Figure 2, using the embedding methods developed in Section 4. As will be justified when we put everything together in Section 5.7, we may assume that $t_1 \leq 2t_2$ throughout these 4 stages. Our starting point is the next result that easily follows from the proof of [21, Theorem 3] by Haxell, Łuczak, and Tingley applied with $\alpha = t_1/t_2$ and $n = (1 - \varepsilon)(t_1 + 2t_2)$. The reason it follows from their proof rather than directly from their theorem statement is that we need to ‘remember’, and later use, the remaining regularity clusters in the graph that is not part of the structure found by Haxell, Łuczak, and Tingley, and thus not included in the statement of their theorem.

Theorem 5.1. *Let $1/n \ll 1/m \ll 1/k \ll \varepsilon \ll 1$. Let t_1 and t_2 satisfy $t_1 + t_2 = n$ and $t_2 \leq t_1 \leq 2t_2$. Let G be any red/blue coloured complete graph on at least $(1 - \varepsilon)(t_1 + 2t_2)$ vertices. Then, there exist*

disjoint subsets $V_0, \dots, V_k \subset V(G)$ that form an ε -regular partition with corresponding red/blue coloured $(\varepsilon, 1/3)$ -reduced graph R , a partition $[k] = I_A \cup I_B \cup I_C$, and a colour $*$ $\in \{\text{red}, \text{blue}\}$ such that the following hold (see **A-situation** in Figure 2).

- $|V_i| = m$ for every $i \in \{0\} \cup I_A \cup I_C$, and $|V_i| = t_1 m / t_2$ for every $i \in I_B$.
- $|I_A| = |I_B| = |I_C| = k$, with $km \geq (1 - 2\varepsilon)t_2$.
- In R_* , 0 is adjacent to every $a \in I_A$.
- $R_{\text{red}}[I_A, I_B]$ contains a perfect matching.

To reduce repetition, we will carry out the 4 stages in reverse order in Sections 5.2–5.5. That is, we start with Lemma 5.4 in Section 5.2 carrying out **Stage 4**, which states that given a **D-situation** obtained in an earlier stage, we can either find a monochromatic copy of T or the reduced graph is extremal (**E-situation**). Then, we similarly proceed through the other stages in reverse order, ending with Lemma 5.9 in Section 5.5 carrying out **Stage 1**, which says that from the **A-situation** structure given by Theorem 5.1, we can either find a monochromatic copy of T , or, combining with known results about the later stages, conclude that the reduced graph is extremal (**E-situation**).

To finish the proof, we still need two further results. In Section 5.1, we formalise what it means for a reduced graph to be extremal, and show in Lemma 5.3 that this implies the original graph is extremal as well in the sense of Definition 2.1. The second is a technical ‘cascading lemma’ about maximum matchings, which is needed in **Stage 2** and proved separately in Section 5.6.

5.1 Extremal regular partitions

Recall from Definition 2.1 what it means for a red/blue coloured complete graph to be Type I (μ, t_1, t_2) -extremal or Type II (μ, t_1, t_2) -extremal, and that if it is extremal of either type then we say it is (μ, t_1, t_2) -extremal. Now, we similarly define a notion of being extremal for the reduced graph of a regular partition.

Definition 5.2. Let $1/n \ll 1/k \ll \varepsilon \ll 1$, let $d, \mu \in [0, 1]$, and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$. We say a red/blue-coloured graph G has a *Type I (μ, t_1, t_2) -extremal (ε, d) -regular partition* if there exists an ε -regular partition $V_1 \cup \dots \cup V_k$ in $V(G)$ with corresponding (ε, d) -reduced graph R , and a partition $[k] = I_A \cup I_B$ such that the following hold.

- $|V_i| = m$ for every $i \in [k]$.
- $|I_A| \geq (1 - \mu)n/m$ and $|I_B| \geq (1 - \mu)t_2/m$.
- Both $R_{\text{red}}[I_A]$ and $R_{\text{blue}}[I_A, I_B]$ are μ -almost empty, or both $R_{\text{blue}}[I_A]$ and $R_{\text{red}}[I_A, I_B]$ are μ -almost empty.

We say G has a *Type II (μ, t_1, t_2) -extremal (ε, d) -regular partition* if there exists an ε -regular partition $V_1 \cup \dots \cup V_k$ in $V(G)$ with corresponding (ε, d) -reduced graph R , and a partition $[k] = I_A \cup I_B$ such that the following hold.

- $|V_i| = m$ for every $i \in [k]$.
- $|I_A|, |I_B| \geq (1 - \mu)t_1/m$.
- All of $R_{\text{red}}[I_A]$, $R_{\text{red}}[I_B]$, and $R_{\text{blue}}[I_A, I_B]$ are μ -almost empty, or all of $R_{\text{blue}}[I_A]$, $R_{\text{blue}}[I_B]$, and $R_{\text{red}}[I_A, I_B]$ are μ -almost empty.

With suitable choices of constants, if a red/blue coloured complete graph G has a Type I or Type II (μ', t_1, t_2) -extremal (ε, d) -regular partition, then it must be (μ, t_1, t_2) -extremal, by the following lemma.

Lemma 5.3. Let $1/n \ll 1/k \ll \varepsilon \ll \mu' \ll d \ll \mu \ll 1$, and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$. If G is a red/blue coloured complete graph on at most $2n$ vertices that contains a Type I or Type II (μ', t_1, t_2) -extremal (ε, d) -regular partition, then G is Type I or Type II (μ, t_1, t_2) -extremal, respectively.

Proof. Suppose G contains a Type I (μ', t_1, t_2) -extremal (ε, d) -regular partition $V_1 \cup \dots \cup V_k$ satisfying the conditions in Definition 5.2, where without loss of generality we assume that both $R_{\text{red}}[I_A]$ and $R_{\text{blue}}[I_A, I_B]$ are μ' -almost empty. Let $V_A = \cup_{i \in I_A} V_i$ and $V_B = \cup_{i \in I_B} V_i$. From assumptions and using $km \leq |G| \leq 2n$, the number of red edges in V_A is at most $\mu'|I_A|^2 m^2 + d|I_A|^2 m^2 + |I_A| m^2 \leq \mu' k^2 m^2 + dk^2 m^2 + km^2 \leq 4(\mu' + d + 1/k)n^2 \leq 5dn^2$. Therefore, by removing at most $\sqrt{5dn}$ vertices from V_A , we can obtain a subset V'_A such that $d_{\text{red}}(u, V'_A) \leq \sqrt{5dn}$ for every $u \in V'_A$.

Similarly, the number of blue edges between V'_A and V_B is at most $\mu'|I_A||I_B|m^2 + d|I_A||I_B|m^2 \leq 4(\mu' + d)n^2 \leq 5dn^2$. Thus, we can remove at most $\sqrt{5dn}$ vertices from each of V'_A and V_B to obtain subsets U_1 and U_2 , respectively, such that for every $i \in [2]$ and every $u \in U_i$, $d_{\text{blue}}(u, U_{3-i}) \leq \sqrt{5dn}$. Since $\mu \gg \sqrt{d}$, U_1 and U_2 show that G is Type I (μ, t_1, t_2) -extremal.

The case when G contains a Type II (μ', t_1, t_2) -extremal (ε, d) -regular partition is similar and thus omitted. $\square$

5.2 Stage 4

Lemma 5.4 (Stage 4). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \ll \mu \leq 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a red/blue coloured graph that contains a coloured ε -regular partition $V_1 \cup \dots \cup V_k$ with $|V_1| = \dots = |V_k| = m$ and corresponding red/blue coloured (ε, d) -reduced graph R . Suppose there is a partition $[k] = I_A \cup I_B$ such that the following hold (see **D-situation** in Figure 2).*

L1 $|I_A| = k_1$ and $|I_B| = k_2$, with $k_1 m, k_2 m \geq (1 - \alpha)t_2$ and $(k_1 + k_2)m \geq (1 - \alpha)(n + t_2)$.

L2 $R_{\text{red}}[I_A, I_B]$ is η -almost complete.

Then, G contains a monochromatic copy of T , or G is Type I (μ, t_1, t_2) -extremal, or $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Proof. Without loss of generality, assume that $k_1 \geq k_2$. Let $A = \cup_{i \in I_A} V_i$ and $B = \cup_{i \in I_B} V_i$. Note that from **L1**, $|A| + |B| \geq (1 + 200\alpha)n$ and $|A| = k_1 m \geq (1 - \alpha)(n + t_2)/2 \geq t_2 + 200\alpha n$.

Case I. $k_2 m \geq t_2 + 100\alpha n$. If there exists an edge ij in $R_{\text{red}}[I_A]$, then we can use **L2** to greedily find a perfect matching between some $I'_A \subset I_A \setminus \{i, j\}$ of size $100\alpha n/m$ and some $I'_B \subset I_B$. This allows us to apply Lemma 4.10 (**EM2a**) to $I_A, I_B \setminus I'_B$, and I'_B to embed T in red. Similarly, we can use Lemma 4.11 (**EM2b**) to embed T in red if there is an edge in $R_{\text{red}}[I_B]$. Thus, we may assume that both $R_{\text{red}}[I_A]$ and $R_{\text{red}}[I_B]$ are empty. This implies that both $R_{\text{blue}}[I_A]$ and $R_{\text{blue}}[I_B]$ contain at most εk^2 non-edges, so we can find $J_A \subset I_A$ and $J_B \subset I_B$, such that $|J_A| \geq (1 - 10\sqrt{\varepsilon})k_1$, $|J_B| \geq (1 - 10\sqrt{\varepsilon})k_2$, and both $R_{\text{blue}}[J_A], R_{\text{blue}}[J_B]$ are $10\sqrt{\varepsilon}$ -almost complete.

If there is any edge ib in $R_{\text{blue}}[J_A, J_B]$ with $i \in J_A$, then by moving i out of J_A and finding an arbitrary neighbour a of i in J_A , we get the structure required to apply Lemma 4.13 (**EM2d**) to find a blue copy of T in G . Therefore, we can assume that $R_{\text{blue}}[J_A, J_B]$ is empty.

If $k_1 m \geq t_1 + 10\alpha n$, then after using Lemma 2.24 to refine the clusters indexed by J_A and J_B , we can embed T in red using Lemma 4.5 (**HLT**). If instead $k_1 m < t_1 + 10\alpha n$, then it follows from $k_1 m \geq (1 - \alpha)(n + t_2)/2$ that $t_1 \geq (2 - 100\alpha)t_2$. Also, we have $k_2 m \geq (1 - \alpha)(n + t_2) - k_1 m \geq (2 - 50\alpha)t_2 \geq (1 - 100\alpha)t_1$. Therefore, G contains a Type II $(200\alpha, t_1, t_2)$ -extremal (ε, d) -regular partition, and so G is Type II (μ, t_1, t_2) -extremal by Lemma 5.3.

Case II. $k_2 m < t_2 + 100\alpha n$. Then $k_1 m \geq (1 - \alpha)(n + t_2) - k_2 m \geq (1 - 102\alpha)n > t_2 + 500\alpha n$.

If the maximum matching M in $R_{\text{red}}[I_A]$ has size at least $101\alpha n/m$, then we can move one side of a submatching of M with size $100\alpha n/m$ out of I_A to obtain the structure needed to apply Lemma 4.10 (**EM2a**) to find a red copy of T . Otherwise, let $I'_A = I_A \setminus M$, $A' = \cup_{i \in I'_A} V_i$, and $k'_1 = |I'_A| \geq k_1 - 202\alpha n/m$. Then, $R_{\text{red}}[I'_A]$ is empty and so $R_{\text{blue}}[I'_A]$ contains at most εk^2 non-edges.

Note that $|A'| = k'_1 m \geq (1 - 400\alpha)n$. If $|A'| = k'_1 m \geq (1 + 5\alpha)n$, then we can easily find the structure to apply Lemma 4.12 (**EM2c**) to embed T in $G_{\text{blue}}[A']$. Thus, we may assume that $|A'| \leq (1 + 5\alpha)n$, and so $|B| \geq (1 - \alpha)(n + t_2) - |A'| - 202\alpha n \geq (1 - 1000\alpha)t_2$. Let $\alpha \ll \beta \ll d$. If at least $2\beta n/m$ vertices in I'_A have at least $2\beta n/m$ blue neighbours in I_B , then we can greedily find the structure to apply Lemma 4.12 (**EM2c**) to embed T in $G_{\text{blue}}[A' \cup B]$. Otherwise, we can find $J_A \subset I'_A$ and $J_B \subset I_B$ such

that $|J_A| \geq (1 - 3\beta)n/m$, $|J_B| \geq (1 - 10\sqrt{\beta})t_2/m$, and $R_{\text{blue}}[J_A, J_B]$ is $10\sqrt{\beta}$ -almost empty. This shows that G contains a Type I $(10\sqrt{\beta}, t_1, t_2)$ -extremal (ε, d) -regular partition, so G is Type I (μ, t_1, t_2) -extremal by Lemma 5.3. $\square$

5.3 Stage 3

Lemma 5.5 (Stage 3). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a red/blue coloured graph that contains a coloured ε -regular partition $V_1 \cup \dots \cup V_k$ with $|V_1| = \dots = |V_k| = m$ and corresponding blue/red coloured (ε, d) -reduced graph R . Suppose there is a partition $[k] = I_A \cup I_B \cup I_C$ such that the following hold (see **C-situation** in Figure 2).*

M1 $|I_A| = k_1$, $|I_B| = k_2$, and $|I_C| = k_3$, with $k_1m, k_2m, k_3m \geq (1 - \alpha)t_2$ and $(k_1 + k_2)m = (1 - \alpha)n$.

M2 $R_{\text{red}}[I_A, I_B]$ is η -almost complete.

Then, G contains a monochromatic copy of T , or G is Type I (μ, t_1, t_2) -extremal, or $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Proof. Let $\alpha \leq \alpha' \ll \beta \ll d$. To avoid repetition, we first prove the following two claims dealing with two commonly occurring structures.

Claim 5.6. *Suppose there exist disjoint $J_A, J_B, J_C \subset [k]$, such that the following hold (see Figure 6).*

N1 $|J_A| = (1 - \alpha')t_1/m$ and $|J_B| = |J_C| = (1 - \alpha')t_2/m$.

N2 $R_{\text{red}}[J_A, J_B]$ is η -almost complete.

N3 $R_{\text{red}}[J_B, J_C]$ contains a matching with size $100\alpha't_2/m$.

Then, G contains a monochromatic copy of T , or G is Type I (μ, t_1, t_2) -extremal, or $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Proof of Claim 5.6. If at least $100\alpha't_2/m$ vertices in J_A have at least $200\alpha't_2/m$ red neighbours in J_C , then we can greedily find a matching of size $100\alpha't_2/m$ in $R_{\text{red}}[J_A, J_C]$ disjoint from the matching given by **N3**. This allows us to apply Lemma 4.12 (**EM2c**) to find a red copy of T in G . Otherwise, at most $100\alpha't_2/m$ vertices in J_A have at least $200\alpha't_2/m$ red neighbours in J_C , so we can find $J'_A \subset J_A$ and $J'_C \subset J_C$ with $|J'_A| \geq (1 - 200\alpha')t_1/m$ and $|J'_C| \geq (1 - 20\sqrt{\alpha'})t_2/m$, such that $R_{\text{blue}}[J'_A, J'_C]$ is $20\sqrt{\alpha'}$ -almost complete.

If at most $200\beta t_2/m$ vertices in J_B have at least $200\beta t_2/m$ blue neighbours in J'_C , then we can find $J'_B \subset J_B$ and $J''_C \subset J'_C$, such that $|J'_B| \geq (1 - 300\beta)t_2/m$, $|J''_C| \geq (1 - 20\sqrt{\beta})t_2/m$, and $R_{\text{red}}[J_A \cup J'_C, J'_B]$ is $20\sqrt{\beta}$ -almost complete. This gives the required structure (**D-situation**) in red to apply Lemma 5.4 (**Stage 4**) with $\eta = 20\sqrt{\beta}$ to finish the proof.

Thus, we may assume that at least $200\beta t_2/m$ vertices in J_B have at least $200\beta t_2/m$ blue neighbours in J'_C , so we can find a blue matching of size $100\beta t_2/m$ in $R_{\text{blue}}[J_B, J'_C]$ disjoint from the red matching given by **N3**. Arbitrarily pick a matching of size $10(\alpha' + \beta)t_2/m$ in $R[J_A]$. By pigeonhole, it either contains a red matching M_{red} of size $10\alpha't_2/m$ or a blue matching M_{blue} of size $10\beta t_2/m$. In the former case, moving vertices on one side of M_{red} out of J_A , and using **N2**, **N3**, we have the structure to apply Lemma 4.12 (**EM2c**) to find a red copy of T in G . In the latter case, moving vertices on one side of M_{blue} out of J_A , and using that $R_{\text{blue}}[J_A, J'_C]$ is $20\sqrt{\alpha'}$ -almost complete, we can again apply Lemma 4.12 (**EM2c**) to find a blue copy of T in G . $\square$

Claim 5.7. *Suppose there exist disjoint $J_A, J_B, J_C \subset [k]$, such that the following hold (see Figure 7).*

O1 $|J_A| \geq (t_2 + 200\alpha'n)/m$, $|J_B| \geq (t_2 - \alpha'n)/m$, $|J_A| + |J_B| = (1 - \alpha')n/m$, and $|J_C| = (1 - \alpha')t_2/m$.

O2 $R_{\text{red}}[J_A, J_B]$ is η -almost complete.

O3 $R_{\text{red}}[J_A, J_C]$ contains a matching with size $100\alpha'n/m$.

Then, G contains a monochromatic copy of T , or G is Type I (μ, t_1, t_2) -extremal, or $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Proof of Claim 5.7. If there is an edge in either $R_{\text{red}}[J_A]$ or $R_{\text{red}}[J_B]$, then we can apply Lemma 4.10 (**EM2a**) or Lemma 4.11 (**EM2b**), respectively, to find a red copy of T in G . Thus, by removing at most $10\sqrt{\varepsilon}$ -fraction of vertices from J_A and J_B , we may assume that $R_{\text{blue}}[J_A]$ and $R_{\text{blue}}[J_B]$ are both $10\sqrt{\varepsilon}$ -almost complete.

Suppose there exists a set of $10\beta t_2/m$ vertices in J_C , each of which has at least $10\beta t_2/m$ blue neighbours in both J_A and J_B . Then, using a similar argument to Claim 4.7, we can find a blue matching M_1 of size $\beta^2 t_2/m$ between J_A and J_C with a vertex in $V(M_1) \cap J_A$ adjacent to every vertex in $J'_C := V(M_1) \cap J_C$. Greedily, we can also find another blue matching between J'_C and J_B covering J'_C . This allows us to apply Lemma 4.13 (**EM2d**) to find a blue copy of T in G .

Thus, we may now assume that there exist two disjoint subsets $J_A^+, J_B^+ \subset J_C$ containing all but at most $10\beta t_2/m$ vertices in J_C , such that every vertex in J_A^+ has at most $10\beta t_2/m$ blue neighbours in J_B , and every vertex in J_B^+ has at most $10\beta t_2/m$ blue neighbours in J_A . Using **O2**, and by removing at most $5\sqrt{\beta} t_2/m$ vertices from each of J_A and J_B to obtain J_A^- and J_B^- , respectively, we can ensure that both $R_{\text{red}}[J_A^-, J_B^- \cup J_B^+]$ and $R_{\text{red}}[J_B^-, J_A^- \cup J_A^+]$ are $5\sqrt{\beta}$ -almost complete. Note that if $|J_A^+| \leq 10\beta t_2/m$, then $|J_A^- \cup J_B^- \cup J_B^+| \geq (1 - 20\beta)(n + t_2)/m$, so J_A^- and $J_B^- \cup J_B^+$ form the required red structure (**D-situation**) to apply Lemma 5.4 (**Stage 4**) with $\eta = 5\sqrt{\beta}$ to finish the proof. Thus, we can assume that $|J_A^+| \geq 10\beta t_2/m$, and similarly $|J_B^+| \geq 10\beta t_2/m$.

As in the beginning of this proof, we may further assume that $R_{\text{blue}}[J_A^- \cup J_A^+]$ and $R_{\text{blue}}[J_B^- \cup J_B^+]$ are both $10\sqrt{\varepsilon}$ -almost complete, as the presence of any red edge within either of them allows us to apply Lemma 4.10 (**EM2a**) or Lemma 4.11 (**EM2b**) to find a red copy of T in G . If there is any blue edge between $J_A^- \cup J_A^+$ and $J_B^- \cup J_B^+$, then we can move one end of this blue edge out and then find a blue copy of T using Lemma 4.13 (**EM2d**). Therefore, $R_{\text{red}}[J_A^- \cup J_A^+, J_B^- \cup J_B^+]$ contains the red structure (**D-situation**) needed to apply Lemma 5.4 (**Stage 4**) to finish the proof. $\square$

Now we can carry out **Stage 3**. If at most $2\beta t_2/m$ vertices in $I_A \cup I_B$ have at least $2\beta t_2/m$ red neighbours in I_C , then we can find $I_{AB} \subset I_A \cup I_B$ and $I'_C \subset I_C$ with $|I_{AB}| \geq (1 - 2\beta)n/m$ and $|I'_C| \geq (1 - 5\sqrt{\beta})t_2/m$, such that $G_{\text{blue}}[I_{AB}, I'_C]$ is $5\sqrt{\beta}$ -almost complete. Therefore, I_{AB} and I'_C form the blue structure (**D-situation**) required to apply Lemma 5.4 (**Stage 4**) to finish the proof. Thus, we may assume that at least $2\beta t_2/m$ vertices in $I_A \cup I_B$ have at least $2\beta t_2/m$ red neighbours in I_C , from which it follows that there is a red matching of size $\beta t_2/m$ either between I_A and I_C , or between I_B and I_C .

Case I. $|I_A|, |I_B| \geq (t_2 + 200\alpha n)/m$. Without loss of generality, assume that there is a red matching of size $\beta t_2/m$ between I_A and I_C . Then, we can apply Claim 5.7 to finish the proof.

Case II. Without loss of generality, assume that $|I_B| < (t_2 + 200\alpha n)/m$, so $|I_A| \geq (t_1 - 201\alpha n)/m \geq (1 - 500\alpha)t_1/m$ from **M1**. If there is a red matching of size $\beta t_2/m$ between I_B and I_C , then we can apply Claim 5.6 with $\alpha' = 500\alpha$ to finish the proof.

Thus, we may assume that there is a red matching of size $\beta t_2/m$ between I_A and I_C . If $t_1 \leq (1 + 1000\alpha)t_2$, then $|I_B| \geq (1 - \alpha)t_2/m \geq (1 - 2000\alpha)t_1/m$, so we can apply Claim 5.6 with $\alpha' = 2000\alpha$ to finish the proof. If instead $t_1 > (1 + 1000\alpha)t_2$, then $|I_A| \geq (1 - 3\alpha)t_1/m \geq (t_2 + 200\alpha n)/m$, so we can apply Claim 5.7 to finish the proof. $\square$

5.4 Stage 2

Lemma 5.8 (Stage 2). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \eta \ll \alpha \ll d \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Let G be a red/blue coloured graph that contains a coloured ε -regular partition $V_0 \cup V_1 \cup \dots \cup V_k$ with $|V_0| = \dots = |V_k| = m$ and corresponding red/blue coloured (ε, d) -reduced graph R . Suppose there is a partition $[k] = I_A \cup I_B \cup I_C$ such that the following hold (see **B-situation** in Figure 2).*

P1 $|I_A| = |I_B| = k_1$ with $k_1 m = (1 - \alpha)t_2$, and $|I_C| = k_2$ with $k_2 m = (1 - \alpha)t_1$.

P2 In R_{red} , 0 is adjacent to every $i \in I_A \cup I_B$.

P3 $R_{\text{red}}[I_A, I_B]$ is η -almost complete.

Then, G contains a monochromatic copy of T , or G is Type I (μ, t_1, t_2) -extremal, or $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Proof. Let M be a maximum red matching in R between $I_A \cup I_B$ and I_C . For any $I \subset [k]$, for brevity we use $V(I)$ to denote $\cup_{i \in I} V_i$, the vertices covered by I . Let $\alpha \ll \beta \ll d$.

Case I. M contains at most $(t_1 - t_2 + 2\beta n)/m$ edges. Let $X = (I_A \cup I_B) \cap V(M)$, $X' = (I_A \cup I_B) \setminus V(M)$, $Y = I_C \cap V(M)$, and $Y' = I_C \setminus V(M)$. By Lemma 5.10, there are partitions $X = X^+ \cup X^- \cup \bar{X}$ and $Y = Y^+ \cup Y^- \cup \bar{Y}$ such that M matches X^+ with Y^- , X^- with Y^+ , and $\bar{X}$ with $\bar{Y}$, and $R_{\text{red}}[X' \cup X^-, Y' \cup Y^- \cup \bar{Y}]$ is empty. From assumption, $|X| = |Y| \leq (t_1 - t_2 + 2\beta n)/m$, and the following hold.

$$|V(X' \cup X^-)| \geq |V(X')| \geq (1 - \alpha)2t_2 - (t_1 - t_2 + 2\beta n) = (3 - 2\alpha - 2\beta)t_2 - (1 + 2\beta)t_1 \geq (1 - 10\beta)t_2.$$

$$|V(Y' \cup Y^- \cup \bar{Y})| = |V(I_C)| - |V(Y^+)| \geq |V(I_C)| - |V(Y)| \geq (1 - \alpha)t_1 - (t_1 - t_2 + 2\beta n) \geq (1 - 10\beta)t_2.$$

$$|V(X' \cup X^- \cup Y' \cup Y^- \cup \bar{Y})| = |V([k] \setminus X)| \geq (1 - \alpha)(t_1 + 2t_2) - (t_1 - t_2 + 2\beta n) \geq (1 - 10\beta)n.$$

Hence, $X' \cup X^-$ and $Y' \cup Y^- \cup \bar{Y}$ form the blue structure (**C-situation**) required to apply Lemma 5.5 (**Stage 3**) to finish the proof.

Case II. M contains at least $(t_1 - t_2 + 2\beta n)/m$ edges. Let $\varepsilon \ll \gamma \ll \alpha$. Refine every cluster V_i with $i \in I_A$ or $i \in I_B$ down to clusters of size γm , and label these new clusters as $\{V'_i : i \in J_A\}$ and $\{V'_i : i \in J_B\}$, respectively. Refine every cluster V_i with $i \in I_C$ down to clusters of size $\gamma t_1 m / t_2$ and label the new clusters as $\{V'_i : i \in J_C\}$. In this refinement process, at most $O(\gamma n)$ covered vertices are lost. Let R' be the new reduced graph on $J_A \cup J_B \cup J_C$, where for each $* \in \{\text{red}, \text{blue}\}$, $ij \in E(R'_*)$ if and only if $G_*[V'_i, V'_j]$ is $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular. Using **P3**, Lemma 2.19, and the matching M , we see that $R'_{\text{red}}[J_A, J_B]$ is also η -almost complete, and we can find a matching M' in $R'_{\text{red}}[J_A \cup J_B, J_C]$ covering $(t_1 - t_2 + \beta n)t_2/t_1$ vertices in $\cup_{i \in J_A \cup J_B} V'_i$ and $t_1 - t_2 + \beta n$ vertices in $\cup_{i \in J_C} V'_i$.

Case II.1. $|V(M') \cap J_A|, |V(M') \cap J_B| \geq 5\beta n / \gamma m$. Then, we can find disjoint subsets $J_{A,1}, J_{A,2} \subset J_A \setminus V(M')$ and $J_{B,1}, J_{B,2} \subset J_B \setminus V(M')$, such that $|J_{A,1}|/|J_{B,2}| = |J_{B,1}|/|J_{A,2}| \in [t_1/t_2, (1 + \alpha)t_1/t_2]$, each of $J_{A,1}, J_{A,2}, J_{B,1}, J_{B,2}$ covers at least αn vertices, and together with $V(M') \cap (J_A \cup J_B)$ they cover at least $(2 - 10\alpha)t_2$ vertices. Moreover, since $R'_{\text{red}}[J_A, J_B]$ is η -almost complete, by choosing the subsets above randomly and using Lemma 2.5, we can ensure that both $R'_{\text{red}}[J_{A,1}, J_{B,2}]$ and $R'_{\text{red}}[J_{A,2}, J_{B,1}]$ are 10η -almost complete. Therefore, for some $\varepsilon \ll \gamma' \ll \gamma$, we can use Lemma 2.24 to further appropriately refine these clusters along with those in M' into two sets of smaller clusters of sizes $\gamma'm$ and $\gamma't_1 m / t_2$, respectively, which can be paired up into $(\sqrt[4]{\varepsilon}, d - 2\sqrt{\varepsilon})$ -regular pairs. Together, these new clusters cover at least

$$(1 - \alpha)(2 - 10\alpha)t_2 + (1 - \alpha)(t_1 - t_2 + \beta n) \geq (1 + \beta/2)n$$

vertices. Finally, note that by **P2**, each new cluster of size $\gamma'm$ forms an $(\sqrt{\varepsilon}, d - \varepsilon)$ -regular pair with V_0 , so we have the structure to apply Lemma 4.5 (**HLT**) to find a red copy of T .

Case II.2. One of $|V(M') \cap J_A|$ and $|V(M') \cap J_B|$ is at most $5\beta n / \gamma m$. Without loss of generality, assume it is $|V(M') \cap J_B|$, and assume that M' is chosen so that $|V(M') \cap J_B|$ is maximised. Let J_D be the vertices in J_C matched by M' to the vertices in J_B . Since M' maximises $|V(M') \cap J_B|$, there is no red edge in R' between $J_B \setminus V(M')$ and $J_C \setminus J_D$, as we can swap any such edge with an edge in M' adjacent to J_A to obtain a matching M'' with $|V(M'') \cap J_B| > |V(M') \cap J_B|$, a contradiction. Thus, $R'_{\text{blue}}[J_B \setminus V(M'), J_C \setminus J_D]$ contains at most $\varepsilon(k+1)^2$ non-edges. Moreover, from $|V(M') \cap J_B| \leq 5\beta n / \gamma m$, it follows that $J_B \setminus V(M')$ and $J_C \setminus J_D$ cover at least $(1 - 20\beta)t_2$ and $(1 - 20\beta)t_1$ vertices, respectively. Thus, after removing clusters with low degrees in $R'_{\text{blue}}[J_B \setminus V(M'), J_C \setminus J_D]$ and refining all remaining clusters down to a smaller common size, we get the required structure (**C-situation**) in blue to apply Lemma 5.5 (**Stage 3**) to finish the proof. $\square$

5.5 Stage 1

Lemma 5.9 (Stage 1). *Let $1/n \ll 1/m \ll c \ll 1/k \ll \varepsilon \ll \alpha \ll d \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 with $t_2 \leq t_1 \leq 2t_2$. Suppose G is a red/blue coloured graph that contains an ε -regular partition $V_0 \cup V_1 \cup \dots \cup V_k$ with corresponding red/blue coloured (ε, d) -reduced graph R , and a partition $[k] = I_A \cup I_B \cup I_C$ such that the following hold (see **A-situation** in Figure 2).*

Q1 $|V_i| = m$ for every $i \in \{0\} \cup I_A \cup I_C$, and $|V_i| = t_1 m / t_2$ for every $i \in I_B$.

Q2 $|I_A| = |I_B| = |I_C| = k$, with $km = (1 - \alpha)t_2$.

Q3 In R_{red} , 0 is adjacent to every $a \in I_A$.

Q4 $R_{\text{red}}[I_A, I_B]$ contains a perfect matching M .

Then, G contains a monochromatic copy of T , or G is Type I (μ, t_1, t_2) -extremal, or $t_1 \geq (2 - \mu)t_2$ and G is Type II (μ, t_1, t_2) -extremal.

Proof. Let $\alpha \ll \beta \ll d$. Let X be the set of vertices in $I_A \cup I_B$ that have at least $20\beta|I_C|$ red neighbours in I_C . Let $I_{B,1} = I_B \cap X$, and let $I_{A,1}$ be the vertices matched with $I_{B,1}$ by M . Let $I_{A,3} = (I_A \cap X) \setminus I_{A,1}$, and let $I_{B,3}$ be the vertices matched with $I_{A,3}$ by M . Let $I_{A,2} = I_A \setminus (I_{A,1} \cup I_{A,3})$, $I_{B,2} = I_B \setminus (I_{B,1} \cup I_{B,3})$, and note that M gives a perfect matching between $I_{A,2}$ and $I_{B,2}$.

If $|I_{B,1}| \leq 2\beta k$, then we can remove at most $5\sqrt{\beta}|I_C|$ vertices from I_C to obtain I'_C , such that $R_{\text{blue}}[I_B \setminus I_{B,1}, I'_C]$ is $5\sqrt{\beta}$ -almost complete, both $I_B \setminus I_{B,1}$ and I'_C cover at least $(1 - 6\sqrt{\beta})t_2$ vertices, and they together cover at least $(1 - 6\sqrt{\beta})n$ vertices. Thus, after refining all clusters down to a common smaller size, $I_B \setminus I_{B,1}$ and I'_C provide the blue structure (**C-situation**) required to apply Lemma 5.5 (**Stage 3**) with $\eta = 5\sqrt{\beta}$ to finish the proof.

Now suppose $|I_{B,1}| \geq 2\beta k$. In this case, assuming there is no red copy of T in G , we make the following four deductions.

i) $|I_{A,1} \cap X| \leq \beta|I_{A,1}|$.

If not, let $I'_{A,1} \subset I_{A,1} \cap X$ have size $\beta^2 k$, and let $I'_{B,1} \subset I_{B,1}$ be the vertices in R_{red} matched with $I'_{A,1}$ by M . From the definition of X , we can find a perfect matching between $I'_{A,1}$ and some $I_{C,1} \subset I_C$, such that if $I'_C = I_C \setminus I_{C,1}$, then every vertex in $I'_{B,1}$ still has at least $10\beta^2|I'_C|$ neighbours in I'_C . After appropriate refinements, this structure allows us to apply Lemma 4.6 (**EM1a**) to find T in red.

ii) At most $\beta^2 k$ vertices in $I_{B,3}$ have at least $\beta^2 k$ red neighbours in $I_{A,1}$.

If not, we can find a perfect matching of size $\beta^2 k$ between some $I'_{B,3} \subset I_{B,3}$ and $I'_{A,1} \subset I_{A,1}$. Let $I'_{B,1} \subset I_{B,1}$ be the vertices matched with $I'_{A,1}$ by M , and let $I'_{A,3} \subset I_{A,3}$ be the vertices matched with $I'_{B,3}$ by M . Since $I'_{A,3} \subset X$, we can find a perfect matching between $I'_{A,3}$ and some $I_{C,3} \subset I_C$, such that if $I'_C = I_C \setminus I_{C,3}$, then every vertex in $I'_{B,1}$ still has at least $10\beta^2|I'_C|$ neighbours in I'_C . After appropriate refinements, this structure allows us to apply Lemma 4.6 (**EM1a**) to find T in red.

iii) At most $\beta|I_{A,1}|$ vertices in $I_{A,1}$ have at least $\beta|I_{A,1}|$ red neighbours in $I_{A,1}$.

If not, we can apply Lemma 4.8 (**EM1b**) to find T in red.

iv) At most βk edges in $M[I_{A,2}, I_{B,2}]$ satisfy that both of their endpoints have at least $\beta|I_{A,1}|$ red neighbours in $I_{A,1}$.

If not, we can apply Lemma 4.9 (**EM1c**) to find T in red.

From **iv)**, we can find a subset $J_{AB,2} \subset I_{A,2} \cup I_{B,2}$ containing a vertex in all but at most βk edges in $M[I_{A,2}, I_{B,2}]$, so that every vertex in $J_{AB,2}$ has at most $\beta|I_{A,1}|$ red neighbours in $I_{A,1}$. Similarly, using **i)**, **ii)**, and **iii)**, we can find $J_{B,3} \subset I_{B,3}$ containing all but at most $\beta^2 k$ vertices in $I_{B,3}$, and $J_{A,1} \subset I_{A,1} \setminus X$ containing all but at most $2\beta|I_{A,1}|$ vertices in $I_{A,1}$, such that every vertex in $J := J_{AB,2} \cup J_{B,3} \cup J_{A,1}$ has at most $2\beta|J_{A,1}|$ red neighbours in $J_{A,1}$, and at most $20\beta|I_C|$ red neighbours in I_C . In particular, by averaging and using that $J_{A,1}$ is non-empty, we can find some $j \in J_{A,1}$ with at most $2\beta|J|$ red neighbours in J and at most $20\beta|I_C|$ red neighbours in I_C . Also, we have

$$|J| \geq (1 - 2\beta)|I_{A,1}| + |I_{A,2}| - \beta k + |I_{A,3}| - \beta^2 k \geq (1 - 4\beta)|I_A| \geq (1 - 5\beta)t_2/m.$$

Therefore, we can find $J' \subset N_{R_{\text{blue}}}(j, J)$ and $I'_C \subset N_{R_{\text{blue}}}(j, I_C)$ with $|J'| \geq (1 - 2\beta)|J| \geq (1 - 10\beta)t_2/m$ and $|I'_C| \geq (1 - 20\beta - \sqrt{20\beta})|I_C| \geq (1 - 5\sqrt{\beta})t_2/m$, such that $R_{\text{blue}}[J', I'_C]$ is $5\sqrt{\beta}$ -almost complete. Therefore, after refining all clusters down to a smaller common size, J' and I'_C provide the blue structure (**B-situation**) required to apply Lemma 5.8 (**Stage 2**) with $\eta = 5\sqrt{\beta}$ to finish the proof. $\square$

5.6 Cascading lemma

Lemma 5.10 (Cascading lemma). *Let G be a bipartite graph with bipartition classes A and B , and let M be a maximum matching in G . Let $A_M = A \cap V(M)$, $B_M = B \cap V(M)$, $A' = A \setminus A_M$, and $B' = B \setminus B_M$. Then, A_M and B_M can be partitioned as $A_M = A^+ \cup A^- \cup \bar{A}$ and $B_M = B^+ \cup B^- \cup \bar{B}$ such that the following hold.*

- M matches vertices in A^+ with vertices in B^- , vertices in A^- with vertices in B^+ , and vertices in $\bar{A}$ with vertices in $\bar{B}$.
- $G[A' \cup A^-, B' \cup B^- \cup \bar{B}]$ and $G[A' \cup A^- \cup \bar{A}, B' \cup B^-]$ are both empty graphs.

Proof. First, by maximality of M , $G[A', B']$ must be an empty graph. Consider the following process. To initialise, set $A_0^+ = A_0^- = B_0^+ = B_0^- = \emptyset$, $\bar{A}_0 = A_M$, and $\bar{B}_0 = B_M$. Throughout this process, we will maintain the following conditions.

- R1** $A_i^+ \cup A_i^- \cup \bar{A}_i$ is a partition of A_M and $B_i^+ \cup B_i^- \cup \bar{B}_i$ is a partition of B_M .
- R2** M matches vertices in A_i^+ with vertices in B_i^- , vertices in A_i^- with vertices in B_i^+ , and vertices in $\bar{A}_i$ with vertices in $\bar{B}_i$.
- R3** $G[A' \cup A_i^-, B' \cup B_i^-]$ is the empty graph.
- R4** For every $a \in A' \cup A_i^-$, there exists an M -alternating path, possibly of length 0, that connects a to some vertex $a' \in A'$, starts with an edge in M , and has all internal vertices in $A_i^- \cup B_i^+$. For every $b \in B' \cup B_i^-$, there exists an M -alternating path, possibly of length 0, that connects b to some vertex $b' \in B'$, starts with an edge in M , and has all internal vertices in $B_i^- \cup A_i^+$.

Note that **R1–R4** are all satisfied when $i = 0$. Suppose for some $i \geq 0$ we have found $A_i^+, A_i^-, \bar{A}_i$ and $B_i^+, B_i^-, \bar{B}_i$ satisfying **R1–R4**. If $G[\bar{A}_i, B' \cup B_i^-]$ and $G[A' \cup A_i^-, \bar{B}_i]$ are both empty graphs, then we are done by letting $A^+ = A_i^+, A^- = A_i^-, \bar{A} = \bar{A}_i, B^+ = B_i^+, B^- = B_i^-$, and $\bar{B} = \bar{B}_i$. Otherwise, we show that we can continue this process.

Indeed, suppose there exists $a \in \bar{A}_i$ adjacent to some $b_1 \in B' \cup B_i^-$. Let $b \in \bar{B}_i$ be the vertex matched to a by M . We claim that b is not adjacent to any vertex in $A' \cup A_i^-$. Indeed, suppose b is adjacent to $a_1 \in A' \cup A_i^-$. Then by **R4**, there exists an M -alternating path P_1 starting with an edge not in M connecting some vertex $a'_1 \in A'$ to a_1 , with its internal vertices in $A_i^- \cup B_i^+$. Also by **R4**, there exists some M -alternating path P_2 starting with an edge in M connecting b_1 to some vertex $b'_1 \in B'$, with its internal vertices in $B_i^- \cup A_i^+$. In particular, P_1 and P_2 are disjoint. Then, the path $P_1 a_1 b a b_1 P_2$ is an M -augmenting path beginning and ending with edges not in M connecting $a' \in A'$ and $b' \in B'$, contradicting that M is a maximum matching. Let $A_{i+1}^+ = A_i^+ \cup \{a\}$, $\bar{A}_{i+1} = \bar{A}_i \setminus \{a\}$, $B_{i+1}^- = B_i^- \cup \{b\}$, $\bar{B}_{i+1} = \bar{B}_i \setminus \{b\}$, and keep A_{i+1}^-, B_{i+1}^+ unchanged. Then $G[A' \cup A_{i+1}^-, B' \cup B_{i+1}^-]$ is still an empty graph, and $b a b_1 P_2$ is an M -alternating path starting with an edge in M connecting b to $b'_1 \in B'$, with internal vertices in $B_{i+1}^- \cup A_{i+1}^+$. It follows that **R1–R4** are all maintained.

If instead there exists $b \in \bar{B}_i$ adjacent to some $a_1 \in A' \cup A_i^-$, then let $a \in \bar{A}_i$ be the vertex matched to b in M , set $A_{i+1}^- = A_i^- \cup \{a\}$, $\bar{A}_{i+1} = \bar{A}_i \setminus \{a\}$, $B_{i+1}^+ = B_i^+ \cup \{b\}$, $\bar{B}_{i+1} = \bar{B}_i \setminus \{b\}$, and keep A_{i+1}^+, B_{i+1}^- unchanged. Like above, **R1–R4** are all maintained, which finishes the proof. $\square$

5.7 Proof of Theorem 2.2

Proof of Theorem 2.2. Let $1/n \ll c \ll \varepsilon \ll \mu \ll 1$ and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$ and $t_1 \geq t_2$. Let G be a red/blue coloured complete graph with $\max\{t_1 + 2t_2, 2t_1\} - 1$ vertices. Suppose that G is not Type I (μ, t_1, t_2) -extremal, and either $t_1 < (2 - \mu)t_2$ or G is not Type II (μ, t_1, t_2) -extremal. Let T be

any n -vertex tree with $\Delta(T) \leq cn$ and bipartition class sizes t_1 and t_2 , we need to show that G contains a monochromatic copy of T .

If $t_1 \leq 2t_2$, let $T' = T$, $t'_1 = t_1$ and $t'_2 = t_2$. If $t_1 \geq 2t_2 + 1$, then by Lemma 2.10, T contains a set L of $2/c$ leaves in V_1 . Let T' be a new tree obtained by attaching $\lceil (t_1 - 2t_2)/2 \rceil$ new leaves to vertices in L , with no vertex in L receiving more than cn of these new leaves. Then, observe that $\Delta(T') \leq c|T'|$, and the bipartition classes of T' have sizes $t'_1 = t_1$ and $t'_2 = \lceil t_1/2 \rceil$. It follows that $t'_1 \leq 2t'_2$ and $|G| = 2t_1 - 1 = 2t'_1 - 1 \geq t'_1 + 2t'_2 - 2 \geq (1 - \varepsilon)(t'_1 + 2t'_2)$. Therefore, Theorem 5.1 implies that we have the structure required to apply Lemma 5.9, which then implies that one of the following is true.

S1 G contains a monochromatic copy of T' .

S2 G is Type I $(\mu/2, t'_1, t'_2)$ -extremal.

S3 $t'_1 \geq (2 - \mu/2)t'_2$ and G is Type II $(\mu/2, t'_1, t'_2)$ -extremal.

Note that each of **S2** and **S3** implies that the same is true with t_1, t_2, μ in place of $t'_1, t'_2, \mu/2$, respectively, which contradicts our assumption. Therefore, it must be that **S1** is true, which implies that G contains a monochromatic copy of T as well, finishing the proof. $\square$

6 Proof of Theorem 2.3: Type I extremal graphs

In this section, we prove Theorem 2.3. We will start by outlining the proof in Section 6.1, breaking it down into different cases that are then proved throughout the rest of the section.

6.1 Proof outline for Type I extremal graphs

We start by recapping the situation in Theorem 2.3, where we have parameters $1/n \ll c \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 satisfying $t_1 \geq t_2$. Let G be a red/blue coloured complete graph on $\max\{2t_1, t_1 + 2t_2\} - 1$ vertices which is Type I (μ, t_1, t_2) -extremal. This means that there are disjoint sets $U_1, U_2 \subset V(G)$ such that $|U_1| \geq (1 - \mu)n$, $|U_2| \geq (1 - \mu)t_2$, $d_{\text{red}}(u, U_1) \leq \mu n$ for every $u \in U_1$, and $d_{\text{blue}}(u, U_{3-i}) \leq \mu n$ for every $i \in [2]$ and $u \in U_i$. We wish to find a monochromatic copy of T in G . As mentioned at the end of Section 2.1, we can and will assume that $t_1 \leq 2t_2 + 1$, so $|G| \in \{n + t_2 - 1, n + t_2\}$.

Let us first partition the vertices outside of $U_1 \cup U_2$. For some β with $\mu \ll \beta \ll 1$, partition $V(G) = U_1^+ \cup U_2^+$ so that vertices in U_2^+ have at least βn red neighbours in U_1 and vertices in U_1^+ do not. In particular, $U_1 \subset U_1^+$, $U_2 \subset U_2^+$, and at least one of $|U_1^+| \geq n$ and $|U_2^+| \geq t_2$ must hold. Say $|U_1^+| = n + k$ for some k satisfying $|k| \leq 2\mu n$, from which it follows that $|U_2^+| \in \{t_2 - k - 1, t_2 - k\}$.

A first approach to proving Theorem 2.3 would be to embed T into $G_{\text{blue}}[U_1^+]$ if $|U_1^+| \geq n$, and into $G_{\text{red}}[U_1^+, U_2^+]$ if $|U_2^+| \geq t_2$. We will be able to do this when the tree has many, say $n/100$, vertex-disjoint bare paths with length 5, which we call **Case I.A**. This is split into **Case I.A.1** and **Case I.A.2** depending on whether $k \geq 0$. To embed such a tree T , we first remove many suitable bare paths with length 4, so that the remaining forest T' can be embedded greedily. Then, we find a collection of vertex-disjoint paths with length 2 in G that will form the middle parts of the missing paths, and make sure that they cover all the low degree vertices in G if necessary. Finally, we attach these length 2 paths to the image of T' to complete a copy of T by verifying the appropriate Hall's matching conditions. These proofs are carried out in Section 6.2.

If T does not have $n/100$ vertex-disjoint bare paths with length 5, we call this **Case I.B**. Note that by Lemma 2.11, T must then have at least $n/20$ leaves. Unlike in **Case I.A** above where no spare vertex is required to embed T , it follows from a famous example of Komlós, Sárközy and Szemerédi [23] that we cannot necessarily embed T into $G_{\text{blue}}[U_1^+]$ even if $|U_1^+| \geq n$, or into $G_{\text{red}}[U_1^+, U_2^+]$ even if $|U_2^+| \geq t_2$. For example, in our setting, it is easy to create an n -vertex tree T with $\Delta(T) \leq cn$ in which there is a set W of at most $2/c$ vertices such that every edge of the tree contains a vertex in W . Thus, $G_{\text{blue}}[U_1^+]$ can only contain a blue copy of T if it has a set W' with size at most $2/c$ for which at most $k = |U_1^+| - n$ vertices in $U_1^+ \setminus W'$ have no blue neighbour in W' . An appropriate red sparse binomial random graph placed on U_1^+ can destroy this property, even when $|U_1^+| = n + \Theta_c(n)$, while having maximum degree at most cn .

(see [23] for a more detailed example). Thus, whether we embed the tree in red or blue will depend not only on k , but also on the number of edges in $G_{\text{red}}[U_1^+]$ and $G_{\text{blue}}[U_1^+, U_2^+]$.

We divide **Case I.B** (where T has many leaves) into two further subcases, using a parameter D , which is defined to be the $30\mu n$ -th biggest value of $d_{\text{blue}}(u, U_2^+)$ across all $u \in U_1^+$. The hope is that this could allow us to embed D leaves of the tree T using the edges in $G_{\text{blue}}[U_1^+, U_2^+]$, thus creating more space to embed the rest of the tree into $G_{\text{blue}}[U_1^+]$. With this reasoning, one might expect that if $k + D \geq 0$, then as $|U_1^+| + D = n + k + D \geq n$, we will be able to embed the tree in blue. However, an example similar to the one mentioned above from [23] shows that simply requiring $k + D \geq 0$ is not enough, as comparatively few red edges in $G[U_1^+]$ can spoil this embedding attempt in blue. However, in Section 6.3, we will show that we can embed the tree in blue as long as $G[U_1^+]$ has at most $10^7(k + D + 1)n$ red edges, which we call **Case I.B.1**.

Our final case, **Case I.B.2**, is when there are more than $10^7(k + D + 1)n$ red edges in $G[U_1^+]$. These red edges will allow us to embed into $G_{\text{red}}[U_1^+]$ some small subtree of T , which contains at least $k + D + 1$ vertices from V_2 , the bipartition class of T with size t_2 . Note that to embed the remaining vertices of V_2 , we now have at least $k + D + 1 - k - 1 \geq D$ spare vertices available in U_2^+ . As most of the vertices in U_1^+ have at most D blue neighbours in U_2^+ , we will be able to embed the rest of the tree essentially greedily, with some slight complications such as vertices in $U_2^+ \setminus U_2$ having fewer red neighbours in U_1^+ .

In Section 6.2, we prove the embedding results used for both **Case I.A.1** and **Case I.A.2**. In Section 6.3, we prove the embedding result used for **Case I.B.1**, stated so that it will also be useful in Section 7. In Section 6.4, we prove the embedding result used for **Case I.B.2**. Finally, we put all of these together in Section 6.5 to prove Theorem 2.3. To finish this outline, we recap the different cases in the proof of Theorem 2.3, noting the main result that takes care of each of them.

I G is Type I extremal.

I.A T has at least $n/100$ vertex-disjoint bare paths with length 5.

I.A.1 $k \geq 0$: T embeds in blue.

Lemma 6.1

I.A.2 $k < 0$: T embeds in red.

Lemma 6.2

I.B T has at least $n/20$ leaves.

I.B.1 $e(G_{\text{red}}[U_1^+]) \leq 10^7(k + D + 1)n$: T embeds in blue.

Lemma 6.3

I.B.2 $e(G_{\text{red}}[U_1^+]) > 10^7(k + D + 1)n$: T embeds in red.

Lemma 6.6

6.2 Case I.A: trees with many bare paths in Type I extremal graphs

In **Case I.A.1**, we will use the following result that allows us to embed spanning trees with many short bare paths into an almost complete graph. A stronger version in which the degree condition is weakened like in Lemma 6.2 also holds, though this is not needed in our proof.

Lemma 6.1. *Let $1/n \ll \mu \ll 1$. Let H be an n -vertex graph with $\delta(H) \geq (1 - \mu)n$, and let T be an n -vertex tree that contains $10\mu n$ vertex-disjoint bare paths with length 4. Then, for any $t \in V(T)$ not on any of these bare paths and any $s \in V(H)$, there is a copy of T in H with t copied to s .*

Proof. From assumption, there is a collection $\mathcal{P} = \{P_1, \dots, P_\ell\}$ of $\ell = 10\mu n$ vertex-disjoint bare paths in T with length 4, such that t is not on any of these bare paths. Let T' be the forest obtained by removing all internal vertices of the paths in $\mathcal{P}$ from T , so that $|T'| = n - 3\ell$.

Since $\delta(H) \geq (1 - \mu)n \geq |T'|$, by Lemma 2.17, we can greedily find a copy S' of T' in H with t copied to s . Let $H' = H - V(S')$. Note that $|H'| = 3\ell = 30\mu n$, so $\delta(H') \geq |H'| - \mu n \geq |H'|/2$. Thus, by Dirac's theorem, H' contains a Hamilton cycle. In particular, we can label the vertices in H' as $w_1, \dots, w_\ell, x_1, \dots, x_\ell, y_1, \dots, y_\ell$, so that for each $i \in [\ell]$, $x_i w_i y_i$ is a path in H' .

For each $i \in [\ell]$, let u_i, v_i be the copies of the endpoints of P_i in S' . Let K be an auxiliary bipartite graph with bipartition classes $A = \{a_1, \dots, a_\ell\}$ and $B = \{b_1, \dots, b_\ell\}$, such that for any $i, j \in [\ell]$, there is an edge $a_i b_j$ in K if and only if both $u_i x_j$ and $v_i y_j$ are edges in H . Since $\delta(H) \geq (1 - \mu)n$, for every $i \in [\ell]$, $d_K(a_i), d_K(b_i) \geq \ell - 2\mu n$. Then, for any $I \subset A$ with $0 < |I| \leq \ell - 2\mu n$, we have $|N_K(I, B)| \geq \ell - 2\mu n \geq |I|$, while for any $I \subset A$ with $|I| > \ell - 2\mu n$, we have $|N_K(I, B)| = |B| \geq |I|$, as any $b \in B \setminus N_K(I, B)$ would satisfy $d_K(b) < 2\mu n < \ell - 2\mu n$, a contradiction. Thus, by Lemma 2.8, there is a perfect matching in K , say

matching a_i with $b_{\sigma(i)}$ for every $i \in [\ell]$. This then implies that S' along with the paths $u_i x_{\sigma(i)} w_{\sigma(i)} y_{\sigma(i)} v_i$ for all $i \in [\ell]$ form a copy of T in H , as required. $\square$

In **Case I.A.2**, we will use the following bipartite version of Lemma 6.1. We prove it in a more flexible form so that we may also use it later in the proof of Lemma 6.6.

Lemma 6.2. *Let $1/n \ll \mu \ll \beta \ll 1$. Let H be a bipartite graph with bipartition classes U_1 and U_2 such that $d(u, U_2) \geq |U_2| - \mu n$ for every $u \in U_1$, $d(u, U_1) \geq \beta n$ for every $u \in U_2$, and $d(u, U_1) \geq |U_1| - \mu n$ for all but a set W of at most μn vertices in U_2 .*

Let T be an n -vertex forest with bipartition classes V_1 and V_2 such that $|V_j| \leq |U_j|$ for each $j \in [2]$. Let R be a subforest of T with $|R| \leq \beta n/2$, such that $T - R$ contains a collection $\mathcal{P}$ of $10\mu n$ vertex-disjoint bare paths of length 4 whose endpoints are all in V_2 . Suppose that $H - W$ contains a copy S of R with vertices in $V(R) \cap V_j$ copied into U_j for each $j \in [2]$, then S can be extended to a copy of T in H , such that W is covered by the central vertices of the bare paths in $\mathcal{P}$.

Proof. Let $\ell = 10\mu n$, and label the paths in $\mathcal{P}$ as $P_1, \dots, P_\ell$. Let T' be the forest obtained by removing the internal vertices of $P_1, \dots, P_\ell$ from T , noting that $|V(T') \cap V_1| = |V_1| - 2\ell$ and $|V(T') \cap V_2| = |V_2| - \ell$.

Let $W = \{w_1, \dots, w_r\}$, where $r \leq \mu n$. Since $d(w, U_1 \setminus V(S)) \geq \beta n - \beta n/2 \geq 2\mu n$ for every $w \in W$, we can greedily find distinct vertices $x_1, \dots, x_r, y_1, \dots, y_r \in U_1 \setminus V(S)$, such that $x_i, y_i \in N(w_i)$ for every $i \in [r]$. Let $W^+ = \{w_i, x_i, y_i : i \in [r]\}$, and $H' = H - W^+$.

Note that every vertex in H' has at most μn non-neighbours in the opposite side of the bipartition. Using $\ell = 10\mu n$ and Lemma 2.18, we can greedily extend the copy S of R to a copy S' of T' in H' , in which $V(T') \cap V_j$ is copied into $U_j \setminus W^+$ for each $j \in [2]$. For every $i \in [\ell]$, let $u_i, v_i \in U_2$ be the copies of the two endpoints of P_i in S' . To complete a copy of T , it suffices to find, for every $i \in [\ell]$, a u_i, v_i -path with length 4 using distinct new internal vertices.

Let $H'' = H' - V(S') = H - W^+ - V(S')$. Then $|V(H'') \cap U_1| \geq |V_1| - 2r - (|V_1| - 2\ell) = 2(\ell - r)$, and similarly $|V(H'') \cap U_2| \geq |V_2| - r - (|V_2| - \ell) \geq \ell - r \geq 5\mu n$. Moreover, every vertex in H'' has at most μn non-neighbours in the opposite side. Arbitrarily pick distinct $w_{r+1}, \dots, w_\ell \in V(H'') \cap U_2$, we claim that there exist distinct $x_{r+1}, \dots, x_\ell, y_{r+1}, \dots, y_\ell \in V(H'') \cap U_1$ such that $x_i, y_i \in N(w_i)$ for every $i \in [\ell] \setminus [r]$. Indeed, by Lemma 2.9, it suffices to show that for every $\emptyset \neq I \subset [\ell] \setminus [r]$, $|N(\{w_i : i \in I\}, V(H'') \cap U_1)| \geq 2|I|$. If $0 < |I| \leq \ell - r - \mu n/2$, then $|N(\{w_i : i \in I\}, V(H'') \cap U_1)| \geq |V(H'') \cap U_1| - \mu n \geq 2(\ell - r) - \mu n \geq 2|I|$. If $|I| > \ell - r - \mu n/2$, then $|N(\{w_i : i \in I\}, V(H'') \cap U_1)| = |V(H'') \cap U_1| \geq 2(\ell - r) \geq 2|I|$, as any $u \in V(H'') \cap U_1$ that is not adjacent to any w_i with $i \in I$ would have at least $|I| \geq 2\mu n$ non-neighbours in $V(H'') \cap U_2$, a contradiction.

Therefore, together with W^+ , we have found distinct $x_1, \dots, x_\ell, y_1, \dots, y_\ell, w_1, \dots, w_\ell \in V(H) \setminus V(S')$, such that $x_i, y_i \in N(w_i)$ for every $i \in [\ell]$. Let K be an auxiliary bipartite graph with bipartition classes $A = \{a_1, \dots, a_\ell\}$ and $B = \{b_1, \dots, b_\ell\}$, such that for any $i, j \in [\ell]$, $a_i b_j \in E(K)$ if and only if both $u_i x_j$ and $v_i y_j$ are in $E(H)$. From construction, for each $i \in [\ell]$, $u_i, v_i \in U_2 \setminus W$, so $d_K(a_i) \geq \ell - 2\mu n$. Similarly, $d_K(b_j) \geq \ell - 2\mu n$ for every $j \in [\ell]$. Like in Lemma 6.1, we can now use these degree conditions to verify that for every $I \subset A$, $|N_K(I, B)| \geq |I|$, so Lemma 2.8 gives a perfect matching in K , say matching a_i with $b_{\sigma(i)}$ for every $i \in [\ell]$. Then, S' together with the paths $u_i x_{\sigma(i)} w_{\sigma(i)} y_{\sigma(i)} v_i$ for all $i \in [\ell]$ form a copy of T in H , with vertices in W covered by the central vertices of the bare paths in $\mathcal{P}$, as required. $\square$

6.3 Case I.B.1: embedding trees in almost complete graphs

We now prove the main result to be used in **Case I.B.1**. This is proved in a slightly stronger form so that we may also use it in Section 7.

Lemma 6.3. *Let $1/n \ll c \ll \mu, \beta \ll 1$, let $|k| \leq \mu n$ and $0 \leq D \leq \mu n$ satisfy $k + D \geq 0$. Let G be a graph with a vertex partition $U_1 \cup U_2$ such that $|U_1| = n + k$ and $\delta(G[U_1]) \geq |U_1| - \beta n$. Let $X \subset U_1$ satisfy $|X| \leq \mu n$ and $d_G(u, U_2) \geq n/10$ for each $u \in X$. Suppose $e(G[U_1 \setminus X]) \leq 10^7(k + D + 1)n$, and there are at least $10\mu n$ vertices in U_1 with at least D neighbours in U_2 .*

Let T be an n -vertex tree with $\Delta(T) \leq cn$ such that, if $D > 0$, then T has at least $n/20$ leaves. Then, G contains a copy of T . Moreover, if $D = 0$ and $X = \emptyset$, then for any $t \in T$ and $s \in U_1$, there is a copy of T in G with t copied to s .

Proof. If $D = 0$ and T has fewer than $n/20$ leaves, then T has at least $n/100$ vertex-disjoint bare paths with length 4 by Lemma 2.11. Since $k \geq -D = 0$, the result follows by applying Lemma 6.1 to $G[U_1]$.

Therefore, we can assume that $D \geq 0$ and T has at least $n/20$ leaves. If $D \neq 0$ or $X \neq \emptyset$, pick $t \in V(T)$ arbitrarily. Then, there exists a set L' of $n/49$ leaves in T , which does not contain t or any neighbours of t , and either all belong to V_1 or all belong to V_2 . Let $P = N_T(L')$ be the set of parents of L' in T , and note that P is an independent set. Let $P_1 \subset P$ be a set of size at most D such that $t \notin P_1$ and $D \leq |N_T(P_1, L')| \leq n/150$, which is possible as $\Delta(T) \leq cn$. Similarly, let $P'_2 \subset P \setminus P_1$ be a set of size at most $\lceil |X|/2 \rceil$ such that $t \notin P'_2$ and $\lceil |X|/2 \rceil \leq |N_T(P'_2, L')| \leq n/150$. Let $P_3 = P \setminus (P_1 \cup P'_2)$, so $|N_T(P_3, L')| \geq n/149$. In particular, we can add a set L'_3 of $\lceil |X|/2 \rceil - |P'_2|$ leaves adjacent to P_3 to the set P'_2 to obtain a set P_2 of size $\lceil |X|/2 \rceil$. Note that $P_1 \cup P_2$ is still an independent set. Let $L = L' \setminus L'_3$ and let $m = |T - L|$. Observe that $n/50 \leq |L| \leq n/49$ and $|N_T(P_3, L)| \geq n/150$. Let $t_1 = t$, and let $t_1, t_2, \dots, t_m$ be an ordering of $V(T - L)$ so that each vertex apart from t_1 has exactly 1 neighbour in T to its left in this ordering. Let $d_i = d_T(t_i, L)$ for every $i \in [m]$.

From assumption, we can take a set Y of $2D$ vertices in $U_1 \setminus X$, each having at least D neighbours in U_2 . Let $U_1^- = \{u \in U_1 : d_G(u, Y) \geq D \text{ and } d_G(u, X) \geq |X|/2\}$. Since $\delta(G[U_1]) \geq |U_1| - \beta n$, if $D > 0$, by double counting there are at most $2D\beta n/D = 2\beta n$ vertices $u \in U_1$ with $d_G(u, Y) < D$, while there is no such vertex if $D = 0$. Similarly there are at most $2\beta n$ vertices $u \in U_1$ with $d_G(u, X) < |X|/2$. Thus, $|U_1 \setminus U_1^-| \leq 4\beta n$. If $D = 0$ and $X = \emptyset$, note that $s \in U_1 = U_1^- \setminus (X \cup Y)$ already. Otherwise, pick $s \in U_1^- \setminus (X \cup Y)$ arbitrarily.

Let $s_1 = s$, so that $s_1 \in U_1^- \setminus (X \cup Y)$. For each $1 < i \leq m$ in turn, embed t_i as follows, where $j_i < i$ is such that $t_{j_i} t_i \in E(T - L)$.

T1 If $t_i \in P_1$, select s_i uniformly at random from $N_G(s_{j_i}, Y \setminus \{s_1, \dots, s_{i-1}\})$.

T2 If $t_i \in P_2$, select s_i uniformly at random from $N_G(s_{j_i}, X \setminus \{s_1, \dots, s_{i-1}\})$.

T3 If $t_i \notin P_1 \cup P_2$, select s_i uniformly at random from $N_G(s_{j_i}, U_1^- \setminus (X \cup Y \cup \{s_1, \dots, s_{i-1}\}))$.

Note that **T1** is always possible as $t_{j_i} \notin P_1 \cup P_2$, so $s_{j_i} \in U_1^-$ has at least D neighbours in Y , and at most $|P_1| - 1 \leq D - 1$ of them have been used. Similarly, **T2** is always possible. **T3** is always possible as

$$\begin{aligned} |N_G(s_{j_i}, U_1^- \setminus (X \cup Y \cup \{s_1, \dots, s_{i-1}\}))| &\geq |U_1^-| - \beta n - |X| - |Y| - |T - L| \\ &\geq n + k - 4\beta n - \beta n - \mu n - 2D - n + |L| \geq n/100 > 0. \end{aligned}$$

Therefore, this random process always succeeds in producing a copy of $T - L$ in $G[U_1]$. Note that, in particular, we have $|X \setminus \{s_1, \dots, s_m\}| = |X| - |P_2| = \lfloor |X|/2 \rfloor$.

Now, let j^* be the smallest integer so that $2^{j^*} > cn$. For each $j \in [j^*]$, let I_j be the set of $\{i : t_i \in P_3\}$ with $2^{j-1} \leq d_i < 2^j$, and say a vertex v is j -bad if

$$|\{s_i : i \in I_j\} \cap N_G(v)| \leq \frac{2}{3}|I_j| - 20(j^* - j + 1).$$

If v is j -bad for some $j \in [j^*]$, then say v is *bad*. For each $v \in U_1 \setminus X$, let m_v be the number of non-neighbours of v in $U_1 \setminus X$, so that $m_v \leq \beta n$ from assumption.

Claim 6.4. For each $v \in U_1 \setminus X$,

$$\mathbb{P}(v \text{ is bad}) \leq \frac{m_v}{10^8 n}.$$

Proof of Claim 6.4. Let $j \in [j^*]$. Note that if $|I_j| < 30(j^* - j + 1)$, then v cannot be j -bad, so $\mathbb{P}(v \text{ is } j\text{-bad}) = 0$. Assume now that $|I_j| \geq 30(j^* - j + 1)$. Note that for every $i \in I_j$, $t_i \notin P_1 \cup P_2$, so by **T3** and conditioning on any choices of $s_1, \dots, s_{i-1}$, we have

$$\mathbb{P}(s_i \notin N_G(v) \mid s_1, \dots, s_{i-1}) \leq \frac{m_v}{|N_G(s_{j_i}, U_1^- \setminus (X \cup Y \cup \{s_1, \dots, s_{i-1}\}))|} \leq \frac{m_v}{n/100}.$$

Therefore, for each $I \subset I_j$, if $E_{v,j,I}$ is the event that $s_i \notin N_G(v)$ for each $i \in I$, then

$$\mathbb{P}(E_{v,j,I}) = \prod_{i \in I} \mathbb{P}(s_i \notin N_G(v) \mid s_{i'} \notin N_G(v) \text{ for each } i' \in I \text{ less than } i) \leq \left(\frac{100m_v}{n} \right)^{|I|}.$$

Since v is j -bad implies that $E_{v,j,I}$ holds for some $I \subset I_j$ with size $k_j := \lfloor |I_j|/3 \rfloor \geq 10(j^* - j + 1)$, we can use a union bound to get

$$\mathbb{P}(v \text{ is } j\text{-bad}) \leq \binom{|I_j|}{k_j} \left(\frac{100m_v}{n} \right)^{k_j} \leq \left(\frac{10^4 m_v}{n} \right)^{k_j} \leq \left(\frac{10^4 m_v}{n} \right)^{10(j^* - j + 1)} \leq \left(\frac{m_v}{10^{10}n} \right)^{(j^* - j + 1)},$$

where we used $m_v \leq \mu n$ and $1/n \ll \mu \ll 1$. Thus,

$$\mathbb{P}(v \text{ is bad}) \leq \sum_{j=1}^{j^*} \mathbb{P}(v \text{ is } j\text{-bad}) \leq \sum_{j=1}^{j^*} \left(\frac{m_v}{10^{10}n} \right)^{(j^* - j + 1)} \leq \frac{m_v}{10^8 n},$$

as required. $\square$

By Claim 6.4, we have

$$\mathbb{E}|\{v \in U_1 \setminus X : v \text{ is bad}\}| \leq \sum_{v \in U_1 \setminus X} \frac{m_v}{10^8 n} \leq \frac{2e(G^c[U_1 \setminus X])}{10^8 n} \leq \frac{k + D + 1}{5},$$

so by Markov's inequality, with probability at least $1/2$ there are at most $\lfloor (k + D + 1)/2 \rfloor$ bad vertices in $U_1 \setminus X$. Thus, we can take a realisation of this random embedding that has at most $\lfloor (k + D + 1)/2 \rfloor$ bad vertices in $U_1 \setminus X$.

Let U'_1 be the set of unused vertices in U_1 , noting that $|U'_1| = n + k - |T - L| = k + |L|$ and $|U'_1 \cap X| = \lfloor |X|/2 \rfloor$. To complete the embedding, we use Lemma 2.9 to embed the leaves in L into $U'_1 \cup U_2$. It suffices to show that $N_G(\{s_i : t_i \in J\}, U'_1 \cup U_2) \geq \sum_{i:t_i \in J} d_i$ for every $\emptyset \neq J \subset P$.

First suppose that $0 < \sum_{i:t_i \in J} d_i \leq 999|L|/1000$, then, as required,

$$|N_G(\{s_i : t_i \in J\}, U'_1)| \geq |U'_1| - \beta n = k + |L| - \beta n \geq 999|L|/1000 \geq \sum_{i:t_i \in J} d_i.$$

Now suppose that $\sum_{i:t_i \in J} d_i > 999|L|/1000$. Let $L_3 = N_T(P_3, L)$, recall that $|L| \leq n/49$ and $|L_3| \geq n/150$. Thus, $\sum_{i:t_i \in J \cap P_3} d_i \geq |L_3| - |L|/1000 \geq 9|L_3|/10$. Then, for any $v \in U_1 \setminus X$ which is not bad,

$$\begin{aligned} \sum_{\substack{i:t_i \in P_3, \\ s_i \notin N_G(v)}} d_i &\leq \sum_{j=1}^{j^*} \sum_{i \in I_j : s_i \notin N_G(v)} 2^j \leq \sum_{j=1}^{j^*} \left(\frac{1}{3}|I_j| + 20(j^* - j + 1) \right) \cdot 2^j \\ &\leq \frac{1}{3} \sum_{j=1}^{j^*} \sum_{i \in I_j} 2d_i + \sum_{j=1}^{j^*} 20(j^* - j + 1) \cdot 2^j \\ &\leq \frac{2|L_3|}{3} + 80 \cdot 2^{j^*} \leq \frac{2|L_3|}{3} + 200cn < \frac{9|L_3|}{10} \leq \sum_{i:t_i \in J \cap P_3} d_i. \end{aligned}$$

Thus, v is adjacent to some vertex in $\{s_i : t_i \in J \cap P_3\}$. Since at most $\lfloor (k + D + 1)/2 \rfloor$ vertices in $U_1 \setminus X$ are bad, this implies that

$$\begin{aligned} |N_G(\{s_i : t_i \in J\}, U'_1)| &\geq |U'_1| - |U'_1 \cap X| - \lfloor (k + D + 1)/2 \rfloor \\ &= |L| + k - \lfloor |X|/2 \rfloor - \lfloor (k + D + 1)/2 \rfloor \geq |L| - D - \lceil |X|/2 \rceil, \end{aligned}$$

so we are done if $\sum_{i:t_i \in J} d_i \leq |L| - D - \lceil |X|/2 \rceil$.

Suppose now that $\sum_{i:t_i \in J} d_i > |L| - D - \lceil |X|/2 \rceil$. If $J \cap P_2 = \emptyset$, then $\sum_{i:t_i \in J} d_i \leq |L| - \lceil |X|/2 \rceil$ as $\sum_{i:t_i \in P_2} d_i \geq \lceil |X|/2 \rceil$. Hence, $J \cap P_1 \neq \emptyset$ as $\sum_{i:t_i \in P_1} d_i \geq D$, and thus

$$\begin{aligned} |N_G(\{s_i : t_i \in J\}, U'_1 \cup U_2)| &= |N_G(\{s_i : t_i \in J\}, U'_1)| + |N_G(\{s_i : t_i \in J\}, U_2)| \\ &\geq |L| - D - \lceil |X|/2 \rceil + D = |L| - \lceil |X|/2 \rceil \geq \sum_{i:t_i \in J} d_i, \end{aligned}$$

as required. If $J \cap P_2 \neq \emptyset$, then we have

$$\begin{aligned} |N_G(\{s_i : t_i \in J\}, U'_1 \cup U_2)| &= |N_G(\{s_i : t_i \in J\}, U'_1)| + |N_G(\{s_i : t_i \in J\}, U_2)| \\ &\geq |L| - D - \lceil |X|/2 \rceil + n/10 \geq |L| \geq \sum_{i:t_i \in J} d_i. \end{aligned}$$

Thus, by Lemma 2.9, we can embed L into $U'_1 \cup U_2$ to finish a copy of T in G . $\square$

6.4 Case I.B.2: embedding trees in almost complete bipartite graphs

In **Case I.B.2**, using the notations in Section 6.1, we want to embed a small subtree into $G_{\text{red}}[U_1^+]$, before embedding the rest of T into $G_{\text{red}}[U_1^+, U_2^+]$. To find this small subtree, we use the following result.

Proposition 6.5. *Let $n, m \in \mathbb{N}$ satisfy $1 \leq m \leq n/18$. Let T be an n -vertex tree with bipartition classes V_1 and V_2 , where $|V_2| \geq |V_1|/3$. Then, T contains a subtree T' with $|T'| \leq 10^4 m$, such that it contains at least m vertices in V_2 and has at most 1 vertex in V_2 with a neighbour outside of $V(T')$ in T .*

Proof. For any tree R on at least $18m$ vertices, by repeated applications of Corollary 2.14, we can find subtrees R_1 and R_2 with a unique common vertex that decompose R , such that $6m \leq |R_1| \leq 18m$. Iterating this in the tree T to find a subtree with size between $6m$ and $18m$ at a time, we obtain a sequence $T_1, \dots, T_\ell$ of subtrees satisfying the following properties.

- $E(T_1), \dots, E(T_\ell)$ partition $E(T)$.
- For every $j \in [\ell]$, $6m \leq |T_j| \leq 18m$.
- For every $j \in [\ell]$, $\cup_{i=1}^j T_i$ is a tree.
- For every $2 \leq j \leq \ell$, T_j shares a unique vertex with $\cup_{i=1}^{j-1} T_i$.

Note that $\ell \leq n/(6m-1)$. We claim that $|V(T_i) \cap V_2| \geq m$ for at least $\ell/100$ indices $i \in [\ell]$. Indeed, if not, then there would be at most $\ell m + (\ell/100) \cdot 18m < n/4$ vertices in V_2 , contradicting the assumption that $|V_2| \geq |V_1|/3$. Let $I = \{i \in [\ell] : |V(T_i) \cap V_2| \geq m\}$.

Consider the auxiliary graph K with vertex set $[\ell]$, where for any $i < j$ in $[\ell]$, $ij \in E(K)$ if and only if i is the smallest index such that $V(T_i) \cap V(T_j) \neq \emptyset$. Note that every $1 < j \leq \ell$ has a unique neighbour in $[j-1]$ in K , so K is a tree and $e(K) = \ell - 1$. It follows that there exists $i \in I$ with $d_K(i) \leq 200$.

Let t_i be the unique vertex shared by T_i and $\cup_{j < i} T_j$. Then, $N_T(V(T_i) \setminus \{t_i\}, \cup_{j > i} V(T_j)) \leq 200 \cdot 18m$. Let T' be the subtree of T induced by $V(T_i)$ and $N_T(V(T_i) \setminus \{t_i\}, \cup_{j > i} V(T_j)) \cap V_1$. Then $|T'| \leq 10^4 m$, $|V(T') \cap V_2| \geq |V(T_i) \cap V_2| \geq m$, and every vertex in $V(T') \cap V_2$, except possibly t_i , has no neighbour in T outside of $V(T')$. $\square$

Using Proposition 6.5, we can now prove the main result used for **Case I.B.2**.

Lemma 6.6. *Let $1/n \ll c \ll \mu \ll \beta \ll 1$, let $0 \leq D \leq \mu n$, and let $|k| \leq \mu n$. Let T be any n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 with sizes t_1 and t_2 , respectively, that satisfy $t_2 \leq t_1 \leq 2t_2 + 1$. Let G be a graph with a vertex partition $U_1 \cup U_2$ such that $1.1t_1 \leq |U_1| \leq 2n$, $|U_2| = t_2 - k - 1$, and the following hold.*

- $d_G(u, U_1) \geq \beta n$ for each $u \in U_2$, and $d_G(u, U_1) \geq |U_1| - \mu n$ for all but at most μn vertices $u \in U_2$.
- $d_G(u, U_2) \geq |U_2| - D$ for all but at most $10\mu n$ vertices $u \in U_1$.
- There exists $X \subset U_1$ with $|X| \leq 2\mu n$ such that $d_G(u, U_2) \geq |U_2| - \mu n$ for each $u \in U_1 \setminus X$ and $e(G[U_1 \setminus X]) \geq 10^7(k + D + 1)n$.

Then, G contains a copy of T .

Proof. First observe that it suffices to prove this in the case when $k + D \geq -1$. Indeed, if $k + D < -1$, then let $k' = -D - 1$ and remove $-k - D - 1$ vertices from U_2 to obtain U'_2 with size $t_2 - k' - 1$. Note that trivially $e(G[U_1 \setminus X]) \geq 10^7(k' + D + 1)n$, and all other assumptions still hold with U'_2 and k' in place of U_2 and k , so G contains a copy of T . Thus, we assume that $k + D \geq -1$ from now on.

Let $Y_1 = \{u \in U_1 : d_G(u, U_2) < |U_2| - D\}$ and $Y_2 = \{u \in U_2 : d_G(u, U_1) < |U_1| - \mu n\}$, so that $|Y_1| \leq 10\mu n$ and $|Y_2| \leq \mu n$. For each $i \in [2]$, let $U_i^- = U_i \setminus Y_i$. Let $Z \subset U_1^-$ be a random subset chosen by including each vertex independently at random with probability β . By Lemma 2.5, with high probability, we have $|Z| \leq 3\beta n$, $d_G(u, Z) \geq \beta^2 n/2$ for each $u \in U_2$, and $e(G[U_1 \setminus (X \cup Z)]) \geq 10^7(k + D + 1)n/2$, noting that the last condition is trivial when $k + D = -1$. Fix a choice of Z with all of these properties.

As $e(G[U_1 \setminus (X \cup Z)]) \geq 10^7(k + D + 1)n/2$, we can find a subgraph $H' \subset G[U_1 \setminus (X \cup Z)]$ with minimum degree at least $10^7(k + D + 1)/4$. Since each vertex $u \in U_1 \setminus (X \cup Z)$ satisfies $d_G(u, U_2^-) \geq |U_2^-| - \mu n$, there are at least $|H'| |U_2^-|/2$ edges in G between H' and U_2^- . Hence, there exists $v \in U_2^-$ with at least $|H'|/2 \geq 10^6(k + D + 1)$ neighbours in $V(H')$.

By Proposition 6.5, there is a subtree T' of T with $|T'| \leq 10^5(k + D + 1)$ that contains at least $10(k + D + 1)$ vertices in V_2 , and has at most 1 vertex in $V(T') \cap V_2$ with a neighbour in $T - E(T')$. Call such a vertex t if it exists. Let F be the tree obtained from T by contracting T' to a single vertex r . Let L_1 be the set of leaves in F that are in $V_1 \setminus \{r\}$, observing that they are also leaves in T . Note that $|F - L_1| \geq |V_2| - 10^5(k + D + 1) \geq n/4$, so by Lemma 2.11, $F - L_1$ either has at least $n/100$ vertex-disjoint bare paths of length 5 or at least $n/100$ leaves excluding r . Note that any such leaf in $F - L_1$ must be in V_2 , and is also a leaf in $T - L_1$. We now separate into several cases.

Case I. $F - L_1$ contains at least $n/100$ vertex-disjoint bare paths with length 5, then $T - T' - L_1$ contains $n/200$ vertex-disjoint bare paths with length 4 with both endpoints in V_2 . Since there are at $n/200$ central vertices on these paths, and $|L_1| \leq n$, by averaging, we can find a set L_2 of $10\mu n$ such central vertices that are adjacent to at most $200 \cdot 10\mu n \leq \beta^2 n/10$ vertices in L_1 . Embed t to v if t exists, then in any case embed the rest of T' greedily into H' . Note that this embeds at least $\max\{0, 10k + 10D + 9\}$ vertices in V_2 into U_1 , so we now have enough room to apply Lemma 6.2 to extend this embedding of T' to an embedding of $T - L_1$, with $T - T' - L_1$ embedded into the rest of $G[U_1^- \setminus Z, U_2]$ such that Y_2 is covered by vertices in L_2 . To finish, greedily embed leaves in L_1 not adjacent to L_2 into $U_1^- \setminus Z$, possible as the parents of these leaves are embedded into U_2^- , and greedily embed leaves in L_1 adjacent to L_2 into Z , using that every vertex in U_2 has at least $\beta^2 n/2 \geq |N_T(L_2, L_1)|$ neighbours in Z .

Case II. $F - L_1$ contains a set L'_2 of $n/200$ leaves not in $N_T(V(T'))$. Then, we can find a set $L_2 \subset L'_2$ of size $10\mu n$ with $|N_T(L_2, L_1)| \leq 200 \cdot 10\mu n \leq \beta^2 n/10$. Embed t to v if t exists, then in any case embed the rest of T' greedily into H' . Note that this embeds at least $\max\{0, 10k + 10D + 9\}$ vertices in V_2 into U_1 . Next, greedily extend this to embed the rest of $T - L_1 - L_2$ into the rest of $G[U_1^- \setminus Z, U_2^-]$ with vertices in V_i going into U_i^- for each $i \in [2]$, which is possible as $|L_2| = 10\mu n$ and $d(u, U_2^-) \geq |U_2^-| - D \geq t_2 - |L_2|$ for every $u \in U_1^-$. We can then greedily embed L_2 into the rest of U_2 , which is possible as there are at least $t_2 - k - 1 - (t_2 - \max\{0, 10k + 10D + 9\} - |L_2|) \geq |L_2| + D$ vertices left in U_2 , and every vertex in U_1^- is adjacent to all but at most D of them. Finally, greedily embed the leaves in L_1 not adjacent to L_2 into $U_1^- \setminus Z$, and greedily embed the leaves in L_1 adjacent to L_2 into Z , using that every vertex in U_2 has at least $\beta^2 n/2 \geq |N_T(L_2, L_1)|$ neighbours in Z .

Case III. $F - L_1$ contains a set L'_2 of at least $n/200$ leaves in $N_T(V(T'))$. By Lemma 2.13, there exists subtrees T_1, T_2 decomposing T' with a unique common vertex t' , such that $|V(T_1) \cap V_2|, |V(T_2) \cap V_2| \geq 3(k + D + 1)$. Without loss of generality, suppose that $|N_T(V(T_2), L'_2)| \geq n/500$, and pick a set $L_2 \subset N_T(V(T_2), L'_2)$ of size $10\mu n$, with none of them adjacent to t' , such that $|N_T(L_2, L_1)| \leq \beta^2 n/10$. Note that at most two vertices in $V(T_1) \cap V_2$ can have a neighbour in $V(T) \setminus V(T_1)$, namely t if it exists, and t' if it is in V_2 .

If t exists and $t' \in V_2$, then view T_1 as rooted at t , let the parent of t' in T_1 be p , and let the parent of p be p' . Embed t to v , then greedily embed the rest of T_1 into H' with the following exception. Let W be the set of neighbours of the image of p' in H' that are still unused. If p' is embedded into H' , then $|W| \geq 10^7(k + D + 1)/4 - |T_1| \geq 2 \cdot 10^5(k + D + 1)$, while if p' coincides with t then it is embedded to v , so $|W| \geq 10^6(k + D + 1) - |T_1| \geq 2 \cdot 10^5(k + D + 1)$ as well. Like before, there exists $v' \in U_2^-$ with at least $|W|/2 \geq 10^5(k + D + 1)$ neighbours in W . Embed t' to v' , p and $N_{T_1}(t')$ into W , then carry on greedily

to finish the embedding of the rest of T_1 inside H' . If t does not exist or $t' \notin V_2$, we can greedily embed T_1 into $G[V(H') \cup \{v\}]$ such that t is embedded to v if it exists, and the same for t' if it is in V_2 .

In any case, we have an embedding of T_1 into G , with all but at most two vertices embedded into H' , and every vertex in $V(T_1) \cap V_2$ that has any neighbour outside of T_1 is embedded into U_2^- . In particular, at least $3(k + D + 1) - 2 \geq k + D + 1$ vertices in V_2 are embedded into U_1 if $k + D \geq 0$, while the same holds trivially if $k + D = -1$. Next, greedily embed the rest of $T - L_1 - L_2$ into the rest of $G[U_1^- \setminus Z, U_2^-]$. We can then greedily embed L_2 into the rest of U_2 , which is possible as there are at least $t_2 - k - 1 - (t_2 - k - D - 1 - |L_2|) = |L_2| + D$ vertices left in U_2 and every vertex in U_1^- is adjacent to all but at most D of them. Finally, greedily embed leaves in L_1 not adjacent to L_2 into $U_1^- \setminus Z$, and greedily embed leaves in L_1 adjacent to L_2 into Z , using that every vertex in U_2 has at least $\beta^2 n/2 \geq |N_T(L_2, L_1)|$ neighbours in Z . $\square$

6.5 Proof of Theorem 2.3

Having proved all of the embedding results necessary for the Type I extremal case, we can now put them together to prove Theorem 2.3, following the outline in Section 6.1.

Proof of Theorem 2.3. Let $1/n \ll c \ll \mu \ll 1$ and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$ and $t_1 \geq t_2$. Let G be a Type I (μ, t_1, t_2) -extremal graph on $\max\{2t_1, t_1 + 2t_2\} - 1$ vertices, and let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1, V_2 of sizes t_1 and t_2 , respectively. From definition, there are disjoint subsets $U_1, U_2 \subset V(G)$ such that $|U_1| \geq (1 - \mu)n$, $|U_2| \geq (1 - \mu)t_2$, $d_{\text{red}}(u, U_1) \leq \mu n$ for every $u \in U_1$, and $d_{\text{blue}}(u, U_{3-i}) \leq \mu n$ for every $i \in [2]$ and every $u \in U_i$.

First assume that $t_1 \leq 2t_2 + 1$, so $|G| \in \{n + t_2 - 1, n + t_2\}$. Let β be such that $\mu \ll \beta \ll 1$. Let $U_2^+ = \{v \in V(G) : d_{\text{red}}(v, U_1) \geq \beta n\}$, $U_1^+ = V(G) \setminus U_2^+$, and note that $U_1 \subset U_1^+$ and $U_2 \subset U_2^+$. Let $k = |U_1^+| - n$, so $|k| \leq 2\mu n$ and $|U_2^+| \in \{t_2 - k - 1, t_2 - k\}$. If T has at least $n/100$ vertex-disjoint bare paths with length 5, then T has $n/100$ vertex-disjoint bare paths with length 4 whose endpoints are all in V_2 . Thus, there is a blue copy of T in G if $k \geq 0$ by Lemma 6.1, and there is red copy of T in G if $k < 0$ by Lemma 6.2.

Suppose, then, that T does not have at least $n/100$ vertex-disjoint bare paths with length 5, then, by Lemma 2.11, T has at least $n/20$ leaves. Let $D \geq 0$ be the $30\mu n$ -th biggest value of $d_{\text{blue}}(u, U_2^+)$ across all $u \in U_1^+$, and note that $D \leq 3\mu n$. Let $X = \{u \in U_1^+ : d_{\text{blue}}(u, U_2^+) \geq n/10\}$, and note that $X \subset U_1^+ \setminus U_1$, so $|X| \leq 2\mu n$. If $e(G_{\text{red}}[U_1^+ \setminus X]) < 10^7(k + D + 1)n$, then $k + D \geq 0$, so there is a blue copy of T in G by Lemma 6.3, while if $e(G_{\text{red}}[U_1^+ \setminus X]) \geq 10^7(k + D + 1)n$, then there is a red copy of T in G by Lemma 6.6.

Finally, if $t_1 \geq 2t_2 + 2$, then we can take $2/c$ leaves of T in V_1 , which are guaranteed to exist by Lemma 2.10, and attach $\lfloor (t_1 - 2t_2)/2 \rfloor$ new leaves to them, with none of them receiving more than cn new leaves. Let T' be the new tree obtained in this way, and note that the bipartition classes of T' have sizes $t'_1 = t_1$ and $t'_2 = \lfloor t_1/2 \rfloor$, with $t'_1 \leq 2t'_2 + 1$ and $\max\{t'_1 + 2t'_2, 2t'_1\} - 1 = 2t'_1 - 1 = 2t_1 - 1 = \max\{t_1 + 2t_2, 2t_1\} - 1$. Therefore, G contains a monochromatic copy of T' from above, and thus also contains a monochromatic copy of T . $\square$

7 Proof of Theorem 2.4: Type II extremal graphs

In this section, we prove Theorem 2.4. We will start by outlining the proof in Section 7.1, breaking it down into different cases that are then proved throughout the rest of this section.

7.1 Proof outline for Type II extremal graphs

We start by recapping the situation in Theorem 2.4, where we have parameters $1/n \ll c \ll \mu \ll 1$. Suppose $n = t_1 + t_2$ with $t_1 \geq (2 - \mu)t_2$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 of sizes t_1 and t_2 , respectively. Let G be a red/blue coloured complete graph on $\max\{2t_1 - 1, t_1 + 2t_2 - 1\}$ vertices which is Type II (μ, t_1, t_2) -extremal, which means that there are disjoint sets $U_1, U_2 \subset V(G)$ such that $|U_1|, |U_2| \geq (1 - \mu)t_1$, and for every $i \in [2]$ and $u \in U_i$, $d_{\text{red}}(u, U_i) \leq \mu n$ and $d_{\text{blue}}(u, U_{3-i}) \leq \mu n$.

For some β with $\mu \ll \beta \ll 1$, we will start by taking maximal disjoint sets $U_1^+, U_2^+ \subset V(G)$ with $U_1 \subset U_1^+$ and $U_2 \subset U_2^+$, such that for every $i \in [2]$ and $u \in U_i^+$, $d_{\text{red}}(u, U_{3-i}) \geq \beta n$. By relabelling if necessary, we can assume that $|U_1^+| \geq |U_2^+|$.

Our first two cases are reasonably easy. First, in **Case II.A**, we assume that there exist two vertices v_1, v_2 with mostly blue neighbours in both U_1^+ and U_2^+ . This will allow us to embed part of T into $G_{\text{blue}}[U_1^+]$ and the rest of T , apart from at most 2 vertices, into $G_{\text{blue}}[U_2^+]$, then connect them together appropriately using v_1 and v_2 . Using two vertices in this way is optimal, as T may have a vertex whose removal creates exactly three subtrees of roughly equal sizes, so we could not easily fit two of them together into one of $G_{\text{blue}}[U_1^+]$ or $G_{\text{blue}}[U_2^+]$.

Next, in **Case II.B**, we assume that there is a vertex w that has at least βn red neighbours in both U_1^+ and U_2^+ . Then, we decompose T into subtrees T_1 and T_2 with a common neighbour t , so that T_1 is a small subtree containing suitably more vertices in V_1 than in V_2 (see Proposition 7.4). We embed t to w , then embed the rest of T_1, T_2 greedily into $G_{\text{red}}[U_1^+, U_2^+]$ by embedding vertices in $V(T_1) \cap V_2$ and $V(T_2) \cap V_1$ into U_1^+ , and vertices in $V(T_1) \cap V_1$ and $V(T_2) \cap V_2$ into U_2^+ . Observe that T_1 and T_2 are embedded in opposite ways, which ‘rebalances’ the vertices in T across U_1^+ and U_2^+ , so that there is enough room in each side for this embedding to be completed greedily.

Assuming neither of these cases hold, then U_1^+ and U_2^+ together cover all but at most one vertex of G , and $\delta(G_{\text{blue}}[U_i^+]) \geq |U_i^+| - \beta n$ for each $i \in [2]$. Assume without loss of generality that $|U_1^+| \geq |U_2^+|$. If there is a vertex v not in $U_1^+ \cup U_2^+$, then it has mostly blue neighbours in both U_1^+ and U_2^+ , and we can use Corollary 2.14 to decompose T into subtrees T_1 and T_2 with a unique common vertex t such that $|U_1^+ \cup \{v\}| \geq |T_1|$ and $|U_2^+| \geq |T_2| + n/100$. In **Case II.C**, we assume that there are at most $10^6 n$ red edges in $G[U_1^+]$. Using our work in Section 6.3, we can embed T_1 into $G_{\text{blue}}[U_1^+ \cup \{v\}]$ with t embedded to v . Then, as v has plenty of blue neighbours in U_2^+ and $|U_2^+|$ is comfortably larger than $|T_2|$, we can greedily embed T_2 into $G_{\text{blue}}[U_2^+ \cup \{v\}]$ to complete a blue copy of T .

Suppose now that we are not in **Cases II.A–II.C**. Let $k = t_1 - |U_1^+|$, and note that $k \leq 1$ as $|U_1^+| \geq |U_2^+|$. Moreover, if $k = 1$, then there exists $v \in V(G) \setminus (U_1^+ \cup U_2^+)$, and so $G[U_1^+]$ contain at least $10^6 n$ red edges since we are not in **Case II.C**. In theory, we have enough space while attempting to embed T in red to embed all but at most 1 vertex of V_1 into U_1^+ , and V_2 into U_2^+ . While this can always be done when T has many vertex-disjoint bare paths, if T has many leaves instead, then similar to **Case I.B** discussed before, a small number of blue edges in $G[U_1^+, U_2^+]$ can prevent this embedding for some trees.

Therefore, we will first try to embed T in blue again, using the following sparse cut structure.

Definition 7.1. Let T be an n -vertex tree. An (ε, d) -sparse cut in T is a partition $V(T) = A \cup B$ such that the following hold.

- $T[A]$ is a tree and $|A|, |B| \leq (2/3 - \varepsilon)n$.
- For each $v \in A$, $d_T(v, B) \leq d$.
- $\{v \in A : d_T(v, B) > 0\}$ is an independent set in T with size at most $2\Delta(T)$.

Such a sparse cut with $\varepsilon \gg \mu$ and $d = \sqrt{n}$ will be found later using Proposition 7.10, and we will also need some additional properties guaranteed by Proposition 7.11. In **Case II.D**, we assume that $T[A, B]$ can be embedded into $G_{\text{blue}}[U_1^+, U_2^+]$, with, say A embedded into U_1^+ and B embedded into U_2^+ . Then, we use the high minimum degree condition to find an embedding of $T[A]$ in $G_{\text{blue}}[U_1^+]$ that matches enough of the embedding of $T[A, B]$. Using the part that matches, we can greedily extend the embedding of $T[A]$ to embed most of the vertices in B into U_2^+ . The number of remaining vertices in B will be small enough that they can be greedily embedded into the remaining part of $G_{\text{blue}}[U_1^+]$.

Finally, in **Case II.E**, we assume that $T[A, B]$ cannot be embedded into $G_{\text{blue}}[U_1^+, U_2^+]$ as described above. The fact that we failed to do so will imply that there exist $U_A \subset U_1^+$ and $U_B \subset U_2^+$ of suitable sizes, such that every vertex in $U_1^+ \setminus U_A$ has at most $\sqrt{n}$ blue neighbours in $U_2^+ \setminus U_B$. Here, we focus on a specific case where $|U_1^+| \geq t_1$ and T has a set L of many leaves in V_1 , the other case is handled similarly. First, we embed $T - L$ essentially randomly into $G_{\text{red}}[U_1^+, U_2^+]$, with vertices in V_i embedded into U_i^+ for each $i \in [2]$, while ensuring that leaves in L have their parents embedded into $U_2^+ \setminus U_B$. Similar to **Case I.B.1**, in U_1^+ we may have some ‘bad’ vertices to which it is hard to embed the vertices in L . However, it will be likely that there is no bad vertex in $U_1^+ \setminus U_A$ as all these vertices have very high

red degrees into $U_2^+ \setminus U_B$ from assumption. To make sure that we have no uncovered bad vertices in U_A , we use that the size of U_A is related to the sparse cut $V(T) = A \cup B$. Roughly speaking, there will be enough components in $T[N_T(B, A) \cup B]$ that contains a vertex in $V_1 \cap B$ for us to use these vertices to cover U_A . Finally, having ensured that there is no bad vertex, we can embed the leaves in L to complete the embedding of T .

In Sections 7.2–7.6, we will prove the main embedding results used for **Cases II.A–II.E**, respectively, before putting these all together in Section 7.7 to prove Theorem 2.4. To finish this outline, we recap the different cases in the proof of Theorem 2.4, noting the main result that takes care of each of them. In what follows, by ‘otherwise’ we mean that none of the previous cases hold.

II G is Type II extremal.

II.A Two vertices have mostly blue neighbours in U_1^+ and U_2^+ : T embeds in blue. *Lemma 7.3*

II.B Some vertex has βn red neighbours in both U_1^+ and U_2^+ : T embeds in red. *Lemma 7.5*

II.C Otherwise, but one vertex has mostly blue neighbours in U_1^+ and U_2^+ , and $G[U_1^+]$ contains at most $10^6 n$ red edges: T embeds in blue. *Lemma 7.6*

II.D Otherwise, but $T[A, B]$ embeds into $G_{\text{blue}}[U_1^+, U_2^+]$: T embeds in blue. *Lemma 7.7*

II.E Otherwise, either $|U_1^+| \geq t_1$, or $|U_1^+| = t_1 - 1$ and there are at least $10^6 n$ red edges in $G[U_1^+]$: T embeds in red. *Lemma 7.8*

7.2 Case II.A

For **Case II.A**, we first use the following result to find a large subtree of T that contains at most two vertices that have neighbours in the rest of the tree. Moreover, if there are two such vertices, then they are not adjacent.

Proposition 7.2. *Let $1/n \ll \varepsilon \ll 1$. Let T be an n -vertex tree. Then, there is a partition $V(T) = A \cup B$ with $|A|, |B| \leq (2/3 - \varepsilon)n$ such that $T[A]$ is a tree and $\{v \in A : d_T(v, B) > 0\}$ is an independent set in T with size at most 2.*

Proof. Using Corollary 2.14, let T_1 and T_2 be subtrees decomposing T with a unique common vertex t , such that $\frac{n}{3} \leq |T_1| \leq |T_2| \leq 1 + \frac{2n}{3}$. Furthermore, assume that T_1 and T_2 are chosen so that $|T_2|$ is minimised subject to these conditions. If $|T_2| \leq (2/3 - \varepsilon)n$, then set $A = V(T_2)$ and $B = V(T) \setminus A$, and note that the conditions in the lemma hold.

Suppose now that $|T_2| > (2/3 - \varepsilon)n$. If t has only one neighbour, say t' , in T_2 , then adding tt' to T_1 and removing t from T_2 gives two trees that contradict the minimality of $|T_2|$. If $T_2 - t$ has a component S with size at most $(1/3 - 2\varepsilon)n$, then, letting $T'_1 = T[\{t\} \cup V(S)] \cup T_1$ and $T'_2 = T_2 - V(S)$ gives a pair of trees (T'_1, T'_2) that again contradicts the minimality of $|T_2|$, as $\max\{|T'_1|, |T'_2|\} < |T_2|$. Thus, t must have exactly two neighbours in T_2 , and $T_2 - t$ is the disjoint union of two trees S_1 and S_2 with $(1/3 - 2\varepsilon)n \leq |S_1|, |S_2| \leq (1/3 + 2\varepsilon)n$. For each $i \in [2]$, let t_i be the neighbour of t in S_i .

Using Corollary 2.14 again, let S'_1 and S'_2 be subtrees decomposing S_1 with a unique common vertex t'_1 , such that $(1 - 6\varepsilon)n/9 \leq |S'_1|, |S'_2| \leq 1 + (2 + 12\varepsilon)n/9$. Relabelling if necessary, assume that S'_1 contains t_1 . Let $A = V(T_1) \cup V(S'_1) \cup \{t_2\}$ and note that $T[A]$ is the tree made by connecting T_1 and S'_1 with the edge tt_1 and adding the edge tt_2 . Let $B = V(T) \setminus A$, and note that the only vertices in A with neighbours in B in T are t_2 and t'_1 , and they are not adjacent in T because they are in different components of $T_2 - t$. Thus, A and B satisfy the required conditions. $\square$

Using Proposition 7.2, it is now straightforward to prove the following result used in **Case II.A**.

Lemma 7.3. *Let $1/n \ll c \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$. Let G be a graph that contains two disjoint vertex sets U_1 and U_2 such that $|U_i| \geq (2/3 - \mu)n$ and $\delta(G[U_i]) \geq |U_i| - \mu n$ for each $i \in [2]$. Suppose there exist $v_1, v_2 \in V(G) \setminus (U_1 \cup U_2)$ such that $d_G(v_i, U_j) \geq |U_j| - \mu n$ for each $i, j \in [2]$. Then, G contains a copy of T .*

Proof. Using Proposition 7.2, let $V(T) = A \cup B$ be a partition with $|A|, |B| \leq (2/3 - 10\mu)n$, such that $T[A]$ is a tree and $A' := \{v \in A : d_T(v, B) > 0\}$ is an independent set in T with $|A'| \leq 2$. For each $i \in [2]$, let $U_i^- = N_G(v_1, U_i) \cap N_G(v_2, U_i)$, then $|U_i^-| \geq (2/3 - 3\mu)n$ and $\delta(G[U_i^-]) \geq |U_i^-| - \mu n$. Embed one vertex in A' to v_1 . Then, greedily extend this to an embedding of the tree $T[A]$ in $G[U_1^- \cup \{v_1, v_2\}]$, such that if there is another vertex in A' , then it is embedded to v_2 . We can then extend this to copy of T by embedding $T[A' \cup B]$ greedily in $G[U_2^- \cup \{v_1, v_2\}]$. $\square$

7.3 Case II.B

For **Case II.B**, we need to find a subtree which has suitably more vertices in V_1 than in V_2 , which we do with the following result.

Proposition 7.4. *Let $1/n \ll \mu \ll 1$, and let T be an n -vertex tree with bipartition classes V_1 and V_2 such that $|V_1| \geq 1.1|V_2|$. Then, there exists a decomposition of T into subtrees T_1 and T_2 with a unique common vertex v , such that $10\mu n \leq |V(T_1) \cap V_1| - |V(T_1) \cap V_2| \leq 25\mu n$.*

Proof. Among all $v \in T$ and all subtrees T_1 and T_2 decomposing T with a unique common vertex v that satisfy $|V(T_1) \cap V_1| - |V(T_1) \cap V_2| \geq 12\mu n$, pick the combination that minimises $|T_1|$. Note that such v, T_1, T_2 exist as $|V_1| \geq 1.1|V_2|$ implies that $|V_1| - |V_2| \geq 12\mu n$, so picking an arbitrary $v \in V(T)$, letting $T_1 = T$ and $T_2 = T \setminus \{v\}$ would satisfy these conditions.

First, consider the case when $\deg(v, T_1) \geq 2$. If there is a component S of $T_1 - v$ that satisfies $|V(S) \cap V_2| - |V(S) \cap V_1| > 0$, then, transferring S and the edge between v and S from T_1 to T_2 gives two subtrees T'_1, T'_2 that still satisfy the required conditions but with $|T'_1| < |T_1|$, a contradiction. Thus, we can assume that every component of $T_1 - v$ has at least as many vertices in V_1 as in V_2 . Then, since there are at least 2 such components, at least one of them, say S' , satisfies $0 \leq |V(S') \cap V_1| - |V(S') \cap V_2| \leq (|V(T_1) \cap V_1| - |V(T_1) \cap V_2| + 1)/2$. In order for transferring S' and the edge between v and S' from T_1 to T_2 to not contradict the minimality of $|T_1|$, we must have that $(|V(T_1) \cap V_1| - |V(T_1) \cap V_2|) - (|V(S') \cap V_1| - |V(S') \cap V_2|) < 12\mu n$, and hence, $|V(T_1) \cap V_1| - |V(T_1) \cap V_2| \leq 25\mu n$, as required.

Suppose, then, that $\deg(v, T_1) = 1$. Let v' be the neighbour of v in T_1 . Let $T'_1 = T_1 - v$ and $T'_2 = T_2 + vv'$. Note that, in order to not get a contradiction, we must have $|V(T'_1) \cap V_1| - |V(T'_1) \cap V_2| < 12\mu n$, so $|V(T_1) \cap V_1| - |V(T_1) \cap V_2| < \mu n + 1 \leq 25\mu n$, as required. $\square$

Using Proposition 7.4, we can now prove the following lemma required in **Case II.B**.

Lemma 7.5. *Let $1/n \ll c \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes of sizes t_1 and t_2 satisfying $t_1 \geq 1.1t_2$. Let G be a graph containing two disjoint subsets U_1, U_2 and an additional vertex w , such that for each $i \in [2]$, $|U_i| \geq t_1 - \mu n$, $d_G(w, U_i) \geq \mu n$, and $d_G(u, U_{3-i}) \geq |U_{3-i}| - \mu n$ for every $u \in U_i$. Then, G contains a copy of T .*

Proof. By Proposition 7.4, we can find subtrees T_1 and T_2 decomposing T with a unique common vertex v , such that $10\mu n \leq |V(T_1) \cap V_1| - |V(T_1) \cap V_2| \leq 25\mu n$. Embed v to w , then we can greedily embed both T_1 and T_2 so that vertices in $V(T_2) \cap V_1$ and $V(T_1) \cap V_2$ go into U_1 and vertices in $V(T_1) \cap V_1$ and $V(T_2) \cap V_2$ go into U_2 . This is possible because

$$|V(T_1) \cap V_2| + |V(T_2) \cap V_1| \leq |V(T_1) \cap V_1| + |V(T_2) \cap V_1| - 10\mu n \leq t_1 + 1 - 10\mu n \leq |U_1| - \mu n,$$

$$|V(T_1) \cap V_1| + |V(T_2) \cap V_2| \leq |V(T_1) \cap V_2| + |V(T_2) \cap V_2| + 25\mu n \leq t_2 + 1 + 25\mu n \leq |U_2| - \mu n,$$

and w has $\mu n \geq \Delta(T)$ neighbours in both U_1 and U_2 . $\square$

7.4 Case II.C

The embedding result for **Case II.C** follows easily from Lemma 6.3 proved earlier for **Case I.B.1**.

Lemma 7.6. *Let $1/n \ll c \ll \mu \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$. Let G be a graph that contains two disjoint vertex sets U_1, U_2 such that $|U_1| \geq \lceil 2n/3 \rceil - 1$, $|U_2| \geq (2/3 - \mu)n$, and $\delta(G[U_i]) \geq |U_i| - \mu n$ for each $i \in [2]$. Suppose $G[U_1]$ contains at most $10^6 n$ non-edges, and that there exists $v \in V(G) \setminus (U_1 \cup U_2)$ with $d_G(v, U_i) \geq |U_i| - \mu n$ for each $i \in [2]$. Then, G contains a copy of T .*

Proof. Using Corollary 2.14, let T_1 and T_2 be a decomposition of T into subtrees with a unique common vertex t , so that $\lceil n/3 \rceil \leq |T_1| \leq |T_2| \leq \lceil 2n/3 \rceil$. Let $D = 0$ and $k = |U_1 \cup \{v\}| - |T_2|$, so that $k \geq 0$ and thus $G[U_1 \cup \{v\}]$ has at most $10^6 n + \mu n \leq 10^7(k + D + 1)|T_2|$ non-edges. Then, by Lemma 6.3, $G[U_1 \cup \{v\}]$ contains a copy of T_2 , in which t is copied to v . Since $|T_1| \leq 1 + n/2 \leq |U_2| - \mu n$, we can complete the embedding of T by greedily finding a copy of T_1 in $G[U_2 \cup \{v\}]$ with t copied to v . $\square$

7.5 Case II.D

We now give the embedding in **Case II.D**, where there are enough blue edges between U_1 and U_2 to embed a large subtree of the tree in U_2 and connect this across to embed the rest into U_1 .

Lemma 7.7. *Let $1/n \ll c \ll \mu \ll \varepsilon \ll 1$. Let G be a graph that contains two disjoint vertex sets U_1, U_2 such that $|U_i| \geq (2/3 - \varepsilon/3)n$ and $\delta(G[U_i]) \geq |U_i| - \mu n$ for each $i \in [2]$. Let T be an n -vertex tree with $\Delta(T) \leq cn$. Suppose T has an $(\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A \cup B$, such that $T[A, B]$ can be embedded into $G[U_1, U_2]$ with A embedded into U_1 and B embedded into U_2 . Then, G contains a copy of T .*

Proof. Let $A' = \{v \in A : d_T(v, B) > 0\}$ and let $t_1 \in A \setminus A'$, so $d_T(t_1, B) = 0$. Let $m = |A|$ and extend t_1 to an ordering $t_1, \dots, t_m$ of the vertices in A so that each vertex t_j except t_1 has exactly one neighbour to its left in $T[A]$ in this ordering.

Let $I = \{i \in [m] : t_i \in A'\}$, so $|I| \leq 2cn$ from the definition of sparse cuts. From assumption, there is an embedding ϕ' of $T[A, B]$ into $G[U_1, U_2]$ with A embedded into U_1 and B embedded into U_2 . Let $s'_i = \phi'(t_i)$ for each $i \in I$.

Pick $s_1 \in U_1 \setminus \{s'_i : i \in I\}$ uniformly at random. Then, for each $1 < i \leq m$, let $j_i < i$ satisfy $t_{j_i} t_i \in E(T[A])$, and embed t_i to some $s_i \in U_1$ randomly as follows.

U1 If $i \in I$, then let $s_i = s'_i$ if $s_{j_i} s'_i \in E(G)$, otherwise pick s_i uniformly at random from $N_G(s_{j_i}, U_1) \setminus (\{s_j : j < i\} \cup \{s'_j : j \in I\})$.

U2 If $i \notin I$, then pick s_i uniformly at random from $N_G(s_{j_i}, U_1) \setminus (\{s_j : j < i\} \cup \{s'_j : j \in I\})$.

For each $i \in I$, let T_i be the component containing t_i in $T[A' \cup B]$, let $X_i = |T_i| - 1$ if $s_i \neq s'_i$ and let $X_i = 0$ otherwise. For each $i \in I$, the probability that s_{j_i} is not in $N_G(s'_i)$ is at most $\mu n / (|U_1| - |A| - \mu n) \leq 3\mu/\varepsilon \leq \sqrt{\mu}$. Thus,

$$\sum_{i \in I} \mathbb{E}(X_i) \leq \sqrt{\mu} \cdot \sum_{i \in I} |T_i| \leq \sqrt{\mu} n,$$

so there is a realisation of $s_1, \dots, s_m$ for which $\sum_{i \in I} X_i \leq \sqrt{\mu} n$. Take such a realisation, and let $I' \subset I$ be the set of $i \in I$ for which $s_i = s'_i$, so that $\sum_{i \in I \setminus I'} (|T_i| - 1) = \sum_{i \in I} X_i \leq \sqrt{\mu} n$.

Let $\phi(t_i) = s_i$ for each $i \in [m]$ and note that this is an embedding of $T[A]$ in $G[U_1]$. Extend this to an embedding of $T[A \cup (\cup_{i \in I'} N_T(t_i))]$ using the embedding ϕ' of $T[A, B]$. Using $\delta(G[U_2]) \geq |U_2| - \mu n \geq |B|$, we can greedily extend ϕ to an embedding of $T[A \cup (\cup_{i \in I'} T_i)]$ by embedding the vertices in $\cup_{i \in I'} (V(T_i) \cap B)$ into U_2 . Then, using $\sum_{i \in I \setminus I'} (|T_i| - 1) \leq \sqrt{\mu} n$,

$$\delta(G[U_1]) \geq |U_1| - \mu n \geq (2/3 - \varepsilon/3)n - \mu n \geq |A| + \sum_{i \in I \setminus I'} (|T_i| - 1),$$

so we can extend ϕ to an embedding of T by greedily embedding the vertices in $\cup_{i \in I \setminus I'} (V(T_i) \cap B)$ into U_1 . Thus, G contains a copy of T . $\square$

7.6 Case II.E

In the last case, **Case II.E**, we aim to embed the tree T into $G_{\text{red}}[U_1^+, U_2^+]$. Our embedding method here is the most involved, but it shares some similarities with **Case I.B.1**. We will remove some leaves in V_1 from the tree T , and aim to embed the rest of the tree so that each remaining vertex in G in the correct side has plenty of neighbours among the vertices that need leaves attached to them, which would allow us to complete the embedding using Lemma 2.9. As mentioned before, the key difficulty is to ensure that the lower degree vertices in G are covered, either in the initial stage by some carefully chosen vertices in

T , or in the last stage by the leaves. An additional complication is that if $|U_1^+| \geq t_1$, then we will embed vertices in V_i into U_i^+ for each $i \in [2]$, but if $|U_1^+| = t_1 - 1$, which implies that $|U_2^+| = t_1 - 1$ as well, we will instead embed vertices in V_i into U_{3-i}^+ for each $i \in [2]$, except for one leaf in V_1 which needs to be embedded into U_1^+ . The last part is possible as there will be some red edges in $G[U_1^+]$ in this case. To avoid repetition, we will prove the following embedding lemma that will later be applied to (U_1^+, U_2^+) in the former case, and to (U_2^+, U_1^+) in the latter case.

Lemma 7.8. *Let $1/n \ll c \ll \mu \ll \alpha \ll \beta \ll \varepsilon \ll 1$. Let G be a graph on at most $2n$ vertices that contains two disjoint vertex sets U_1, U_2 with $|U_1| \geq t_1 - 1$ and $|U_2| \geq (1 - \mu)t_1$. Moreover, if $|U_1| = t_1 - 1$ then $e(G[U_2]) \geq 10^6 n$. Suppose $\delta(G[U_1, U_2]) \geq \beta n$ and, for each $i \in [2]$, all but at most μn vertices $u \in U_i$ satisfy $d_G(u, U_{3-i}) \geq |U_{3-i}| - \mu n$.*

Let $0 \leq \ell \leq 2cn$. Suppose there exist subsets $U_A \subset U_1$ and $U_B \subset U_2$ with $|U_A| < \ell$ and $|U_B| \leq (2/3 - \varepsilon)n$, such that $|N_G(u, U_2 \setminus U_B)| \geq |U_2 \setminus U_B| - \sqrt{n}$ for each $u \in U_1 \setminus U_A$.

Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 of sizes t_1 and t_2 , respectively, such that $t_1 \geq (2 - \mu)t_2$, and one of the following holds.

V1 *There is a set L of αn leaves of T in V_1 such that $d_T(u, L) \leq \sqrt{n}$ for each $u \in N_T(L)$.*

V2 *T contains a set V'_1 of ℓ vertices in V_1 and a disjoint set L of αn leaves in V_1 such that vertices in V'_1 have no common neighbour, $|N_T(V'_1)| \leq \varepsilon n$, and $N_T(V'_1) \cap N_T(L) = \emptyset$.*

Then, G contains a copy of T .

Proof. Let $V'_1 = \emptyset$ if T does not satisfy **V2**. In both cases, let s_1 be a leaf of T in $V_1 \setminus V'_1$, which exists by Lemma 2.10 using $t_1 \geq (2 - \mu)t_2$. Let s_2 be the neighbour of s_1 in T , and view T as being rooted at s_1 .

Let $L' = L \setminus N_T(s_2)$, and let P be the set of parents of L' . Note that $|L'| \geq \alpha n/2 \gg \mu n$. Partition $P = P_1 \cup P_2$, so that for each $j \in [2]$, $L_j := N_T(P_j, L')$ has size at least $\alpha n/10$. If **V2** holds, then take an injection $\phi : U_A \rightarrow V'_1$ such that no vertex in $\phi(U_A)$ is adjacent to s_2 , which is possible as $|U_A| < \ell = |V'_1|$ and vertices in V'_1 share no common neighbour. For every $u \in U_A$, let $\phi'(u)$ be the parent of $\phi(u)$ in the rooted tree T , and let $P' = \{\phi'(u) : u \in U_A\}$. Note that from the assumption in **V2**, $P' \cap P = \emptyset$. If **V2** does not hold then let both ϕ and ϕ' be the empty function, and let $P' = \emptyset$.

For each $j \in [2]$, let $U_j^- = \{u \in U_j : d_G(u, U_{3-j}) \geq |U_{3-j}| - \mu n\}$, so that $|U_j \setminus U_j^-| \leq \mu n$. Select a random subset $Z \subset U_2^-$ by including each vertex independently at random with probability β . Using Lemma 2.5, with positive probability we have $|Z| \leq 2\beta n$, $d_G(u, Z) \geq \beta^2 n/2$ for each $u \in U_1$, and $e(G[U_2] - Z) > 0$ if $|U_1| = t_1 - 1$. Fix a choice of Z with these properties. Note that

$$|U_2^- \setminus (U_B \cup Z)| \geq (1 - \mu)t_1 - \mu n - (2/3 - \varepsilon)n - 2\beta n \geq 2\alpha n.$$

By adding vertices to U_B if necessary, we may assume that $|U_2^- \setminus (U_B \cup Z)| = 2\alpha n$.

Let $T' = T - L'$ and $m = |T'|$. Extend s_1, s_2 to an ordering $s_1, \dots, s_m$ of the vertices in T' , such that for every $2 \leq i \leq m$, s_i has a unique neighbour in T' to its left in this ordering. If $|U_1| = t_1 - 1$, pick $v_1 v_2 \in E(G[U_2] - Z)$ arbitrarily. Otherwise, arbitrarily pick $v_1 v_2 \in E(G[U_1, U_2] - Z)$ with $v_1 \in U_1$ and $v_2 \in U_2$. Embed s_1 to v_1 and s_2 to v_2 . For each $3 \leq i \leq m$, suppose s_j has been embedded to v_j for all $j < i$ with only vertices in V'_1 embedded into U_A , and let $j_i < i$ satisfy $s_{j_i} s_i \in E(T')$. Note that if $s_i \in V_2$ and $v_{j_i} \in U_A$, then $s_{j_i} \in V'_1$, so $s_i \notin P \cup P'$ as vertices in V'_1 share no common neighbour and $N_T(V'_1) \cap N_T(L') = \emptyset$. Embed s_i randomly to some v_i as follows.

W1 If $s_i \in P_1$, then randomly select v_i from $N_G(v_{j_i}, U_2^- \setminus (U_B \cup Z \cup \{v_1, \dots, v_{i-1}\}))$.

W2 If $s_i \in P_2$, then randomly select v_i from $N_G(v_{j_i}, Z \setminus \{v_1, \dots, v_{i-1}\})$.

W3 If $s_i \in P'$, say $s_i = \phi'(u)$, then randomly select v_i from $N_G(v_{j_i}, N_G(u, Z) \setminus \{v_1, \dots, v_{i-1}\})$.

W4 If $s_i \in V_2 \setminus (P \cup P')$ and $v_{j_i} \notin U_A$, then randomly select v_i from $N_G(v_{j_i}, (U_2^- \cap U_B) \setminus (Z \cup \{v_1, \dots, v_{i-1}\}))$.

W5 If $s_i \in V_2 \setminus (P \cup P')$ and $v_{j_i} \in U_A$, then randomly select v_i from $N_G(v_{j_i}, Z \setminus \{v_1, \dots, v_{i-1}\})$.

W6 If $s_i \in V_1$ and there is some $u \in U_A$ such that $\phi(u) = s_i$, then let $v_i = u$.

W7 If $s_i \in V_1$ and there is no $u \in U_A$ such that $\phi(u) = s_i$, randomly select v_i from $N_G(v_{j_i}, U_1^- \setminus (U_A \cup \{v_1, \dots, v_{i-1}\}))$.

Note that **W1–W7** can always be carried out to obtain an embedding of T' . Moreover, from **W6**, U_A is covered by this embedding if **V2** holds.

For each $i \in [m]$, let $d_{i,1} = d_T(s_i, L_1)$, $d_{i,2} = d_T(s_i, L_2)$, and $d_i = d_{i,1} + d_{i,2}$.

Claim 7.9. *With high probability, for every vertex $v \in U_1$ not yet covered by the embedding of T' ,*

$$\sum_{i \in [m]: v_i \in N_G(v)} d_i \geq 10\mu n.$$

Proof of Claim 7.9. Note that $d_1 = d_2 = 0$. For each $v \in U_1 \setminus U_A$ and $i \in [m]$, let $X_{v,i} = d_{i,1}$ if $v_i \in N_G(v)$ and $X_{v,i} = 0$ otherwise. For each $v \in U_A$ and $i \in [m]$, let $X_{v,i} = d_{i,2}$ if $v_i \in N_G(v)$ and $X_{v,i} = 0$ otherwise.

Regardless of whether **V1** or **V2** holds, for each $v \in U_1 \setminus U_A$ and $i \in [m]$, let s_{j_i} be the unique neighbour of s_i to its left, then, using **W1** and $|U_2^- \setminus (U_B \cup Z)| = 2\alpha n$,

$$\begin{aligned} \mathbb{P}(X_{v,i} \neq d_{i,1} \mid X_{v,1}, \dots, X_{v,i-1}) &\leq \frac{|N_G(v_{j_i}, U_2^- \setminus (N_G(v) \cup U_B \cup Z \cup \{v_1, \dots, v_{i-1}\}))|}{|N_G(v_{j_i}, U_2^- \setminus (U_B \cup Z \cup \{v_1, \dots, v_{i-1}\}))|} \\ &\leq \frac{\sqrt{n}}{2\alpha n - \alpha n - \mu n} \leq \frac{1}{n^{1/3}}. \end{aligned}$$

Let $\ell_1 = |L_1| = \sum_{i: s_i \in P_1} d_{i,1} \geq \alpha n/10$. Let $P'_1 = \{s_i \in P_1 : d_{i,1} \leq n/\log^2 n\}$. By Lemma 2.6, if $\sum_{i: s_i \in P'_1} d_{i,1} \geq \ell_1/2$, then for every $v \in U_1 \setminus U_A$,

$$\mathbb{P}\left(\sum_{i: s_i \in P'_1} X_{v,i} < \alpha n/40\right) \leq \exp\left(-\frac{\ell_1^2}{10^4 \sum_{i: s_i \in P'_1} d_{i,1}^2}\right) \leq \exp\left(-\frac{\ell_1^2}{10^4 (\frac{\ell_1 \log^2 n}{n}) (\frac{n}{\log^2 n})^2}\right) \leq \frac{1}{n^2}. \quad (7.1)$$

Otherwise, $\sum_{i: s_i \in P_1 \setminus P'_1} d_{i,1} \geq \ell_1/2$. Since $|P_1 \setminus P'_1| \leq \log^2 n$, for each $v \in U_1 \setminus U_A$, the probability that there is some $J \subset P_1 \setminus P'_1$ with $|J| \geq 10$ and $X_{v,i} \neq d_{i,1}$ for each $i \in J$ is at most

$$(\log^2 n)^{10} \cdot \left(\frac{1}{n^{1/3}}\right)^{10} \leq \frac{1}{n^2},$$

so with probability at least $1 - 1/n^2$, $\sum_{i: s_i \in P_1 \setminus P'_1} X_{v,i} \geq \ell_1/2 - 10cn \geq \alpha n/40$. Combined with (7.1) and using a union bound, we have with high probability that $\sum_{i \in [m]: v_i \in N_G(v)} d_i \geq \sum_{i: s_i \in P_1} X_{v,i} \geq \alpha n/40 \geq 10\mu n$ for all $v \in U_1 \setminus U_A$.

Now let $v \in U_A$. If **V2** holds, then v is covered by the embedding of T' so there is nothing to prove. Suppose now that **V1** holds. For every $i \in [m]$, note that $d_{i,2} \leq d_i \leq \sqrt{n}$, and let s_{j_i} be the unique neighbour of s_i to its left. Then, using **W2**,

$$\begin{aligned} \mathbb{P}(X_{v,i} = d_{i,2} \mid X_{v,1}, \dots, X_{v,i-1}) &\geq \frac{|N_G(v_{j_i}, (Z \cap N_G(v)) \setminus \{v_1, \dots, v_{i-1}\})|}{|N_G(v_{j_i}, Z \setminus \{v_1, \dots, v_{i-1}\})|} \\ &\geq \frac{\beta^2 n/2 - \alpha n - \mu n}{2\beta n} \geq \beta/10. \end{aligned}$$

Let $\ell_2 = |L_2| = \sum_{i: s_i \in P_2} d_{i,2} \geq \alpha n/10$, then by Lemma 2.6,

$$\mathbb{P}\left(\sum_{i: s_i \in P_2} X_{v,i} < \alpha \beta n/200\right) \leq \exp\left(-\frac{\beta^2 \ell_2^2}{10^4 \sum_{i: s_i \in P_2} d_{i,2}^2}\right) \leq \exp\left(-\frac{\beta^2 \ell_2^2}{10^4 (\ell_2/\sqrt{n}) \cdot (\sqrt{n})^2}\right) \leq \frac{1}{n^2}.$$

By a union bound, with high probability, all $v \in U_A$ satisfy $\sum_{i \in [m]: v_i \in N_G(v)} d_i \geq \sum_{i: s_i \in P_2} X_{v,i} \geq \alpha \beta n/200 \geq 10\mu n$. $\square$

Finally, it remains to embed L' . Let U be the set of unused vertices in U_1 , note that $|U| \geq |L'|$, and every $v \in U$ satisfies $\sum_{i \in [m]: v_i \in N_G(v)} d_i \geq 10\mu n$ by Claim 7.9. We verify that Hall's condition holds between the set $\{v_i : s_i \in P\}$ of images of parents of L' and U . Indeed, for any non-empty $J \subset P$, if $0 < \sum_{i: s_i \in J} d_i \leq |L'| - \mu n$, then as $v_i \in U_2^-$ for each $i \in J$ by **W1** and **W2**, we have

$$|N_G(\{v_i : s_i \in J\}, U)| \geq |U| - \mu n \geq |L'| - \mu n \geq \sum_{i: s_i \in J} d_i.$$

If instead $\sum_{i: s_i \in J} d_i > |L'| - \mu n$, then $|N_G(\{v_i : s_i \in J\}, U)| = |U| \geq |L'| \geq \sum_{i: s_i \in J} d_i$, as any $v \in U \setminus N_G(\{v_i : s_i \in J\})$ would satisfy $\sum_{i \in [m]: v_i \in N_G(v)} d_i \leq \mu n$, a contradiction. Therefore, by Lemma 2.9, we can embed L' into U to finish a copy of T in G . $\square$

7.7 Proof of Theorem 2.4

Our final step before we can prove Theorem 2.4 is to prove the following results which give the desired sparse cut (see Definition 7.1) used in **Case II.D** and **Case II.E**.

Proposition 7.10. *Let $1/n \ll c \ll \mu \ll \varepsilon \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 satisfying $|V_1| \geq (2 - \mu)|V_2|$. Then, T has an $(\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A \cup B$, such that at most two vertices in $\{v \in A : d_T(v, B) > 0\}$ are adjacent to leaves of T in A .*

Proof. By Lemma 2.15, there is a vertex $v \in V(T)$ such that each component of $T - v$ has size at most $n/2$. Let $\ell \leq cn$ be the number of components in $T - v$ and let $T_1, \dots, T_\ell$ be these components in order of decreasing size. Let r be maximal subject to $\sum_{i=1}^r |T_i| \leq (2/3 - 2\varepsilon)n$, and note that $r \geq 1$.

If $|T_r| \leq \sqrt{n}$, then as $|T_i| \leq \sqrt{n}$ for all $i \geq r$, there exists $r' > r$ such that $(1/3 + 3\varepsilon/2)n \leq \sum_{i=r'}^\ell |T_i| \leq (1/3 + 3\varepsilon/2)n + \sqrt{n}$, so $A = \{v\} \cup N_T(v) \cup (\cup_{i < r'} V(T_i))$ and $B = V(G) \setminus A$ form an $(\varepsilon, \sqrt{n})$ -sparse cut.

Assume now that $|T_r| > \sqrt{n}$, then as $|T_i| > \sqrt{n}$ for all $i \in [r]$, we have $r \leq \sqrt{n}$. If, moreover, $\sum_{i \in [r]} |T_i| \geq (1/3 + \varepsilon)n$, then $A = \{v\} \cup (\cup_{i=r+1}^\ell V(T_i))$ and $B = V(G) \setminus A$ form an $(\varepsilon, \sqrt{n})$ -sparse cut. Thus, we can assume that $\sum_{i \in [r]} |T_i| < (1/3 + \varepsilon)n$.

Now, by the maximality of r , we have $|T_{r+1}| \geq (2/3 - 2\varepsilon)n - (1/3 + \varepsilon)n = (1/3 - 3\varepsilon)n$. As $|T_1|, |T_r| \geq |T_{r+1}| \geq (1/3 - 3\varepsilon)n$, we must have $r = 1$, and $(1/3 - 3\varepsilon)n \leq |T_2| \leq |T_1| \leq (1/3 + \varepsilon)n$. For each $j \in [2]$, let v_j be the unique neighbour of v in T_j , and view T_j as a tree rooted at v_j . Let v'_j be a vertex farthest away from v_j in T_j subject to the condition that the subtree $T'_j \subset T_j$ containing v'_j and all of its descendants in T_j has size at least $(1/3 - 10\varepsilon)n$.

Let $\{S_i : i \in I_1\}$ be the components of $T'_1 - v'_1$ and let $\{S_i : i \in I_2\}$ be the components of $T'_2 - v'_2$. If $\sum_{i \in I_1: |S_i| < \sqrt{n}} |S_i| \geq 5\varepsilon n$, then we can find $I'_1 \subset I_1$ such that $|S_i| \leq \sqrt{n}$ for each $i \in I'_1$ and $5\varepsilon n \leq \sum_{i \in I'_1} |S_i| \leq 6\varepsilon n$. Then, $B = V(T_2) \cup (\cup_{i \in I'_1} V(S_i) \setminus N_T(v'_1))$ and $A = V(G) \setminus B$ form an $(\varepsilon, \sqrt{n})$ -sparse cut. Similarly, an $(\varepsilon, \sqrt{n})$ -sparse cut exists if $\sum_{i \in I_2: |S_i| < \sqrt{n}} |S_i| \geq 5\varepsilon n$. Thus, we can assume that $\sum_{i \in I_1: |S_i| \geq \sqrt{n}} |S_i|$ and $\sum_{i \in I_2: |S_i| \geq \sqrt{n}} |S_i|$ are both at least $(1/3 - 20\varepsilon)n$. Let $I \subset \{i \in I_1 \cup I_2 : |S_i| \geq \sqrt{n}\}$ be minimal subject to $\sum_{i \in I} |S_i| \geq (1/3 + \varepsilon)n$. Then, minimality and the choices of v'_1, v'_2 imply that $\sum_{i \in I} |S_i| < (1/3 + \varepsilon)n + (1/3 - 10\varepsilon)n = (2/3 - 9\varepsilon)n$, so $B = \cup_{i \in I} V(S_i)$ and $A = V(T) \setminus B$ form an $(\varepsilon, \sqrt{n})$ -sparse cut.

Therefore, T always contains an $(\varepsilon, \sqrt{n})$ -sparse cut $A \cup B$. Finally, it is easy to verify that in all cases above, there are at most two vertices in $\{v \in A : d_T(v, B) > 0\}$ that are adjacent to leaves of T in A . $\square$

Proposition 7.11. *Let $1/n \ll c \ll \mu \ll \alpha \ll \varepsilon \ll 1$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 satisfying $|V_1| \geq (2 - \mu)|V_2|$. Then, T has an $(\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A \cup B$ such that at least one of the following holds.*

X1 *There is a set L of at least αn leaves of T in V_1 such that $d_T(u, L) \leq \sqrt{n}$ for each $u \in N_T(L)$.*

X2 *T contains a set V'_1 of $|\{v \in A : d_T(v, B) > 0\}|$ vertices in V_1 , and a disjoint set L' of αn leaves in V_1 , such that vertices in V'_1 have no common neighbour, $|N_T(V'_1)| \leq \varepsilon n$, and $N_T(V'_1) \cap N_T(L') = \emptyset$.*

Proof. Let $c \ll \gamma \ll \varepsilon$. We begin with the following claim.

Claim 7.12. Let L be a set of leaves of T in V_1 . Suppose that $V(T) = A' \cup B'$ is a $(2\varepsilon, \sqrt{n})$ -sparse cut of T , and let $A_1 = \{a \in A' : d_T(a, B') > 0\}$. For each $a \in A_1$, let R_a be the component of $T - (A' \setminus A_1)$ containing a , and suppose there exists $r_a \in V(R_a - a) \cap V_1$ with $d_T(r_a) \leq 1/\gamma$. Moreover, assume that at least $1/3$ of the vertices in $\cup_{a \in A_1} R_a$ are in L , then T has an $(\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A \cup B$ and some corresponding L', V'_1 such that **X2** holds.

Proof of Claim 7.12. Arrange the components $R_a, a \in A_1$, in decreasing order of $|V(R_a) \cap L|/|R_a|$. Take a minimal collection $A'_1 \subset A_1$, starting from the elements with the highest ratio, such that $\sum_{a \in A'_1} |V(R_a) \cap L| \geq 2\alpha n$.

Suppose first that $\sum_{a \in A'_1} |V(R_a) \cap L| \leq 4\alpha n$. Noting that $(\sum_{a \in A_1} |V(R_a) \cap L|)/(\sum_{a \in A_1} |R_a|) \leq \max_{a \in A_1} |V(R_a) \cap L|/|R_a|$, we have $\sum_{a \in A'_1} |R_a| \leq 12\alpha n \ll \varepsilon n$. Let $L' \subset \cup_{a \in A'_1} (V(R_a) \cap L)$ have size αn , $B = \cup_{a \in A_1 \setminus A'_1} V(R_a - a)$, $A = V(G) \setminus B$, $V'_1 = \{r_a : a \in A_1 \setminus A'_1\}$. Note that $|N_T(V'_1)| \leq |A_1|/\gamma \leq 2cn/\gamma \ll \varepsilon n$ and $N_T(V'_1)$ is disjoint from $N_T(L')$, so **X2** holds with respect to the $(\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A \cup B, L'$, and V'_1 .

If $\sum_{a \in A'_1} |V(R_a) \cap L| > 4\alpha n$, then the minimality of A'_1 implies that there exists $a^* \in A'_1 \subset A_1$ satisfying $|V(R_{a^*}) \cap L| \geq 2\alpha n$. In particular, we can find a set $L' \subset V(R_{a^*}) \cap L$ of size αn , and some $r'_{a^*} \in V(R_{a^*}) \cap (L \setminus L')$, such that $N_T(r'_{a^*}) \cap N_T(L') = \emptyset$. Then, let $A = A'$ and $B = B'$, and note that **X2** holds with respect to the $(\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A \cup B, L'$, and $V'_1 = \{r_a : a \in A_1 \setminus \{a^*\}\} \cup \{r'_{a^*}\}$. $\square$

Now, Proposition 7.10 gives a $(100\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A' \cup B'$, such that at most two vertices in $\{v \in A' : d_T(v, B') > 0\}$ are adjacent to leaves of T in A' . Let $A_1 = \{v \in A' : d_T(v, B') > 0\}$. For each $a \in A_1$, let R_a be the component of $T - (A' \setminus A_1)$ containing a . Let A_2 be the set of $a \in A_1$ such that $R_a - a$ contains at least $\gamma|R_a|$ vertices in V_1 . Then, for each $a \in A_2$, there exists $r_a \in V(R_a - a) \cap V_1$ such that $d_T(r_a) \leq 1/\gamma$.

Case I. $\sum_{a \in A_2} |R_a - a| \geq (1/3 + 2\varepsilon)n$. By Lemma 2.10, T contains at least $t_1 - t_2 \geq (1/3 - 10\mu)n$ leaves in V_1 . If $A' \setminus A_1$ contains at least $10\alpha n \geq \alpha n + 2cn$ leaves of T in V_1 , then there is a set L' of αn such leaves in $A' \setminus A_1$ with their parents not in A_1 . Then, let $B = \cup_{a \in A_2} V(R_a - a)$, $A = V(G) \setminus B$, and note that A and B form an $(\varepsilon, \sqrt{n})$ -sparse cut. Set $V'_1 = \{r_a : a \in A_2\}$, then **X2** holds with respect to A, B, L' , and V'_1 .

If instead $A' \setminus A_1$ contains at most $10\alpha n$ leaves of T in V_1 , then at least $(1/3 - \varepsilon)n$ leaves of T in V_1 are in $\cup_{a \in A_1} R_a$, so $\cup_{a \in A_2} R_a$ contains a set L_1 of at least $(1/3 - 10\varepsilon)n$ of leaves of T in V_1 , as $\cup_{a \in A_1 \setminus A_2} R_a$ contains at most γn vertices in V_1 . In particular, at least $1/3$ of the vertices in $\cup_{a \in A_2} R_a$ are in L_1 . Let $B'' = \cup_{a \in A_2} V(R_a - a)$ and $A'' = V(G) \setminus B''$, then we can apply Claim 7.12 to the $(2\varepsilon, \sqrt{n})$ -sparse cut $V(T) = A'' \cup B''$ to finish the proof.

Case II. $\sum_{a \in A_2} |R_a - a| < (1/3 + 2\varepsilon)n$, so $\sum_{a \in A_1 \setminus A_2} |R_a - a| \geq 97\varepsilon n$. Then, $\sum_{a \in A_1 \setminus A_2} (|V(R_a - a) \cap V_2| - |V(R_a - a) \cap V_1|) \geq \sum_{a \in A_1 \setminus A_2} ((1 - 2\gamma)|R_a| - 1) \geq 95\varepsilon n$. Consider the subtree T' obtained by removing $R_a - a$ from T for each $a \in A_1 \setminus A_2$, and note that at most $|A_1 \setminus A_2| \ll \varepsilon n$ leaves in T' are not leaves in T . By Lemma 2.10, T' contains at least $|V(T') \cap V_1| - |V(T') \cap V_2| \geq t_1 - t_2 + 95\varepsilon n \geq (1/3 + 92\varepsilon)n$ leaves in V_1 . Thus, T contains at least $(1/3 + 91\varepsilon)n$ leaves in V_1 .

If there is a set L of at least αn leaves of T in V_1 whose parents are all adjacent to at most $\sqrt{n}$ leaves of T , then **X1** holds and we are done, so suppose otherwise. By Lemma 2.15, there is a vertex v in T such that every component of $T - v$ has size at most $n/2$. View T as being rooted at v . Then, there is a set L_2 of at least $(1/3 + 90\varepsilon)n$ leaves in V_1 , each of whose parent in T is adjacent to at least $\sqrt{n}$ leaves in T , and no leaf in L_2 or parent of leaf in L_2 is equal to v . Let $p_1, \dots, p_k$ be the parents of L_2 in T , and note that $k \leq \sqrt{n}$. Then, for every $i \in [k]$, let P_i be the subtree of T induced by p_i and all of its descendants in T , and note that $|P_i| \leq n/2$ as $p_i \neq v$. Let $I \subset [k]$ be the set of indices $i \in [k]$, such that p_i is not a descendant of any p_j with $j \neq i$. Let $I' \subset I$ be minimal subject to $\sum_{i \in I'} |P_i| \geq (1/3 + 2\varepsilon)n$.

Case II.1. $\sum_{i \in I'} |P_i| \leq (2/3 - 2\varepsilon)n$. Let $\bar{B} = \cup_{i \in I'} V(P_i)$ and $\bar{A} = V(T) \setminus \bar{B}$. Note that $T[\bar{A}]$ is a tree, and $V(T) = \bar{A} \cup \bar{B}$ is a $(2\varepsilon, \sqrt{n})$ -sparse cut. Let $\bar{A}_1 = \{a \in \bar{A} : d_T(a, \bar{B}) > 0\}$. For each $a \in \bar{A}_1$, let $\bar{R}_a$ be the component of $T - (\bar{A} \setminus \bar{A}_1)$ containing a , pick any $p_i \in N_T(a, \bar{B})$, and then pick r_a to be any children of p_i in L_2 .

If $|\overline{B} \cap L_2| \leq (1/3 + 3\varepsilon)n$, then at least εn leaves in L_2 are in $\overline{A}$, so we can pick a set L' of αn such leaves. Then, $\overline{A}$, $\overline{B}$, L' , and $V'_1 = \{r_a : a \in \overline{A}_1\}$ satisfy **X2**.

If $|\overline{B} \cap L_2| > (1/3 + 3\varepsilon)n$, then at least $1/3$ of the vertices in $\cup_{a \in \overline{A}_1} R_a$ are leaves in L_2 , so we are done by Claim 7.12.

Case II.2. $\sum_{i \in I'} |P_i| > (2/3 - 2\varepsilon)n$. Then, by minimality, $|P_i| \geq (1/3 - 4\varepsilon)n$ for each $i \in I$, so $|I| = 2$ as $|P_i| \leq n/2$. Without loss of generality, say $I = \{1, 2\}$, and note that $|P_1|, |P_2| \leq (1/3 + 2\varepsilon)n$ by minimality. Since at most $(1/3 + 2\varepsilon)n$ vertices in T are outside of $P_1 \cup P_2$, $P_1 \cup P_2$ contains at least $80\varepsilon n$ leaves in L_2 , so we may assume that, say, P_1 contains at least $40\varepsilon n$ such leaves.

Let $\{q_1, \dots, q_m\} \subset \{p_3, \dots, p_k\}$ be the set of descendants of p_1 in T , and assume they are ordered so that if q_i is a descendant of q_j in T , then $i \leq j$. Let Q_i be the subtree of T induced by q_i and all of its descendants in T , and let $m' \in [m]$ be minimal subject to $|L_2 \cap (\cup_{i=1}^{m'} V(Q_i))| \geq 20\varepsilon n$. Note that by minimality and $\Delta(T) \leq cn$, $|L_2 \cap (\cup_{i=1}^{m'} V(Q_i))| \leq 21\varepsilon n$, and so $|(L_2 \cap V(P_1)) \setminus (\cup_{i=1}^{m'} V(Q_i))| \geq 19\varepsilon n$. Let $\overline{B} = V(P_2) \cup (\cup_{i=1}^{m'} V(Q_i))$, and observe that $(1/3 - 4\varepsilon)n + 20\varepsilon n \leq |\overline{B}| \leq (2/3 + 4\varepsilon)n - 19\varepsilon n$. Let $\overline{A} = V(G) \setminus \overline{B}$, and note that $V(T) = \overline{A} \cup \overline{B}$ form a $(2\varepsilon, \sqrt{n})$ -sparse cut. Thus, we are now in the same situation as in **Case II.1**, and can finish the proof in the same way. $\square$

Finally, we can put all the work of this section together to prove Theorem 2.4, following the outline in Section 7.1.

Proof of Theorem 2.4. Let $1/n \ll c \ll \mu \ll 1$ and let $t_1, t_2 \in \mathbb{N}$ satisfy $t_1 + t_2 = n$ and $t_1 \geq (2 - \mu)t_2$. Let G be a Type II (μ, t_1, t_2) -extremal graph on $\max\{2t_1, t_1 + 2t_2\} - 1$ vertices, so from definition there are disjoint subsets $U_1, U_2 \subset V(G)$ such that $|U_1|, |U_2| \geq (1 - \mu)t_1$, and for each $i \in [2]$ and $u \in U_i$, $d_{\text{red}}(u, U_i) \leq \mu n$ and $d_{\text{blue}}(u, U_{3-i}) \leq \mu n$. Let T be an n -vertex tree with $\Delta(T) \leq cn$ and bipartition classes V_1 and V_2 with $|V_i| = t_i$ for each $i \in [2]$. We need to find a monochromatic copy of T in G .

Let $\mu \ll \beta \ll 1$. Let $U_1^+, U_2^+ \subset V(G)$ be maximal disjoint sets with $U_1 \subset U_1^+$, $U_2 \subset U_2^+$, and $d_{\text{red}}(u, U_{3-i}) \geq \beta n$ for every $i \in [2]$ and $u \in U_i^+$. Note that $|U_i^+ \setminus U_i| \leq 2\mu n$ for each $i \in [2]$. By relabelling if necessary, we can assume that $|U_1^+| \geq |U_2^+|$.

First (for **Case II.A**), suppose that there are distinct vertices $v_1, v_2 \in V(G) \setminus (U_1^+ \cup U_2^+)$. By the maximality of U_1^+ and U_2^+ , we have that $d_{\text{blue}}(v_i, U_j) \geq |U_j| - \beta n$ for each $i, j \in [2]$. As $t_1 \geq (2 - \mu)t_2$, we have $|U_i| \geq t_1 - \mu n \geq (2/3 - 10\mu)n$ for each $i \in [2]$. Since $\delta(G_{\text{blue}}[U_i]) \geq |U_i| - \mu n$ for each $i \in [2]$, we can apply Lemma 7.3 to find a copy of T in G_{blue} .

Thus, we can assume that $|V(G) \setminus (U_1^+ \cup U_2^+)| \leq 1$. It follows that $|U_1^+| \geq \lceil (|G| - 1)/2 \rceil \geq \lceil 2n/3 \rceil - 1$, and $|U_1^+| \geq t_1 - 1$. Next (for **Case II.B**), suppose there is some vertex $v \in V(G)$ with at least βn red neighbours in both U_1^+ and U_2^+ . Then, v has at least $\beta n - 2\mu n \geq \mu n$ red neighbour in both U_1 and U_2 . Therefore, by Lemma 7.5, G contains a red copy of T .

Hence, we can assume there is no such vertex v , from which we get $\delta(G_{\text{blue}}[U_i^+]) \geq |U_i^+| - \beta n$ for each $i \in [2]$. Next (for **Case II.C**), suppose $G[U_1^+]$ contains at most $10^6 n$ red edges and there is exactly one vertex w in $V(G) \setminus (U_1^+ \cup U_2^+)$. Using the maximality like above, $d_{\text{blue}}(w, U_i^+) \geq |U_i^+| - 2\beta n$ for each $i \in [2]$, so we can find a blue copy of T in G using Lemma 7.6.

Thus, we can assume that either $G[U_1^+]$ has more than $10^6 n$ red edges, or $V(G) \setminus (U_1^+ \cup U_2^+) = \emptyset$. Let $\beta \ll \varepsilon \ll 1$. Let $V(T) = A \cup B$ be an $(\varepsilon, \sqrt{n})$ -sparse cut given by Proposition 7.11. Suppose (for **Case II.D**) that $T[A, B]$ can be embedded into $G_{\text{blue}}[U_1^+, U_2^+]$, either with A embedded into U_1^+ and B embedded into U_2^+ , or the other way around, then G contains a blue copy of T by Lemma 7.7.

Finally, suppose (for **Case II.E**) that $T[A, B]$ cannot be embedded into $G_{\text{blue}}[U_1^+, U_2^+]$. Let $\ell = |\{v \in A : d_T(v, B) > 0\}|$ and list the elements in $\{v \in A : d_T(v, B) > 0\}$ as $a_1, \dots, a_\ell$. For each $i \in [\ell]$, let $d_i = d_T(a_i, B) \leq \sqrt{n}$.

If $|U_1^+| \geq t_1$, let $I \subset [\ell]$ be a maximal set for which there are distinct vertices $\{w_i : i \in I\} \subset U_1^+$ and disjoint subsets $\{W_i \subset U_2^+ : i \in I\}$, such that $W_i \subset N_{\text{blue}}(w_i)$ and $|W_i| = d_i$ for each $i \in I$. As we are not in **Case II.D**, $|I| < \ell$. Let $U_A = \{w_i : i \in I\}$ and $U_B = \cup_{i \in I} W_i$. Then, the maximality of I implies that every vertex in $U_1^+ \setminus U_A$ has at most $\sqrt{n}$ blue neighbours in $U_2^+ \setminus U_B$. Therefore, by Lemma 7.8, G contains a copy of T in red.

If $|U_1^+| = t_1 - 1$, then we must have $|V(G) \setminus (U_1^+ \cup U_2^+)| = 1$ and $|U_2^+| = t_1 - 1$. Since we are not in **Case II.C**, $G[U_1^+]$ has at least $10^6 n$ red edges. We now proceed as above but swapping the role of U_1^+ and U_2^+ . Let $I \subset [\ell]$ be a maximal set for which there are distinct vertices $\{w_i : i \in I\} \subset U_2^+$ and disjoint

subsets $\{W_i \subset U_1^+ : i \in I\}$, such that $W_i \subset N_{\text{blue}}(w_i)$ and $|W_i| = d_i$ for each $i \in I$. Let $U_A = \{w_i : i \in I\}$ and $U_B = \cup_{i \in I} W_i$. Then, $|I| < \ell$ and the maximality of I implies that every vertex in $U_2^+ \setminus U_A$ has at most $\sqrt{n}$ blue neighbours in $U_1^+ \setminus U_B$. Therefore, by Lemma 7.8, G contains a copy of T in red. This completes the proof of Theorem 2.4. $\square$

References

- [1] P. Balister, B. Bollobás, M. Campos, S. Griffiths, E. Hurley, R. Morris, J. Sahasrabudhe, and M. Tiba. Upper bounds for multicolour Ramsey numbers. *arXiv:2410.17197*, 2024.
- [2] G. Besomi, M. Pavez-Signé, and M. Stein. Degree conditions for embedding trees. *SIAM Journal on Discrete Mathematics*, 33(3):1521–1555, 2019.
- [3] B. Bollobás. *Modern Graph Theory*. Springer, 1998.
- [4] J. Bondy and P. Erdős. Ramsey numbers for cycles in graphs. *Journal of Combinatorial Theory, Series B*, 14(1):46–54, 1973.
- [5] S. A. Burr. Generalized Ramsey theory for graphs - a survey. In *Graphs and Combinatorics*, pages 52–75. Springer, 1974.
- [6] S. A. Burr and P. Erdős. On the magnitude of generalized Ramsey numbers for graphs. In *Infinite and finite sets*, pages 215–240. János Bolyai Mathematical Society, 1975.
- [7] S. A. Burr and P. Erdős. Extremal Ramsey theory for graphs. *Utilitas Mathematica*, 9:247–258, 1976.
- [8] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. An exponential improvement for diagonal Ramsey. *arXiv:2303.09521*, 2023.
- [9] V. Chvátal, V. Rödl, E. Szemerédi, and W. T. Trotter Jr. The Ramsey number of a graph with bounded maximum degree. *Journal of Combinatorial Theory, Series B*, 34(3):239–243, 1983.
- [10] F. F. Dubó and M. Stein. On the Ramsey number of the double star. *Discrete Mathematics*, 348(1):114227, 2025.
- [11] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.
- [12] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. Ramsey numbers for brooms. *Congressus Numerantium*, 35:283–293, 1982.
- [13] P. Erdős, Z. Füredi, M. Loeb, and V. T. Sós. Discrepancy of trees. *Studia Scientiarum Mathematicarum Hungarica*, 30(1-2):47–57, 1995.
- [14] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.
- [15] R. J. Faudree and R. H. Schelp. All Ramsey numbers for cycles in graphs. *Discrete Mathematics*, 8(4):313–329, 1974.
- [16] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae, Sectio Mathematica*, 10:167–170, 1967.
- [17] J. W. Grossman, F. Harary, and M. Klawe. Generalized Ramsey theory for graphs, x: double stars. *Discrete Mathematics*, 28(3):247–254, 1979.
- [18] P. Gupta, N. Ndiaye, S. Norin, and L. Wei. Optimizing the CGMS upper bound on Ramsey numbers. *arXiv:2407.19026*, 2024.

- [19] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30, 1935.
- [20] F. Harary. Recent results on generalized Ramsey theory for graphs. In *Graph Theory and Applications*, pages 125–138. Springer, 1972.
- [21] P. E. Haxell, T. Łuczak, and P. W. Tingley. Ramsey numbers for trees of small maximum degree. *Combinatorica*, 22(2):287–320, 2002.
- [22] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience, 2000.
- [23] J. Komlós, G. N. Sárközy, and E. Szemerédi. Spanning trees in dense graphs. *Combinatorics, Probability and Computing*, 10(5):397–416, 2001.
- [24] J. Komlós and M. Simonovits. Szemerédi’s Regularity Lemma and its applications in graph theory. In *Combinatorics, Paul Erdős is eighty, Volume 2*, pages 295–352. János Bolyai Mathematical Society, 1996.
- [25] M. Krivelevich. Embedding spanning trees in random graphs. *SIAM Journal on Discrete Mathematics*, 24(4):1495–1500, 2010.
- [26] C. Lee. Ramsey numbers of degenerate graphs. *Annals of Mathematics*, 185(3):791–829, 2017.
- [27] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics, 1989*, pages 148–188. Cambridge University Press, 1989.
- [28] R. Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.
- [29] S. Norin, Y. R. Sun, and Y. Zhao. Asymptotics of Ramsey numbers of double stars. *arXiv:1605.03612*, 2016.
- [30] A. Pokrovskiy. Hyperstability in the Erdős-Sós conjecture. *arXiv:2409.15191*, 2024.
- [31] S. Radziszowski. Small Ramsey numbers. *The Electronic Journal of Combinatorics*, DS1, 2024.
- [32] F. P. Ramsey. On a problem of formal logic. *Proceedings of The London Mathematical Society*, s2-30(1):264–286, 1930.
- [33] V. Rosta. On a Ramsey-type problem of J. A. Bondy and P. Erdős. II. *Journal of Combinatorial Theory, Series B*, 15(1):105–120, 1973.
- [34] M. Stein. Tree containment and degree conditions. In *Discrete Mathematics and Applications*, pages 459–486. Springer, 2020.
- [35] N. C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, pages 73–155. Polish Scientific Publishers, 1999.
- [36] Y. Zhao. Proof of the $(n/2 - n/2 - n/2)$ conjecture for large n . *The Electronic Journal of Combinatorics*, 18(1):P27, 2011.